Understanding How the Flash Clashes are Affected in an Asymmetric Informational Market with Agent-based Modelling

Sun, Yan

Awarding institution:
King's College London

Understanding the Impacts of Flash Crashes in a Market under Asymmetric Information

with Agent-based Modelling

Yan Sun

Supervisors: Prof. Peter McBurney

Dr. Frederik Mallmann-Trenn

at The Department of Informatics

King's College London

This dissertation is submitted for the degree of

Doctor of Philosophy

November 2023

Declaration

I hereby declare that except where specific reference is made to the work of others, the contents of this dissertation are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other university. This dissertation is my own work and contains nothing which is the outcome of work done in collaboration with others, except as specified in the text and Acknowledgements. This dissertation contains fewer than 45,000 words including appendices, bibliography, footnotes, tables and equations and has fewer than 150 figures.

Yan Sun

November 2023

Acknowledgements

Firstly, I would like to thank my supervisor, Prof. Peter McBurney, for his supervision and encouragement throughout the four years of my PhD. I am also grateful to Dr. Frederik Mallmann-Trenn for being my second supervisor, and his valuable advice and comments gave me significant insights into my research. I would also like to thank my first two supervisors, Dr. Steve Phelps and Prof. Patrick McCorry. Although they left King's College London and no longer work for universities for different reasons, their supervision has been important to me. Steve also supervised me for my Master's program, and what he taught sparked my interest in doing a PhD in computational finance; Patrick, as a specialist in blockchain, also taught me a lot about cryptocurrencies.

The outbreak of COVID-19 in 2020 was a challenge to us international students both mentally and physically. Many thanks to my friends who supported me during the tough period. I would like to thank Kejia Wu, Wanwen Liang, Yuxuan Li and Jinglei Xu for encouraging me when I was down and providing material support. I am also grateful to all my friends I haven't mentioned, and I am not alone in my studies because of you.

I am extremely grateful to King's College London and all the academics and staff I have met since I started my Master's course. The fantastic 5-year study at King's College London has been the most enjoyable time of my student life, and I am honoured forever to be a student at King's College London.

I would like to thank the China Scholarship Council for providing me with financial support for my PhD studies so I have not been financially burdened. I would also like to thank all the staff at the Embassy of the People's Republic of China in the UK who have been so helpful throughout my PhD.

Finally, I would like to deeply thank my parents for their mental and financial support, and for their selfless dedication throughout my life, which has enabled me to pursue my ambitions. I cannot thank both of you enough for everything you have done for me.

Abstract

This thesis explores the impact of flash crashes on the dynamics of financial markets with asymmetric information. We built, implemented, and analysed an agent-based model of an extended information-sequential trading framework inspired by the models of Das and Glosten-Milgrom, where an exogenous fake shock is added into the system to disturb the actions of some traders where there is informational asymmetry. The key modelled agents include fundamental traders, who place orders at preferred prices; zero-intelligence traders, who place orders randomly; a market maker, who provides liquidity; and an exchange matching all orders under continuous auctions or batch auctions. To this end, by Monte-Carlo methods, we implement the model and examine the dynamics of the market under information asymmetry in the following aspects: the market structure, market risk, the network topology of agents and market mechanisms. Our results demonstrate that, an uninformed fundamental trader (UFT) in a messy network is highly likely to suffer a major loss due to the significant price crash in a strongly UFT-dominated market (the informed traders only account for less than 20%), in which case the market efficiency is also negatively affected; Applying batch auctions helps reallocate the profits among the agents to reduce the information advantage between informed and uninformed traders, but it has limited effect on mitigating flash crashes; Building an information-sharing connection between agents is effective to reducing flash crashes and narrows the information advantage gap between informed and uninformed traders, but a complete network with full information exposure could mislead uninformed traders to make biased decisions. These findings generated by an agent-based simulation model give us insights into real-world financial markets under asymmetric information, and the framework proposed in this thesis can be extended for future studies of asymmetric-information markets.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Problem Statement and Motivaltion

On May 6, 2010, US stock markets suffered a rapidly dramatic decline followed by a recovery in about 30 minutes. Some major equity indices, like the S&P 500 and Dow Jones Industrial Average (DJIA) and Nasdaq Composite, went down over 4% compared with the last end price. DJIA collapsed by 998.5 points within minutes followed by a recovery of most parts of the loss. After a five-month investigation, the Securities and Exchange Commission (SEC) and the Commodity Futures Trading Commission (CFTC) blamed the crash on Waddell & Reed selling $4.1 billion futures in a single order hitting the market (SEC and CFTC, 2010). Such unusually large selling made most available buyers exit, while high-frequency trading accelerated the effect of the mutual funds' selling, amplifying the price declines that day.

However, several critics (CME Group, 2010; Lauricella, 2010) argued against the SEC/CFTC blaming a single order from Waddell & Reed as the main cause. For example, CME Group Inc. said that 75,000 contracts sold by W&R only represented 1.3% of the total E-mini volumes and less than 9% of the volume during the period in which the orders were executed (CME Group, 2010). Also, the public sentiment was reasonable based on this unusual order. There is no evidence to show the orders and the activity were illegal or inconsistent with market practices.

In April 2015, the US Department of Justice issued a warrant for the arrest of Navinder Singh Sarao, a London-based trader, and alleged that Sarao played a significant role in the Flash Crash that happened in May 2010. Sarao was accused of manipulating the market by using an automated program to generate large sell orders to push down prices and then cancelling these orders to buy at the lower market prices. However, a research paper written by Eric Aldrich, Joseph Grundfest and Gregory Laughlin found that Sarao should not be blamed as the culprit of the Flash Crash in 2010 (Aldrich et al., 2017). They said a more likely explanation was the unsettled market and the unusual orders placed by Waddell & Reed, making the market unstable.

Until now, the exact cause of the Flash Crash has been debated among scholars, but, in short, the Flash Crash in 2010 can be explained by unusual actions combined with the algorithm traders who were in lack of information but were driven by programs, amplifying the risks. Flash crashes are often difficult to predict and can be caused by various factors. While it is important to examine the causes of flash crashes, it is also worth investigating how markets behave after a flash crash and how we can defend and mitigate the effects of unusual actions on markets. Motivated by this background, this thesis aims to build a model and investigate the performance of all sectors in an information-asymmetric market, especially with respect to introducing shocks to create a flash crash. Another aim of this work is to advance an understanding of a market with the occurrence of flash crashes for the reality and to what directions the market can be improved.

1.2 Research Topics and Research Questions

Sarao spoofed the market by placing a huge number of sell orders and then quickly selling or amending the orders. Such malicious action released a fake signal to the market that the market prices were about to decline, and this did deceive some of the high-frequency traders to place sell orders as the algorithms indicated and pushed the price down. Therefore, Sarao had private market information that the prices would rebound after the spoofing activity. He consequently obtained substantial profits by strategically buying low and selling high. In essence, this situation can be summarised as follows:

due to unequal access to market information after the fake shock by spoofing, different traders respond by placing orders in different directions, ultimately leading to market price changes and differences in the profits among the traders. This thesis concerns the following preliminary hypotheses to investigate the research questions regarding spoofing and flash crash.

Hypothesis 1*: Spoofing can cause flash crashes or market fluctuations, resulting in altered equilibrium prices, and the spoofer profits from other traders.*

Spoofing leads to varying levels of knowledge on the market among different traders, which in turn affects their trading strategies. The market consists of various types of traders, and it is reasonable to assume that some traders are less susceptible to being deceived by spoofing, while others with limited information access may be the most affected. If spoofing affects the market, how does the impact manifest under different market structures? Additionally, it is reasonable to speculate that market volatility can influence the price fluctuations caused by spoofing, which can be also investigated.

Hypothesis 2*: Traders can learn from public information after spoofing and subsequently recover the market from the flash crashes.*

Some indices, after experiencing a flash crash, rebound shortly thereafter. For example, the Dow Jones Industrial Average (DJIA) dropped by almost 1,000 points within 10 minutes but then recovered almost 600 points in the next 30 minutes. It is reasonable to speculate that this is because some traders discovered they were deceived and subsequently altered their trading strategies, leading to a recovery in prices. These traders have the ability to learn from public information flows or order flows to adjust their behaviour. We are interested in whether traders who are spoofed but have the ability to learn can successfully mitigate their losses through learning, and how this learning behaviour affects the market.

Hypothesis 3*: Batch auctions can serve as an alternative to continuous double auctions to reduce the impact of adverse selection, such as flash crashes caused by spoofing, in the CDA market.*

Some studies found that batch auctions can mitigate adverse selection compared to continuous double auctions (Aquilina et al., 2020; Budish et al., 2014; Foucault et al.,

2017). In batch auctions where orders are gathered and cleared periodically, the information disclosure and traders' response speed could be different. Therefore, batch auctions may also have an impact on flash crashes caused by spoofing.

Hypothesis 4*: Building an information-sharing mechanism can mitigate flash crashes and reduce the losses of the traders who are spoofed.*

Let us consider a scenario where different traders can share information, which was also introduced in Goldstein et al. (2021), allowing learning traders to make inferences with more information on whether they are being deceived. Our interest lies in understanding the impact on both traders and the market when information-sharing connections are established among traders. However, different information-sharing mechanisms may result in diverse topological structures, and the counterparties with whom different traders share information may vary. Thus, how the asymmetry in information sharing affects individuals and the market also interests me to investigate more deeply.

Through computational simulations addressing these questions, we aim to gain a clearer understanding of markets with asymmetric information, providing valuable guidance on enhancing real-world marketplaces.

1.3 Methodology

Glosten (1987); Glosten and Milgrom (1985) proposed an information-sequential trading model as a protocol for an order-driven market to study information asymmetry. In the Glosten-Milgrom model, there are informed traders and uninformed traders interacting in a market with a risk-neutral market maker. The market maker provides liquidity but each time gets a premium, which is the bid-ask spread as the compensation to trade with the informed traders. The bid-ask spreads also reflect the adverse selection risk of the market maker. However, the Glosten-Milgrom model has sophisticated hierarchies and structures, so it was challenging to implement it until Das (2005); Das and Magdon-Ismail (2008) proposed an algorithm to convert the Glosten-Milgrom model into a computational model. Das presented the solutions of the market maker's quotes and allowed them to update by iterations. This leads the way of agent-based modelling in studies of market

micro-structure for further studies. Inspired by the Glosten-Milgrom model and Das's extensions, we built, implemented and assessed an agent-based model of an extended information-sequential trading framework following the methodology of Das's model.

In this model, we have extended and redefined the compositions of all agents, including fundamental traders, who place orders at preferred prices; zero-intelligence traders, who place orders randomly like some individual traders; and a market maker, who provides liquidity. To create a possible flash crash by an unusual event, we add an exogenous fake shock into the system to disturb some traders' actions. Fundamental traders can be divided into informed fundamental traders and uninformed fundamental traders, depending on whether they are able to infer the shock is fake without any extra information. Our extended model also allows the uninformed fundamental traders to learn by Bayesian inference from past order flows. To the end, by Monte-Carlo methods, we implement the model and examine the market with information asymmetry between informed and uninformed fundamental traders by sensitivity analysis and graphical analysis in the following aspects: the market structure, market risk, the network topology of agents and market mechanisms.

1.4 Thesis Structure

Following the introduction in Chapter 1 and a review of relevant literature in Chapter 2, the remaining chapters are summarized as follows:

Chapter 3 provides a comprehensive introduction to the fundamental components of the flash crash model proposed in this thesis. It covers essential topics such as the continuous double auction and batch auction mechanisms, the process of price determination, and the foundation models – Glosten-Milgrom mode and Das's model. Additionally, this chapter outlines the basic framework of the agent-based model designed for studying flash crashes, and it specifies the essential components required for constructing the model and experimental analyses. Chapter 3 establishes the foundation upon which the simulations and experiments in the subsequent chapters are built, forming the cornerstone of the entire research.

Chapter 4 builds, implements and assesses a basic model in which informed fundamental traders, uninformed fundamental traders, zero-intelligence traders, and a market maker interact under the fake shock in a continuous auction market. This basic model presents how the market differs among various market structures and volatility parameters.

In the real world, during the flash crashes in 2010, many high-frequency traders found out they had been deceived and suffered losses by the fake orders soon after being spoofed. They, therefore, changed their trading strategies in time to stop their losses. Accordingly, Chapter 5 extends the framework of the basic model in Chapter 4 by adding learning abilities to the uninformed fundamental traders to alter trading strategies upon past order flows. We show that most of the intelligent uninformed fundamental traders are able to make a correct inference on whether being deceived then pulling up the price after a crash, but they still suffer significant losses when uninformed fundamental traders dominate the market.

Chapter 6 considers an alternative market mechanism batch auction, where a few orders are batched to be executed simultaneously instead of sequentially, one by one. Apart from the model in Chapter 5 applied in a continuous auction, we rebuild and implement the model with batch auction mechanisms where batch sizes range over $\{2,4,6\}$. We have found that the batch auction does reduce the fake shock impact on the market from the aspect of the crash size. However, a batch auction can also narrow the distribution of agents' profits, which may lower the incentive of the fundamental traders to trade.

Chapter 7 considers two further network topologies for agents' sharing information. Compared with the model in Chapter 5, where all agents are independent of each other, connections between two agents are built in two different networks: a cycle network, where each agent only has the connection with the neighbours, and a complete network, where each agent has full connections with all other agents. The agents with connections can share information about their actions each round. In the cycle network, we find that uninformed fundamental traders prefer to share information with informed fundamental traders. Also, compared to the messy networks, complete and cycle networks can significantly reduce the impact of flash crashes caused by the fake shock. Still, these two network structures

lower the accuracy of uninformed fundamental traders' inferences, thereby increasing the variation in the distribution of their profits.

Chapter 8 summarises the analysis of the previous chapters and provides guidance for real-world policy based on the experimental results found in this research. We have also considered the limitations of this research and what future work may be possible. This thesis extends the theoretical framework of information asymmetric markets from many aspects, but due to the necessary simplification of the model, it currently only implements short-term simulations in a single market with a single instrument. However, this framework could pave the road for further studies in the future.

1.5 Contribution

Our results contribute to the growing literature on using ABMs to study the complex dynamics of financial systems. The complicated phenomena caused by interactions of heterogeneous, boundedly-rational, autonomous agents are not predictable simply from knowledge of their individual behaviour. We built, implemented, and assessed an agent-based model of an extended information-sequential trading framework inspired by Das's model and the Glosten-Milgrom model. The artificial market with information asymmetry is investigated under the following aspects: the market structure, market risk, the network topology of agents, and market mechanisms.

Our model provides a basic framework which can be easily extended for future studies of asymmetric-information markets. For example, this thesis extends it in the following two ways: 1. the difference in the topology of agents linked with each other due to information sharing, and 2. the change in market mechanisms from continuous auctions to batch auctions. This model also gives insights into future studies on dynamic game-theoretic models of asymmetric-information markets to study the equilibrium of more complex systems.

Chapter 2

Literature Review

2.1 Double Auction

An auction market is a place where buyers and sellers submit offers to compete to make a trade. There are various types and categories of auctions. Depending on the market structure, the three typical categories are forward auction, reverse auction, and double auction. The forward auction is the most common type in which a single seller offers items and waits for the bidder who offers the highest price, and this bidder wins the auction. The reverse auction is the case that, as the name suggests, one buyer expects the lowest price of the items offered by many sellers. Forward auction and reverse auction are both examples of one-sided auctions.

As opposed to the one-side auction, a double auction has multiple buyers and sellers in the market. The double auction mechanism is commonly used in an order-driven market. The double auction allows buyers and sellers to participate in transactions by submitting orders simultaneously. An order generally contains two elements: price (quote) and quantity. The orders submitted by sellers are called ask (sell) orders, while the ones from buyers are called bid (buy) orders. A transaction occurs only when the bid price exceeds or is equal to the ask price. Given the supply side of the ask orders and the demand side of the bid orders, a double auction is typified by (1) how to match asks and bids, and (2) how to determine the trade price of each ask-bid pair. A possible mechanism could be the average mechanism, where the buyers and sellers are ranked in the natural order to find

the breakeven pair[1]. The trade price for all matched traders can then be calculated as the average of the quotes of the two pairs, regardless of time priority (Brewer et al., 2002).

To avoid confusing the market, the auctioneer, acting like a broker, often plays a role in the double auction market to ensure a balance between buyers and sellers. An auctioneer hosts the market with a guide price to make all participants interested in offering. Also, the auctioneers should keep the market active to avoid "dead spots", which requires them to set prices and spreads reasonably (Cassady, 1967). The auctioneer sets the prices paid to the sellers and the prices paid from the buyers and earns the price difference as the profits. In the securities market, the auctioneer is also called the market maker.

2.1.1 Continuous Double Auctions

Depending on whether the orders are matched continuously or in a batch, two types of double auction are designed, called **Continuous Double Auction (CDA)** and **Batch Auction (BA)**. Continuous double auctions and batch auctions are the two of the most widely-used auction mechanisms in stock exchanges. The market always opens with the call auction, where all submitted orders are batched together until the execution point. In a continuous double auction market, transactions take place as soon as the current bid price equals or exceeds the ask price. Priority in trading is determined by a combination of offering a better price and the timing of the order. Participants who either offer a more favourable price or submit their orders earlier have an advantage in executing their trades. With the random activities of traders, the market price is driven to be volatile, bringing about the complexity and uncertainty of the CDA market. CDA markets are difficult to study but applied broadly in the real world; thus many scholars are investigating the CDA market from several aspects, such as auction theory, mechanism design and trading strategy, for decades.

In some earlier studies, researchers have studied simplified double auction markets mainly through game equilibrium analysis. Chatterjee and Samuelson (1983) first applied

[1]The buyers can be arranged in descending order of their bids, denoted as $b_1 \geq b_2 \geq \ldots \geq b_n$. Similarly, we organize sellers in ascending order of their bids, represented as $s_1 \leq s_2 \leq \ldots \leq s_n$. Then, we can identify the largest index k where $b_k \geq s_k$; this index is referred to as the "breakeven index", and we call (b_k, s_k) "breakeven pair".

Bayesian Nash Equilibrium (BNE) to the double auction. Their theoretical bargain model comes with a simple case where only one buyer and a seller trade with each other with incomplete information. Myerson and Satterthwaite (1983) demonstrated that linear[2] strategy equilibrium produces a higher net surplus level than other Bayesian Nash equilibrium strategies. Friedman (1991) simplified double auction as a Bertrand game process, where traders compete in prices leading to the equilibrium prices convergence to marginal costs, and concluded that there is an optimal auction strategy for both parties. McAfee (1992) studied the extended double auction model with m buyers and m sellers and demonstrated that the dominant strategy for participants is to place bids or offers at prices corresponding to their true cost evaluations.

The game analysis of double auctions is based on the assumption of "perfect rationality" and the deduction of rigorous mathematics. Under several traditional economic models, full rationality or perfect rationality, complete information and homogeneous traders are assumed, in order to find an equilibrium for a decision-making problem (Gigerenzer and Selten, 2002). However, there exist "cognitive limitations" for different entities, "information asymmetry" in the whole system and other restrictions so that agents are not able to always make an optimal decision. In another word, the assumption of perfect rationality is not always true in the real world. To revise the notion of "perfect rationality", Simon (1997) proposed the bounded rationality theorem as an alternative basis for decision-making problems. Bounded rationality allows people to make decisions regarding the knowledge and information they have, only if the decision is satisfactory. Therefore, bounded rationality could serve as a good basis for experimental studies in double auction markets.

Challenging the traditional view, Vernon Smith, who was awarded the 2002 Nobel Prize, initially built an experimental model to study the empirical economics and market mechanism (Smith, 1962, 1982, 1994). Through a series of laboratory experiments, he claimed that trade prices can also converge to the equilibrium with bounded-rational traders, which laid the foundation for Agent-based modelling and simulation (ABMS) in

[2]Consider a scenario where two bidders want to buy an object through an auction, and each submits a bid $b_i \geq 0$ to win. Assume that the value of the object for bidder i is v_i, and then "linear" here means that the bidding strategy b is a linear function of v_i: $b_i(v_i) = a + cv_i$, where a and c are constants.

economics and finance. Gode and Sunder (1993) defined the zero-intelligence (ZI) trader as an agent that "has no intelligence, does not seek or maximize profits, and does not observe, remember or learn." (p. 121). They proposed two types of ZI trading strategies: Zero-Intelligence unconstrained (ZIU) and Zero-Intelligence with Constraint (ZIC). The assumptions of the two types of zero-intelligence traders are not intended to create a descriptive model of individual behaviour; instead, they introduce significant cognitive limitations that are expected to influence market performance. Assume the minimum and maximum quotes allowed to be submitted are $Q_{\min}$ and $Q_{\max}$, and ZIU agent's quote is uniformly selected between $Q_{\min}$ and $Q_{\max}$. With the unconstrained zero-intelligence assumption, agents accept all matched quotes from the counter-party even if the agent gets negative profits. ZIC strategy guarantees that the trader only accepts non-negative profits by letting buyers (sellers) submit quotes between their evaluation (cost) and $Q_{\min}$ ($Q_{\max}$), which is so-called bounded rationality. Gode and Sunder conducted simulations with those two strategies and compared with human traders. With the highly random and simple model, the simulations had human-like market dynamics, and they concluded that the allocation efficiency in CDA markets is mainly affected by the market discipline rather than traders' rationality, motivation, intelligence and learning. Gode and Sunder's surprising results have drawn much attention from both the economic and agent-based communities. However, Cliff and Bruten (1997) argued against Gode's conclusion and claimed that the price convergence and the high allocation efficiency in the Gode and Sunder's model were ensured by the model settings like the symmetry of the buyers and sellers. To improve the ZI strategy, Cliff and Bruten (1997) proposed a new Zero-Intelligence Plus (ZIP) strategy with self-adaption and learning ability. Das et al. (2001) studied the performance of ZIP robot strategy in a series of ExpEcon market experiments, where they applied a LOB-based market simulator to run six experiments. The surprising result shown in that report was that the ZIP strategy could significantly outperform human traders.

As computational power increases with the advance of technology, more scholars started to introduce machine learning into trading strategy design and studied high-frequency trading in double auction markets. Erev and Roth (1998) designed a reinforcement learning model to mimic human behaviours in games. The Roth-Erev model

allows traders to learn from the feedback from interacting with the trading mechanism. Such a model can be generally applied to disregard the type of market mechanism. Gjerstad and Dickhaut (1998) proposed another strategy which is based on a belief function to maximise traders' expected profits. GD strategy makes use of the historical market data and builds a probability model to figure out which quote is most possibly accepted. Das et al. (2001) also examined the GD model and similarly concluded that it was superior to human traders.

High-frequency trading, driven by high-performance computers with algorithms to seize fleeing profitable opportunities, has emerged in every essential exchange, including NYSE, NASDAQ and LSE, in the last decade. Some research on HFT illustrates the positive effects, like increasing the liquidity (Hendershott et al., 2009; Jovanovic and Menkveld, 2016) and enhancing the price discovery on the market (Brogaard et al., 2010). However, some opponents argue that HFT has weaknesses and should be replaced by batch auctions (Budish et al., 2015).

2.1.2 Batch Auctions

Compared to continuous-time trading in CDA markets, batch auctions collect (batches of) the orders over a certain time interval and execute when batching ends. Therefore, the orders over the batching period have the same matching priority, which can reduce the speed advantage and lead to competition upon price instead of speed (Budish et al., 2015). The batch auction, or call auction, was initially established as a pricing tool used when the market opens and closes, as London Stock Exchange calls the opening batch auction from 7.50am to 8.00am to determine the open price, and the closing batch auction is called from 4.30pm to 4.35pm to determine the close price (LSE, 2020). Madhavan (1992) initially studied the effect of the pooling periodic auctioning system, which was the prototype concept of the batch auction, and claimed that higher market efficiency was produced than that in the continuous order-driven trading market. Budish et al. (2014) found that batch auctions can be an alternative of the continuous double auction and initially built the theoretical model (referred to as BCS) of the frequent batch auction (FBA). They

found that the FBA can mitigate the latency arbitrage[3], which occurred in high-frequency trading, causing adverse selection risks for traders in CDA markets (Aquilina et al., 2020; Foucault et al., 2017). In FBA markets, all traders have the same speed priority to make trades to avoid auction sniping[4]. From then on, an increasing number of scholars have studied the advantages of batch auctions. Riccò and Wang (2020) examined the trading data on the Taiwan Stock Exchange in March 2020 and concluded that trading spreads in FBA markets are lower than those in CDA, but FBA decreases market liquidity as the trading volumes shrink. Zhang and Ibikunle (2021) investigated the cross-market latency arbitrage opportunities (LAOs) among different levels of FBA markets, and they found that short durations of LAOs, which implies a more powerful sniping, induce the use of FBA mechanism. Jagannathan (2022) claimed that FBA is a potential mechanism to reduce the flash crashes of stock prices only if a properly wide time interval, such as one minute, is set between auctions. Also, Wah et al. (2016), Wah and Wellman (2016), and Aldrich and López Vargas (2020) applied agent-based modelling to show that trading in FBA markets has lower transaction costs than in CDA markets.

Although many scholars believe that batch auctions have many advantages and can address some problems that arise in continuous double auctions, some argue that batch auctions also expose some problems to be discussed. Du and Zhu (2017) also built a theoretical model running a periodic auction with a finite number of traders and a diversity of private information. They found that the optimal auction frequency maximising the allocative efficiency depends on trade-offs between auction frequency and the strategic behaviour of trading. Haas et al. (2021) illustrated another trade-off in FBA: even though batch auctions can mitigate sniping attacks, informed traders have fewer incentives to trade, thereby leading to a decline in liquidity. Also, Budish et al. (2019) pointed out another

[3]Latency arbitrage is a high-frequency trading strategy that centers on front-running trading orders. Front-running occurs when a trader or trading algorithm exploits advanced knowledge of pending orders to make a profit. Traders and algorithms aim to capitalize on the small time delays (latency) that occur during order executions in financial markets. For example, a high-frequency trading algorithm rapidly places buy and sell orders ahead of a detected pending order, then benefits from the anticipated price movement resulting from the impending large order.

[4]"sniping" is a process in which a trader places a large number of "immediate-or-cancel" orders within milliseconds before stock prices realign. As the name implies, an immediate-or-cancel order can only be filled immediately or cancelled. Therefore, a high-frequency trader can stumble upon a large hidden order to earn by placing hundreds of "immediate-or-cancel" orders giving other traders no time to reflect. (Adrian, 2015)

reason why frequent batch auctions have not been widely discussed: the exchanges have little incentive to change the market design from CDA to FDA because they can profit by encouraging the trading firms in the arms race by selling high-speed technologies.

2.2 Market Micro-structure

The study of a double auction market inevitably involves examining the trading process and the price discovery in which market micro-structure is concerned. Garman (1976) initially introduced the term "market structure" to describe the trading behaviours and built a bridge linking the empirical statistics of the market with the aggregate trading activities. Precisely, as Cohen (1986) specified, market micro-structure analyzes the agents' interactions, market mechanism, the security price dynamics, bid-ask spreads, the shapes of the order book, how the market and trading process are affected, and so forth.

Easley and O'hara (1987) investigated how trade size affects the dealer market. They find that informed traders tending to trade large volumes bring adverse selection problems to the market maker.

2.2.1 Market Making

Market makers, also known as dealers, are key players in the micro-structure of financial markets, and their role is closely intertwined with how markets function. The market makers act as the auctioneer to balance the buy and sell sides and provide liquidity by quoting bid and ask prices. In the 1980s, a few scholars built essential frameworks to study the market micro-structure theoretically. Stoll (1978) developed a theoretical model for a market with a single dealer making the market for a single stock and illustrates the positive relation between illiquidity and volatility. Ho and Stoll (1981) extended Stoll's model and introduced the uncertainty effect. The extended model constructs a framework for the multi-period market-making strategy and price risks on inventory. Glosten and Milgrom (1985) described an information asymmetric dealer market where informed traders and uninformed traders trade with the market maker for a single asset. They found the market maker suffers the adverse selection risk and widens the bid-ask spreads to protect against

the risk and possible losses. Another important information-based model is the model of Kyle (1985). Kyle also introduced a strategical informed trader to maximize profits with private information. Kyle showed that the market prices will converge to the level reflecting full information after many trading rounds. Inspired by the Glosten-Milgrom model, Das (2005) designed a framework for the Bayesian learning market-making algorithm (BMM), which learns from the current market stats and the incoming orders to update the bid-ask spreads and quotes. Furthermore, there are several extensions based on Das's framework. Brahma et al. (2012) introduced a new adaptive Bayesian market maker for binary outcome markets, building upon the work of Das (2005) and Das and Magdon-Ismail (2008). In addition, Li and Das (2016) developed an agent-based model to explore the competition between a call market and a CDA market, incorporating the Bayesian market maker. My thesis also builds on the framework of the Glosten-Milgrom model (Glosten and Milgrom, 1985) and BMM algorithm (Das, 2005) with the introduction of Bayesian-style uninformed traders.

2.2.2 Price Formation

In market micro-structure research, the order flow plays a significant role in the study of price formation. Some research works proposed mathematical models as the benchmark for the limit order book. Traditionally, the stock price is often modelled as a Geometric Brownian Motion (GBM) by Bachelier (1900). GBM is a stochastic mathematical model commonly used in finance to describe the continuous and random evolution of asset prices over time. It captures the uncertainty and continuous nature of price movements in financial markets and has been empirically observed to offer a reasonable approximation of actual stock price behaviour (Reddy and Clinton, 2016). Bachelier illustrated that stock prices evolve as the following differential equation formula:

$$dS_t = \mu S_t dt + \sigma_t dW_t$$

where $S(t)$ stands for the stock price, μ is called drift, σ is the constant volatility of stock returns and W_t is a Wiener process. The solution can be obtained by Ito's Lemma (Itô,

1946) coming to $S(t) = S_0 \exp((\mu - \sigma^2/2)t + \sigma W_t)$. Akansu and Torun (2015) considered a discrete-time form of the GBM model where stock prices evolve as:

$$S(n) = S(n-1) + \mu + \sigma z(n)$$

where $S(n)$ is the log price, n is the n-th time step, μ is a constant referring to the drift, and $z(n)$ is normally distributed: $z(n) \sim \mathcal{N}(0, \sigma)$. With reference to the Akansu-Torun model (Akansu and Torun, 2015), in the short-term flash crash model presented in this thesis, the model has been modified by redefining μ as a shock effect to determine the actual stock prices:

$$\mu = \begin{cases} 0, & \text{if no shock comes} \\ r, & \text{if a shock comes} \end{cases} \tag{2.1}$$

where r refers to the shock effect causing the stock price changing. For example, if $r = -1$, then the stock price of the next time step is reduced by 1 unit with the addition of a random disturbance term.

In double auction markets, both buyers and sellers will set price limitations for their quotes, i.e., the highest accepted sell price for buyers or the lowest accepted buy price for sellers. Those orders that are set with price limitations are called limit orders. The limit order book (LOB) is a pair of lists containing buy limit orders and sell limit orders, respectively, of the asset. A limit order is not always matched with an order from the other side, and then all unmatched orders will be listed in the limit order book. In the limit order book, orders are aggregated by a price and time priority rule, and bid orders are sorted by price in descending order of price, while ask orders are sorted by price in ascending order of price. Therefore, in a static state, the ask prices are always higher than the bid prices in the limit order book. When a new order comes into the marker, the order will seek an order that exists in the limit order book to achieve the transaction. The continuous stream of buy and sell orders placed by traders is the order flow, it provides a snapshot of how various traders are engaging in the market, offering the participants insight into the present market conditions and price trends.

2.2.3 Bid-ask Spreads

Analyzing the factors affecting bid-ask spreads has been of interest in the time-series analysis of micro-structure models. Harris (1995) claimed that three factors, order processing, adverse selection and inventory cost[5], are the most important components considered in bid-ask spread models. As defined by the difference between the best ask and bid prices, it is straightforward that the bid-ask spreads can reflect the order processing costs to some degree. Some researchers, e.g. Amihud and Mendelson (1980) and Ho and Stoll (1981, 1983), claimed that the inventory-holding cost is another component of bid-ask spreads. In addition, Glosten and Milgrom (1985) have focused on the market maker's adverse selection costs with an information-based model.

There have been several statistical models to estimate the components of the bid-ask spreads. Roll (1984) was a pioneer of building a bid-ask spread model where it represents that the spreads rely on the covariance of the transaction prices. However, this model only considers the order processing cost as a component of bid-ask spreads, and the expected return of the security is assumed to be constant. Glosten (1987) argued that Roll's measure would underestimate the true spread with the inclusion of the adverse selection component, and the covariance is always time-varying, leading to a downward bias of the spread measure. Glosten and Harris (1988) and Madhavan, Richardson, and Roomans (1997, hereafter MRR) built linear models with trade indicators to estimate the components of the bid-ask spreads on the atheoretical model and empirical model, respectively. In 1989, Stoll (1989) proposed a more general model with an extension that includes adverse selection costs and inventory holding costs with empirical tests. Lin et al. (1995) introduced the trade size as an effect on the adverse selection component of the bid-ask spread. Furthermore, Hasbrouck (1988, 1991) built an autoregressive framework to model the time series of quotes and derive the sources of the bid-ask spreads. In this thesis, to estimate the power

[5]Order processing costs are commonly defined as the costs of arranging, recording, and clearing transactions. Higher order processing costs can lead to wider spreads as traders seek to cover their expenses through larger spreads. Adverse selection cost represents the cost or risk incurred by market makers when trading with informed traders or counterparties with superior information or market insight. Consequently, spreads could be widened to offset this risk. Inventory costs arise when market makers or dealers hold positions in assets. If they accumulate inventory, they may widen spreads to mitigate the risks associated with price fluctuations while holding these assets.

of order processing and adverse selection on the bid-ask spreads, we apply the framework proposed in the Glosten-Harris model and MRR model.

2.3 High-Frequency Trading

High-frequency trading is typically a kind of algorithmic trading with high-speed automated programs for trading in an electronic market (Jones, 2013; US Securities and Exchange Comission, 2010). High-frequency traders or firms do not typically invest a massive amount of capital or keep their positions overnight to reduce the risk of overnight positions, but they trade with high volumes within an extremely short period to earn a tiny profit per trade (Goldstein et al., 2014). High-frequency traders often act as market makers that provide more liquidity by adding more limit orders and narrowing the ask-bid spreads to reduce the transaction costs of other participants (Jones, 2013).

High-frequency trading (HFT) has seen a remarkable rise in prominence worldwide in recent years (Zaharudin et al., 2022). In the United States, it escalated from approximately 20% in 2005 to a peak of 60% in 2009, maintaining stability at 52% since 2018. Similarly, in Europe, HFT's share of total equity trading surged to around 40% in 2010 and settled at approximately 35% in 2014. In Australia, ASIC's data revealed that HFT constituted roughly 27% of equity market turnover in specific securities from January 2012 to March 2015. With the rise of high-frequency trading, there has been a debate among all market participants. On the one hand, high-frequency traders, such as market makers, make the market more active and liquid by reducing price inefficiencies and creating more trades (Hoffmann, 2014). Another benefit of HFT is that it can maintain asset prices to be stable and consistent. Chaboud et al. (2014) discovered that algorithmic traders can immediately detect arbitrage opportunities among different currencies in the foreign exchange market and then adjust the rates to the levels of no arbitrage. However, apart from market making, some high-frequency traders can act as speculative traders (Budish et al., 2015). Such a type of high-frequency trader could have the opposite effect on the markets. Breckenfelder (2019) applied the difference-in-differences analysis by comparing an environment with no high-frequency trading competition and an environment with high-

frequency trading competition. The results showed that speculative activities increase by 11% when HFTs compete. Also, speculative trading reduces market liquidity but increases intraday volatility. Jones (2013) also claimed that high-frequency trading introduces adverse selection and potentially harms individual traders by leveraging its speed advantage to anticipate and execute orders before slower participants. This advantage can lead to HFT firms consistently obtaining more favourable prices, while individual traders, often unaware of rapidly changing market conditions, may experience less advantageous executions, ultimately resulting in losses.

Since high-frequency trading has increasingly developed these years, it causes the arms race on trading on speed (Budish et al., 2015). In the continuous double auction market, many traders probably receive the same positive signal simultaneously. However, only the fastest traders can earn the most because the orders in continuous double auctions match in serial. Aquilina et al. (2020) claimed that the fastest firms, which only offer 42% of liquidity of the market, can take 80% of liquidity in races. For instance, if traders see the future market experiencing a jump in a short time, which suggests a potential increase in the stock market soon, a high-frequency trader/market maker would withdraw the old sell quotes and set some new ones with higher sell prices to widen the spread. If the trader trades fast enough, the trader's old buy orders can be executed before the stock price is adjusted, and the trader can make more profits from the other traders by widening the spread with the short-time fluctuation, which is called "sniping" or "latency arbitrage" (Budish et al., 2015; Haas et al., 2021). While Budish et al. (2015) argued that the continuous double auction should be blamed for the arms race in high-frequency trading. They claimed that the arm race profits should be a constant value instead of a high-value reward for the arms race winner. They also showed that frequent batch auctions could eliminate latency arbitrage and the arms race in high-frequency trading.

2.4 The Flash Crash

On 6th May 2010, U.S. stock and futures markets simultaneously experienced a rapidly dramatic price fall within 30 minutes. Some major indices, like S&P 500, DJIA and

NASDAQ Composite, went down over 4% compared with the end price of the previous trading day. As far as individual securities, some suffered a more significant decline before rebounding. For instance, Apple went down by 22%, and Procter & Gamble dropped over 37% from around $60 to approximately $39 per share, while Accenture Plc fell from $40.13 to even 1 cent before it rebounded to $39.57, showcasing the rapid and severe market disruptions caused by high-frequency trading algorithms and swift order executions (Wahba, 2010).

Such a remarkable phenomenon, called the flash crash, occurred again on 5 February 2018, with a greater impact on the market. Stocks had been declining throughout the day, and shortly before 3:00pm, the Dow Jones Industrial Average experienced an 800-point drop within a 10-minute span, reaching a low of 1,579 points down before a slight rebound at the closing bell. Throughout that day, the Dow Jones Industrial Average experienced a significant 4.6% decline, with a total drop of 1,175 points, marking the largest single-day point decline in history up to that point (Wiener-Bronner, 2018). As those two significant flash crashes occurred, it drew increasing attention to investigate the reasons for the flash crash. CFTC and SEC (2010) claims that the Flash Crash in 2010 originated in a large number of selling contracts of E-mini S&P 500 Futures. The report interprets the Flash Crash as the impact of high-frequency traders. High-frequency traders, equipped with trading algorithms, automatically absorbed a large number of initial future orders. However, restricted with the long position limit, the high-frequency traders began selling orders when their constraint was reached, leading to prices decreasing. A price decline also scared the fundamental traders and market makers, so they started to exit consecutively, reducing the market liquidity. Eventually, the high-frequency traders, driven by their algorithms, traded with each other, further driving the price down. Kirilenko et al. (2017) applies the CFTC data to track transaction-level data of future contracts and shows that the algorithm trading and order cancellation are large contributors to the flash crash. Paddrik et al. (2012) built a zero-intelligence agent-based model to simulate the flash crash from the perspective of market micro-structure. To examine whether and how different categories and styles of traders contribute to flash-crash-type events, Paddrik et al. designed six categories of trader types: fundamental buyers and sellers, market makers, opportunistic traders,

high-frequency traders and small traders. These categories were determined based on four key variables: trade speed, position limits, order price selection, and order size. The experiment illustrates that high-frequency traders and large sell orders are responsible for the flash crash.

Apart from putting the blame on the high-frequency traders, there are several scholars interested in how the market micro-structure may affect the flash crash. In real financial markets, varieties of financial instruments have been created and invested, and investors will not put all their eggs in a single bucket. Allen and Babus (2009) illustrated that modern financial systems are not isolated. Network analysis helps us comprehend financial systems and interpret certain financial phenomena. Paulin et al. (2019) proposed a hybrid micro-macro agent-based model to characterize system risks in terms not only of system stability but also of the distress propagation speed. From the micro-aspect, this model simulates a single asset and a single fund based on the market micro-structure model in (Paddrik et al., 2012). They found the occurrence of a flash crash strongly depends on the price impact of the distressed selling agent's orders. From the macro-aspect, they identified that the behaviours of the algorithm traders, leverage management, and the network topology strongly affect systemic risks. To amplify their profits, funds often leverage the capital invested by traders to borrow from a bank or brokerage. However, when the price of the portfolio goes down, the funds will suffer a larger loss due to this margin policy. With the connection of the assets in a fund, if a fund is forced to liquidate, it results in an amplification mechanism (margin spiral) where further selling actions taken by other funds further depress prices.

2.5 Artificial Financial Markets

Due to the complexity of financial markets, an applicable method of analysing how trading activities affect the market is to study agent-based artificial markets (Chen et al., 2012; Chiarella et al., 2009; LeBaron, 2006). Agent-based models (ABMs), also called multi-agent models, have recently drawn increasing attention from the perspective of computational simulation in financial markets. Different from traditional economic mod-

elling, ABMs are simulations driven by programmed agents under certain mechanisms to simulate the actions and interactions of the participants. Each agent is designed with specific acting rules (e.g. buying when the price is underestimated and selling when the price is overestimated) and interacts with other agents (e.g. a buyer submits a new order to the exchange). With the rules and interactions, the system is autonomous over time, and the dynamics of the system runs can be observed. At the same time, the agents' behaviours can also be examined to see how they affect and drive the market. Like the asset prices moving discretely, an agent-based model in finance is often implemented in a discrete-time form with explicit mechanisms in the simulation. With different agents gaming with each other, the simulation is like a turn-based game, where the agents decide when they arrive and what actions to take.

With the bounded rationality theorem, an agent's behaviour could be simply defined under certain programmable rules without the multiple restrictive assumptions of mainstream theoretical models. As the complexity of the model structure increases, the status of the whole system could be dynamic over time, while pure mathematical or economic analysis is limited in such a dynamic sequential problem. Farmer and Foley (2009) and Kirman (2010) propose that the traditional economic and financial models could not work well for extreme situations, like financial crises, bubbles or crashes, which indeed are barely researched under a classically economic framework. However, agent-based models can specify complex features and dynamics to emulate and explain unstable and unusual occurrences in the real world (Jackson, 2007).

During the past few years, as an increasing number of scholars have attached the importance of agent-based modelling in the study of economic-finance networks. In agent-based modelling of financial markets, researchers program these software agents with specific rules, strategies, and decision-making processes to simulate the behaviour of actual market participants. These simulated environments, known as artificial financial markets, are used for studying and analysing the complex interactions, behaviours, and emergent properties that occur in real financial markets (LeBaron, 2001). These artificial markets are designed to achieve different goals, ranging from market prediction (Wolfers and Zitzewitz, 2004) to market-making mechanism design (Das, 2005; Das and Magdon-Ismail, 2008)

and to construct warning systems of crises (Fievet and Sornette, 2018). Fagiolo and Roventini (2016) applied the agent-based model as a competitive alternative to DSGE models; agent-based models are regarded as a more effective method to study trading activities in financial markets than are classical models with the efficient-market hypothesis (EMH) assumptions (Franke and Westerhoff, 2012; Leal et al., 2016).

While there has been a growing focus on Agent-Based Modelling (ABM) approaches, it is unavoidable that some critical discussions revolve around such kind of emerging methodologies in economics and finance. Some critics argue that agent-based modelling has many more limitations compared with mainstream models: Critics argue that ABMs are often too narrowly focused, particularly in financial markets, where they primarily explain specific observed phenomena without addressing broader financial research questions. Furthermore, there is a lack of clarity regarding how ABMs can be effectively used for policymaking (Iori and Porter, 2012). Although agent-based modelling has gained remarkable success in financial markets, there is less generality for variant situations as the traditional financial models apply. Farmer and Foley (2009) acknowledged that there is a need for more comprehensive models that include more markets and economic factors to model large parts of an economy. Haldane and Turrell (2019) compared ABMs with a traditional method, the Dynamic Stochastic General Equilibrium (DSGE) model, in macroeconomics. DSGE models have core assumptions, such as rational expectations, which make them analytical but potentially overly restrictive. In contrast, ABMs offer a flexible method for modelling complex issues with various agents and assumptions. However, this flexibility can result in "bespoke", criticised as "black box" models that may lack analytical explanation. Also, the absence of a representative ABM framework poses challenges in establishing a standardised macroeconomic framework using ABMs.

Chapter 3

Fundamentals of Flash Crash Model

3.1 Introduction

Flash crashes are a real-world issue with systemic risk potential, particularly after the notorious 2010 Flash Crash. Despite increased attention to flash crashes, they remain inadequately understood. The research aim of this thesis stems from an incident during the 2010 Flash Crash when Sarao was accused of spoofing the market. He employed an automated program to generate large sell orders, driving down prices, and then cancelled these orders to buy at the lower market prices. Although prices later rebounded, some traders incurred substantial losses during that flash crash. This context inspired my investigation into spoofing, flash crashes, and their impacts on markets. Therefore, we have proposed two primary hypotheses regarding the research questions:

- **Hypothesis 1:** Spoofing can cause flash crashes or market fluctuations, resulting in altered equilibrium prices, and the spoofer profits from other traders.
- **Hypothesis 2:** Traders can learn from public information to stop losses after spoofing and subsequently recover the market from flash crashes.

The Glosten-Milgrom model and Das's model provide a valuable foundation for studying individual and systemic behaviour in incomplete markets. Inspired by these models, we have extended and designed an innovative model to address the questions posed in this thesis. In addition, some studies suggest that batch auctions could serve as an

alternative mechanism to reduce risk in incomplete markets, including addressing adverse selection issues like flash crashes caused by spoofing (Aquilina et al., 2020; Budish et al., 2014; Foucault et al., 2017). Also, Goldstein et al. (2021) introduces the concept of an information-sharing mechanism for an in-depth investigation of incomplete markets..

We aim to explore whether the application of batch auctions and information sharing within the structure of this thesis can yield similar results, providing a better understanding of the impacts of flash crashes. Therefore, we have put forward two additional hypotheses:

- **Hypothesis 3:** Batch auctions can serve as an alternative to continuous double auctions to reduce the impact of adverse selection, such as flash crashes caused by spoofing, compared to the continuous double auction market.
- **Hypothesis 4:** Building an information-sharing mechanism can mitigate flash crashes and reduce the losses of the traders who are spoofed.

To address these research questions, we have further designed extended models, complete with implementation and analysis. All experimental chapters, Chapters 4-7, will follow the structure depicted in Fig. 3.1:

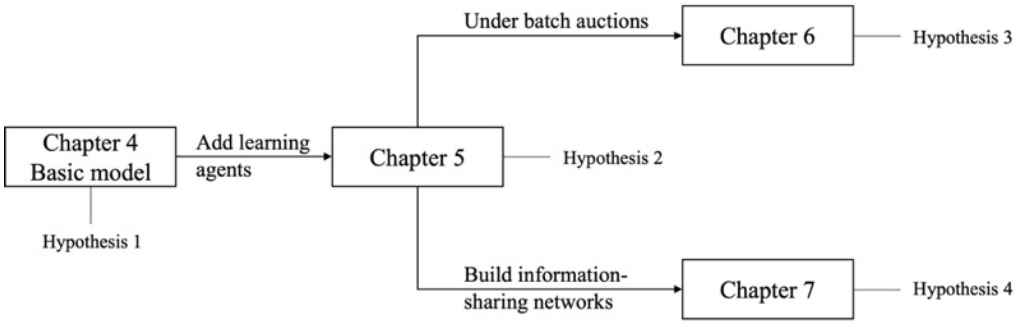


Figure 3.1 The Structure of Experimental Chapters

For better model construction, this chapter offers a comprehensive and detailed introduction to the core components of the flash crash model presented in this thesis. It introduces the market mechanisms, including the continuous double auction and batch auction mechanisms, and illustrates how trades occur under such mechanisms. Furthermore, it specifies the foundational models for incomplete markets, namely the Glosten-Milgrom

model and Das's model, providing detailed explanations of their key settings and frameworks that will be referenced in this thesis. Additionally, this chapter introduces the fundamental framework of the basic flash crash model applied and examined in Chapter 4. By defining and detailing the essential metrics required for our experimental analyses, it provides us with a solid understanding of the tools and methodologies that will be employed to investigate the dynamics of flash crashes and market behaviour.

3.2 Market Mechanisms

3.2.1 Continuous Double Auctions

Definition

A continuous double auction (CDA) is a market mechanism allowing buyers and sellers to submit bids and asks continuously, and a trade can occur once there is a matching pair. This contrasts with a batch or call auction where orders are accumulated over a fixed period and then matched at a single price at the end of the period. Prices are determined by the interaction of supply and demand and can fluctuate frequently during trading periods as new orders are placed with old orders being filled or cancelled. The New York Stock Exchange (NYSE) operates primarily under a continuous double auction mechanism.

Limit Order Book

A Limit Order Book (LOB) is a fundamental component of modern financial markets, representing the current state of supply and demand for an underlying asset. It holds the list of outstanding buy and sell orders in a market.

In a continuous double auction market, buyers and sellers submit orders to trade, specifying the prices at which they are willing to buy or sell. There are two types of orders: market orders and limit orders. A **market order** is a straightforward type of order where traders offer to buy or sell at the best prices. Market orders are typically executed immediately, provided there are willing counterparties, and they remove liquidity from the market, as they match with existing orders on the opposite side of the transaction.

Contrasting with a market order, a **limit order** specifies both the direction (buy or sell) and the price at which a trader is willing to transact. For example, a trader placing a limit buy order agrees to purchase at the specified price or lower, while a limit sell order will execute at the specified price or higher. Therefore, each side can be ordered like:

Bid Side: $B_{t1} > B_{t2} > B_{t3} > ...$

Ask Side: $A_{t1} < A_{t2} < A_{t3} < ...$

Consequently, there is no guarantee that a limit order will be executed at all, as there could be some orders with higher priorities or there may be no suitable counter-party orders. Limit orders that are not executed immediately are stored in the limit order book. The limit order book records and displays all the outstanding bids and asks. They are ranked by price levels and arrival time, with the highest bid and the lowest ask arriving earlier visible at the top.

If there is a trade occurring, the trade price is determined by the prices of the matched orders. Typically, when a new order matches with an existing order, the trade is executed at the price of the existing order.

Limit Order Book Dynamics and Trade Execution Illustrations

Here are examples showing the trade executions in the continuous double auction:

We start with a limit order book that has an existing best ask at £101.50 for 3 shares and a best bid at £100.00 for 4 shares. Time 0 stands for the initial state of the limit order book.

Limit Order Book – Time 0			
	Shares	Price	Time
Asks	3	£ 102.00	0
	2	£ 101.50	0
Bids	4	£ 100.00	0

Figure 3.2 The Limit Order Book at Time 0

At time 1, a new limit buy order is placed at a price of £101.00 for 2 shares. Traders and market makers examine the limit order book to identify a suitable match for this order.

Upon review, it becomes evident that there are no existing sell orders at or below £101.00. As a result, the new buy order cannot be matched with any existing sell orders and is thus positioned at the top of the buy limit order book, becoming the new best bid, given its higher price compared to the existing best bid.

Limit Order Book - Time 1			
	Shares	Price	Time
Asks	3	£ 102.00	0
	2	£ 101.50	0
Bids	2	£ 101.00	1
	4	£ 100.00	0

Figure 3.3 The Limit Order Book at Time 1

The limit order book is now populated with the unmatched orders at time 0 with the best ask still at £101.50 for 2 shares and the new best bid at £101.00 for 2 shares.

At time 2, another buy order is placed, this time a limit sell order, with an ask of £101.00 for 3 shares, as shown in Fig. 3.4. Once again, market participants review the limit order book to find a suitable match.

Limit Order Book - Time 2 (before clearing)			
	Shares	Price	Time
Asks	3	£ 102.00	0
	2	£ 101.50	0
	3	£ 101.00	2
Bids	2	£ 101.00	1
	4	£ 100.00	0

Figure 3.4 The Limit Order Book at Time 2 (Before Clearing)

Limit Order Book – Time 2 (after clearing)			
	Shares	Price	Time
Asks	3	£ 102.00	0
	2	£ 101.50	0
	1	£ 101.00	2
Bids	4	£ 100.00	0

Figure 3.5 The Limit Order Book at Time 2 (After Clearing)

The new limit sell order is reviewed, and it is identified that there is a matching buy order at £101.00. Therefore, a transaction occurs immediately, and 2 shares are traded at £101.00. Since the incoming sell order quantity is more than the quantity of the best bid, the incoming sell order becomes the best ask but with a reduced quantity of 1 share, as shown in Fig. 3.5.

This detailed illustration shows the dynamics of the limit order book in a continuous double auction market and how new orders can interact with existing ones, either being stored in the book or executing transactions when matching orders are found.

3.2.2 Batch Auctions

Definition

Compared with continuous double actions, batch auctions are a distinct market mechanism where orders are accumulated over a fixed period and then matched and processed at specific times. All buy and sell orders are matched and executed at a single price, which is determined by finding the price that maximises the trading volume of the batch. Therefore, trades are not continuous but occur at discrete time intervals. For instance, a batch auction might occur every second, minute, or even once a day. Based on such features, it is less susceptible to price manipulation and high-frequency trading strategies, leading to a more stable and fair trading environment. IEX, the Investors Exchange, applies a form of batch auction known as the Discretionary Peg (D-Peg) Order to prevent predatory trading (IEX Exchange, 2023).

Limit Order Book

In batch auction markets, the behaviour of the limit order book is different from that in CDA markets, primarily characterised by the accumulation of orders and their subsequent periodic clearing. Despite this difference, the order priority within the limit order book remains consistent with that of continuous double auctions.

In both types of auctions, the execution of orders generally follows the price-time priority. This means that orders with the best prices—highest for buys and lowest for

sells—are executed first. When multiple orders share the same price, the one submitted earliest has the priority. However, the different feature of batch auctions is that all orders placed during a single batch period are treated with equal time priority.

Given the clearing period T, during the k-th batch interval $[(k-1)T, kT]$, orders (bids and asks) are accumulated in the limit order book but are not immediately matched or executed. The accumulation batch can be represented as $F(A,B,T,k) = \{(A_t, B_t | t \in [(k-1)T, kT])\}$, where A_t and B_t denote the accumulated ask and bid orders during the k-th batch interval in the limit order book..

At the end of the k-th interval $[(k-1)T, kT]$, all accumulated orders are analysed to determine a single clearing price p^* at which the maximum number of buy and sell orders can be matched. This can be derived to calculate the equilibrium price where market demand meets supply.

To determine the clearing price p^*, we introduce the Marshallian path (Brewer et al., 2002), which, in a double auction market, is illustrated by sorting buyers' (demand side) valuations from highest to lowest and sellers' (supply side) costs from lowest to highest. Then the market starts pairing the highest buyer with the lowest seller and continues pairing in descending order for buyers and ascending order for sellers. The equilibrium price can be determined as the price at which the last pair of orders match, and the equilibrium quantity is represented by the number of trades that have been made. Fig. 3.6 illustrates how the batch auction works on a single batch:

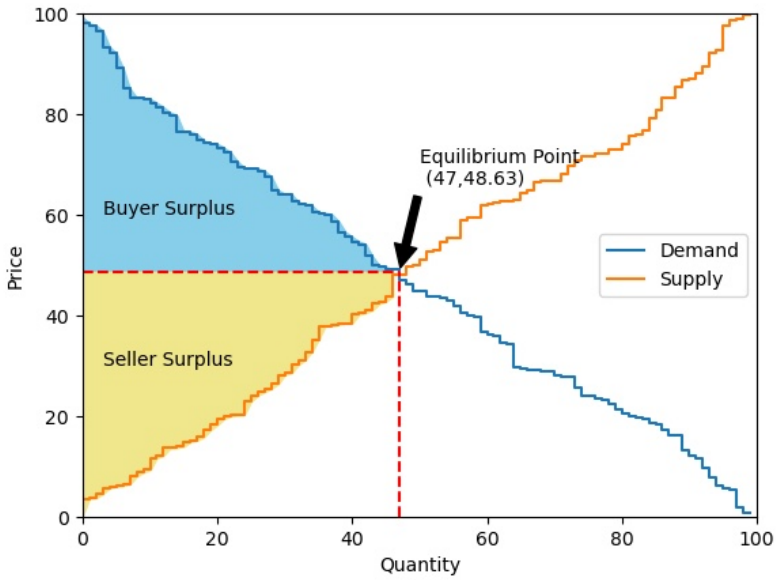


Figure 3.6 Illustration of Supply, Demand, and Auction Outcomes

Mathematically, if we denote the ordered list of bid side's quotes as $\mathbf{B} = [B_1, B_2, ..., B_n]$ with $B_1 \geq B_2 \geq ... \geq B_n$ and the ordered list of ask side's quotes as $\mathbf{A} = [A_1, A_2, ..., A_m]$ with $A_1 \leq A_2 \leq ... \leq A_m$, the market-clearing price p^* and quantity q^* are determined by:

- $q^* = \min(i, j)$ for all (i, j) satisfying $B_i \geq A_j$ for $1 \leq i, j \leq \min(n, m)$
- p^* can be set as the average of B_{q^*} and A_{q^*}

In other words, it represents the point at which no more reasonable trades are feasible. The equilibrium quantity, denoted as q^*, corresponds to the number of such order pairs.

Limit Order Book Dynamics and Trade Execution Illustrations

Now we illustrate how matching works under the batch auction mechanism give a simple example. We start with the initial limit order book under the assumption of clearing executed every two orders. During the batch period 0, a sell order of £103.00 for 2 shares submitted at time 0 and a buy order of £100.50 for 4 shares submitted at time 1 already exist in the market.

Limit Order Book – Batch Period 0			
	Shares	Price	Time
Asks	2	£ 103.00	0
Bids	4	£ 100.00	1

Figure 3.7 The Limit Order Book at Batch Period 0

During the batch period 1, there are another two orders arriving in the market. A sell order of £102.00 for 3 shares at time 2 and a buy order of £99.00 for 2 shares at time 3 are submitted. However, there are no matched pairs for clearing, so the two new orders are also stored in the limit order book.

Limit Order Book - Batch Period 1			
	Shares	Price	Time
Asks	2	£ 103.00	0
	3	£ 102.00	2
Bids	4	£ 100.00	1
	2	£ 99.00	3

Figure 3.8 The Limit Order Book at Batch Period 1

During the batch period 2, another two orders continue to be submitted in the market. A new buy order arrives at time 4, willing to pay £101.00 for 2 shares, and a new sell order arrives at time 5, asking for £100.00 for 3 shares. Then the limit order book, before processing the batch, is updated as follows:

Limit Order Book – Batch Period 2 (before clearing)			
	Shares	Price	Time
Asks	2	£ 103.00	0
	3	£ 102.00	2
	3	£ 100.00	5
Bids	2	£ 101.00	4
	4	£ 100.50	1
	2	£ 99.00	3

Figure 3.9 The Limit Order Book at Batch Period 2 (Before Clearing)

The color-shaded orders in Fig. 3.9 are ready to be paired. At the end of the batch period, the limit order book is reviewed and the market mechanism starts the matching process as the following happens:

- The 2 shares from the new buy order at £101.00 are matched with 2 shares from the new sell order at £100.00. Now, the new sell order has 1 share remaining.
- The remaining 1 share from the ask order at £100.00 continues to match with 1 share from the buy order of £100.50 submitted at time 1, leaving the buy order with 3 shares remaining.

Since the order submitted at time 1 has the time priority, the trade price for this batch clearing is determined to be £100.50. The final state of the limit order book would be:

Limit Order Book - Batch Period 2 (after clearing)			
	Shares	Price	Time
Asks	2	£ 103.00	0
	3	£ 102.00	2
Bids	3	£ 100.50	1
	2	£ 99.00	3

Figure 3.10 The Limit Order Book at Batch Period 2 (After Clearing)

This example shows a simplistic illustration of a batch auction's matching process, where orders are matched at a single price at the end of the period, depending on the accumulated orders' prices and quantities.

3.3 Glosten-Milgrom-Das model

3.3.1 Classic Glosten-Milgrom Model

Glosten-Milgrom-Das model is a well-known information-based model describing how prices form in a single-asset market under asymmetric information. In 1985, Glosten and Milgrom (1985) introduced the Glosten-Milgrom model where all participants are divided into three types: informed traders, liquidity (uninformed) traders and the market maker.

The informed traders know the signal of the true underlying value of the stock and place orders if non-negative profits are expected, while the liquidity traders do not know the information of the true stock prices and randomly place orders irrespective of the current price. The market maker sets the bid and ask prices but has no clue about whether a trader is informed or uninformed.

In the classic Glosten-Milgrom model, the market maker is monopolistic and risk-neutral based on the behaviours of all traders, and the market maker aims to set bid and ask prices to balance the buy and sell orders, but in doing so, they face the risk of trading with informed traders under the adverse selection. As a result, they adjust their bid and ask prices at a reasonable level to compensate for this risk. With the assumption of the risk-neutral market maker, the expected profit of this agent at each instant should be zero. Specifically, when a buy order arrives at time t, the market maker's expected profit is $A_t - \mathbb{E}[V|\text{Buy}]$. Similarly, the expected profit is $\mathbb{E}[V|\text{Sell}] - B_t$ if the new order is a sell order. Therefore, at any time, the market maker will set the quotes as follows:

$$\begin{cases} A_t = \mathbb{E}[V|\text{Buy}] \\ B_t = \mathbb{E}[V|\text{Sell}] \end{cases} \tag{3.1}$$

To obtain the expectation of the price, the market maker should know the probability distribution of the fundamental prices. In the basic Glosten-Milgrom model, the price follows a binary distribution, which means it only has two values which can be denoted as V_H for higher value and V_L for lower value, respectively, and the corresponding probabilities are $(\pi, 1-\pi)$.

For informed traders, they decide whether to buy or sell the stocks based on whether the stock price can bring profits to them. For example, at time t an informed trader will buy if this trader believes the true stock price is higher than the ask price and can earn $V_t - A_t$. Similarly, they will sell the stocks if they believe the bid price is overestimated, and the corresponding profit is $B_t - V_t$.

We denote the proportion of informed traders is μ, so $1-\mu$ of the trading crowd are uninformed traders. The uninformed traders have an equal probability of $1/2$ for buying and selling a stock. Fig.3.11 shows how different traders act upon different conditions.

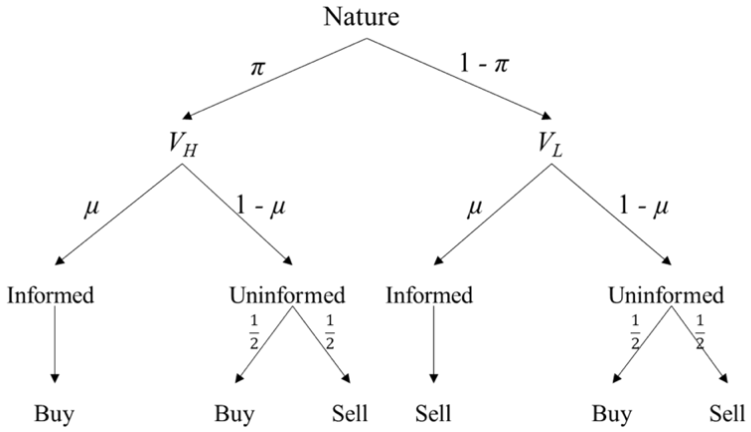


Figure 3.11 Basic Glosten-Milgrom Model Structure

Therefore, we can do the further calculation based on Eq. 3.1:

$$A_t = \mu\pi(A_t - V_H) + \frac{1}{2}(1-\mu)\left(A_t - \bar{V}\right) \tag{3.2}$$

where $\bar{V} = \pi V_H + (1-\pi)V_L$.

Then, the ask price setting strategy is based on the following equation:

$$A_t = \bar{V} + \frac{\mu\pi}{\mu\pi + 1/2(1-\mu)}(V_H - \bar{V}) \tag{3.3}$$

$$= \bar{V} + \frac{\mu\pi(1-\mu)}{\mu\pi + 1/2(1-\mu)}(V_H - V_L) \tag{3.4}$$

Similarly. the bid price can be set as

$$B_t = \bar{V} - \frac{\mu\pi(1-\mu)}{(1-\mu\pi) + 1/2(1-\mu)}(V_H - V_L) \tag{3.5}$$

Therefore, the zero-profit market maker can set bid and ask prices according to Eqs. 3.2 and 3.5 under the Glosten-Milgrom model. It is worth noting that, beyond the classic Glosten-Milgrom model, which assumes a binary distribution for the stock price V, various

extensions introduce different forms for the distribution of stock prices. For example, in Das's extension, the distribution of stock prices at each time follows a normal distribution.

3.3.2 Das's Extension

Das (2005) extended the classic Glosten-Milgrom model by adding a learning component, allowing the market maker to adapt and improve their pricing strategies over time based on the observed order flows. This sequential trade model could be particularly insightful for understanding how automated trading systems behave in real-world financial markets under information asymmetry.

Similar to the basic Glosten-Milgrom model, there are two types of traders: informed traders and uninformed traders. Uninformed traders, when chosen to trade, will place either a buy or a sell order for one unit with equal probability, and there is also a certain probability that they will place no order. Let δ denote the probability of an uninformed trader placing a buy order, so δ must be no more than 0.5. The decision of an uninformed trader can be simplified as follows:

$$\begin{cases} \text{place a buy order,} & \text{with a probability } \delta \\ \text{place a sell order,} & \text{with a probability } \delta \\ \text{no order,} & \text{with a probability } 1-2\delta \end{cases} \tag{3.6}$$

On the other hand, informed traders have knowledge of V_t but with noises. At time t, the noisy estimation w_t of the fundamental price V_t is $w_t = V_t + \varepsilon_t$. The random variable ε_t represents the stochastic noise term in this model, which is presumed to follow an independent and identically distributed (i.i.d) normal distribution. The noise is characterized by a mean of 0 and a standard deviation denoted by σ_n. Therefore, the order-placing decision of an informed trader can be simplified as follows:

$$\begin{cases} \text{place a buy order,} & \text{if } w_t > A_t \\ \text{place a sell order,} & \text{if } w_t < B_t \\ \text{no order,} & \text{otherwise} \end{cases} \tag{3.7}$$

The market maker in Das's model remains risk-neutral and determines the bid and ask prices according to Eq. 3.1. However, to calculate the expectation of the price, the market maker should update the estimation of the probability distribution of the fundamental prices, which is implemented by the Bayesian learning algorithm proposed in Das's model.

3.3.3 Related Computations of Das's Model

In Das's model, an essential feature is that the market maker can adjust the bid and ask prices by learning from the order flows as the market evolves over time. The model assumes that the market maker knows the shares of informed traders in all trading crowd and the distribution of the noise. To calculate the new quotes according to Eq. 3.1, with the introduction of Bayesian Inference in Appendix C, we aim to determine the expression of $\Pr(V_t = v|I_t)$, which is the posterior probability according to the following Bayesian Theorem:

$$\Pr(V_t = v|I_t) = \frac{\Pr(I_t|V_t = v)\Pr(V_t = v)}{\Pr(I_t)} \tag{3.8}$$

where I_t represents the information available to the market maker at time t. Specifically, this information indicates whether the market maker is matched with a buy order, a sell order, or a null order (indicating that no trader has placed an order). Eq. 3.8 serves as the foundation for the market maker to do learning. thus, all mathematical calculations aim to deduce the posterior term by calculating $\Pr(I_t|V_t = v)$ and $\Pr(I_t)$.

Conditional Probabilities of the New Information Arriving

Let $\mathbb{V}_t$ represent the space containing all possible values of the fundamental prices at time t. Then the conditional expectation is

$$\mathbb{E}[V|\text{Sell}] = \sum_{v \in \mathbb{V}_t} v\Pr(V_t = v|\text{Sell}) \tag{3.9}$$

By applying the Bayesian Theorem,

$$\Pr(V_t = v|\text{Sell}) = \frac{\Pr(\text{Sell}|V_t = v)\Pr(V_t = v)}{\Pr(\text{Sell})} \tag{3.10}$$

Then, we obtain

$$\mathbb{E}[V|\text{Sell}] = \sum_{v \in \mathbb{V}_t} v\frac{\Pr(\text{Sell}|V_t = v)\Pr(V_t = v)}{\Pr(\text{Sell})} \tag{3.11}$$

$$= \sum_{v \in \mathbb{V}_t} v\frac{\Pr(\text{Sell}|V_t = v)\Pr(V_t = v)}{\sum_{v \in \mathbb{V}_t}\Pr(\text{Sell}|V_t = v)\Pr(V_t = v)} \tag{3.12}$$

Similarly,

$$\mathbb{E}[V|\text{Buy}] = \sum_{v \in \mathbb{V}_t} v\frac{\Pr(\text{Buy}|V_t = v)\Pr(V_t = v)}{\sum_{v \in \mathbb{V}_t}\Pr(\text{Buy}|V_t = v)\Pr(V_t = v)} \tag{3.13}$$

When informed traders receive a signal that the noisy fundamental price, w_t is lower than the bid price at time t, they opt to sell a unit of stock. Conversely, they will buy a unit of stock when the noisy price signal w_t exceeds the ask price. In all other instances, informed traders retain their stocks, awaiting more favourable opportunities. However, uninformed traders, as mentioned before, place their orders randomly. Consequently, the conditional probabilities can be calculated as

$$\Pr(\text{Sell}|V_t = v) = \Pr(\text{Sell}|V_t = v, v < B_t)\Pr(v < B_t) + \Pr(\text{Sell}|V_t = v, v \geq B_t)\Pr(v \geq B_t) \tag{3.14}$$

$$\Pr(\text{Buy}|V_t = v) = \Pr(\text{Buy}|V_t = v, v > A_t)\Pr(v > A_t) + \Pr(\text{Buy}|V_t = v, v \leq A_t)\Pr(v \leq A_t) \tag{3.15}$$

Thus, we can calculate each component of the conditional probabilities based on the decisions made by both informed and uninformed traders under various circumstances. The informed traders only place orders when profitable:

$$w_t = V_t + \varepsilon_t < B_t \Rightarrow \varepsilon_t < B_t - V_t$$

$$w_t = V_t - \varepsilon_t < B_t \Rightarrow \varepsilon_t > V_t - B_t$$

As defined in the previous section, ε_t is a random variable, normally distributed with a mean of 0 and a standard deviation of σ_n. Consequently, the conditional probabilities mentioned in Eq. 3.14 and 3.15 can be expressed as

$$\Pr(\text{Sell}|V_t = v, v < B_t) = \mu \Pr(N(0,\sigma_n^2) < B_t - v) + (1-\mu)\delta \tag{3.16}$$

$$\Pr(\text{Sell}|V_t = v, v \geq B_t) = \mu \Pr(N(0,\sigma_n^2) > v - B_t) + (1-\mu)\delta \tag{3.17}$$

where μ stands for the shares of the informed traders in the trading crowd. Thus,

$$\begin{aligned}
\Pr(\text{Sell}|V_t = v) =& \Pr(\text{Sell}|V_t = v, v < B_t)\Pr(v < B_t) + \Pr(\text{Sell}|V_t = v, v \geq B_t)\Pr(v \geq B_t) \\
=& (\mu \Pr(N(0,\sigma_n^2) < B_t - v) + (1-\mu)\delta)\Pr(v < B_t) + \\
& (\mu \Pr(N(0,\sigma_n^2) > v - B_t) + (1-\mu)\delta)\Pr(v \geq B_t) \\
=& \big(\mu \Pr(N(0,\sigma_n^2) < B_t - v) + (1-\mu)\delta\big)\big(\Pr(v < B_t) + \Pr(v \geq B_t)\big) \\
=& \mu \Pr(N(0,\sigma_n^2) < B_t - v) + (1-\mu)\delta \\
=& \mu P_{\varepsilon_t}(B_t - v) + (1-\mu)\delta
\end{aligned} \tag{3.18}$$

where P_{ε_t} refers to the cumulative distribution function of the noise's distribution. Similarly,

$$\Pr(\text{Buy}|V_t = v) = \mu \Pr(N(0,\sigma_n^2) < v - A_t) + (1-\mu)\delta \tag{3.19}$$

$$= \mu P_{\varepsilon_t}(v - A_t) + (1-\mu)\delta \tag{3.20}$$

Also, in some cases, no order comes in the market since both informed traders and uninformed traders decide to do nothing. Then we can similarly obtain the conditional

probability as

$$\Pr(\text{No order}|V=v) = (1-\mu)(1-2\delta) + \mu P_{\varepsilon_t}(v-B_t) + \mu P_{\varepsilon_t}(A_t-v) \tag{3.21}$$

Unconditional Probabilities of the New Information Arriving

According to the law of total probability,

$$\Pr(I_t) = \sum_{v\in\mathbb{V}_t} \Pr(I_t|V_t=v)\Pr(V_t=v) \tag{3.22}$$

At each time, there are three possible cases that the market maker could face: the new order is a buy order, the new order is a sell order or no order comes into the market. Therefore, we can calculate the corresponding probabilities as follows:

$$\Pr(\text{Sell}) = \sum_{v\in\mathbb{V}_t} (\mu P_{\varepsilon_t}(B_t-v) + (1-\mu)\delta))\Pr(V_t=v) \tag{3.23}$$

$$\Pr(\text{Buy}) = \sum_{v\in\mathbb{V}_t} (\mu P_{\varepsilon_t}(v-A_t) + (1-\mu)\delta))\Pr(V_t=v) \tag{3.24}$$

$$\Pr(\text{No order}) = \sum_{v\in\mathbb{V}_t} (\mu(1-\mu)(1-2\delta) + \mu P_{\varepsilon_t}(v-B_t) + \mu P_{\varepsilon_t}(A_t-v))\Pr(V=v) \tag{3.25}$$

Posterior Probabilities

As soon as a new order comes to the market, it provides new information to the market maker regarding the traders' reflection on the current bid and ask prices, so the market maker can calculate the posterior to update the probability density function of the fundamental prices V_t. The updated estimation of the density function can be derived by

Bayesian learning according to the formulas in Eq. 3.8.

$$\Pr^*(V=v)=\begin{cases}\dfrac{(\mu P_{\varepsilon_t}(B_t-v)+(1-\mu)\delta))\Pr(V=v)}{\sum_{v\in\mathbb{V}_t}(\mu P_{\varepsilon_t}(B_t-v)+(1-\mu)\delta))\Pr(V=v)}, \\ \qquad\qquad \text{if a sell order arrives} \\ \dfrac{(\mu P_{\varepsilon_t}(v-A_t)+(1-\mu)\delta))\Pr(V=v)}{\sum_{v\in\mathbb{V}_t}(\mu P_{\varepsilon_t}(v-A_t)+(1-\mu)\delta))\Pr(V=v)}, \\ \qquad\qquad \text{if a buy order arrives} \\ \dfrac{(\mu(1-\mu)(1-2\delta)+\mu P_{\varepsilon_t}(v-B_t)+\mu P_{\varepsilon_t}(A_t-v))\Pr(V=v)}{\sum_{v\in\mathbb{V}_t}((1-\mu)(1-2\delta)+\mu P_{\varepsilon_t}(v-B_t)+\mu P_{\varepsilon_t}(A_t-v))\Pr(V=v)}, \\ \qquad\qquad \text{if no order arrives}\end{cases} \tag{3.26}$$

Where $\Pr^*$ refers to the posterior probability. At the same time, the vector $\mathbb{V}_t$ will also be updated representing the new probability density function of the fundamental prices.

Then the new ask and bid prices can be updated as

$$\begin{aligned} B_{t+1}&=\mathbb{E}[V|\text{Sell}]=\frac{\sum_{v\in\mathbb{V}_t}(\mu P_{\varepsilon_t}(B_t-v)+(1-\mu)\delta))\Pr(V=v)v}{\sum_{v\in\mathbb{V}_t}(\mu P_{\varepsilon_t}(B_t-v)+(1-\mu)\delta))\Pr(V=v)} \\ A_{t+1}&=\mathbb{E}[V|\text{Sell}]=\frac{\sum_{v\in\mathbb{V}_t}(\mu P_{\varepsilon_t}(v-A_t)+(1-\mu)\delta))\Pr(V=v)v}{\sum_{v\in\mathbb{V}_t}(\mu P_{\varepsilon_t}(v-A_t)+(1-\mu)\delta))\Pr(V=v)} \end{aligned} \tag{3.27}$$

Eq. 3.27 shows how the learning market maker in Das's model sets quotes to balance the buy and sell sides after the probability density function estimation of $\mathbb{V}_t$ is updated.

Pseudo code

The market maker functions to balance the buy and sell side of the traders by setting reasonable ask and bid prices, which are analytically derived in Eq. 3.27. In Das's model, the market maker adjusts the bid and ask price by Bayesian learning as soon as a new order arrives. Algorithm 1 shows how the market maker is modelled and acts in the framework of Das (2005).

Algorithm 1 The Learning Market Maker

Require: $\mu, \delta, V_0, B_0, A_0,$ time steps, noise

Ensure: The strategy of quoting

```
function LEARNING
    for t in time steps do
        observe the new information
        if a buy/sell order arrives then
            match the order
        end if
        update the distribution of the prices
        update B_t, A_t
    end for
end function
```

The basic model structure, mathematical computations, and related agents' behaviours involving Glosten-Milgrom-Das model mentioned in this section will serve as the foundation for extending models in the following chapters.

3.4 Flash Crash Model Framework

In this thesis, we designed a flash crash model based on the agent-based modelling extended from the classic market micro-structure model, the Glosten-Milgrom-Das (GMD) model, where traders with different information levels are included. The structure of the artificial market we are using is extended from the introduction of the GMD model in Section 3.3. In order to better understand the structure and the mechanism of the model, it is necessary to describe the framework and components in the modelled system in detail for simulations.

3.4.1 The Marketplace Assumptions

We design that all activities take place in an artificial market under the double auction mechanism. To support our market model, here are the key assumptions:

- Only a single stock is traded in the unique exchange in the model.
- The stock pays no dividends and has its own true price, which is exogenous and independent of any activities of the market.
- Trading is frictionless, in which case there are no costs, such as commissions and exchange fees, nor constraints related to transactions.
- A zero-profit market maker exists in the market to maintain liquidity. To guarantee the liquidity of the market, the market maker simultaneously submits a unit sell order and a unit buy at each trading round.
- Semi-strong efficiency: prices reflect all public information, but there exists some private information that only a subset of traders are aware of.

3.4.2 Trading Agents

The artificial market includes three different kinds of trading agents: fundamental traders, zero-intelligence traders and a market maker. Referred to Das (2005), the behaviour of different trading agents in the dealer market is modelled as follows:

All traders are modelled to submit a unit share of the stock for each submission. Inventory limits are ignored in this model.

1. **ZIT, zero-intelligence trader:** Zero-intelligence traders place a market buy order or a market sell order at the same probability, acting regardless of their profits. Zero-intelligence traders is referred to as the definition of "uninformed traders" in Das's model, which is illustrated in Section 3.3.2.
2. **FT, fundamental trader:** Fundamental traders have access to information signals of "true" stock prices, and will only trade when expecting positive yield. The features of fundamental traders is consistent with the definition of "informed traders" in Das's model.
3. **MM, marker maker:** The market maker, with the settings similar to that in Das's model, posts buy and sell orders, follows a risk-neutral model according to Eq.

3.31, thus this agent plays a role in providing liquidity and capturing the asset price unknown to him.

Next, we will provide a more detailed explanation of the settings and behaviours of these three traders proposed in this thesis.

Zero-intelligence Traders

The settings of zero-intelligence traders are referred to as that of informed traders in Eq. 3.6. Such modelled traders do not know the true price of the stock and any information about the shock. They only randomly submit buy market orders or sell market orders with the same probability of 0.5. This kind of trader is aggressive in the market as they randomly take the liquidity from the opposing order book. Introducing zero-intelligence traders could help the market become more active. Otherwise, there would be almost no trade if the market price is around the true stock price. By setting δ in Eq. 3.6 to be 0.5, the decisions of ZI trader can be described as follows:

$$\begin{cases} \text{place a buy order at the best ask price,} & \text{with a probability 0.5} \\ \text{place a sell order at the best bid price,} & \text{with a probability 0.5} \end{cases} \tag{3.28}$$

Fundamental Traders

Fundamental traders are in possession of their own estimates of the stock through the information they can get. They only place orders when it comes to positive profits. Under the assumption that short selling is allowed, there is no inventory limit or capital limit for each fundamental trader, which means a fundamental trader can always place an order when recognising the stock is mispriced in the market.

However, under the circumstance that there are few differences between the market price and the trader's estimated price, the fundamental trader will tend to take no action and wait for better opportunities. In other words, the market is more inactive if the stock is well-priced. From the micro perspective, there is no trade in a single round under such a condition. For unified expression, a trader taking no action is equivalent to placing a null order into the market.

Each fundamental trader, denoted as i, can have evaluations of the true stock price, denoted as $w_{i,t}$. A fundamental trader will post the buy order only if the trader believes that the stock price is underestimated, under which circumstance the trader's estimated true price is higher than the best ask price at time t. Similarly, a trader will place a sell order only if the estimated true price is lower than the best bid price at which the stock price is overestimated. The fundamental traders always place limit orders at the prices same as the price signals they receive. The strategy of fundamental traders can be summarised mathematically as follows:

$$\begin{cases} \text{place a buy order,} & \text{if } w_t > A_t \\ \text{place a sell order,} & \text{if } w_t < B_t \\ \text{not place an order,} & \text{otherwise} \end{cases} \tag{3.29}$$

In this thesis, fundamental traders are categorised into two types: **informed fundamental traders (IFT)** and **uninformed fundamental traders (UFT)** based on their awareness of private information. Informed fundamental traders have access to all information, allowing them to receive accurate price signals, while uninformed fundamental traders could be spoofed by misinformation, such as a "fake shock", then they would receive biased price signals. More details are specified in Section 3.4.3 regarding the "fake shock" and spoofing.

The Market Maker

Unlike the other traders, the market maker simultaneously posts limit orders on both sides of the limit order book and provides liquidity. The market maker also aims to capture the true stock price, and Bayesian learning is employed referred to as the application of the Glosten-Milgrom-Das model shown in Section 3.3. The Glosten-Milgrom-Das model is a theoretical model that studies the informational asymmetry effect on market making. In this model, the market maker is assumed to be risk-neutral. Therefore, the expected profit of setting a bid and an ask for each submission should be zero. If a buy order arrives at time t, the market maker's expected profit can be calculated as $A_t - \mathbb{E}[V|\text{Buy}]$. Similarly,

the expected profit is $\mathbb{E}[V|\text{Sell}] - B_t$ when a sell order arrives. Therefore, to make the expected profit zero at any time, the market maker will set the quotes as Eq. 3.1 shows:

$$\begin{cases} A_{t+1} = \mathbb{E}[V|\text{Buy}] \\ B_{t+1} = \mathbb{E}[V|\text{Sell}] \end{cases} \tag{3.30}$$

In order to get the conditional expectation of the price, the market maker should estimate the probability distribution over the entire possible range of the true price, which can be derived from the Bayesian learning algorithm. The market marker is assumed to know the market structure and the distribution of the price noise.

Referring to Eq. 3.27 and after a series of mathematical computations, we derive that the market maker updates the bid and ask prices using the following formulas:

$$\begin{aligned} B_{t+1} = \mathbb{E}[V|\text{Sell}] &= \frac{\sum_{v\in\mathbb{V}_t}\left(\mu\Phi\left(\frac{B_t-v}{\sigma_{\varepsilon_n}}\right)+\frac{1}{2}(1-\mu)\right)\Pr(V_t=v)v}{\sum_{v\in\mathbb{V}_t}\left(\mu\Phi\left(\frac{B_t-v}{\sigma_{\varepsilon_n}}\right)+\frac{1}{2}(1-\mu)\right)\Pr(V_t=v)} \\ A_{t+1} = \mathbb{E}[V|\text{Buy}] &= \frac{\sum_{v\in\mathbb{V}_t}\left(\mu\Phi\left(\frac{v-A_t}{\sigma_{\varepsilon_n}}\right)+\frac{1}{2}(1-\mu)\right)\Pr(V_t=v)v}{\sum_{v\in\mathbb{V}_t}\left(\mu\Phi\left(\frac{v-A_t}{\sigma_{\varepsilon_n}}\right)+\frac{1}{2}(1-\mu)\right)\Pr(V_t=v)} \end{aligned} \tag{3.31}$$

where $\Phi(\cdot)$ refers to the cumulative distribution function of a standard normal distribution, and σ_{ε_n} denotes the standard deviation of the price noise. μ represents the proportion of fundamental traders in the total number of traders in the system, expressed as $n_{\text{IFT}}+n_{\text{UFT}}/n_{\text{Total}}$. $\mathbb{V}_t$ represents the set of possible true price values at time t, indicating that the true price could be any value within this set. $\Pr(V_t = v)$ is used to calculate the probability of the true price being equal to v. Eq. 3.31 is the extended form of Eq. 3.27, enabling the market maker to update bid and ask prices when a new order arrives. The complete mathematical computation can be found in Appendix D.

3.4.3 Trading Model

Price Dynamics

Let us assume the following trading process: we separate the whole trading period into several divisions with different time length scales. Assume the market is running in the time interval $[0, T]$, which we call the trading "era" for the whole period. The trading era is composed of sequential trading epochs, as shown in Fig. 3.12. To simplify the model, we assume that there are a fixed number of trading rounds M in each trading epoch throughout the trading era.

The true stock price, also called the fundamental price as in Das (2005), is exogenous to the model as Section 3.4.1 mentioned. Assume that the stock is continuously traded. The true stock price evolves as a "special" jump process[1].

We assume that in this jump process, in each epoch there can be at most one jump. Therefore, in each trading epoch, there are two situations of the true stock prices:

- With a jump: $V_{t+1} = V_t + \delta V$
- Without a jump: $V_{t+1} = V_t$

where V_t denote the true stock price at time t. We describe how the simulation works with the illustration of Fig. 3.12.

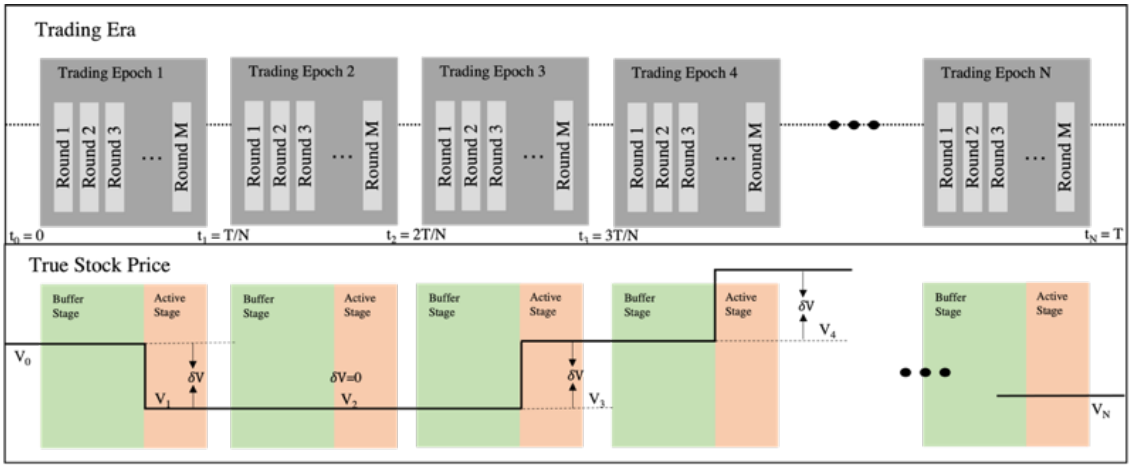


Figure 3.12 How the Simulation Works

[1] A jump process is a mathematical model used in probability theory and stochastic processes to describe a system that undergoes sudden, discontinuous changes, which are called "jumps", at certain random times.

We assume that the jump occurs at the k-th round. For convenience, the trading epoch can be divided into two stages: **the buffer stage**, containing round 1 to round k, and **the active stage**, containing round k+1 to round M.

In this thesis, we assume that all trading activities occur within a single trading epoch, so we consequently do short-term experiments of the model. This simplified timeframe allows us to focus on specific interactions and behaviors within a constrained period. It's important to acknowledge that this is a simplifying assumption, as real-world trading often extends over longer periods, involving multiple trading sessions and diverse market conditions. For a more comprehensive analysis, future research could consider scenarios that encompass longer timeframes and account for the dynamics of trading activities over time.

Informed Traders and Spoofing

In the semi-strong efficient market hypothesis, all public information is used in the determination of a stock's price. However, if there is some private information indicating that the stock price to a level different from that of the public information, the parties who have information superiority could make excess profits against the parties who lack information. For example, under the circumstance of the market stock price being £100, if the informed traders get the price signal to be £110, but the uninformed traders believe the stock price should be £90, an informed trader chooses to buy the stock at £100 and grabs £10 per share from the uninformed party who sells the stocks. The information asymmetry leads to the party who lacks information making a bad decision, in which case we call it **adverse selection**.

Therefore, regardless of integrity, a profit-pursuing informed trader may attempt to enlarge the information gap between the informed party and the uninformed party to increase profits from adverse selection. One approach is "spoofing". Some malicious traders may make a spoofing strategy causing a piece of "fake" information to mislead the uninformed traders, in which case the true price is not affected. In 2010, according to the US government, a trader named Sarao used the spoofing practice to make tens of millions of dollars, causing a dramatic drop in the Dow Jones Industrial Average Index.

He placed a large number of orders before quickly cancelling or amending them, in which case he created the "fake" demand or supply on the asset. This practice could drive the stock price to move in one direction, but Sarao can earn profits from the adverse selection by just placing genuine orders.

This thesis simplifies the scenario by considering a case that there exists a trader spoofing by adding a "fake shock" into the market. We assume that a fundamental trader is confident that there will be no price jump in the current trading epoch. However, this trader spreads a fake shock signal, causing some traders to be deceived to mistakenly believe that the stock price will experience a jump, thus altering their price estimates, as shown in Fig. 3.13.

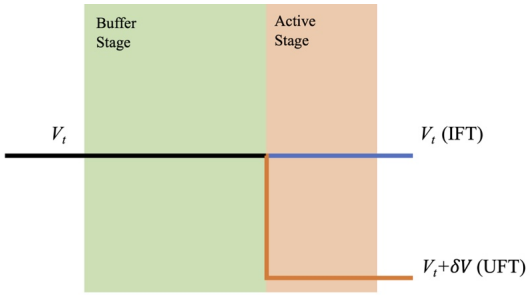


Figure 3.13 Price Estimates of Different Fundamental Traders

Based on this scenario, we categorize fundamental traders into two groups: **Informed fundamental traders (IFTs)** have both public and enough private information so that they can always well estimate the true stock price. An informed fundamental trader's estimate can be described as $w_{i,t}^{\text{IFT}} = v_t^{\text{IFT}} + \varepsilon_{n,i}$, where w_t^{IFT} is the signal of the true stock price that an informed fundamental trader receives, and $\varepsilon_{n,i}$ is signal noise with a normal distribution $\left(\varepsilon_i \sim \mathcal{N}(0, \sigma_{\varepsilon_n})\right)$. In contrast, the other fundamental traders are in lack of private information of the "fake shock", and we call these traders **uninformed fundamental traders (UFTs)**. An uninformed fundamental trader's estimate can be described as $w_{i,t}^{\text{UFT}} = v_t^{\text{UFT}} + \varepsilon_{n,i}$ and may take bad trading strategies, where w_t^{UFT} is the signal of the true stock price that an uninformed fundamental trader receives. The experiments in this thesis will investigate how the market will perform under the circumstance of a fake shock.

Agents' Interactions

In our model, there are four different types of trading agents: informed fundamental traders (IFT), uninformed fundamental traders (UFT), zero-intelligence traders (ZIT), and a market maker (MM), whose trading behaviour drives the dynamics of the market moving forwards. However, for the trading epochs without jumps, we assume that there may be unusual events not affecting the real stock prices but disturbing the public information, such as spoofing by an informed fundamental trader, misleading some uninformed fundamental traders to take bad trading strategies, which leads the market prices to deviate significantly from the true prices. For simplification, we call such spoofing event the "fake shock". When the "fake shock" hits the market, the true stock price remains to be v, but the fake information could bring a biased price signal v'. Thus, the uninformed fundamental traders change their estimates as $w_{i,t}^{\mathrm{UFT}} = v' + \varepsilon_{n,t}$, while the informed fundamental traders' estimates are $w_{i,t}^{\mathrm{IFT}} = v + \varepsilon_{n,t}$, where $w_{i,t}$ represents the price signal that the i-th trader receives at the trading round t, and $\varepsilon_{n,t}$ is the noise term for the price signal at time t. The noise term varies for different traders at different trading rounds.

As the previous sections illustrate, in the CDA market, buyers and sellers submit their orders to specify the amount and price they would like to buy or sell. A trade happens only if the highest bid price is no lower than the lowest ask price, and the trade price is determined by the trading pair. It equals the quote of the earlier order in the trader pair or the midpoint of the quotes of the two orders if they arrive simultaneously. For all outstanding orders, they will be kept in the limit order book, which is divided into two priority queues: the buy order book for bids and the sell order book for asks.

In this model, based on the assumptions set out in Section 3.4.1, only one trade is executed each round. Fig. 3.14 illustrates how the agents interact to drive the market price of the stock round by round. Before each trading round starts, the fundamental traders get the signal of the true stock price from the market and wait for the ask and bid from the market maker. After the trading round starts, there is an order from the competition of all traders coming to the exchange and matching with the buy order or sell order from the market maker. When the trade is completed, the exchange will broadcast the trade information to the market maker and the traders. As soon as the market maker knows

which kind of order has been executed, the market maker adjusts the bid and ask prices for the next round.

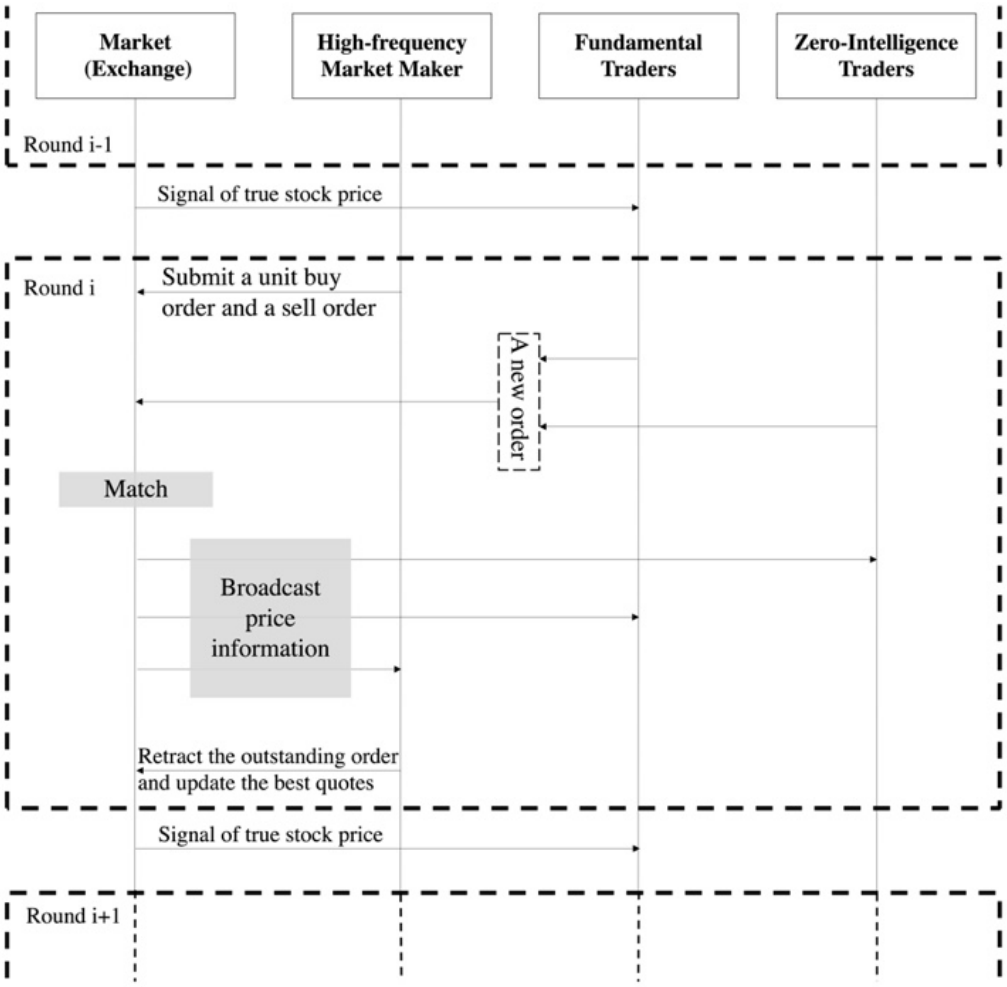


Figure 3.14 The Trading Process during a Non-spoofing Round

In this model, since the real spoofing strategy and its interaction in the order book are both relatively complicated, we simplify the spoofing strategy as traders spreading the disinformation that may affect a trader's estimation, which is regarded as a "fake shock". Assume the informed fundamental traders can estimate the stock price based on the "real" true stock price v_t, while spoofing is implemented by giving uninformed fundamental traders a biased stock price signal v_t' rather than true signal v_t. As the spoofing starts, different traders receive the stock price signals according to the following representations:

$$\begin{cases} w_{i,t}^{\text{IFT}} = v_t^{\text{IFT}} + \varepsilon_{n,i} = v_t + \varepsilon_{n,i} \\ w_{i,t}^{\text{UFT}} = v_t^{\text{UFT}} + \varepsilon_{n,i} = v_t' + \varepsilon_{n,i} \end{cases} \tag{3.32}$$

Fig. 3.14 illustrates how the agents interact to drive the market price during the spoofing round:

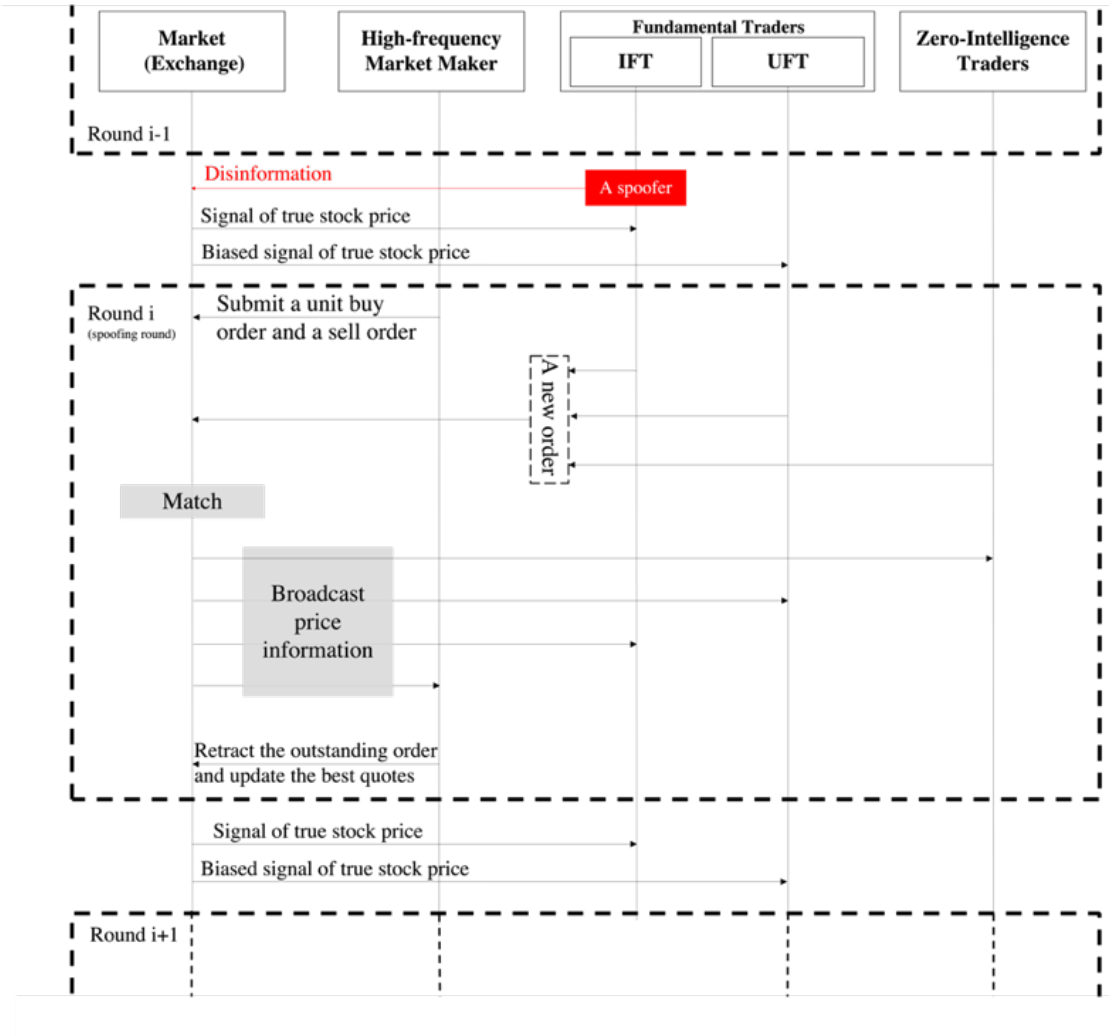


Figure 3.15 The Trading Process during a Spoofing Round

This thesis considers a situation, in which an informed fundamental trader is taking the spoofing strategy over the public so that the uninformed fundamental traders receive the

biased estimate of the stock price. However, informed fundamental traders can identify such disinformation and receive more accurate price signals from the market. Consequently, a significant difference may arise between v_t^{IFT} and v_t^{UFT} when information asymmetry exists, leading to substantial fluctuations, such as flash crashes, in market prices. Fig. 3.16 shows an example of how market prices crash after the "fake shock", and such cases are thoroughly examined in Chapter 4. Furthermore, Chapters 5, 6, and 7 investigate whether and how flash crashes can be mitigated in the presence of intelligent UFTs, within batch auction markets, and within information-sharing networks, respectively.

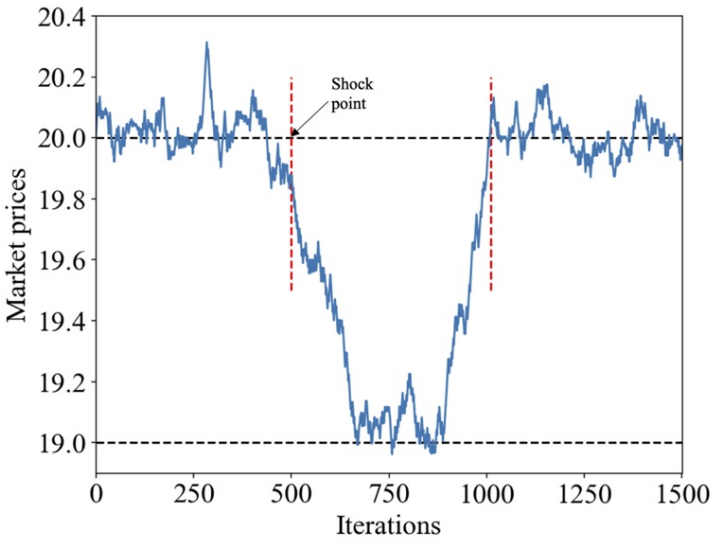


Figure 3.16 Market Prices Movement after the "Fake Shock" (with Intelligent UFTs, in Chapter 5)

3.5 Simulation Metrics

3.5.1 Agents' Profits

The profits of each agent could be a good measurement to reflect different agents' performance in our model. As we run simulations among sets of parameters regarding the market structure, in order to rule out the effect of different agents' shares on the profits, we

focus on the average profits on a certain type of agents in the trading crowd as follows:

$$\bar{\text{Prof}}^{\text{ZI}} = \frac{\sum \text{Prof}_i^{\text{ZI}}}{n_{\text{ZIT}}}, \quad \bar{\text{Prof}}^{\text{IFT}} = \frac{\sum \text{Prof}_i^{\text{IFT}}}{n_{\text{IFT}}}, \quad \bar{\text{Prof}}^{\text{UFT}} = \frac{\sum \text{Prof}_i^{\text{UFT}}}{n_{\text{UFT}}} \tag{3.33}$$

Next, we briefly explain how each agent works and makes profits.

1. The market maker: in our model, the market maker is set to be risk-neutral/zero-profit, so this agent simply adjusts the quotes upon the direction of new orders – if the upcoming is a buy (sell) order, then the market maker moves the price up (down) – to change the offer price. Therefore, all the agent earns is basically the cumulative bid/ask spreads if the number of buy and sell orders in the market is about the same. However, when the number of buy orders in the market is much larger than the number of sell orders or when the number of buy orders is much smaller than the number of sell orders, this will result in long or short stock positions for the market maker. In this case, the profits regarding the inventory are related to redemption values in the end.

2. Fundamental traders: Based on their own estimates of the true price, all fundamental traders are able to determine whether the market price is accurately priced. Also, fundamental traders will submit buy (sell) orders when the market is under (over) priced. After a fake shock, the informed fundamental traders will have a different price estimate from the uninformed fundamental traders, leading them to quote in the opposite direction. This information asymmetry may also advantage the informed fundamental traders grabbing profits from the market maker and the uninformed fundamental traders.

3. Zero-intelligence traders: In this model, they are completely random agents with minimal information advantage. Therefore, the ZIT usually has a negative profit in this zero-sum model.

In this thesis, we will conduct and investigate simulations in both continuous double auction markets and batch auction markets. As we focus on the simulations within a single trading epoch, the agents' profits are calculated as the expected profits they can obtain at the end of the trading epoch. In other words, traders are presumed to clear their positions at the price of the market price as the trading epoch ends. Therefore, for the k-th trade occurring during the trading epoch between a buyer b_i and a seller s_j, the profits of each

party can be calculated as:

$$\bar{\text{Prof}}_k^{b_i} = v_M - v_k^*, \quad \bar{\text{Prof}}_k^{s_j} = v_k^* - v_M \tag{3.34}$$

where v_k^* stands for the trade price of the k-th trade, and v_N stands for the market price at the ending point M of the trading epoch. Agents' profits will be calculated differently depending on the ways of determining trade prices under different market mechanisms applied in the model.

Continuous Double Auction Section 3.2.1 introduces that, in a continuous double auction market, the matching mechanism is executed continuously. If there is a trade occurring, the trade price is determined by matching the new order with the existing order at the best available price. Specifically:

- If a new buy order matches with a sell order, the trade occurs at the best ask price.
- If a new sell order matches with a buy order, the trade occurs at the best bid price.

Batch Auction As Section 3.2.2 specifies, in batch auctions, orders are gathered over a certain time interval and matched and executed periodically. All buy and sell orders within a batch are processed at a uniform price, determined by identifying the price that optimises the total trading volume for the batch. The uniform trade price for a batch is determined through Marshallian Paths shown in Fig. 3.6.

Appendix A also shows examples of matching and price determination in continuous double auction markets and batch auction markets.

3.5.2 Market Prices

Crashes

As soon as the external signal arrives in the market, some traders could alter their estimate from v to v', driving the market price away. For example, with the assumption of that the shock causes a price down from $v = 20$ to $v' = 19$, as the fake shock occurs at time t_s, only uninformed fundamental traders change the estimate to around $v' = 19$, while

the informed fundamental traders remain the same as $v = 20$. Therefore, the two types of traders exert opposite forces when the market price is between 19 and 20 by placing orders in opposite directions.

Assume that there exists a price $\bar{v}$ making the buy-side and the sell-side balanced, then it is possible that the market price converges to $\bar{v}$ only with small fluctuations around the level. We call the first point satisfying such condition in the active stage **the early stop point** t_e and the corresponding price **the steady-state price** $\bar{v}$. In this thesis, we employ MSER-5 method to determine the early stop point which can be shown in Appendix B. Therefore, we can further determine the **crash size** and **crash duration** by calculating the differences between the shock point and the early stop point in prices and time-steps, respectively. Fig. 3.17 shows a graphic explanation of the flash crash metrics.

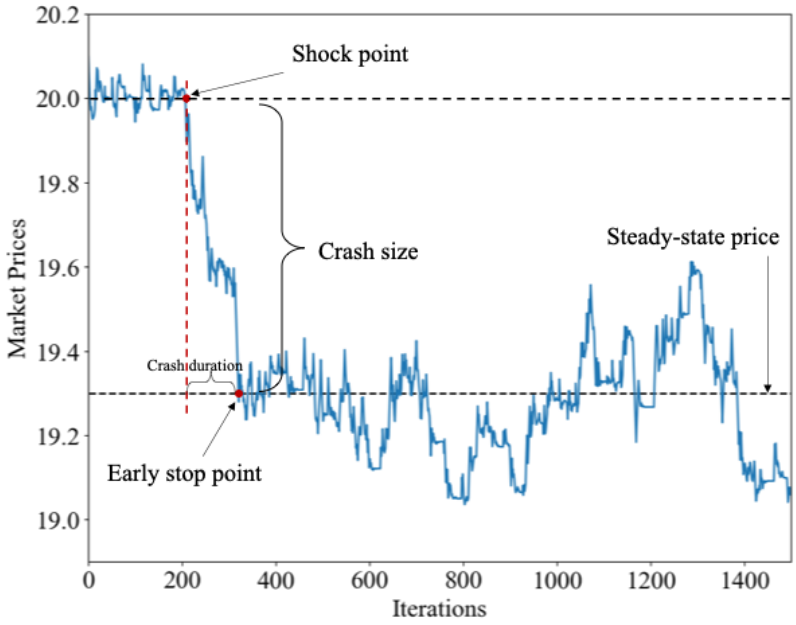


Figure 3.17 A Graphic Explanation of the Metrics of Crashes

Note that the early stop point (truncation point) is an integer random variable, and multiple replications of the simulation generate different samples of early stop points. To avoid the effect of extreme values, the median of the sample MSER early stop points is used for analysis.

Bid-ask Spreads

In a model under asymmetric information, one of the most important frictions to be studied in the micro-structure model is adverse selection, which is reflected in the bid-ask spreads. A single simulation can be divided into three phases: before-crash, in-crash, and after-crash. We will investigate the bid-ask spreads and the components in these three phases. The bid-ask spreads can be decomposed into adverse selection, inventory holding and order processing. In our model, the inventory is unconstrained, and we focus on the effects of adverse selection and order processing on the spreads. Glosten-Harris model (Glosten and Harris, 1988) and MRR model (Madhavan et al., 1997) combine the inventory effect and order processing as an integrated component, which will be applied in this chapter.

In the monopoly dealer model, the trades are at either the best bid (the order is initiated by a seller) or the best ask (the order is initiated by a buyer), so the bid-ask spread model can be expressed as

$$\Delta v_t = Z_t Q_t + \frac{1}{2} C_t \Delta Q_t + e_t \tag{3.35}$$

where

v_t is the market price;

Z_t is the adverse selection spread component;

C_t is the order processing components;

Q_t = +1 if the trade is at the best ask, -1 if the trade is at the best bid, 0 if no trade occurs.

ΔQ_t is 0 if two consecutive orders are in the same direction. However, ΔQ_t is +2 if a buy order is initiated followed by a seller order and -2 if a sell order comes after a buy order. It is straightforward that the price is highly likely to experience a wider change with two consecutive buy orders rather than with a sell order followed by a buy order. Therefore, the impact of the attributes Z_t could be with different signs. We run 100 simulations for

each given parameter setting and take the regression on all three phases in each simulation. To the final, for each set of parameters simulated, we obtain clusters of adverse selection and order processing factor distributions.

3.5.3 Variability

In order to express the variability and deviation of the repeated simulations, we adopt two measures of the Coefficient of Variability (CV) in this essay: inter-run CV and intra-run CV. Besides, Since the intra-run variability and inter-run variability interact with each other in a well-behaved simulation, we can easily combine the two indicators to make a ratio measurement generating a Variability Ratio (VR) reflecting the overall variability of the simulations which can be expressed by the following math:

$$VR = \frac{\text{Sum of Intra-run CV}}{\text{Sum of Inter-run CV}} = \frac{\sum_N CV_T^n}{\sum_T CV_t^N} \tag{3.36}$$

where CV_t^N refers to the inter-run CV at time t over a set of runs, and CV_T^n refers to the intra-run CV of the n-th run over the iterations. If VR is much smaller than 1, then the ensemble variability accounts for more, and the internal variability accounts for more if VR is much larger than 1.

3.5.4 Stationarity

With agent-based simulation on the artificial double auction market, the groups of sequential information have been generated and contain plenty of important characteristics of the system. Discovering the properties of the sequential-information-series or time-series data is important for understanding the model and the interaction of agents. To achieve such goals, this section presents the quantitative analysis of the statistics of the artificial data to study stationarity. As the agent-based model runs, the interaction among agents pushes the system to converge towards a state that could be “stable” or “unstable”. An equilibrium state refers to a stable condition where key indicators remain relatively stationary over time. If the artificial system turns out to be stationary, the theoretical model allows a better understanding of the convergences and the variants of the system as the

market conditions change (Grazzini, 2011). This section utilizes non-parametric tests, specifically the Wald-Wolfowitz test (Wald and Wolfowitz, 1940) to assess stationarity.

A stationary time series or process $\{X_t\}_t$ meets the condition that its statistics do not change as a function of time. Mathematically, the property of a stationary process can be defined as

Definition 1: A stochastic process $\{X_t\}_t$ is strictly stationary if $\forall k$, $\forall t_1,...,t_k$, $\forall c_1,...,c_k$ constants we have $P(X_{t_1} \leq c_1,...,X_{t_k} \leq c_k) = P(X_{t_{1+h}} \leq c_1,...,X_{t_{k+h}} \leq c_k)$, $\forall h$ time shifts.

Definition 2: A stochastic process is weakly stationary if 1)its mean function is constant and independent of t: $E(X_t) = \mu$, and 2)its autocovariance function depends on s,t only via $|t-s|$: $\gamma(s,t) = \gamma(|t-s|,0)$.

There have been several useful methods applied to test the stationarity of time series. Dickey-Fuller test (Dickey and Fuller, 1979), focusing on the unit root test, plays a big role in statistics and has built a foundation on the parametric test for the stationarity of time series. However, the Dickey-Fuller test is subject to autoregressive processes and limited assumptions on the noise term. Said and Dickey (1984) extended the previous model to the ARIMA(p,1,q) process considering the effect of serial correlation. Another alternative of the unit root test was Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test (Kwiatkowski et al., 1992) to figure out whether a time series is stationary or with a linear trend. Unlike the Dickey-Fuller test with the alternative hypothesis of stationary data, the null hypothesis for the KPSS test is that the data is stationary. Although these methods have been powerful in statistics, as they are parametric methods, the time series to be tested are subject to assumptions and limitations about specific stochastic processes, so nonparametric test could be a better way in artificial times series. Inspired by Grazzini (2012), we will use the Wald-Wolfowitz test or called Run Test to test the stationarity and ergodicity, which are described in the next sections.

Stationarity Test

Wald-Wolfowitz test (or Runs Test) used to test stationarity, and Gibbons and Chakraborti (2014) extended the method to figure out if a time series fits the given function. If so, the times series should evolve randomly around the function regardless of the noises. If the

data was located above the function, it is labelled as "+", and the others that are distributed below the function are labelled as "-". A run can be defined as a non-empty maximal sub-sequence composed of consecutively adjacent equal elements of the sequence. To specify the randomness, on the one hand, the number of runs consisting of "+" and "-" should be close, and, on the other hand, randomness reflects that the number of runs should not be too many or too few. The null hypothesis of the Run Test is that data in time series are independent and identically distributed given a distribution function, which implies stationarity. Let N denote the sample time series size where the number of "+" data points is N_+ and the number of "-" data points is N_-, then we can create the U-statistics based on the total number of runs. Under the null hypothesis, the number of runs is approximately normal with the mean and variance:

$$\begin{aligned} &\text{mean: } \mu = \frac{2N_+N_-}{N} + 1 \\ &\text{variance: } \sigma^2 = \frac{2N_+N_-(2N_+N_- - N)}{N^2(N-1)} = \frac{(\mu-1)(\mu-2)}{N-1} \end{aligned} \tag{3.37}$$

Then the Run Test is given by the following Z-score:

$$z = \frac{R-\mu}{\sigma} \tag{3.38}$$

where R is the observed number of runs.

If the null hypothesis is true, it means that the sample time series well fits the given function. However, it concludes that the sample time series does not evolve according to the function in the case of the null hypothesis is rejected.

To apply the Wald-Wolfowitz test for testing the stationarity, define that the k-th order non-central moment as

$$M_k = \frac{1}{T}\sum_{t=1}^{T} x_t^k$$

and stationarity test checks whether the first k-th order moments are constant in time. To implement the Wald-Wolfowitz test, we need to divide the time series to be tested into w windows and calculate the k-th order non-central moment for each window (window moments), then we test the window moments on the overall moments to figure out whether the window moments are randomly "sampled" by the overall moments. The stationarity

implies that the true null hypothesis that the k-th moment is stationary cannot be rejected. In this essay, we only check the stationarity of the first-order moments.

3.6 Statistical Testing

Based on the flash crash models designed and examined in this thesis to investigate the four main hypotheses for the experimental chapters, we would like to perform statistical testing on the artificial markets introduced learning agents within two aspects: to what extent flash crashes can be mitigated and to what degree the losses of the uninformed fundamental traders can be reduced.

To assess differences between two sets of results under varying circumstances, we will employ independent samples t-tests. These t-tests will rigorously test the hypotheses that there are no significant differences (H_0) or that significant differences exist (H_1) in flash crash sizes or agents' profits. To complement the t-test results, we will calculate Cohen's (1988) d effect sizes. Cohen's d provides a measure for quantifying the difference between two groups, allowing us to assess changes in the metrics under different market mechanisms (Chapter 6) or different information-sharing mechanisms (Chapter 7).

3.6.1 T-Test

A t-test is a statistical test used to determine if there is a significant difference between the means of two groups. It is commonly used when comparing two groups to see if they are different from each other. There are different types of t-tests, and this thesis will apply the independent samples t-test to compare lash crash sizes or agents' profits of different models.

Independent Samples T-Test is also known as the unpaired t-test, it's used to compare the means of two independent groups to see if they are statistically different from each other.

- H_0: no significant difference between the two groups;
- H_1: there is a significant difference between the two groups.

The formula for the t-statistic in an independent samples t-test is:

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \tag{3.39}$$

where:

- $\bar{X}_1$ and $\bar{X}_2$ are the sample means of the two groups.
- S_1^2 and S_2^2 are the sample variances of the two groups.
- n_1 and n_2 are the sample sizes of the two groups.

Once we calculate the t-statistic, we can find the p-value by comparing the t-statistic to the t-distribution. The p-value is a measure of the probability that the results are obtained if there were no real difference between the two groups. It serves as an indicator of the level of statistical significance, which is set to be 5% in this thesis. A smaller p-value (less than 5%) suggests stronger evidence against the null hypothesis H_0.

3.6.2 Cohen's d

Cohen's (1988) d is a widely used and valuable statistic for quantifying the effect size, and it can be suitable to include with t-test. It quantifies the difference between two groups by standardising the difference in means. The formula for Cohen's d is as follows:

$$d = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{S_1^2 + S_2^2}{2}}} \tag{3.40}$$

Where $\bar{X}_1$ and $\bar{X}_2$ are the means of the two groups; S_1^2 and S_2^2 are the variances of the two groups.

Cohen's d doesn't not provide information about whether the difference is statistically significant; instead, it offers information about the magnitude of the difference. Typically, Cohen's d can be interpreted as follows:

- When d is approximately 0, it indicates a very small mean difference between the two data groups, resulting in a small effect size.

- When d falls in the range of approximately 0.2 to 0.5, it suggests a medium effect size.
- When d is greater than 0.5, it signifies a large effect size.
- When d exceeds 1.0, it reflects a very large effect size.

Cohen's d values can be positive or negative. A positive Cohen's d suggests that the first group has a higher mean, while a negative Cohen's d suggests that the second group has a higher mean. By combining the t-test and Cohen's d, this thesis aims to provide a comprehensive evaluation of the differences in certain metrics under various circumstances, such as crash sizes between CDA and batch auction markets. This dual approach helps not only in determining if there are significant differences but also in gaining a better understanding of the impacts of flash crashes in incomplete markets.

Chapter 4

Basic Flash Crash Model

Chapter 3 serves as an introduction to the flash crash model developed for this thesis, which extends the Glosten-Milgrom model and Das's model. Within this basic framework, four different categories of trading agents participate: informed fundamental traders (IFT), uninformed fundamental traders (UFT), zero-intelligence traders (ZIT), and a market maker (MM). These agents engage in interactions within double auction markets.

This chapter primarily focuses on agent-based simulations that build upon the foundational framework established in Chapter 3. To investigate the hypothesis: "spoofing can cause flash crashes or market fluctuations, resulting in altered equilibrium prices, and the spoofer profits from other traders", the core objective of these simulations is to explore the impact of a spoofing event, the fake shock, on market dynamics. By doing so, this research aims to contribute valuable insights into the understanding of market behaviour and the effects of such disturbances.

4.1 Experiments

In the short-term experiments, we assume all trading activities happen in a single trading epoch. This experiment will investigate how the market will perform under the circumstance of a fake shock. Besides, we will discuss the effect of the market structure on the market performance by analysing the stylised facts.

4.1.1 Basic Settings

We assume that there are n_{total} traders excluding the market marker, and the number of ZI traders, IFTs and UFTs are n_{ZI}, n_{IFT}, n_{UFT}, respectively, so $n_{\text{total}} = n_{\text{ZI}} + n_{\text{IFT}} + n_{\text{UFT}}$. A single trading epoch contains $M = 1500$ trading rounds, where we add the external "fake shock" signal at the k-th round. Referred to the setups in Section 3.4.3, the trading epoch can be divided into two stages: **the buffer stage**, containing round 1 to round k, and **the active stage**, containing round k+1 to round M.

4.1.2 Parameters

Volatility Parameters: **The standard deviation (σ_{ε_n}) of the price noise (ε_n)**: each time the fundamental traders can get the signal of stock price. For example, if the true stock price is v_i, each fundamental trader will receive the price signal $w_i = v_i + \varepsilon_n$ as the estimate. Also, although the market maker does not know the price signal, this agent is aware of how the price signal varies, that is, the standard deviation of the price signal ε_n. A higher level of the s.t.d of the stock price signal could suggest a lower validity of the information, and the s.t.d of the price noise is also referred to **information bias rate (IBR)** in some parts of this thesis.

Market-structure Parameters: **The number of informed fundamental traders ($n_{\mathbf{IFT}}$)** and **The number of uninformed traders ($n_{\mathbf{UFT}}$)**. The market structure can be determined as the two parameters are set. If the sum of n_{IFT} and n_{UFT} is too low – the market is mainly occupied by ZI traders, the market prices could be much more random so that the market is less predictable.

4.1.3 Performance Metrics

Based on the introduction in Chapter 3, this chapter will discuss the following metrics in the basic flash crash model and study the stationarity and variability of market prices.

Metrics	Explanation
Crash size	The biggest price decline of the market prices after the fake shock and before all UFTs finish inference
Crash duration	The duration between the shock point and the one reaching the biggest price decline
Bid-ask Spreads	The average bid-ask spreads during before-crash, in-crash and after-crash phases
Cumulative profits	Average cumulative profits of different types of trading agents

Table 4.1 Performance Metrics

4.1.4 Market Shock Experimental Setup

The market price is initialised at 20, and the initial best bid and ask prices are set as 20.1 and 19.9, respectively. All fundamental traders can get the price signal around 20 until additional information is introduced.

The power of shock is -1. Namely, all fundamental traders who believe the shock is real will revise down their estimates on the stock price by 1. The shock power being +1 is symmetric to that of -1, so we only investigate the one-way shock with shock size -1 in this experiment representing a price decline caused by the shock. The following, in detail, illustrates the effect of different external signals on the market.

1. A real shock: all fundamental traders change their estimates from 20 to 19;
2. A fake shock: all informed fundamental traders remain the previous estimates 20, while all uninformed fundamental traders change their estimates from 20 to 19;

The simulation length is set to be 1500 steps, and the fake shock signal point is 500. Namely, the steps interval [1,500] is the buffer stage, while the active stage is [501, 1500], as shown in Fig. 4.1.

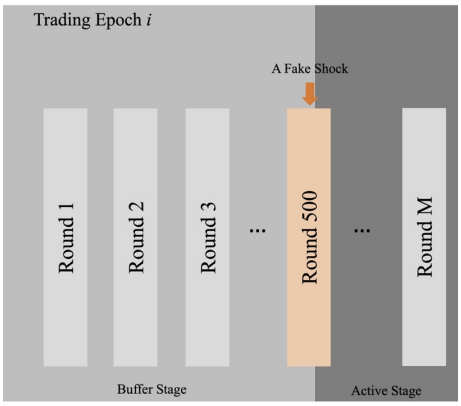


Figure 4.1 A Fake Shock Hitting the Market

The total number of traders is 10, i.e., $n_{\text{total}} = n_{\text{ZI}} + n_{\text{IFT}} + n_{\text{UFT}} = 10$, and the number of the informed fundamental traders n_{IFT} varies from 1 to 9, and the number of uninformed fundamental traders is set to be from [1, n_{total}-n_{IFT}], which are 54 sets of parameters in total. To investigate the market performance among different market noises, we set three groups of simulations where the value of the standard deviation of the noise is selected from $\{0.05, 0.1, 0.25\}$. We apply Monte-Carlo methods for each simulation which contains 100 replications given a set of parameters.

4.2 Results

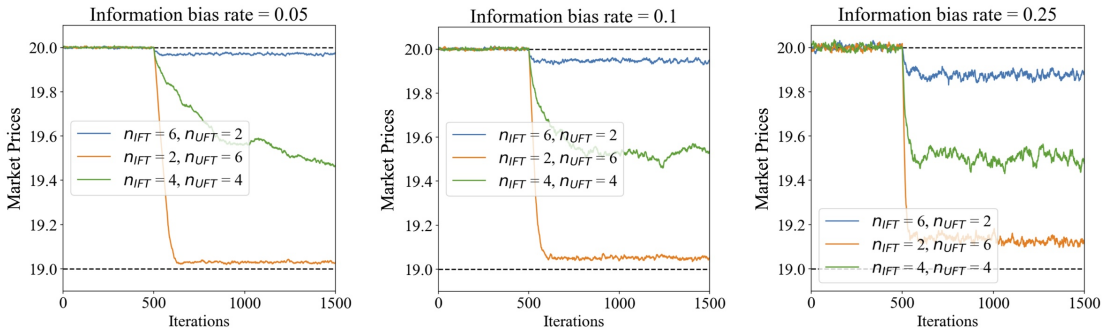
4.2.1 Global Statistics Overview

In this agent-based flash crash model, to start with investigating the global statistics, the most straightforward indicator is the market prices to evaluate the convergence. Sequences of transactions in each simulation drive the market prices, which may experience some level of price flash crash after a fake shock affects the market. We consider three different configurations of market structures in the 10-agent system:

- **IFT-dominated market:** the informed traders dominate the market: 6 informed traders, 2 uninformed traders and 2 zero-intelligence traders are in the market;
- **UFT-dominated market:** the uninformed traders dominate the market: 2 informed traders, 6 uninformed traders and 2 zero-intelligence traders are in the market;

- **Information-balanced market:** uninformed and informed traders have the same number: 4 informed traders, 4 uninformed traders and 2 zero-intelligence traders are in the market.

Price Movements



(a) IBR $\sigma_{\varepsilon_n} = 0.05$ (b) IBR $\sigma_{\varepsilon_n} = 0.1$ (c) IBR $\sigma_{\varepsilon_n} = 0.25$

Figure 4.2 The Average Market Prices Movements

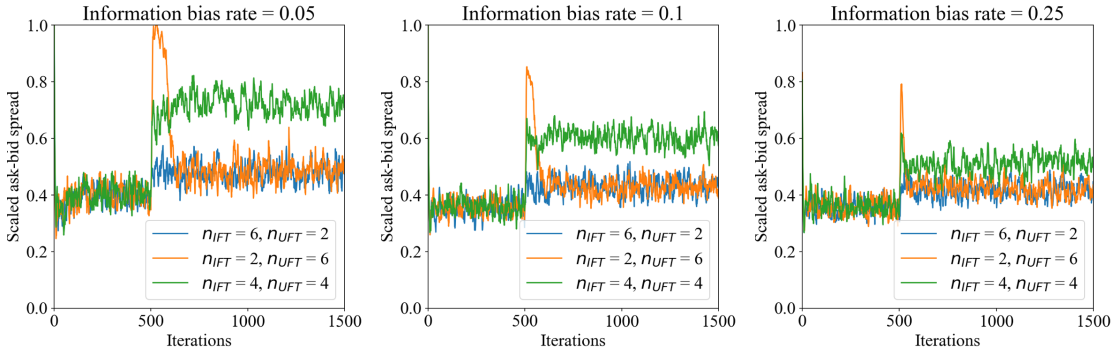
The information bias rates are selected over $\{0.05, 0.1, 0.25\}$. Fig. 4.2 shows the examples of the average market prices among all simulations under the given parameter settings with the comparison among different information bias rates. From Fig. 4.2, the market prices seem to converge to certain levels, although the stable prices after the fake shock differ among different market structure settings. A more IFT-dominated or less UFT-dominated market is less affected by the fake shock. In addition, the information bias rates seem also to affect the stable prices after the shock – a noisier IFT-dominated market suffers a more significant crash. Therefore, the global volatility of a market may have two sides. Noise will exaggerate the negative impact on an IFT-dominated market but shrink the crash in a UFT-dominated market.

Bid-ask Spreads

To a certain extent, the bid-ask spreads measure how much a market maker benefits from the market. However, a noisier market could also cause a wider spread due to higher volatility. In this section, to investigate how the market structure affects the market maker's

profits, the bid-ask spreads are scaled by the information bias rate to eliminate the market volatility effect.

$$\text{Scaled bid-ask spread} = \frac{\text{Bid-ask spread}}{\sigma_{\varepsilon_n}}$$



(a) IBR $\sigma_{\varepsilon_n} = 0.05$ (b) IBR $\sigma_{\varepsilon_n} = 0.1$ (c) IBR $\sigma_{\varepsilon_n} = 0.25$

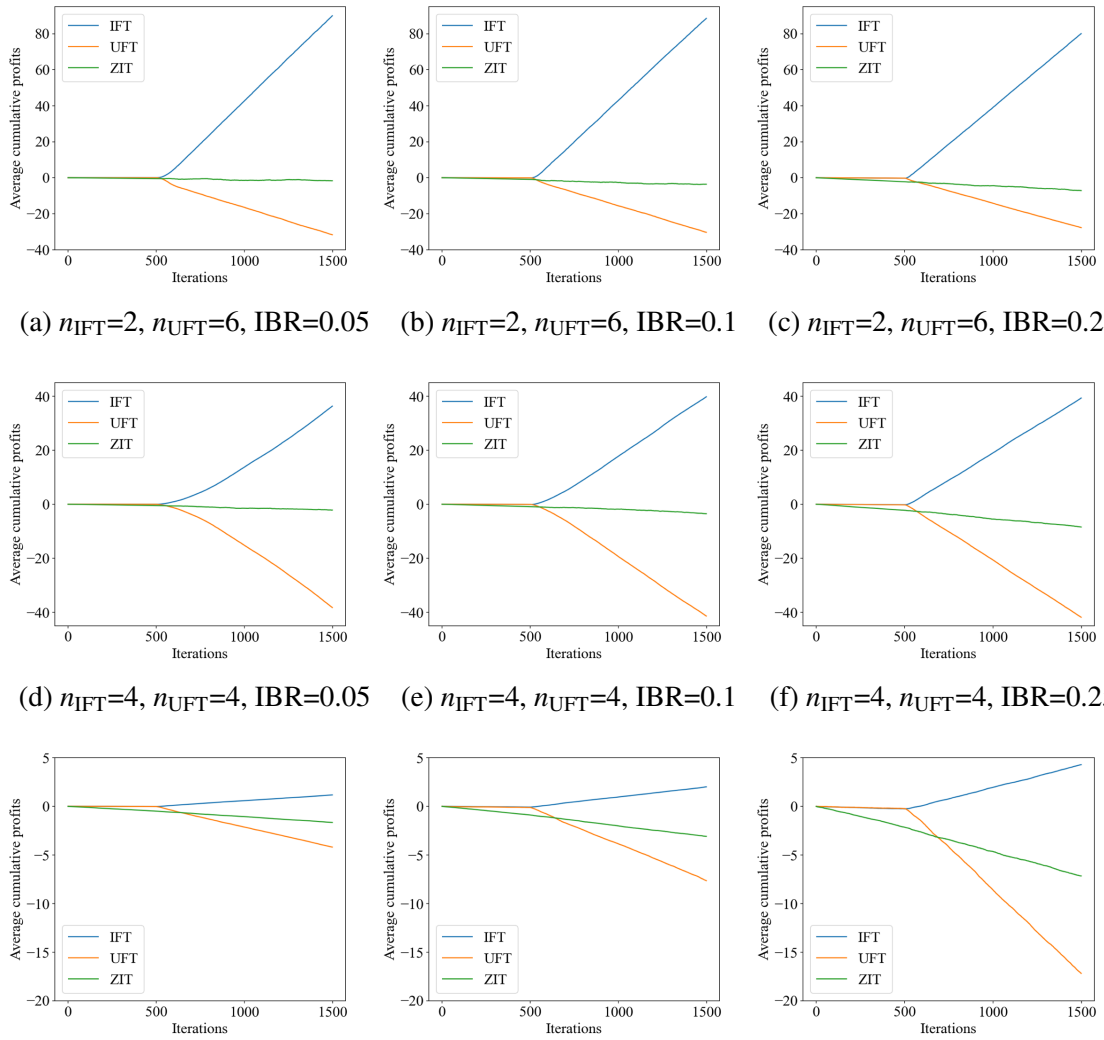
Figure 4.3 The Average Bid-Ask Spreads Changes

From Fig.4.3, an interesting finding is that before the fake shock happens, the scaled bid-ask spreads are similar among all different parameter settings. In contrast, the spreads show apparent differences after the fake shock if the parameter settings vary, while the differences narrow as the information bias rate increases. Plus, both informed and uninformed markets have less wide scaled bid-ask spreads than the balanced markets, in which case the market maker could earn more. However, the market maker could be more challenging to benefit from market volatility in a highly noisy market. In addition, we can observe a sudden spike in the bid-ask spread followed by a decrease in the spread for UFT-dominated markets. This pattern is similar to the results in Das (2005) shortly after a price jump occurs. We reasonably infer that in a UFT-dominated market, many UFTs adjust their order placement strategy after the fake shock, prompting the market maker to widen the bid-ask spread in defense against adverse selection risk. Subsequently, the market maker continues to adjust quotes by learning from order flows, narrowing the bid-ask spreads.

Profits of Agents

Fig. 4.4 compares the average cumulative profits per trader among different types of traders. Market volatility does not have a structural impact on allocating total profits but can amplify their profits and losses. However, we can observe that, for different configurations of market structures, the profit structure differs clearly, both in terms of average cumulative profits and total cumulative profits for a certain trader type.

For example, in the case of IBR = 0.25, the average cumulative profits of an IFT drop dramatically from 40 in the scenario of $n_{\text{IFT}} = 4, n_{\text{UFT}} = 4$ to 5 in the scenario of $n_{\text{IFT}} = 6, n_{\text{UFT}} = 2$. However, the total cumulative profits for IFTs in both scenarios of $n_{\text{IFT}} = 2, n_{\text{UFT}} = 6$, and $n_{\text{IFT}} = 4, n_{\text{UFT}} = 4$ are similar, approximately 160. Additionally, UFTs experience fewer losses in UFT-dominated markets, both in terms of average and total losses. These preliminary findings provide insights into how market structure influences the agents' profit structure.



(a) n_{IFT}=2, n_{UFT}=6, IBR=0.05 (b) n_{IFT}=2, n_{UFT}=6, IBR=0.1 (c) n_{IFT}=2, n_{UFT}=6, IBR=0.25

(d) n_{IFT}=4, n_{UFT}=4, IBR=0.05 (e) n_{IFT}=4, n_{UFT}=4, IBR=0.1 (f) n_{IFT}=4, n_{UFT}=4, IBR=0.25

(g) n_{IFT}=6, n_{UFT} = 2, IBR=0.05 (h) n_{IFT}=6, n_{UFT} = 2, IBR=0.1 (i) n_{IFT}=6, n_{UFT} = 2, IBR=0.25

Figure 4.4 Agents' Profits Dynamics

4.2.2 Sensitivity Analysis

In this section, compared to the analysis of the global and dynamic statistics in Section 4.2.1, we investigate the sensitivity analysis of the indicators in the steady state of the agent-based flash crash model. Sensitivity analysis can show which parameters of a model, or combinations of parameters, have the most significant impact on the metrics and how strongly the interaction of the parameters affect the model performance.

Symbol	Description	Value
n_{IFT}	the number of the informed fundamental traders	{1,2,..,9}
n_{UFT}	the number of the uninformed fundamental traders	{1,2,...,10 - n_{IFT}}
n_{ZIT}	the number of the zero-intelligence traders	10 - n_{IFT} - n_{UFT}
σ_{ε_n}	the information bias rate	{0.05, 0.1, 0.25}

Table 4.2 Parameters

In the 10-agent model, we mainly examine the effect of these two types of parameters on the system: information parameters and market-structure parameters. The relative parameters are specified as Table 4.2:

As the sum of the numbers of informed fundamental traders (IFTs), uninformed fundamental traders (UFTs), and zero-intelligence traders (ZITs) is fixed to be 10, we can visualize how the market structure affects the system with the ternary plot. As Fig. 4.5 shows, in a ternary plot, there are three axes representing the number of IFTs, UFTs, and ZITs on the top axis, left axis and right axis, respectively. Point A represents the parameters setting as $n_{\text{IFT}} = 3$, $n_{\text{UFT}} = 2$ and $n_{\text{ZIT}} = 5$, with which the heat colour in the ternary heatmap indicates the level of the corresponding metrics.

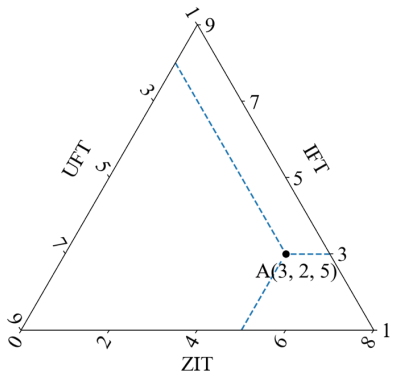
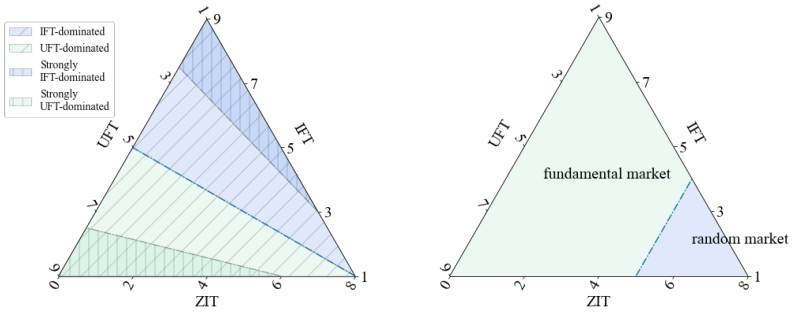


Figure 4.5 Ternary Plot Example

To better explain the market structure, we have made divisions of the ternary plot to stand for different situations. Firstly, we have divided the market into four categories based on the imbalance between the number of informed fundamental traders and uninformed fundamental traders in the market: strongly IFT-dominated markets ($n_{\text{IFT}} \geq 3n_{\text{UFT}}$), weakly

IFT-dominated markets ($n_{\mathrm{UFT}} < n_{\mathrm{IFT}} < 3n_{\mathrm{UFT}}$), strongly UFT-dominated markets ($n_{\mathrm{UFT}} \geq 3n_{\mathrm{IFT}}$) and weakly UFT-dominated markets ($n_{\mathrm{IFT}} < n_{\mathrm{UFT}} < 3n_{\mathrm{IFT}}$), as Fig. 4.6 shows.

Furthermore, based on the difference between the number of fundamental traders and zero-intelligence traders, all markets can be divided into two types: the fundamental markets, if the fundamental traders dominate, and the random markets, if the zero-intelligence traders dominate.



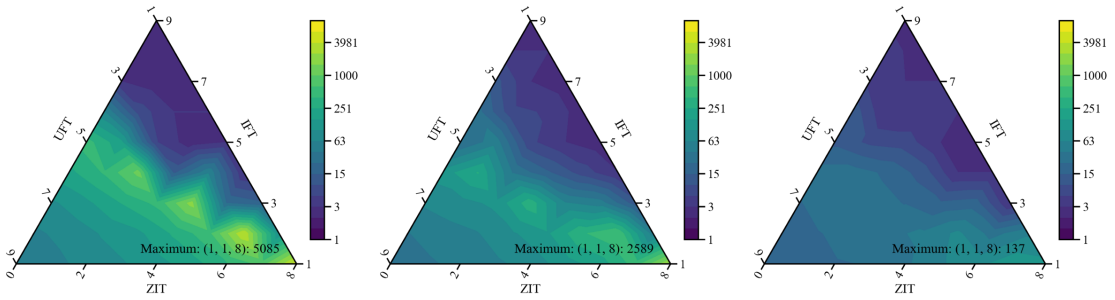
(a) (b)

Figure 4.6 Graphic Explanation of the Market Structure

Crash Sizes and Crash Duration

Fig. 4.2 reveals that the market prices crash to a certain level in the active stage. To determine the steady state for each run of simulations, we should figure out at which point the market prices move to the stable level, called the early stop point in Section 3.5.2. Therefore, we use the MSER-5 method which is specified in Chapter 3 to determine the global early stop points d for all simulations. The initial market price is set to be 20, and the shock point is 500, so the crash size and duration can be respectively calculated as $|v(d) - v(0)|$ and $d - 500$.

Based on the calculation of the crash size and duration specified in Section 3.5.2, Fig. 4.7 and Fig. 4.8 show the triangular contour heatmaps of the the average crash duration and the crash sizes in a trading epoch under different market structures as the information bias rate varies.

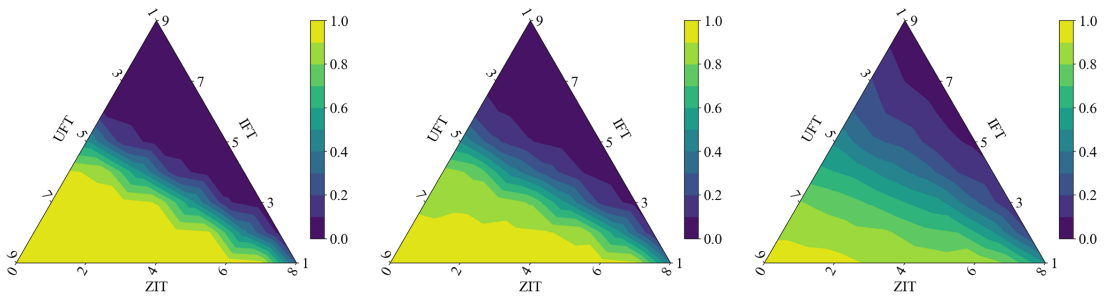


(a) IBR $\sigma_{\varepsilon_n} = 0.05$ (b) IBR $\sigma_{\varepsilon_n} = 0.1$ (c) IBR $\sigma_{\varepsilon_n} = 0.25$

Figure 4.7 The Crash Duration under Different Parameter Settings

Determined by MSER-5 method, the crash duration varies dramatically at different levels of market structure. In the case of IBR $\sigma_{\varepsilon_n} = 0.05$, when both the numbers of informed and uninformed traders equal to be 1, the market prices drop very slowly, making the crash duration last 5085 steps on average, while the minimised crash duration is only 3 when the numbers of informed and uninformed traders are 5 and 1 respectively. Therefore, we apply a logarithm log_{10} on the crash size for better visualisation of the contour plots. As the market noise (information bias rate, IBR) increases from 0.05 to 0.1 to 0.25, the maximised crash duration dramatically decreases from 5085 to 2589 to 137, respectively, and all corresponding market structure parameters are where the numbers of informed and uninformed traders are very small.

Another interesting finding is that the more "imbalanced" the numbers of informed and uninformed traders are, the shorter the crash lasts. However, when informed traders dominate the market, the crash duration for each market seems to drop faster as the "imbalance" increases. This will be further discussed by calculating the crash speed.



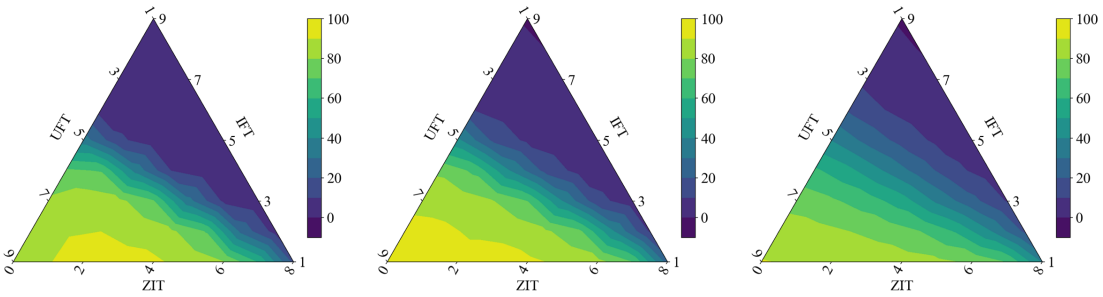
(a) IBR $\sigma_{\varepsilon_n} = 0.05$ (b) IBR $\sigma_{\varepsilon_n} = 0.1$ (c) IBR $\sigma_{\varepsilon_n} = 0.25$

Figure 4.8 The Crash Sizes under Different Parameter Settings

According to Fig. 4.8, we can see that contour lines are denser in a balanced market when the information bias rate is low. Thus, when the market has a low level of volatility, if the number of informed traders dominates the market, the crash size will be tiny and not sensitive to the market structure changes; however, if the number of uninformed traders in the market exceeds the number of informed traders, the price decline will suffer a large and abrupt decline.

However, as the market noise (information bias rate, IBR) increases, the area referring to a balanced market becomes less clear, which means the crash size is less sensitive to the market structure changes in the case of the number of informed traders close to the number of uninformed traders. However, the crash size becomes more sensitive to the market structure changes when the numbers of informed and uninformed traders apparently differ.

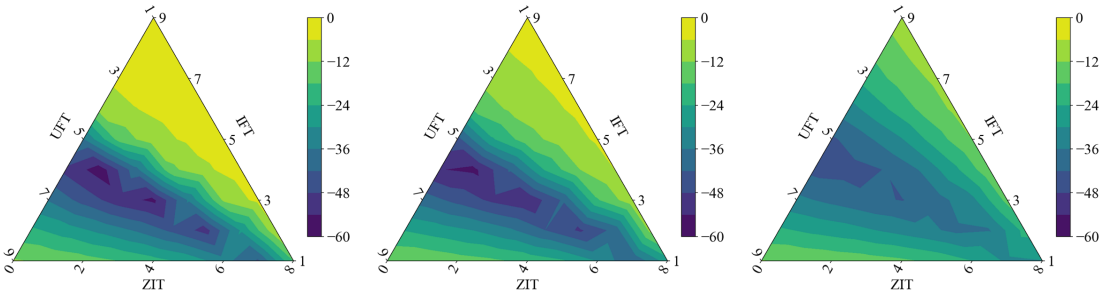
Agents' Average Profits



(a) IBR $\sigma_{\varepsilon_n} = 0.05$ (b) IBR $\sigma_{\varepsilon_n} = 0.1$ (c) IBR $\sigma_{\varepsilon_n} = 0.25$

Figure 4.9 The Average Profits Earned of a Single Informed Fundamental Trader

Fig. 4.4 shows that when the fake shock arrives, The informed fundamental traders can grab benefits from the uninformed fundamental traders and ZI traders. Through sensitivity analysis in this section, we find that the effect of market structure on the average profits of a single informed fundamental trader is clearly revealed shown by Fig. 4.9. If the number of IFTs remains constant, the average profits of an IFT increase as the number of UFTs increases. In contrast, if the number of UFTs remains constant, the average profits of an IFT increase as the number of IFTs decreases.



(a) IBR $\sigma_{\varepsilon_n} = 0.05$ (b) IBR $\sigma_{\varepsilon_n} = 0.1$ (c) IBR $\sigma_{\varepsilon_n} = 0.25$

Figure 4.10 The Average Profits Earned of a Single Uninformed Fundamental Trader

Unlike the negatively correlated monotonic relationship between the number of IFTs and the average profits of a single IFT, the average profits of a UFT could experience bigger losses in a more balanced market. While the average profits increase as the market is increasingly UFT-dominated or IFT-dominated.

Bid-ask spreads

As Das (2005); Das and Magdon-Ismail (2008) specifies, the bid-ask spread reflects to what degree the market maker is uncertain about the true stock price. As the market has different kinds of traders with different levels of information advantages, the market maker cannot recognise which kind of trader is traded with at each trading round, causing an adverse selection issue. The market maker will have a loss if trading with the informed fundamental traders, so the market maker will set a wider spread to cover the possible loss. When the fake shock comes, the UFTs change their previous actions into selling, and the market maker, with the risk-neutral algorithm, has to set the quotes down consecutively to reach a new break-even state. As the trading goes, the market maker can fully use the information from the previous order flows to set proper quotes to balance both sides.

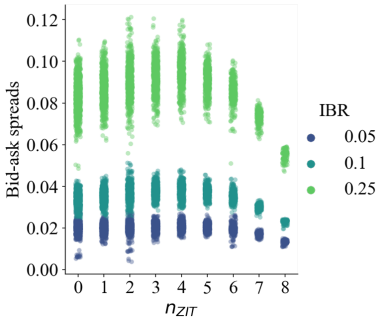
In this model, a single simulation can be divided into three phases:

- **Before-crash phase**: all trading rounds from the simulation start to the fake shock step;
- **In-crash phase**: all trading rounds between the fake shock step to the early stop point (clarified in the section);

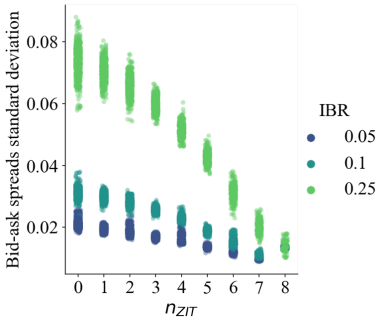
- **After-crash phase**: all trading rounds from the early stop point to the end.

Before-crash phase: In the before-crash phase, the informed and uninformed fundamental traders act indifferently, so only two parameters affect the bid-ask spreads. Fig. 4.11 shows the distribution of averaged bid-ask spreads in the before-crash phase among all simulations in the markets with different numbers of zero-intelligence traders.

Fig. 4.11(a) shows that, in the before-crash phase, the bid-ask spreads are higher in a noisier market; the spreads decrease as the number of zero-intelligence traders increases. More specifically, if the zero-intelligence highly dominates the market (when $n_{\text{ZIT}} \geq 6$), the bid-ask spreads drop more quickly as n_{ZIT} increases. Also, Fig. 4.11(b) shows that the bid-ask spread standard deviation has a negative relationship with the number of ZITs. In summary, the market maker faces higher adverse selection risks if more fundamental traders are in the market.



(a) Bid-ask spreads distribution

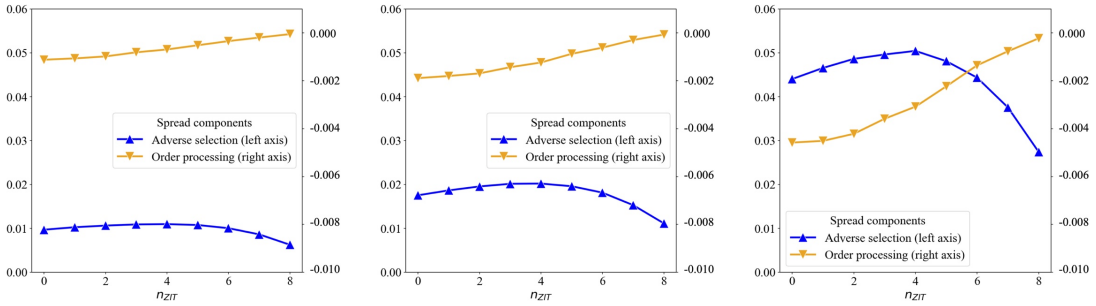


(b) Bid-ask spreads standard deviation distribution

Figure 4.11 The Bid-ask Spreads Distribution in the Before-crash Phase

As specified in Section 3.5.2, we apply the Glosten-Harris (Glosten and Harris, 1988)/MRR model (Madhavan et al., 1997) to decompose the bid-ask spreads into adverse selection and order processing costs. The order processing costs refer to the premiums for market-making services, and the adverse selection costs refer to the premiums of trading with the traders who have information advantages (Glosten and Milgrom, 1985; Stoll, 1989). By applying the linear regression model to obtain the attributes of adverse selection $\{Z_t\}$ and order processing $\{C_t\}$, we can analyse how the two components account for the bid-ask spreads. Fig. 4.12 shows how the levels of adverse selection costs and order

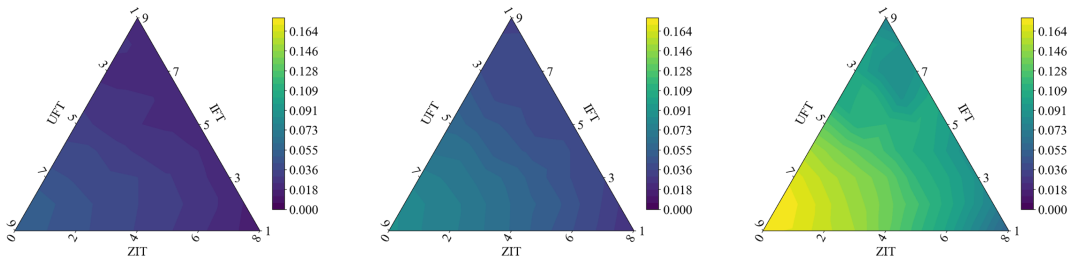
processing costs vary as the numbers of zero-intelligence traders change in the different markets. First, we find that the effect of the number of ZITs on the order processing cost is monotonic: the order processing cost decreases as the number of ZITs increases. However, the adverse selection cost is inverted U-shaped, reaching the highest point as the number of ZITs is about 4. As the information bias rate of the market increases, both adverse selection and order processing in terms of absolute values increase. The adverse selection cost accounts more for the bid-ask spread in our model, making the bid-ask spreads curve also inverted U-shaped shown in Fig. 4.11(a).



(a) IBR $\sigma_{\varepsilon_n} = 0.05$ (b) IBR $\sigma_{\varepsilon_n} = 0.1$ (c) IBR $\sigma_{\varepsilon_n} = 0.25$

Figure 4.12 The Levels of the Bid-ask Spreads Components in the Before-crash Phase

In-crash phase: During the crash, the fundamental traders are differentiated into two types: informed fundamental traders and uninformed fundamental traders. The two different types of traders receive different levels of price signals, leading them to submit orders in opposite directions. The contours of Fig. 4.13 show the patterns of the bid-ask spreads during the crash period when the information bias rate varies.



(a) IBR $\sigma_{\varepsilon_n} = 0.05$ (b) IBR $\sigma_{\varepsilon_n} = 0.1$ (c) IBR $\sigma_{\varepsilon_n} = 0.25$

Figure 4.13 The Average Bid-ask Spreads in the In-crash Phase

Firstly, as is consistent with our intuition, the bid-ask spreads are wider when the market becomes noisier. What is more, we can see that the contour lines are roughly parallel with respect to the right side of each triangle plot. It means that bid-ask spreads are sensitive to the number of UFTs rather than to the number of IFTs, and the spreads increase as the number of UFTs increases. By calculating the correlation, we can obtain that the correlation between n_{UFT} and spreads exceeds 0.95 for each market (see Fig. 4.14), while the correlation between n_{IFT} and spreads is less than 0.5 (see Fig. 4.15). In fact, in the in-crash phase, only the uninformed fundamental traders change their actions compared to the before-crash phase. In contrast, the informed fundamental traders' and zero-intelligence traders' submissions remain unchanged under the set trading rules.

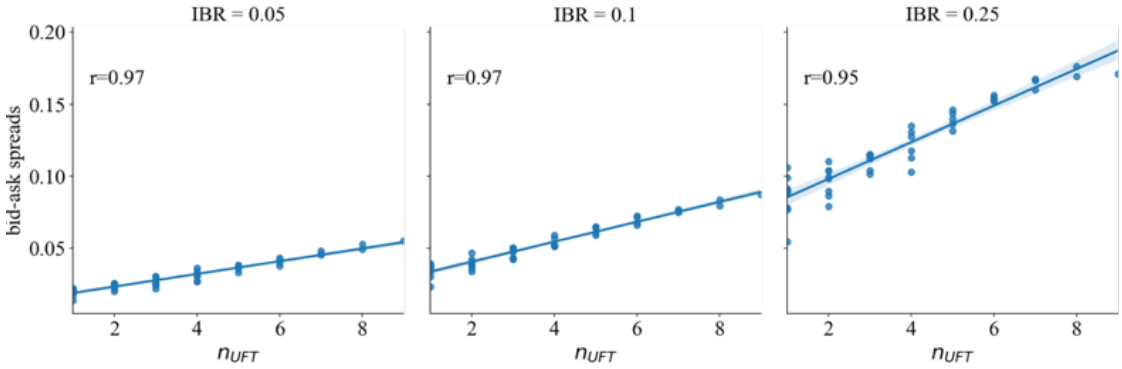


Figure 4.14 Regression Plots of Bid-ask Spreads on n_{UFT}

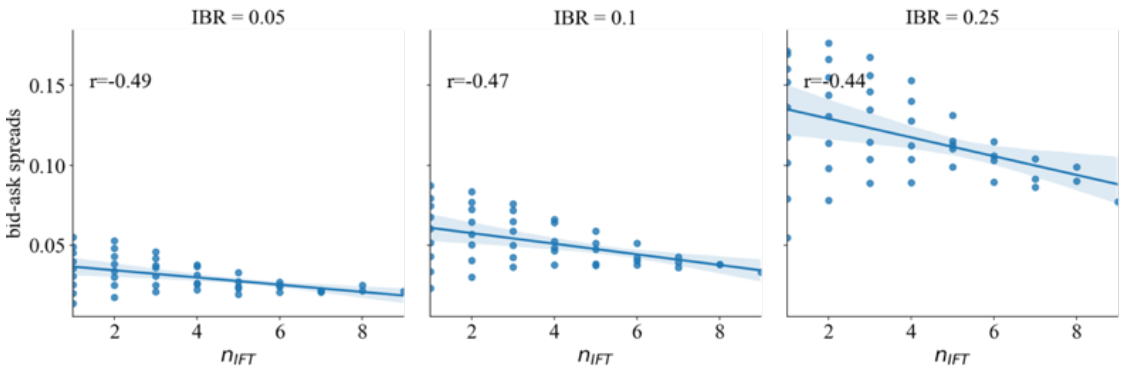
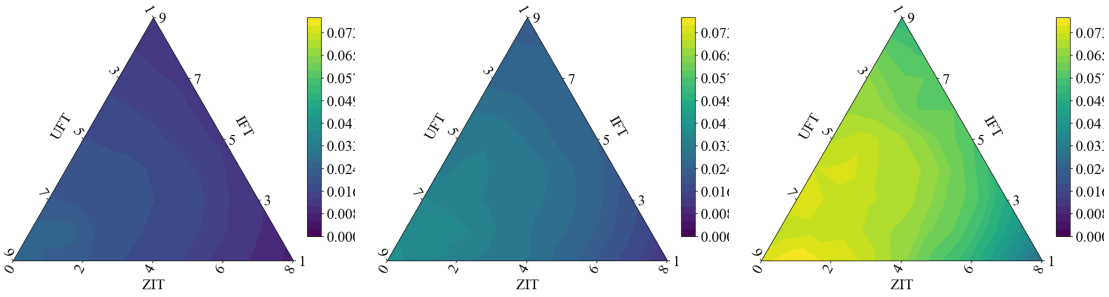


Figure 4.15 Regression Plots of Bid-ask Spreads on n_{IFT}

We still decompose the bid-ask spreads in the in-crash phase into two components: adverse selection cost and order processing cost by the GH model. Similar to Fig. 4.12 shown, the patterns of adverse selection cost and order processing cost are opposite, but

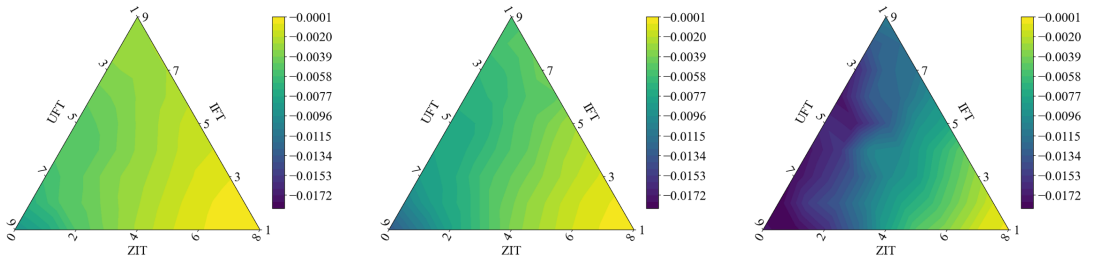
the adverse selection cost still dominates. According to the heatmaps in Fig. 4.16, both the adverse selection and order processing costs depend on the market structure. Also, by observing the trends of the contour lines of the heatmaps regarding adverse selection, we find that, with the point $(n_{\mathrm{IFT}}, n_{\mathrm{UFT}}) = (1, 9)$ as the centre, the contour represents a lower value if it is further away from the centre point.



(a) IBR $\sigma_{\varepsilon_n} = 0.05$ (b) IBR $\sigma_{\varepsilon_n} = 0.1$ (c) IBR $\sigma_{\varepsilon_n} = 0.25$

Figure 4.16 The Adverse Selection Cost of the Market Maker in the In-crash Phase

In contrast, the contour lines of the order processing cost heatmaps are relatively left-right symmetric. The further the contour lines are to the left (low n_{ZIT} and high n_{UFT}), the lower values they represent, and the further the contour lines are to the right (high n_{ZIT} and low n_{IFT}). Compared with the adverse selection affected by the combination of n_{IFT}, n_{UFT} and n_{ZIT}, the order processing cost is mainly affected by the number of zero-intelligence traders or the total number of fundamental traders.

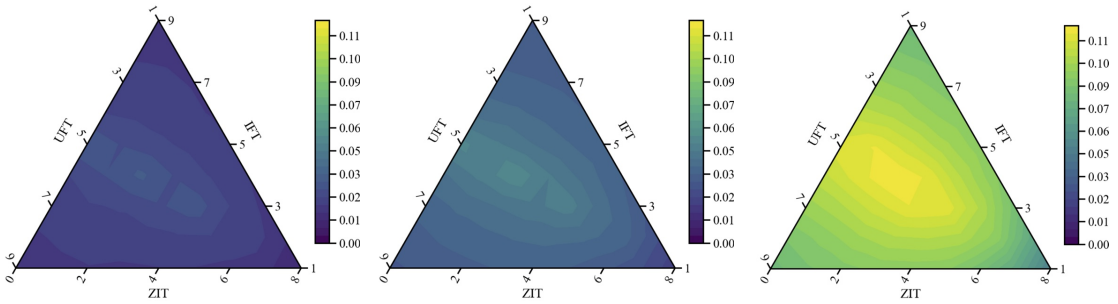


(a) IBR $\sigma_{\varepsilon_n} = 0.05$ (b) IBR $\sigma_{\varepsilon_n} = 0.1$ (c) IBR $\sigma_{\varepsilon_n} = 0.25$

Figure 4.17 The Order Processing Cost of the Market Maker in the In-crash Phase

After-crash phase: In the after-crash phase, the market prices move into a new steady state, where the informed traders and uninformed traders submit orders in opposite direc-

tions. Similarly, we plot the heatmaps of the bid-ask spreads when the market structures and information bias rates change. Unlike the patterns in the in-crash phase, the contours show that a maximum point is around $(n_{\text{IFT}}, n_{\text{UFT}}) = (4,4)$, and the contours further away from such point represent lower values.

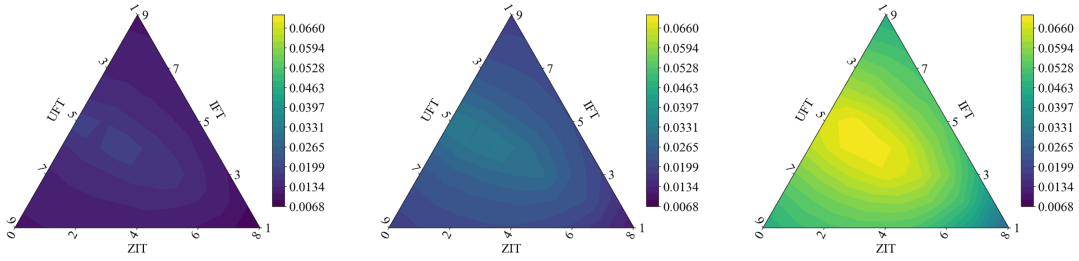


(a) IBR $\sigma_{\varepsilon_n} = 0.05$ (b) IBR $\sigma_{\varepsilon_n} = 0.1$ (c) IBR $\sigma_{\varepsilon_n} = 0.25$

Figure 4.18 The Average Bid-ask Spreads in the After-crash Phase

Another interesting finding is that the contours are almost symmetrical along the line $n_{\text{IFT}} = n_{\text{UFT}}$, which implies a relationship between the bid-ask spreads and market imbalance.

Similar to the analysis in the previous two parts, we decompose the bid-ask spreads into the adverse selection costs and order processing costs. The patterns of adverse selection costs are similar to those of bid-ask spreads – the maximum is around $(n_{\text{IFT}}, n_{\text{UFT}}) = (4,4)$ for all markets with different information bias rates.

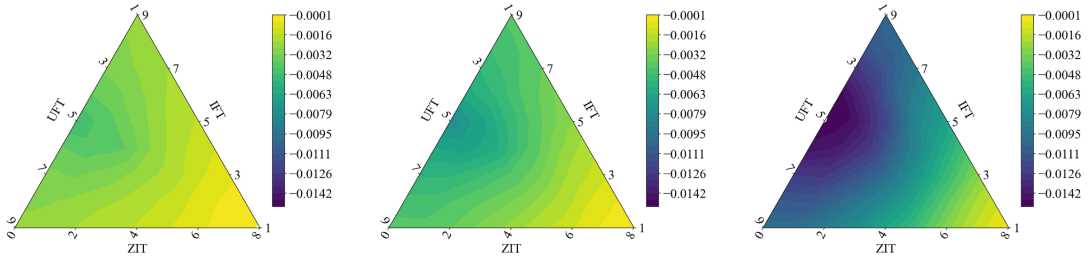


(a) IBR $\sigma_{\varepsilon_n} = 0.05$ (b) IBR $\sigma_{\varepsilon_n} = 0.1$ (c) IBR $\sigma_{\varepsilon_n} = 0.25$

Figure 4.19 The Adverse Selection Costs in the After-crash Phase

As the adverse selection costs dominate in the bid-ask spreads, the patterns imply similar findings where the contours of adverse selection costs are almost symmetrical along the line $n_{\text{IFT}} = n_{\text{UFT}}$.

However, the patterns of order processing are quite different from those of bid-ask spreads or adverse selection costs. The contours show a minimum point around $(n_{\text{IFT}}, n_{\text{UFT}}) = (5,5)$, and the contours further away from such point represent higher order processing costs. We can also observe that the contours of order processing costs are also almost symmetrical along the line $n_{\text{IFT}} = n_{\text{UFT}}$, which implies that the order processing costs of two markets could be at the same level if they are similarly imbalanced.



(a) IBR $\sigma_{\varepsilon_n} = 0.05$ (b) IBR $\sigma_{\varepsilon_n} = 0.1$ (c) IBR $\sigma_{\varepsilon_n} = 0.25$

Figure 4.20 The Order Processing Costs in the After-crash Phase

4.2.3 Equilibrium Analysis

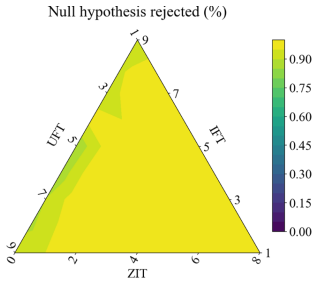
Stationarity Test

Stationarity tests assess the robustness of the results in terms of stability. The experiments are carried out by the two-tail test, where the confidence level is 95%. The null hypothesis is that the moment of the time series is constant, and the moments of window-resized subseries are also fitted by the overall moment. Therefore, the rejection of the null hypothesis means that a set of observations are not randomly allocated around a certain value or function, implying that the time series is not stationary.

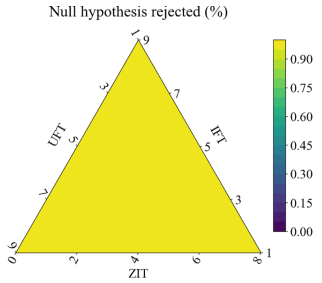
To run the stationarity, we must decide the window size to calculate the window moments. The null hypothesis implies that the estimation could be better in a longer window in the Wald-Wolfowitz test. However, fewer windows and worse normal distribution approximation could also reduce the accuracy of the estimation. Therefore, we vary the window sizes from short to long and check the estimation performance under all different settings. The time series are chosen to be the observations during the after-crash phase, and the window sizes are selected to be 1, 5, 10, 25 and 50, in which case the numbers

of samples are $n_i^{ac}, n_i^{ac}/5, n_i^{ac}/10$ and $n_i^{ac}/25$, respectively, where n_i^{ac} is the number of iterations in the after-crash phases of the i-th simulation.

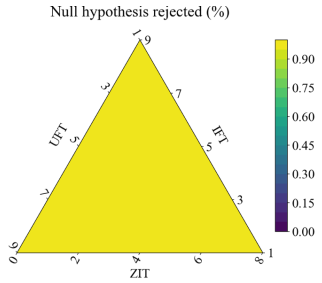
We run Monte-Carlo simulations with 100 runs for each test to check the performance of the Wald-Wolfowitz test. In this experiment, the first-order moment is applied. The following triangle contours show the Wald-Wolfowitz test results of the time series in after-crash phases with different values of information bias rates. The contours show the different levels of percentages of null-hypothesis-rejected cases. With high percentages of null-hypothesis rejection, the stock prices under a particular parameter setting after the crash could be highly likely non-stationary. In contrast, low percentages of null-hypothesis rejection could imply that the stock prices will converge to a stable level after the crash.



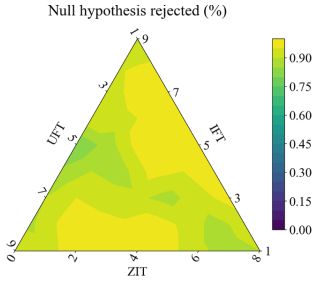
(a) window size = 1, IBR=0.05



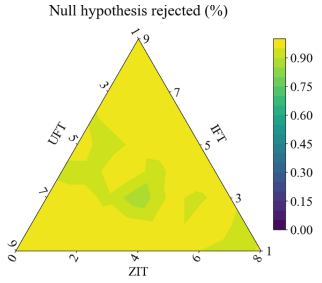
(b) window size = 1, IBR=0.1



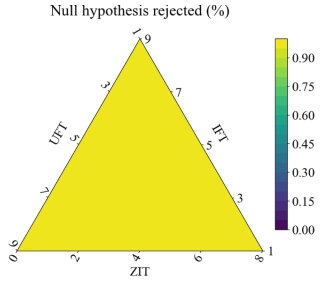
(c) window size = 1, IBR=0.25



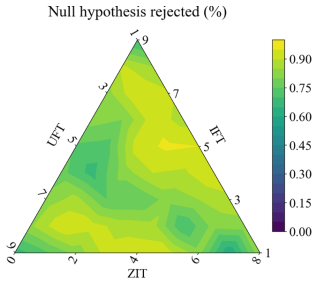
(d) window size = 5, IBR=0.05



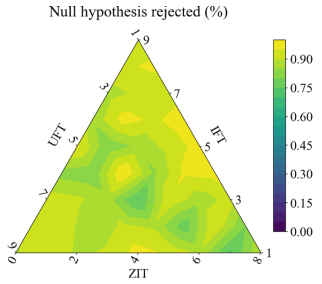
(e) window size = 5, IBR=0.1



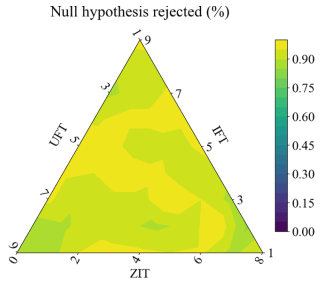
(f) window size = 5, IBR=0.25



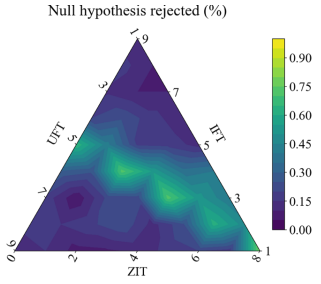
(g) window size = 10, IBR=0.05



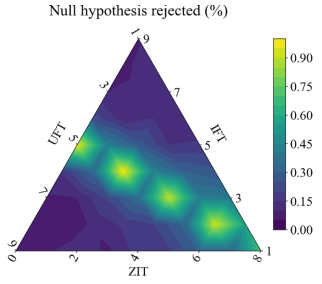
(h) window size = 10, IBR=0.1



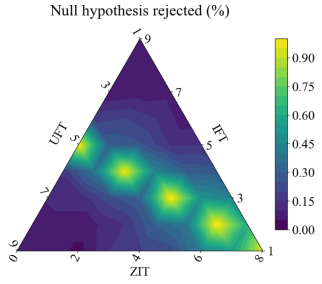
(i) window size = 10, IBR=0.25



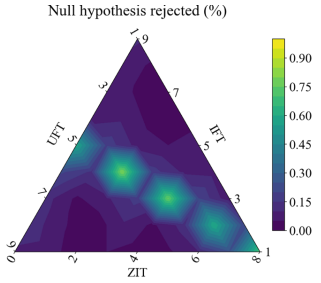
(j) window size = 25, IBR=0.05



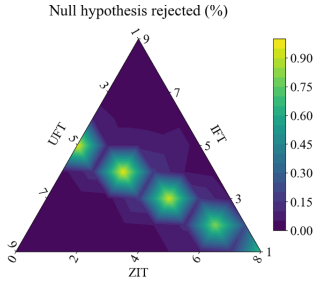
(k) window size = 25, IBR=0.1



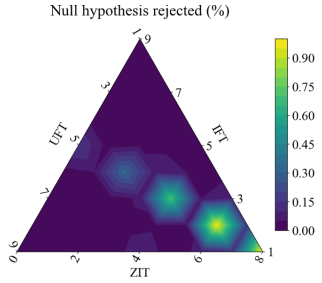
(l) window size = 25, IBR=0.25



(m) window size = 50, IBR=0.05



(n) window size = 50, IBR=0.1



(o) window size = 50, IBR=0.25

Figure 4.21 Stationarity Test Results

By slicing and resampling the time series according to the window size, we are actually investigating the stationarity of the time series at different frequencies. The original time series (window size 1) is high-frequency, and the window sizes of 5 and 10 are referred to as mid-frequency. The window sizes of 25 and 50 imply low-frequency time series.

The heatmaps show that the after-crash stock prices in high frequency and mid frequency are highly likely non-stationary, even though the global average prices converge after the crash. The areas where $n_{\text{IFT}} = n_{\text{UFT}}$ are in dark green when the window size is 10, but the corresponding percentage is still as high as 70%, showing much uncertainty in the stock prices.

However, the patterns differ in the case of low frequency. When the window sizes are 25 and 50, there is a bright "band" in the triangle heatmaps representing the points which are close to the ones representing $n_{\text{IFT}} = n_{\text{UFT}}$. Therefore, from the perspective of low frequency, the time series are highly likely stationary when there is a large difference between the quantities of informed fundamental traders and uninformed fundamental traders, in which case the low-frequency data could be easy to predict. However, when the numbers of informed traders and uninformed traders are almost the same, the time series after the crashes are still non-stationary. Instead, the balance of informed and uninformed traders brings more uncertainty to the market.

Variability Test

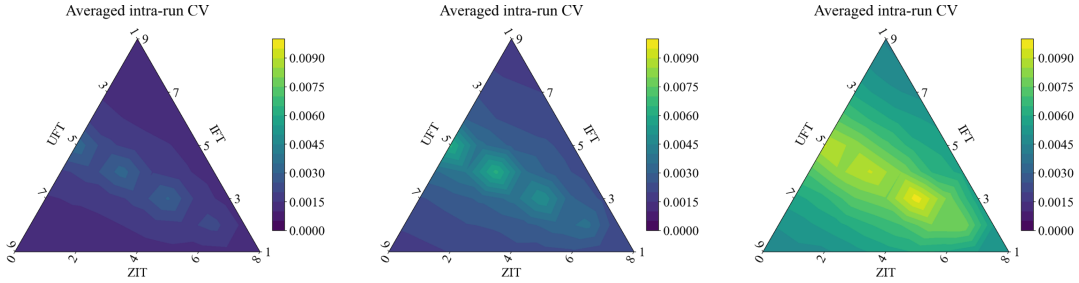
The Coefficient of Variability (CV), defined as the sample standard deviation divided by the sample mean, measures the time series variability well. In Monte-Carlo simulations, we often apply two measures of the Coefficient of Variability: intra-run CV and inter-run CV.

In a well-behaved and stationary time series, the intra-run CV and the inter-run CV account similarly, giving a ratio measure variability ratio defined by

$$VR = \frac{\text{Sum of Intra-run CV}}{\text{Sum of Inter-run CV}} = \frac{\sum_N CV_T^n}{\sum_T CV_t^N} \tag{4.1}$$

In this section, the three different types of measures are investigated.

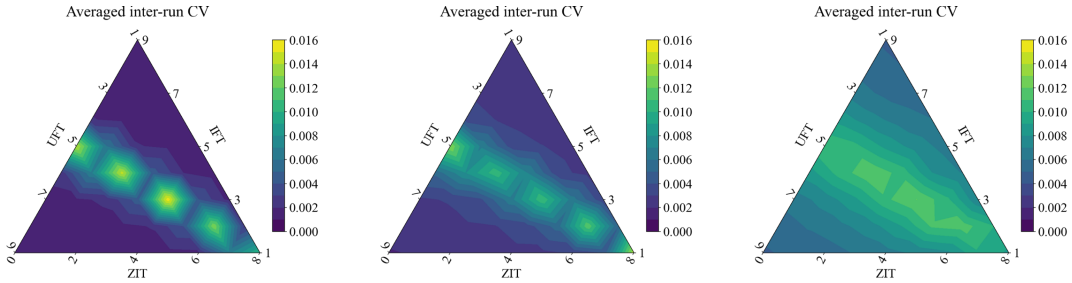
From the perspective of the intra-run CV, it measures the internal effect on the variation of the time series. Fig. 4.22 shows the heatmaps of the averaged intra-run CVs under different parameter settings, revealing clear market structure effects on the variance. The patterns are approximately symmetric along the line $n_{\mathrm{IFT}} = n_{\mathrm{UFT}}$, and the areas closer to the line $n_{\mathrm{IFT}} = n_{\mathrm{UFT}}$ have more significant intra-run CVs. Also, as the information bias rate increases, the overall colours of the three heatmaps become lighter, implying increases in the intra-run CVs.



(a) IBR $\sigma_{\varepsilon_n} = 0.05$ (b) IBR $\sigma_{\varepsilon_n} = 0.1$ (c) IBR $\sigma_{\varepsilon_n} = 0.25$

Figure 4.22 The Averaged Intra-run CV

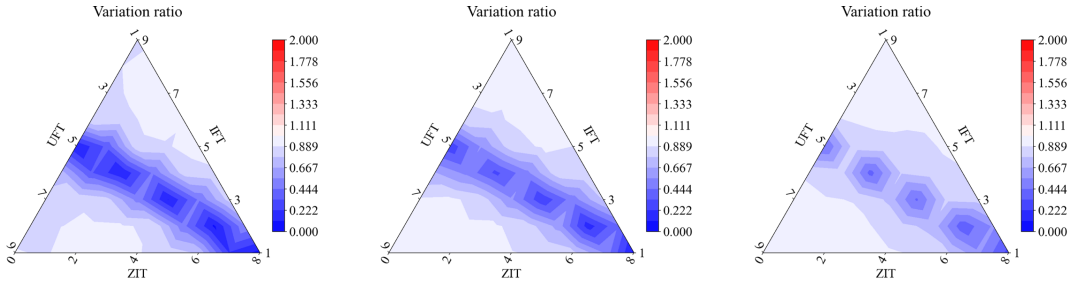
Similarly to the patterns of intra-run CVs, the heatmaps of inter-run CVs are also approximately symmetric along the line $n_{\mathrm{IFT}} = n_{\mathrm{UFT}}$. For every single heatmap in Fig. 4.23, the areas closer to the line $n_{\mathrm{IFT}} = n_{\mathrm{UFT}}$ have the higher values of inter-run CVs. However, an interesting finding is that the colours of the areas away from the line $n_{\mathrm{IFT}} = n_{\mathrm{UFT}}$ become lighter as the information bias rate increases, which means the corresponding inter-run CVs increase. However, the inter-run CVs of the areas close to the line $n_{\mathrm{IFT}} = n_{\mathrm{UFT}}$ drop as the information bias rate increases. In other words, in the case of the number of informed traders and uninformed traders being equal, the ensemble variability accounts more significantly, while the internal variability affects the simulation more. For the cases other than $n_{\mathrm{IFT}} = n_{\mathrm{UFT}}$, both ensemble variance and internal variability increase.



(a) IBR $\sigma_{\varepsilon_n} = 0.05$ (b) IBR $\sigma_{\varepsilon_n} = 0.1$ (c) IBR $\sigma_{\varepsilon_n} = 0.25$

Figure 4.23 The Averaged Inter-run CV

Fig. 4.24 shows the variance ratio heatmaps combining the two measures mentioned above. We can find that the top and bottom left areas of the triangle heatmaps are both close to and smaller than 1. In contrast, the VRs in areas around $n_{\text{IFT}} = n_{\text{UFT}}$ are significantly below 1.



(a) IBR $\sigma_{\varepsilon_n} = 0.05$ (b) IBR $\sigma_{\varepsilon_n} = 0.1$ (c) IBR $\sigma_{\varepsilon_n} = 0.25$

Figure 4.24 The Variation Ratio

This also means that, when the market is approximately information-balanced, the dynamics of market price changes caused by each fake shock can be highly uncertain and unpredictable.

4.3 Conclusion

Based on the models of Glosten and Milgrom (1985) and Das (2005), we innovatively constructed an extended model in this chapter such that both informed and uninformed fundamental traders are in the market. The two types of traders have opposite trading strategies when facing a fake shock due to information asymmetry. Upon the hypothesis, spoofing can cause flash crashes or market fluctuations, resulting in altered equilibrium

prices, and the spoofer profits from other traders, this chapter constructed, implemented and analysed a basic artificial market with asymmetric information, and we have made some novel findings:

Due to information asymmetry in the market, a fake shock will lead to the uninformed and informed fundamental traders changing the estimates of the true stock prices. Therefore, the market price will fall to a new level of equilibrium which differs depending on market structures. We refer to the degree of the price decline as the maximum price changes under that particular parameter setting. We find that maximum price changes increase as the market is increasingly UFT-dominated. In low-risk markets, maximum price changes are more sensitive to changes in the market structure in balanced markets but less sensitive to highly unbalanced markets. The circumstance of high-risk markets is the opposite, as shown in Fig. 4.8.

There is a clear correlation between the profits of the fundamental traders and market structure. For informed fundamental traders, as Fig. 4.9 shows, they can earn more profits from a more UFT-dominated market. However, there is no monotonic relationship between the profits of the uninformed fundamental trader and the market structure balance. Fig. 4.10 shows that the average profits of a UFT may encounter significant declines in a market with a balanced configuration in the market structure. Conversely, the average profits tend to rise as the market structure configuration towards either UFT-dominated or IFT-dominated conditions.

In addition to the above findings, we found that when the market is more information-balanced, the variability of the market, shown in Fig. 4.24, becomes larger. Also, a highly balanced market structure comes with higher bid-ask spreads, which implies that the dynamics of prices in an information-balanced market can be highly uncertain and unpredictable under asymmetric information.

Chapter 5

Flash Crash Model with Intelligent Uninformed Traders

5.1 Introduction

Consider the circumstance extended from the basic flash crash model constructed in Chapter 4. There are three different kinds of trading agents in the trading crowd: informed fundamental traders, uninformed fundamental traders and zero-intelligence traders. They all observe the additional price signals arriving at the market, but the uninformed could doubt the authenticity of the stock price value v' after the "shock signal". As Chapter 4 discussed, the credulous uninformed traders will lose profits taken by the informed traders when the shock is fake. However, an uninformed fundamental trader can change their bidding strategy to reduce their losses if they can recognise this misjudgement of the market movements. We assume that uninformed traders can learn from public information after spoofing and recover the market from flash crashes. Therefore, in this chapter, upon the basic model in Chapter 4, we build an extended framework representing flash crashes and recovery of the market prices by introducing learning capabilities for uninformed fundamental traders and discuss the performance of the different types of agents in the market under asymmetric information.

5.2 Model Framework with Intelligent Uninformed Fundamental Traders

In this chapter, we extend the model constructed in Section 3.4 by endowing the uninformed fundamental traders with learning ability, and such IFTs are called **Intelligent Uninformed Fundamental Traders**. They are allowed to learn from past information flow to infer whether a shock deceives them. Therefore, the uninformed fundamental traders could adjust their trading strategies in case of recognising the fake shock in order to drive stock prices back to a reasonable level. As this chapter does not incorporate information sharing, uninformed fundamental traders rely solely on public information for their decision-making. In this context, we consistently refer to this as the "messy network"[1]. Apart from the new intelligent traders, all other agents are modelled with the same settings as in Section 3.4, and most marketplace assumptions remain the same. With the learning ability, the uninformed fundamental traders are modelled with the following extra assumptions:

5.2.1 Additional Assumptions

- The uninformed informed traders all know about the market structure but not the allocation of the other agents;
- When an uninformed fundamental trader observes an incoming shock, the UFT will initially trust the information but simultaneously adjust their beliefs regarding whether they are being deceived. The initial belief of each UFT regarding whether they are being deceived by the shock is set randomly;
- Each uninformed fundamental trader updates their beliefs on whether is deceived as soon as a new trade is made;

[1]Chapter 7 specifies the three different networks: messy network, cycle network and complete network based on the information-sharing topologies.

- Each uninformed fundamental trader will stop learning once they are supposed to recognise whether the information is real based on the learning process shown in Fig. 5.2.

5.2.2 Learning Strategy

Based on the assumption proposed in Chapter 4, the intelligent UFTs still lack inside information regarding the fake shock, so they could not instantly evaluate the stock price well if the new information is from "fake" news, rumour or some traders' malicious actions. However, once an intelligent UFT acquires the public information for each round, the trader can learn from the past information to determine if he has been deceived. The trading process for a single trading round is shown in Fig. 5.1.

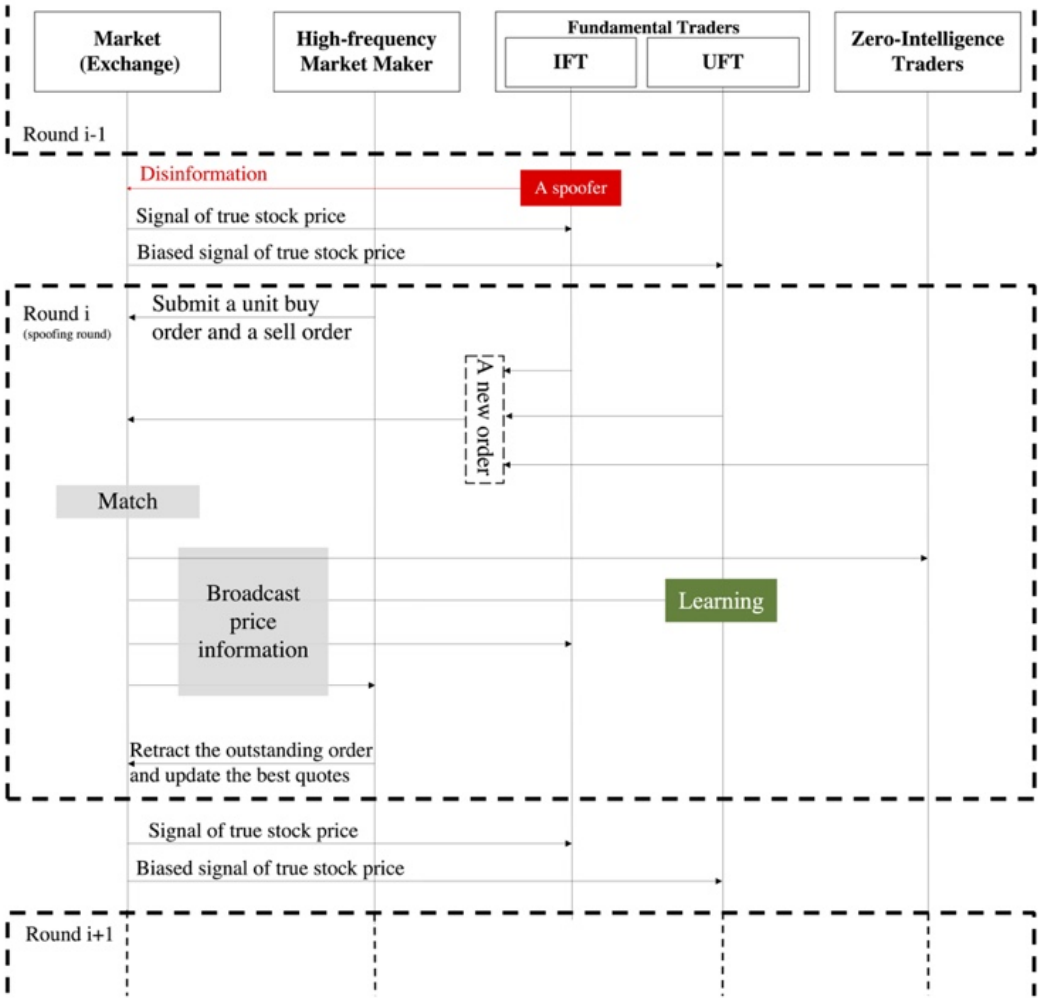


Figure 5.1 The Trading Process in a Spoofing Round

Therefore, when a new shock arrives in the system, a UFT receives a new signal of the fundamental prices, and there are two possible states of the system θ_0 and θ_1 that the intelligent UFTs are facing, where

1. θ_0: The shock is real, so the true stock price moves from v to v'. Accordingly, the price signal that a fundamental trader receives changes from w to w';

2. θ_1: The shock is fake, so the true stock price remains to be v;

Each uninformed fundamental trader has the beliefs about the state of the system $P_{x_i,t}(\Theta) = \left[P_{x_i,t}(\theta_0), P_{x_i,t}(\theta_1)\right]$ at the i-th trading round after the shock arrives. $P_{x_i,0}(\Theta)$

refers to the initial beliefs on different states as soon as the shock occurs. In this extended model, we allow the UFTs to update their beliefs Θ by using the Bayesian learning method based on orders entered into the market on whether they are deceived. The full details of learning can be shown in Appendix D.2, and it can be summarized as follows:

- Once a shock arrives, a UFT will change the fundamental price estimate with his limited information regardless of whether the shock is fake or not. In other words, the uninformed fundamental trader believes the shock is real initially.
- Before making the Bayesian inference, the UFT sets the initial prior probabilities of his beliefs $\left[P(\theta_0), P(\theta_1)\right]$ and the threshold c of stopping inferring. The stopping rule of the Bayesian inference is simply set as when

$$\min\{P(\theta_0), P(\theta_1)\} < c \tag{5.1}$$

- Given a new order arriving, a UFT will apply the Bayesian inference (detailed works shown in the Appendix D) to update his beliefs.
- Check whether the stopping rule is satisfied; otherwise, the UFT will repeat the last step.
- If $P(\theta_0) < c$, it means it is highly likely that the UFT makes a correct inference of adjusting the estimate. However, if $P(\theta_1) < c$, the UFT recognises that he is highly likely to be deceived, and he will correct his estimate of the true price to the level before observing the shock.

The following flow chart (Fig. 5.2) better illustrates the Bayesian inference process of the uninformed fundamental traders.

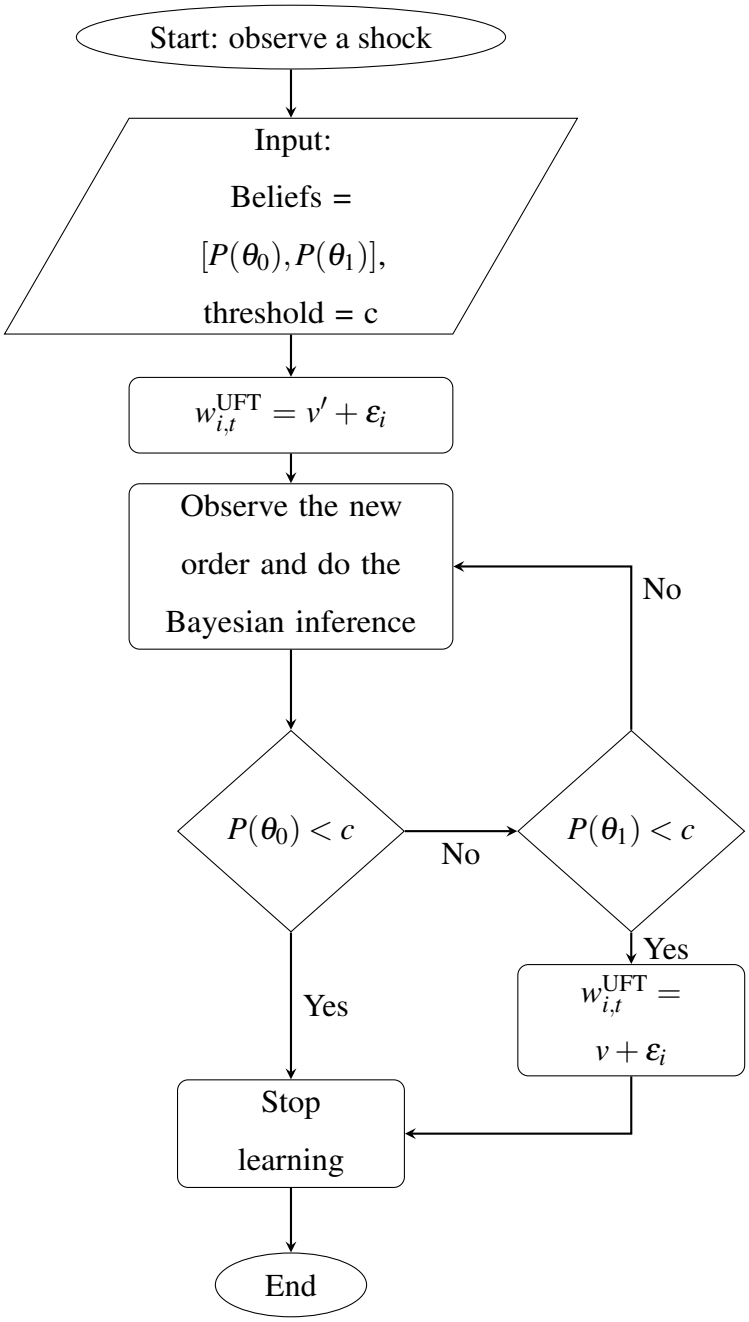


Figure 5.2 Flow Chart of Learning Process for the Uninformed Fundamental Traders

As Chapter 4 specifies, in each trading round, there are three possible actions $\mathscr{A}_1, \mathscr{A}_2, \mathscr{A}_3 \in \mathbb{A}$ for a fundamental trader. A fundamental trader can take the following actions at each trading round:

1. $\mathscr{A}_1$: Submit a buy order, if $w_i' > A_i$;

2. $\mathscr{A}_2$: Submit a sell order, if $w_i' < B_i$;

3. $\mathscr{A}_3$: Submit a null order – does not submit an order, if $B_i \leq w_i' \leq A_i$;

Denote the space of all possible allocations of the other agents that from the perspective of the i-th agent as $\mathbf{T}_i$, given the numbers of informed fundamental traders n_{IFT} and uninformed fundamental traders n_{UFT} (n_{IFT} and n_{UFT} are re-denoted as a and b, respectively, in mathematical formulas) are determined. Then, in a n-agent system, the number of allocations is $N = \binom{n-1}{b-1}\binom{n-b}{a}$, $\mathbf{T}_i = \bigcup_{k=1}^{N} \Omega_k$. Ω_k is a possible agent allocation describing that for each agent, which party, IFT, UFT or ZIT, the agent belongs to.

Based on Bayesian inference, this model endows the uninformed fundamental traders with learning capacity. Consider the situation of a 10-agent network, in which there are $n_{\text{IFT}} = a$ and informed fundamental traders and $n_{\text{UFT}} = b$ uninformed fundamental traders. We denote $x_i \in X =$ UFTs as the i-th trading agent, and his beliefs on the system states can be presented as $P_{x_i}(\Theta) = \left[(P_{x_i}(\theta_0), P_{x_i}(\theta_1)\right]$.

According to the symmetry, without loss of generality, we only consider the situation that $v' < v$ and assume that $w' < w$. If $\Theta = \theta_0$, all fundamental traders take the $\mathscr{A}_2$ until the market price moves to around v', which is the steady state. If $\Theta = \theta_1$, the informed fundamental traders initially take the $\mathscr{A}_1$ when the market price is down to v, while the uninformed fundamental traders take the $\mathscr{A}_2$.

$I_{i,t}$ stands for the information the i-th trading agents can receive at time t (the t-th trading round), then the learning strategy can be described as

$$\begin{aligned} P_{x_j,t}(\Theta = \theta_i | I_{i,t}) &= \frac{P_{x_i,t}(I_{i,t}|\Theta = \theta_j) P_{x_i,t}(\Theta = \theta_j)}{P_{x_i,t}(I_{i,t})} \\ &= \frac{P_{x_i,t}(I_{i,t}|\Theta = \theta_j) P_{x_i,t}(\Theta = \theta_j)}{\sum_{j=0,1} P_{x_i,t}(I_{i,t}|\Theta = \theta_j) P_{x_i,t}(\Theta = \theta_j)} \end{aligned} \tag{5.2}$$

The key to solving the learning Eq. 5.2 is calculating the conditional probability $P_{x_i,t}(I_{i,t}|\Theta = \theta_j)$ for each step, and further computations are specified in Appendices F and G . Referred to the settings in Zollman (2007), the network finishes learning when all uninformed fundamental traders believe that they are in θ_1 or θ_2 with the probability larger than 0.9999, i.e., $\max\{P_{x_i}(\theta_1), P_{x_i}(\theta_2)\} > 0.9999$, in which case $c = 10^{-5}$.

5.3 Experiments

This chapter still runs the short-term experiments where all trading activities are involved in a single trading epoch, and all basic settings remain the same as the previous chapter specifies. In a single trading epoch, there is only one shock occurs at a certain point. The shock could be either real or fake and influence the market differently. In this chapter, with the extension of intelligent uninformed fundamental traders, the model uncertainty is rising in the extended artificial market due to the randomness of the initial beliefs and order flows. Each UFT could spend different time periods making inferences and may also make incorrect inferences. In the simulation with such high uncertainty, we will study market dynamics from the perspective of the changes in flash crashes. Also, the cumulative profits of each agent are examined for the impact of information asymmetry on agents.

5.3.1 Basic Setups

Referring to the settings in Chapter 4, we assume that there are n_{total} traders excluding the market marker, and the number of ZI traders, IFTs and UFTs are n_{ZI}, n_{IFT}, n_{UFT}, respectively. A single trading epoch remains to have M trading rounds, where we add the external signal at the k-th round. The experiment has been mainly investigated in an artificial market affected by a "fake" shock, and we also examine the control group of the markets with a "real" shock as a support in some sections. For convenience, the trading epoch can be divided into two stages: the buffer stage, containing the range from round 1 to round k, and the active stage, containing the range from round k to round M.

5.3.2 Parameters

The settings of informational (σ_{ε_n}) and the market-structure parameters (n_{IFT} and n_{UFT}), remain unchanged as introduced in Section 4.1.2.

5.3.3 Performance Metrics

The important metrics proposed in Section 3.5, including crash size, crash duration and agents' profits, will also be investigated in this chapter. As this chapter introduced

intelligent uninformed fundamental traders, the market prices could recover after the flash crash. Hence, in this chapter, there are slight differences in the metrics to be examined compared to those presented in Chapter 4.

Crash Duration and Recovery Duration

Unlike the one set by the basic model, due to the introduction of intelligent uninformed fundamental traders in this chapter, the market price will bounce back as the UFTs finish their inferences and change strategies. Therefore, the whole market price movement can be divided into two phases – the crash and recovery.

The crash duration is redefined as the time period between the time step at which the shock occurs and the time step at which the market price reaches the lowest point. The recovery in this model is defined as the period during which market prices start to recover from their lowest point to a new steady state, so the recovery duration is defined as how long the recovery lasts.

In order to calculate the duration of the crash and recovery, we take the following approach to calculate two key points: the lowest price point and the new steady point. We first apply the MSER mentioned in Chapter 3 to calculate the MSER-5 truncation point as the new steady point, representing where the recovery ends. Next, we calculate the point standing for the lowest market price between the fake shock point and the new steady point, and use this division point to define the interval from the fake shock to the lowest price point as the crash phase and the interval from the lowest price point to the new steady point as the recovery phase, respectively. The crash and recovery are calculated in every single simulation.

Information Advantage

From Fig. 4.4, what we can find, in the simulation with a fake shock, the cumulative profit curves of the two different fundamental traders show bifurcations after the shock, which implies that the informed fundamental traders gain additional profits from the uninformed fundamental traders by the information advantage. The area between the two curves is calculated to be the cumulative profit difference between an informed fundamental

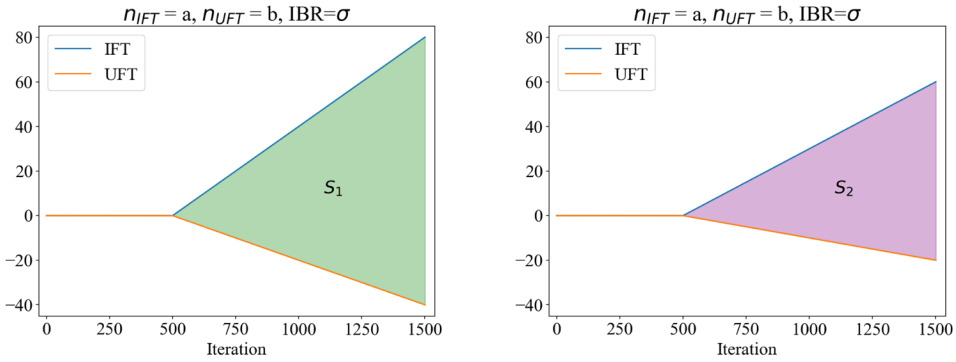
trader and an uninformed fundamental trader on average. To measure the information advantage, we now define an indicator as follows:

Under a specific set of parameters (including the market-structure parameters and informational parameter) and a specific extended model, the **information advantage (IA)** is defined to be the ratio of the average profit difference in the market with intelligent uninformed traders to that in the market with non-intelligent uninformed traders. Mathematically, the **informational advantage** can be written as

$$IA^{T_i}(n_{\mathrm{IFT}}, n_{\mathrm{UFT}}, \sigma_{\varepsilon_n}) = \frac{PD^{T_i}(n_{\mathrm{IFT}}, n_{\mathrm{UFT}}, \sigma_{\varepsilon_n}, T_i)}{PD^{Basic}(n_{\mathrm{IFT}}, n_{\mathrm{UFT}}, \sigma_{\varepsilon_n})}$$

.

Fig. 5.3 shows the graphic explanation of the informational advantage. The area in green of (a) shows the profit difference in the market with non-intelligent uninformed fundamental traders, referred to as S_1, and, in Fig. 5.3(b), the area in purple shows the cumulative profit difference in the market with intelligent uninformed fundamental traders under a certain model extension, referred to as S_2. Then the informational advantage can be defined as the ratio S_2/S_1.



(a) Average Cumulative Profits in a Basic Model (b) Average Cumulative Profits in a Learning Model

Figure 5.3 Graphic Explanation of Informational Advantage Calculation

In a market with only non-intelligent UFTs, who do not adapt their actions through learning, all non-intelligent UFTs are deceived by the fake shock, leading to the adoption of bad order placement strategies. Theoretically, this scenario results in the maximum average

losses for a UFT, while an IFT can gain the maximum average profits. We use this "worst-case" scenario, the market with non-intelligent UFTs, as the benchmark. However, when UFTs can adjust their behaviour through learning, they can change their order placement directions to prevent further losses. In Fig. 5.3, the colour-shaded areas represent the average cumulative profit differences between an IFT and a UFT under the information asymmetry caused by spoofing. Therefore, by calculating S_2/S_1, we can determine, in the context of introducing intelligent UFTs, the advantage score of an IFT over a UFT relative to the worst-case scenario (S_1). This value theoretically falls between 0 and 1, with the two ends respectively representing no information advantage and complete information advantage. If S_2/S_1 approaches 1, it indicates a larger information advantage of IFTs over UFTs; conversely, if S_2/S_1 approaches 0, the advantage of IFTs over UFTs tends to diminish.

Inference Accuracy

In this model, all traders are independent of each other, so the learning process could be also different for each intelligent UFT.

As the uninformed fundamental traders are enabled to have the learning ability to infer whether a shock is fake or not, an uninformed fundamental trader can either obtain correct or incorrect inference. We have kinds of modelled markets where a real shock or a fake shock separately occurs in the two markets. Therefore, the correct and incorrect inference of an uninformed fundamental trader of a simulation can be defined as follows:

In a fake-shock market:

- Correct inference: if the UFT infers that the shock is fake before the simulation ends;
- Incorrect inference: if the UFT infers that the shock is real or fails to complete the inference before the simulation ends;

In a real-shock market:

- Correct inference: if the UFT infers that the shock is real or fails to complete the inference before the simulation ends;

- Incorrect inference: if the UFT infers that the shock is fake before the simulation ends.

Therefore, in a single simulation, there could be several different UFTs who come to different conclusions after making a series of inferences. For example, trader A finds that he has been deceived, but trader B believes that the shock in the market is real. This section will calculate the average accuracy of traders' inference accuracy in different systems.

Let $D = (X_i, \Theta_i)_{i=1}^{N_D}$ be a series of data for learning, where N_D is the data sample size. X_i, Θ_i are commonly called input and output, respectively. Suppose that Θ_i is a binary classification variable, that is $\Theta_i \in 0, 1$.

Therefore, the goal of learning is to apply the data D to learn the classification model $f : X \to \Theta$. Once a new input is applied in the classifier, there comes a predicted value $\hat{\theta} = f(x)$. Denote the bayesian classifier as I is $I = \hat{f}$, and $I : x \to [0, 1]$. Based on this setting, the bayesian classifier is a model that generates a belief score on how much the output is 1. A threshold γ is defined as a confidence score, which means that the predicted output is 1 if the belief score $I(x) > \gamma$.

In order to measure the quality of the prediction, we define the following loss function $\Delta : \Theta \times \Theta \to TP, TN, FP, FN$. Let $\theta \in 0, 1$ be the true values and $\hat{\theta} \in 0, 1$ be the predicted values, then the mapping can be defined as follows:

- if $\theta = 1$ and $\hat{\theta} = 1$, then $\Delta(\theta, \hat{\theta})$ = TP(True Positive)
- if $\theta = 0$ and $\hat{\theta} = 0$, then $\Delta(\theta, \hat{\theta})$ = TN(True Negative)
- if $\theta = 0$ and $\hat{\theta} = 1$, then $\Delta(\theta, \hat{\theta})$ = FP(False Positive)
- if $\theta = 1$ and $\hat{\theta} = 1$, then $\Delta(\theta, \hat{\theta})$ = FN(False Negative)

In this model, the null hypothesis ($\theta = 0$) of the bayesian inference is that the shock is real/ the uninformed fundamental traders are not deceived, then the alternative hypothesis ($\theta = 1$) is that the shock is fake/ the uninformed fundamental traders are deceived.

Suppose that we run N times in the Monte-Carlo simulations, with the simplification of the notation $\Delta_i = \Delta(\theta_i, \hat{\theta_i})$, we can include all loss function values in a vector $M = (\Delta_1, ..., \Delta_N)$ and calculate the numbers of TP, TN, FP and FN, respectively.

The numbers of TP, TN, FP and FN can be summarized in a 2×2 matrix as follows, which is called the confusion matrix.

		Predicted	
		Yes	No
Actual	Yes	TP	FN
	No	FP	TN

This thesis mainly examines the market under asymmetric information, in which case a fake shock occurs, so the main performance indicators are true negative rate and accuracy to show how well an uninformed fundamental trader can defend the fake shock.

$$\text{True Negative Rate} = \frac{\#TN}{\#TN+\#FP}$$

$$\text{Accuracy} = \frac{\#TP+\#TN}{\#TP+\#FP+\#FN+\#TN}$$

Other common performance indicators are

$$\text{Precision} = \frac{\#TP}{\#TP+\#FP}$$

$$\text{Recall} = \frac{\#TP}{\#TP+\#FN}$$

$$\text{F1 Scores} = 2 \cdot \frac{\text{precision} \cdot \text{recall}}{\text{precision}+\text{recall}}$$

These indicators can be used to compare market performance in different parameters to support this chapter's studies. The additional experimental results are shown in Appendix H.

5.4 Experimental Results

5.4.1 Global Statistics Overview

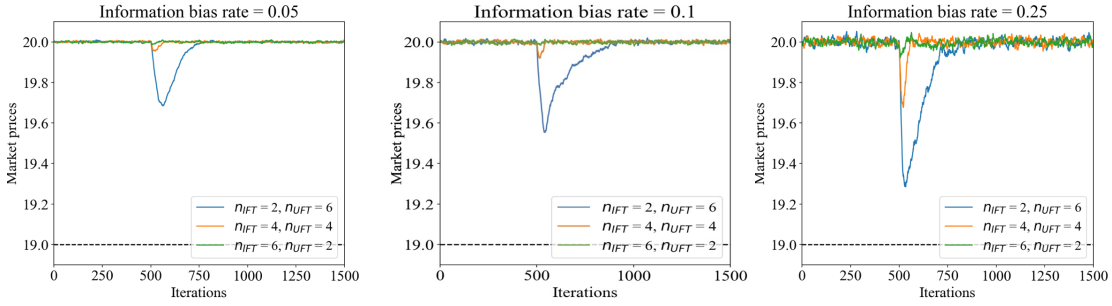
At the beginning of our analysis of the experimental results, we focus on some global statistics on price movements and agents' profits as an overview. Similar to the setups in Chapter 4, we consider three distinct market structures: an IFT-dominated market, a UFT-dominated market, and an information-balanced market. These market structures maintain identical configurations to the 10-agent system described in Section 4.2.1.

Prices Movements

If the shock is fake Fig. 5.4 shows the average market prices among all simulations under the given three parameter settings with the comparison among different information bias rates in the **messy network**. Compared with the corresponding fake-shock simulations in Fig. 4.2, the average prices all experience a bounce after the shock point in the simulations with intelligent uninformed fundamental traders. However, the patterns differ among different market structure settings. The cases of $\{n_{\text{IFT}} = 6, n_{\text{UFT}} = 2\}$, $\{n_{\text{IFT}} = 4, n_{\text{UFT}} = 4\}$ and $\{n_{\text{IFT}} = 2, n_{\text{UFT}} = 6\}$ respectively represent three kinds of markets: IFT-dominated market, balanced market and UFT-dominated market, which are shown in green, orange and blue lines in Fig. 5.4.

In all three cases under different information bias rate settings, the UFT-dominated market experiences a more significant and longer-lasting crash, and the market with more informed fundamental traders experiences a weaker crash under the fake-shock system. However, the reason for such a small crash size is that, even though the uninformed fundamental traders are not intelligent (in the basic model from Chapter 4), the market prices still drop not too much when the informed fundamental traders dominate. More analysis is further discussed in the following sections.

In addition, we can see that the market's average price has fully rebounded and converged to the level before the fake shock, implying that, in CDA markets, the intelligent uninformed fundamental traders have likely made correct inferences after the fake shock.



(a) IBR $\sigma_m = 0.05$ (b) IBR $\sigma_m = 0.1$ (c) IBR $\sigma_m = 0.25$

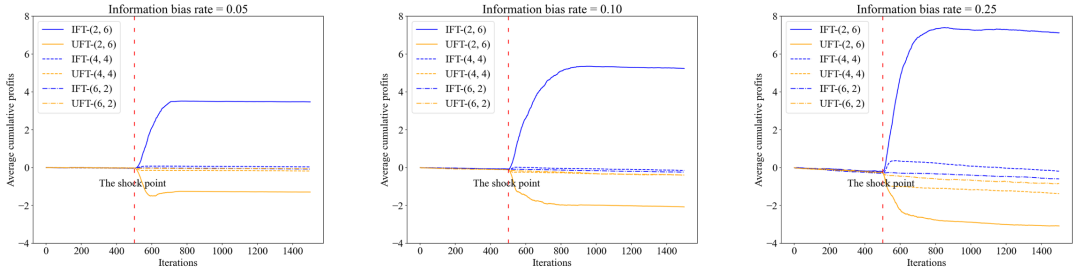
Figure 5.4 The Average Market Prices Movements in the Continuous Auction Market

Profits of Agents

The average cumulative profits graphs show how many profits different kinds of agents can accumulate on average as the iteration goes. The slope refers to the average profits a trader obtains at a certain iteration point.

The shock is set to occur at the 500th iteration. Before the shock, we notice that the average cumulative profits curves of the fundamental traders have slightly negative slopes meaning that they hardly make profits during the equilibrium of the no-shock periods. After the shock point, the curves of informed and uninformed fundamental traders' profits begin to diverge to some degree, which means that the informed fundamental traders begin to take advantage of the information to grab profits from the uninformed fundamental traders.

Fig. 5.5 illustrates the average cumulative profits of the fundamental traders during a trading epoch of 1500 trading rounds.



(a) n_{IFT}=2, n_{UFT}=6, IBR=0.05 (b) n_{IFT}=2, n_{UFT}=6, IBR=0.1 (c) n_{IFT}=2, n_{UFT}=6, IBR=0.25

Figure 5.5 Fundamental Traders' Average Cumulative Profits

Similar to the previous chapter, the cumulative profits of IFT and UFT diverge after the shock point. The graph below shows the difference in divergence among different information bias rates and market structures. A clear finding is that when the market has a higher information bias rate, which implies higher market risk, the cumulative profit difference between IFT and UFT becomes larger.

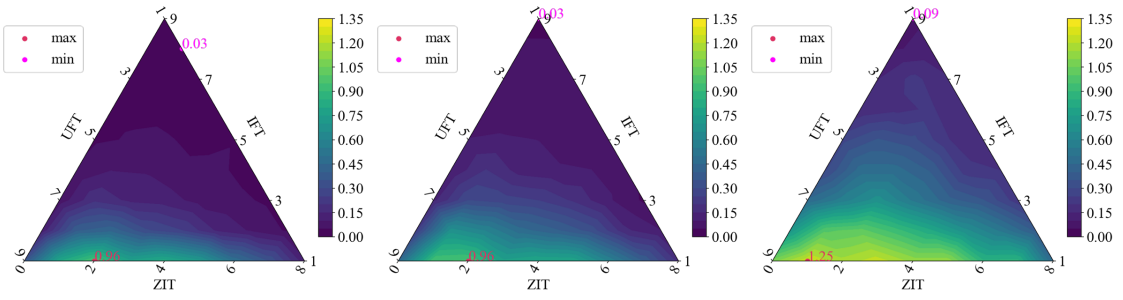
Another preliminary finding is that there could be a much wider divergence between the profits of the informed fundamental traders and the uninformed fundamental traders if the market is more UFT-dominated.

5.4.2 Sensitivity Analysis with Fake Shocks

Crash Size

This section discusses the differences in crash size in different networks. Based on the setting of this thesis, uninformed traders and informed traders may act in opposite order-placing behaviour due to information asymmetries, thus, causing the market price to deviate from its true value. In this chapter, the uninformed fundamental traders are given the ability to learn, which is hence called intelligent UFTs, to infer whether they have been deceived. As soon as intelligent UFTs have made a correct inference after the shock, they can change their order-placing behaviour to bring the market price back to its true level. A v-pattern will appear in the market price trend to indicate a flash crash. The crash size measures the degree of market price deviation and also reflects the ability of intelligent UFTs to detect false information.

Fig. 5.6 illustrates the crash size in the CDA market under various parameter configurations. The colour bars suggest that the colours approaching yellow refer to a large flash crash, while the colours approaching navy refer to small crash sizes. According to Fig. 5.6(a), (b) and (c), all light yellow colours are clustered around the points where there is a large portion of UFTs with a small portion of IFTs and ZITs, referring a strongly UFT-dominated market. The simulation results show that, regardless of the information bias rate, the average crash size reaches the maximum point in the case of $n_{\text{IFT}} = 1, n_{\text{UFT}} = 7, n_{\text{ZIT}} = 2$.



(a) IBR $\sigma_m = 0.05$ (b) IBR $\sigma_m = 0.1$ (c) IBR $\sigma_m = 0.25$

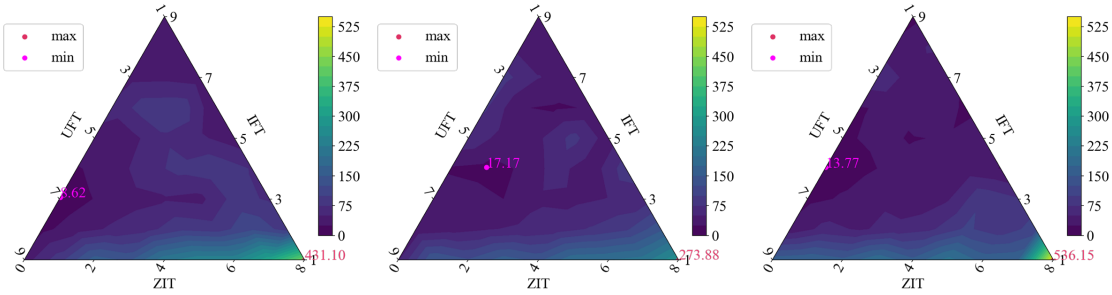
Figure 5.6 Mean of Crash Sizes

Crash and Recovery

In addition to studying the crash size to analyze the market dynamics after the fake shock, we will also examine the crash duration and recovery duration during the price variation phase.

We have studied the zones of crash size under different parameter settings, shown as triangular heat maps. Unlike the crash size pattern, the average crash size increases asymptotically as the market becomes more UFT-dominated, the crash and recovery durations are only significantly long when the number of IFTs is 1, while the crash and recovery durations are significantly lower as the market structure is otherwise.

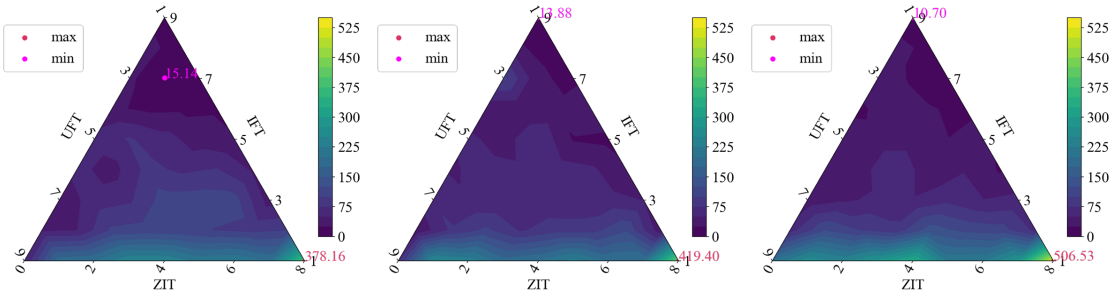
We have studied our basic model in markets without intelligent UFTs showing that the crash durations of the market prices decline due to a fake shock decreasing significantly with the increase in systemic risk. In other words, the faster the price changes when the market risk is higher. This finding is also observed in the model with intelligent UFTs.



(a) IBR $\sigma_m = 0.05$ (b) IBR $\sigma_m = 0.1$ (c) IBR $\sigma_m = 0.25$

Figure 5.7 Mean of Crash Durations

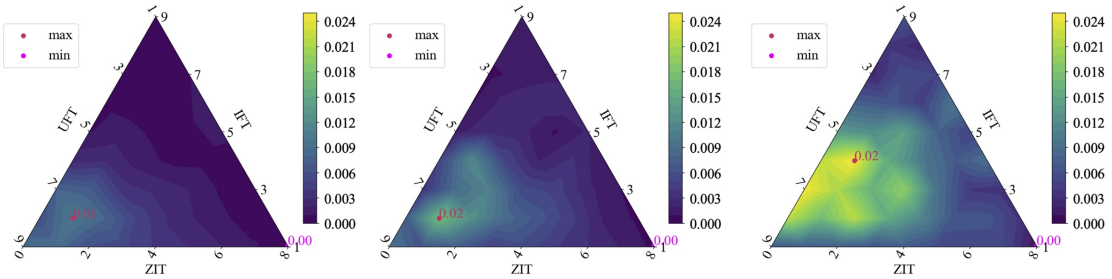
However, it is worth noting that while crash duration and recovery duration have similar patterns under different market structure settings, they vary in terms of systemic risk factors. If we focus on the case of strongly UFT-dominated markets, we find that as systemic risk increases, the crash duration gradually decreases while the recovery duration gradually increases.



(a) IBR $\sigma_m = 0.05$ (b) IBR $\sigma_m = 0.1$ (c) IBR $\sigma_m = 0.25$

Figure 5.8 Mean of Recovery Durations

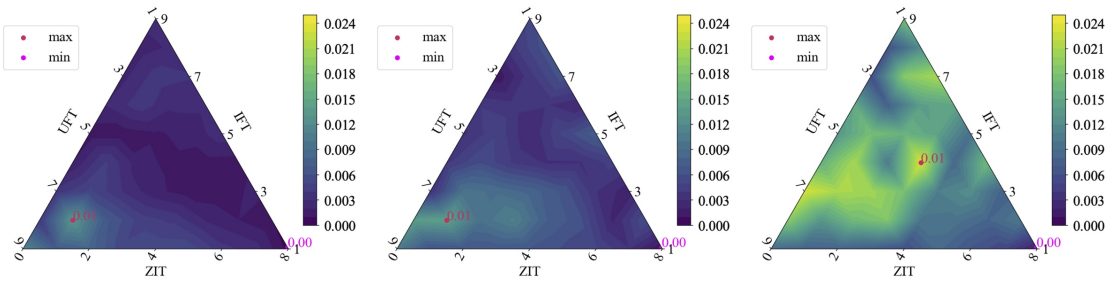
We then obtain the crash speed and the recovery speed by dividing the crash size by the crash duration and the recovery duration, respectively. By comparing the contour plots in Fig. 5.9, we find that crash speed increases with the market volatility increasing, and the market prices in the UFT-dominated fundamental markets drop faster.



(a) IBR $\sigma_m = 0.05$ (b) IBR $\sigma_m = 0.1$ (c) IBR $\sigma_m = 0.25$

Figure 5.9 Average Crash Speed

The recovery speed is slower than the crash speed because it takes time for the uninformed fundamental traders to change their trading strategies, and each UFT needs different time periods to finish inference.

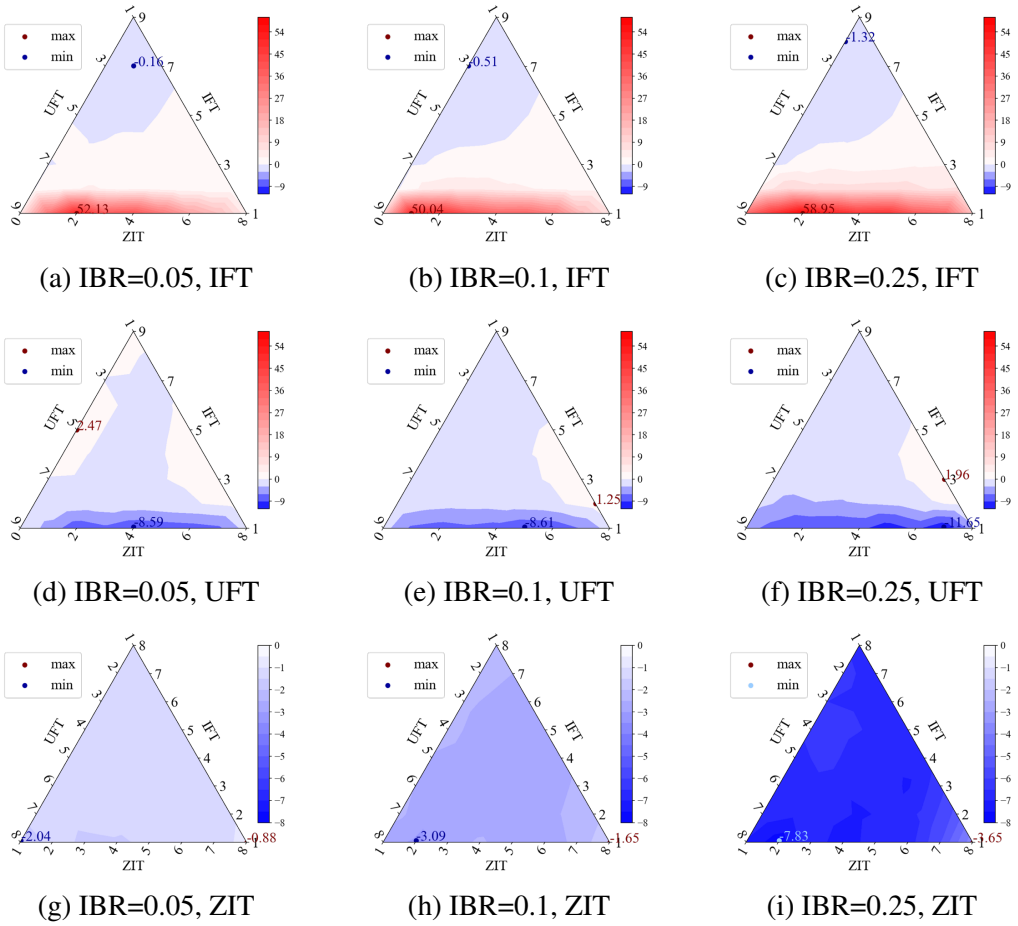


(a) IBR $\sigma_m = 0.05$ (b) IBR $\sigma_m = 0.1$ (c) IBR $\sigma_m = 0.25$

Figure 5.10 Average Recovery Speed

Agents' Average Profits

From Fig. 5.11(a), (b) and (c), we can clearly see that, if the market is strongly UFT-dominated, especially when the number of IFTs is 1, each IFT can obtain significant profits on average. The average profit per IFT reaches the highest when the shares of agents in the market are $n_{\text{IFT}} : n_{\text{UFT}} : n_{\text{ZIT}} = 1 : 7 : 2$ or $n_{\text{IFT}} : n_{\text{UFT}} : n_{\text{ZIT}} = 1 : 8 : 1$. However, as the number of IFTs in the market increases and the number of UFTs decreases, the average profit per IFT decreases rapidly.



(a) IBR=0.05, IFT (b) IBR=0.1, IFT (c) IBR=0.25, IFT

(d) IBR=0.05, UFT (e) IBR=0.1, UFT (f) IBR=0.25, UFT

(g) IBR=0.05, ZIT (h) IBR=0.1, ZIT (i) IBR=0.25, ZIT

Figure 5.11 Agents' Average Profits

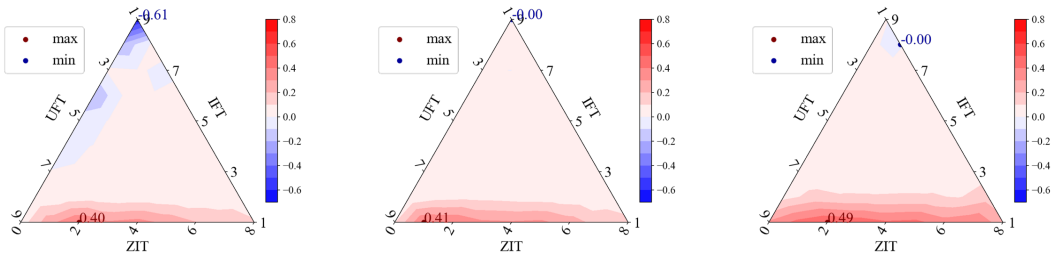
However, when the market structure is, roughly speaking, in the range of $n_{\text{ZIT}} < 0.5 * n_{\text{IFT}}$ referring to the figures, IFTs only achieve negative profits on average. While it is very inaccurate to approximate such a range just based on the contours, it does show that when the number of random traders in the market is small, and the number of IFTs is relatively high compared to the number of random traders, there are few possibilities for IFTs to profit from fake shock by spoofing.

The profits of uninformed fundamental traders are poor compared to those of informed fundamental traders. As UFTs are on the spoofed side, the contour figures indicate that the average profits per UFT are negative in most cases. This is especially apparent in a strongly UFT-dominated market, where each UFT takes a huge loss and, comparatively, each IFT has a high average profit. In other words, in this case, a high amount of profits of each UFT is taken by the IFT.

Compared to the fundamental traders, the earnings for zero-intelligence traders are below par, as expected. Due to the setting of entirely random quoting, ZITs have the weakest information advantage in the market. As a result, the profit distributions of ZITs are negative in all cases. However, from a market structure perspective, the points at which ZITs achieve the lowest average negative profits are around $n_{\text{IFT}} : n_{\text{UFT}} : n_{\text{ZIT}} = 1 : 1 : 8$. This conclusion is also reasonable because, in this case, the market has the largest number of ZITs and the smallest number of fundamental traders, so the loss of each IFT caused by the lack of information advantage could be the least.

Information Advantage

Fig. 5.12 represents the differences in information advantage between an IFT and a UFT when the market structures and market volatility differ. We find that, in strongly UFT-dominated markets, each IFT has a high level of information advantage on average over each UFT. Each IFT gains an information advantage of approximately 10%-50% when the number of IFTs in the market is only 1, with the largest information advantage in the case of $n_{\text{UFT}} = 7$ or 8. When the number of IFTs in the market is greater than 1, the information advantage for IFTs is, in most cases, less than 10%. Also, the information advantage of IFT is greater as the market is more volatile. It is worth noting that when the market information bias rate is 0.05, IFTs lose their information advantage (negative values) in strongly IFT-dominated markets, i.e. an IFT gains fewer profits than a UFT on average in this case.



(a) IBR $\sigma_m = 0.05$ (b) IBR $\sigma_m = 0.1$ (c) IBR $\sigma_m = 0.25$

Figure 5.12 Information Advantage

5.4.3 Inference Measurement

With the addition of learning abilities to the uninformed fundamental traders, a UFT can adjust his behaviour by digging the public information to determine if he is being deceived. All intelligent uninformed fundamental traders are modelled to adopt the same learning strategy with the same rules based on Bayesian inference. To measure the inference power of the uninformed fundamental traders under different parameter settings, we introduce the confusion matrix and calculate the precision, recall, true negative rate and accuracy defined in Section 5.3.3. According to the definition,

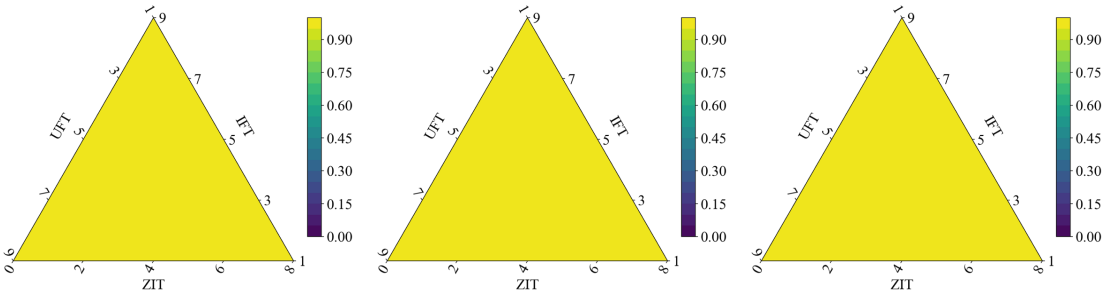
- TP: if an uninformed fundamental trader infers that the shock is real in the simulation where a real shock hits;
- FP: if an uninformed fundamental trader infers that the shock is real in the simulation where a fake shock hits;
- TN: if an uninformed fundamental trader infers that the shock is fake in the simulation where a fake shock hits;
- FN: if an uninformed fundamental trader infers that the shock is fake in the simulation where a fake shock hits

When we focus on the flash crash situation in the asymmetric informational market, the true negative rate is the most important indicator in this model. The true negative rate is defined to be $\#TN/\#TN+\#FP$ describing that the correct rate under the circumstance of the market with a fake shock. A true negative inference means that an uninformed fundamental trader successfully avoids being deceived by the fake shock. However, only considering the performance of the inference is incomplete, and we also need to integrate the inference duration and the crash duration to investigate the performance of UFTs in the case of fake shock.

True Negative Rate

The true negative rate measures the average probability of each uninformed fundamental trader successfully inferring that a shock is fake from past order flows under the

circumstance that the market shock is fake. A high true negative rate indicates that the majority of UFTs realise they have been deceived, so that they change their estimate of the true stock price and quoting strategy to bring the market price back to the true value level; a low true negative rate indicates that some UFTs do not believe they are being deceived and maintain their previous quoting strategy but take greater losses.



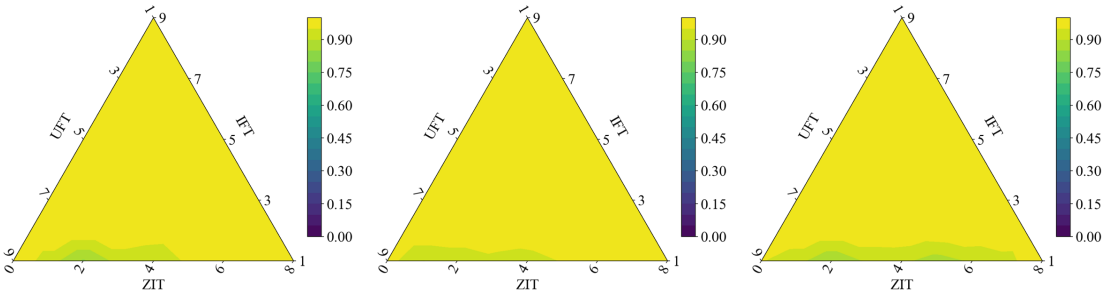
(a) IBR $\sigma_m = 0.05$ (b) IBR $\sigma_m = 0.1$ (c) IBR $\sigma_m = 0.25$

Figure 5.13 The Inference True Negative Rate of the Uninformed Fundamental Traders (Messy Network)

According to Fig. 5.13, we find that the true negative value reaches almost 100% in all cases, regardless of the overall market risk or the structure of the agents in the market. That is, although the crash size and duration were relatively large, UFTs were able to detect that they were being deceived by fake shock with almost 100% rate.

Accuracy

When we look at the accuracy of inferences, we find that the accuracy of each UFT's inference decreases significantly when the market is very strongly UFT-dominated, due to the fact that UFTs could make incorrect inferences when the market is affected by a real shock.



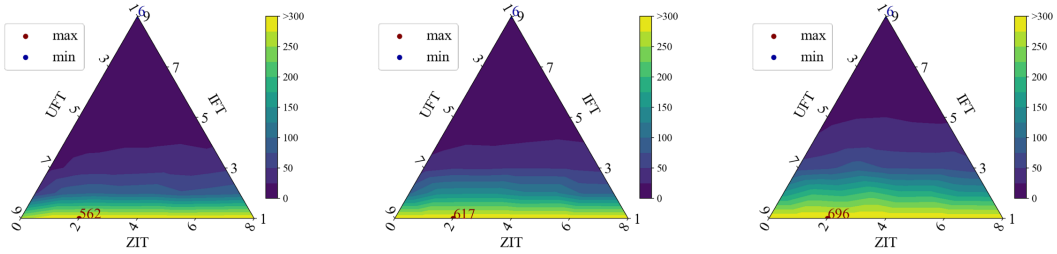
(a) IBR $\sigma_m = 0.05$ (b) IBR $\sigma_m = 0.1$ (c) IBR $\sigma_m = 0.25$

Figure 5.14 The Inference Accuracy of the Uninformed Fundamental Traders (Messy Network)

However, a real shock does not create the information asymmetry between the UFTs and IFTs. Therefore, with the learning ability set in this chapter, each UFT may overlearn from past orders, leading to a bad order-placing strategy in a non-asymmetric-information market with a strongly UFT-dominated structure.

Inference Duration

Fig .5.15 shows the average duration of each uninformed fundamental trader finishing the inference. We found that, under the circumstance of a fake shock hitting, while almost all UFTs are still able to accurately infer that they have been spoofed, the duration of completing the inference for each UFT is significantly high in a strongly UFT-dominated market, in which case a UFT suffers a big loss on average.



(a) IBR $\sigma_m = 0.05$ (b) IBR $\sigma_m = 0.1$ (c) IBR $\sigma_m = 0.25$

Figure 5.15 The Average Inference Duration of an Uninformed Fundamental Trader

5.4.4 Statistical Testing

In this chapter, we aim to understand how the introduction of intelligent uninformed fundamental traders in our model affects market prices during crashes. Additionally, we seek to explore the differences in profits between intelligent UFTs and non-intelligent UFTs in Chapter 4. To address these questions, we utilise a combination of t-tests and Cohen's d for statistical testing, as explained in Section 3.6, which also serves to address the primary hypothesis of this chapter. In statistical testing here, we set the dataset of the crash sizes from the basic model as Group 1, and the dataset from the model with intelligent UFTs specified in this chapter as Group 2.

Cohen's d is calculated as a complement to see how significantly the two datasets are different. A positive Cohen's d means that Group 1 has a higher mean than Group 2, implying the crash sizes are reduced by introducing intelligent UFTs. Conversely, a negative Cohen's d means that Group 2 has a higher mean. The effect size is referred to as mentioned in Section 3.6.2:

Value	[0, 0.2)	[0.2, 0.5)	[0.5, 1.0)	≥ 1.0
Effect	Small	Medium	Large	Very large

Table 5.1 Cohen's d Effect Size Reference Table

Crash Sizes

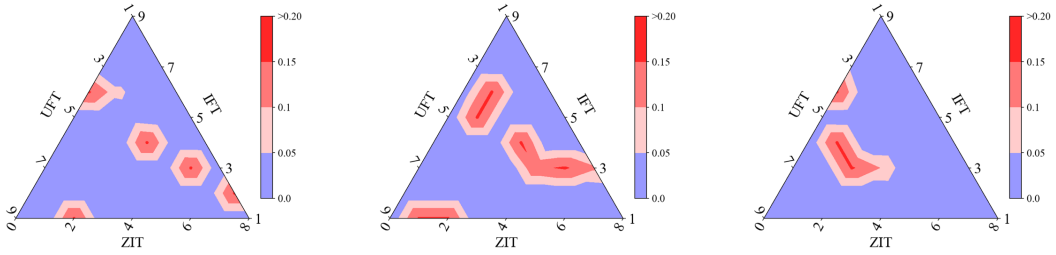
We have the following null hypothesis:

Null Hypothesis H_0**:** *The crash sizes in the simulations where intelligent UFTs are introduced are not significantly different from the crash sizes in the basic model with non-intelligent UFTs.*

Therefore, if we obtain a small p-value (less than 0.05), then the null hypothesis is rejected, implying that the crash sizes are significantly affected by the UFTs who are able to learn from order flows. However, a large p-value (higher than 0.05) does not show the significance.

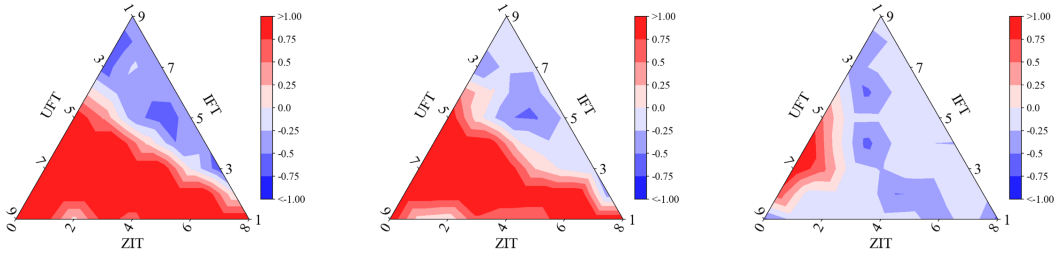
Fig. 5.16 shows p-values in contour plots under different parameter settings. According to the colour bar, the blue colour stands for the p-values less than 0.05, and the contour

plots show that crash sizes in the market with the introduction of intelligent UFTs are significantly different from those derived in the basic model.



(a) IBR $\sigma_m = 0.05$ (b) IBR $\sigma_m = 0.1$ (c) IBR $\sigma_m = 0.25$

Figure 5.16 p-values in Crash Sizes



(a) IBR $\sigma_m = 0.05$ (b) IBR $\sigma_m = 0.1$ (c) IBR $\sigma_m = 0.25$

Figure 5.17 Cohen's d in Crash Sizes

Positive Cohen's d values mean crash reduction, while negative values stand for crash amplification. Through calculating Cohen's d of the dataset obtained from the two different models, we can clearly see that, when the market volatility is low (IBR=0.05 or 0.1), the crash size reduction is significant in the UFT-dominated markets, while there is no clear evidence to show that the flash crashes are mitigated in the IFT-dominated markets, but it shows that crash sizes may be slightly amplified. However, the crash sizes are already small as the market is IFT-dominated, so it is reasonable to observe such results.

When the market volatility is relatively high, Cohen's d values are almost small everywhere in the triangular plot except the configurations around $5 \leq n_{\text{UFT}} \leq 8, 0 \leq n_{\text{UFT}} \leq 2$. This means that the crash sizes could not be reduced significantly.

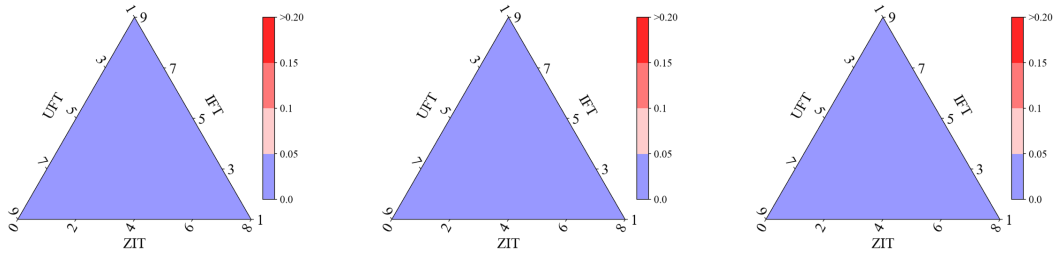
UFT's Profits

We have conducted several investigations, revealing that the uninformed fundamental traders can adapt their order-placing strategies by learning from the order flows, effectively mitigating their losses as the market prices recover. This led us to propose an additional hypothesis for statistical testing, focusing on whether the UFTs' profits can be reduced through their learning. The null hypothesis under examination is as follows:

Null Hypothesis H_0**:** *Intelligent UFTs are unable to reduce their losses by learning from order flows and adjusting their strategies.*

We then proceeded to calculate p-values and Cohen's d values for the datasets associated with **UFTs' cumulative profits** in the models presented in Chapters 4 and 5.

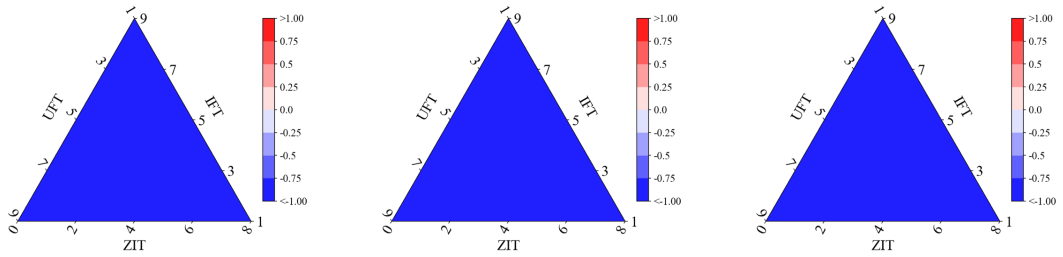
Then, similarly, we calculated the p-values and Cohen's d values of the dataset regarding UFTs' profits from the two models proposed in Chapter 4 and Chapter 5. The results are visually represented in Fig. 5.18 and Fig. 5.19, depicted as triangular contour plots.



(a) IBR $\sigma_m = 0.05$ (b) IBR $\sigma_m = 0.1$ (c) IBR $\sigma_m = 0.25$

Figure 5.18 p-values in UFT's Profits

The p-value and Cohen's d plots provide clear evidence that enabling UFTs with learning capabilities significantly reduces their losses. p-values fall below 0.05 for every market-structure configuration, while Cohen's d values indicate very large negative effect sizes (Cohen's d < -1) in all cases, representing that the profits data in Group 2 (model with intelligent UFTs) has a higher mean, which also confirms the substantial reduction in UFTs' losses. These findings align with the observations in Fig. 4.10 and Fig. 5.11, which clearly illustrate the notable changes in UFTs' losses.



(a) IBR $\sigma_m = 0.05$ (b) IBR $\sigma_m = 0.1$ (c) IBR $\sigma_m = 0.25$

Figure 5.19 Cohen's d in UFT's Profits

5.5 Conclusion

Chapter 4 built a basic framework of an artificial market with asymmetric information to study markets with flash crashes under asymmetric information. The uninformed fundamental traders in the basic model are set to remain with their order-placing strategies after being deceived. However, in reality, traders always adjust their trading strategies instead of being completely passive, especially in the case of uninformed traders. They could keep updating the estimates and understanding of the market in response to the past dynamics. Hence, this leads to the hypothesis that *traders can learn from public information after spoofing and, consequently, facilitate the recovery of the market from flash crashes*. In this chapter, we built an extended framework based on the basic model in Chapter 4, and this framework allowed uninformed fundamental traders to learn from the historical order flows to infer whether the market shock is real or fake.

Based on the new framework, we implemented the computational model and examined the impact of market structure and volatility/risk on the market performance with a fake shock. We finally obtained the following findings:

- Fig. 5.13 shows that the intelligent UFTs were able to detect that they were being deceived by fake shock with almost 100% rate, and, compared to the scenario that UFTs do not learn, the intelligent UFTs can reduce their losses during flash crashes, represented in Fig. 5.18 and Fig. 5.19. However, it takes different time steps for UFTs to learn and infer under different market structures, so they are more likely to suffer big losses in strongly UFT-dominated markets.

In addition to the investigation of the hypothesis, we have also obtained other findings through simulation and analysis.

- While the intelligent UFTs can effectively recognise the fake shock in Fig. 5.14, we still observe flash crashes under different parameter settings, but the prices recover shortly after the decline. The market with high volatility suffers a larger flash crash.
- When the market is strongly UFT-dominated, the flash crash durations are more affected than other cases. When $n_{\text{IFT}} \leq 2$, it takes much longer for the market prices to recover, as shown in Fig. 5.15, which implies that the market is riskier in such cases under the information asymmetry.
- Fig. 5.6 shows that the crash sizes are sensitive to the market structure. More specifically, the crash sizes are more significant as the market is increasingly UFT-dominated. The market prices could suffer the biggest price decline in the case of $n_{\text{IFT}} = 1$ and $n_{\text{UFT}} = 8$ or 9.
- IFTs can still grab high profits from UFTs in strongly UFT-dominated markets (shown in Fig. 5.11), where the information advantage of each IFT is also at a high value (shown in Fig. 5.12).

In summary, this chapter introduces an innovative model extended based on Chapter 4 by enabling uninformed fundamental traders with the capacity to learn. Intelligent UFTs use Bayesian learning to adjust their order-placing strategies in response to a fake shock. The results demonstrate that these intelligent UFTs can effectively detect spoofing, and adjust their strategies to stop losses, ultimately leading to market price recovery. Because of the varying inference durations, the flash crashes still occur under different market structure configurations. We find that when the share of informed fundamental traders in the market is roughly less than 20%, the market as a whole is highly vulnerable under information asymmetry in the aspects of market price crashes and profit allocation. Also, the decision-making of uninformed traders themselves is also negatively affected.

Based on the framework of this chapter, we will introduce two extensions regarding the batch auction mechanism; and the topology of agent connections of information sharing in

the next two chapters. We will examine whether some of these changes can mitigate flash crashes under the impact of fake shocks in the market.

Chapter 6

Flash Crash under Batch Auctions

6.1 Introduction

Chapters 4 and 5 explored the market performance of the flash crash model in a continuous double auction market with diverse types of agents. We found that in an asymmetric informational market, fake shocks bring adverse selection among traders, and the information-advantaged agents (informed fundamental traders) benefit from the information-disadvantaged agents (uninformed fundamental traders). However, in Chapter 5, we found that the information asymmetry is weakened by intelligent UFTs, and therefore the market price can return to the true price level. However, the risk of adverse selection still lasts a long time when the market is strongly UFT-dominated, causing a major loss for each information-disadvantaged agent (uninformed fundamental trader).

Several studies have found that batch auctions can serve as an alternative to continuous double auctions to mitigate the impact of adverse selection in the CDA market (Aquilina et al., 2020; Budish et al., 2014; Foucault et al., 2017). Therefore, within the context of flash crashes, this chapter explores the hypothesis that batch auctions can serve as an alternative to continuous double auctions to reduce the impact of adverse selection impact, particularly in situations like flash crashes caused by spoofing, within the CDA market.

In Chapter 3, we introduced the differences between the mechanisms of continuous double auction and batch auction, and in this chapter, we expand upon the flash crash model presented in Chapter 5 by replacing the CDA market mechanism with batch auctions.

This chapter will investigate the differences in market performance under the two market mechanisms

6.2 The Model Framework of the Batch Auction Market in the Messy Network

This chapter investigates the impact of batch auctions on the flash crash model in a messy network. The most significant difference between the batch and the continuous double auction is whether submitted orders are executed immediately or held until a certain point for matching. In this section, we apply the volume-based batch auction and fix the batch size for each simulation.

6.2.1 Batch Auction Setups

In our batch auction market, the exogenous batch size is fixed to be k during all simulation rounds, which means the exchange will execute the matching mechanism for every k traders arriving. For each trading round, the trader could take one of the following three kinds of actions: submitting a buy order; submitting a sell order; or not submitting an order. We also call the trading batch a trading window. As the order flow is composed of several batches, the whole trading period is divided into several sequential trading windows.

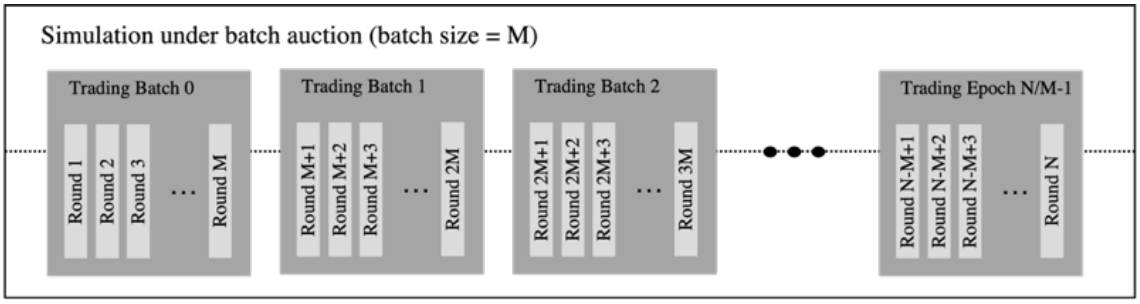


Figure 6.1 The Trading Flow under the Batch Auction Mechanism

The assumptions of the market with batch auction mechanism are referred to the ones in Chapters 4 and 5. In addition, in the batch auction, the batch size remains fixed during

the whole simulation as Fig. 6.1 shows. Once an order comes into the exchange, the order will be kept in a batch. As soon as the batch is released, the public will know how many buy orders and sell orders have been executed and at what trade price.

In the case of the batch size being 1, the batch auction is equivalent to the CDA market specified in Chapters 4 and 5.

6.2.2 The Market Maker

Similar to the assumptions set in the CDA model, the market maker is a monopolistic market maker to provide enough liquidity to the market, so the market maker guarantees the other traders' orders to be matched by simultaneously placing adequate buy and sell orders at reasonable prices. At the beginning of a trading window, the market maker will broadcast the bid and ask prices, and then k agents in the trading crowd are sequentially selected to submit their orders which are stored in the market but not executed immediately. All buy (sell) orders submitted during the trading windows are no less (higher) than the best ask (bid) price, and they are not broadcast publicly until the end of the window. As soon as the batch is released, there will be three cases that the market maker will consider about:

1. If the quantity of the buy orders is equal to that of the sell orders, the market maker does not need to take orders from any sides because the orders of the two sides will be cleared automatically with each other at the market price.

2. If the quantity of buy orders is larger than that of the sell orders, it means the buy side is dominated in the current trading window. All outstanding orders are buy-side at the end of the trading window, and the market maker must take these orders by consuming the inventory at the shouted ask price.

3. If the quantity of sell orders is larger than that of the but orders, it means the sell side is dominated in the current trading window. At the end of the trading window, all outstanding orders are sell-side, and the market maker must take these orders by paying the extra cost at the shouted bid price.

For example, if there are 3 buy orders and 2 sell orders in one batch, at the moment that batch is released, 2 of 3 buy orders will match with the 2 sell orders, but the rest of the buy orders, which are outstanding, will be taken by the market maker. In our model, the market maker can update the best ask and bid prices towards a level which possibly balances the buy side and sell side. As soon as the batch is released, the market maker can observe the order flows in the latest batch to adjust their quotes with the formulas shown in Appendix E.

We hold the assumption that the last section mentioned that the market maker is risk neutral. Therefore, the market maker has a zero-profit expectation which means the expected profits EP_i should be 0. Now assume that there are m buy orders and n sell orders in the i-th batch, under such an assumption, then we can know the market maker will adjust the bid and ask prices according to the following formulas:

$$\begin{cases} A_{i+1} = \mathrm{E}[V_i | Q_i^b = m, Q_i^s = n], \text{ if } m > n \\ B_{i+1} = \mathrm{E}[V_i | Q_i^b = m, Q_i^s = n], \text{ if } m < n \end{cases} \tag{6.1}$$

The updated bid and ask prices are applied to the outstanding orders after the next batch.

Again, to determine the updated quotes, we must start with the conditional expectation on the right-hand side of Equation 6.1.

$$\mathrm{E}[V_i | Q_i^b = m, Q_i^s = n] = \sum_{v \in \mathbb{V}_n} v \mathrm{Pr}_{MM}(V_i = v | Q_i^b = m, Q_i^s = n) \tag{6.2}$$

Appendix D illustrates all detailed computations about the deduction of conditional probability.

6.2.3 Uninformed Fundamental Traders in the Batch Auction Markets

This chapter also considers the case of intelligent uninformed fundamental traders. Apart from the market maker and the uninformed fundamental traders, all other trading

agents are modelled as Chapters 4 and 5 specify. Similar to the settings in Chapter 4, the intelligent UFTs receive biased price signals once the fake shock comes into the trading crowd. In the case of batch auctions, the UFTs will not get any public trade information until a batch is released, but they are allowed to know the status of the submitted orders in the batch once the trade information is released to the public. The detailed assumptions are illustrated as follows.

Assumptions

- The uninformed fundamental traders all know about the market structure, but there is no connection between any two agents in a messy network;
- When an uninformed fundamental trader observes an incoming shock, the UFT will initially trust the information but simultaneously adjust their beliefs regarding whether they are being deceived. The initial belief of each UFT regarding whether they are being deceived by the shock is set randomly;
- Each uninformed fundamental trader can only update the beliefs on whether they are deceived as soon as a trading batch opens to the public;
- Each uninformed fundamental trader will stop learning once the they are supposed to recognise whether the information (the shock causing a price decline) is real based on the learning process shown in Fig. 6.3.

Learning Strategy

Referring to the learning process specified in Chapter 5, an intelligent UFT lacks inside information, so the trader needs to infer whether him or her received a biased or unbiased price signal at every trading round by observing the past order flows or called order batched in this chapter, more specifically.

Similar to the setups in Chapter 5, once a new shock arrives in the market, an intelligent UFT adjust the beliefs $\Theta = [\theta_0, \theta_1]$, where

1. θ_0: The shock is real, so the true stock price moves from v to v'. Accordingly, the price signal that a fundamental trader receives changes from w to w';

2. θ_1: The shock is fake, so the true stock price remains to be v;

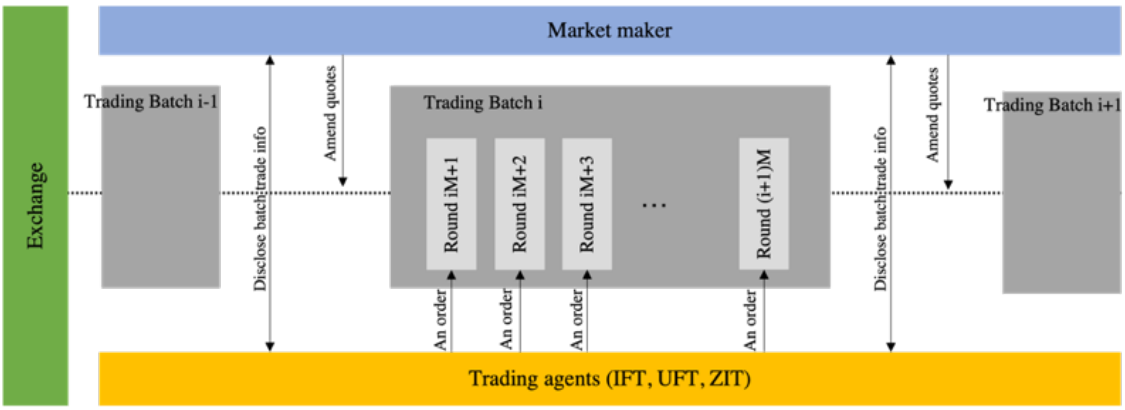


Figure 6.2 The Interaction of Agents in the Batch Auction Market

After the i-th batch ends, each uninformed fundamental trader adjusted the beliefs about the state of the system $P_{x_i,t}(\Theta) = \left[P_{x_i,t}(\theta_0), P_{x_i,t}(\theta_1)\right]$. $P_{x_i,0}(\Theta)$, which is randomly set, refers to the initial beliefs on different states as soon as the shock occurs. Referring the Fig. 6.2, the UFTs update their beliefs as follows:

- Once a shock arrives, a UFT will change the fundamental price estimate with the limited information regardless of whether the shock is fake or not and set the initial prior probabilities to be $[P(\theta_0), P(\theta_1)]$.
- Each UFT obtains the price signal at each trading round. However, no other information is used for inference until a batch opens.
- Once the current trading window opens, each UFT applies the Bayesian inference (detailed works shown in Appendix D) to update the beliefs.
- Given a threshold c, the UFT stops inferring, once $\min\{P(\theta_0), P(\theta_1)\} < c$.

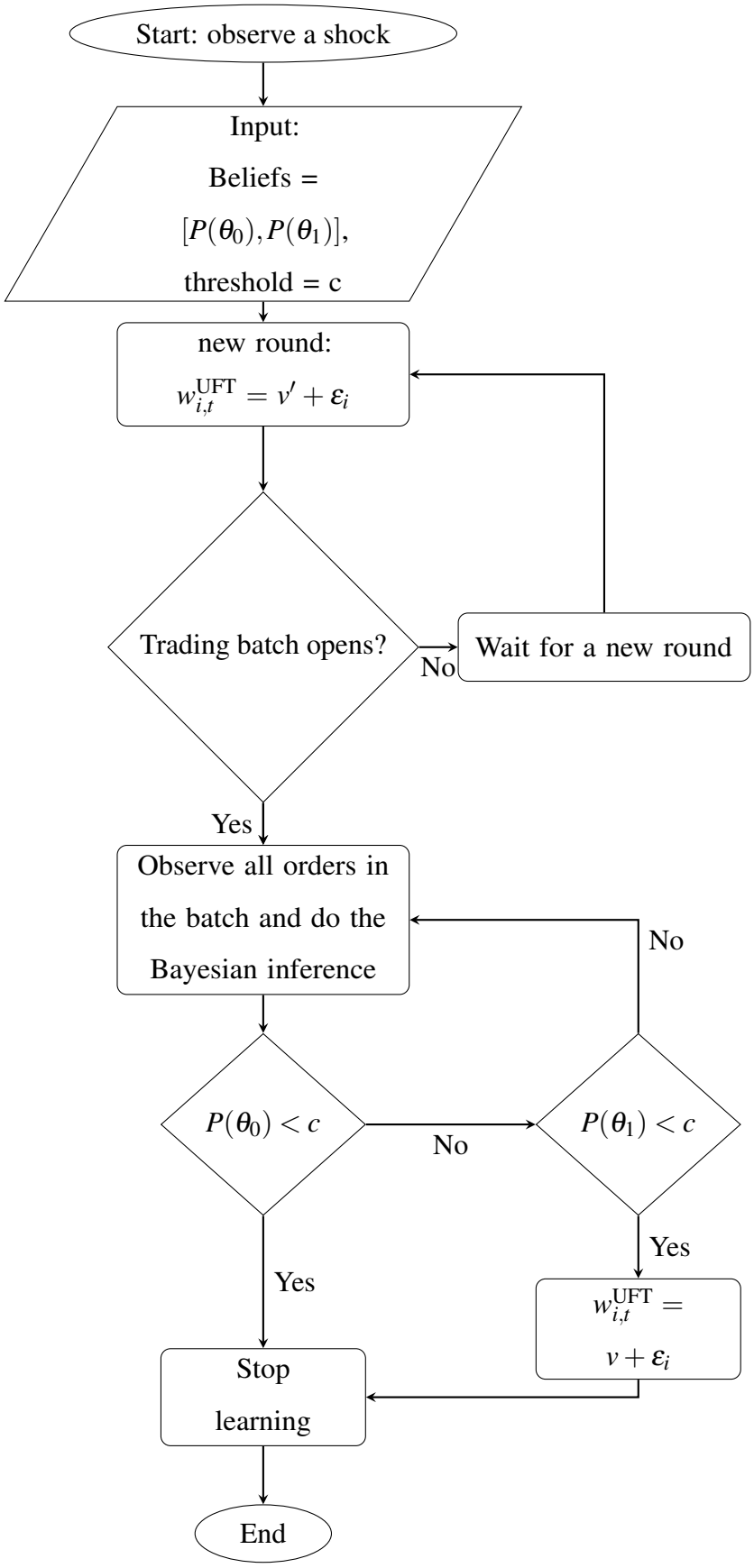


Figure 6.3 Flow Chart of Learning Process for the Uninformed Fundamental Traders in Batch Auction Markets

Different from the framework in Chapter 5, in the messy network, the intelligent uninformed fundamental traders infer by observing the number of each kind of order in the batch. In an M-size batch auction market, denote the numbers of buy and sell orders as Q_b and Q_s, respectively. Let $I_{i,t}$ stand for the information that the i-th trading agents can receive after trading batch k then the learning strategy can be described as the same as mentioned in Eq. 5.2:

$$P_{x_j,t}(\Theta = \theta_i | I_{i,t}) = \frac{P_{x_i,t}(I_{i,t}|\Theta = \theta_j) P_{x_i,t}(\Theta = \theta_j)}{\sum_{j=1,2} P_{x_i,t}(I_{i,t}|\Theta = \theta_j) P_{x_i,t}(\Theta = \theta_j)} \tag{6.3}$$

In the case of the M-size batch auction, the information $I_{i,t}$ can be specifically expressed as $Q_b = m, Q_s = n$, so the probability $P_{x_i,t}(I_{i,t}|\Theta = \theta_j)$ can be explained as the probability of that there are m buy orders and n sell orders in the batch under the belief $\Theta = \theta_j$. The full maths of an intelligent fundamental trader updating the beliefs are illustrated in Appendix D.2.

6.3 Experiments

This chapter still runs according to the short-term experiment framework extended from Chapters 4 and 5. Namely, all trading activities happen in a single trading epoch where only one shock hits the market.

6.3.1 Basic Settings

This chapter discusses the flash crash model with batch auction when the batch sizes vary. Combined with the experimental results in Chapters 4 and 5, this chapter will investigate how the batch auction mechanism affects. Similar to the settings in the previous chapters, the artificial market contains n_{total} traders interacting with the market maker and the exchange within a trading epoch of N rounds. A shock hits at k-th round, dividing the whole trading epoch into the buffer stage over the range from round 1 to round k and the active stage over the range from round $k+1$ to round M.

6.3.2 Parameters

The main parameter to be investigated is the **batch size (k)**, and k is selected from $\{2,4,6\}$. To compare with the results obtained in Chapters 4 and 5, the settings of other parameters, informational (σ_{ε_n}) and market-structure parameters (n_{IFT} and n_{UFT}), remain unchanged as introduced in Section 4.1.2.

6.3.3 Performance Metrics

The discussion of metrics remains consistent with Chapter 5.

Metrics	Explanation
Crash size	The biggest price decline of the market prices after the fake shock and before all UFTs finish inference
Crash duration	The duration between the shock point and the one reaching the biggest price decline
Recovery duration	The duration between the point reaching the biggest price decline and the turning point of a new steady state
Cumulative profits	Average cumulative profits of different types of trading agents
Information advantage	A quantitative indicator measuring the information advantage of a single IFT to a single UFT on average
Inference accuracy	Some indicators to measure how accurate a single UFT can make a correct inference

Table 6.1 Performance Metrics for the Studies of Batch Auction Markets

6.4 Experimental Results of the Market with Intelligent UFTs

6.4.1 Setups

Chapters 4 and 5 discuss how the model behaves under information asymmetry in a continuous double auction market. The difference between CDA and batch auction lies in whether each submitted order is executed immediately or waits to be batched with other orders so that they can be executed simultaneously. In fact, CDA can be regarded as a special case of the batch auction, in which case the batch size is equal to 1.

In this part, we set the batch size to three different values: 2, 4, and 6, and all other setups are referred to those in Chapters 4 and 5. We examine what happens when the market runs different sizes of batch auctions.

6.4.2 Global Statistics Overview

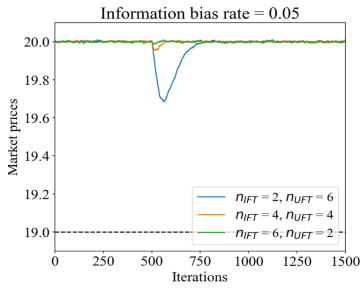
Similar to what was studied in the previous chapters, we first examine the global statistics on price movements and the agents' profit trends under different parameter settings. Similarly, in our 10-agent system, we still consider three cases, an IFT-dominated market, a UFT-dominated market, and an information-balanced market, with the same market structure configurations as described in Section 4.2.1.

Here we also include the simulation results in the CDA market where the messy agents' network is involved for comparison.

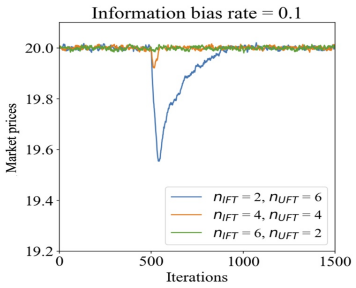
Prices Movements

We first briefly show, if a fake shock comes, how the average market price varies in the UFT-dominated, IFT-dominated and balanced markets, respectively, under different batching frequencies in the batch auction models. Firstly, we can find preliminary results regarding the batch auction market consistent with those in the CDA market of the messy network - the market price crashes more significantly in the markets with larger risks.

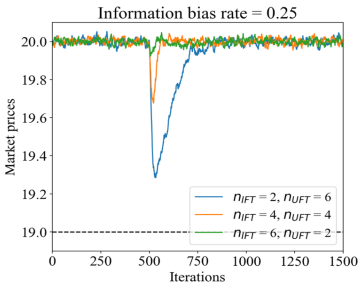
Furthermore, in batch auctions, the crash size also tends to increase as the market becomes increasingly UFT-dominated.



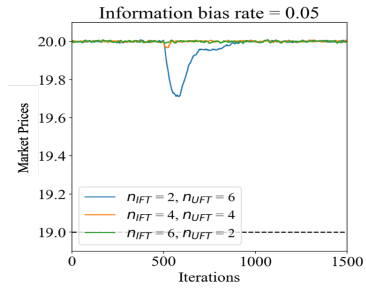
(a) IBR $\sigma_m = 0.05$, CDA



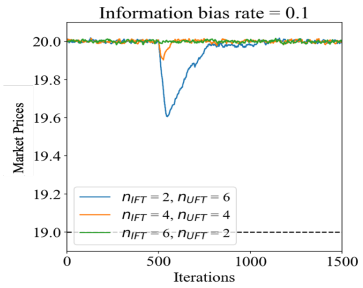
(b) IBR $\sigma_m = 0.1$, CDA



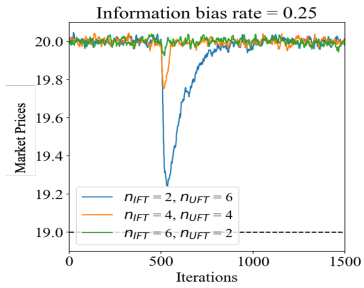
(c) IBR $\sigma_m = 0.25$, CDA



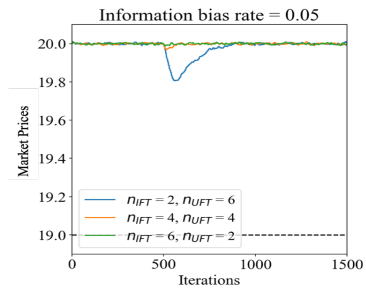
(d) IBR $\sigma_m = 0.05$, batch size=2



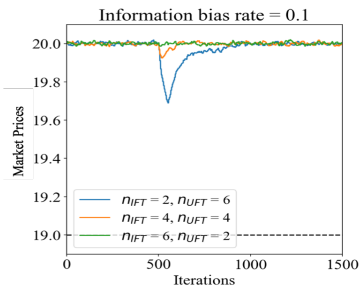
(e) IBR $\sigma_m = 0.1$, batch size=2



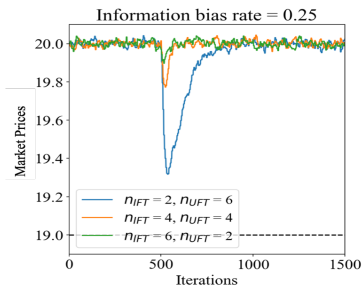
(f) IBR $\sigma_m = 0.25$, batch size=2



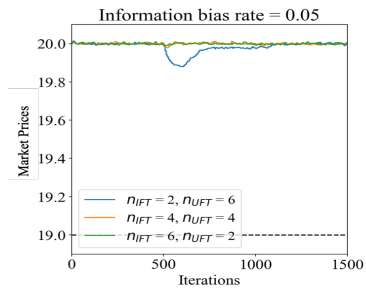
(g) IBR $\sigma_m = 0.05$, batch size=4



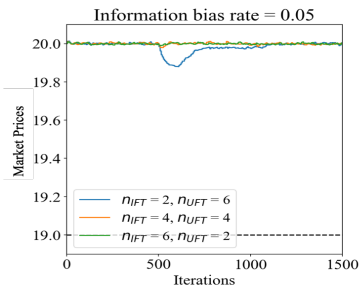
(h) IBR $\sigma_m = 0.1$, batch size=4



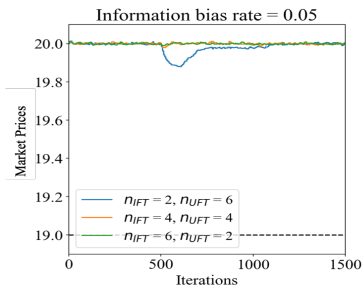
(i) IBR $\sigma_m = 0.25$, batch size=4



(j) IBR $\sigma_m = 0.05$, batch size=6



(k) IBR $\sigma_m = 0.1$, batch size=6



(l) IBR $\sigma_m = 0.25$, batch size=6

Figure 6.4 Batch Auction: The Average Market Prices Movements (Fake Shock)

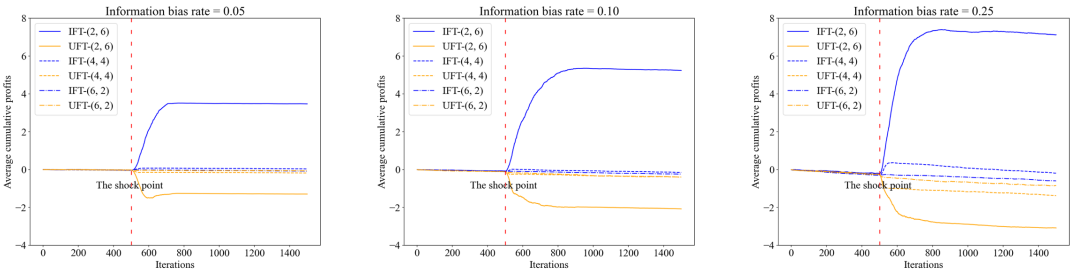
In addition, although few clear changes in the crash size are observed, we have discovered a finding that is specific to the batch auction model: As the frequency of batching becomes lower, i.e. as the batch size increases, the market price crashes less significantly, which is examined by statistical testing in Section 6.4.6. This result also suggests that the batch auction could weaken flash crashes to some degree.

Profits of Agents

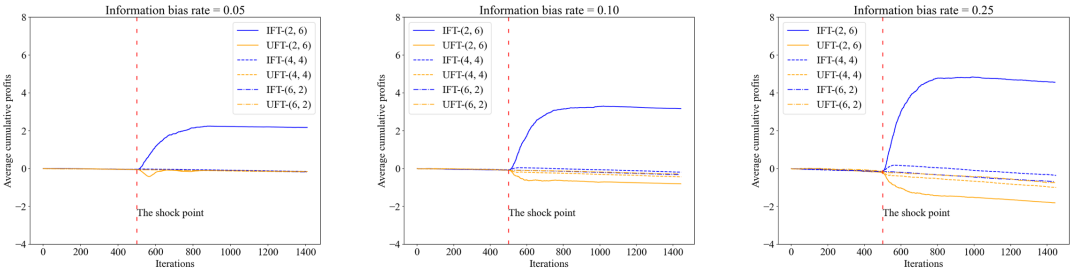
We move on to examine the dynamics from the perspective of agents. Fig. 6.5 shows the average profit dynamics for the fundamental traders, for certain parameter settings.

With Fig. 6.5 (d-l) the profit dynamics of different agents in different batch auction markets, we find that the patterns of profit dynamics for fundamental traders are also similar to that in the previous messy network study. The average profit curves of IFTs and UFTs bifurcate after the fake shock arrives. However, an interesting finding is that the batch auction could affect this bifurcation size.

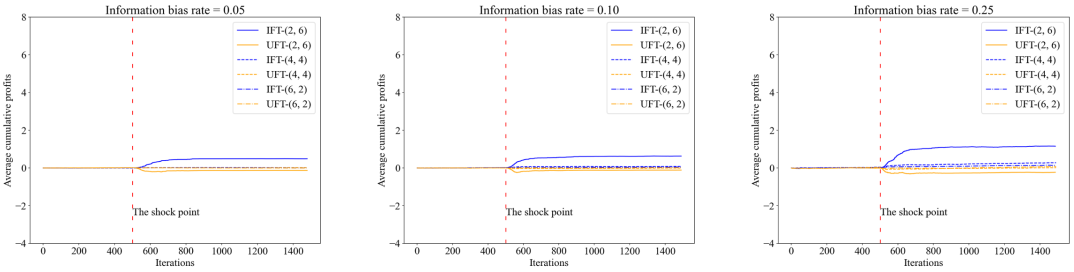
As the batch size increases, we can clearly see that the profit curves of IFTs and UFTs bifurcate more narrowly if the market structure and market risk are kept fixed. Moreover, when other parameters remain unchanged, the average cumulative profit of IFTs decreases as the batch size increases, while, in contrast, the average cumulative profit of relative UFTs increases. In other words, the batch auction can potentially reduce the profit differences between the two kinds of fundamental traders, leading to a less informational advantage for informed fundamental traders.



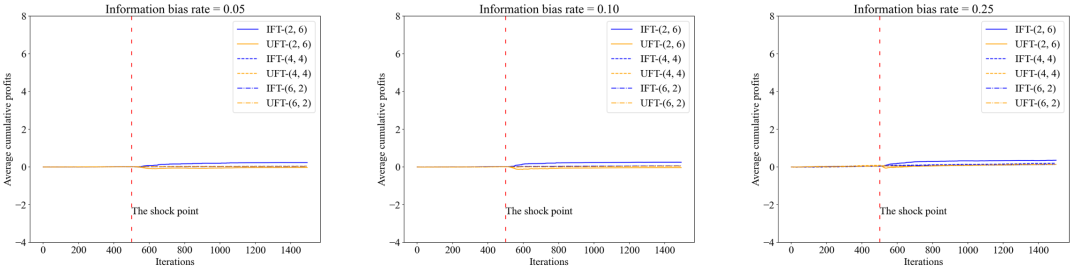
(a) IBR $\sigma_m = 0.05$, CDA (b) IBR $\sigma_m = 0.1$, CDA (c) IBR $\sigma_m = 0.25$, CDA



(d) IBR $\sigma_m = 0.05$, batch size=2 (e) IBR $\sigma_m = 0.1$, batch size=2 (f) IBR $\sigma_m = 0.25$, batch size=2



(g) IBR $\sigma_m = 0.05$, batch size=4 (h) IBR $\sigma_m = 0.1$, batch size=4 (i) IBR $\sigma_m = 0.25$, batch size=4



(j) IBR $\sigma_m = 0.05$, batch size=6 (k) IBR $\sigma_m = 0.1$, batch size=6 (l) IBR $\sigma_m = 0.25$, batch size=6

Figure 6.5 Batch Auction: The Average Cumulative Profits of Fundamental Traders (Fake Shock)

Moreover, when other parameters remain unchanged, the average cumulative profit of IFTs decreases as the batch size increases, while, in contrast, the average cumulative profit of relative UFTs increases. An interesting finding is that, when the batch size=6, the

average cumulative profit curves of IFTs and UFTs almost coincide. This implies that the batch auction mechanism can greatly reduce the information advantage of informed traders in the case the batch size is large. However, on the other hand, the batching mechanism leads the informed traders not to obtain great excess profits, which may also reduce their trading enthusiasm.

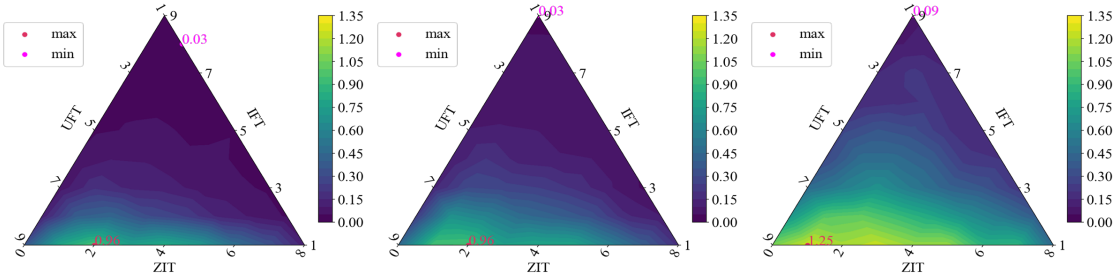
The slope of the cumulative profits curve represents the expected profit obtained when a trader places an order. Now we define the "natural ordering profit" as the slope of the cumulative return curve in the buffer stage where the market prices are stable, and no shock is included. Therefore, the negative slope of the cumulative return curve of fundamental traders means that the natural returns of fundamental traders are negative.

Another finding is that the slope of the profit curve of both fundamental traders decreases with the batch size increasing when other parameters are set constant. In other words, the batch auction mechanism could reduce the negative "natural ordering profit" of the fundamental traders.

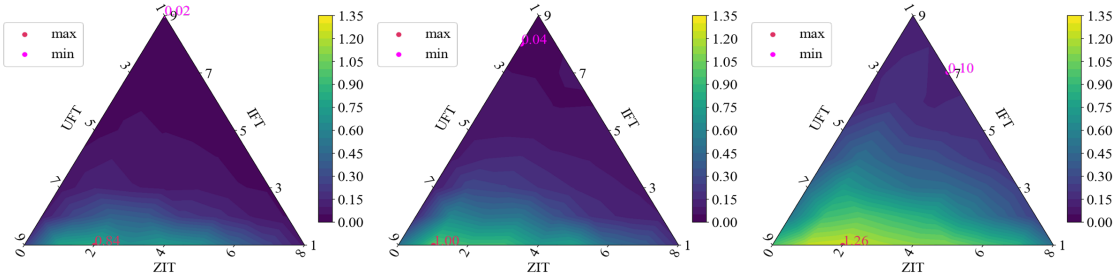
6.4.3 Sensitivity Analysis with Fake Shocks

Crash Size

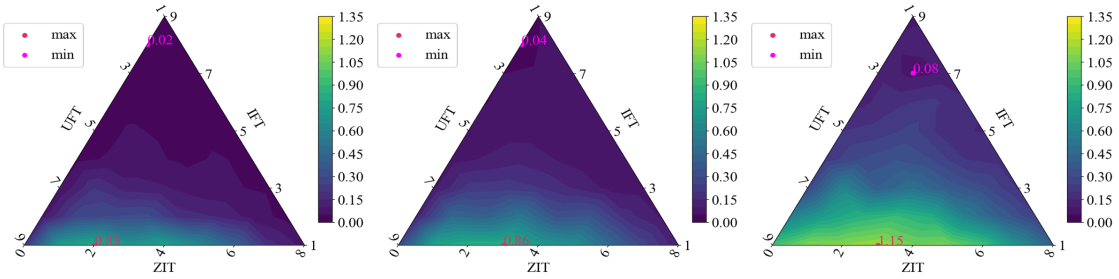
This section presents a complete sensitivity analysis of the crash sizes in the batch auction model. Together with the studies in the CDA market, we finally examine four different batch auction models with the batch size varying from 1, 2, 4, 6. In Section 6.4.2, according to the snapshots of the average market price with some specific parameters, we preliminarily find that the crash sizes could be reduced as the batch size increases. In this chapter, we further examine the effects of the batch auction on crash sizes by contour heatmaps for the simulations with all parameters and under different batch auction mechanisms.



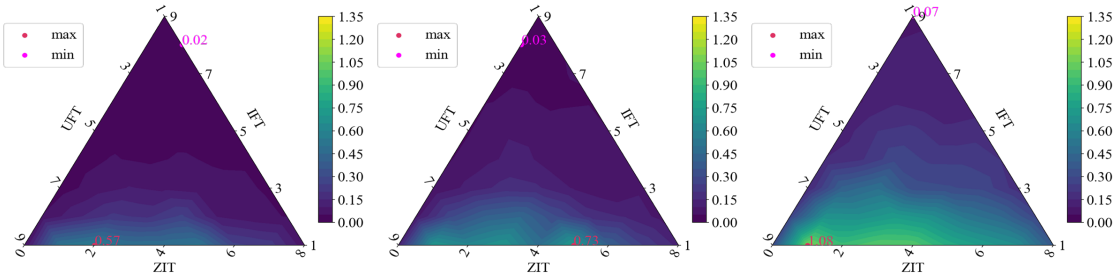
(a) IBR $\sigma_m = 0.05$, CDA (b) IBR $\sigma_m = 0.1$, CDA (c) IBR $\sigma_m = 0.25$, CDA



(d) IBR $\sigma_m = 0.05$, batch size=2 (e) IBR $\sigma_m = 0.1$, batch size=2 (f) IBR $\sigma_m = 0.25$, batch size=2



(g) IBR $\sigma_m = 0.05$, batch size=4 (h) IBR $\sigma_m = 0.1$, batch size=4 (i) IBR $\sigma_m = 0.25$, batch size=4



(j) IBR $\sigma_m = 0.05$, batch size=6 (k) IBR $\sigma_m = 0.1$, batch size=6 (l) IBR $\sigma_m = 0.25$, batch size=6

Figure 6.6 Batch Auction: Average Crash Sizes (Fake Shock)

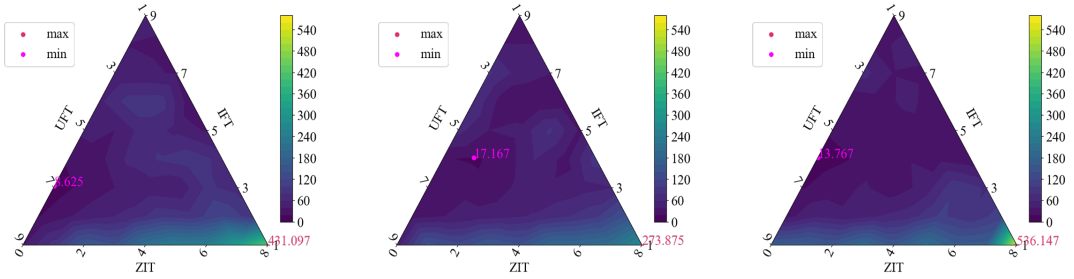
The crash size patterns of the batch auction's contour graphs are similar to those studied in the CDA market — a more UFT-dominated market could suffer a more significant crash. Also, it is obvious that the areas standing for UFT-dominated markets become darker as the batch size increases. This also coincides with our previous speculation: batch auction can

weaken flash crash, and the larger the batch size affects more greatly. Furthermore, we find that batch auction has a greater impact on crash size in low-risk markets than in high-risk markets. Numerically, when the market structure is $n_{\text{IFT}} = 1, n_{\text{UFT}} = 7$, the average crash size of the batch auction market with a batch size of 6 is 40.63% lower than that in the CDA market, when the market risk is 0.05. However, under the same market structure, as the batch size is 6, the simulation results of the BA market at market risks of 0.1 and 0.25 show that the corresponding average crash sizes are decreased by 35.60% and 21.37% compared with the CDA market, respectively.

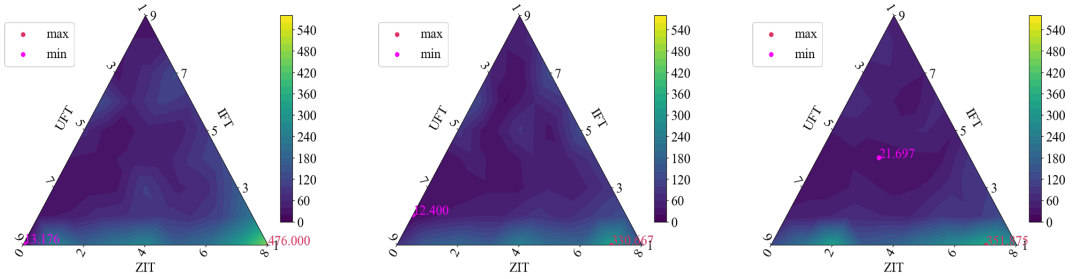
Crash and Recovery

After a fake shock arrives, there will be a price decline (called a crash) and a rebound (called a recovery) in market prices. To investigate the market's reflection of a fake shock, we explored the average crash and recovery duration and the average price change speed in different batch auctions at various batching frequencies.

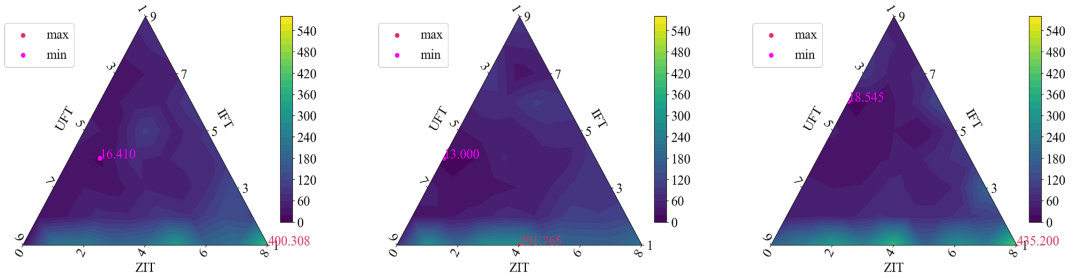
Fig. 6.7 and Fig. 6.8 show the duration of crashes and recoveries that occur under different batch auctions. However, by comparing the contour plots at different batch sizes, we do not find obvious pattern differences from the duration perspective on the effect of the batch auctions.



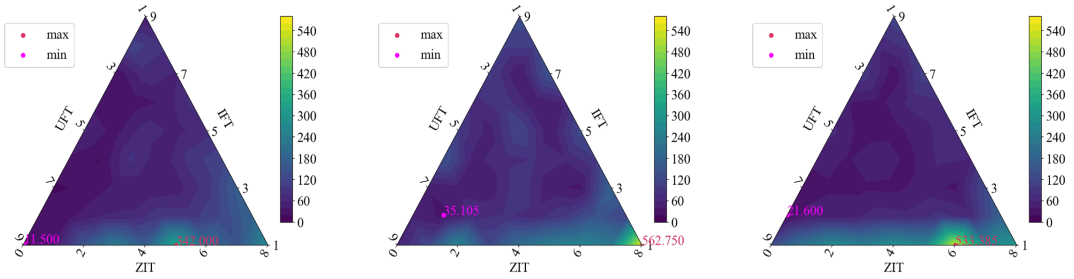
(a) IBR $\sigma_m = 0.05$, CDA (b) IBR $\sigma_m = 0.1$, CDA (c) IBR $\sigma_m = 0.25$, CDA



(d) IBR $\sigma_m = 0.05$, batch size=2 (e) IBR $\sigma_m = 0.1$, batch size=2 (f) IBR $\sigma_m = 0.25$, batch size=2



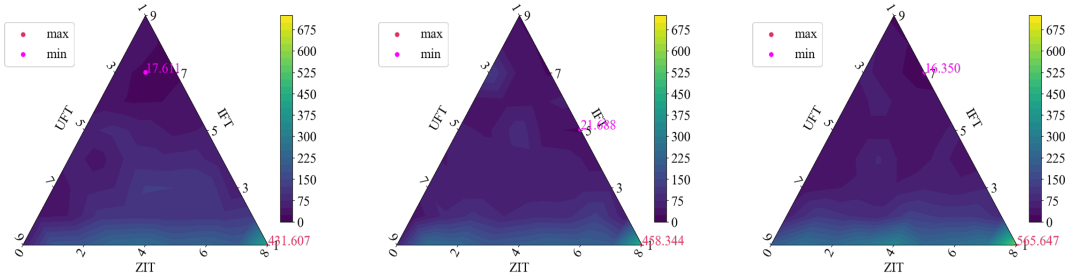
(g) IBR $\sigma_m = 0.05$, batch size=4 (h) IBR $\sigma_m = 0.1$, batch size=4 (i) IBR $\sigma_m = 0.25$, batch size=4



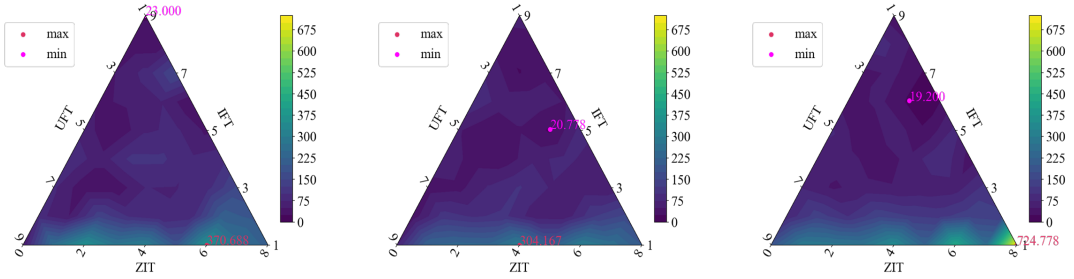
(j) IBR $\sigma_m = 0.05$, batch size=6 (k) IBR $\sigma_m = 0.1$, batch size=6 (l) IBR $\sigma_m = 0.25$, batch size=6

Figure 6.7 Batch Auction: Average Crash Duration (Fake Shock)

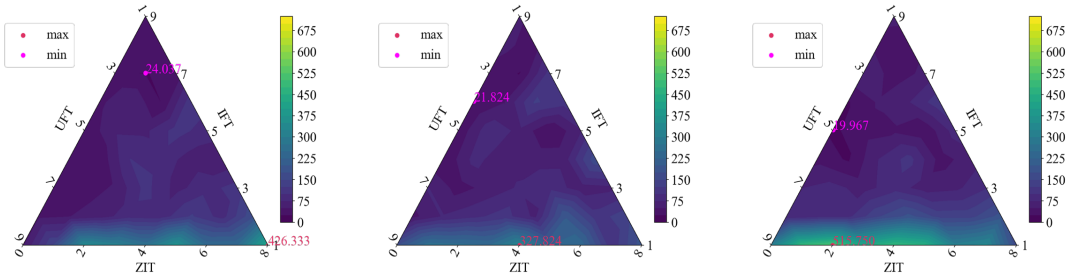
Namely, it is not clear to show that setting the batch auction can shorten the flash crash duration or recovery duration. Section 6.4.5 also investigates the inference duration for an uninformed fundamental trader, showing that the inference duration may not be related to batch auctions.



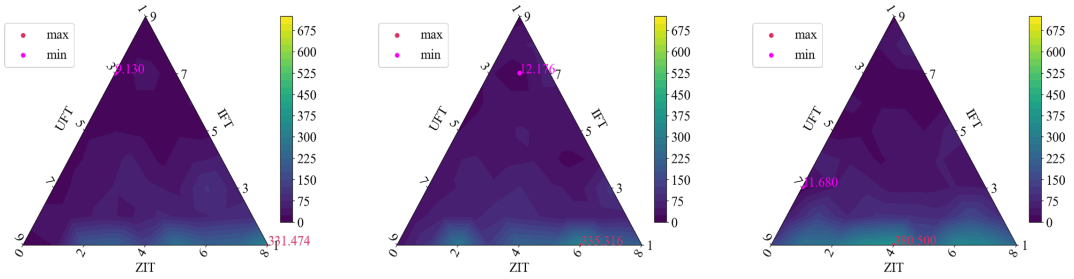
(a) IBR $\sigma_m = 0.05$, CDA (b) IBR $\sigma_m = 0.1$, CDA (c) IBR $\sigma_m = 0.25$, CDA



(d) IBR $\sigma_m = 0.05$, batch size=2 (e) IBR $\sigma_m = 0.1$, batch size=2 (f) IBR $\sigma_m = 0.25$, batch size=2



(g) IBR $\sigma_m = 0.05$, batch size=4 (h) IBR $\sigma_m = 0.1$, batch size=4 (i) IBR $\sigma_m = 0.25$, batch size=4

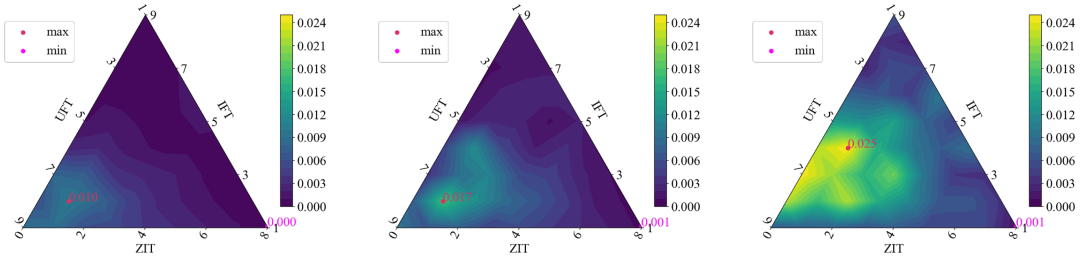


(j) IBR $\sigma_m = 0.05$, batch size=6 (k) IBR $\sigma_m = 0.1$, batch size=6 (l) IBR $\sigma_m = 0.25$, batch size=6

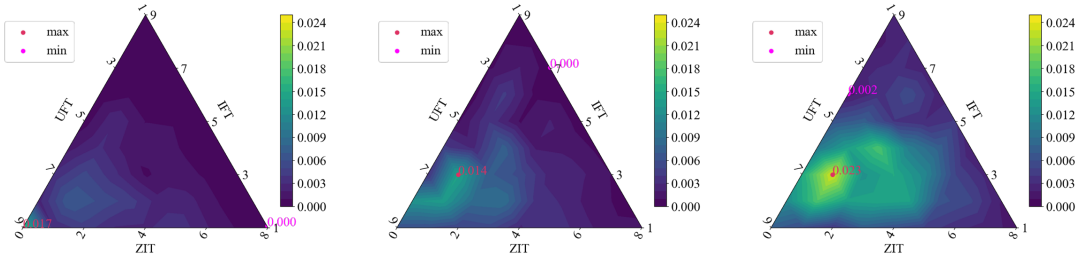
Figure 6.8 Batch Auction: Average Recovery Duration (Fake Shock)

As Chapter 4 examines, the market's maximum price decline differs across market structures in the non-intelligent-UFT model. Therefore, we can capture the effect of the market on fake shock by calculating the average speed of price decline and recovery. Fig. 6.9 and Fig. 6.10 show the crash and recovery speed contour plots, respectively. We find

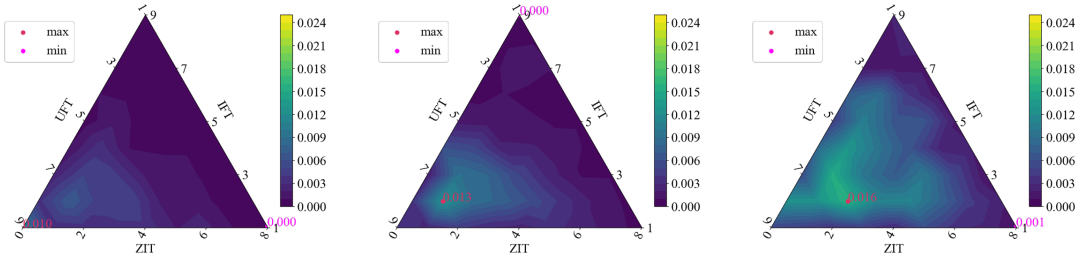
that a low-frequency batch auction can slow down the speed of price decline. The strength of the contour plot also shows that when the market is UFT-dominated and fundamental, i.e. when there are a large number of fundamental traders in the market, a majority of which are UFTs, the market price changes at the fastest rate. In other words, market prices are most sensitive to the fake shock in such cases.



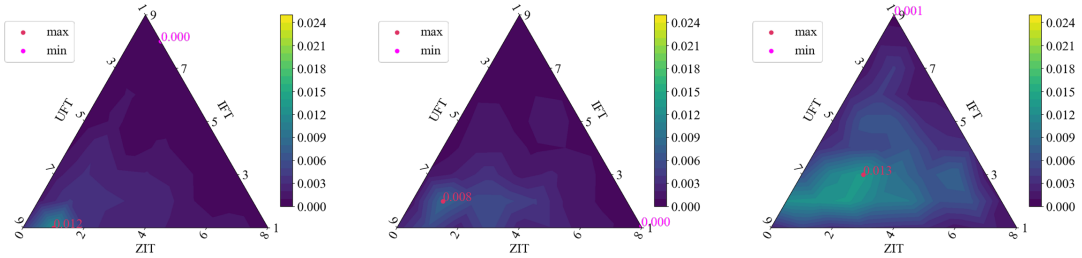
(a) IBR $\sigma_m = 0.05$,CDA (b) IBR $\sigma_m = 0.1$,CDA (c) IBR $\sigma_m = 0.25$,CDA



(d) IBR $\sigma_m = 0.05$,batch size=2 (e) IBR $\sigma_m = 0.1$,batch size=2 (f) IBR $\sigma_m = 0.25$,batch size=2



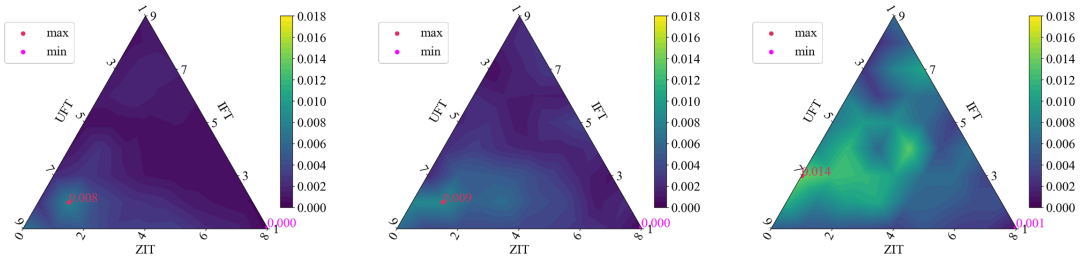
(g) IBR $\sigma_m = 0.05$,batch size=4 (h) IBR $\sigma_m = 0.1$,batch size=4 (i) IBR $\sigma_m = 0.25$,batch size=4



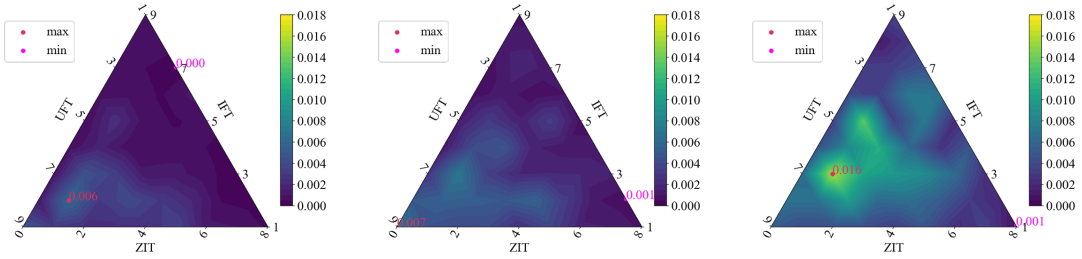
(j) IBR $\sigma_m = 0.05$,batch size=6 (k) IBR $\sigma_m = 0.1$,batch size=6 (l) IBR $\sigma_m = 0.25$,batch size=6

Figure 6.9 Batch Auction: Average Crash Speed

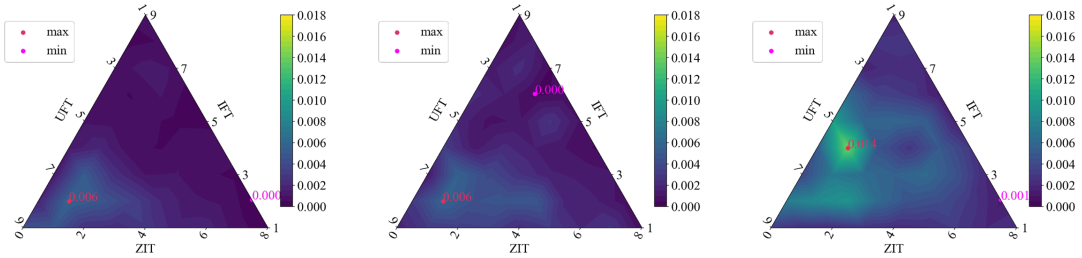
The recovery duration plots show roughly similar patterns to the crash plots, but, according to Fig. 6.10, the recovery speed seems to be slower than the crash speed because it takes time for the uninformed fundamental traders to change their trading strategies, and each UFT needs different time periods to finish the inference.



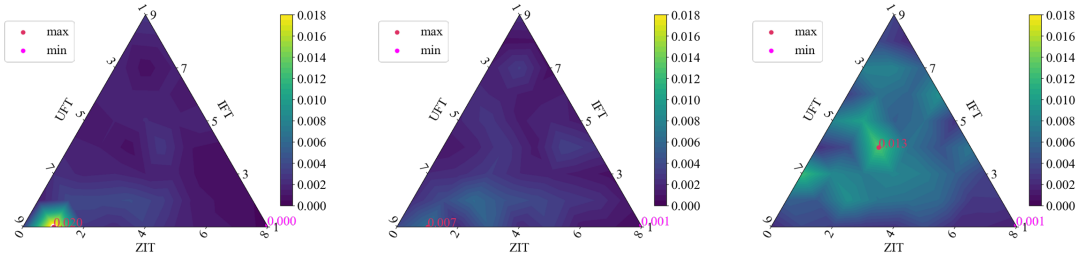
(a) IBR $\sigma_m = 0.05$,CDA (b) IBR $\sigma_m = 0.1$,CDA (c) IBR $\sigma_m = 0.25$,CDA



(d) IBR $\sigma_m = 0.05$,batch size=2 (e) IBR $\sigma_m = 0.1$,batch size=2 (f) IBR $\sigma_m = 0.25$,batch size=2



(g) IBR $\sigma_m = 0.05$,batch size=4 (h) IBR $\sigma_m = 0.1$,batch size=4 (i) IBR $\sigma_m = 0.25$,batch size=4

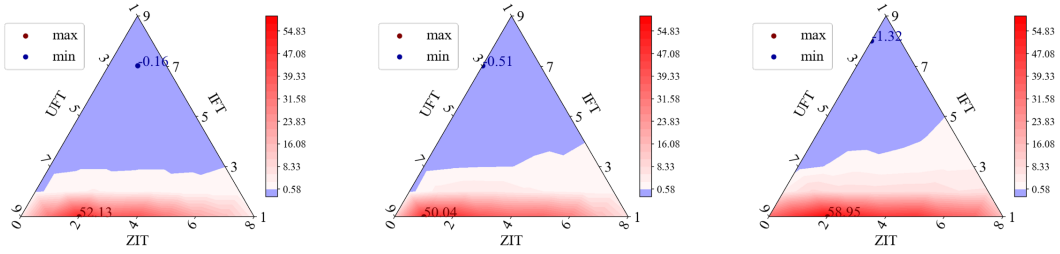


(j) IBR $\sigma_m = 0.05$,batch size=6 (k) IBR $\sigma_m = 0.1$,batch size=6 (l) IBR $\sigma_m = 0.25$,batch size=6

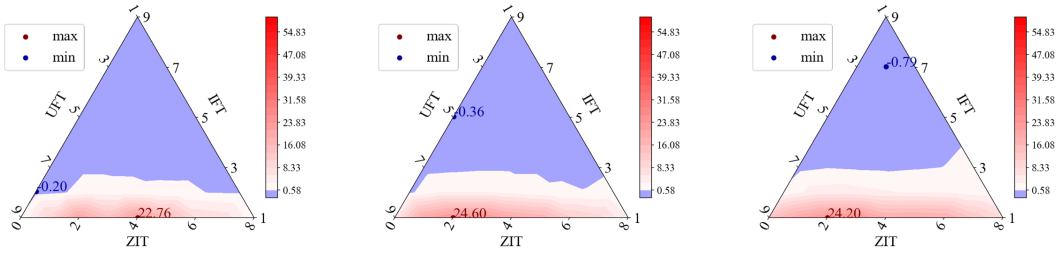
Figure 6.10 Batch Auction: Average Recovery Speed

Agents' Average Profits

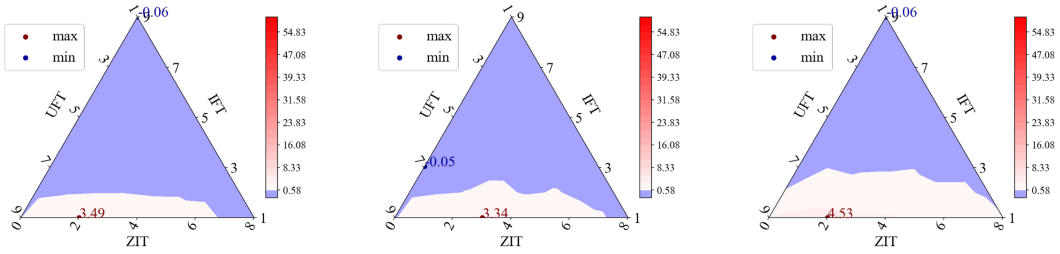
In the CDA market, the simulations have shown the average profits obtained by informed fundamental traders affect the market structurally: when the share of IFTs in the market accounts for more than 40%, each IFT only obtains a negative profit on average. However, as the market becomes increasingly UFT-dominated, the average profit of each IFT rises dramatically, especially in the case of $n_{\text{IFT}} = 1$ in the market.



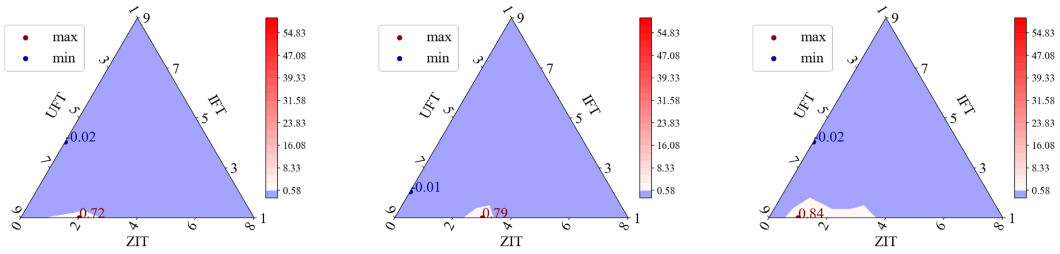
(a) IBR $\sigma_m = 0.05$,CDA (b) IBR $\sigma_m = 0.1$,CDA (c) IBR $\sigma_m = 0.25$,CDA



(d) IBR $\sigma_m = 0.05$, batch size=2 (e) IBR $\sigma_m = 0.1$, batch size=2 (f) IBR $\sigma_m = 0.25$, batch size=2



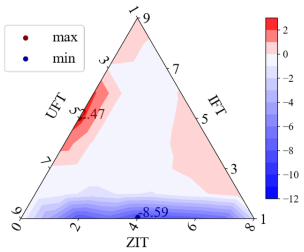
(g) IBR $\sigma_m = 0.05$, batch size=4 (h) IBR $\sigma_m = 0.1$, batch size=4 (i) IBR $\sigma_m = 0.25$, batch size=4



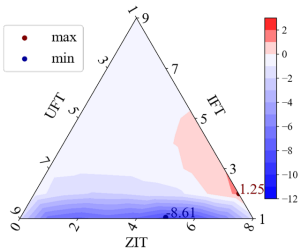
(j) IBR $\sigma_m = 0.05$, batch size=6 (k) IBR $\sigma_m = 0.1$, batch size=6 (l) IBR $\sigma_m = 0.25$, batch size=6

Figure 6.11 Batch Auction: IFTs' Average Cumulative Profits

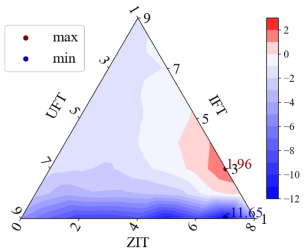
However, the batch auction reduces the maximum average maximum profits that each informed fundamental trader could obtain. With an example in the case of $n_{\text{IFT}} = 1, n_{\text{UFT}} = 7, \text{IBR} = 0.05$, compared with the CDA market, the average profits of a single IFT in a batch auction market with the batch size of 2 has dropped by 58.95% from 58.95 to 24.20. Also, when the batch size rises to 4, the average profits drop by 92.31% compared with the CDA market, and each IFT has almost no profits as the batch size is set up to 6.



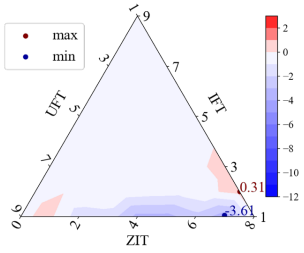
(a) IBR $\sigma_m = 0.05$,CDA



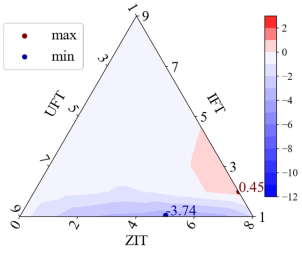
(b) IBR $\sigma_m = 0.1$,CDA



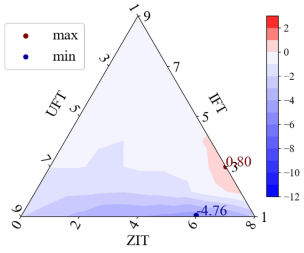
(c) IBR $\sigma_m = 0.25$,CDA



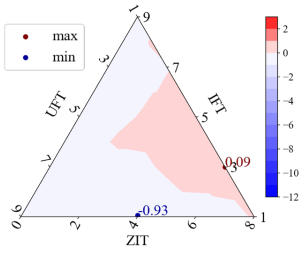
(d) IBR $\sigma_m = 0.05$, batch size=2



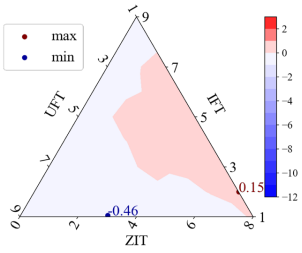
(e) IBR $\sigma_m = 0.1$, batch size=2



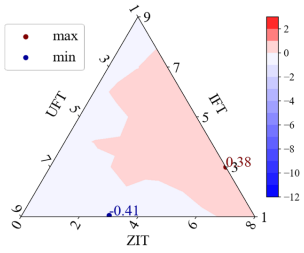
(f) IBR $\sigma_m = 0.25$, batch size=2



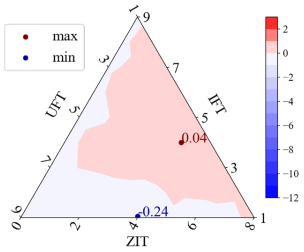
(g) IBR $\sigma_m = 0.05$, batch size=4



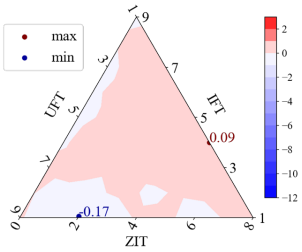
(h) IBR $\sigma_m = 0.1$, batch size=4



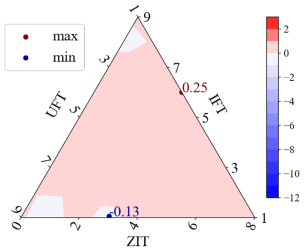
(i) IBR $\sigma_m = 0.25$, batch size=4



(j) IBR $\sigma_m = 0.05$, batch size=6



(k) IBR $\sigma_m = 0.1$, batch size=6



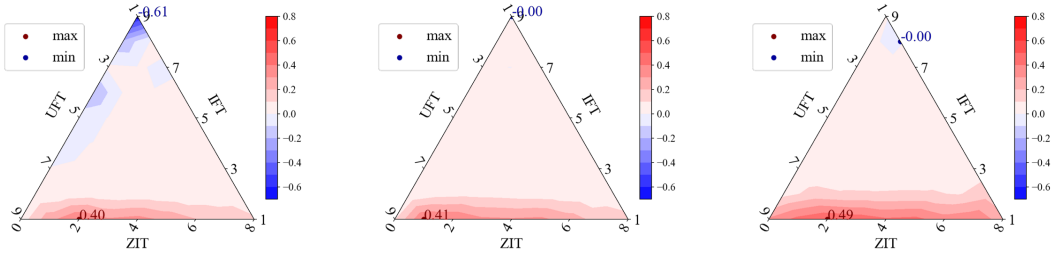
(l) IBR $\sigma_m = 0.25$, batch size=6

Figure 6.12 Batch Auction: UFTs' Average Cumulative Profits (Fake Shock)

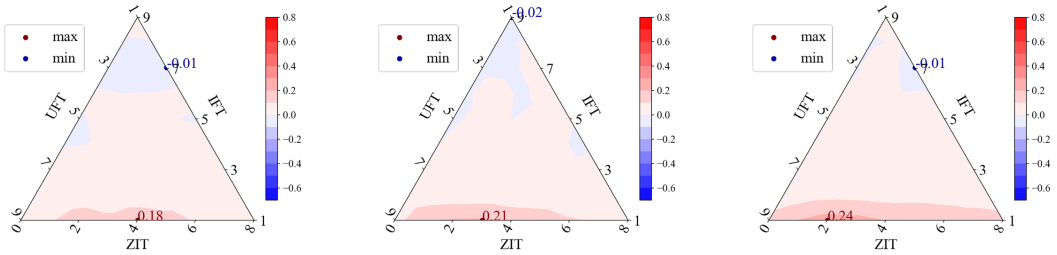
Similarly, the batch auction has significantly changed UFT's profit patterns. We can observe that in the UFT-dominated CDA market, UFT's profits were heavily grabbed by the IFTs. However, such UFTs' loss is greatly mitigated in the batch auction market with larger batch sizes. On the other hand, the positive profits that each UFT could have when the market structure is close to $n_{\text{IFT}} = 3, n_{\text{UFT}} = 1$ are also reduced under the batch auction mechanism.

Information Advantage

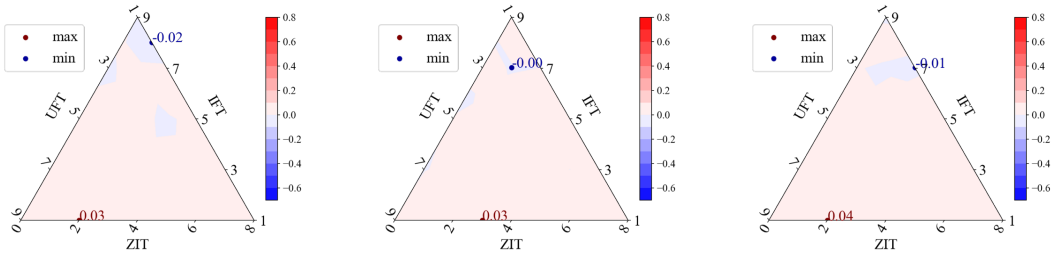
After investigating agents' profits, we found that each agent's profit margin is cut down in the low-frequency batch auction market according to Figs. 6.11 and 6.12, so the information advantage is also reduced as a result.



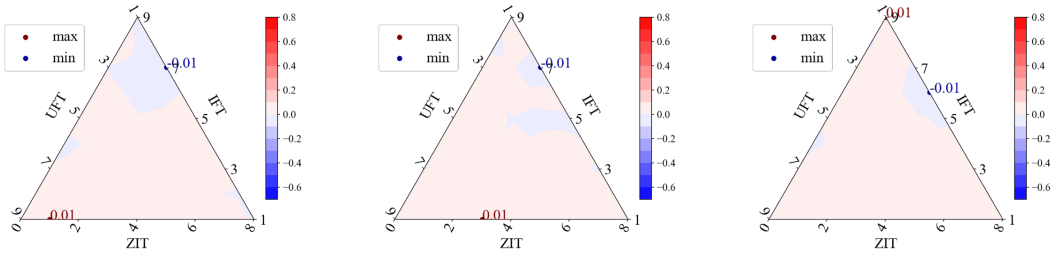
(a) IBR $\sigma_m = 0.05$, CDA (b) IBR $\sigma_m = 0.1$, CDA (c) IBR $\sigma_m = 0.25$, CDA



(d) IBR $\sigma_m = 0.05$, batch size=2 (e) IBR $\sigma_m = 0.1$, batch size=2 (f) IBR $\sigma_m = 0.25$, batch size=2



(g) IBR $\sigma_m = 0.05$, batch size=4 (h) IBR $\sigma_m = 0.1$, batch size=4 (i) IBR $\sigma_m = 0.25$, batch size=4



(j) IBR $\sigma_m = 0.05$, batch size=6 (k) IBR $\sigma_m = 0.1$, batch size=6 (l) IBR $\sigma_m = 0.25$, batch size=6

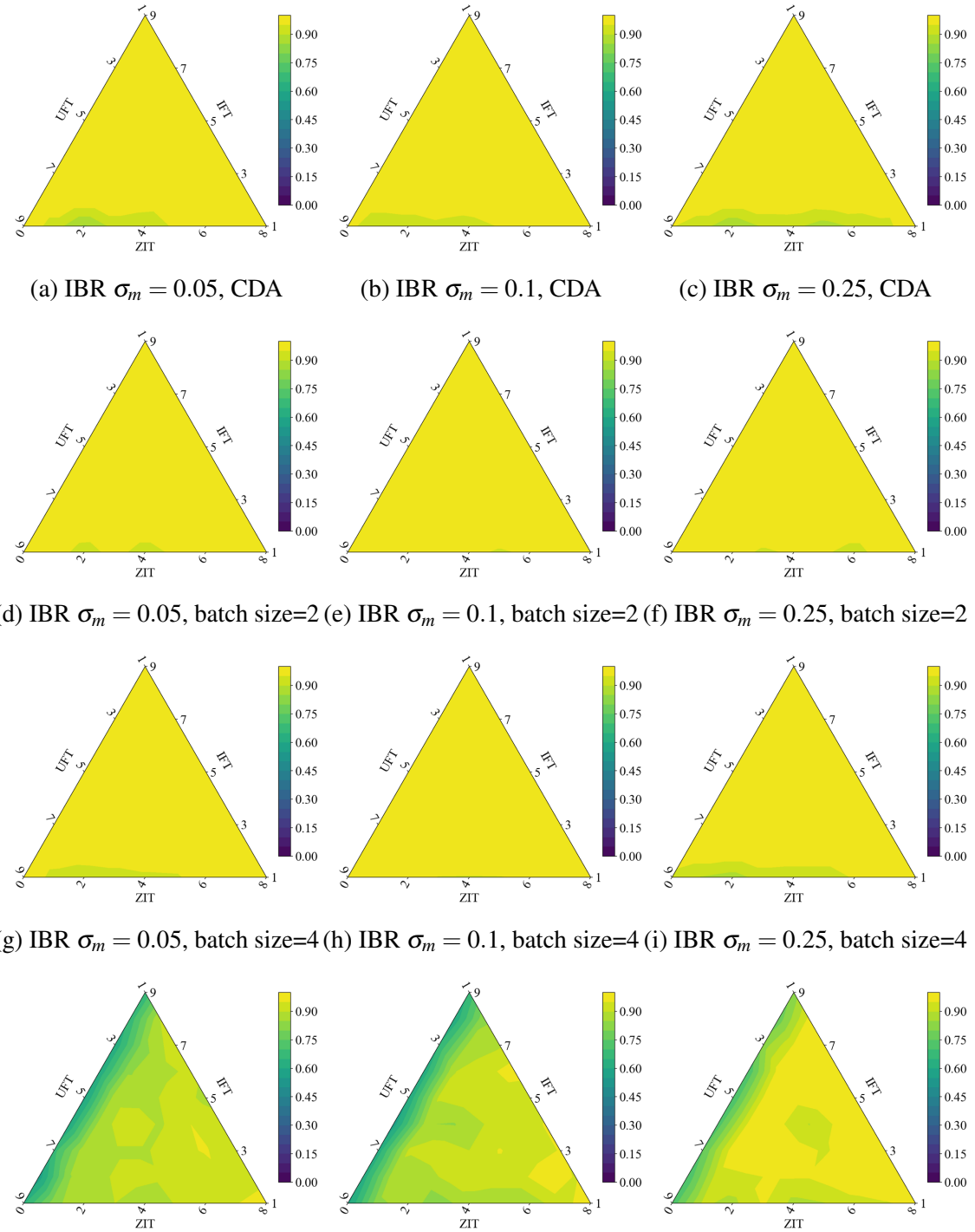
Figure 6.13 Information Advantage of IFTs in the BA Markets When Batch Sizes Differ

The information advantage is significant in the CDA and the batch auction market with batch size 2 in the strongly UFT-dominated market, but the informed fundamental traders have little information advantage as the batch size is larger than 4. In cases where the batch size is equal to 6, there is little difference in information advantage between an informed fundamental trader and an uninformed fundamental trader, on average. The reason for this result may be that, in batch auctions with larger batch sizes, informed fundamental traders

obtain significantly lower maximum profits compared to themselves in CDA markets as Fig. 6.11 shows.

6.4.4 Inference Measurement

Fig. 6.14 shows the true negative rate of uninformed fundamental traders in different batch auction markets with fake shocks. An interesting finding is that, in the case of batch size being 1, 2, or 4, all uninformed fundamental traders can make correct inferences, but it is significantly different in the case of the batch size equal to 6, in which case more uninformed fundamental traders could make incorrect inferences in a less random market (when n_{ZIT} is decreasing). Especially when there is no ZIT trader, only approximately 60% of uninformed fundamental traders can make correct inferences.



(a) IBR $\sigma_m = 0.05$, CDA (b) IBR $\sigma_m = 0.1$, CDA (c) IBR $\sigma_m = 0.25$, CDA

(d) IBR $\sigma_m = 0.05$, batch size=2 (e) IBR $\sigma_m = 0.1$, batch size=2 (f) IBR $\sigma_m = 0.25$, batch size=2

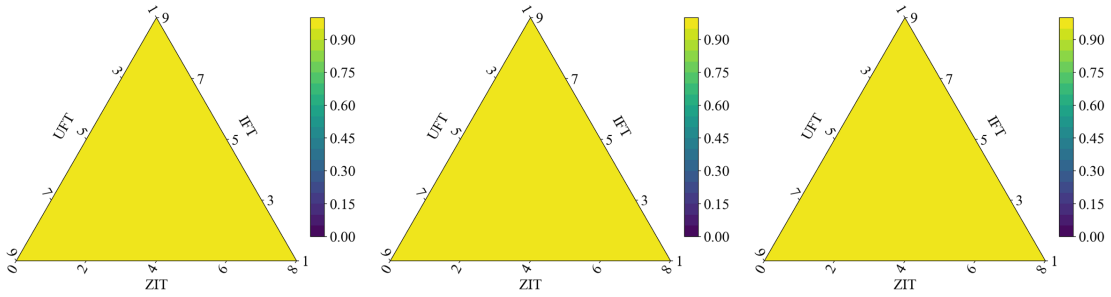
(g) IBR $\sigma_m = 0.05$, batch size=4 (h) IBR $\sigma_m = 0.1$, batch size=4 (i) IBR $\sigma_m = 0.25$, batch size=4

(j) IBR $\sigma_m = 0.05$, batch size=6 (k) IBR $\sigma_m = 0.1$, batch size=6 (l) IBR $\sigma_m = 0.25$, batch size=6

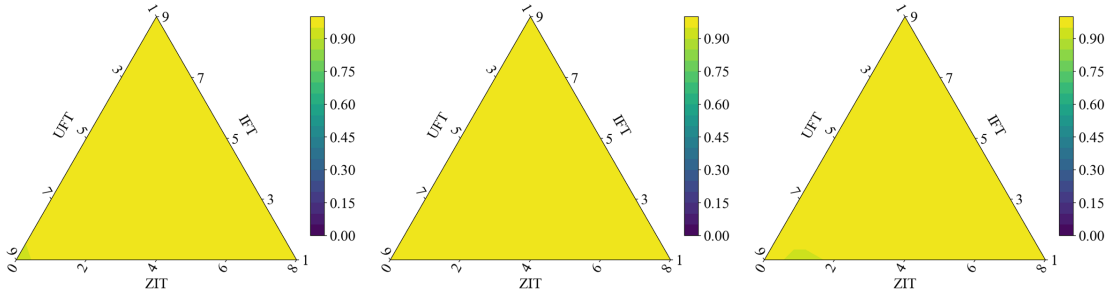
Figure 6.14 The Inference Accuracy of the Uninformed Fundamental Traders in the Messy Network

The inference accuracy shown in Figs. 6.15(d)-(i) are similar to the contour plots in that of CDA markest shown in Figs. 6.15(a)-(c). However, we find that a proper batch auction

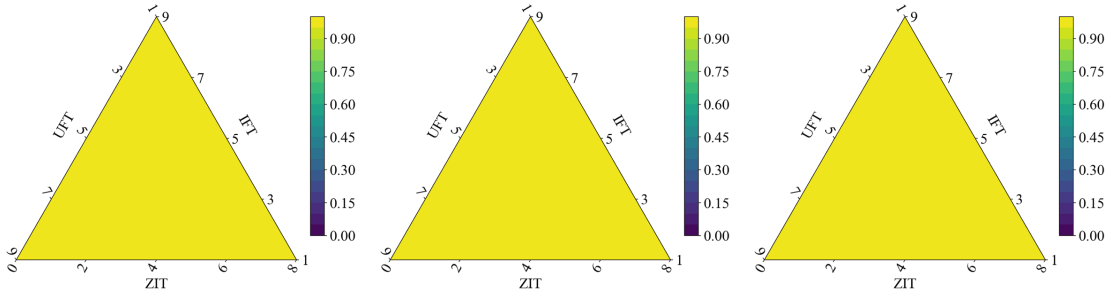
can alleviate the risk of uninformed fundamental traders making incorrect inferences in a strongly UFT-dominated market without asymmetric information. If the batch size is set to be too high (e.g. 6 in this model), the uninformed fundamental traders may make biased inferences.



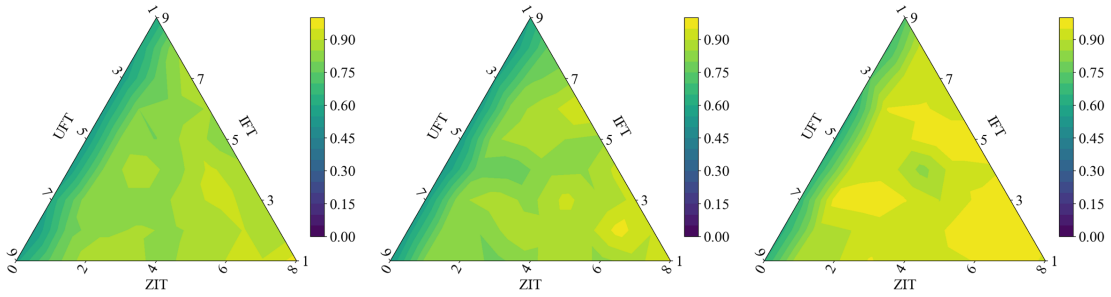
(a) IBR $\sigma_m = 0.05$, CDA (b) IBR $\sigma_m = 0.1$, CDA (c) IBR $\sigma_m = 0.25$, CDA



(d) IBR $\sigma_m = 0.05$, batch size=2 (e) IBR $\sigma_m = 0.1$, batch size=2 (f) IBR $\sigma_m = 0.25$, batch size=2



(g) IBR $\sigma_m = 0.05$, batch size=4 (h) IBR $\sigma_m = 0.1$, batch size=4 (i) IBR $\sigma_m = 0.25$, batch size=4

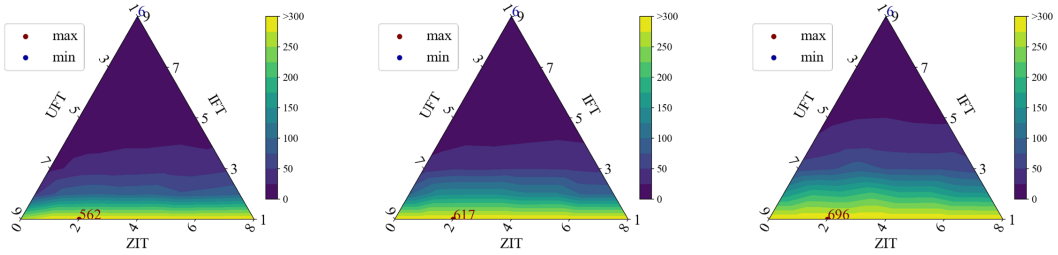


(j) IBR $\sigma_m = 0.05$, batch size=6 (k) IBR $\sigma_m = 0.1$, batch size=6 (l) IBR $\sigma_m = 0.25$, batch size=6

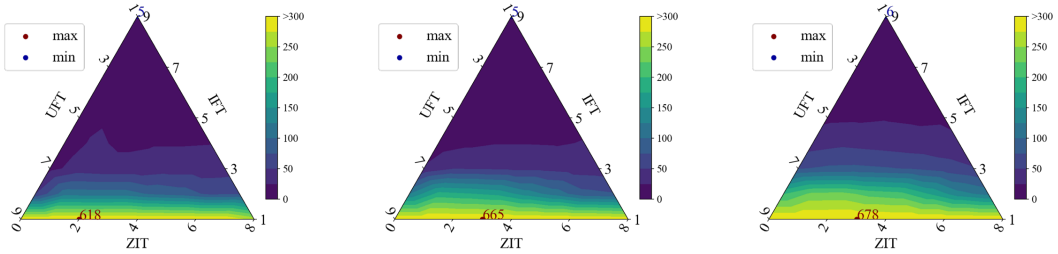
Figure 6.15 Batch Auction: The True Negative Rate of the Uninformed Fundamental Traders

6.4.5 Inference Duration

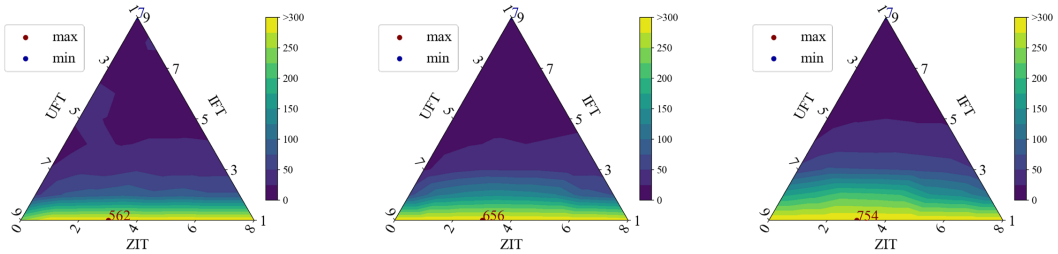
Fig. 6.16 shows the average duration of an uninformed fundamental trader making the inference. However, all contour plots in Fig. 6.16 look similar, so there is little evidence to show that batch auctions can shorten the inference duration.



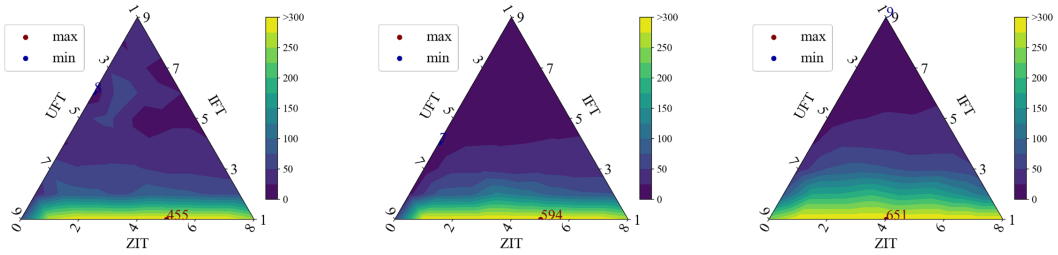
(a) IBR $\sigma_m = 0.05$, CDA (b) IBR $\sigma_m = 0.1$, CDA (c) IBR $\sigma_m = 0.25$, CDA



(d) IBR $\sigma_m = 0.05$, batch size=2 (e) IBR $\sigma_m = 0.1$, batch size= 2 (f) IBR $\sigma_m = 0.25$, batch size=2



(g) IBR $\sigma_m = 0.05$, batch size=4 (h) IBR $\sigma_m = 0.1$, batch size=4 (i) IBR $\sigma_m = 0.25$, batch size=4



(j) IBR $\sigma_m = 0.05$, batch size=6 (k) IBR $\sigma_m = 0.1$, batch size=6 (l) IBR $\sigma_m = 0.25$, batch size=6

Figure 6.16 Batch Auction: Inference Duration of the Uninformed Fundamental Traders

6.4.6 Statistical Testing

In this chapter, we aim to understand how batch auctions affect flash crashes in our model. Again, we apply t-tests and Cohen's d for statistical testing, similar to Section 5.4.4, to explore the differences in crash sizes and UFTs' cumulative profits among different market mechanisms.

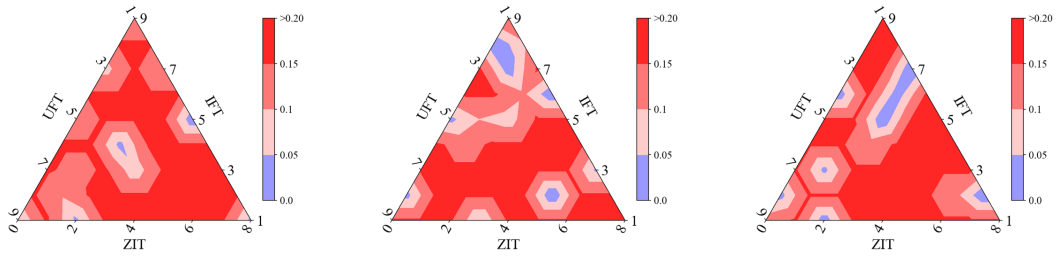
In this section, we set the dataset obtained in the simulations of CDA market (Chapter 5) as Group 1, and the dataset obtained in the simulations of the batch auction market in this chapter as Group 2. The explanations of the statistics p-value and Cohen's d are specified in Section.

Crash Sizes

To do the statistical testing on crash sizes, we proposed the null hypothesis as follows:

Null Hypothesis H_0: *The crash sizes that occurred in the batch auction markets are not different from the crash sizes in the CDA markets.*

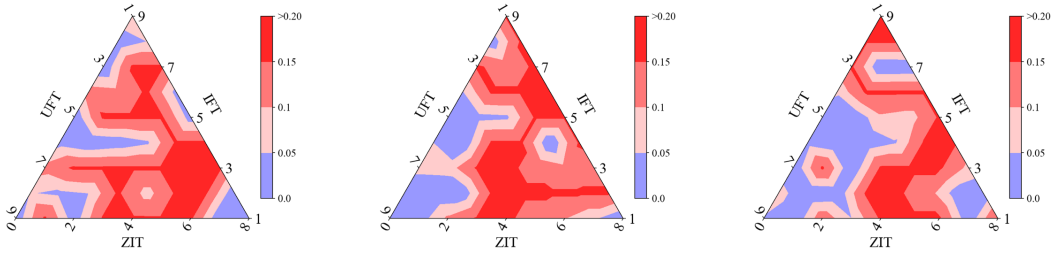
Figs. 6.17, 6.18, and 6.19 depict triangular contour plots displaying the p-values obtained from statistical testing on crash sizes across different market mechanisms. Fig. 6.17 does not present clear evidence to reject the null hypothesis as p-values are large almost everywhere. However, if the batch size increases to 4, there is significance in rejecting the null hypothesis as the market is UFT-dominated and with small number of ZITs, in which case the crash sizes are significantly affected. When the batch size increases to 6, Fig. 6.19 shows a clearer significance in rejecting the null hypothesis.



(a) IBR $\sigma_m = 0.05$ (b) IBR $\sigma_m = 0.1$ (c) IBR $\sigma_m = 0.25$

Figure 6.17 p-values in Crash Sizes between CDA and Batch Auctions (Batch Size 2)

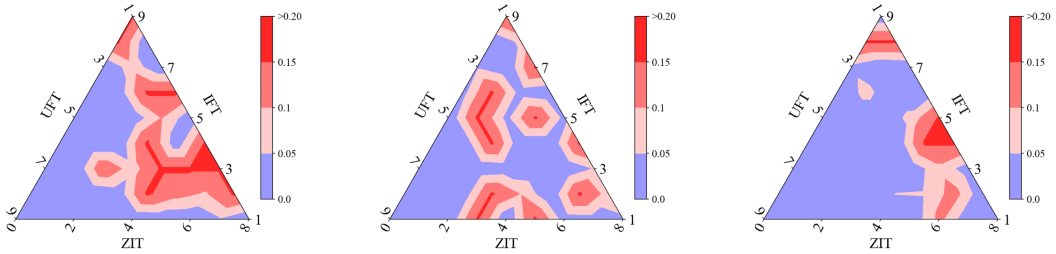
Fig. 6.17 does not provide clear evidence to reject the null hypothesis, as the p-values are large (> 0.05) across the board. However, when the batch size increases to 4, we observe a significant rejection of the null hypothesis. This occurs in UFT-dominated markets with a small number of zero-intelligence traders, where crash sizes are significantly impacted.



(a) IBR $\sigma_m = 0.05$ (b) IBR $\sigma_m = 0.1$ (c) IBR $\sigma_m = 0.25$

Figure 6.18 p-values in Crash Sizes between CDA and Batch Auctions (Batch Size 4)

Further, with the batch size increasing to 6, Fig. 6.19 clearly demonstrates a notable significance in rejecting the null hypothesis.



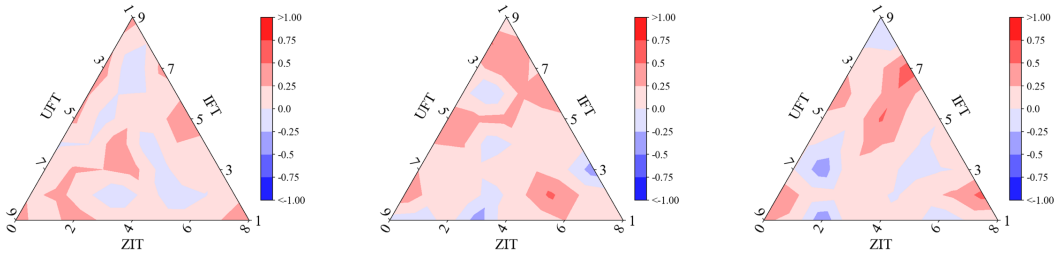
(a) IBR $\sigma_m = 0.05$ (b) IBR $\sigma_m = 0.1$ (c) IBR $\sigma_m = 0.25$

Figure 6.19 p-values in Crash Sizes between CDA and Batch Auctions (Batch Size 6)

In terms of the statistical testing results for Cohen's d, when combined with the t-tests, we can derive three key findings from Figs. 6.20, 6.21, and 6.22:

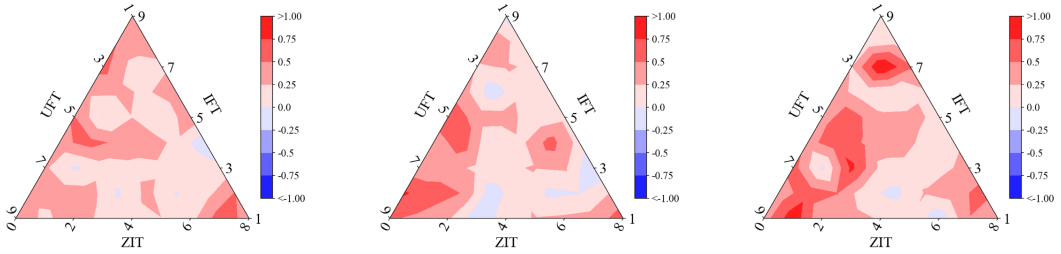
- In most cases, nearly all contour plots appear in red, indicating positive Cohen's d values;
- With an increase in batch size, the deepening colours signify that effect sizes are also increasing;

- In regions where the null hypotheses are rejected, strong evidence demonstrates that batch auctions have the potential to reduce crash sizes (Cohen's d > 1).



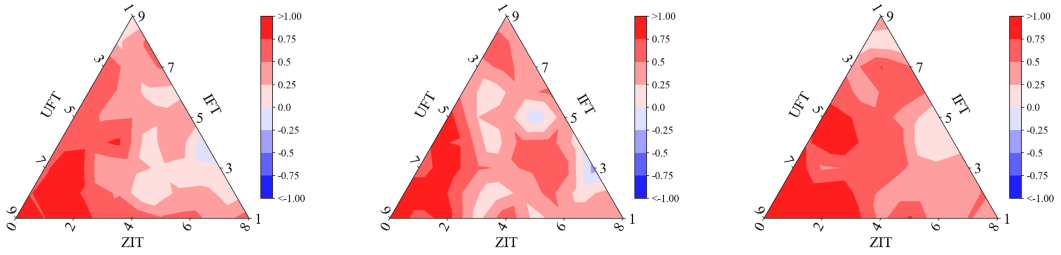
(a) IBR $\sigma_m = 0.05$ (b) IBR $\sigma_m = 0.1$ (c) IBR $\sigma_m = 0.25$

Figure 6.20 Cohen's d in Crash Sizes between CDA and Batch Auctions (Batch Size 2)



(a) IBR $\sigma_m = 0.05$ (b) IBR $\sigma_m = 0.1$ (c) IBR $\sigma_m = 0.25$

Figure 6.21 Cohen's d in Crash Sizes between CDA and Batch Auctions (Batch Size 4)



(a) IBR $\sigma_m = 0.05$ (b) IBR $\sigma_m = 0.1$ (c) IBR $\sigma_m = 0.25$

Figure 6.22 Cohen's d in Crash Sizes between CDA and Batch Auctions (Batch Size 6)

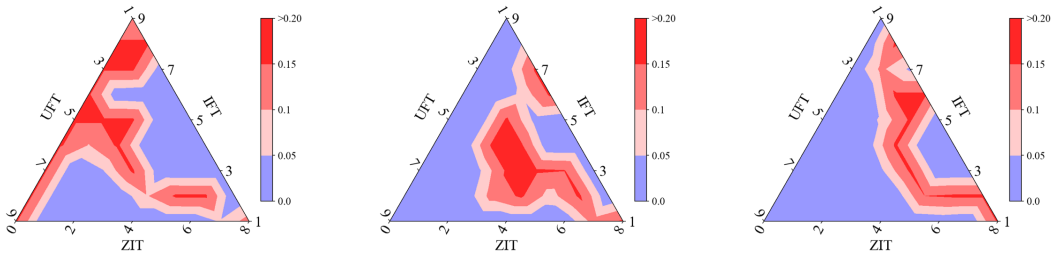
By comparison of the results shown above, we found that batch auctions can, to some extent, reduce flash crashes caused by spoofing, but their impact is significant only under specific market structure configurations. Additionally, larger batch sizes lead to a greater reduction in crash sizes.

UFT's Profits

To do the statistical testing on UFT's profits, we proposed the null hypothesis as follows:

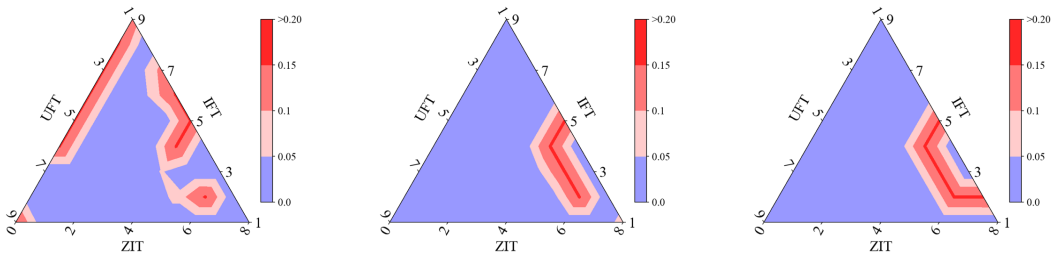
Null Hypothesis H_0: *The UFT's profits/losses in the batch auction markets are not different from those in the CDA markets.*

Figs. 6.23, 6.24 and 6.25 show the triangular contour plots regarding the p-values of statistical testing on UFTs' cumulative profits among different market mechanisms. Fig. 6.23 does not clearly show significance to reject the null hypothesis when the market is IFT-dominated. However, most areas representing the UFT-dominated markets are in blue, implying that the UFT's profits in UFT-dominated batch auction markets are significantly different from those in CDA markets. The p-values become less as the batch size increases.



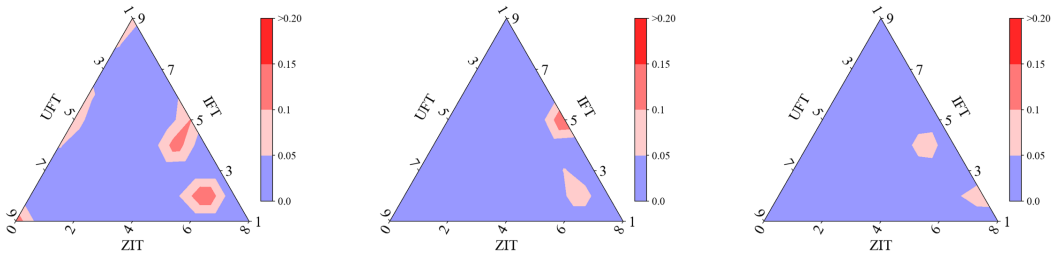
(a) IBR $\sigma_m = 0.05$ (b) IBR $\sigma_m = 0.1$ (c) IBR $\sigma_m = 0.25$

Figure 6.23 p-values in UFT's Profits between CDA and Batch Auctions (Batch Size 2)



(a) IBR $\sigma_m = 0.05$ (b) IBR $\sigma_m = 0.1$ (c) IBR $\sigma_m = 0.25$

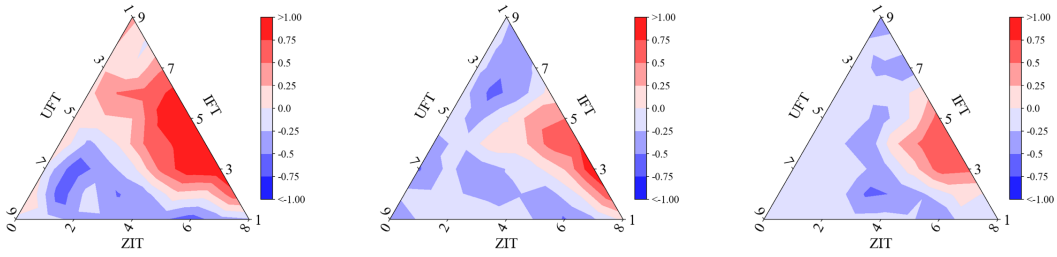
Figure 6.24 p-values in UFT's Profits between CDA and Batch Auctions (Batch Size 4)



(a) IBR $\sigma_m = 0.05$ (b) IBR $\sigma_m = 0.1$ (c) IBR $\sigma_m = 0.25$

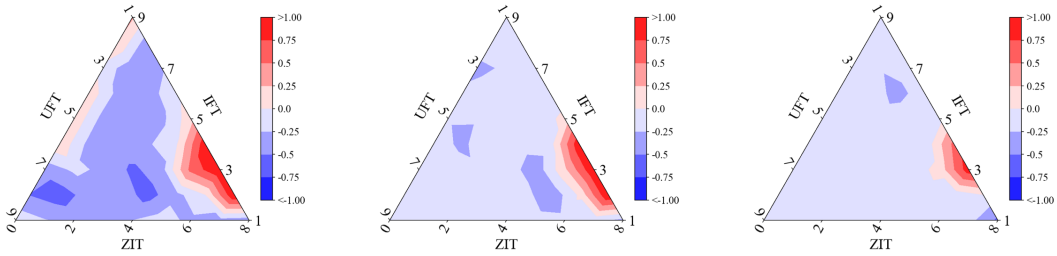
Figure 6.25 p-values in UFT's Profits between CDA and Batch Auctions (Batch Size 6)

Based on the p-value results, our primary focus is on Cohen's d values within UFT-dominated markets, as illustrated in Figs. 6.26, 6.27, and 6.28. One significant finding is that these areas display negative Cohen's d values, indicating increased profits or mitigated losses in batch auctions. Another notable observation is that the effect size appears to be more substantial when the market has low volatility.



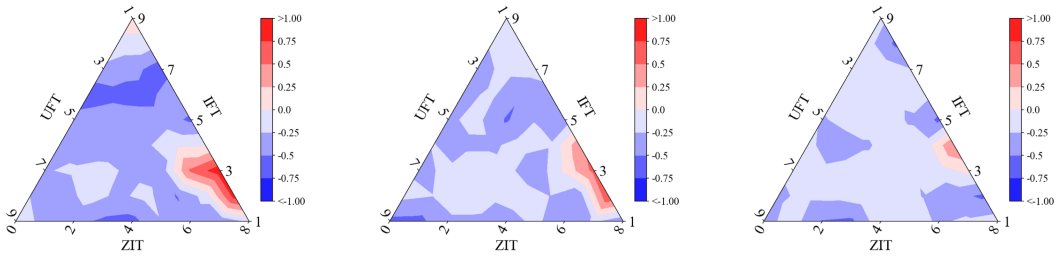
(a) IBR $\sigma_m = 0.05$ (b) IBR $\sigma_m = 0.1$ (c) IBR $\sigma_m = 0.25$

Figure 6.26 Cohen's d in UFT's Profits between CDA and Batch Auctions (Batch Size 2)



(a) IBR $\sigma_m = 0.05$ (b) IBR $\sigma_m = 0.1$ (c) IBR $\sigma_m = 0.25$

Figure 6.27 Cohen's d in UFT's Profits between CDA and Batch Auctions (Batch Size 4)



(a) IBR $\sigma_m = 0.05$ (b) IBR $\sigma_m = 0.1$ (c) IBR $\sigma_m = 0.25$

Figure 6.28 Cohen's d in UFT's Profits between CDA and Batch Auctions (Batch Size 6)

6.5 Conclusion

In this chapter, we introduced batch auctions and extended the flash crash model of information asymmetry to the circumstances of different market mechanisms, where, more specifically, batch auctions and continuous auctions are examined. Here are our findings in this chapter:

- We find that in markets utilizing batch auctions, as the frequency of batching decreases (corresponding to larger batch sizes), the price crashes caused by fake shocks in the market slow down, resulting in weaker flash crashes, particularly in UFT-dominated markets (Fig. 6.6, Figs. 6.17-6.19). However, the decrease in crash speed is not significantly different due to changes in the mechanisms Fig. 6.9.

- Nonetheless, batch auctions prove to be more effective in mitigating the impact of fake shocks in lower-risk markets compared to higher-risk markets, especially in terms of crash sizes (Figs. 6.20-6.22). Additionally, UFTs' profits are also mitigated in batch auctions when compared to CDA markets.

- Due to the batch auction mechanism itself, as shown in Fig. 6.5 and Figs. 6.11-6.12 the volatile market prices in the continuous auction market can be smoothed, potentially reducing the maximum possible profit margin for all traders. As shown in Fig. 6.11, when the batch size equals 4, the profit margin of a single IFT can be cut by more than 90%, while when the batch size is 6, all traders have almost no profit margin, and, in this case, the information advantage of the IFT is almost zero.

- Madhavan (1992) claimed that batch auctions reduce information asymmetry since all traders are given access to the same prices at the same time. In this chapter, we obtained a similar result, showing that the information advantage of each IFT is significantly reduced as the batch size increases, and little information advantage exists if the batch size is larger than 4 (Fig. 6.13).

- As Fig. 6.15 shows, setting a high value of batch size has a negative effect on each uninformed fundamental for making accurate inferences. The results show that the accuracy of a UFT making correct inferences decreases dramatically when the batch size increases from 4 to 6.

In summary, we find that batch auctions can, to some extent, reduce the impact of the fake shock on market prices in markets with asymmetric information. However, over-low-frequency batch auctions not only do not substantially alleviate flash crashes but also reduce the profit margins of all traders and thus may affect traders' motivation to trade. In addition, over-low-frequency batch auctions may affect the decision-making accuracy of uninformed fundamental traders.

Chapter 7

Flash Crash in Different Networks

7.1 Introduction

In Chapter 5, we explored the extension of intelligent uninformed fundamental traders into the flash crash model. Intelligent UFTs consider two opposing hypotheses about the authenticity of the shock and determine which one they believe. In the simulation, UFTs update their beliefs by observing the past order flows.

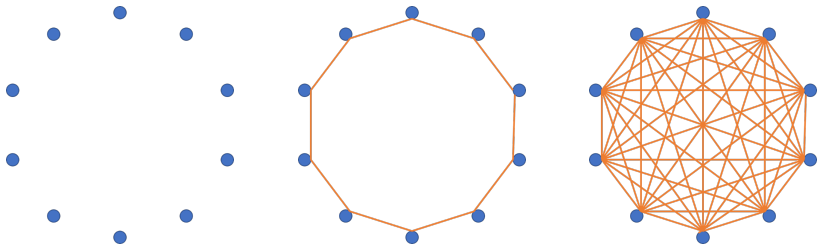
However, the research reported in Chapter 5 is based on a completely messy network; namely, traders act independently based on the public information only. In fact, in real-world markets, information sharing among traders is quite possible and realistic. Goldstein et al. (2021) proposes a theory to study the motivations and consequences of information sharing among informed investors and uninformed or imperfectly informed investors. They concluded that uninformed investors might voluntarily share information with informed investors. In contrast, informed ones do not share their information because information sharing would reduce their information advantage and likely cause their profits to shrink.

Two economists, Bala and Goyal (1998) present a framework where agents can update their beliefs on different states by iteratively learning from observing actions or signals. Inspired by Bala and Goyal (1998), Zollman (2007, 2010) generalise the model by employing bandit problems, where the gamblers aim to maximise their payoffs when confronted with many bandits with different success probabilities. They also investigate the scientist's problems: a group of scientists are confronted with two opposing hypotheses and try to

find out the best one. In the simulation, the scientists update their beliefs by observing the actions of agents which are connected with each other in a social network. To investigate whether building an information-sharing mechanism can mitigate flash crashes and reduce the losses of the traders who are spoofed, in this chapter, referring to Zollman's model (Zollman, 2007, 2010), we build connections for the agent in the flash crash model in Chapter 5 by constructing two additional networks: a cycle network and a complete network shown in Fig. 7.1.

7.2 Social Networks

Zollman (2007, 2010) proposed three different networks, messy, cycle, and complete networks, to study the effects of network topology in an agent-based model. Fig. 7.1 specifies how the connections work in different networks. The blue nodes represent the agents, and two nodes with a connection of the orange line can share some information. In a messy network, nobody shares their own information with others, while each agent only shares information with their two neighbors in the cycle network. However, in the complete network, all shareable information is released to the public since every node connects with all other nodes.



(a) Messy Network (b) Cycle Network (c) Complete Network

Figure 7.1 Types of Social Networks

Assume that trader x^* is selected and matches with the market maker to trade at round t, and the trade information is broadcast to the public as soon as the trade is made. We denote the action of the trader x^* as $\mathbb{A}_{x^*}$. The differences among the three networks are regarding the information that an uninformed fundamental trader obtained each round. As Fig. 7.1 shows, the blue nodes represent the trading agents, and the orange line linking

two nodes means the two linked agents both know the actions from each other. Since the trade information is broadcast publicly, every trader knows $\mathbb{A}_{x^*}$ no matter what network they are linked as. However, if the winning trader x^* is trader number i, it means the trade information does not provide extra information for their learning. The following illustrates in detail the information differences among the three networks:

- Messy: $I_{i,t} = \begin{cases} \{\mathbb{A}_{x_t^*}\} & \text{if } x^* \neq x_i \\ \emptyset & \text{if } x^* = x_i \end{cases}$

In the messy network, the agent x_i will be notified of the trade order's information after each trading round. If the i-th trader himself did not place the order, this trader could update the beliefs on different states based on the order direction. However, if the i-th trader made a trade in this trading round, there is no additional information for the trader to update beliefs representing $\emptyset$.

- Cycle: $I_{i,t} = \begin{cases} \{\mathbb{A}_{x_t^*}, \mathbb{A}_{x_{i-1},t}, \mathbb{A}_{x_{i+1},t}\} & \text{if } x^* \neq x_i \\ \mathbb{A}_{x_{i-1},t}, \mathbb{A}_{x_{i+1},t} & \text{if } x^* = x_i \end{cases}$

In the cycle network, x_i will be notified of the trade order's information and the neighbours' actions after each trading round. However, x_i is not told what types of the neighbours are, but the additional information from the neighbours helps the trader better recognise not only the shock but the neighbours' identities. Appendix F shows the full details of how an uninformed trader learns to understand the neighbours and the information in the cycle network. Even x_i made a trade in the current trading round, the neighbours' information still allows the trader to conduct the inference.

- Complete: $I_{i,t} = \begin{cases} \bigcup_{j \neq i}\{\mathbb{A}_{x_j,t}\} \bigcup \{\mathbb{A}_{x_t^*}\} & \text{if } x^* \neq x_i \\ \bigcup_{j \neq i}\{\mathbb{A}_{x_j,t}\} & \text{if } x^* = x_i \end{cases}$

In the complete network, every two nodes have a connection, meaning all traders' actions are public after each trading round. Trader x_i has sufficient information and can take advantage of this information to learn the shock and estimate what kind of other traders are. The Appendix G illustrates the full work regarding how an intelligent uninformed trader acts in the complete network.

7.3 Experiments

This chapter still runs the short-term experiments where all trading activities are involved in a single trading epoch, and all basic settings remain the same as the study of the CDA market in Chapters 4 and 5.

7.3.1 Parameters

In this chapter, the key parameter is Topology Parameters, **Network type ($\mathscr{T}$):**, and informational (σ_{ε_n}) and the market-structure parameters (n_{IFT}) and (n_{UFT}) are still considered and remain unchanged as introduced in Section 4.1.2.

Topology Parameters: Networks ($\mathscr{T}$): there are three kinds of networks, messy network(T_1), cycle network(T_2) and complete network(T_3). The details of the network are specified in Appendices F and G.

7.3.2 Performance Metrics

The important metrics proposed in Chapter 6 are also investigated in this chapter. As this chapter introduced intelligent uninformed fundamental traders, the market prices could recover after the flash crash. Therefore, we will investigate a few more metrics regarding such more complicated models.

Metrics	Explanation
Crash size	The biggest price decline of the market prices after the fake shock and before all UFTs finish inference
Crash duration	The duration between the shock point and the one reaching the biggest price decline
Recovery duration	The duration between the point reaching the biggest price decline and the turning point of a new steady state
Cumulative profits	Average cumulative profits of different types of trading agents
Information advantage	A quantitative indicator measuring the information advantage of a single IFT to a single UFT on average
Inference accuracy	Some indicators to measure how accurate a single UFT can make a correct inference

Table 7.1 Performance Metrics

7.4 Experimental Results

7.4.1 Setups

This section studies the experiments of the flash crash model in different networks. In the messy network, all traders are modelled to be completely independent, which means they do not know others' actions during trading. Different from the messy network, the traders in the cycle network have and only have two connected traders. The connected traders share their actions with each other when submitting orders. In the complete network, all traders are connected with each other, so every trader has information on other traders' actions. Most setups remain consistent with the basic model specified in Chapter 5.

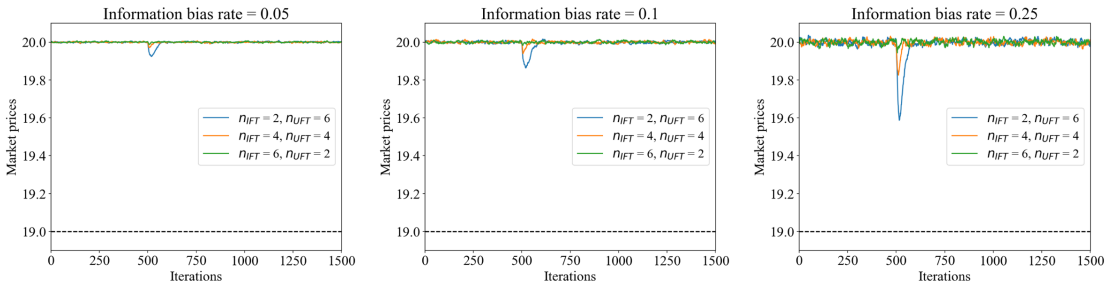
7.4.2 Global Statistics Overview

At the beginning of analysing the experimental results, we focus on some global statistics on price movements and agents' profits as an overview. Similar to the setups in Chapter 4, we consider three distinct market structures: an IFT-dominated market, a UFT-dominated market, and an information-balanced market. These market structures maintain identical configurations to the 10-agent system described in Chapter 4.

Prices Movements

If the shock is fake Fig. 5.4 shows the average market prices among all simulations under the given three parameter settings with the comparison among different information bias rates in the **messy network**.

In the **cycle network**, the price movement patterns slightly differ from the messy network's circumstances. Compared to the messy network patterns, the crash size and duration are apparently reduced. According to Fig. 7.2, the market price under each parameter setting experiences flash crashes when the shock comes, also followed by an almost full price recovery. Moreover, similar to the messy networks, the crash sizes differ among different parameter settings, in which the crash size in the case of $n_{\text{IFT}} = 2, n_{\text{UFT}} = 6$ is larger than the case of $n_{\text{IFT}} = 4, n_{\text{UFT}} = 4$. When $n_{\text{IFT}} = 6, n_{\text{UFT}} = 2$, the market price has the smallest crash size among the three market-structure settings.

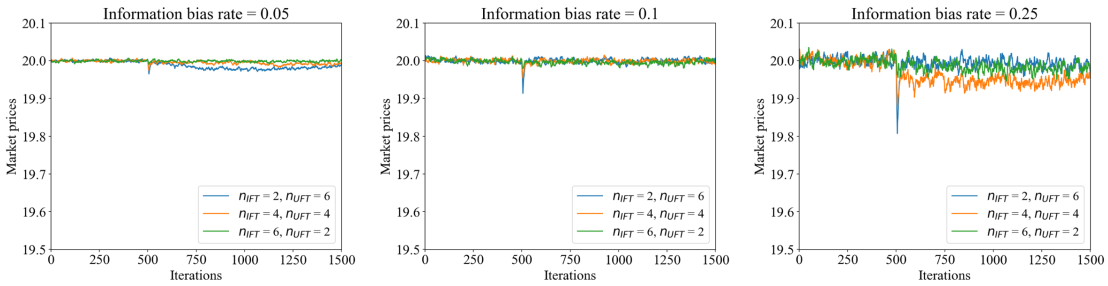


(a) IBR $\sigma_{\varepsilon_n} = 0.05$ (b) IBR $\sigma_{\varepsilon_n} = 0.1$ (c) IBR $\sigma_{\varepsilon_n} = 0.25$

Figure 7.2 Cycle Network: The Average Market Prices Movements (Fake Shock)

In the **complete network**, the price movements among different sample simulations are less different from those in the cycle network. Even though a price decline occurs right

after the shock point, it is very difficult to observe the price declines and rebounds from the perspective of the average market prices. It is not surprising to obtain such preliminary results because the uninformed traders connected more agents can acquire more information for the inference leading to a quicker decision after the inference. It is worth noting that when the information bias rate is as high as IBR = 0.25, it is more apparent that the price crashes more deeply and lasts longer in the case of $n_{\text{IFT}} = 2, n_{\text{UFT}} = 6$. The feature that market price drops more greatly if the uninformed fundamental traders dominate is similar to the cases in the messy network and complete network.



(a) IBR $\sigma_{\varepsilon_n} = 0.05$ (b) IBR $\sigma_{\varepsilon_n} = 0.1$ (c) IBR $\sigma_{\varepsilon_n} = 0.25$

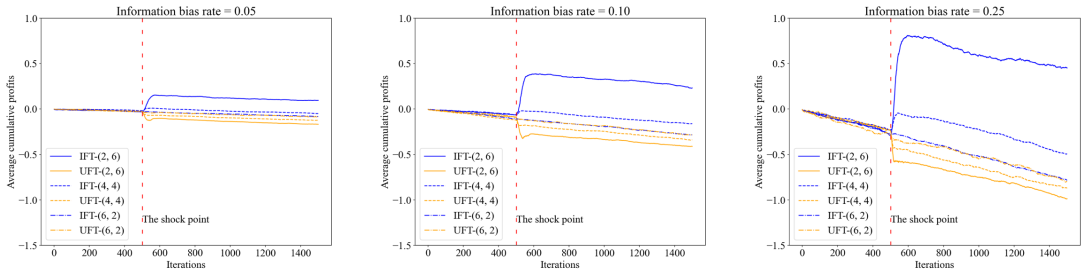
Figure 7.3 Complete Network: The Average Market Prices Movements (Fake Shock)

However, we found that the market price curves in the complete network do not always converge to a level close to the true stock price. It implies that some uninformed fundamental traders may have made incorrect inferences in the complete network.

Profits of Agents

The average cumulative profits graphs show how the different kinds of agents can accumulate profits on average as the iteration goes on.

Cycle Network We next take a preliminary look at the profit changes of the agents in the cycle network. Fig. 7.4 shows the average cumulative profit of fundamental traders under different information bias rate settings. Similar to the messy network, the profit dynamics of different types of fundamental traders also diverge after the shock point.

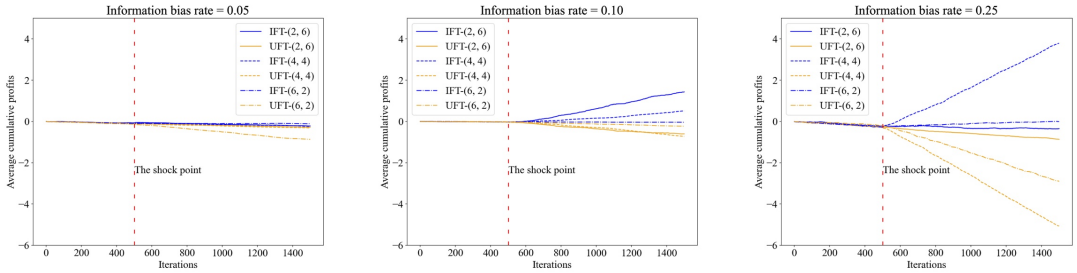


(a) n_{IFT}=2, n_{UFT}=6, IBR=0.05 (b) n_{IFT}=2, n_{UFT}=6, IBR=0.1 (c) n_{IFT}=2, n_{UFT}=6, IBR=0.25

Figure 7.4 Cycle Network: Fundamental Traders' Average Cumulative Profits in the Fake-shock Simulations

From the perspective of the divergence size of the two kinds of fundamental traders' profits curves, compared with the messy network, the narrower divergence in the cycle network implies that the informed fundamental traders grab fewer profits from the uninformed fundamental traders under the fake shock.

Complete Network Fig. 7.5 presents the average cumulative curves of the fundamental traders in the fake-shock market under the complete network. The figures also present the bifurcation of the profit curves as shown in the messy and cycle networks. However, even with higher complexity, the profit bifurcation size in the complete network is relatively larger than that of the same parameters in the cycle network. In the subsequent analysis, we will find that the inference accuracy of UFTs in the complete network is not as high as that in the cycle network. Therefore, these preliminary results may indicate that some UFTs have not successfully adjusted their strategies through learning to prevent losses. Consequently, on average, the information advantage of IFTs over UFTs in the complete network may be greater than in the cycle network.



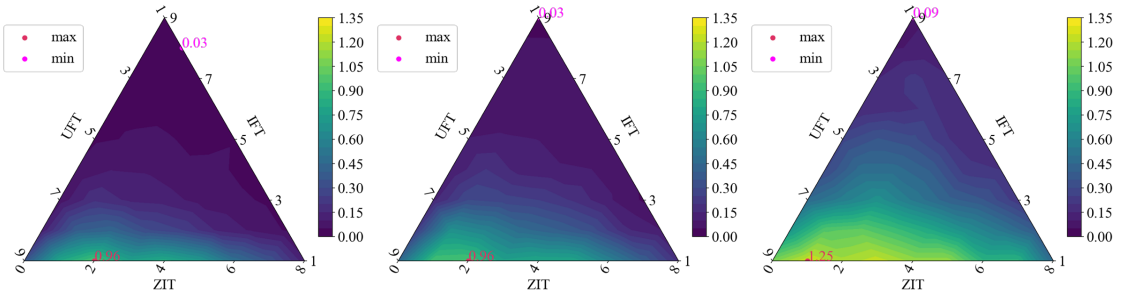
(a) n_{IFT}=2, n_{UFT}=6, IBR=0.05 (b) n_{IFT}=2, n_{UFT}=6, IBR=0.1 (c) n_{IFT}=2, n_{UFT}=6, IBR=0.25

Figure 7.5 Complete Network: Fundamental Traders' Average Cumulative Profits in the Fake-shock Simulations

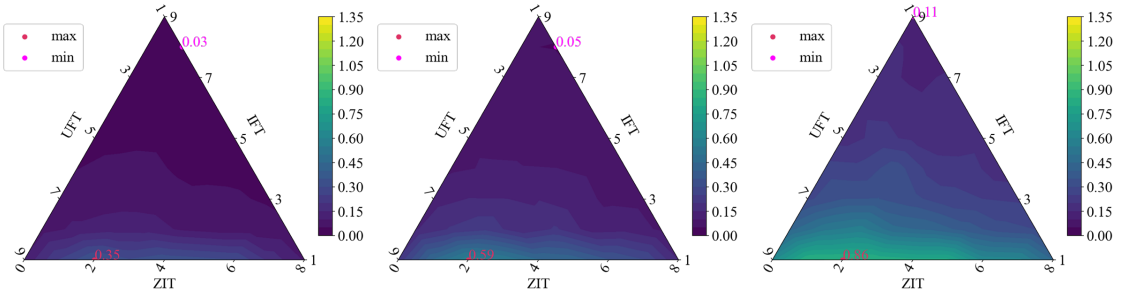
7.4.3 Sensitivity Analysis with Fake Shocks

Crash Size

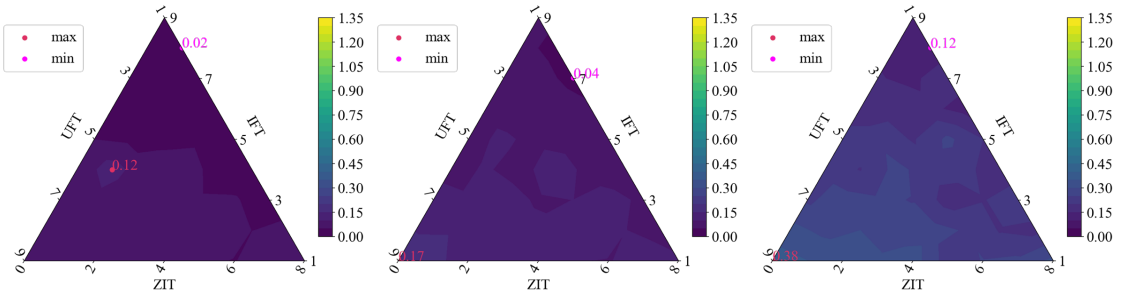
We have shown that the crash sizes are large in a highly UFT-dominated CDA market from Fig. 7.2 and Fig. 7.3, and the market price suffers the largest crash on average in the case of $n_{\text{IFT}} = 1, n_{\text{UFT}} = 7, n_{\text{ZIT}} = 2$.



(a) IBR $\sigma_{\varepsilon_n} = 0.05$, messy network (b) IBR $\sigma_{\varepsilon_n} = 0.1$, messy network (c) IBR $\sigma_{\varepsilon_n} = 0.25$, messy network



(d) IBR $\sigma_{\varepsilon_n} = 0.05$, cycle network (e) IBR $\sigma_{\varepsilon_n} = 0.1$, cycle network (f) IBR $\sigma_{\varepsilon_n} = 0.25$, cycle network



(g) IBR $\sigma_{\varepsilon_n} = 0.05$, complete network (h) IBR $\sigma_{\varepsilon_n} = 0.1$, complete network (i) IBR $\sigma_{\varepsilon_n} = 0.25$, complete network

Figure 7.6 Crash Sizes (mean) in Different Networks

As Fig. 7.6 shows, under the cycle network, we can clearly find that the overall level of crash size is lower than that in the messy network. However, the pattern related to the market structure is still similar to that in the messy network—when the market is highly UFT-dominated, the crash size has a relatively large value. Especially in the case of n_{IFT}=1, n_{UFT}=7, and n_{ZIT}=2, the crash sizes reach the highest level among all systematic risk settings. Compared to the messy network, where market prices experience almost full crashes in the case of highly UFT-dominated markets, the average maximum crash in

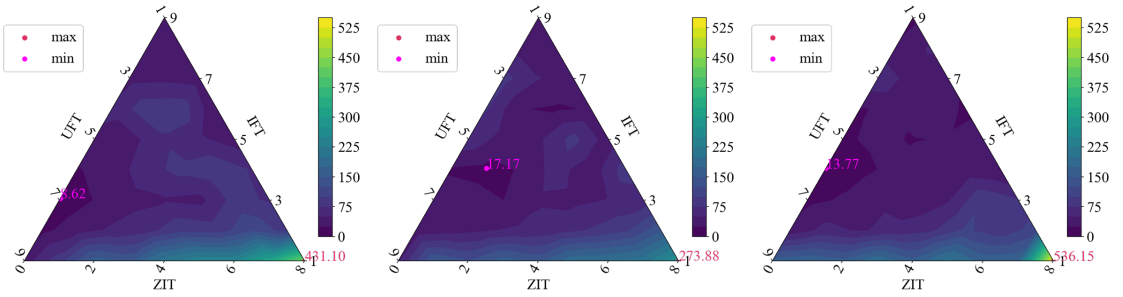
the cycle network is only around 0.4 to 0.5, which means the size of the crashes has been reduced considerably in the cycle network.

In contrast, the complete network greatly mitigates the impact of a flash crash. When fake shock occurs, under different market structures and systemic risks, the crash sizes are all at a very low level - the average maximum drop is no more than 0.2 in all cases specified. However, it is still clear, from the contour figures, that the UFT-dominated markets still have weak price crashes and that the crash sizes still have the highest levels, especially when the systematic risk level is at 0.1 and 0.25.

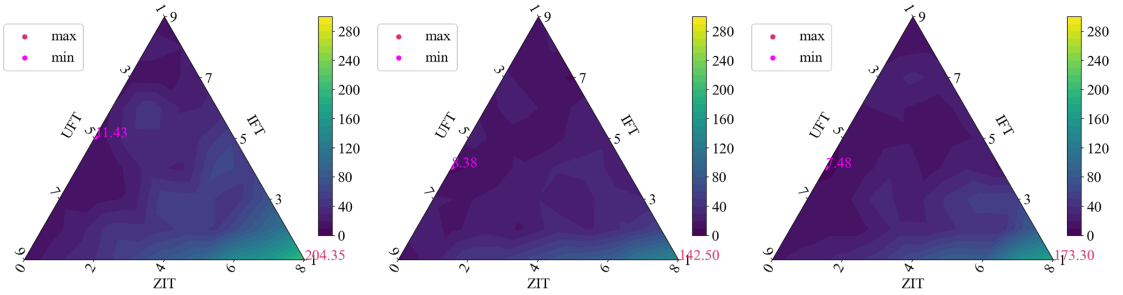
Crash and Recovery

In addition to studying the crash size to analyze the market dynamics after the fake shock, we will also examine the crash duration and recovery duration during the price variation phase.

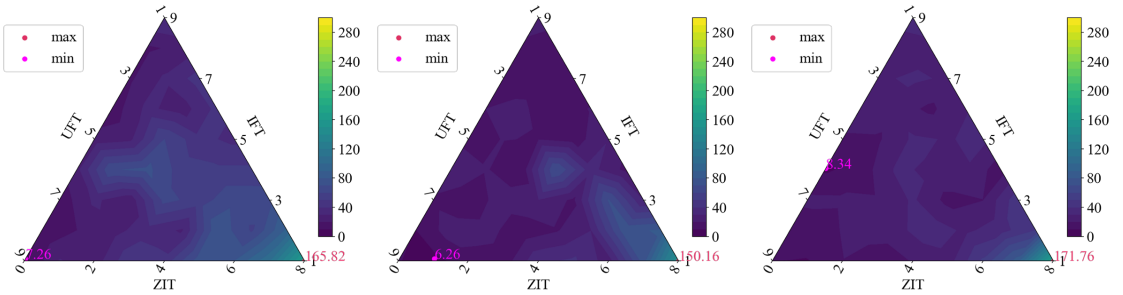
We have studied the patterns of crash and recovery duration under the messy CDA market in Chapter 5. The crash and recovery durations in the messy network are only clearly long if $n_{\text{IFT}} = 1$, while the crash and recovery durations are lower as the market structure is otherwise.



(a) IBR $\sigma_{\varepsilon_n} = 0.05$, messy network (b) IBR $\sigma_{\varepsilon_n} = 0.1$, messy network (c) IBR $\sigma_{\varepsilon_n} = 0.25$, messy network



(d) IBR $\sigma_{\varepsilon_n} = 0.05$, cycle network (e) IBR $\sigma_{\varepsilon_n} = 0.1$, cycle network (f) IBR $\sigma_{\varepsilon_n} = 0.25$, cycle network



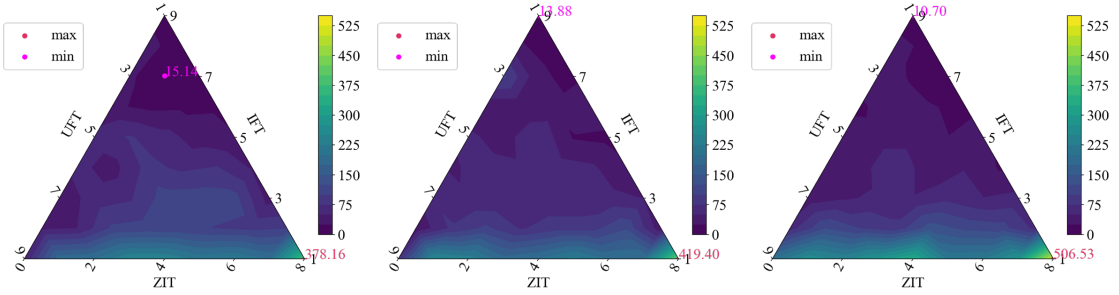
(g) IBR $\sigma_{\varepsilon_n} = 0.05$, complete network (h) IBR $\sigma_{\varepsilon_n} = 0.1$, complete network (i) IBR $\sigma_{\varepsilon_n} = 0.25$, complete network

Figure 7.7 Crash Durations (mean) in Different Networks

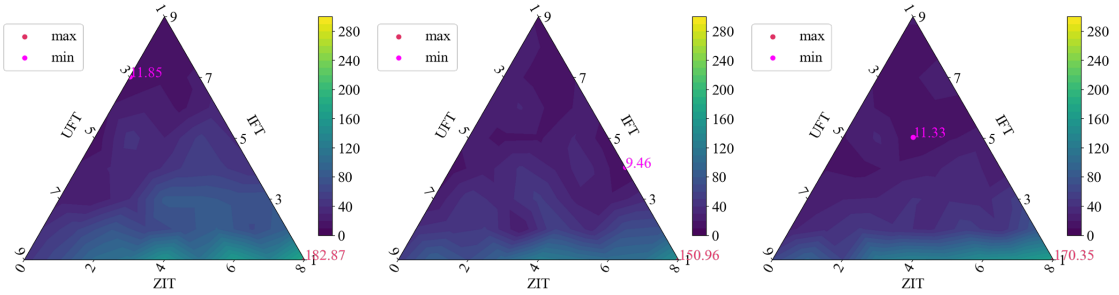
The crash duration and recovery duration pattern in the cycle network differs from that in the messy network. In the cycle network, simulations with the longer average crash and recovery durations correspond to markets with a small number of IFTs, a small number of UFTs and a large number of ZITs, which are defined as random markets.

Our previous study on crash sizes found that the cycle network could reduce the crash size compared to the messy network. However, compared to the crash sizes, the reduction in crash and recovery durations is not as remarkable as that of the crash size. It is particularly

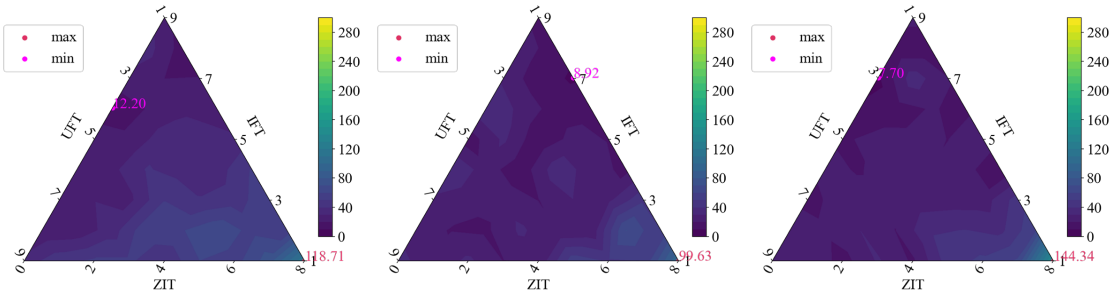
noteworthy that, unlike crash size, which increases monotonically with systemic risk, in the study of crash and recovery duration, the simulations under the information bias rate 0.1 have lower maximum values compared to the cases IBR $= 0.05$ and 0.25 according to the triangular contours.



(a) IBR $\sigma_{\varepsilon_n} = 0.05$, messy network (b) IBR $\sigma_{\varepsilon_n} = 0.1$, messy network (c) IBR $\sigma_{\varepsilon_n} = 0.25$, messy network



(d) IBR $\sigma_{\varepsilon_n} = 0.05$, cycle network (e) IBR $\sigma_{\varepsilon_n} = 0.1$, cycle network (f) IBR $\sigma_{\varepsilon_n} = 0.25$, cycle network



(g) IBR $\sigma_{\varepsilon_n} = 0.05$, complete network (h) IBR $\sigma_{\varepsilon_n} = 0.1$, complete network (i) IBR $\sigma_{\varepsilon_n} = 0.25$, complete network

Figure 7.8 Recovery Durations (mean) in Different Networks

Furthermore, a large crash size does not mean a longer crash duration for a given information bias rate. As an example, Fig. 7.7 shows how the simulated average crash size and duration change as the number of UFTs varies, given the number of IFTs is fixed at

1. The crash size tends to rise and then fall as the number of UFTs increases, reaching its highest point at roughly $n_{\text{IFT}} = 7$. In contrast, the crash duration shows a clear monotonic downward trend as the number of UFTs increases.

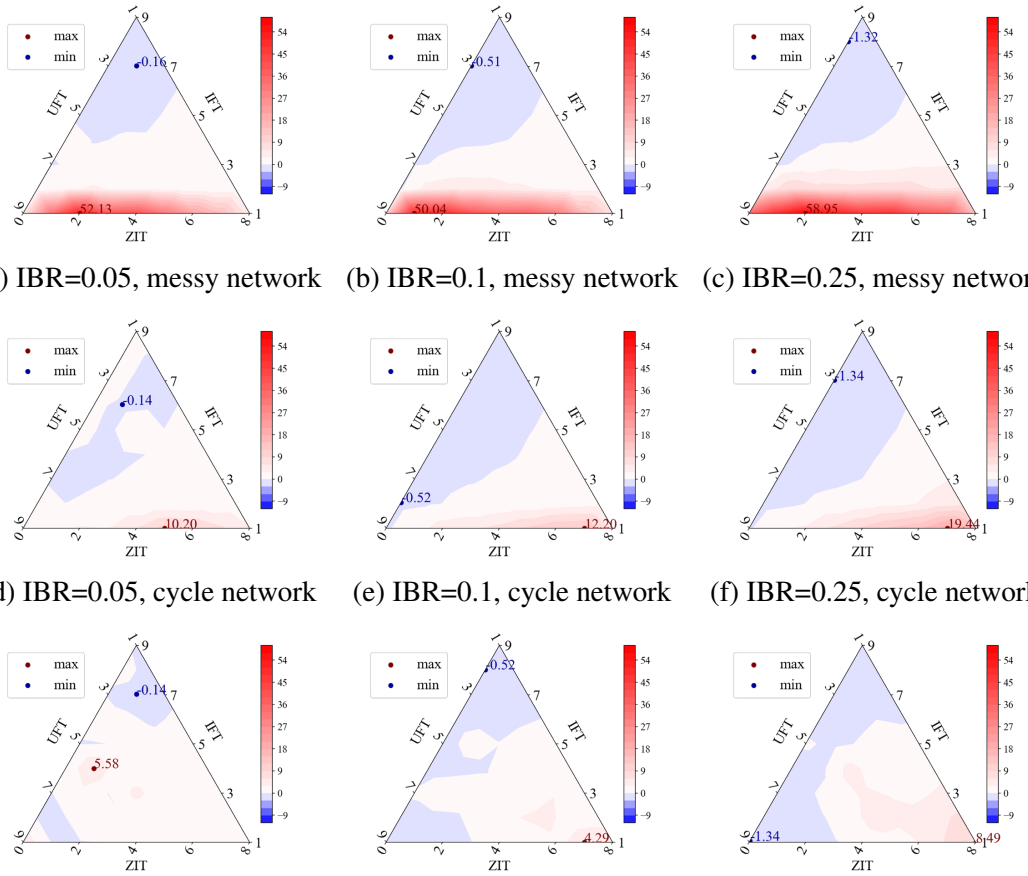
In the complete network, accompanied by the reduction in the crash size of the market prices, the crash and recovery durations are also reduced greatly overall compared to the other two networks. Although we can see in Fig. 7.7 and 7.8 that when the numbers of IFTs and UFTs are set to be small, the crash and recovery sizes are at high levels. In fact, such high values are not meaningful because it is difficult to determine the crash phase and recovery phase when the flash crash size is very small. As a result, the crash duration calculated by a given formula is subject to a large bias.

Agents' Average Profits

By examining the market regarding the market price crashes, we know that the different network clearly influences the changes in market prices and dynamics. In addition, the greater information exposure resulting from a more complex topology network can also reduce the impact of flash crashes. In this section, we focus on each agent by calculating their profit distribution to investigate the impact of the fake shock on the individual agents.

Each contour graph represents the cumulative profits of the agent in each simulation on average. Positive returns are shown in red and negative returns are in blue.

Informed Fundamental Traders In the study of the CDA market, we have found that if the market is strongly UFT-dominated, especially when the number of IFTs is 1, each IFT can grab significant profits from the UFTs. However, as the number of IFTs increases and the number of UFTs decreases, the average profit per IFT decreases rapidly.



(a) IBR=0.05, messy network (b) IBR=0.1, messy network (c) IBR=0.25, messy network

(d) IBR=0.05, cycle network (e) IBR=0.1, cycle network (f) IBR=0.25, cycle network

(g) IBR=0.05, complete network (h) IBR=0.1, complete network (i) IBR=0.25, complete network

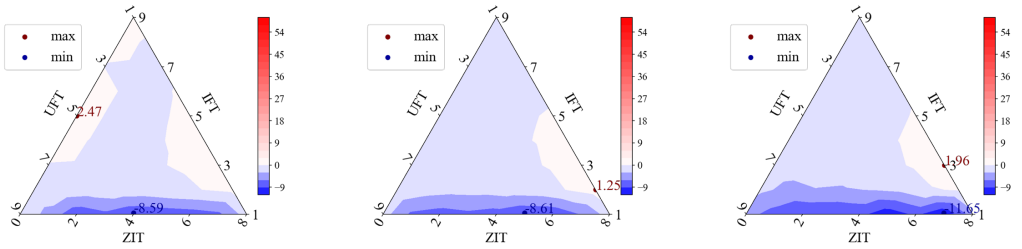
Figure 7.9 Agents' Average Cumulative Profits of IFTs

The overall structures of the contours of IFTs' average profits are similar across the three different networks. The main difference is that when the market is highly UFT-dominated, the average maximum profit that each IFT can earn decreases as the network becomes increasingly complex.

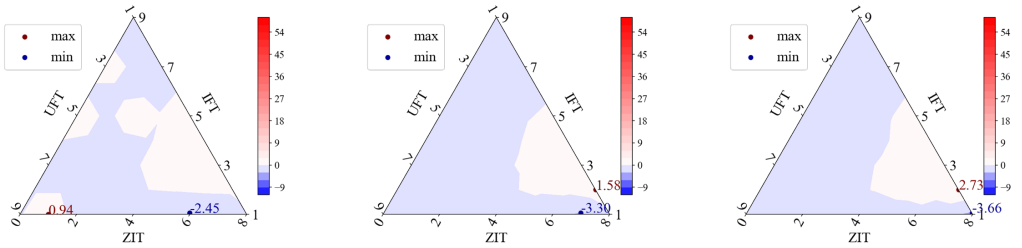
We also found another interesting result. Although an IFT in the market of $n_{\text{IFT}} = 1$ is usually able to achieve the highest average profits among all networks, the structure of UFT and ZIT differs at the peak point. By comparing the nine contours plots in the Fig. 7.9, we find that when the market topology is a messy network, the share of UFT at the highest profit point is much larger than that of ZIT ($n_{\text{UFT}} : n_{\text{ZIT}} = 7 : 2$), and that the peak points vary less for different levels of market risk settings. In contrast, in the cycle and complete networks, the share of UFT at the highest point is much smaller than that of ZIT

($n_{\text{UFT}} : n_{\text{ZIT}} = 2 : 7$). That is, in the messy network, each IFT's profit comes mainly from grabbing profits of UFT, while in the cycle and complete networks, the IFT's profit mainly stems from the increased randomness of the market.

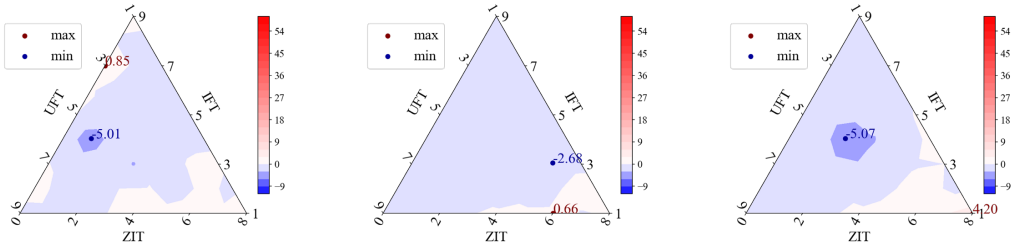
Uninformed Fundamental Traders Due to the lack of information advantage, the uninformed fundamental traders in the messy CDA market suffer a huge loss if the market is highly UFT-dominated. However, when the market topology is cycle or complete, the loss of UFTs is substantially moderated.



(a) IBR=0.05, messy network (b) IBR=0.1, messy network (c) IBR=0.25, messy network



(d) IBR=0.05, cycle network (e) IBR=0.1, cycle network (f) IBR=0.25, cycle network



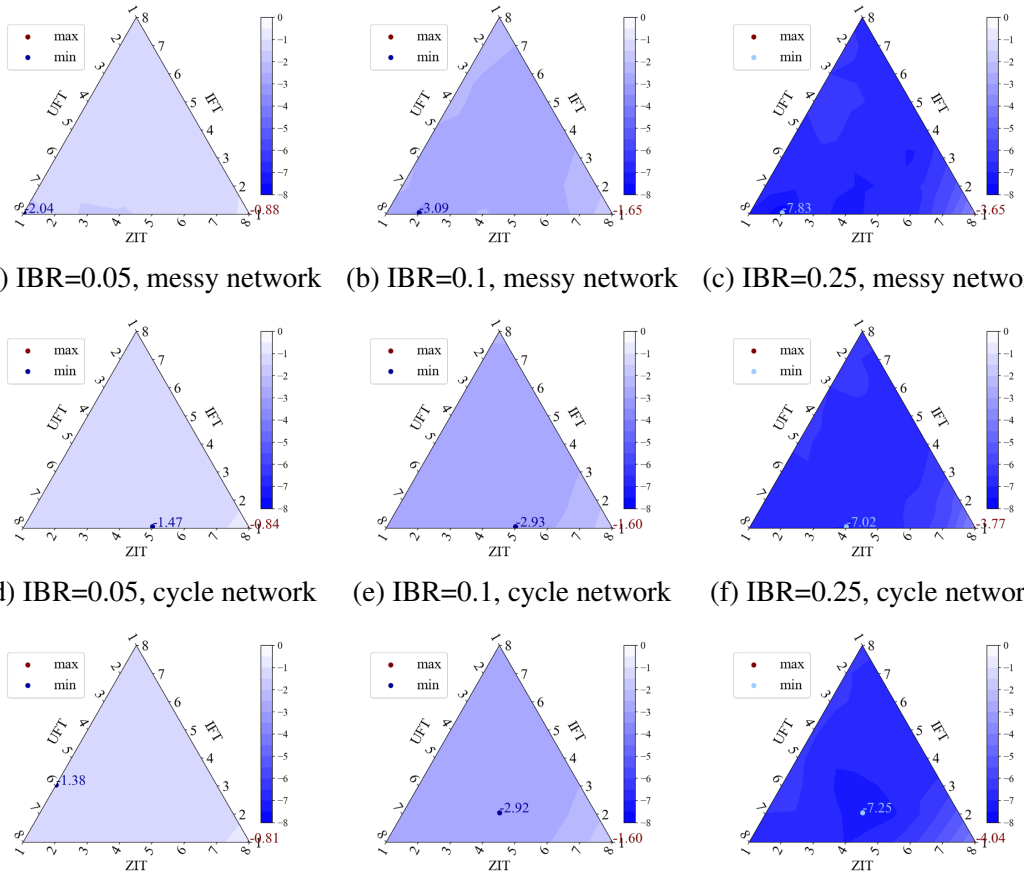
(g) IBR=0.05, complete network (h) IBR=0.1, complete network (i) IBR=0.25, complete network

Figure 7.10 Agents' Cumulative Average Profits of UFTs

In addition, several more features can be identified from these contour charts. In both the messy and the cycle networks, there are small triangular areas in the contour plots roughly representing the same market structure settings ($n_{\text{IFT}} > 1, n_{\text{ZIT}} \geq 4$ and

$n_{\text{UFT}} \leq 4$), and the UFTs in such areas can obtain positive profits. The average profits of IFTs within these areas are also positive, implying that both UFTs and IFTs can profit from an information advantage over ZITs. Also, in a complete network, UFTs can achieve small positive average profits when the market is dominated by randomness.

Zero-intelligence Traders By comparing with the contours of Fig. 7.11, their profit distributions are only related to the overall market risk - the information bias rate - while the market topology has little influence on the profits.



(a) IBR=0.05, messy network (b) IBR=0.1, messy network (c) IBR=0.25, messy network

(d) IBR=0.05, cycle network (e) IBR=0.1, cycle network (f) IBR=0.25, cycle network

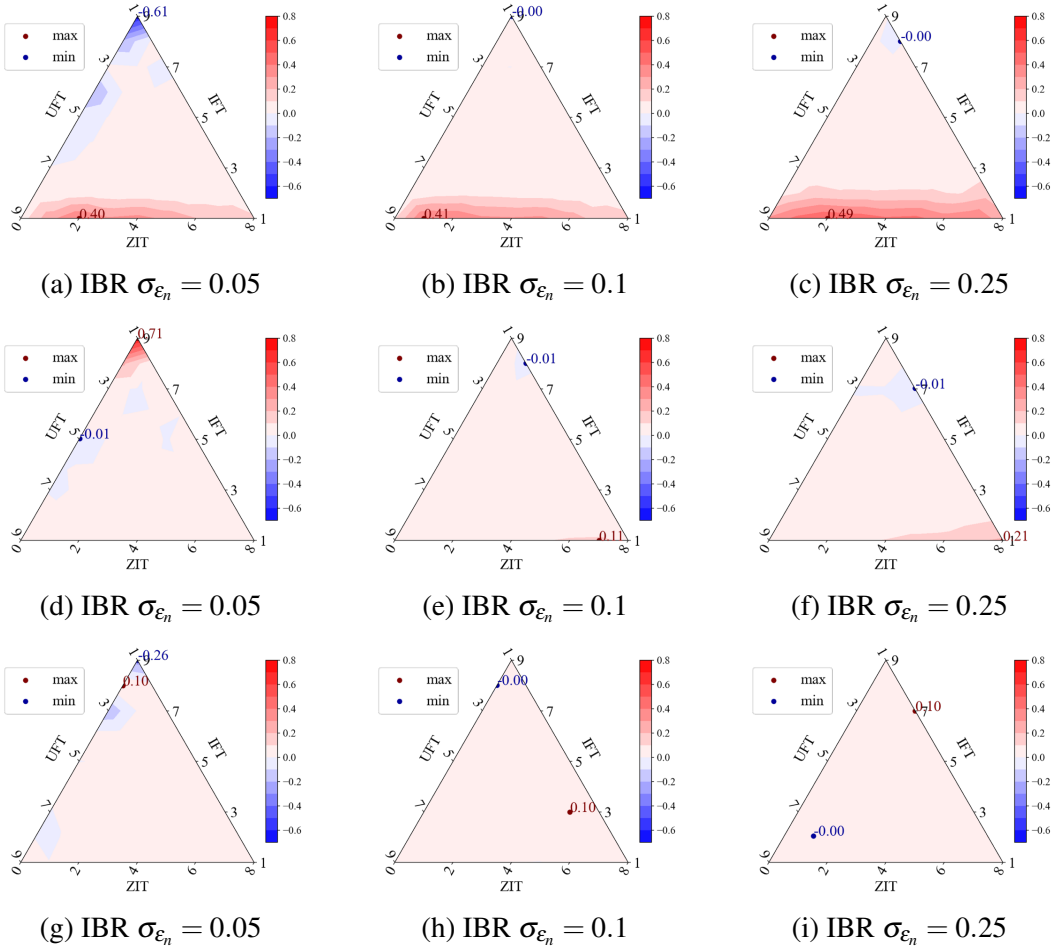
(g) IBR=0.05, complete network (h) IBR=0.1, complete network (i) IBR=0.25, complete network

Figure 7.11 Agents' Average Cumulative Profits of ZITs

Information Advantage

A previous study in Chapter 5 found that in CDA markets, when the agents' topology network is messy, IFTs have significant information advantage over UFTs when the market

is UFT-dominated. Fig. 7.12 shows that the information advantage of IFT over UFT is clearly reduced in a UFT-dominated market under more complicated networks.



(a) IBR $\sigma_{\varepsilon_n} = 0.05$ (b) IBR $\sigma_{\varepsilon_n} = 0.1$ (c) IBR $\sigma_{\varepsilon_n} = 0.25$

(d) IBR $\sigma_{\varepsilon_n} = 0.05$ (e) IBR $\sigma_{\varepsilon_n} = 0.1$ (f) IBR $\sigma_{\varepsilon_n} = 0.25$

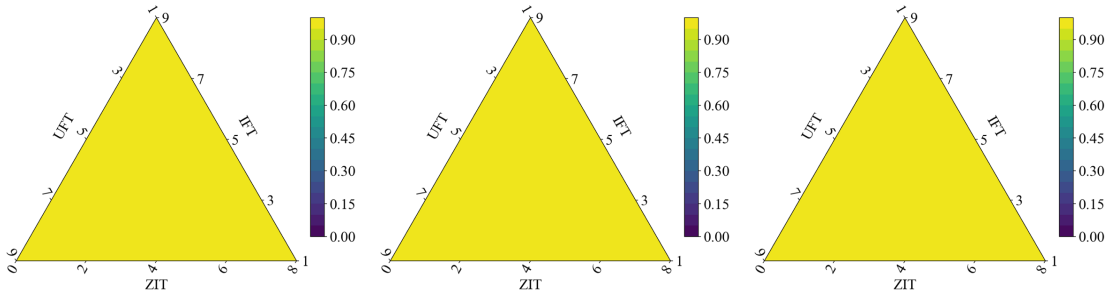
(g) IBR $\sigma_{\varepsilon_n} = 0.05$ (h) IBR $\sigma_{\varepsilon_n} = 0.1$ (i) IBR $\sigma_{\varepsilon_n} = 0.25$

Figure 7.12 Information Advantage of IFTs in the CDA Market with Different Networks

7.4.4 Inference Measurement

True Negative Rates

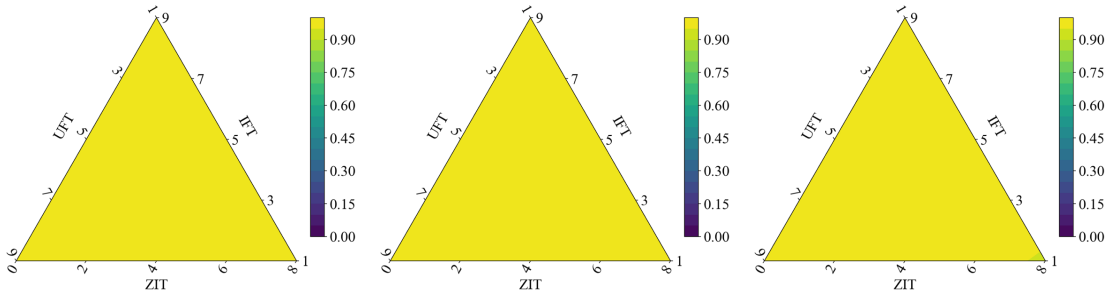
In the messy network, Fig. 7.13 shows that there are very high true negative rates of the inferences among all parameter settings.



(a) IBR $\sigma_{\varepsilon_n} = 0.05$ (b) IBR $\sigma_{\varepsilon_n} = 0.1$ (c) IBR $\sigma_{\varepsilon_n} = 0.25$

Figure 7.13 The Inference True Negative Rate of the Uninformed Fundamental Traders in the Messy Network

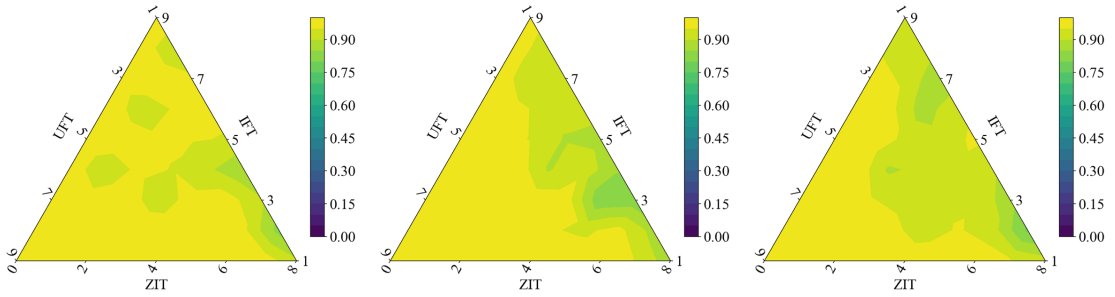
Fig.7.14 shows that the contour plots of the corresponding true negative rate in the cycle network are similar to that of the messy network with high rates.



(a) IBR $\sigma_{\varepsilon_n} = 0.05$ (b) IBR $\sigma_{\varepsilon_n} = 0.1$ (c) IBR $\sigma_{\varepsilon_n} = 0.25$

Figure 7.14 The Inference True Negative Rate of the Uninformed Fundamental Traders in the Cycle Network

In the complete network, each UFT can receive information about all other agents' quoting strategies at each time point. However, even with full information exposure, such sufficient information does not always lead to accurate inferences for the UFTs as Fig. 7.15 shows. The contour plots show that the true negative rates only reach around 75% in the case of only a few UFTs and IFTs in the market. It is reasonable to infer that when the number of ZITs in the market is high, although each UFT knows everyone else's strategy information, most of the other agents' information is not superior to the UFTs' own. In other words, such information is little effective and can even be misleading to the UFTs.

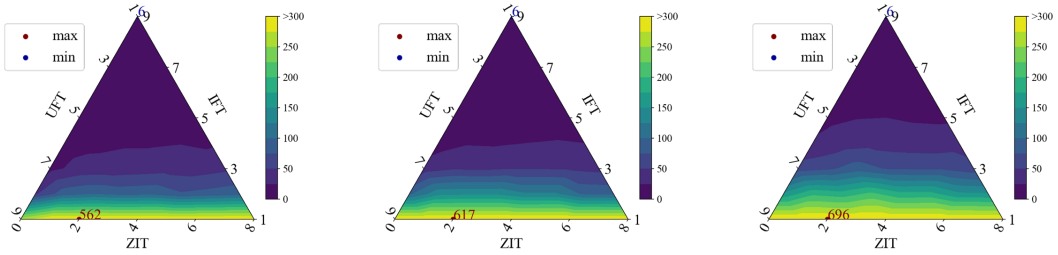


(a) IBR $\sigma_{\varepsilon_n} = 0.05$ (b) IBR $\sigma_{\varepsilon_n} = 0.1$ (c) IBR $\sigma_{\varepsilon_n} = 0.25$

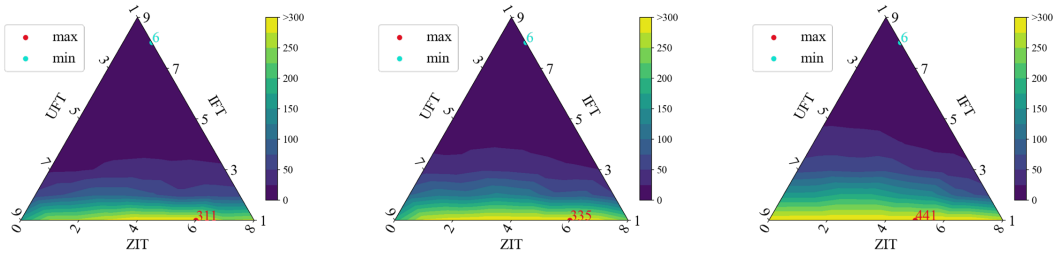
Figure 7.15 The Inference True Negative Rate of the Uninformed Fundamental Traders in the Complete Network

Inference Duration

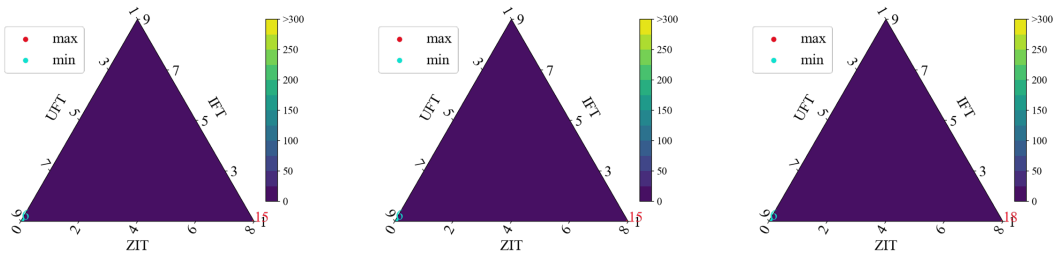
Fig. 7.16 presents the average duration of an uninformed fundamental trader finishing the inference. We can see that the UFTs in the complete network spend an extremely short period on finishing the inferences. However, although the complete network shortens the inference time for the UFTs, it also implies that the UFTs only take advantage of limited information to infer, which could cause biases and inaccuracy of inferences, as Section 7.4.4 shows. Therefore, with the order placing strategy this thesis sets, each UFT's decision-making could be misled by over-shared information in the complete network.



(a) IBR $\sigma_m = 0.05$, messy network (b) IBR $\sigma_m = 0.1$, messy network (c) IBR $\sigma_m = 0.25$, messy network



(d) IBR $\sigma_m = 0.05$, cycle network (e) IBR $\sigma_m = 0.1$, cycle network (f) IBR $\sigma_m = 0.25$, cycle network



(g) IBR $\sigma_m = 0.05$, complete network (h) IBR $\sigma_m = 0.1$, complete network (i) IBR $\sigma_m = 0.25$, complete network

Figure 7.16 The Average Inference Duration of an Uninformed Fundamental Trader in Different Networks

Although in a complete network, all information is shared among all other agents, the information shared by all other agents is not necessarily better than the information owned by a UFT himself. Informed fundamental traders with inside information can provide extra information to UFTs, but the randomness of zero-intelligence traders may mislead UFTs to make the right inference. Also, too short inferring duration, to a certain extent, does not allow uninformed fundamental traders to correct their beliefs.

7.4.5 Statistical Testing

In this chapter, our goal is to comprehend the impact of information-sharing mechanisms on flash crashes within our model. We continue to utilize t-tests and Cohen's d for statistical testing to investigate disparities in crash sizes and the cumulative profits of UFTs across different information-sharing networks.

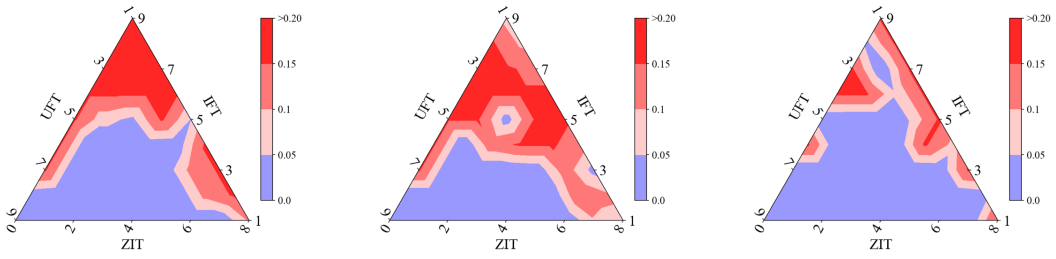
In this section, we designate the dataset obtained from the simulations of the messy network (Chapter 5) as Group 1, while the dataset derived from simulations of cycle networks or complete networks, as outlined in this chapter, is assigned to Group 2.

Crash Sizes

As part of our investigation, we put forth the following null hypothesis regarding crash sizes:

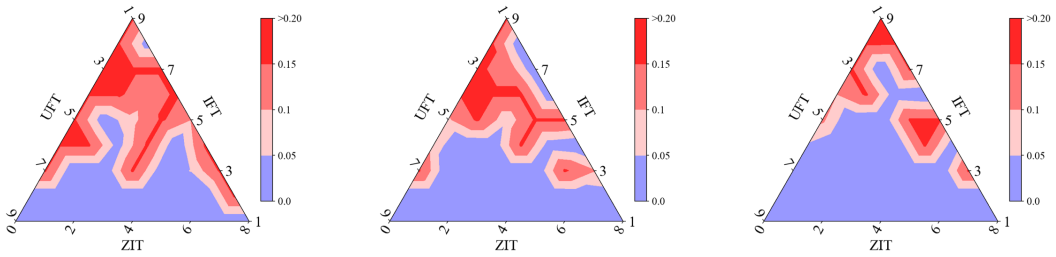
Null hypothesis H_0: *The crash sizes are not significantly different in a more complete network, such as a cycle network or complete network.*

Based on the triangular contour plots of p-values across different networks, we make a notable observation: significant differences in crash sizes (indicated by p-values < 0.05) are evident in both the cycle network and the complete network. This finding highlights the substantial impact of network structure on flash crashes. However, this significance is particularly pronounced in UFT-dominated markets, where the p-values are consistently small. This insight suggests that information-sharing networks have a profound effect on crash sizes in specific market structures.



(a) IBR $\sigma_m = 0.05$ (b) IBR $\sigma_m = 0.1$ (c) IBR $\sigma_m = 0.25$

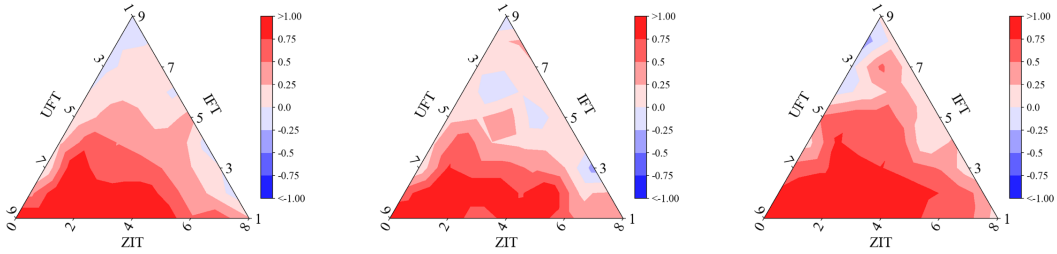
Figure 7.17 p-values in Crash Sizes between Messy and Cycle Networks



(a) IBR $\sigma_m = 0.05$ (b) IBR $\sigma_m = 0.1$ (c) IBR $\sigma_m = 0.25$

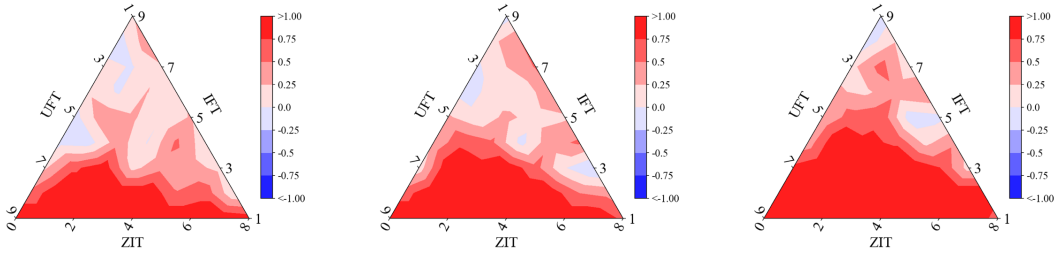
Figure 7.18 p-values in Crash Sizes between Messy and Complete Networks

As shown in Figs. 7.19 and 7.20, Cohen's d values also demonstrate a notable positive effect, signifying that Group 1 exhibits a higher mean, especially in the UFT-dominated markets. In other words, the crash sizes in Group 2 are significantly reduced when compared to those in Group 1. This compelling observation highlights the efficacy of information sharing in cycle networks or complete networks in the context of mitigating flash crashes in the market.



(a) IBR $\sigma_m = 0.05$ (b) IBR $\sigma_m = 0.1$ (c) IBR $\sigma_m = 0.25$

Figure 7.19 Cohen's d in Crash Sizes between Messy and Cycle Networks



(a) IBR $\sigma_m = 0.05$ (b) IBR $\sigma_m = 0.1$ (c) IBR $\sigma_m = 0.25$

Figure 7.20 Cohen's d in Crash Sizes between Messy and Complete Networks

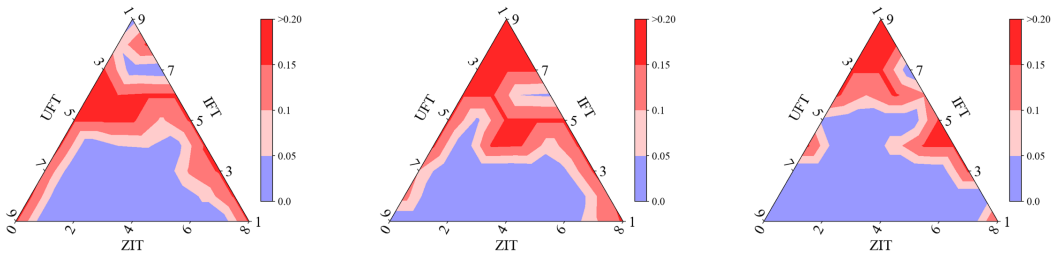
These findings, as indicated by Cohen's d, underscore the valuable role that information-sharing networks can play in enhancing market stability and reducing the impact of flash crashes. These findings have implications for both market participants and regulatory bodies regarding the construction of connections between agents.

UFTs' Profits

We introduce another null hypothesis related to UFT's profits:

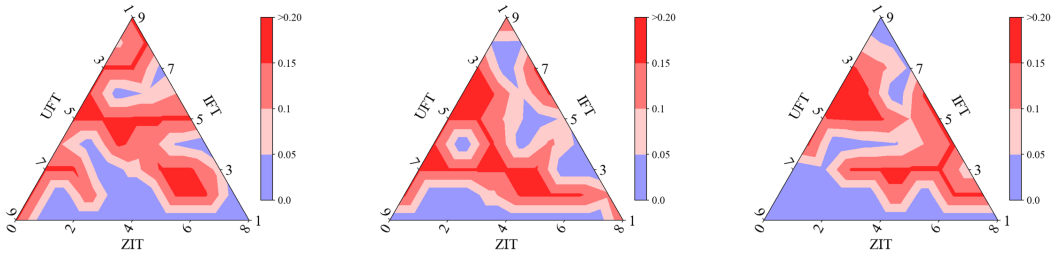
Null Hypothesis H_0**:** *The profits of UFTs cannot be diminished through the establishment of information-sharing connections with other agents.*

With the triangular contour plots of p-values across different networks shown in Fig. 7.21, we make a noteworthy observation that only UFT-dominated markets demonstrate small p-values in the cycle network. Moreover, in strongly UFT-dominated regions within the complete network in Fig. 7.22, there is substantial evidence supporting the rejection of the null hypothesis.



(a) IBR $\sigma_m = 0.05$ (b) IBR $\sigma_m = 0.1$ (c) IBR $\sigma_m = 0.25$

Figure 7.21 p-values in UFT's Profits between Messy and Cycle Networks

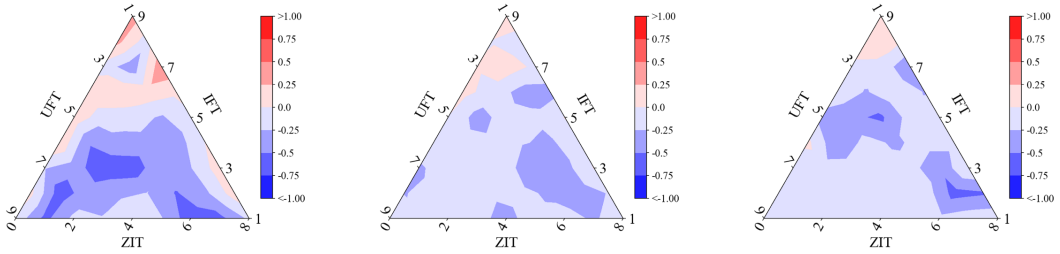


(a) IBR $\sigma_m = 0.05$ (b) IBR $\sigma_m = 0.1$ (c) IBR $\sigma_m = 0.25$

Figure 7.22 p-values in UFT's Profits between Messy and Complete Networks

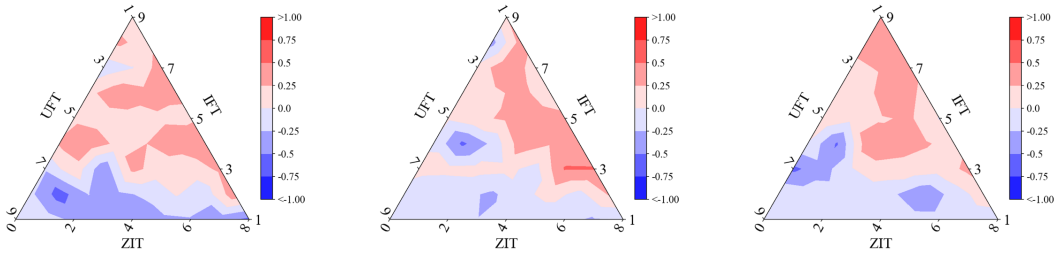
This insight provides valuable implications for understanding the role of information-sharing networks and their potential effects on UFTs' profits in different market structures. To gain a deeper understanding of the effects, we proceed to calculate Cohen's d values, which will reveal the directions and magnitudes of these effects.

Fig. 7.23 and Fig. 7.24 present Cohen's d values in contour plots under various parameter settings related to the testing on cycle networks and complete networks, respectively. Combined with the p-value results that emphasise the significance in UFT-dominated markets in the last section, we have discovered that both the cycle network and the complete network demonstrate a negative effect in Cohen's d values within these UFT-dominated markets. This suggests that information-sharing can assist uninformed fundamental traders in increasing profits or reducing losses.



(a) IBR $\sigma_m = 0.05$ (b) IBR $\sigma_m = 0.1$ (c) IBR $\sigma_m = 0.25$

Figure 7.23 Cohen's d in UFT's Profits between Messy and Cycle Networks



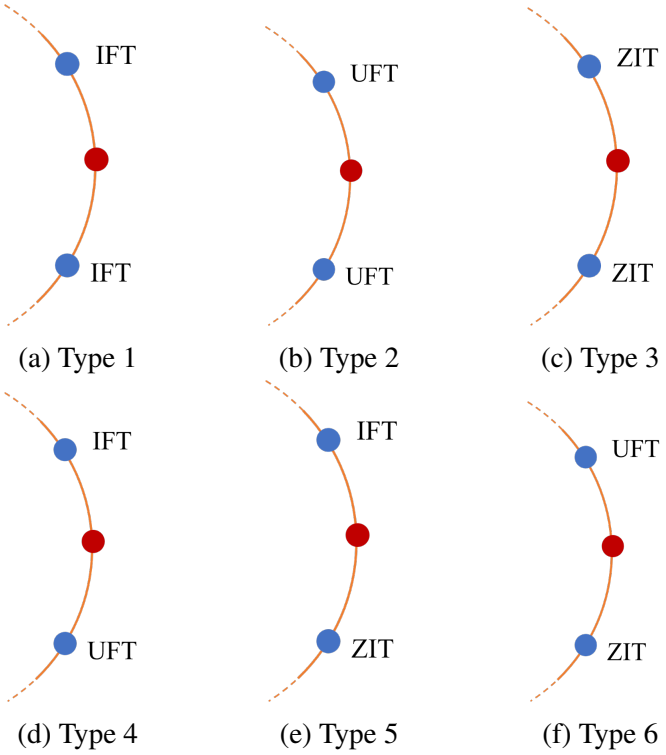
(a) IBR $\sigma_m = 0.05$ (b) IBR $\sigma_m = 0.1$ (c) IBR $\sigma_m = 0.25$

Figure 7.24 Cohen's d in UFT's Profits between Messy and Complete Networks

7.4.6 Asymmetry in Cycle Networks

When constructing the information-sharing network, we actually introduced information asymmetry artificially. Unlike the messy network and the complete network, all

agents in the cycle network are not symmetrical with each other. In a cycle network, each uninformed fundamental trader only knows the two neighbours' quoting information so that, in detail, there are six possible types for each UFT's position in the cycle network shown in Fig. 7.25:



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

(d) Type 4 (e) Type 5 (f) Type 6

Figure 7.25 All Types of Positions of the Uninformed Fundamental Traders in the Cycle Network

Type 1: both two neighbours are informed fundamental traders.

Type 2: both two neighbours are uninformed fundamental traders.

Type 3: both two neighbours are zero-intelligence traders.

Type 4: one neighbour is an informed fundamental trader, and the other one is an uninformed fundamental trader

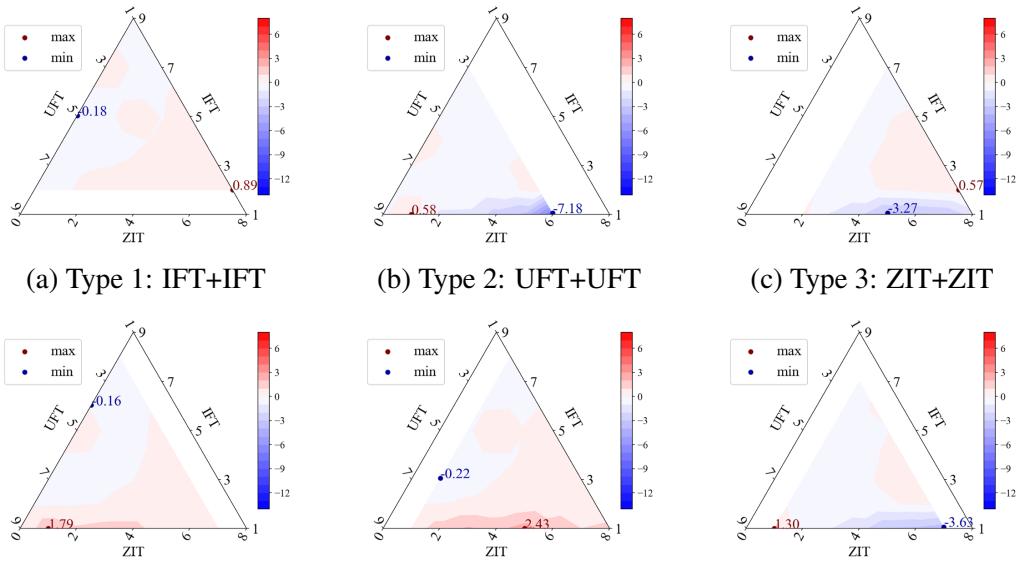
Type 5: one neighbour is an informed fundamental trader, and the other one is a zero-intelligence trader

Type 6: one neighbour is an uninformed fundamental trader, and the other is a zero-intelligence trader

In the trading process, each UFT could likely behave differently within a simulation because of the varied positions. Therefore, we also propose an additional hypothesis: in a cycle network, the trader types of a UFT's neighbours influence the UFT's decisions and profits, and the UFTs are more willing to establish information-sharing connections with IFTs. Here we analyse the asymmetry of uninformed fundamental traders in the Cycle network from the aspects of profits.

Profits

First, Fig. 7.26 shows the contours of the expected profits of UFTs with different position types for different market structures and risks, where red represents positive profits and blue for negative returns. When the market risk is 0.05, it is clear that the contour plots for types 1, 4 and 5 have large areas of red, especially in the strongly UFT-dominated area, where the average profits for the UFTs are still positive. In contrast, the pattern of the profit contour plots for types 2, 3 and 6 is similar to the overall contour plots in Fig. 7.10, with most of the areas indicating negative profits.

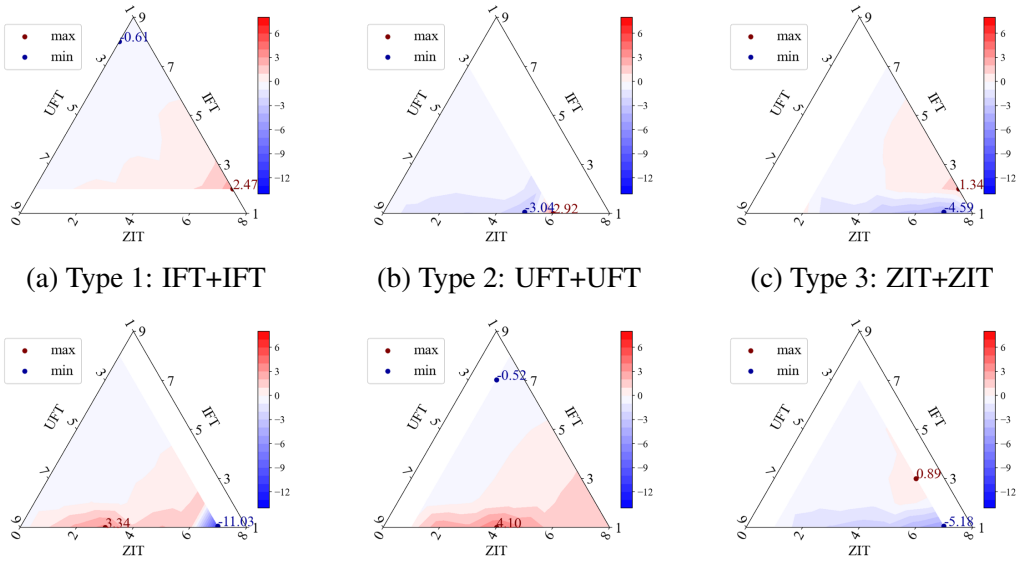


(a) Type 1: IFT+IFT (b) Type 2: UFT+UFT (c) Type 3: ZIT+ZIT

(d) Type 4: IFT+UFT (e) Type 5: IFT+ZIT (f) Type 6: UFT+ZIT

Figure 7.26 UFT's Expected Profits among All Types of Positions in the Cycle Network (IBR=0.05)

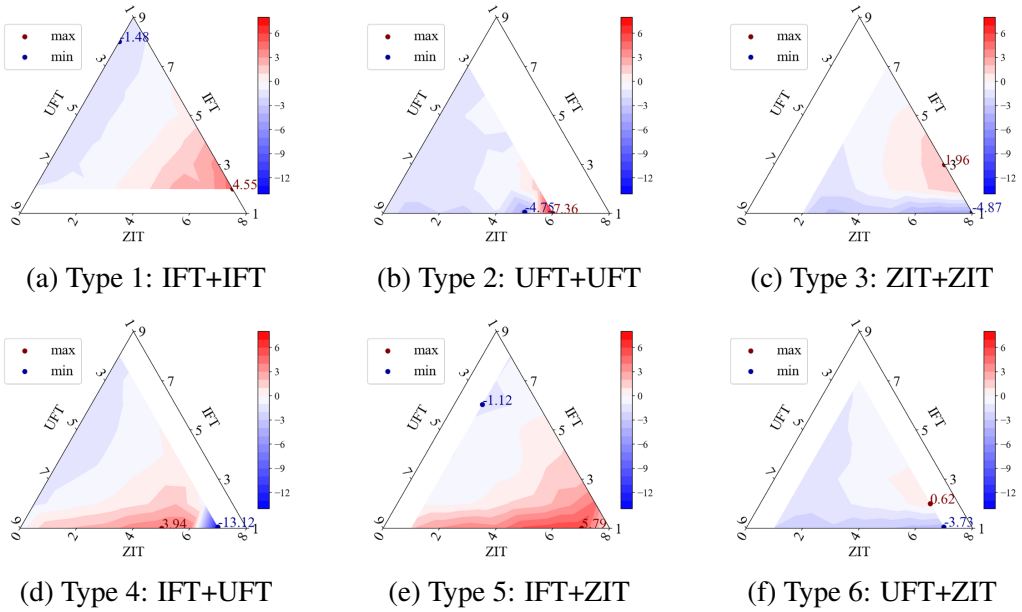
When the market risk is 0.1 and 0.25, the UFTs adjacent to IFTs significantly earn higher profits than those not adjacent to IFTs. When adjacent to at least one IFT, a single UFT can obtain relatively high positive profits when the market is random or strongly UFT-dominated, and the risk of the market can magnify the profits, but it also magnifies the loss suffered by the UFTs not adjacent to any IFTs.



(a) Type 1: IFT+IFT (b) Type 2: UFT+UFT (c) Type 3: ZIT+ZIT

(d) Type 4: IFT+UFT (e) Type 5: IFT+ZIT (f) Type 6: UFT+ZIT

Figure 7.27 UFT's Expected Profits among All Types of Positions in the Cycle Network (IBR=0.1)



(a) Type 1: IFT+IFT (b) Type 2: UFT+UFT (c) Type 3: ZIT+ZIT

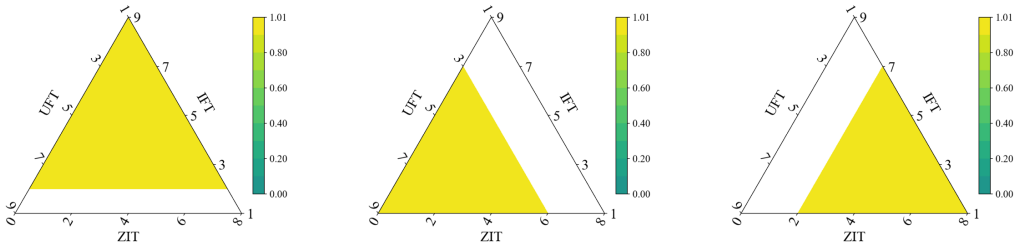
(d) Type 4: IFT+UFT (e) Type 5: IFT+ZIT (f) Type 6: UFT+ZIT

Figure 7.28 UFT's Expected Profits among All Types of Positions in the Cycle Network (IBR=0.25)

The long-time inferring and learning have led to such UFTs losing significant profits to IFTs and other UFTs who have made correct inferences and changed their strategies early.

Inference Accuracy

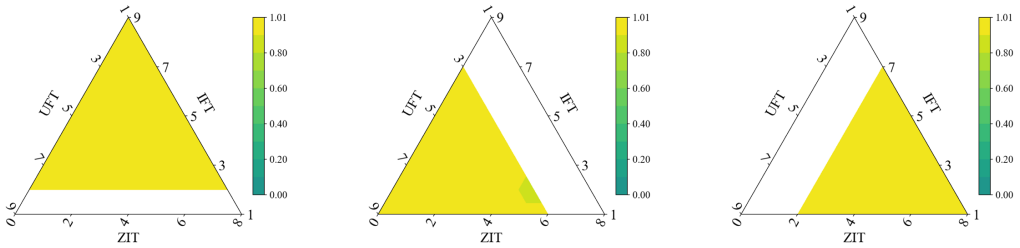
We first calculated the probability of each uninformed fundamental trader making the correct inference when the fake shock occurred, as measured by the true negative rate.



(a) Type 1: IFT+IFT (b) Type 2: UFT+UFT (c) Type 3: ZIT+ZIT

(d) Type 4: IFT+UFT (e) Type 5: IFT+ZIT (f) Type 6: UFT+ZIT

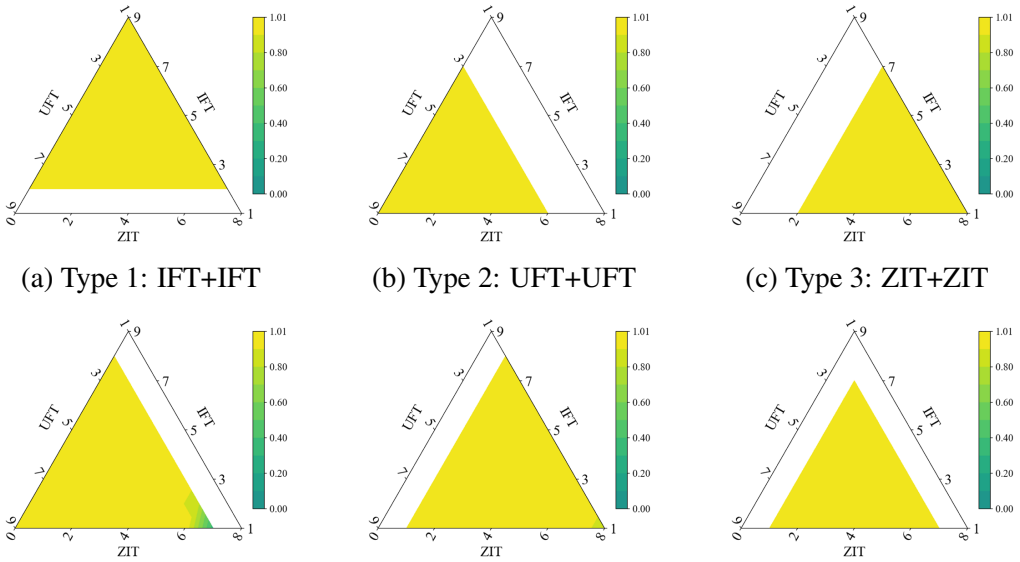
Figure 7.29 UFT's True Negative Rate among All Types of Positions in the Cycle Network (IBR=0.05)



(a) Type 1: IFT+IFT (b) Type 2: UFT+UFT (c) Type 3: ZIT+ZIT

(d) Type 4: IFT+UFT (e) Type 5: IFT+ZIT (f) Type 6: UFT+ZIT

Figure 7.30 UFT's True Negative Rate among All Types of Positions in the Cycle Network (IBR=0.1)



(a) Type 1: IFT+IFT (b) Type 2: UFT+UFT (c) Type 3: ZIT+ZIT

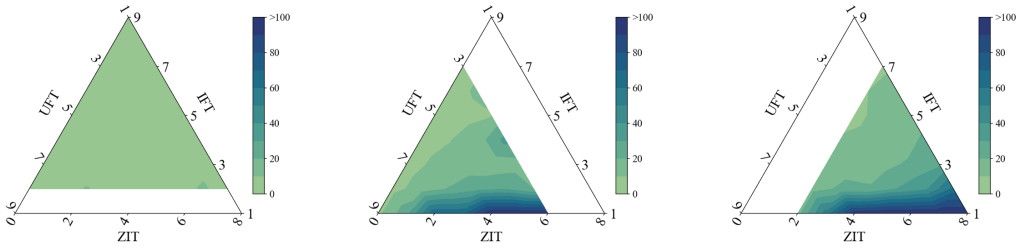
(d) Type 4: IFT+UFT (e) Type 5: IFT+ZIT (f) Type 6: UFT+ZIT

Figure 7.31 UFT's True Negative Rate among All Types of Positions in the Cycle Network (IBR=0.25)

We find that an uninformed fundamental trader can almost make a correct inference with a high probability, regardless of the position where the trader is. Therefore, the difference in profits of the UFTs in different positions is likely to depend on the inference durations.

Inference Duration

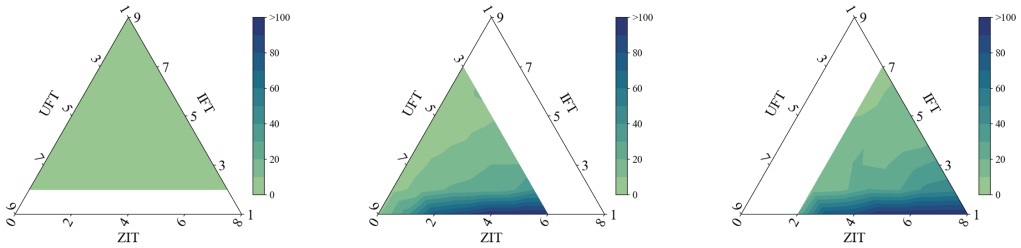
The reason for this situation can actually be seen by examining how these UFTs behave when making inferences. After the occurrence of the fake shock, although almost every UFT was able to infer that they had been deceived, the durations required to make correct inferences differed among UFTs in different positions. As we can see from Fig. 7.32, when a UFT shares information with at least one IFT, (Type 1, 4 and 5) the UFT only needs less than 20 iterations on average to learn until making a correct inference. When a UFT does not share information with an IFT (Type 2, 3 and 6), it takes much longer to make an inference than when a UFT shares information with an IFT, especially when the market is strongly UFT-dominated, where hundreds of iterations are required to make an inference.



(a) Type 1: IFT+IFT (b) Type 2: UFT+UFT (c) Type 3: ZIT+ZIT

(d) Type 4: IFT+UFT (e) Type 5: IFT+ZIT (f) Type 6: UFT+ZIT

Figure 7.32 UFT's Average Inference Duration among All Types of Positions in the Cycle Network (IBR=0.05)



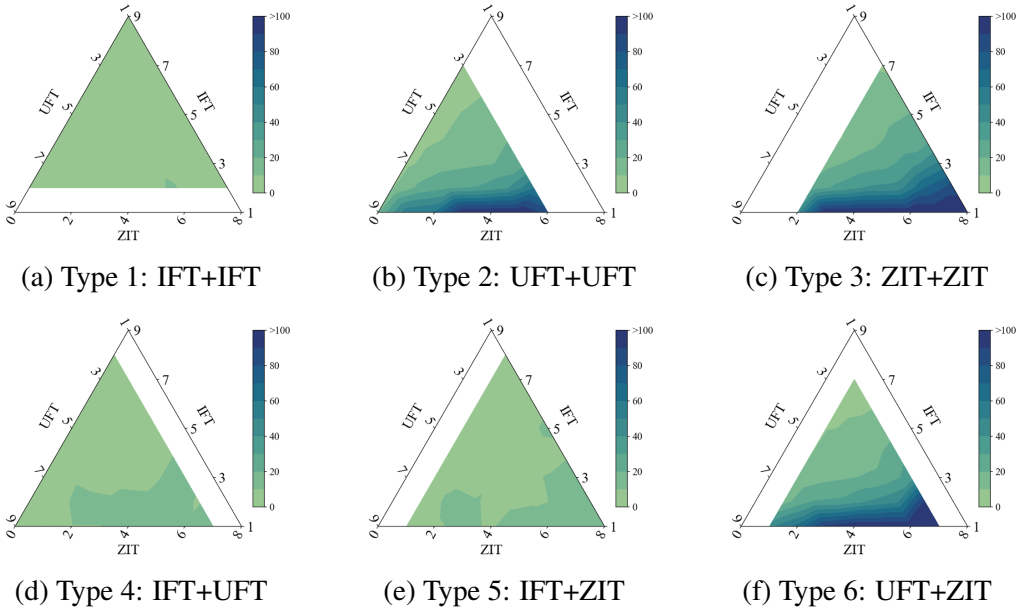
(a) Type 1: IFT+IFT (b) Type 2: UFT+UFT (c) Type 3: ZIT+ZIT

(d) Type 4: IFT+UFT (e) Type 5: IFT+ZIT (f) Type 6: UFT+ZIT

Figure 7.33 UFT's Average Inference Duration among All Types of Positions in the Cycle Network (IBR=0.1)

This section shows that uninformed fundamental traders, who are information disadvantaged, have the incentive to share information with informed fundamental traders. However, as we studied before, when the market is shifted from the messy network to the

cycle network with the addition of information sharing, the average profit of the informed fundamental traders decreases as a result. Therefore, for IFTs, there is no incentive to share information with other agents who lack information. This finding is also consistent with the conclusion in Goldstein et al. (2021).



(a) Type 1: IFT+IFT (b) Type 2: UFT+UFT (c) Type 3: ZIT+ZIT

(d) Type 4: IFT+UFT (e) Type 5: IFT+ZIT (f) Type 6: UFT+ZIT

Figure 7.34 UFT's Average Inference Duration among All Types of Positions in the Cycle Network (IBR=0.25)

7.5 Conclusion

This chapter introduced the concept of information-sharing between agents and extended the flash crash model of information asymmetry by exploring different network topologies for agent connections. Combined with the CDA model in Chapter 5, we classify the agent topologies into: the messy network; the cycle network; and the complete network, according to the different complexity of the topologies. We have obtained the following findings in this chapter:

- By constructing more complex topologies of agents, cycle and complete, by building different information sharing mechanisms between agents, we find that the flash crash caused by the fake shock can be significantly alleviated in more complicated

networks from the perspectives of both crash sizes and duration, as illustrated in Section 7.4.5.

- In addition, the cycle network and complete network significantly change the profit structure of fundamental traders, but not the profit structure of the market maker or the zero-intelligence traders. In the messy network, IFTs take remarkable benefits from UFTs in strongly UFT-dominated markets, but such a profit gap has been significantly weakened in the cycle network and the complete network. In the complete network, IFTs have little information advantages over UFTs in response to the fake shock shown in Fig. 7.12.

- The cycle and complete networks can significantly reduce the losses of each UFT, as shown in Figs. 7.21-7.24, especially when the market is strongly UFT-dominated.

- However, we found that a complete network is not conducive to making the correct decisions for each UFT. It implies that building a market with full information exposure could not be helpful to uninformed traders.

- We also study the asymmetry among the uninformed fundamental traders for the cycle network in Section 7.4.6. We find that UFTs that create information-sharing links with IFTs can speed up the time required for inference. Thus they can make higher profits than the UFTs not connected to IFTs, even if the UFTs themselves are not aware of the types of the traders they are connected with.

In summary, this chapter's study of agent information-sharing network topologies shows that more complex network structures can greatly weaken the impact of the fake shock on the market, but an over-complex network could have negative effects on decision-making by agents. It also suggests that uninformed traders are more willing to share information with informed traders.

Chapter 8

Conclusions

8.1 Conclusion Summary

Our experimental results reported in this thesis contribute to the analysis and understanding of informational asymmetry in financial markets under the circumstances of flash crashes. This thesis aims to investigate four kinds of research questions:

- How does spoofing impact flash crashes and market fluctuations?
- How does the market and agent behaviour change when uninformed traders can learn from public information?
- How and to what extent do batch auctions serve as an alternative to continuous double auctions in mitigating adverse selection in incomplete markets?
- How and to what extent does the implementation of an information-sharing mechanism mitigate flash crashes and reduce losses for traders affected by spoofing?

To answer these questions, we built, implemented, and assessed an agent-based model of an extended information-sequential trading framework inspired by Das. Starting with the fundamental model settings of (Das, 2005; Das and Magdon-Ismail, 2008; Glosten and Milgrom, 1985), this study has made following key extensions:

- The types of trading agents considered included: fundamental traders, both informed and uninformed; zero-intelligence traders; and a market maker; and these different types of agents were allowed to interact with each other;
- Fundamental traders were divided into informed fundamental traders and uninformed fundamental traders based on whether the trader is able to correctly recognise the shock as fake;
- Uninformed fundamental traders were enabled with learning abilities to infer whether the shock is fake, based on data from past order flows;
- The framework was extended with different network topologies for agent connections, so as to explore the effects of information-sharing between agents;
- The framework was extended with both continuous auctions and frequent batch auctions.

We implemented and assessed the model from the following aspects: market structure; market volatility; the complexity of agent networks; and market mechanisms. Interrupted by a fake shock, agents differ in actions, leading to various dynamics in terms of market prices, agents' profits, and so forth.

Our key findings are:

- In the basic model without intelligent uninformed traders, the fake shock decreases market prices to different degrees as the market structures and volatility vary. However, the variability of the marker prices after the shock is clearly related to the balance of the fundamental traders: As the numbers of IFTs and UFTs get closer, the market becomes more unstable. This instability is also reflected in the other metrics. Compared to IFTs, who achieve higher average returns in more UFT-dominated markets, UFTs suffer the greatest losses in more balanced markets with great variability.
- Next, we allowed uninformed fundamental traders to have the learning ability to infer whether they are deceived so that the uninformed fundamental traders can

re-estimate stock prices. The results show that each UFT is highly likely to make the correct inference to change strategy, making the market price recover. However, a more UFT-dominated market suffers a larger price decline, and, in this case, each IFT can earn a lot from UFTs in the market because of their high information advantage. Also, when the share of IFTs is lower than 20%, an uninformed trader suffers a big loss.

- If we change the market mechanism of the framework from continuous auctions to batch auctions with different matching frequencies, we found that flash crashes can be weakened, but the decrease in crash speed is not significant when batch auctions are applied. However, batch auctions can better defend against the fake shock for lower-risk markets than higher-risk markets.

- Chapter 7 considers the connections between two agents for information sharing in two additional topologies: a cycle network and a complete network. A more complicated agent network can narrow the profit difference between informed and uninformed traders. At the same time, in our study of the cycle network, we found that when a UFT establishes an information-sharing connection with an IFT, even if that UFT does not know the identity of the IFT, the UFT can make the inference more accurately and faster than a UFT that does not link with an IFT, resulting in a higher cumulative profit. The results also suggest that uninformed traders are more willing to share information with informed traders. Although there is full information exposure in the complete network, it is not superior to the cycle network for a UFT from the perspective of inference accuracy.

8.2 Insights into Real-world Markets

In the real world, informed and uninformed fundamental traders and zero-intelligence represent different types of traders. According to the difference in information advantage, zero-intelligence traders often represent individual investors who have limited information sources. Uninformed fundamental traders represent smaller institutional investors who have the ability to estimate market prices and mine information from markets. However, they

lack large numbers or diversity of sources of information or any insider information. Or they could be programmatic traders who find it hard to adjust quickly to market movements. Informed fundamental traders represent the larger, more sophisticated institutional investors with a wealth of technical expertise, processing resources, and information sources. The key findings of this study also provide insights into real-world markets. For example,

- When the market is dominated by large institutions, the market as a whole is still relatively stable; in this case, investors have no incentive to undertake spoofing, as the spoofer can only get expect a limited amount of profits. Smaller institutions are relatively safe at this time.

- When small-scale institutional investors dominate the market, or, in other words, when there is only a small group of large institutions (accounting for lower than 20%), the market is relatively most vulnerable, and small-scale institutional investors are likely to suffer losses due to a lack of information. At this point, small institutional investors can avoid the information disadvantage by purchasing additional information sources or building cooperation with larger institutions.

- Altering the market mechanism from continuous auction to frequent batch auction reduces, to some extent, the effect of crashes on market movements, but it is not wise to set a low match frequency. On the one hand, batch auctions only have a limited effect on alleviating flash crashes, and on the other hand, setting a large batch size can affect traders' motivation to trade.

The equity market in China over the past 20 years experienced great growth and, with a market of $7.2 trillion, became the second-largest equity market in the world in March 2019. However, unlike the stock markets in developed markets, including the US and the UK, dominated by institutional investors, the stock markets in China are dominated by individual investors. In 2017 retail investors in China accounted for 82% of turnover, while this figure is below 20% for the same year in the US (Lockett, 2019; Wigglesworth and Martin, 2021). This thesis also provides insights into the reform of the stock markets in China. Although Fig. 8.1 shows that individual investors have become less dominant in the past few years, the stock market in China is still quite a distance from a developed

market. The high concentration of individual investors suggests the high volatility and long information spread time. Cakici et al. (2017) found that the predictability of the stock market returns in China is generally weak, with a few conflicting results out of expectations.

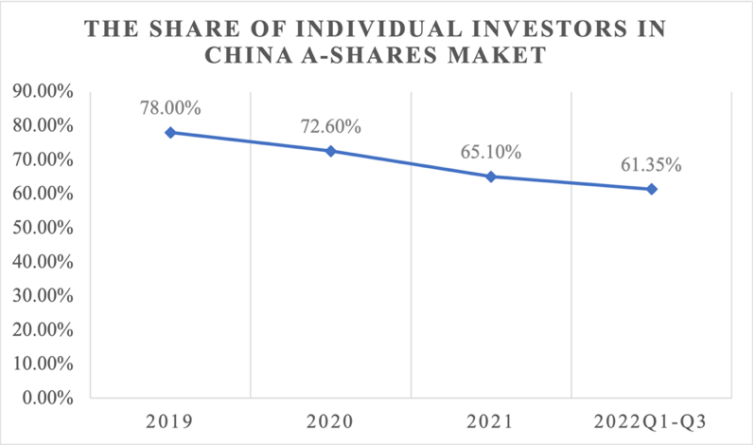


Figure 8.1 The Share of Individual Investors in China A-shares Market in the Past Four Years

8.3 Limitations and Future Work

Due to the limits of computational power, some parameters in this research have been treated as discrete variables and not part of a continuum. For example, we set a 10-agent model regarding the market structure settings, in which case the market-structure parameters are represented by the numbers of informed and uninformed fundamental traders rather than the market shares of those agents. Because of this limitation, the research reported in this thesis does not include sensitivity analysis for extreme cases, such as the case of an extremely UFT-dominated market with a share of IFTs/UFTs less than 10%. Future work could consider such extreme cases.

In addition, the costs associated with gaining an information advantage have not been considered in this research. In the real world, upgrading from zero-intelligence to fundamental traders or from uninformed traders to informed traders would only be achieved at some cost. Therefore, when examining the expected profits of individual agents, we can

determine the optimal strategies for each kind of agent by investigating different levels of information costs. Again, this could be the subject of future research.

While we conducted an analysis on numerous metrics, we have only selectively chosen those that are deemed valuable. An agent-based model is inherently complex, and the datasets resulting from simulations encompass a wide array of information. Beyond the analysis presented in this paper, it is inevitable that certain important data analysis findings may have been ignored. These will be subject to further exploration in future research.

As we are not able to exhaustively search the entire learning strategy space and because we used a specific model framework, the findings presented in this thesis may not apply to other strategies or to other model settings. Indeed, qualitatively distinct findings are always possible. Our learning strategy for the uninformed fundamental traders is limited, with all UFTs updating their beliefs under a certain rule. One unrealistic limitation of our work is that UFTs cannot update the beliefs again once they finish the inferences. As before, future work could explore these questions with more realistic assumptions about learning.

Interesting extensions of the work in this thesis could include multiple markets with different market mechanisms, which permit traders to select adaptively which market to participate in, or to formulate an iterated form of dynamic game-theoretic models. Similarly, broader explorations of different types of choice games, for instance, whether upgrading from an uninformed trader to an informed trader by paying for the inside information or whether building an information-sharing connection with another trader, could provide more insights into the understanding of the behaviours of agents and the market under information asymmetry.

Bibliography

Adrian, J. (2015). Informational inequality: How high frequency traders use premier access to information to prey on institutional investors. *Duke L. & Tech. Rev.*, 14:256.

Akansu, A. N. and Torun, M. U. (2015). *A primer for financial engineering: financial signal processing and electronic trading*. Academic Press.

Aldrich, E. M., Grundfest, J., and Laughlin, G. (2017). The flash crash: A new deconstruction. *Available at SSRN 2721922*.

Aldrich, E. M. and López Vargas, K. (2020). Experiments in high-frequency trading: comparing two market institutions. *Experimental Economics*, 23(2):322–352.

Allen, F. and Babus, A. (2009). Networks in finance. *The network challenge: strategy, profit, and risk in an interlinked world*, 367.

Amihud, Y. and Mendelson, H. (1980). Dealership market: Market-making with inventory. *Journal of financial economics*, 8(1):31–53.

Aquilina, M., Budish, E., and O'Neill, P. (2020). Quantifying the high-frequency trading "arms race": A simple new methodology and estimates. *Chicago Booth Research Paper*, 20:16.

Bachelier, L. (1900). Théorie de la spéculation. In *Annales scientifiques de l'École normale supérieure*, volume 17, pages 21–86.

Bala, V. and Goyal, S. (1998). Learning from neighbours. *The review of economic studies*, 65(3):595–621.

Brahma, A., Chakraborty, M., Das, S., Lavoie, A., and Magdon-Ismail, M. (2012). A bayesian market maker. In *Proceedings of the 13th ACM Conference on Electronic Commerce*, pages 215–232.

Breckenfelder, J. (2019). Competition among high-frequency traders, and market quality.

Brewer, P. J., Huang, M., Nelson, B., and Plott, C. R. (2002). On the behavioral foundations of the law of supply and demand: Human convergence and robot randomness. *Experimental economics*, 5:179–208.

Brogaard, J. et al. (2010). High frequency trading and its impact on market quality. *Northwestern University Kellogg School of Management Working Paper*, 66.

Budish, E., Cramton, P., and Shim, J. (2014). Implementation details for frequent batch auctions: Slowing down markets to the blink of an eye. *American Economic Review*, 104(5):418–24.

Budish, E., Cramton, P., and Shim, J. (2015). The high-frequency trading arms race: Frequent batch auctions as a market design response. *The Quarterly Journal of Economics*, 130(4):1547–1621.

Budish, E., Lee, R. S., and Shim, J. J. (2019). A theory of stock exchange competition and innovation: Will the market fix the market? Technical report, National Bureau of Economic Research.

Cakici, N., Chan, K., and Topyan, K. (2017). Cross-sectional stock return predictability in china. *The European Journal of Finance*, 23(7-9):581–605.

Cassady, R. (1967). *Auctions and auctioneering*. Univ of California Press.

CFTC, S. and SEC, U. (2010). Findings regarding the market events of may 6, 2010. *Report of the Staffs of the CFTC and SEC to the Joint Advisory Committee on Emerging Regulatory Issues*, 104.

Chaboud, A. P., Chiquoine, B., Hjalmarsson, E., and Vega, C. (2014). Rise of the machines: Algorithmic trading in the foreign exchange market. *The Journal of Finance*, 69(5):2045–2084.

Chatterjee, K. and Samuelson, W. (1983). Bargaining under incomplete information. *Operations research*, 31(5):835–851.

Chen, S.-H., Chang, C.-L., and Du, Y.-R. (2012). Agent-based economic models and econometrics. *The Knowledge Engineering Review*, 27(2):187–219.

Chiarella, C., Iori, G., and Perelló, J. (2009). The impact of heterogeneous trading rules on the limit order book and order flows. *Journal of Economic Dynamics and Control*, 33(3):525–537.

Cliff, D. and Bruten, J. (1997). Zero is not enough: On the lower limit of agent intelligence for continuous double auction markets.

CME Group (2010). What happened on may 6th?

Cohen, K. J. (1986). *The microstructure of securities markets*. Prentice Hall.

Cohen, S. (1988). Perceived stress in a probability sample of the united states.

Das, R., Hanson, J. E., Kephart, J. O., and Tesauro, G. (2001). Agent-human interactions in the continuous double auction. In *International joint conference on artificial intelligence*, volume 17, pages 1169–1178. Lawrence Erlbaum Associates Ltd.

Das, S. (2005). A learning market-maker in the glosten–milgrom model. *Quantitative Finance*, 5(2):169–180.

Das, S. and Magdon-Ismail, M. (2008). Adapting to a market shock: Optimal sequential market-making. In *NIPS*, volume 22, pages 361–368. Citeseer.

Dickey, D. A. and Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American statistical association*, 74(366a):427–431.

Du, S. and Zhu, H. (2017). What is the optimal trading frequency in financial markets? *The Review of Economic Studies*, 84(4):1606–1651.

Easley, D. and O'hara, M. (1987). Price, trade size, and information in securities markets. *Journal of Financial economics*, 19(1):69–90.

Erev, I. and Roth, A. E. (1998). Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria. *American economic review*, pages 848–881.

Fagiolo, G. and Roventini, A. (2016). Macroeconomic policy in dsge and agent-based models redux: New developments and challenges ahead. *Available at SSRN 2763735*.

Farmer, J. D. and Foley, D. (2009). The economy needs agent-based modelling. *Nature*, 460(7256):685–686.

Fievet, L. and Sornette, D. (2018). Calibrating emergent phenomena in stock markets with agent based models. *PloS one*, 13(3):e0193290.

Foucault, T., Kozhan, R., and Tham, W. W. (2017). Toxic arbitrage. *The Review of Financial Studies*, 30(4):1053–1094.

Franke, R. and Westerhoff, F. (2012). Structural stochastic volatility in asset pricing dynamics: Estimation and model contest. *Journal of Economic Dynamics and Control*, 36(8):1193–1211.

Friedman, D. (1991). A simple testable model of double auction markets. *Journal of Economic Behavior & Organization*, 15(1):47–70.

Garman, M. B. (1976). Market microstructure. *Journal of financial Economics*, 3(3):257–275.

Gibbons, J. D. and Chakraborti, S. (2014). *Nonparametric statistical inference*. CRC press.

Gigerenzer, G. and Selten, R. (2002). *Bounded rationality: The adaptive toolbox*. MIT press.

Gjerstad, S. and Dickhaut, J. (1998). Price formation in double auctions. *Games and economic behavior*, 22(1):1–29.

Glosten, L. R. (1987). Components of the bid-ask spread and the statistical properties of transaction prices. *The Journal of Finance*, 42(5):1293–1307.

Glosten, L. R. and Harris, L. E. (1988). Estimating the components of the bid/ask spread. *Journal of financial Economics*, 21(1):123–142.

Glosten, L. R. and Milgrom, P. R. (1985). Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *Journal of financial economics*, 14(1):71–100.

Gode, D. K. and Sunder, S. (1993). Allocative efficiency of markets with zero-intelligence traders: Market as a partial substitute for individual rationality. *Journal of political economy*, 101(1):119–137.

Goldstein, I., Xiong, Y., and Yang, L. (2021). Information sharing in financial markets. *Available at SSRN 3632315*.

Goldstein, M. A., Kumar, P., and Graves, F. C. (2014). Computerized and high-frequency trading. *Financial Review*, 49(2):177–202.

Grazzini, J. (2011). Estimating micromotives from macrobehavior. *University of Turin Department of Economics Working Papers Series*.

Grazzini, J. (2012). Analysis of the emergent properties: Stationarity and ergodicity. *Journal of Artificial Societies and Social Simulation*, 15(2):7.

Haas, M., Khapko, M., and Zoican, M. (2021). Speed and learning in high-frequency auctions. *Journal of Financial Markets*, 54:100583.

Haldane, A. G. and Turrell, A. E. (2019). Drawing on different disciplines: macroeconomic agent-based models. *Journal of Evolutionary Economics*, 29:39–66.

Harris, J. H. (1995). *The cost components of bid-ask spreads: an intraday analysis*. The Ohio State University.

Hasbrouck, J. (1988). Trades, quotes, inventories, and information. *Journal of financial economics*, 22(2):229–252.

Hasbrouck, J. (1991). Measuring the information content of stock trades. *The Journal of Finance*, 46(1):179–207.

Hendershott, T., Riordan, R., et al. (2009). Algorithmic trading and information. *Manuscript, University of California, Berkeley*.

Ho, T. and Stoll, H. R. (1981). Optimal dealer pricing under transactions and return uncertainty. *Journal of Financial economics*, 9(1):47–73.

Ho, T. S. and Stoll, H. R. (1983). The dynamics of dealer markets under competition. *The Journal of finance*, 38(4):1053–1074.

Hoffmann, P. (2014). A dynamic limit order market with fast and slow traders. *Journal of Financial Economics*, 113(1):156–169.

IEX Exchange (2023). Discretionary Peg™ | IEX Exchange | IEX — iexexchange.io. https://www.iexexchange.io/order-types/d-peg. [Accessed 14-10-2023].

Iori, G. and Porter, J. (2012). Agent-based modelling for financial markets.

Itô, K. (1946). On a stochastic integral equation. *Proceedings of the Japan Academy*, 22(1-4):32–35.

Jackson, M. O. (2007). The study of social networks in economics. *The missing links: Formation and decay of economic networks*, 76:210–225.

Jagannathan, R. (2022). On frequent batch auctions for stocks. *Journal of Financial Econometrics*, 20(1):1–17.

Jones, C. M. (2013). What do we know about high-frequency trading? *Columbia Business School Research Paper*, (13-11).

Jovanovic, B. and Menkveld, A. J. (2016). Middlemen in limit order markets. *Available at SSRN 1624329*.

Kirilenko, A., Kyle, A. S., Samadi, M., and Tuzun, T. (2017). The flash crash: High-frequency trading in an electronic market. *The Journal of Finance*, 72(3):967–998.

Kirman, A. (2010). The economic crisis is a crisis for economic theory. *CESifo Economic Studies*, 56(4):498–535.

Kwiatkowski, D., Phillips, P., Schmidt, P., and Shin, Y. (1992). Distribution of the estimators for autoregressive time series with a unit root. *J. Econometrics*, 54(1-3):159–178.

Kyle, A. S. (1985). Continuous auctions and insider trading. *Econometrica: Journal of the Econometric Society*, pages 1315–1335.

Lauricella, T. (2010). Debate on 'crash'and its causes. *Wall Street Journal*.

Leal, S. J., Napoletano, M., Roventini, A., and Fagiolo, G. (2016). Rock around the clock: An agent-based model of low-and high-frequency trading. *Journal of Evolutionary Economics*, 26(1):49–76.

LeBaron, B. (2001). A builder's guide to agent-based financial markets. *Quantitative finance*, 1(2):254.

LeBaron, B. (2006). Agent-based financial markets: Matching stylized facts with style. *Post Walrasian Macroeconomics: Beyond the DSGE Model*, pages 221–235.

Li, Z. and Das, S. (2016). An agent-based model of competition between financial exchanges: Can frequent call mechanisms drive trade away from cdas? In *Proceedings of the 2016 International Conference on Autonomous Agents & Multiagent Systems*, pages 50–58.

Lin, J.-C., Sanger, G. C., and Booth, G. G. (1995). Trade size and components of the bid-ask spread. *The Review of Financial Studies*, 8(4):1153–1183.

Lockett, H. (2019). Chinese stock rebound has fingerprints of retail investors.

LSE (2020). What is an auction?

Madhavan, A. (1992). Trading mechanisms in securities markets. *the Journal of Finance*, 47(2):607–641.

Madhavan, A., Richardson, M., and Roomans, M. (1997). Why do security prices change? a transaction-level analysis of nyse stocks. *The Review of Financial Studies*, 10(4):1035–1064.

McAfee, R. P. (1992). A dominant strategy double auction. *Journal of economic Theory*, 56(2):434–450.

Myerson, R. B. and Satterthwaite, M. A. (1983). Efficient mechanisms for bilateral trading. *Journal of economic theory*, 29(2):265–281.

Paddrik, M., Hayes, R., Todd, A., Yang, S., Beling, P., and Scherer, W. (2012). An agent based model of the e-mini s&p 500 applied to flash crash analysis. In *2012 IEEE Conference on Computational Intelligence for Financial Engineering & Economics (CIFEr)*, pages 1–8. IEEE.

Paulin, J., Calinescu, A., and Wooldridge, M. (2019). Understanding flash crash contagion and systemic risk: A micro–macro agent-based approach. *Journal of Economic Dynamics and Control*, 100:200–229.

Reddy, K. and Clinton, V. (2016). Simulating stock prices using geometric brownian motion: Evidence from australian companies. *Australasian Accounting, Business and Finance Journal*, 10(3):23–47.

Riccò, R. and Wang, K. (2020). Frequent batch auctions vs. continuous trading: Evidence from taiwan. *Continuous Trading: Evidence from Taiwan (November 19, 2020)*.

Roll, R. (1984). A simple implicit measure of the effective bid-ask spread in an efficient market. *The Journal of finance*, 39(4):1127–1139.

Said, S. E. and Dickey, D. A. (1984). Testing for unit roots in autoregressive-moving average models of unknown order. *Biometrika*, 71(3):599–607.

SEC and CFTC (2010). Findings regarding the market events of may 6, 2010. *Washington DC*.

Simon, H. A. (1997). *Models of bounded rationality: Empirically grounded economic reason*, volume 3. MIT press.

Sivia, D. and Skilling, J. (2006). *Data analysis: a Bayesian tutorial*. OUP Oxford.

Smith, V. L. (1962). An experimental study of competitive market behavior. *Journal of political economy*, 70(2):111–137.

Smith, V. L. (1982). Microeconomic systems as an experimental science. *The American economic review*, 72(5):923–955.

Smith, V. L. (1994). Economics in the laboratory. *Journal of Economic Perspectives*, 8(1):113–131.

Spratt, S. (1998). Heuristics for the startup problem. *Department of Systems Engineering, University of Virginia*.

Stoll, H. R. (1978). The supply of dealer services in securities markets. *The Journal of Finance*, 33(4):1133–1151.

Stoll, H. R. (1989). Inferring the components of the bid-ask spread: Theory and empirical tests. *the Journal of Finance*, 44(1):115–134.

US Securities and Exchange Comission (2010). Concept Release on Equity Market Structure; Proposed Rule 17 CFR Part 242. *Federal Register*, Part III:1–22.

Wah, E., Hurd, D., and Wellman, M. (2016). Strategic market choice: Frequent call markets vs. continuous double auctions for fast and slow traders. *EAI Endorsed Transactions on Serious Games*, 3(10).

Wah, E. and Wellman, M. P. (2016). Latency arbitrage in fragmented markets: A strategic agent-based analysis. *Algorithmic Finance*, 5(3-4):69–93.

Wahba, P. (2010). NYSE says stock plunge validates its hybrid system.

Wald, A. and Wolfowitz, J. (1940). On a test wether two samples are from the same distribution. *Ann. Math. Stat*, 11:147–162.

White Jr, K. P. (1997). An effective truncation heuristic for bias reduction in simulation output. *Simulation*, 69(6):323–334.

Wiener-Bronner, D. (2018).

Wigglesworth, R. and Martin, K. (2021). Rise of the retail army: The amateur traders transforming markets.

Wolfers, J. and Zitzewitz, E. (2004). Prediction markets. *Journal of economic perspectives*, 18(2):107–126.

Zaharudin, K. Z., Young, M. R., and Hsu, W.-H. (2022). High-frequency trading: Definition, implications, and controversies. *Journal of Economic Surveys*, 36(1):75–107.

Zhang, Z. and Ibikunle, G. (2021). Latency arbitrage and frequent batch auctions. *Available at SSRN 3984570*.

Zollman, K. J. (2007). The communication structure of epistemic communities. *Philosophy of science*, 74(5):574–587.

Zollman, K. J. (2010). The epistemic benefit of transient diversity. *Erkenntnis*, 72(1):17–35.

Appendix A

Continuous Double Auction v.s. Batch Auction in a Simple Case

To start with the Batch Auction design, we will see how the batching mechanism change or improve the market compared with the Continuous Double Auction.

A.1 Continuous Double Auction Market

Assume a double auction market where six traders have submitted their orders and wait to be matched. The details of the six traders are shown in Table A.1, and their orders are kept temporarily in the limit order book which is shown in Figure A.1. The limit order book structure of the buy-side and sell-side is symmetric. The best bid price and the best ask price are 99.9 and 100.0, respectively.

Now, there are three incoming orders entering the market, whose details are shown below:

Trader	Direction	Price	Quantity
G	B	100.1	1
H	B	100.1	1
I	A	100.0	1

Firstly, we will see the dynamics of the limit order book in the CDA market.

Trader ID	Direction	Price	Quantity
A	A	100.0	1
B	B	99.9	1
C	A	100.1	1
D	B	99.8	1
E	A	100.2	1
F	B	99.7	1

Table A.1 Existed Orders in the Market

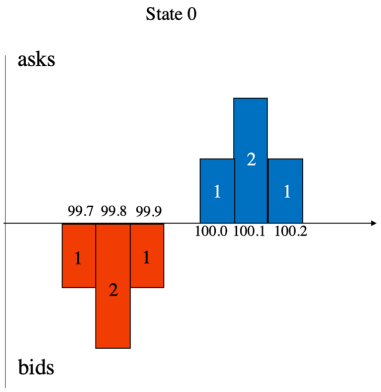


Figure A.1 The Limit Order Book before Three Orders Entering

Case 1: G $\rightarrow$ H $\rightarrow$ I (H $\rightarrow$ G $\rightarrow$ I)

If all traders are in the CDA market, every time a trader submits an order, the exchange will execute the matching mechanism immediately. The following figure shows the dynamics of the limit order book under the CDA market.

At Stage 1, trader G submits a buy order at a price of 100.1, which becomes the best bid price at the moment. Therefore, trader G's order will match with A's sell order (quote: 100.0). According to the trading rule proposed by Gode and Sunder, the trade price is 100.0. Trader G's profit is $(100.1 - 100.0) \times 1 = 0.1$, while trader A's profit is 0. Thus, the market surplus increases by 0.1. The best ask price will become 100.1.

At Stage 2, buyer H arrives, and his shout is also 100.1, Since his shout is higher than other bid prices, his bid becomes the best bid and will match with the best ask which is 100.1 at the moment. Thus the trader price is 100.1, and the profits of both traders G and A are 0. The best ask and bid prices remain to be the same as in the last stage.

At Stage 3, seller I arrives with an offer of 100.0. Since all bid prices are lower than his offer, his order is kept in the limited order book and no transaction happens at this stage. The best ask and best bid are 99.9 and 100.0, respectively.

Case 2: G $\rightarrow$ I $\rightarrow$ H (H $\rightarrow$ I $\rightarrow$ G)

Stage 1: same as Case 1.

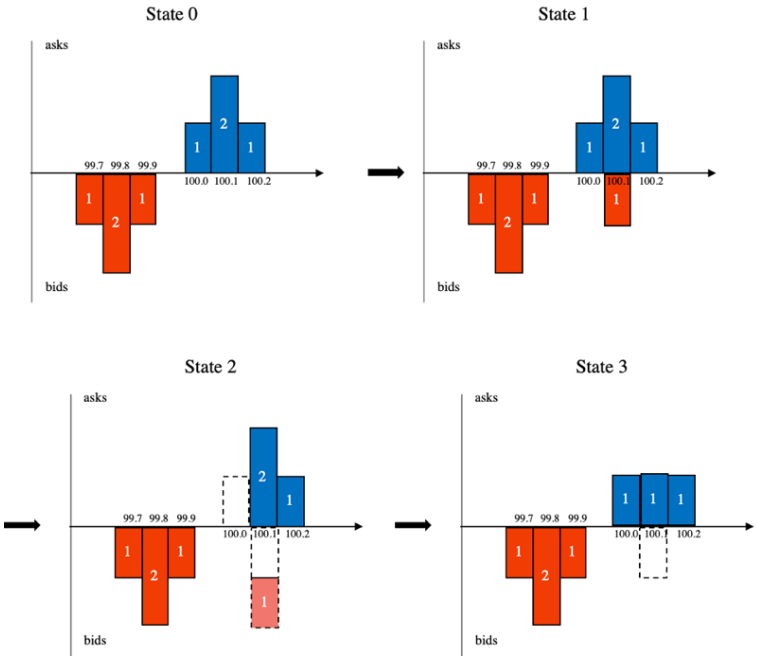


Figure A.2 The Limit Order Book Dynamics in the CDA Market: Case 1

At Stage 2, a sell order from trader I comes into the market and becomes the best ask. This order will not match with others, causing no transaction at this stage.

At Stage 3, buyer H enters and matches with seller I. H's shout is higher than the best ask, and he arrives later than seller I, so the trade price is 100.0, and traders H and I earn 0.1 and 0, respectively.

Case 3: I $\rightarrow$ G $\rightarrow$ H (I $\rightarrow$ H $\rightarrow$ G)

At Stage 1, seller I arrives, and his order is kept in the limit order book without being traded.

At Stage 2, buyer G arrives and matches with seller A. The trade price is 100.0, and buyer G's profit is 0.1. The best ask and best bid prices are not changed.

Stage 3: same as Case 2.

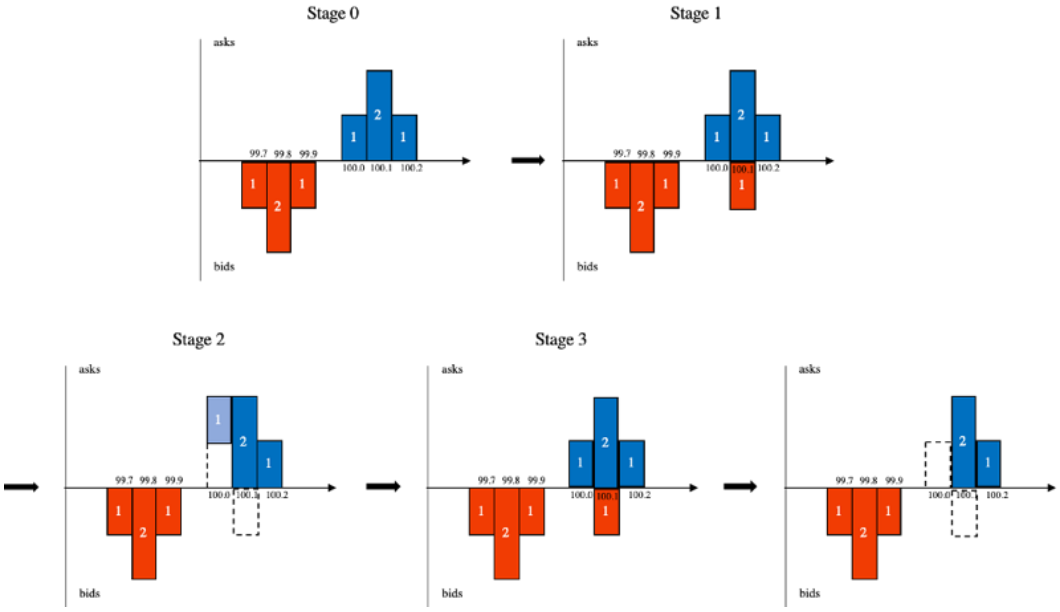


Figure A.3 The Limit Order Book Dynamics in the CDA Market: Case 2

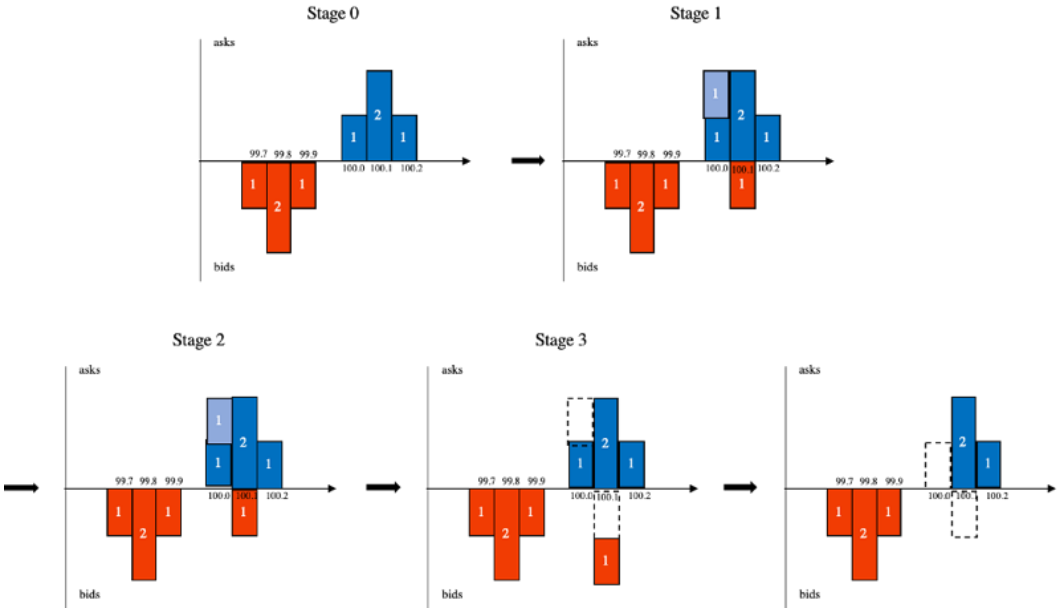


Figure A.4 The Limit Order Book Dynamics in the CDA Market: Case 3

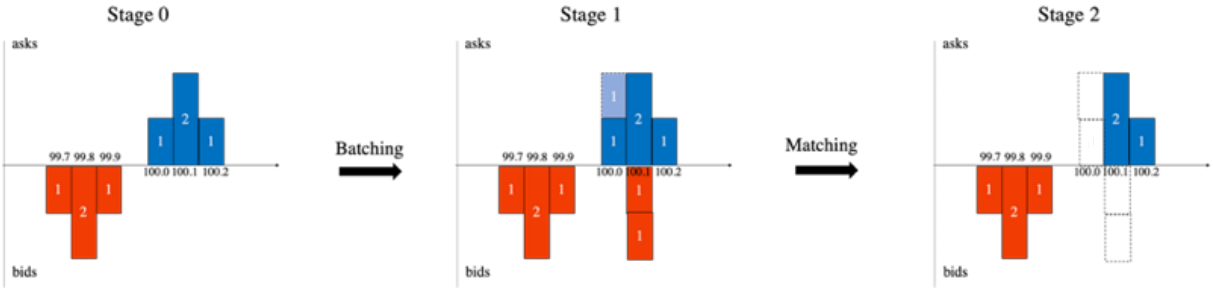


Figure A.5 The Limit Order Book Dynamics in the Batch Auction Market

A.2 Batch Double Auction Market

In the Batch Auction market, assume the orders from the traders G H and I are batched, and the matching mechanism will be executed after batching. Under this circumstance, the dynamics of the limit order book are illustrated as follows:

At Stage 0, the limit order book structure of the buy-side and sell-side is balanced. The best bid price and the best ask price are 99.9 and 100.0, respectively.

At the batching stage, all incoming offers are included in the limit order book. The exchange will batch all book orders and incoming orders under the price-time priority rule. The matching mechanism will be executed later, and there is no transaction at this stage. The best bid and best ask prices are 100.1 and 100.0, respectively.

At the matching stage, the matching mechanism is executed. Traders G and H are matched with traders A and I. Assume trader G is matched with A, while H is matched with I. Since the order of A existed before batching, the trade price between A and G is 100.0. Trader G will earn $100.1 - 100.0 = 0.1$, and trader A will earn 0. Since traders H and trader I come in this batch, the trade price is the average of the quotes of those orders. Thus the trade price is $0.5 \times (100.0 + 100.0) = 100.05$, and both traders H and I earn 0.05. The best ask and best bid prices become 99.9 and 100.1 after matching.

Table A.2 shows the simple comparison between the circumstances in the CDA and Batch Auction markets. In the Batch Auction market, the market surplus increment is no less than any circumstances in the CDA market, which means the market efficiency could be improved by the batching mechanism. Trader I earns no profit in the CDA market but

has an opportunity to earn in the BA market. However, in the BA market, the expected profits of traders G and H are lower than they can earn in the CDA market, while trader I, who has a less competitive quote, earn positive profit in the BA market. This probably indicates that under the Batch Auction mechanism, the profits can be reallocated, making the profits more balanced among traders. In the experiment, we will conduct a further experiment to investigate the batching effect.

	CDA Case 1	CDA Case 2	CDA Case 3	Batch Auction
Number of trades	2	2	2	2
Average trade price	100.05	100	100	100.025
Δ of best bid price	0	0	0	0
Δ of best ask price	0	0.1	0.1	0.1
Δ of mid price	0	0.05	0.05	0.05
Δ of market surplus	0.1	0.2	0.2	0.2
Trader G's profit	0.1(0)	0.1(0.1)	0.1(0.1)	0.1(0.05)
Trader H's profit	0(0.1)	0.1(0.1)	0.1(0.1)	0.05(0.1)
Trader I's profit	0	0	0	0.05

Table A.2 Comparison of the Two Circumstances

Appendix B

MSER-m Method

Considerable attention has been devoted to the study of steady-state and discrete-event simulation. An issue arises from the fact that a steady-state operating regime does not provide clear boundary conditions for its initiation or conclusion. To determine this point, the MSER (White Jr, 1997) and MSER-5 (Spratt, 1998) rules offer effective methods for identifying the truncation point where the standard error (test statistic) in the data is minimised.

In the Monte-Carlo simulation of the model, we run M numbers of replications of a simulation. In each replication, the market price observations could change from the initial level to a new stable level after the fake shock hits. We refer to the moment when market prices reach this new stable level as the "early stop point" (also known as the 'truncation point' in White Jr (1997)). White Jr (1997) describes an effective rule – marginal standard error rules (MSERs) – to determine truncation points. The following sections show how the algorithm works and how the early stop point is determined.

Let us assume that there are M replications of a simulation. In each replication, the data series observations may transition from their initial level to a new stable level at some point. We refer to the moment when data series reach this new stable level as the "early stop point" (also known as the "truncation point" in White Jr (1997)). Suppose that all replications have the same initial point x_0 and the same time series size $N_j = N, \forall j$. Consider the j-th replication $\{x_{ij} : i = 1,2,...,N; j = 1,2,...,M\}$, and there presumably exists the truncation point dividing the whole time series into two segments

$\{x_{ij} : i = 1,2,...,d_j;\}$ and $\{x_{ij} : i = d_j + 1,2,...,N\}$, in which d_j is the truncation point of the j-th replication. The optimal truncation point for this series is selected to minimize the MSER statistics:

$$\begin{aligned} d_j^* &= \underset{N>>d\geq 0}{\operatorname{argmin}} \left[MSER(N,d|x_0)\right] \\ &= \underset{N>>d\geq 0}{\operatorname{argmin}} \left[\frac{1}{(N-d)^2} \sum_{i=d+1}^{N} (x_{ij} - \bar{x}_{N,d}(j))^2 \right] \end{aligned} \tag{B.1}$$

where

$$\bar{x}_{N,d}(j) = \frac{1}{N-d} \sum_{i=d+1}^{N} x_{ij} \tag{B.2}$$

MSER-m works under the same scheme to the batch averages

$$y_{ij} = \frac{1}{m} \sum_{p=(i-1)m+1}^{im} x_{pj} \tag{B.3}$$

instead of a single observation x_{ij}. MSER-5, proposed by Spratt (1998), applies with batches of every 5 data points to determine the early stop point.

However, the MSER-5 statistics could be extremely small at the end of the time series, causing the misleading result of the truncation points. The reason is that the calculation of the MSER-5 is sensitive to a small number of observations. This can be easily avoided by excluding some MSER-5 statistics close to the end, and we let the last 100 statistics invalid in this thesis. That means if the iteration steps come into the last 100 steps, the data will not reach the "steady" state. Fig. B.1 shows an example of the MSER-5 method to determine the early stop point at work.

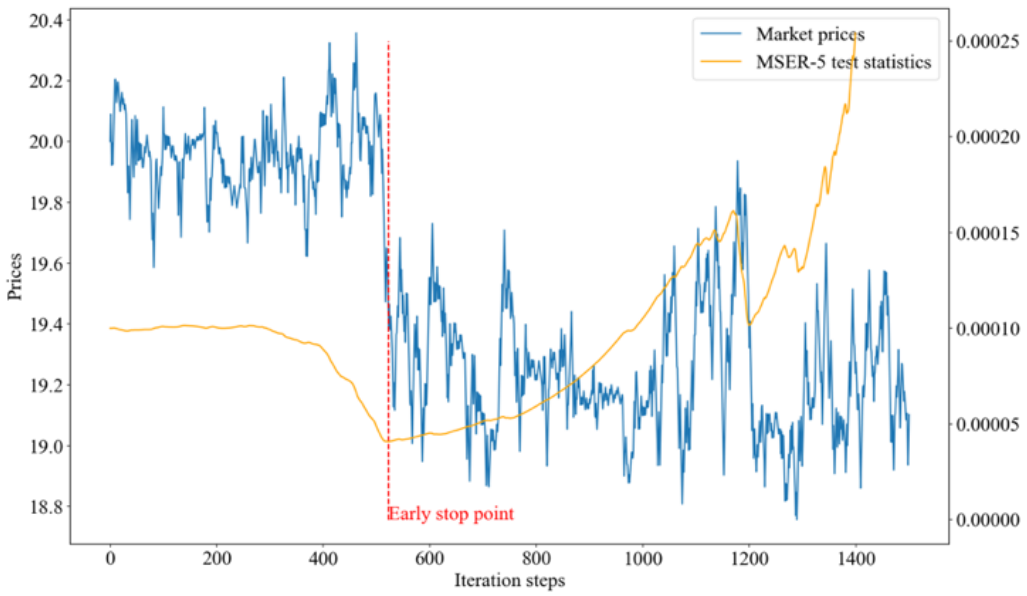


Figure B.1 An Example of the MSER-5 Method at Work

Appendix C

Introduction of Bayesian Inference

According to the definition of conditional probability,

$$P(A,B) = P(A|B)P(B) = P(B|A)P(A) \tag{C.1}$$

$$\text{i.e. } P(A|B) = \frac{P(B|A)P(A)}{P(B)} \tag{C.2}$$

The formula above is called Bayesian Theorem (Sivia and Skilling, 2006).

If we let θ and E denote the hypothesis and evidence respectively, then the Bayesian theorem can be rewritten as follows:

$$P(\theta|E) = \frac{P(E|\theta)P(\theta)}{P(E)} \tag{C.3}$$

- θ represents the hypothesis of a certain random variable, and E represents the evidence.
- P(θ) is the prior probability, which stands for the estimated probability of the hypothesis θ before the current evidence E is observed.
- P($\theta|E$) is the posterior probability representing the updated belief of the hypothesis after new evidence is observed.
- P($E|\theta$) is a conditional probability of observing the evidence E given the hypothesis θ. It is also called the likelihood.

- P(E) is called marginal probability or model evidence. This is an unconditional probability which is a constant for all possible hypotheses.
- $\frac{P(E|\theta)}{P(E)}$ stands for the impact of the evidence on the probability of the hypothesis θ.

We can apply Bayesian inference to derive the posterior by likelihood and prior. Based on the Bayesian Theorem, the posterior is proportional to the prior, where the factor is $\frac{P(E|\theta)}{P(E)}$. Prior is the probability of the hypothesis before new data comes to light, while posterior stands for the predictive probability given the data observed. We can repeatedly apply Bayesian inference by using the posterior for the prior of the next inference, which is called Bayesian learning.

Appendix D

Related Computation of Bayesian Inference

In this model, the UFTs and the market maker are endowed with the ability of Bayesian inference. The aim of the market maker is to set proper bid and ask prices to balance both the buy and sell sides. The aim of UFTs is to infer whether the shock is real or fake so that they can well estimate the true price of the stock.

D.1 The Market Maker

D.1.1 Conditional Probabilities

The market maker is modelled to be risk-neutral, so each time he shouts prices based on the following formulas:

$$\begin{cases} A_t = \mathbb{E}[V|\text{Buy}] \\ B_t = \mathbb{E}[V|\text{Sell}] \end{cases} \tag{D.1}$$

Therefore, to determine the best bid and ask prices, we have to start with the conditional expectation. Denote $\mathbb{V}_t$ as the vector space of the true prices at time t, then by the definition of the conditional expectation, taking calculating the bid price as an example, we will

obtain the expected stock price provided that the new order is a sell order:

$$\begin{aligned}
\mathbb{E}[V|\text{Sell}] &= \sum_{v\in\mathbb{V}_t} v\Pr(V_t = v|\text{Sell}) \\
\mathbb{E}[V|\text{Buy}] &= \sum_{v\in\mathbb{V}_t} v\Pr(V_t = v|\text{Buy}) \\
\mathbb{E}[V|\text{No order}] &= \sum_{v\in\mathbb{V}_t} v\Pr(V_t = v|\text{No order})
\end{aligned} \tag{D.2}$$

In some cases, there is no order coming since both fundamental traders and ZI traders decide to do nothing. By applying Bayes Theorem, we get

$$\begin{aligned}
\Pr(V_t = v|\text{Sell}) &= \frac{\Pr(\text{Sell}|V_t = v)\Pr(V_t = v)}{\Pr(\text{Sell})} \\
\Pr(V_t = v|\text{Buy}) &= \frac{\Pr(\text{Buy}|V_t = v)\Pr(V_t = v)}{\Pr(\text{Buy})} \\
\Pr(V_t = v|\text{No order}) &= \frac{\Pr(\text{No order}|V_t = v)\Pr(V_t = v)}{\Pr(\text{No order})}
\end{aligned} \tag{D.3}$$

By substituting the equation D.3 into the equation D.2, then

$$\mathbb{E}[\text{V}|\text{Sell}] = \sum_{v\in\mathbb{V}_t} v\frac{\Pr(\text{Sell}|V_t = v)\Pr(V_t = v)}{\Pr(\text{Sell})} \tag{D.4}$$

$$= \sum_{v\in\mathbb{V}_t} v\frac{\Pr(\text{Sell}|V_t = v)\Pr(V_t = v)}{\sum_{v\in\mathbb{V}_t}\Pr(\text{Sell}|V_t = v)\Pr(V_t = v)} \tag{D.5}$$

Similarly,

$$\mathbb{E}[\text{V}|\text{Buy}] = \sum_{v\in\mathbb{V}_t} v\frac{\Pr(\text{Buy}|V_t = v)\Pr(V_t = v)}{\sum_{v\in\mathbb{V}_t}\Pr(\text{Buy}|V_t = v)\Pr(V_t = v)} \tag{D.6}$$

$$\mathbb{E}[\text{V}|\text{No order}] = \sum_{v\in\mathbb{V}_t} v\frac{\Pr(\text{No order}|V_t = v)\Pr(V_t = v)}{\sum_{v\in\mathbb{V}_t}\Pr(\text{No order}|V_t = v)\Pr(V_t = v)} \tag{D.7}$$

Next, we need to calculate the conditional probability of $\Pr(\text{Sell}|V_t = v)$ and $\Pr(\text{Buy}|V_t = v)$.

A market maker has limited information. He only knows the proportions of fundamental traders and ZI traders. However, he is not able to distinguish an informed trader from an uninformed trader. Namely, all fundamental traders are identical to the market maker. At time t, when the fundamental traders receive a signal of the true stock price w_t lower than the bid price, they will place sell orders. Similarly, they will buy stocks if w_t is higher than the ask price. In other cases, the fundamental traders will be inactive. However, for ZI traders, their actions are independent of the market situation. Upon these settings, the conditional probability can be calculated as

$$\begin{aligned}\Pr(\text{Sell}|V_t = v) = \\ \Pr(\text{Sell}|V_t = v,\text{FT})\Pr(\text{FT}) + \Pr(\text{Sell}|V_t = v,\text{ZI})\Pr(\text{ZI})\end{aligned} \tag{D.8}$$

$$\begin{aligned}\Pr(\text{Buy}|V_t = v) = \\ \Pr(\text{Buy}|V_t = v,\text{FT})\Pr(\text{FT}) + \Pr(\text{Buy}|V_t = v,\text{ZI})\Pr(\text{ZI})\end{aligned} \tag{D.9}$$

For a fundamental trader, he will place the sell order only if

$$w_t = v_t + \varepsilon_t < B_t \Rightarrow \varepsilon_t < B_t - v_t$$

Similarly, the fundamental trader will buy the stock if

$$w_t = v_t + \varepsilon_t > A_t \Rightarrow \varepsilon_t > A_t - v_t$$

Thus,

$$\begin{aligned}\Pr(\text{Sell}|V_t = v,\text{FT}) &= \Pr(\varepsilon_t < B_t - v_t) \\ &= \Pr(\mathcal{N}(0,\sigma_{\varepsilon_n}) < B_t - v) \\ &= \Phi\Big(\frac{B_t - v}{\sigma_{\varepsilon_n}}\Big)\end{aligned} \tag{D.10}$$

and

$$\begin{aligned}\Pr(\text{Buy}|V_t = v,\text{FT}) &= \Pr(\varepsilon_t > A_t - v_t) \\ &= \Pr(\mathcal{N}(0,\sigma_{\varepsilon_n}) < v_t - A_t) \\ &= \Phi\Big(\frac{v_t - A_t}{\sigma_{\varepsilon_n}}\Big)\end{aligned} \tag{D.11}$$

Also, the conditional probability of a fundamental trader not taking action is

$$\begin{aligned}\Pr(\text{NO}|V_t = v, \text{FT}) &= \Pr(B_t - v_t \leq \varepsilon_t \leq A_t - v_t) \\ &= \Pr(B_t - v_t \leq \mathcal{N}(0, \sigma_{\varepsilon_n}) \leq A_t - v_t) \\ &= 1 - \Phi\Big(\frac{B_t - v}{\sigma_{\varepsilon_n}}\Big) - \Phi\Big(\frac{v_t - A_t}{\sigma_{\varepsilon_n}}\Big)\end{aligned} \tag{D.12}$$

where $\Phi(\cdot)$ refers to the cumulative distribution function of a standard normal distribution. To simplify the notation, we denote μ as ${}^{(n_{\text{IFT}}+n_{\text{UFT}})}/_{n_{\text{total}}}$ representing Pr(FT). Now we substitute the equations D.10 and D.11 into D.9 and D.8,

$$\begin{aligned}\Pr(\text{Sell}|V_t = v) &= \mu\Phi\Big(\frac{B_t - v}{\sigma_{\varepsilon_n}}\Big) + \frac{1}{2}(1-\mu) \\ \Pr(\text{Buy}|V_t = v) &= \mu\Phi\Big(\frac{v - A_t}{\sigma_{\varepsilon_n}}\Big) + \frac{1}{2}(1-\mu) \\ \Pr(\text{No order}|V_t = v) &= \mu\left(1 - \Phi\Big(\frac{v - A_t}{\sigma_{\varepsilon_n}}\Big) - \Phi\Big(\frac{B_t - v}{\sigma_{\varepsilon_n}}\Big)\right)\end{aligned} \tag{D.13}$$

D.1.2 Updating Probability Distribution

The market maker has little information about the stock price, so he has to estimate the stock price to set his quotes. Before all trades happen, the market maker has initialised the probability distribution for all possible values of the true stock prices. The market maker will apply the Bayesian learning algorithm once a new order arriving, and the probability distribution of the stock price will be repeatedly adjusted. By using the law of total probability,

$$\begin{aligned}\Pr(\text{Sell}) &= \sum_{v \in \mathbb{V}_t} \Pr(\text{Sell}|V_t = v)\Pr(V_t = v) \\ &= \sum_{v \in \mathbb{V}_t} \left(\mu\Phi\Big(\frac{B_t - v}{\sigma_{\varepsilon_n}}\Big) + \frac{1}{2}(1-\mu)\right)\Pr(V_t = v)\end{aligned} \tag{D.14}$$

Similarly,

$$\begin{aligned}\Pr(\text{Buy}) &= \sum_{v \in \mathbb{V}_t} \Pr(\text{Sell}|V_t = v)\Pr(V_t = v) \\ &= \sum_{v \in \mathbb{V}_t} \left(\mu\Phi\Big(\frac{v - A_t}{\sigma_{\varepsilon_n}}\Big) + \frac{1}{2}(1-\mu)\right)\Pr(V_t = v)\end{aligned} \tag{D.15}$$

$$\begin{aligned}\Pr(\text{No order}) &= \sum_{v\in\mathbb{V}_t} \Pr(\text{No order}|V_t = v)\Pr(V_t = v) \\ &= \sum_{v\in\mathbb{V}_t} \mu\left(1-\Phi\Big(\frac{v-A_t}{\sigma_{\varepsilon_n}}\Big)-\Phi\Big(\frac{B_t-v}{\sigma_{\varepsilon_n}}\Big)\right)\Pr(V_t = v)\end{aligned} \tag{D.16}$$

We can substitute the expressions into the equations $\Pr(V_t = v|\text{Sell})$, $\Pr(V_t = v|\text{Buy})$ and $\Pr(V_t = v|\text{No order})$ obtained in D.3. As a new order comes to light, the market maker can update the posterior of the fundamental prices $\Pr^*(V = v)$ to

$$\Pr^*(V = v) = \begin{cases} \Pr(V_t = v|\text{Sell}), & \text{if a sell order arrives} \\ \Pr(V_t = v|\text{Buy}), & \text{if a buy order arrives} \\ \Pr(V_t = v|\text{No order}), & \text{if no order arrives} \end{cases} \tag{D.17}$$

$$\begin{aligned}\Pr(V_t = v|\text{Sell}) &= \frac{\Pr(\text{Sell}|V_t = v)\Pr(V_t = v)}{\Pr(\text{Sell})} \\ &= \frac{\left(\mu\Phi\left(\frac{B_t-v}{\sigma_{\varepsilon_n}}\right)+\frac{1}{2}(1-\mu)\right)\Pr(V_t = v)}{\sum_{v\in\mathbb{V}_t}\left(\mu\Phi\left(\frac{B_t-v}{\sigma_{\varepsilon_n}}\right)+\frac{1}{2}(1-\mu)\right)\Pr(V_t = v)}\end{aligned} \tag{D.18}$$

Similarly,

$$\begin{aligned}\Pr(V_t = v|\text{Buy}) &= \frac{\Pr(\text{Buy}|V_t = v)\Pr(V_t = v)}{\Pr(\text{Buy})} \\ &= \frac{\left(\mu\Phi\left(\frac{v-A_t}{\sigma_{\varepsilon_n}}\right)+\frac{1}{2}(1-\mu)\right)\Pr(V_t = v)}{\sum_{v\in\mathbb{V}_t}\left(\mu\Phi\left(\frac{v-A_t}{\sigma_{\varepsilon_n}}\right)+\frac{1}{2}(1-\mu)\right)\Pr(V_t = v)}\end{aligned} \tag{D.19}$$

$$\begin{aligned}\Pr(V_t = v|\text{No order}) &= \frac{\Pr(\text{No order}|V_t = v)\Pr(V_t = v)}{\Pr(\text{No order})} \\ &= \frac{\mu\left(1-\Phi\left(\frac{v-A_t}{\sigma_{\varepsilon_n}}\right)-\Phi\left(\frac{B_t-v}{\sigma_{\varepsilon_n}}\right)\right)\Pr(V_t = v)}{\sum_{v\in\mathbb{V}_t}\mu\left(1-\Phi\left(\frac{v-A_t}{\sigma_{\varepsilon_n}}\right)-\Phi\left(\frac{B_t-v}{\sigma_{\varepsilon_n}}\right)\right)\Pr(V_t = v)}\end{aligned} \tag{D.20}$$

D.1.3 Updating Quotes

We have obtained that

$$\begin{cases} A_t = \mathbb{E}[V|\text{Buy}] \\ B_t = \mathbb{E}v[V|\text{Sell}] \end{cases} \tag{D.21}$$

Combined with the equations in D.2 and D.3, a market maker's quotes can be updated by the following formulas:

$$B_t = \mathbb{E}[V|\text{Sell}] = \frac{\sum_{v\in\mathbb{V}_t}\left(\mu\Phi\left(\frac{B_t-v}{\sigma_{\varepsilon_n}}\right)+\frac{1}{2}(1-\mu)\right)\Pr(V_t=v)v}{\sum_{v\in\mathbb{V}_t}\left(\mu\Phi\left(\frac{B_t-v}{\sigma_{\varepsilon_n}}\right)+\frac{1}{2}(1-\mu)\right)\Pr(V_t=v)} \tag{D.22}$$

$$A_t = \mathbb{E}[V|\text{Buy}] = \frac{\sum_{v\in\mathbb{V}_t}\left(\mu\Phi\left(\frac{v-A_t}{\sigma_{\varepsilon_n}}\right)+\frac{1}{2}(1-\mu)\right)\Pr(V_t=v)v}{\sum_{v\in\mathbb{V}_t}\left(\mu\Phi\left(\frac{v-A_t}{\sigma_{\varepsilon_n}}\right)+\frac{1}{2}(1-\mu)\right)\Pr(V_t=v)} \tag{D.23}$$

$$\mathbb{E}[V|\text{No order}] = \frac{\sum_{v\in\mathbb{V}_t}\mu\left(1-\Phi\left(\frac{v-A_t}{\sigma_{\varepsilon_n}}\right)-\Phi\left(\frac{B_t-v}{\sigma_{\varepsilon_n}}\right)\right)\Pr(V_t=v)v}{\sum_{v\in\mathbb{V}_t}\mu\left(1-\Phi\left(\frac{v-A_t}{\sigma_{\varepsilon_n}}\right)-\Phi\left(\frac{B_t-v}{\sigma_{\varepsilon_n}}\right)\right)\Pr(V_t=v)} \tag{D.24}$$

D.1.4 The Expected Stock Price

A market maker should estimate the expected stock price E$[V]$ for each time. His estimated probability density

$$\mathbb{E}[V] = \sum_{v\in\mathbb{V}_t}\Pr(V=v)v \tag{D.25}$$

the expectation can be re-expressed in the following form:

$$\begin{aligned}\mathbb{E}[V] = &\sum_{v\in\mathbb{V}_t}\Big(\Pr(V=v|\text{Sell})\Pr(\text{Sell})+\Pr(V=v|\text{Buy})\Pr(\text{Buy})\\ &+\Pr(V=v|\text{No order})\Pr(\text{No order})\Big)v\\ =&\Pr(\text{Sell})\sum_{v\in\mathbb{V}_t}\Pr(V=v|\text{Sell})v+\Pr(\text{Buy})\sum_{v\in\mathbb{V}_t}\Pr(V=v|\text{Buy})v\\ &+\Pr(\text{No order})\sum_{v\in\mathbb{V}_t}\Pr(V=v|\text{No order})v\\ =&\Pr(\text{Sell})\mathbb{E}[\text{V}|\text{Sell}]+\Pr(\text{Buy})\mathbb{E}[\text{V}|\text{Buy}]+\Pr(\text{No order})\mathbb{E}[\text{V}|\text{No order}]\end{aligned} \tag{D.26}$$

D.2 Uninformed Fundamental Traders

As uninformed trader has two beliefs on his estimate of the true stock price:

Belief 0 ($\Theta=\theta_0$):

- He was not deceived;
- All fundamental traders have unbiased estimates of the true stock price;

Belief 1 ($\Theta=\theta_1$):

- He was deceived and had a biased estimate;
- all informed fundamental traders have unbiased estimates, while the rest, the uninformed fundamental traders, are having biased estimates;

The market maker updates his beliefs according to the following formulas:

$$\begin{aligned}\Pr(\Theta=\theta_i|\text{Sell}) &= \frac{\Pr(\text{Sell}|\Theta=\theta_i)\Pr(\Theta=\theta_i)}{\Pr(\text{Sell})}\\ \Pr(\Theta=\theta_i|\text{Buy}) &= \frac{\Pr(\text{Buy}|\Theta=\theta_i)\Pr(\Theta=\theta_i)}{\Pr(\text{Buy})}\\ \Pr(\Theta=\theta_i|\text{No order}) &= \frac{\Pr(\text{No order}|\Theta=\theta_i)\Pr(\Theta=\theta_i)}{\Pr(\text{No order})}\end{aligned} \tag{D.27}$$

where i = 0 or 1. To obtain the posterior beliefs on whether the UFT is deceived, we now focus on the likelihood part $\Pr(\cdot|\Theta=\theta_i)$:

$$\begin{aligned}\Pr(\cdot|\Theta=\theta_i) =& \Pr(\cdot|\Theta=\theta_i,\text{IFT})\Pr(\text{IFT})+\\ & \Pr(\cdot|\Theta=\theta_i,\text{UFT})\Pr(\text{UFT})+\\ & \Pr(\cdot|\Theta=\theta_i,\text{ZI})\Pr(\text{ZI})\end{aligned} \tag{D.28}$$

If the new sell order is submitted by a fundamental trader, it occurs only if

$$\begin{aligned} w_{i,t}^{\text{IFT}} &= v_t^{\text{IFT}} + \varepsilon_i < B_t \Rightarrow \varepsilon_i < B_t - v_t^{\text{IFT}} \\ w_{i,t}^{\text{UFT}} &= v_t^{\text{UFT}} + \varepsilon_i < B_t \Rightarrow \varepsilon_i < B_t - v_t^{\text{UFT}} \end{aligned}$$

and the new buy order is submitted by a fundamental trader, it occurs only if

$$\begin{aligned} w_{i,t}^{\text{IFT}} &= v_t^{\text{IFT}} + \varepsilon_i > A_t \Rightarrow \varepsilon_i > A_t - v_t^{\text{IFT}} \\ w_{i,t}^{\text{UFT}} &= v_t^{\text{UFT}} + \varepsilon_i > A_t \Rightarrow \varepsilon_i > A_t - v_t^{\text{UFT}} \end{aligned}$$

To simplify the notation, we denote μ, α as ${}^{(n_{\text{IFT}}+n_{\text{UFT}})}/_{n_{\text{total}}}$ and ${}^{n_{\text{UFT}}}/_{(n_{\text{IFT}}+n_{\text{UFT}})}$, respectively. Thus, the Eq. D.27 can be expressed as follows:

$$\begin{aligned}\Pr(\text{Sell}|\Theta=\theta_i) =& \mu\alpha\Pr(\varepsilon_i < B_t - v_t^{\text{IFT}})+\\ & \mu(1-\alpha)\Pr(\varepsilon_i < B_t - v_t^{\text{UFT}}) + \frac{1}{2}(1-\mu)\\ =& \mu\alpha\Pr(\mathcal{N}(0,\sigma_i^2) < B_t - v_t^{\text{IFT}})+\\ & \mu(1-\alpha)\Pr(\mathcal{N}(0,\sigma_i^2) < B_t - v_t^{\text{UFT}}) + \frac{1}{2}(1-\mu)\\ =& \mu\alpha\Phi\Big(\frac{B_t - v_t^{\text{IFT}}}{\sigma_{\varepsilon_i}}\Big) + \mu(1-\alpha)\Phi\Big(\frac{B_t - v_t^{\text{UFT}}}{\sigma_{\varepsilon_i}}\Big) + \frac{1}{2}(1-\mu)\end{aligned} \tag{D.29}$$

Similarly,

$$\begin{aligned}\Pr(\text{Buy}|\Theta=\theta_i) =& \mu\alpha\Pr(\varepsilon_i < v_t^{\text{IFT}} - A_t) + \\ & \mu(1-\alpha)\Pr(\varepsilon_i < v_t^{\text{UFT}} - A_t) + \frac{1}{2}(1-\mu) \\ =& \mu\alpha\Pr(\mathcal{N}(0,\sigma_i^2) < v_t^{\text{IFT}} - A_t) + \\ & \mu(1-\alpha)\Pr(\mathcal{N}(0,\sigma_i^2) < v_t^{\text{UFT}} - A_t) + \frac{1}{2}(1-\mu) \\ =& \mu\alpha\Phi\Big(\frac{v_t^{\text{IFT}} - A_t}{\sigma_{\varepsilon_i}}\Big) + \mu(1-\alpha)\Phi\Big(\frac{v_t^{\text{UFT}} - A_t}{\sigma_{\varepsilon_i}}\Big) + \frac{1}{2}(1-\mu)\end{aligned} \tag{D.30}$$

and

$$\Pr(\text{No order}|\Theta=\theta_i) = 1 - \Pr(\text{Sell}|\Theta=\theta_i) - \Pr(\text{Buy}|\Theta=\theta_i) \tag{D.31}$$

With the law of total probability,

$$\Pr(\text{Sell}) = \sum_{j=\{0,1\}} \Pr(\text{Sell}|\Theta=\theta_j)\Pr(\Theta=\theta_j) \tag{D.32}$$

$$\Pr(\text{Buy}) = \sum_{j=\{0,1\}} \Pr(\text{Buy}|\Theta=\theta_j)\Pr(\Theta=\theta_j) \tag{D.33}$$

$$\Pr(\text{No order}) = \sum_{j=\{0,1\}} \Pr(\text{No order}|\Theta=\theta_j)\Pr(\Theta=\theta_j) \tag{D.34}$$

Therefore,

$$\Pr^*(\Theta=\theta_i) = \begin{cases} \dfrac{\Pr(\text{Sell}|\Theta=\theta_i)\Pr(\Theta=\theta_i)}{\sum_{j=\{0,1\}}\Pr(\text{Sell}|\Theta=\theta_j)\Pr(\Theta=\theta_j)}, & \text{if a sell order arrives} \\ \dfrac{\Pr(\text{Buy}|\Theta=\theta_0)\Pr(\Theta=\theta_i)}{\sum_{j=\{0,1\}}\Pr(\text{Buy}|\Theta=\theta_j)\Pr(\Theta=\theta_j)}, & \text{if a buy order arrives} \\ \dfrac{\Pr(\text{No order}|\Theta=\theta_0)\Pr(\Theta=\theta_i)}{\sum_{i=\{0,1\}}\Pr(\text{No order}|\Theta=\theta_j)\Pr(\Theta=\theta_j)}, & \text{if no order arrives} \end{cases} \tag{D.35}$$

Appendix E

Maths Regarding Batch Auction Markets in Messy Networks

E.1 The Market Maker's Quoting

In the n-agent messy network under the k-size batch auction mechanism, the market maker updates his quotes every other k trading rounds. The market maker is assumed to be risk-neutral, which means that the market maker updates the quotes based on the zero-expected-profit rule. Denote the numbers of buy orders and sell orders as Q^b and Q^s, respectively. Hence, the expected profits of the market maker after the i-th batch can be described mathematically as follows:

$$EP_i = \begin{cases} \mathbb{E}[A_i - V_i] \cdot (Q_i^b - Q_i^s), \text{ if } Q_i^b > Q_i^s \\ \mathbb{E}[V_i - B_i] \cdot (Q_i^s - Q_i^b), \text{ if } Q_i^b < Q_i^s \\ 0, \qquad\qquad\qquad\qquad \text{if } Q_n^{Buy} = Q_n^{Sell} \end{cases} \tag{E.1}$$

$$\begin{cases} A_{i+1} = \mathbb{E}[V_i | Q_i^b = m, Q_i^s = n], \text{ if } m > n \\ B_{i+1} = \mathbb{E}[V_i | Q_i^b = m, Q_i^s = n], \text{ if } m < n \end{cases} \tag{E.2}$$

The updated bid and ask prices are applied to the outstanding orders after the next batch.

Again, to determine the updated quotes, we must start with the conditional expectation on the right-hand side of the equations E.2.

$$\mathbb{E}[V_i|Q_i^b = m, Q_i^s = n] = \sum_{v\in\mathbb{V}_n} v\mathrm{Pr}_{\mathrm{MM}}(V_i = v|Q_i^b = m, Q_i^s = n) \tag{E.3}$$

To compute the conditional probability $\mathrm{Pr}_{\mathrm{MM}}(V_i = v|Q_i^b = m, Q_i^s = n)$, we can simply use the Bayesian Theorem leading to

$$\begin{aligned}
&\mathrm{Pr}_{\mathrm{MM}}(V_i = v|Q_i^b = m, Q_i^s = n) \\
&\qquad = \frac{\mathrm{Pr}_{\mathrm{MM}}(Q_i^b = m, Q_i^s = n|V_i = v)\mathrm{Pr}(V_i = v)}{\mathrm{Pr}_{\mathrm{MM}}(Q_i^b = m, Q_i^s = n} \\
&\qquad = \frac{\mathrm{Pr}_{\mathrm{MM}}(Q_i^b = m, Q_i^s = n|V_i = v)\mathrm{Pr}(V_i = v)}{\sum_{v\in\mathbb{V}_i}\mathrm{Pr}_{\mathrm{MM}}(Q_i^b = m, Q_i^s = n|V_i = v)\mathrm{Pr}(V_i = v)}
\end{aligned} \tag{E.4}$$

The ask and bid prices can be updated based on the following formulas:

$$\mathbb{E}[V_i|Q_i^b = m, Q_i^s = n] = \frac{\sum_{v\in\mathbb{V}_i}\mathrm{Pr}_{\mathrm{MM}}(Q_i^b = m, Q_i^s = n|V_i = v)\mathrm{Pr}(V_i = v)v}{\sum_{v\in\mathbb{V}_i}\mathrm{Pr}_{\mathrm{MM}}(Q_i^b = m, Q_i^s = n|V_i = v)\mathrm{Pr}(V_i = v)} \tag{E.5}$$

As soon as the batched orders are released to the public, the market maker will update the probability estimation of the stock prices by the Bayesian learning method shown in Eq. E.4. Denote the batch size as k, k is fixed during a single simulation, and $k \geq 2$. Therefore, the conditional probability $\mathrm{Pr}_{\mathrm{MM}}(Q_i^b = m, Q_i^s = n|V_i = v)$ is the key part for updating the bid and ask prices as well as the probability estimation of stock prices. To derive the presentation, we employ the conditional probabilities obtained in the Appendix D.13:

$$\begin{aligned}
\mathrm{Pr}_{\mathrm{MM}}(\mathrm{Sell}|V_n = v) &= \mu\Phi\Big(\frac{B_n - v}{\sigma_\varepsilon}\Big) + \frac{1}{2}(1-\mu) \\
\mathrm{Pr}_{\mathrm{MM}}(\mathrm{Buy}|V_n = v) &= \mu\Phi\Big(\frac{v - A_n}{\sigma_\varepsilon}\Big) + \frac{1}{2}(1-\mu) \\
\mathrm{Pr}_{\mathrm{MM}}(\mathrm{No\ order}|V_n = v) &= \mu\left(1 - \Phi\Big(\frac{v - A_n}{\sigma_\varepsilon}\Big) - \Phi\Big(\frac{B_n - v}{\sigma_\varepsilon}\Big)\right)
\end{aligned} \tag{E.6}$$

The conditional probabilities given a certain market price value can be regarded as the functions of the market price. Therefore, to simplify the computation, we denote the

different conditional probabilities of a sell order and a buy order arriving as β and γ, then

$$\begin{aligned}
&\Pr_{\mathrm{MM}}(\text{Sell}|V_i = v) = \beta_i(v) \\
&\Pr_{\mathrm{MM}}(\text{Buy}|V_i = v) = \gamma_i(v) \\
&\Pr_{\mathrm{MM}}(\text{No order}|V_i = v) = 1 - \beta_i(v) - \gamma_i(v)
\end{aligned} \tag{E.7}$$

where β and γ are the functions of the true stock prices v. If an uninformed fundamental trader observes that there are m buy orders and n sell orders in a batch, then such a case refers to $\binom{k}{m}\binom{k-m}{n}$ possible occurrences. Thus, we can straightly get the likelihood $\Pr_{\mathrm{MM}}(Q_i^b = m, Q_i^s = n|V_i = v)$:

$$\Pr_{\mathrm{MM}}(Q_i^b = m, Q_i^s = n|V_i = v) = \binom{k}{m}\binom{k-m}{n}\gamma_i^m(v)\beta_i^n(v)(1-\beta_i(v)-\gamma_i(v))^{k-m-n} \tag{E.8}$$

By using the law of total probability,

$$\begin{aligned}
&\Pr_{\mathrm{MM}}(Q_i^b = m, Q_i^s = n) = \\
&\qquad \sum_{v\in\mathbb{V}_n}\binom{k}{m}\binom{k-m}{n}\gamma_i^m(v)\beta_i^n(v)(1-\beta_i(v)-\gamma_i(v))^{k-m-n}\Pr_{\mathrm{MM}}(V_n = v)
\end{aligned} \tag{E.9}$$

where $\Pr_{\mathrm{MM}}(V_n = v)$ is referred to the prior probability distribution function. Then, the posterior probability calculated according to Eq. E.4 can be written as

$$\begin{aligned}
&\Pr_{\mathrm{MM}}(V_i = v|Q_i^b = m, Q_i^s = n) \\
&\qquad = \frac{\binom{k}{m}\binom{k-m}{n}\gamma_i^m(v)\beta_i^n(v)(1-\beta_i(v)-\gamma_i(v))^{k-m-n}\Pr(V_i = v)}{\sum_{v\in\mathbb{V}_n}\binom{k}{m}\binom{k-m}{n}\gamma_i^m(v)\beta_i^n(v)(1-\beta_i(v)-\gamma_i(v))^{k-m-n}\Pr_{\mathrm{MM}}(V_i = v)}
\end{aligned} \tag{E.10}$$

The derived posterior probability distributed function will be applied as the prior for the updating quotes after the next batch of trading. The zero-profit market maker always maintains the quotes equal to the conditional expected market prices, then the conditional expectation that is shown in Eq. E.5 is re-written as

$$\mathbb{E}[V|Q_i^b = m, Q_i^s = n] = \frac{\sum_{v\in\mathbb{V}_n}\binom{k}{m}\binom{k-m}{n}\gamma_i^m(v)\beta_i^n(v)(1-\beta_i(v)-\gamma_i(v))^{k-m-n}\mathrm{Pr}_{\mathrm{MM}}(V_n=v)v}{\sum_{v\in\mathbb{V}_n}\binom{k}{m}\binom{k-m}{n}\gamma_i^m(v)\beta_i^n(v)(1-\beta_i(v)-\gamma_i(v))^{k-m-n}\mathrm{Pr}_{\mathrm{MM}}(V_n=v)} \tag{E.11}$$

Also, the expectation of the true stock price could be expressed as

$$\mathbb{E}(V) = \sum_{n=0}^{k}\sum_{m=0}^{k-n}\mathrm{Pr}_{\mathrm{MM}}(Q_i^b = m, Q_i^s = n)\mathrm{E}[\mathrm{V}|Q_i^b = m, Q_i^s = n] \tag{E.12}$$

E.2 The Uninformed Fundamental Traders

Same as the settings in Chapter 5, the uninformed fundamental traders do not know the private information but are allowed to adjust their beliefs by observing past order flows. However, due to the batching mechanism, the order flow information is only broadcast at the end of each trading window, so the UFTs adjust beliefs less frequently but spend a longer time on inference in the batch auction market.

An uninformed trader has two beliefs on his estimate of the true stock price:

Belief 0 ($\Theta = \theta_0$):

- He was not deceived;
- All fundamental traders have unbiased estimates of the true stock price;

Belief 1 ($\Theta = \theta_1$):

- He was deceived and had a biased estimate;
- All informed fundamental traders have unbiased estimates, while the rest of all uninformed fundamental traders have biased estimates;

According to the Bayesian Learning formula in Eq. 6.3, an uninformed traders update his beliefs as follows:

$$\Pr(\Theta=\theta_0|Q_i^b=m,Q_i^s=n)=\frac{\Pr(Q_i^b=m,Q_i^s=n|\Theta=\theta_0)\Pr(\Theta=\theta_0)}{\Pr(Q_i^b=m,Q_i^s=n)} \tag{E.13}$$

Now we focus on the likelihood part. As we have got the conditional probability upon different beliefs in the CDA market in the Appendix D.31:

$$\begin{aligned}
\Pr(\text{Sell}|\Theta=\theta_i)&=\mu\alpha\Phi\Big(\frac{B_t-v_t^{\text{IFT}}}{\sigma_{\varepsilon_i}}\Big)+\mu(1-\alpha)\Phi\Big(\frac{B_t-v_t^{\text{UFT}}}{\sigma_{\varepsilon_i}}\Big)+\frac{1}{2}(1-\mu)\\
\Pr(\text{Buy}|\Theta=\theta_i)&=\mu\alpha\Phi\Big(\frac{v_t^{\text{IFT}}-A_t}{\sigma_{\varepsilon_i}}\Big)+\mu(1-\alpha)\Phi\Big(\frac{v_t^{\text{UFT}}-A_t}{\sigma_{\varepsilon_i}}\Big)+\frac{1}{2}(1-\mu)\\
\Pr(\text{No order}|\Theta=\theta_0)&=1-\Pr(\text{Sell}|\Theta=\theta_0)-\Pr(\text{Buy}|\Theta=\theta_0)
\end{aligned} \tag{E.14}$$

To simplify the computation, we denote

$$\begin{aligned}
\Pr(\text{Sell}|\Theta=\theta_p)&=\beta_{\theta_p}\\
\Pr(\text{Buy}|\Theta=\theta_p)&=\gamma_{\theta_p}\\
\Pr(\text{No order}|\Theta=\theta_p)&=1-\beta_{\theta_p}-\gamma_{\theta_p}
\end{aligned} \tag{E.15}$$

where $p\in\{0,1\}$. Similar to what we have obtain in the equation Eq. E.9, we can depose $\Pr(Q_i^b=m,Q_i^s=n|\Theta=\theta_p)$ as

$$\begin{aligned}
&\Pr(Q_i^b=m,Q_i^s=n|\Theta=\theta_p)=\\
&\quad\binom{k}{m}\binom{k-m}{n}\gamma_{\theta_p,i}^m\beta_{\theta_p,i}^n(1-\beta_{\theta_p,i}-\gamma_{\theta_p,i})^{k-m-n}
\end{aligned} \tag{E.16}$$

By using the law of total probability,

$$\begin{aligned}
&\Pr_{\text{UFT}}(Q_i^b=m,Q_i^s=n)\\
&\quad=\sum_{p\in\{0,1\}}\Pr_{\text{UFT}}(Q_i^b=m,Q_i^s=n|\Theta=\theta_p)\Pr_{\text{UFT}}(\Theta=\theta_p)\\
&\quad=\sum_{p\in\{0,1\}}\binom{k}{m}\binom{k-m}{n}\gamma_{\theta_p,i}^m\beta_{\theta_p,i}^n(1-\beta_{\theta_p,i}-\gamma_{\theta_p,i})^{k-m-n}\Pr_{\text{UFT}}(\Theta=\theta_p)
\end{aligned} \tag{E.17}$$

Therefore, the uninformed traders can probability distribution according to the following formula:

$$\begin{aligned}
&\Pr^{*}_{\mathrm{UFT}}(\Theta=\theta_j) \\
&\quad=\frac{\binom{k}{m}\binom{k-m}{n}\gamma^{m}_{\theta_j,i}\beta^{n}_{\theta_j,i}(1-\beta_{\theta_j,i}-\gamma_{\theta_j,i})^{k-m-n}\Pr_{\mathrm{UFT}}(\Theta=\theta_j)}{\sum_{p\in\{0,1\}}\binom{k}{m}\binom{k-m}{n}\gamma^{m}_{\theta_p,i}\beta^{n}_{\theta_p,i}(1-\beta_{\theta_p,i}-\gamma_{\theta_p,i})^{k-m-n}\Pr_{\mathrm{UFT}}(\Theta=\theta_p)}
\end{aligned} \tag{E.18}$$

Appendix F

Maths Regarding Learning Strategy in Cycle Networks

F.1 Bayesian Learning Formula

As Chapter 4 specifies, in each trading round, there are three possible actions $\mathscr{A}_1, \mathscr{A}_2, \mathscr{A}_3 \in \mathbb{A}$ for a fundamental trader. A fundamental trader can take the following actions at each trading round:

1. $\mathscr{A}_1$: Submit a buy order, if $w_i' > A_i$;
2. $\mathscr{A}_2$: Submit a sell order, if $w_i' < B_i$;
3. $\mathscr{A}_3$: Submit a null order – does not submit an order, if $B_i \leq w_i' \leq A_i$;

In the n-agent cycle network, assume that the numbers of informed fundamental traders n_{IFT} and uninformed fundamental traders n_{UFT} are a and b, respectively. $I_{i,t} = \{\mathbb{A}_{x_t^*} = \mathscr{A}_i, \mathbb{A}_{x_{i-1,t}} = \mathscr{A}_k, \mathbb{A}_{x_{i+1,t}} = \mathscr{A}_l\}$ if $x_t^* \neq x_{i,t}$, or $I_{i,t} = \{\mathbb{A}_{x_{i-1,t}} = \mathscr{A}_k, \mathbb{A}_{x_{i+1,t}} = \mathscr{A}_l\}$ if $x_t^* = x_{i,t}$, where $\mathbb{A}_{x_{i-1}}$ and $\mathbb{A}_{x_{i+1}}$ refer to the two neighbours of the agent i.

Let $\mathbf{\Omega} = \{\Omega_{\text{IFT}}, \Omega_{\text{UFT}}, \Omega_{\text{ZIT}}\}$. In order to calculate the posterior of Eq. 5.2,

(Referred to Eq. 5.2)

$$P_{x_j,t}(\Theta = \theta_i | I_{i,t}) = \frac{P_{x_i,t}(I_{i,t}|\Theta = \theta_j) P_{x_i,t}(\Theta = \theta_j)}{\sum_{j=1,2} P_{x_i,t}(I_{i,t}|\Theta = \theta_j) P_{x_i,t}(\Theta = \theta_j)} \tag{F.1}$$

we need to obtain the likelihood by the following formula:

$$\begin{aligned}
&P_{x_i,t}(I_{i,t}|\Theta=\theta_j)\\
=&\sum_{\Omega_1\in\mathbf{\Omega}}\sum_{\Omega_2\in\mathbf{\Omega}} P_{x_i,t}(I_{i,t}|\Theta=\theta_j, x_{i-1,t}\in\Omega_1, x_{i+1,t}\in\Omega_2)P(x_{i-1,t}\in\Omega_1, x_{i+1,t}\in\Omega_2|\Theta=\theta_j)
\end{aligned} \tag{F.2}$$

The right-hand side of the equation F.2 consists of two terms: the probability of the information occurs $P_{x_i,t}(I_{i,t}|\Theta=\theta_j, x_{i-1,t}\in\Omega_1, x_{i+1,t}\in\Omega_2)$ and the probability of the neighbours with a certain allocation $P(x_{i-1,t}\in\Omega_1, x_{i+1,t}\in\Omega_2|\Theta=\theta_j)$.

F.1.1 The First Term

The first term can be decomposed as follows:

If $x_t^*\neq x_{i,t}$,

$$\begin{aligned}
&P_{x_i,t}\left(I_{i,t}|\Theta=\theta_0, x_{i-1,t}\in\Omega_1, x_{i+1,t}\in\Omega_2\right)\\
=&P_{x_i,t}\left(\mathbb{A}_{x_t^*}=\mathscr{A}_i|\Theta=\theta_j, x_{i-1,t}\in\Omega_1, x_{i+1,t}\in\Omega_2, \mathbb{A}_{x_{i-1,t}}=\mathscr{A}_k, \mathbb{A}_{x_{i+1,t}}=\mathscr{A}_l\right)\\
&\cdot P(\mathbb{A}_{x_{i-1,t}}=\mathscr{A}_k, \mathbb{A}_{x_{i+1,t}}=\mathscr{A}_l|\Theta=\theta_j, x_{i-1,t}\in\Omega_1, x_{i+1,t}\in\Omega_2)\\
=&P_{x_i,t}\left(\mathbb{A}_{x_t^*}=\mathscr{A}_i|\Theta=\theta_j, x_{i-1,t}\in\Omega_1, x_{i+1,t}\in\Omega_2, \mathbb{A}_{x_{i-1,t}}=\mathscr{A}_k, \mathbb{A}_{x_{i+1,t}}=\mathscr{A}_l\right)\\
&\cdot P(\mathbb{A}_{x_{i-1,t}}=\mathscr{A}_k|\Theta=\theta_j, x_{i-1,t}\in\Omega_1)P(\mathbb{A}_{x_{i+1,t}}=\mathscr{A}_l|\Theta=\theta_j, x_{i-1,t}\in\Omega_2)
\end{aligned} \tag{F.3}$$

If $x_t^*=x_{i,t}$,

$$\begin{aligned}
&P_{x_i,t}\left(I_{i,t}|\Theta=\theta_j, x_{i-1,t}\in\Omega_1, x_{i+1,t}\in\Omega_2\right)\\
=&P(\mathbb{A}_{x_{i-1,t}}=\mathscr{A}_k, \mathbb{A}_{x_{i+1,t}}=\mathscr{A}_l|\Theta=\theta_j, x_{i-1,t}\in\Omega_1, x_{i+1,t}\in\Omega_2)\\
=&P(\mathbb{A}_{x_{i-1,t}}=\mathscr{A}_k|\Theta=\theta_j, x_{i-1,t}\in\Omega_1)P(\mathbb{A}_{x_{i+1,t}}=\mathscr{A}_l|\Theta=\theta_j, x_{i-1,t}\in\Omega_2)
\end{aligned} \tag{F.4}$$

This equations hold since all agents' actions are independent of each other. More detailedly, given the belief $\Theta=\theta_0$,

$$
P(\mathbb{A}_{x_{p,t}} = \mathscr{A}_1 | \Theta = \theta_0, x_{p,t} \in \mathbf{\Omega}) = \begin{cases} \Phi\left(\frac{v'_t - A_t}{\sigma_{\varepsilon_n}}\right), & \text{if } x_{p,t} \in \Omega_{\text{IFT}} \\ \Phi\left(\frac{v'_t - A_t}{\sigma_{\varepsilon_n}}\right), & \text{if } x_{p,t} \in \Omega_{\text{UFT}} \\ \frac{1}{2}, & \text{if } x_{p,t} \in \Omega_{\text{ZIT}} \end{cases}
$$

$$
P(\mathbb{A}_{x_{p,t}} = \mathscr{A}_2 | \Theta = \theta_0, x_{p,t} \in \mathbf{\Omega}) = \begin{cases} \Phi\left(\frac{B_t - v'_t}{\sigma_{\varepsilon_n}}\right), & \text{if } x_{p,t} \in \Omega_{\text{IFT}} \\ \Phi\left(\frac{B_t - v'_t}{\sigma_{\varepsilon_n}}\right), & \text{if } x_{p,t} \in \Omega_{\text{UFT}} \\ \frac{1}{2}, & \text{if } x_{p,t} \in \Omega_{\text{ZIT}} \end{cases} \tag{F.5}
$$

$$
P(\mathbb{A}_{x_{p,t}} = \mathscr{A}_3 | \Theta = \theta_0, x_{p,t} \in \mathbf{\Omega}) = \begin{cases} \Phi\left(\frac{A_t - v'_t}{\sigma_{\varepsilon_n}}\right) - \Phi\left(\frac{B_t - v'_t}{\sigma_{\varepsilon_n}}\right), & \text{if } x_{p,t} \in \Omega_{\text{IFT}} \\ \Phi\left(\frac{A_t - v'_t}{\sigma_{\varepsilon_n}}\right) - \Phi\left(\frac{B_t - v'_t}{\sigma_{\varepsilon_n}}\right), & \text{if } x_{p,t} \in \Omega_{\text{UFT}} \\ 0, & \text{if } x_{p,t} \in \Omega_{\text{ZIT}} \end{cases}
$$

When belief $\Theta = \theta_1$,

$$
P(\mathbb{A}_{x_{p,t}} = \mathscr{A}_1 | \Theta = \theta_1, x_{p,t} \in \mathbf{\Omega}) = \begin{cases} \Phi\left(\frac{v_t - A_t}{\sigma_{\varepsilon_n}}\right), & \text{if } \mathbb{A}_{x_{p,t}} \in \Omega_{\text{IFT}} \\ \Phi\left(\frac{v'_t - A_t}{\sigma_{\varepsilon_n}}\right), & \text{if } x_{p,t} \in \Omega_{\text{UFT}} \\ \frac{1}{2}, & \text{if } x_{p,t} \in \Omega_{\text{ZIT}} \end{cases}
$$

$$
P(\mathbb{A}_{x_{p,t}} = \mathscr{A}_2 | \Theta = \theta_1, x_{p,t} \in \mathbf{\Omega}) = \begin{cases} \Phi\left(\frac{B_t - v_t}{\sigma_{\varepsilon_n}}\right), & \text{if } x_{p,t} \in \Omega_{\text{IFT}} \\ \Phi\left(\frac{B_t - v'_t}{\sigma_{\varepsilon_n}}\right), & \text{if } x_{p,t} \in \Omega_{\text{UFT}} \\ \frac{1}{2}, & \text{if } x_{p,t} \in \Omega_{\text{ZIT}} \end{cases} \tag{F.6}
$$

$$
P(\mathbb{A}_{x_{p,t}} = \mathscr{A}_3 | \Theta = \theta_1, x_{p,t} \in \mathbf{\Omega}) = \begin{cases} \Phi\left(\frac{A_t - v_t}{\sigma_{\varepsilon_n}}\right) - \Phi\left(\frac{B_t - v_t}{\sigma_{\varepsilon_n}}\right), & \text{if } x_{p,t} \in \Omega_{\text{IFT}} \\ \Phi\left(\frac{A_t - v'_t}{\sigma_{\varepsilon_n}}\right) - \Phi\left(\frac{B_t - v'_t}{\sigma_{\varepsilon_n}}\right), & \text{if } x_{p,t} \in \Omega_{\text{UFT}} \\ 0, & \text{if } x_{p,t} \in \Omega_{\text{ZIT}} \end{cases}
$$

To fully obtain the expression of the first term, we need to calculate the probability $P_{x_i,t}\left(\mathbb{A}_{x_t^*} = \mathscr{A}_i | \Theta = \theta_j, x_{i-1,t} \in \Omega_1, x_{i+1,t} \in \Omega_2, \mathbb{A}_{x_{i-1,t}} = \mathscr{A}_k, \mathbb{A}_{x_{i+1,t}} = \mathscr{A}_l\right)$, which is shown in the next section.

F.1.2 The Second Term

Initial Probabilities

At the beginning of learning, the probability $P(x_{i-1,t} \in \Omega_1, x_{i+1,t} \in \Omega_2)$ is only based on the number of the informed fundamental traders $n_{\text{IFT}} = a$ and the uninformed fundamental traders $n_{\text{UFT}} = b$.

$$P(x_{i-1,t} \in \Omega_1, x_{i+1,t} \in \Omega_2) = \begin{cases} \frac{\binom{n-3}{a-2}}{\binom{n-1}{a}} = \frac{a(a-1)}{(n-1)(n-2)} & \text{if } \Omega_1 = \Omega_2 = \Omega_{\text{IFT}} \\ \frac{\binom{n-3}{b-3}}{\binom{n-1}{b-1}} = \frac{(b-1)(b-2)}{(n-1)(n-2)} & \text{if } \Omega_1 = \Omega_2 = \Omega_{\text{UFT}} \\ \frac{\binom{n-3}{b-2}\binom{n-1-b}{a-1}}{\binom{n-1}{b-1}\binom{n-b}{a}} = \frac{a(b-1)}{(n-1)(n-2)} & \text{if } \Omega_1 = \Omega_{\text{UFT}}, \Omega_2 = \Omega_{\text{IFT}} \\ & (\text{or } \Omega_1 = \Omega_{\text{IFT}}, \Omega_2 = \Omega_{\text{UFT}}) \\ \frac{\binom{n-3}{a-1}\binom{n-2-a}{b-1}}{\binom{n-1}{a}\binom{n-1-a}{b-1}} = \frac{a(n-a-b)}{(n-1)(n-2)} & \text{if } \Omega_1 = \Omega_{\text{IFT}}, \Omega_2 = \Omega_{\text{ZIT}} \\ & (\text{or } \Omega_1 = \Omega_{\text{ZIT}}, \Omega_2 = \Omega_{\text{IFT}}) \\ \frac{\binom{n-3}{b-2}\binom{n-1-b}{a}}{\binom{n-1}{a}\binom{n-1-a}{b-1}} = \frac{(b-1)(n-a-b)}{(n-1)(n-2)} & \text{if } \Omega_1 = \Omega_{\text{UFT}}, \Omega_2 = \Omega_{\text{ZIT}} \\ & (\text{or } \Omega_1 = \Omega_{\text{ZIT}}, \Omega_2 = \Omega_{\text{UFT}}) \\ \frac{\binom{n-3}{a+b-1}}{\binom{n-1}{a+b-1}} = \frac{(n-a-b)(n-1-a-b)}{(n-1)(n-2)} & \text{if } \Omega_1 = \Omega_2 = \Omega_{\text{ZIT}} \end{cases} \tag{F.7}$$

Updating Probabilities in the Following Trading Rounds

As the trading activities run, the i-th agent can learn to know which party of his neighbours are likely to be from their actions each trading round. The learning formula can be derived by Bayesian Inference:

The above probability can be solved out in combination with the equations shown in F.5, F.6 and F.7.

F.2 More Detailed Maths in Difference Cases

This section we will discuss the probability $P_{x_i,t}\left(\mathbb{A}_{x_t^*}=\mathscr{A}_i|\Theta=\theta_j, x_{i-1,t}\in\Omega_1, x_{i+1,t}\in\Omega_2, \mathbb{A}_{x_{i-1,t}}=\mathscr{A}_k, \mathbb{A}_{x_{i+1,t}}=\mathscr{A}_l\right)$ in different cases. The i-th trader needs to update above probability only if the he is not selected to match with the market maker, i.e., $x_{i,t}\neq x_t^*$

F.2.1 Belief $\Theta=\theta_0$

When $\Theta=\theta_0$, the uninformed traders believe that all fundamental traders are indifferent that will react to adjust their estimates on stock price to v' when watching the shock.

Case 1: $\Omega_1,\Omega_2\in\{\Omega_{\mathbf{IFT}},\Omega_{\mathbf{UFT}}\}$

Action 1[1]:

$$\begin{aligned}&P_{x_i,t}\left(\mathbb{A}_{x_t^*}=\mathscr{A}_1|\Theta=\theta_0, x_{i-1,t}\in\Omega_1, x_{i+1,t}\in\Omega_2, \mathbb{A}_{x_{i-1,t}}=\mathscr{A}_k, \mathbb{A}_{x_{i+1,t}}=\mathscr{A}_l\right)\\ =&\frac{a+b-3}{n-1}\Phi\Big(\frac{v_t'-A_t}{\sigma_{\varepsilon_n}}\Big)+\frac{1}{2}\frac{n-a-b}{n-1}+\frac{1}{n-1}(\delta_{\mathscr{A}_k=\mathscr{A}_1}+\delta_{\mathscr{A}_l=\mathscr{A}_1})\end{aligned}\tag{F.8}$$

Action 2:

$$\begin{aligned}&P_{x_i,t}\left(\mathbb{A}_{x_t^*}=\mathscr{A}_2|\Theta=\theta_0, x_{i-1,t}\in\Omega_1, x_{i+1,t}\in\Omega_2, \mathbb{A}_{x_{i-1,t}}=\mathscr{A}_k, \mathbb{A}_{x_{i+1,t}}=\mathscr{A}_l\right)\\ =&\frac{a+b-3}{n-1}\Phi\Big(\frac{B_t-v_t'}{\sigma_{\varepsilon_n}}\Big)+\frac{1}{2}\frac{n-a-b}{n-1}+\frac{1}{n-1}(\delta_{\mathscr{A}_k=\mathscr{A}_2}+\delta_{\mathscr{A}_l=\mathscr{A}_2})\end{aligned}\tag{F.9}$$

Action 3:

$$\begin{aligned}&P_{x_i,t}\left(\mathbb{A}_{x_t^*}=\mathscr{A}_3|\Theta=\theta_0, x_{i-1,t}\in\Omega_1, x_{i+1,t}\in\Omega_2, \mathbb{A}_{x_{i-1,t}}=\mathscr{A}_k, \mathbb{A}_{x_{i+1,t}}=\mathscr{A}_l\right)\\ =&\frac{a+b-3}{n-1}\left(\Phi\Big(\frac{A_t-v_t'}{\sigma_{\varepsilon_n}}\Big)-\Phi\Big(\frac{B_t-v_t'}{\sigma_{\varepsilon_n}}\Big)\right)+\frac{1}{n-1}(\delta_{\mathscr{A}_k=\mathscr{A}_3}+\delta_{\mathscr{A}_l=\mathscr{A}_3})\end{aligned}\tag{F.10}$$

[1] $\delta_{[\text{condition}]}\equiv\begin{cases}1 & \text{if [condition] is true.}\\ 0 & \text{otherwise}\end{cases}$

Case 2: $\Omega_1 \in \{\Omega_{\mathbf{IFT}}, \Omega_{\mathbf{UFT}}\}, \Omega_2 = \Omega_{\mathbf{ZIT}}$ **(or** $\Omega_2 \in \{\Omega_{\mathbf{IFT}}, \Omega_{\mathbf{UFT}}\}, \Omega_1 = \Omega_{\mathbf{ZIT}}$**)**

Action 1:

$$\begin{aligned}&P_{x_i,t}\big(\mathbb{A}_{x_t^*} = \mathscr{A}_1 | \Theta = \theta_0, x_{i-1,t} \in \Omega_1, x_{i+1,t} \in \Omega_2, \mathbb{A}_{x_{i-1,t}} = \mathscr{A}_k, \mathbb{A}_{x_{i+1,t}} = \mathscr{A}_l\big)\\ =&\frac{a+b-2}{n-1}\Phi\Big(\frac{v_t' - A_t}{\sigma_{\varepsilon_n}}\Big) + \frac{1}{2}\frac{n-1-a-b}{n-1} + \frac{1}{n-1}(\delta_{\mathscr{A}_k=\mathscr{A}_1} + \delta_{\mathscr{A}_l=\mathscr{A}_1})\end{aligned} \tag{F.11}$$

Action 2:

$$\begin{aligned}&P_{x_i,t}\big(\mathbb{A}_{x_t^*} = \mathscr{A}_2 | \Theta = \theta_0, x_{i-1,t} \in \Omega_1, x_{i+1,t} \in \Omega_2, \mathbb{A}_{x_{i-1,t}} = \mathscr{A}_k, \mathbb{A}_{x_{i+1,t}} = \mathscr{A}_l\big)\\ =&\frac{a+b-2}{n-1}\Phi\Big(\frac{B_t - v_t'}{\sigma_{\varepsilon_n}}\Big) + \frac{1}{2}\frac{n-1-a-b}{n-1} + \frac{1}{n-1}(\delta_{\mathscr{A}_k=\mathscr{A}_2} + \delta_{\mathscr{A}_l=\mathscr{A}_2})\end{aligned} \tag{F.12}$$

Action 3:

$$\begin{aligned}&P_{x_i,t}\big(\mathbb{A}_{x_t^*} = \mathscr{A}_3 | \Theta = \theta_0, x_{i-1,t} \in \Omega_1, x_{i+1,t} \in \Omega_2, \mathbb{A}_{x_{i-1,t}} = \mathscr{A}_k, \mathbb{A}_{x_{i+1,t}} = \mathscr{A}_l\big)\\ =&\frac{a+b-2}{n-1}\left(\Phi\Big(\frac{A_t - v_t'}{\sigma_{\varepsilon_n}}\Big) - \Phi\Big(\frac{B_t - v_t'}{\sigma_{\varepsilon_n}}\Big)\right) + \frac{1}{n-1}\delta_{\mathscr{A}_l=\mathscr{A}_3}\end{aligned} \tag{F.13}$$

Case 3: $\Omega_1 = \Omega_{\mathbf{ZIT}}, \Omega_2 = \Omega_{\mathbf{ZIT}}$

Action 1:

$$\begin{aligned}&P_{x_i,t}\big(\mathbb{A}_{x_t^*} = \mathscr{A}_1 | \Theta = \theta_0, x_{i-1,t} \in \Omega_1, x_{i+1,t} \in \Omega_2, \mathbb{A}_{x_{i-1,t}} = \mathscr{A}_k, \mathbb{A}_{x_{i+1,t}} = \mathscr{A}_l\big)\\ =&\frac{a+b-1}{n-1}\Phi\Big(\frac{v_t' - A_t}{\sigma_{\varepsilon_n}}\Big) + \frac{1}{2}\frac{n-2-a-b}{n-1} + \frac{1}{n-1}(\delta_{\mathscr{A}_k=\mathscr{A}_1} + \delta_{\mathscr{A}_l=\mathscr{A}_1})\end{aligned} \tag{F.14}$$

Action 2:

$$\begin{aligned}&P_{x_i,t}\big(\mathbb{A}_{x_t^*} = \mathscr{A}_2 | \Theta = \theta_0, x_{i-1,t} \in \Omega_1, x_{i+1,t} \in \Omega_2, \mathbb{A}_{x_{i-1,t}} = \mathscr{A}_k, \mathbb{A}_{x_{i+1,t}} = \mathscr{A}_l\big)\\ =&\frac{a+b-1}{n-1}\Phi\Big(\frac{B_t - v_t'}{\sigma_{\varepsilon_n}}\Big) + \frac{1}{2}\frac{n-2-a-b}{n-1} + \frac{1}{n-1}(\delta_{\mathscr{A}_k=\mathscr{A}_2} + \delta_{\mathscr{A}_l=\mathscr{A}_2})\end{aligned} \tag{F.15}$$

Action 3:

$$\begin{aligned}&P_{x_i,t}\big(\mathbb{A}_{x_t^*} = \mathscr{A}_3 | \Theta = \theta_j, x_{i-1,t} \in \Omega_1, x_{i+1,t} \in \Omega_2, \mathbb{A}_{x_{i-1,t}} = \mathscr{A}_k, \mathbb{A}_{x_{i+1,t}} = \mathscr{A}_l\big)\\ =&\frac{a+b-1}{n-1}\left(\Phi\Big(\frac{A_t - v_t'}{\sigma_{\varepsilon_n}}\Big) - \Phi\Big(\frac{B_t - v_t'}{\sigma_{\varepsilon_n}}\Big)\right)\end{aligned} \tag{F.16}$$

F.2.2 Belief $\Theta = \theta_1$

When $\Theta = \theta_1$, the uninformed fundamental traders believe that they are deceived and that the shock is fake to mislead them into estimating the biased value of the stock price. As soon as the shock comes, the informed fundamental traders receive the price signal $w^{\text{IFT}} = v + \varepsilon$, while the uninformed fundamental traders receive the price signal $w^{\text{UFT}} = v' + \varepsilon$.

Case 1: When $\Omega_1 = \Omega_{\text{IFT}}, \Omega_2 = \Omega_{\text{IFT}}$

In this case $a \geq 2$.

Action 1:

$$\begin{aligned}
&P_{x_i,t}\left(\mathbb{A}_{x_t^*} = \mathscr{A}_1 | \Theta = \theta_2, x_{i-1,t} \in \Omega_1, x_{i+1,t} \in \Omega_2, \mathbb{A}_{x_{i-1,t}} = \mathscr{A}_k, \mathbb{A}_{x_{i+1,t}} = \mathscr{A}_l\right) \\
=&\frac{a-2}{n-1}\Phi\left(\frac{v_t - A_t}{\sigma_{\varepsilon_n}}\right) + \frac{b-1}{n-1}\Phi\left(\frac{v_t' - A_t}{\sigma_{\varepsilon_n}}\right) + \frac{1}{2}\frac{n-a-b}{n-1} + \frac{1}{n-1}(\delta_{\mathscr{A}_k = \mathscr{A}_1} + \delta_{\mathscr{A}_l = \mathscr{A}_1})
\end{aligned} \tag{F.17}$$

Action 2:

$$\begin{aligned}
&P_{x_i,t}\left(\mathbb{A}_{x_t^*} = \mathscr{A}_2 | \Theta = \theta_j, x_{i-1,t} \in \Omega_1, x_{i+1,t} \in \Omega_2, \mathbb{A}_{x_{i-1,t}} = \mathscr{A}_k, \mathbb{A}_{x_{i+1,t}} = \mathscr{A}_l\right) \\
=&\frac{a-2}{n-1}\Phi\left(\frac{B_t - v_t}{\sigma_{\varepsilon_n}}\right) + \frac{b-1}{n-1}\Phi\left(\frac{B_t - v_t'}{\sigma_{\varepsilon_n}}\right) + \frac{1}{2}\frac{n-a-b}{n-1} + \frac{1}{n-1}(\delta_{\mathscr{A}_k = \mathscr{A}_2} + \delta_{\mathscr{A}_l = \mathscr{A}_2})
\end{aligned} \tag{F.18}$$

Action 3:

$$\begin{aligned}
&P_{x_i,t}\left(\mathbb{A}_{x_t^*} = \mathscr{A}_3 | \Theta = \theta_j, x_{i-1,t} \in \Omega_1, x_{i+1,t} \in \Omega_2, \mathbb{A}_{x_{i-1,t}} = \mathscr{A}_k, \mathbb{A}_{x_{i+1,t}} = \mathscr{A}_l\right) \\
=&\frac{a-2}{n-1}\left(\Phi\left(\frac{A_t - v_t}{\sigma_{\varepsilon_n}}\right) - \Phi\left(\frac{B_t - v_t}{\sigma_{\varepsilon_n}}\right)\right) + \frac{b-1}{n-1}\left(\Phi\left(\frac{A_t - v_t'}{\sigma_{\varepsilon_n}}\right) - \Phi\left(\frac{B_t - v_t'}{\sigma_{\varepsilon_n}}\right)\right) + \\
&\frac{1}{n-1}(\delta_{\mathscr{A}_k = \mathscr{A}_3} + \delta_{\mathscr{A}_l = \mathscr{A}_3})
\end{aligned} \tag{F.19}$$

Case 2: When $\Omega_1 = \Omega_{\text{UFT}}, \Omega_2 = \Omega_{\text{UFT}}$

In this case $b \geq 3$.

Action 1:

$$\begin{aligned}&P_{x_i,t}\left(\mathbb{A}_{x_t^*}=\mathscr{A}_1|\Theta=\theta_2,x_{i-1,t}\in\Omega_1,x_{i+1,t}\in\Omega_2,\mathbb{A}_{x_{i-1,t}}=\mathscr{A}_k,\mathbb{A}_{x_{i+1,t}}=\mathscr{A}_l\right)\\=&\frac{a}{n-1}\Phi\left(\frac{v_t-A_t}{\sigma_{\varepsilon_n}}\right)+\frac{b-3}{n-1}\Phi\left(\frac{v_t'-A_t}{\sigma_{\varepsilon_n}}\right)+\frac{1}{2}\frac{n-a-b}{n-1}+\frac{1}{n-1}(\delta_{\mathscr{A}_k=\mathscr{A}_1}+\delta_{\mathscr{A}_l=\mathscr{A}_1})\end{aligned}\tag{F.20}$$

Action 2:

$$\begin{aligned}&P_{x_i,t}\left(\mathbb{A}_{x_t^*}=\mathscr{A}_2|\Theta=\theta_j,x_{i-1,t}\in\Omega_1,x_{i+1,t}\in\Omega_2,\mathbb{A}_{x_{i-1,t}}=\mathscr{A}_k,\mathbb{A}_{x_{i+1,t}}=\mathscr{A}_l\right)\\=&\frac{a}{n-1}\Phi\left(\frac{B_t-v_t}{\sigma_{\varepsilon_n}}\right)+\frac{b-3}{n-1}\Phi\left(\frac{B_t-v_t'}{\sigma_{\varepsilon_n}}\right)+\frac{1}{2}\frac{n-a-b}{n-1}+\frac{1}{n-1}(\delta_{\mathscr{A}_k=\mathscr{A}_2}+\delta_{\mathscr{A}_l=\mathscr{A}_2})\end{aligned}\tag{F.21}$$

Action 3:

$$\begin{aligned}&P_{x_i,t}\left(\mathbb{A}_{x_t^*}=\mathscr{A}_3|\Theta=\theta_j,x_{i-1,t}\in\Omega_1,x_{i+1,t}\in\Omega_2,\mathbb{A}_{x_{i-1,t}}=\mathscr{A}_k,\mathbb{A}_{x_{i+1,t}}=\mathscr{A}_l\right)\\=&\frac{a}{n-1}\left(\Phi\left(\frac{A_t-v_t}{\sigma_{\varepsilon_n}}\right)-\Phi\left(\frac{B_t-v_t}{\sigma_{\varepsilon_n}}\right)\right)+\frac{b-3}{n-1}\left(\Phi\left(\frac{A_t-v_t'}{\sigma_{\varepsilon_n}}\right)-\Phi\left(\frac{B_t-v_t'}{\sigma_{\varepsilon_n}}\right)\right)+\\&\frac{1}{n-1}(\delta_{\mathscr{A}_k=\mathscr{A}_3}+\delta_{\mathscr{A}_l=\mathscr{A}_3})\end{aligned}\tag{F.22}$$

Case 3: When $\Omega_1=\Omega_{\mathbf{IFT}},\Omega_2=\Omega_{\mathbf{UFT}}$ (or $\Omega_1=\Omega_{\mathbf{UFT}},\Omega_2=\Omega_{\mathbf{IFT}}$)

In this case $b\geq 3$.

Action 1:

$$\begin{aligned}&P_{x_i,t}\left(\mathbb{A}_{x_t^*}=\mathscr{A}_1|\Theta=\theta_2,x_{i-1,t}\in\Omega_1,x_{i+1,t}\in\Omega_2,\mathbb{A}_{x_{i-1,t}}=\mathscr{A}_k,\mathbb{A}_{x_{i+1,t}}=\mathscr{A}_l\right)\\=&\frac{a-1}{n-1}\Phi\left(\frac{v_t-A_t}{\sigma_{\varepsilon_n}}\right)+\frac{b-2}{n-1}\Phi\left(\frac{v_t'-A_t}{\sigma_{\varepsilon_n}}\right)+\frac{1}{2}\frac{n-a-b}{n-1}+\frac{1}{n-1}(\delta_{\mathscr{A}_k=\mathscr{A}_1}+\delta_{\mathscr{A}_l=\mathscr{A}_1})\end{aligned}\tag{F.23}$$

Action 2:

$$\begin{aligned}&P_{x_i,t}\left(\mathbb{A}_{x_t^*}=\mathscr{A}_2|\Theta=\theta_j,x_{i-1,t}\in\Omega_1,x_{i+1,t}\in\Omega_2,\mathbb{A}_{x_{i-1,t}}=\mathscr{A}_k,\mathbb{A}_{x_{i+1,t}}=\mathscr{A}_l\right)\\=&\frac{a-1}{n-1}\Phi\left(\frac{B_t-v_t}{\sigma_{\varepsilon_n}}\right)+\frac{b-2}{n-1}\Phi\left(\frac{B_t-v_t'}{\sigma_{\varepsilon_n}}\right)+\frac{1}{2}\frac{n-a-b}{n-1}+\frac{1}{n-1}(\delta_{\mathscr{A}_k=\mathscr{A}_2}+\delta_{\mathscr{A}_l=\mathscr{A}_2})\end{aligned}\tag{F.24}$$

Action 3:

$$
\begin{aligned}
&P_{x_i,t}\big(\mathbb{A}_{x_t^*}=\mathscr{A}_3|\Theta=\theta_j, x_{i-1,t}\in\Omega_1, x_{i+1,t}\in\Omega_2, \mathbb{A}_{x_{i-1,t}}=\mathscr{A}_k, \mathbb{A}_{x_{i+1,t}}=\mathscr{A}_l\big)\\
=&\frac{a-1}{n-1}\left(\Phi\Big(\frac{A_t-v_t}{\sigma_{\varepsilon_n}}\Big)-\Phi\Big(\frac{B_t-v_t}{\sigma_{\varepsilon_n}}\Big)\right)+\frac{b-2}{n-1}\left(\Phi\Big(\frac{A_t-v_t'}{\sigma_{\varepsilon_n}}\Big)-\Phi\Big(\frac{B_t-v_t'}{\sigma_{\varepsilon_n}}\Big)\right)+\\
&\frac{1}{n-1}(\delta_{\mathscr{A}_k=\mathscr{A}_3}+\delta_{\mathscr{A}_l=\mathscr{A}_3})
\end{aligned}
\tag{F.25}
$$

Case 4: When $\Omega_1=\Omega_{\mathbf{IFT}},\Omega_2=\Omega_{\mathbf{ZIT}}$ (or $\Omega_1=\Omega_{\mathbf{ZIT}},\Omega_2=\Omega_{\mathbf{IFT}}$)

In this case $b\geq 3$.

Action 1:

$$
\begin{aligned}
&P_{x_i,t}\big(\mathbb{A}_{x_t^*}=\mathscr{A}_1|\Theta=\theta_2, x_{i-1,t}\in\Omega_1, x_{i+1,t}\in\Omega_2, \mathbb{A}_{x_{i-1,t}}=\mathscr{A}_k, \mathbb{A}_{x_{i+1,t}}=\mathscr{A}_l\big)\\
=&\frac{a-1}{n-1}\Phi\Big(\frac{v_t-A_t}{\sigma_{\varepsilon_n}}\Big)+\frac{b-1}{n-1}\Phi\Big(\frac{v_t'-A_t}{\sigma_{\varepsilon_n}}\Big)+\frac{1}{2}\frac{n-1-a-b}{n-1}+\frac{1}{n-1}(\delta_{\mathscr{A}_k=\mathscr{A}_1}+\delta_{\mathscr{A}_l=\mathscr{A}_1})
\end{aligned}
\tag{F.26}
$$

Action 2:

$$
\begin{aligned}
&P_{x_i,t}\big(\mathbb{A}_{x_t^*}=\mathscr{A}_2|\Theta=\theta_2, x_{i-1,t}\in\Omega_1, x_{i+1,t}\in\Omega_2, \mathbb{A}_{x_{i-1,t}}=\mathscr{A}_k, \mathbb{A}_{x_{i+1,t}}=\mathscr{A}_l\big)\\
=&\frac{a-1}{n-1}\Phi\Big(\frac{B_t-v_t}{\sigma_{\varepsilon_n}}\Big)+\frac{b-1}{n-1}\Phi\Big(\frac{B_t-v_t'}{\sigma_{\varepsilon_n}}\Big)+\frac{1}{2}\frac{n-1-a-b}{n-1}+\frac{1}{n-1}(\delta_{\mathscr{A}_k=\mathscr{A}_2}+\delta_{\mathscr{A}_l=\mathscr{A}_2})
\end{aligned}
\tag{F.27}
$$

Action 3:

$$
\begin{aligned}
&P_{x_i,t}\big(\mathbb{A}_{x_t^*}=\mathscr{A}_3|\Theta=\theta_2, x_{i-1,t}\in\Omega_1, x_{i+1,t}\in\Omega_2, \mathbb{A}_{x_{i-1,t}}=\mathscr{A}_k, \mathbb{A}_{x_{i+1,t}}=\mathscr{A}_l\big)\\
=&\frac{a-1}{n-1}\left(\Phi\Big(\frac{A_t-v_t}{\sigma_{\varepsilon_n}}\Big)-\Phi\Big(\frac{B_t-v_t}{\sigma_{\varepsilon_n}}\Big)\right)+\frac{b-1}{n-1}\left(\Phi\Big(\frac{A_t-v_t'}{\sigma_{\varepsilon_n}}\Big)-\Phi\Big(\frac{B_t-v_t'}{\sigma_{\varepsilon_n}}\Big)\right)+\\
&\frac{1}{n-1}\delta_{\mathscr{A}_k=\mathscr{A}_3}
\end{aligned}
\tag{F.28}
$$

Case 5: When $\Omega_1=\Omega_{\mathbf{UFT}},\Omega_2=\Omega_{\mathbf{ZIT}}$ (or $\Omega_1=\Omega_{\mathbf{ZIT}},\Omega_2=\Omega_{\mathbf{UFT}}$)

In this case $b\geq 3$.

Action 1:

$$
\begin{aligned}
&P_{x_i,t}\big(\mathbb{A}_{x_t^*}=\mathscr{A}_1|\Theta=\theta_2,x_{i-1,t}\in\Omega_1,x_{i+1,t}\in\Omega_2,\mathbb{A}_{x_{i-1,t}}=\mathscr{A}_k,\mathbb{A}_{x_{i+1,t}}=\mathscr{A}_l\big)\\
=&\frac{a}{n-1}\Phi\Big(\frac{v_t-A_t}{\sigma_{\varepsilon_n}}\Big)+\frac{b-2}{n-1}\Phi\Big(\frac{v_t'-A_t}{\sigma_{\varepsilon_n}}\Big)+\frac{1}{2}\frac{n-1-a-b}{n-1}+\frac{1}{n-1}(\delta_{\mathscr{A}_k=\mathscr{A}_1}+\delta_{\mathscr{A}_l=\mathscr{A}_1})
\end{aligned}
\tag{F.29}
$$

Action 2:

$$
\begin{aligned}
&P_{x_i,t}\big(\mathbb{A}_{x_t^*}=\mathscr{A}_2|\Theta=\theta_2,x_{i-1,t}\in\Omega_1,x_{i+1,t}\in\Omega_2,\mathbb{A}_{x_{i-1,t}}=\mathscr{A}_k,\mathbb{A}_{x_{i+1,t}}=\mathscr{A}_l\big)\\
=&\frac{a}{n-1}\Phi\Big(\frac{B_t-v_t}{\sigma_{\varepsilon_n}}\Big)+\frac{b-2}{n-1}\Phi\Big(\frac{B_t-v_t'}{\sigma_{\varepsilon_n}}\Big)+\frac{1}{2}\frac{n-1-a-b}{n-1}+\frac{1}{n-1}(\delta_{\mathscr{A}_k=\mathscr{A}_2}+\delta_{\mathscr{A}_l=\mathscr{A}_2})
\end{aligned}
\tag{F.30}
$$

Action 3:

$$
\begin{aligned}
&P_{x_i,t}\big(\mathbb{A}_{x_t^*}=\mathscr{A}_3|\Theta=\theta_2,x_{i-1,t}\in\Omega_1,x_{i+1,t}\in\Omega_2,\mathbb{A}_{x_{i-1,t}}=\mathscr{A}_k,\mathbb{A}_{x_{i+1,t}}=\mathscr{A}_l\big)\\
=&\frac{a}{n-1}\left(\Phi\Big(\frac{A_t-v_t}{\sigma_{\varepsilon_n}}\Big)-\Phi\Big(\frac{B_t-v_t}{\sigma_{\varepsilon_n}}\Big)\right)+\frac{b-2}{n-1}\left(\Phi\Big(\frac{A_t-v_t'}{\sigma_{\varepsilon_n}}\Big)-\Phi\Big(\frac{B_t-v_t'}{\sigma_{\varepsilon_n}}\Big)\right)+\\
&\frac{1}{n-1}\delta_{\mathscr{A}_k=\mathscr{A}_3}
\end{aligned}
\tag{F.31}
$$

Case 6: When $\Omega_1=\Omega_{\mathbf{ZIT}},\Omega_2=\Omega_{\mathbf{ZIT}}$

In this case $b\geq 3$.

Action 1:

$$
\begin{aligned}
&P_{x_i,t}\big(\mathbb{A}_{x_t^*}=\mathscr{A}_1|\Theta=\theta_2,x_{i-1,t}\in\Omega_1,x_{i+1,t}\in\Omega_2,\mathbb{A}_{x_{i-1,t}}=\mathscr{A}_k,\mathbb{A}_{x_{i+1,t}}=\mathscr{A}_l\big)\\
=&\frac{a}{n-1}\Phi\Big(\frac{v_t-A_t}{\sigma_{\varepsilon_n}}\Big)+\frac{b-1}{n-1}\Phi\Big(\frac{v_t'-A_t}{\sigma_{\varepsilon_n}}\Big)+\frac{1}{2}\frac{n-1-a-b}{n-1}+\frac{1}{n-1}(\delta_{\mathscr{A}_k=\mathscr{A}_1}+\delta_{\mathscr{A}_l=\mathscr{A}_1})
\end{aligned}
\tag{F.32}
$$

Action 2:

$$
\begin{aligned}
&P_{x_i,t}\big(\mathbb{A}_{x_t^*}=\mathscr{A}_2|\Theta=\theta_2,x_{i-1,t}\in\Omega_1,x_{i+1,t}\in\Omega_2,\mathbb{A}_{x_{i-1,t}}=\mathscr{A}_k,\mathbb{A}_{x_{i+1,t}}=\mathscr{A}_l\big)\\
=&\frac{a}{n-1}\Phi\Big(\frac{B_t-v_t}{\sigma_{\varepsilon_n}}\Big)+\frac{b-1}{n-1}\Phi\Big(\frac{B_t-v_t'}{\sigma_{\varepsilon_n}}\Big)+\frac{1}{2}\frac{n-1-a-b}{n-1}+\frac{1}{n-1}(\delta_{\mathscr{A}_k=\mathscr{A}_2}+\delta_{\mathscr{A}_l=\mathscr{A}_2})
\end{aligned}
\tag{F.33}
$$

Action 3:

$$
\begin{aligned}
&P_{x_i,t}\left(\mathbb{A}_{x_t^*}=\mathscr{A}_3|\Theta=\theta_2, x_{i-1,t}\in\Omega_1, x_{i+1,t}\in\Omega_2, \mathbb{A}_{x_{i-1,t}}=\mathscr{A}_k, \mathbb{A}_{x_{i+1,t}}=\mathscr{A}_l\right)\\
=&\frac{a}{n-1}\left(\Phi\Big(\frac{A_t-v_t}{\sigma_{\varepsilon_n}}\Big)-\Phi\Big(\frac{B_t-v_t}{\sigma_{\varepsilon_n}}\Big)\right)+\frac{b-1}{n-1}\left(\Phi\Big(\frac{A_t-v_t'}{\sigma_{\varepsilon_n}}\Big)-\Phi\Big(\frac{B_t-v_t'}{\sigma_{\varepsilon_n}}\Big)\right)
\end{aligned}
\tag{F.34}
$$

Appendix G

Maths Regarding Learning Strategy in Complete Network

G.1 Bayesian Learning Formula

In the n-agent complete network, assume that the numbers of informed fundamental traders n_{IFT} and uninformed fundamental traders n_{UFT} are a and b, respectively. $I_{i,t} = \bigcup_{j\neq i}\{\mathbb{A}_{x_j,t}\}\bigcup\{\mathbb{A}_{x_t^*}\}$, or $I_{i,t} = \bigcup_{j\neq i}\{\mathbb{A}_{x_j,t}\}$ if $x_t^* = x_{i,t}$.

Denote the space of all possible allocations of the other agents that from the perspective of the i-th agent as $\mathbf{T}_i$, given the numbers of informed fundamental traders n_{IFT} and uninformed fundamental traders n_{UFT} are a and b are determined. Then the number of the possibilities is $N = \binom{n-1}{b-1}\binom{n-b}{a}$, $\mathbf{T}_i = \bigcup_{n=1}^{N}\Omega_n$

Then

$$\begin{aligned} &P_{x_i,t}(I_{i,t}|\Theta=\theta_j) \\ =&\sum_{\Omega_n\in\mathbf{T}_i} P_{x_i,t}(I_{i,t}|\Theta=\theta_j,\Omega_n)P(\Omega_n) \end{aligned} \tag{G.1}$$

Similar to the illustration in F.1, the right hand side of the above consists two terms: the probability of the information occurs $P_{x_i,t}(I_{i,t}|\Theta=\theta_j,\Omega_n)$ and the probability of the other agents' identities with a certain allocation $P(\Omega_n)$. Following parts present detailed maths of the two terms.

G.1.1 The First Term

Let $\mathscr{F}_i$ refer to the all decisions that the i-th agent is notified of from the other agents excluding the trade information, that is, $\mathscr{F}_i = \bigcup_{j\neq i}\{\mathbb{A}_{x_j,t}\}$. As the assumption of the model is set, all trading agents are independent, so

If $x_t^* \neq x_{i,t}$,

$$\begin{aligned} & P_{x_i,t}\big(I_{i,t}|\Theta=\theta_j,\Omega_n\big) \\ =& P_{x_i,t}\big(\mathbb{A}_{x_t^*}=\mathscr{A}_i|\Theta=\theta_j,\Omega_n,\mathscr{F}_i\big)P(\mathscr{F}_i|\Theta=\theta_j,\Omega_n) \\ =& P_{x_i,t}\big(\mathbb{A}_{x_t^*}=\mathscr{A}_i|\Theta=\theta_j,\Omega_n,\mathscr{F}_i\big)\Pi_{p=1,p\neq i}^{n}P(\mathbb{A}_{x_{p,t}}=\mathscr{A}_{x_p}|\Theta=\theta_j,\Omega_n) \end{aligned} \tag{G.2}$$

If $x_t^* = x_{i,t}$,

$$P_{x_i,t}\big(I_{i,t}|\Theta=\theta_j,\Omega_n\big) = P(\mathscr{F}_i|\Theta=\theta_j,\Omega_n) = \Pi_{p=1,p\neq i}^{n}P(\mathbb{A}_{x_{p,t}}=\mathscr{A}_{x_p}|\Theta=\theta_j,\Omega_n) \tag{G.3}$$

where

$$P_{x_i,t}\big(\mathbb{A}_{x_t^*}=\mathscr{A}_i|\Theta=\theta_1,\Omega_n,\mathscr{F}_i\big) = \frac{1}{n-1}\Big(\sum_{j\neq i}\delta_{\mathscr{A}_j=\mathscr{A}_i}\Big) \tag{G.4}$$

The term $P(\mathbb{A}_{x_{p,t}}=\mathscr{A}_{x_p}|\Theta=\theta_j,\Omega_n)$ can be derived from the Eq. F.5 and Eq. F.6.

G.1.2 The Second Term

Initial Probabilities

At the beginning of learning, the probability $P(\Omega_n)$ is only based on the number of the informed fundamental traders $n_{\mathrm{IFT}} = a$ and the uninformed fundamental traders $n_{\mathrm{UFT}} = b$.

$$P(\Omega_n) = 1/N = \frac{1}{\binom{n-1}{b-1}\binom{n-b}{a}} \tag{G.5}$$

Updating Probabilities in the Following Trading Rounds

$$\begin{aligned} P(\Omega_n|\mathscr{F}_i) =& \frac{P(\mathscr{F}_i|\Omega_n)P(\Omega_n)}{\sum_{\Omega_n\in\mathbf{T}_i} P(\mathscr{F}_i|\Omega_n)P(\Omega_n)} \\ =& \frac{\Pi^n_{p=1,p\neq i} P(\mathbb{A}_{x_{p,t}} = \mathscr{A}_{x_p}|\Omega_n)P(\Omega_n)}{\sum_{\Omega_n\in\mathbf{T}_i} \Pi^n_{p=1,p\neq i} P(\mathbb{A}_{x_{p,t}} = \mathscr{A}_{x_p}|\Omega_n)P(\Omega_n)} \end{aligned} \tag{G.6}$$

Appendix H

Simulation Data and More Figures

To access the simulation data and see more figures, please visit https://github.com/ysun94/data-and-figure.