Impact of False Information from Spoofing Strategies: An ABM Model of Market Dynamics

1st HaoHang Li
School of Business
Stevens Institute of Technology
Hoboken, NJ, US
hli113@stevens.edu

2nd Steve Y. Yang
School of Business
Stevens Institute of Technology
Hoboken, NJ, US
steve.yang@stevens.edu

***Abstract*—Spoofing has been identified a form of market manipulation, and it is harmful to the stability of the financial market. However, the effect of spoofing activity is hard to analyze due to its complex interactions within the market and lack of data. This paper presents an agent-based simulation model of the continuous double auction market to replicate and analyze the market dynamics under spoofing conditions. The simulated market consists of fundamentalist, chartist, zero intelligence agents, and spoofing agents where several existing market stylized facts are reproduced and validated. The results show that in the presence of the spoofing agents and their market manipulation activities, the market volatility would increase, and spoofing activities would exacerbate the price variations. The fundamentalist agents would suffer a loss during the spoofing period but would be able to make profit during the price recovery phase. The chartist agents would suffer a loss when the spoofing agent realized its profit and the price recovers to its normal condition where they falsely believed the price movement trend would continue. The Sharpe ratio analysis also indicates the market manipulation activities of the spoofing agent would give themselves an unfair advantage resulting in a significantly higher Sharpe ratio than the other agents.**

***Index Terms*—Agent-based model, spoofing, fundamentalist, chartist, LOB**

I. Introduction

Algorithmic trading is a double-edged sword. The financial market has well acknowledged its ability to provide more liquidity [1] and improve the price discovery process [2]. But, at the same time, it may also make the financial market vulnerable to other new market manipulation practices driven by algorithms. In the spoofing case involving algorithmic trading, the algorithms inject fraudulent orders into the limit order book. Other traders who extract the market information from the limit order book may be misled by the false information and alternate their trading decision and place their order in favor of spoofers [3]. This paper aims to investigate the impact of the spoofing strategies on both other market participants and the market quality in general.

The spoofing's effect to undermine the market stability and exert negative impact on investors was first brought to public sight on May 6, 2010, during which the Dow Jones Industrial Average experienced the largest one-hour price decline in its history by falling down by 9 percent and recovered by 5 percent at the end of the day [4]. The chain-reaction also caused "many individual securities and ETFs experienced extreme price fluctuations and traded in a disorderly fashion" [5]. This event is commonly known as the "Flash Crash". In the later years, it was proven that spoofing trading activities was attributed to the "Flash Crash". Despite its detrimental effect on the financial market, the complex causal factors and irreproduciblility make the spoofing case hard to analyze.

This paper aims to reproduce the spoofing in an agent-based simulation model and study its impact on other market participants. In the model, the continuous double auction market with single security is implemented, and agents interact by placing bid or ask orders with specified prices and sizes in the limit order book. At each simulation time, the market price will be determined by the average of the highest bid price and the lowest ask price. The arrival and cancellation of the orders of agents follows the exponential distribution. Upon each trading time, the different groups of agents' trading actions will be characterized by their own heuristic functions. By applying the Monte Carlo simulation method, the simulated results are verified with stylized facts of financial time series. A major contribution of the model into the existing ABM literature is that the Chartists are not only making decisions based on their heuristic function but also are responsive to limit order book dynamics. Moreover, the model is able to replicate the upward or downward price movement and the price recovery process under the spoofing market conditions. In addition, the market dynamics, volatility and agents' profit are analyzed. We show that spoofing strategies would increase the market volatility i.e. prolong the normal price discovery process. At the same time, the spoofing agents bear much lower risk in making profit.

The rest of the paper is organized as follows. Section II introduces the agent-based simulation model, its significance in financial market simulations, and the previous research review. Section III demonstrates the design of the continuous double auction market and the behaviors of agents in the market. Section IV summarizes the normal market simulation, the verification of stylized facts, and section V discusses spoofing simulation, market dynamics and other measurements during the simulation time. Lastly, section VI provides the summary and concluding remarks of the paper.

II. Literature Review

The agent-based simulation model (ABM), also known as the individual-based simulation model (IBM), is a computational simulation in which artificial entities interact over time within customized environments [6]. Rules will determine the agents' actions, but the interactions of the different agents will be capable of producing behaviors that resemble real-world complex systems. Because of its ability to simulate dynamics in sophisticated systems, agent-based simulation is widely used in many scientific domains, including biology, epidemiology, social science, economy, and finance.

A. Agent-based Model in Artificial Market Simulation

Because of its ability to replicate the realistic market conditions, the agent-based simulation models gain extensive attention in the discussion of market design and regulations. Mizuta et al. (2015) [7] build an artificial market model to investigate the effectiveness of price variation limit and up-tick rule in limiting the volatile price variation caused by erroneous orders and the dark pool's ability to stabilize the market. They conclude that the time span of the price variation limit should be shorter than the life span of erroneous orders, and the up-tick rule is effective in limiting price fluctuations. The dark pool mechanism is able to improve the stability of the market, but overusing the dark pool would lead to an increase in the kurtosis of returns. Nagumo et al. (2017) [8] presents an artificial two-market model to examine the share-deprivation effect. In their simulation, the market with a smaller tick size would have more advantages and deprive the trading volume of the market with a larger tick size. Yang et al. (2020) [9]'s model further expands the discussion. They found the smaller tick size would lead to smaller bid-ask spread and larger trading volume but at the expense of smaller market depth. Compared to the uniform tick size system, the combination stepwise tick system based on price and volume would improve the market quality in terms of liquidity, volatility, and price efficiency.

Due to the ability to replicate the financial market, agent-based simulation models are often used as a tool to investigate financial market properties. Bookstaber and Paddrik (2015) [10] present a micro-structured liquidity dynamic model within the agent-based simulation framework. Within the model, the interactions among market making, liquidity supplying, and liquidity demanding agents are able to reproduce realistic liquidity crisis events. Yang et al. (2014) [11] simulates social media networks to investigate the propagation of Twitter information in the financial market and its impact. Their results show the spread of malicious messages would influence the agents' trading actions and could trigger an extreme price fluctuation event. Reducing the critical nodes with the highest betweenness centrality may be an effective preventive policy to limit the market manipulation activities through social networks.

B. Agent-Based Simulation Model for Spoofing

Despite its detrimental effect to disrupt the financial market, the flash crash event and spoofing market manipulation behavior are hard to analyze because of its irreproducibility and the complex interactions within the market. Later findings prove the flash crash event is not an isolated case where spoofing market manipulation behavior influences the market. In 2020, the CFTC brought up a dozen spoofing enforcement actions and filed complaints to several financial institutions. Therefore, an agent-based simulation model reflecting the statistical properties of realistic financial markets will provide more insight into microstructure dynamics and guidance to regulate spoofing activities.

Paddrik et al. (2012) [12] suggest an agent-based simulation model simulating the flash crash event, in which the market is composed of the interactions among fundamental investors, market makers, opportunistic, high-frequency traders, and small traders. Different types of agents are characterized by their trade speed, position limit, order size, and order price selections. The characteristics of the flash crash event are accurately recreated, and the authenticity of the model is verified with stylized facts. Hayes et al. (2012) [13]'s following work further improves the model by adding intelligent high-frequency market makers. The high-frequency market makers passively participate in the market with inventory limits and cancel their orders when the bid-ask volume difference exceeds a specified threshold.

While using attributes to define agents' behaviors in the simulation, heuristic functions enable agents to actively evaluate the simulated markets and give them autonomy to react accordingly, which makes agents' interactions more complex and enhances the model's ability to describe real-world systems. Westerhoff (2016) [14] proposes a fundamentalist-chartist's approach to simulating the artificial financial market. Within the model, the log price of the asset is driven by the demand of the agents participating in fundamental or chartist strategy. Each simulation time, the number of orders of fundamental strategy is formalized by the deviation between the current asset price and its own fundamental value and the charist strategy' orders correlated to the short-term price movement. The attractiveness of one strategy will be calculated based on the previous round's return, and the softmax function normalizes the weights of strategies. Following the fundamentalist-chartist's framework, Leal et al. (2016) [15] 's work further extends the model in the context of high-frequency trading and flash crash. The low-frequency agents constantly switch their strategy between fundamentalist and chartist depending on the strategy's profitability in the last time. The probability low-frequency agents behave in a specific strategy is determined by the similar fashion of Westerhoff (2016) [14]'s model. The high-frequency agents are activated when the last price change in the market is above a fixed threshold sampled from a uniform distribution with bounded support. The volatile price fluctuation similar to the flash crash event is endogenously generated within simulations, and the model is able to replicate some of the main stylized facts of financial markets. In the further discussion of Leal and Mauro (2019) [16], the model is applied to investigate the effectiveness of regularizations on the occurrence and duration of flash crashes. The simulation result suggests that the implementation of minimum resting time and cancellation fees would lead to lower market volatility during incidents. The introduction of transaction tax would also increase market stability but with lower effectiveness. On the other side, restricting the high-frequency traders with regulatory policies would weaken the market resilience because they potentially positively affect liquidity restoration and price recovery after the crash. The work of Wang et al. (2021) [17] presents a computation model of spoofing. The background agents apply a heuristic belief learning strategy to evaluate the information in the limit order book to determine trading with the intent to maximize their surplus. The spoofing agents continuously cancel and submit a large order one tick behind the best bid price, to feign a strong buying power. The simulation analysis shows the heuristic belief learning agents are vulnerable to spoofing attacks, and spoofing decreases the market surplus. A cloaking mechanism, a mechanism that symmetrically hides a fixed number of price levels from the best bid and ask, is purposed to curb the spoofing, with the effect to diminish spoofing but reduce the surplus.

III. Model Design

In this section, we discuss the design of the agent-based model using the standard design process. First we outline the market settings, and then we describe different types of agents and their behaviors.

A. Market Topology

The simulation market setting in the model is a continuous double auction market with a single security. The security will have a constant fundamental value (F) at each simulation time. Agents will interact by submitting limit orders with specified order sizes and order prices to the limit order book. Different types of agents will use different heuristic functions to evaluate the information in the order book and decide their trading actions. Every simulation period agents may submit an order, cancel the unmatched orders, or do nothing. A matching engine in the limit order book checks for trade execution between each period based on buy and sell orders.

In the simulation model, the limit order submitted by agents will be characterized by:

- **Order Size (S)**: The empirical finding suggests the size of limit orders in the limit order book can be well fitted by a lognormal distribution [18]. Therefore, in the simulation model, the order size of agents will be described by:

$$S = int(Z_s) + \gamma \tag{1}$$

  The random variable Z_s is sampled from a truncated lognormal distribution with $ln(Z_s) \sim N(0, \sigma_s)$. The sampled variable Z_s will be rounded to integer with in range$[Z_l, Z_u]$, and the γ is a variable to shift the entire sample distribution.

- **Order Prices**: Following the liquidity model of Bookstaber and Paddrik (2015) [10], the limit order price will follow a lognormal distribution as well. At each simulation time, an agent will decide its trading direction via its own heuristic function. With the trading action, the price of the limit order will be determined by:

$$P_{bid} = P_t(2.0 - Z_p/\lambda_p) \tag{2}$$

$$P_{ask} = P_t Z_p/\lambda_p \tag{3}$$

  In the above equations, the P_t is the current market mid-price, and the Z_p is a lognormal random variable with $ln(Z_p) \sim N(0, 0.25)$. The price movement ratio will be truncated within $[0, 2]$ The λ_p parameter will govern the overlapping bid and ask orders.
  From the bid and ask order price distributions:
  - The agent will have a higher probability of placing the order price near the current mid-price and a lower probability of placing away from the current mid-price. Therefore, after matching the crossed orders, the remaining orders will accumulate around the price level close to the mid-price, and fewer orders accumulate around the price level away from the current mid-price.
  - Sampled order prices will center around the current mid-price, with part of the order across the current mid-price. The overlapping of the crossed bid and ask orders will be matched with another side.
- **Order Arrival and Order Life Span**: The duration of the agent's order arrival and order cancellation follows the exponential distribution. When a new order is generated, it will be assigned a life span. If the order in the limit order book is not fully fulfilled in its life span, the agents will cancel it. The order arrival rate is unique to different types of agents. In the simulation model, the chartists will monitor the market more closely than the fundamentalist to look for possible trading opportunities. Thus chartists will have a shorter average order arrival time than and tend to trade more frequently than the fundamentalist.

B. Agents

1) Fundamentalist Agents

Following to Westerhoff (2016) [14]'s fundamentalist-charitist's approach, the fundamentalists have an informed view of security. They places orders to stabilize the market and reduce mispricing in the market. The heuristic function of fundamentalist is defined as:

$$J_F(t) = a(F - P_t) + \sigma_F \epsilon \tag{4}$$

The parameter a and parameter σ_F are greater than 0, and are sampled from uniform distribution for each individual agent. The ϵ is a standard normal random variable. Fundamentalists decide their trading actions based on the difference between the current market price and asset fundamental value. The variable ϵ represents the randomness introduced in agents' decision process. Agents will react accordingly when the mispricing signals outperform the random observations. Based on the heuristic value $J_F(t)$, the fundamentalist agents will place:

- Bid orders when the mid price is less than the fundamental value , thus $J_F(t) > 0$
- Ask orders when the mid price is greater than the fundamental value , thus $J_F(t) < 0$

2) *Chartist Agents*

The chartist agents' trading actions are made by exploiting the past price changes and information in the book shape. The chartist agents follow the momentum trading strategy and seek to benefit from the potential price movement direction. The heuristic function of the agents are defined as:

$$J_C(t) = c\frac{1}{W}\sum_{i=0}^{W}(P_{t-i} - P_{t-i-1}) + dI + \sigma_C\epsilon \tag{5}$$

where the I is the agent's measurement of order book imbalance and is defined as:

$$I = \begin{cases} max(\frac{Vol_{Bid}}{Vol_{Ask}} - 1 - T_I, 0), & Vol_{Bid} > Vol_{Ask} \\ -max(\frac{Vol_{Ask}}{Vol_{Bid}} - 1 - T_I, 0), & Vol_{Ask} > Vol_{Bid} \end{cases} \tag{6}$$

The parameter c, d and σ_C are greater than 0, and are sampled from uniform distribution for each individual agent. The ϵ is a standard normal random variable. The heuristic of chartist agents consists of two parts. First, the agents will evaluate the past price changes with an evaluation window of size W. The evaluation window size is sampled from the uniform distribution, and each agent would make their trading decision based on different lengths of price history. Since chartist agents tend to follow the market trend, agents will place bid orders if the upward movements of prices are presented in the market and will place ask orders if the markets move downward. Second, the chartist agents will exploit the order book imbalance information to make their trading direction. If the imbalance volume ratio of bid and ask sides are significant than the thresholds, the agents will place orders accordingly due to buying or selling pressures.

3) *Zero Intelligence Agents*

The zero intelligence agents serve as the background agents during the simulation. Unlike the fundamentalist or chartist agents, the zero intelligence agents do not have a heuristic function to evaluate and exploit the information in the price history or order book. At each simulation time, the zero intelligence agents will have equal probability to place either a bid or ask order with a specified constant order size (S_Z) and order life span (L_Z). The unfulfilled orders will be canceled after reaching their life span.

4) *Spoofing Agent*

By submitting the spoofing orders to the limit order book, the spoofing agents intend to feign the buying or selling pressure and affect other agents' trading decisions. The spoofing agent arrives at the market at a certain simulation time (S_t) with a clear intention to manipulate the market price upward or downward. To profit from the market manipulation activity, the spoofing agent first establishes its position at arrival time and places a limit bid or ask order with its desired price (S_p). Then the agent will keep injecting the spoofing orders to the opposite direction of that limit order to create fallacious buying or selling pressures. The size of the spoofing orders is proportional to the current market volume and gradually increases with step (S_s) until it reaches the specified ratio limit (S_l). From the empirical analysis, the spoofers are high-frequency traders who frequently modify their spoofing order to avoid transactions. In the simulation model, the spoofing agents will cancel all their previous spoofing orders and place new orders to ensure orders are anchored to the price level relative to the current best bid price or best ask price.

IV. Experimentation

A. *Normal Market Simulation*

Parameter	Value	Description
σ_s	2.5	Underlying normal standard deviation of order size lognormal random variable Z_s
γ	3.0	Parameter for order size function
$[Z_l, Z_u]$	[1, 10]	Truncated order size lower bound and upper bound
σ_p	0.25	Standard deviation of order price lognormal random variable Z_p
λ_p	0.8	Parameter for order price function
N_Z	100	Number of zero intelligence agents
S_Z	1	Constant order size for zero intelligence agents
L_Z	2	Order life span for zero intelligence agents
N_F	200	Number of fundamentalist agents
A_F	15	Average order arrival time of fundamentalist agents
L_F	5	Average order life span of fundamentalist agents
$[a_l, a_u]$	[0.1, 0.15]	Lower and upper bound of parameter a of fundamentalist agents heuristic function
$[\sigma_F^l, \sigma_F^u]$	[1.0, 2.0]	Lower and upper bound of parameter σ_F of fundamentalist agents heuristic function
N_C	200	Number of chartist agents
A_F	5	Average order arrival time of chartist agents
L_C	5	Average Order life span of chartist agents
$[c_l, c_u]$	[0.0025, 0.01]	Lower and upper bound of parameter c of chartist agents heuristic function
$[W_l, W_u]$	[3, 10]	Lower and upper bound of evaluation window size of chartist agents
$[d_l, d_u]$	[0.08, 0.1]	Lower and upper bound of parameter d of chartist agents heuristic function
$[T_I^l, T_I^u]$	[0.4, 1.0]	Lower and upper bound of imbalance ratio T_I of chartist agents heuristic function
$[\sigma_C^l, \sigma_C^u]$	[0.1, 0.2]	Lower and upper bound of parameter σ_C of chartist agents heuristic function

TABLE I: Parameters for the simulation model

The parameters of the model are summarized in Table I. In this study, there are 500 agents in the simulation model. The agents' heuristic parameters are sampled from uniform distributions to ensure the behavior heterogeneity of agents. Compared with fundamentalist agents, the chartist agents as technical traders would have a faster trading speed to profit from the short-term price variations.

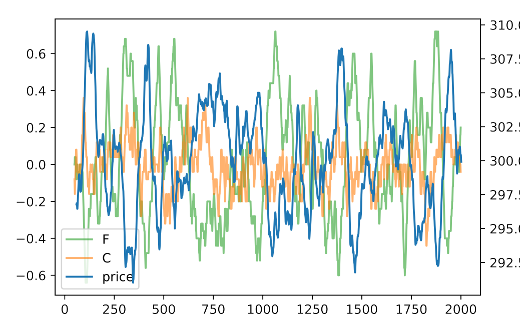


Fig. 1: Market dynamic

From 2000 simulation results, the equilibrium price formation process is observed with occasional extreme fluctuations.

Figure 1 presents the rolling average of agents' trading direction and market price with a window size of 50. At each time point, the average heuristic value of each agents group is calculated. If the average heuristic value is larger than 0, the group of agents' trading direction will be recorded as 1.0 to denote a buy action. If the average heuristic value is less than 0, the group of agents' trading direction will be recorded as -1.0 to denote a sell action. From the plot, it is clear that the trading actions of fundamentalist agents are inverse to the price movement directions, and the chartist agents would follow the sharp price movement in the market. The market dynamic can be summarized as the fundamentalist agents stabilize the market and chartist agents exacerbate the price variations [14].

B. Stylized Facts and Model Validation

The financial time series have well-established stylized facts, which reflect the common statistical properties of prices or returns. Therefore, to verify whether the simulated market is able to replicate the characteristics of the financial market, the price and return series generated by the simulation model should be checked to see if they are aligned with empirical properties. The tests are carried out with the Monte-Carlo simulation with 200 iterations, and each iteration is simulated for 2000 steps. The mid-prices of each simulation time are recorded, and the return of each simulation time is calculated as log returns, which is defined by $r_t = log(P_t) - log(P_{t-1})$ where P_t is the price at time t.

For the the i-th result of the Monte-Carlo simulation with N simulations, the measurements (Y_i) of the return series are estimated and averaged to calculate the sample mean ($\hat{Y}$) the 95% confidence interval by:

$$\hat{Y} \pm z_{0.95}\frac{S}{\sqrt{N}}, \text{ and } \hat{Y} = \frac{1}{N}\sum_{i=1}^{N} Y_i \tag{7}$$

with normal score $z_{0.95} = 1.645$ and S as sampling standard deviation defined by $S = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(Y_i - \hat{Y})^2}$.

1) Absence of Autocorrelation

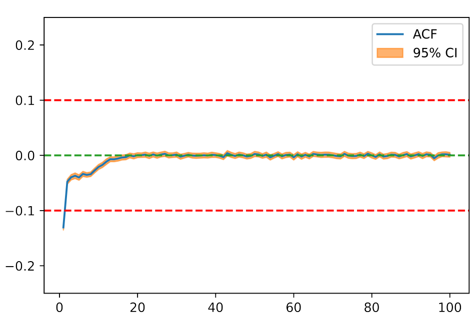


Fig. 2: Autocorrelation function

It is well established that the autocorrelation of the return series of the financial market is insignificant except for the short time interval where the microstrcture effect present, which demonstrates the unpredictability of the future movement of the price series [19]. The autocorrelations of simulated return series are calculated by:

$$C(\tau) = corr(r(t), r(t+\tau)) \tag{8}$$

where $corr(a, b)$ is the correlation function and τ is the lag. Figure 2 shows the autocorrelation function and its 95% confidence interval of simulated returns series for 100 lags. The autocorrelation for the short lags is relatively larger than those of long lags and gradually decays to zero as lag increases, which shows our simulation result is consistent with empirical findings.

2) Aggregational Gaussianity

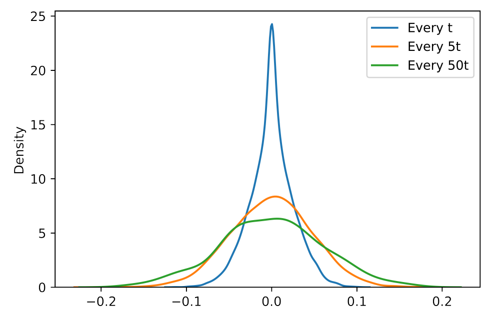


Fig. 3: Return distribution for different time scale

From the empirical study, the distribution of returns will gradually approach a normal distribution shape as the return calculation interval increases. Instead of using Monte Carlo simulation, a single market simulation is run with an extended period of 30000 steps. The return distributions of log return calculated in different intervals are shown in Figure 3. As shown in the picture, the distribution will gradually approach the normal form and become less fat-tailed form with the increase of the return calculation interval.

3) Heavy Tail Distribution of Returns

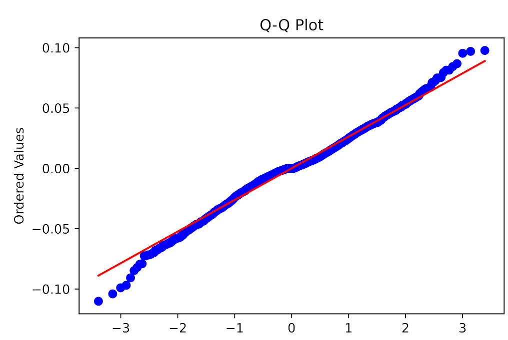


Fig. 4: Q-Q plot for one simulation returns, with steps=2000

Figure 4 shows the Q-Q plot of one simulation returns with simulation steps equal to 2000. From the plot, the fat-tailed distribution is witnessed in our simulation model. Another way to test for fat-tailed distribution is to calculate the kurtosis of the return series. The kurtosis is defined as:

$$Kurt(x) = \frac{E[(x-\mu)^4]}{(E[(x-\mu)^2])^2} \tag{9}$$

where μ is the mean of the distribution. For the distribution with fat-tailed property, the kurtosis value will be larger than 3. Calculating the kurtosis for 200 return series simulated with the Monte Carlo method, a one-sided t-test can be conducted with the following hypotheses:

H_0: The kurtosis values of simulated return series is equal to 3.0;

H_a: The kurtosis values of simulated return series is larger than 3.0.

The results of return series kurtosis and t-test result is summarized in Table II, which indicates the fat-tailed property is statistically significant across all simulated return series.

Measurements	Value
Average kurtosis values	3.766
Standard Deviation of kurtosis values	0.189
p-value of t-test	6.879×10^{-126}

TABLE II: Monte Carlo simulated kurtosis observations

4) Gain/Loss Asymmetry

The gain/loss asymmetry suggests the drawdown in stock prices will be larger than the upward movement. The property can be verified by examing if the skewness:

$$Skew(x) = \frac{E[(x-\mu)^3]}{(E[(x-\mu)^2])^{\frac{3}{2}}} \tag{10}$$

of the simulated return series distribution is negative. Similar to the kurtosis test in fat-tailed property, the skewness is calculated for the 200 return series generated by the Monte Carlo simulation. A one-sided t-test is conducted to ensure negative skewness property is statistically significant across all simulated return series with hypotheses:

H_0: The skewness values of simulated return series is equal to 0;

H_a: The skewness values of simulated return series is less than 0.

The result of table III verifies the gain/loss asymmetry property in the simulated return series.

Measurements	Value
Average skewness values	-0.064
Standard Deviation of kurtosis values	0.069
p-value of t-test	3.614×10^{-29}

TABLE III: Monte Carlo simulated skewness observations

5) Leverage Effect

The leverage effect states the correlation between returns and subsequent squared returns:

$$C_l(\tau) = corr(r(t), r(t+\tau)^2) \tag{11}$$

would start from negative value and gradually diminish to values close to zero [19]. Figure 5 shows the correlation between simulated log-returns and their subsequent squared returns for lag from 1 to 20. From the plot, the correlation starts from -0.038 and gradually decreases to 0 after lag 3, which indicates the leverage effect is presented in the simulated return series.

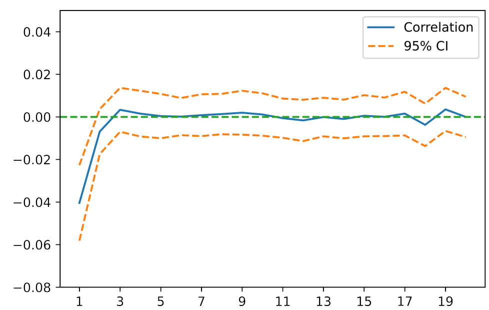


Fig. 5: Correlation between returns and subsequent squared returns

V. Spoofing Simulation

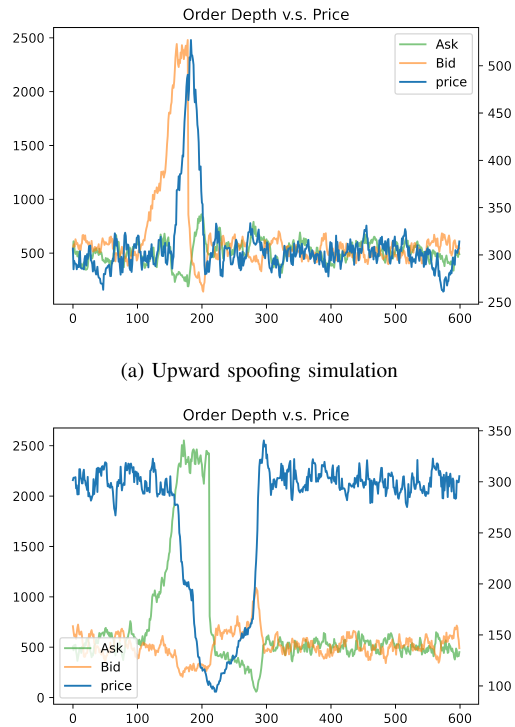
A. Simulation Result

In the previous part, the ability of the model to replicate the realistic financial time series is verified by checking the stylized facts. To simulate the market under spoofing conditions, the spoofing agents with parameters are listed in Table IV are introduced to produce upward price manipulation and downward price manipulation.

	Spoofing Agent A	Spoofing Agent B
Price Manipulating Direction	Upward	Downward
Arrival Time (S_t)	1000	1000
Spoofing Limit Order Price (S_p)	500	100
Trading Volume Ratio Limit (S_l)	0.6	0.6
Trading Volume Increase Step (S_s)	0.01	0.01
Spoofing Order Relative Price Level (S_i)	0.7	1.3

TABLE IV: Spoofing agent parameters

At the arrival time S_t, the spoofing agent will enter the market with the purpose of manipulating the price to either upward or downward. In order to profit from the market manipulation activity, the spoofing agent will establish a position at the arrival time and place a spoofing limit order at the price level (S_p) which it intended to manipulate and cancel after the limit order is traded. At each simulation time after arrival point, the spoofing agent will feign spurious buying or selling pressure by supplying the spoofing order to the bid or ask side at best bid price or best ask price multiply S_i. It is common to assume that the spoofers would be high-frequency traders who rapidly modify and cancel the submitted orders. To reflect this characteristic in the modeling, at each simulation time, the spoofing agent will cancel all submitted orders and place new orders to avoid the risk of spoofing orders being traded with other agents. The volume of the spoofing orders is proportional to the current market volume with an upper ratio limit S_l, and the ratio gradually increases by a step of S_s. The upward and downward manipulations are performed during the simulation by spoofing agent A and spoofing agent B, respectively. The two agents are only different at their price manipulation direction and the price level they decided to influence the market.



(a) Upward spoofing simulation

(b) Downward spoofing simulation

Fig. 6: Upward and downward spoofing simulation

Figure 6 presents the results of spoofing simulation. Once the spoofing agent arrives at the market, it will submit the spoofing orders to the corresponding side and gradually exert the influence on the order book imbalance. On another side of the market, the liquidity begins to decrease as the price changes. After the limit order of the spoofing agent is traded and the price goes beyond its desired price level, the spoofing agent would stop updating its spoofing orders, and the order depth of spooging side suddenly dropdown. With the interaction of fundamentalist and chartist agents, the price eventually recovers to the price level of normal market condition.

B. Market Dynamics

Figure 7 shows the heuristic curve of the fundamentalist and chartist agents. At each simulation time, the average value of the heuristic across all agents is calculated to represent the overall trading actions for the type of agents. For both types of agents, the positive heuristic value would signal a buy order, and the negative heuristic value would signal a sell order.

When the spoofing agents arrive at the market and artificially create an order imbalance, the fake increasing buying

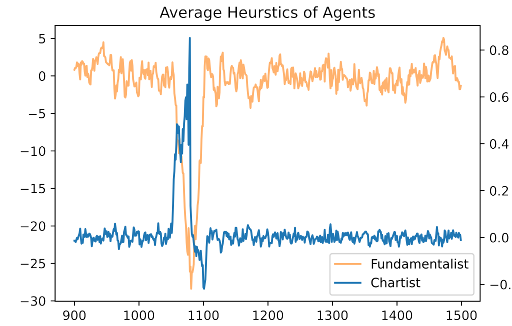


Fig. 7: Average heuristics values for agents during spoofing

pressure would coordinate a part of the chartist agents with lower order imbalance thresholds to place the buy orders and initiate a price divergence. In the meantime, the initiated price increasing trend would further activate more chartist agents. The joint forces of the increasing order book imbalance and upward price movement trend eventually caused the rapid increase in the average heuristic value of chartist agents. The fundamentalist agents perceive the price variation and keep the false belief the price would return to its fundamental value and place sell orders. Due to the trading speed difference, the buying power of coordinated chartist agents dominates the market, and price increases.

After the spoofing limit order is traded and the profit is obtained, the spoofing agent cancels all spoofing orders. With the absence of misleading buying power and the price correction exerted by fundamentalist agents, the coordination in chartist agents rapidly disappeared, which can be observed by their heuristic value quickly reversing to zero. The price correction force of fundamentalist agents will initiate the price recovery process. With the continuous price decreasing trend identified in the market, some chartist agents would place the sell orders to join the movement and exacerbate the trend. The synchronized trading actions would create an order book imbalance reversed to the spoofing time (the order book imbalance can be observed in Figure 6), which in turn activates more chartist agents. In the end, the price gradually enters equilibrium and returns to normal market condition.

C. Market Volatility

The volatility estimation of the market is measured by the standard deviation of the return of the market price. At each simulation time t, the return of the market price is measured by:

$$R_t = \frac{P_t - P_{t-1}}{P_{t-1}} \tag{12}$$

where P_t is the current market price. The volatility during the spoofing time is measured within the spoofing agent arrival time (S_t), and when the price first recovery back around its fundamental value. The recovery time S_r is defined by:

$$S_r = \underset{t}{\operatorname{argmax}}(|P_t - F| < T) \tag{13}$$

where $F = 300$ and $T = 10$ for the simulation where we will consider the price has returned to the normal price level.

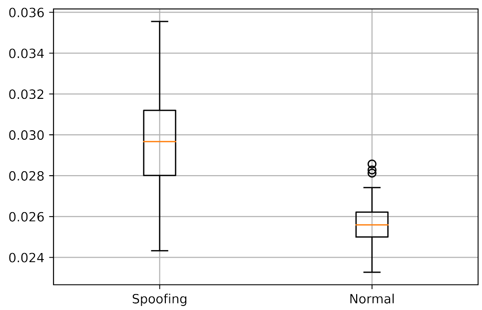


Fig. 8: Volatility estimation

Figure 8 shows the volatility estimations during the spoofing and normal market condition for 50 independent simulations. The volatility estimation of normal market condition is calculated with a pooled return series with a window size of 500. The plots show that volatility during the spoofing time diverges from the normal market condition. The presence of spoofer and its market manipulation activity would exacerbate the market volatility and increase the price variations. Therefore, the presence of spoofing agent would also prolong the normal market price discovery process.

D. Agents' Profit

To understand how spoofing affects different types of agents' profit, we need first to discuss how the profit is obtained for different agents. The agents profit by making an expectation about the future price movement. And we will only take the traded limit order into our consideration. For a given traded order, the profit (Ω) of agents' limit order at time t is calculated by:

$$\Omega(t) = (P_t - \hat{P}_{t-1})Sign(J(t^*)) \tag{14}$$

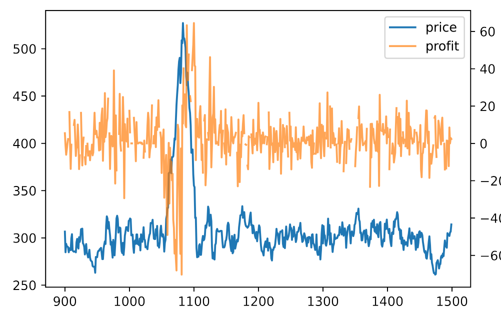
where P_t is current market price, the $\hat{P}_{t-1}$ is the limit order price traded at time $t-1$ and the $J(t^*)$ is the agent heuristic value when the limit order is placed. The $Sign(x)$ function is defined as:

$$Sign(x) = \begin{cases} 1, & \text{if } x > 0. \\ -1, & \text{otherwise} \end{cases} \tag{15}$$

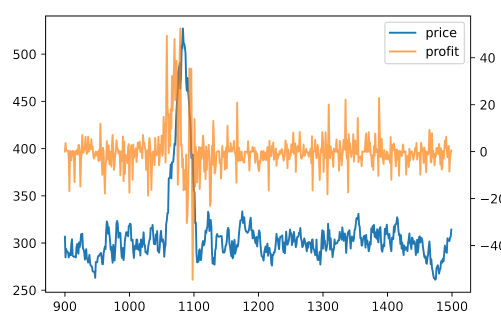
For both fundamentalist and chartist agents, their trading action is correlated with their heuristic by:

- A bid order if $J(t) > 0$
- A ask order if $J(t) < 0$

If the limit order is a bid order, when the order is traded, the agent's expectation about the price movement will give it profit with long position by $P_t - \hat{P}_{t-1}$(with $Sign(J(t^*)) = 1$). If the limit order is ask order, the agent will achieve profit with short position by $\hat{P}_{t-1} - P_t$(with $Sign(J(t^*)) = -1$). According to the order price sampling distribution, the agents may sample an order price across the current best bid or ask price to express their intention to trade, and the orders will be matched immediately. Therefore, these orders will be regarded as market orders, and the order price will be the current market price. Figure 9 presents the instantaneous profit of fundamentalist and chartist agents during the spoofing simulation. The fundamentalist agents expect the price would move toward the fundamental value when the price diverges. But with the influence of spoofing agent, the equilibrium of price dynamics no longer exists. As a result, the fundamentalist agents continuously make false expectations of price movement and suffer losses from their trading activity. After spoofing, the price gradually reverses to the normal price level in the price recovery process, and the fundamentalist agents would be able to make profits. For the chartist agents, its heuristic function enables them to capture and follow the sharp price movement. Thus, they would be able to make profits from rapid price upward movement and downward movement. But they would suffer a loss when spoofing limit order is traded and the price recovery process started, at which they falsely believe the upward price movement will continue.



(a) Fundamentalist agents



(b) Chartist agents

Fig. 9: Profit during spoofing simulation

E. Agents' Sharpe Ratios

The Sharpe ratio is a measurement of an investment return compared to its risk. The higher Sharpe ratio indicates the investor can obtain more return with a given risk level. Before calculating the Sharpe ratio for different agents, the return series of agents needs to be determined first. For a given agent, its cumulative wealth is measured by its cumulative position and the price the agent cost to establish that position. At a given time t, the agent accumulative wealth($W(t)$) is:

$$W(t) = \sum_{i=\hat{t}}^{t} O_i \cdot \hat{P}_i \cdot Sign(J(i)) \tag{16}$$

where O_i is the order size for limit order placed at time i, the $\hat{P}_i$ is the limit order price, and the $J(i)$ is the agent heuristic value when the limit order is placed. As in the previous part, if the limit order's price is across the best bid/ask price, it will be regarded as a market order, and the order price will be the current market price. And the agent's cumulative net position ($N(t)$) at time t is calculated by:

$$N(t) = \sum_{i=\hat{t}}^{t} O_i \cdot Sign(J(i)) \tag{17}$$

At each time t, the agent would re-evaluate the cost ($W^{'}(t)$) to establish the same net position with the current market price by:

$$W^{'}(t) = N(t) \cdot P_t \tag{18}$$

and the agent's return at time t is measured by:

$$R(t) = \begin{cases} \frac{W^{'}(t)-W(t)}{W(t)}, & \text{if } N(t) \geq 0. \\ \frac{W(t)-W^{'}(t)}{W(t)}, & \text{if } N(t) < 0. \end{cases} \tag{19}$$

In other words, at each time t, the agent's return is measured in an unrealized manner. If the cost to establish the same cumulative position with the current market price is greater than the agent's actual cost (wealth) with its trading activities, the agent will have a positive return. Otherwise, the agent would have a negative return at time t. And if the agent's net position is negative, we will assume the agent is shorting the security, and its return would be the opposite of the agent's return as it takes a long position. With the return series $R(t)$, the agent's Sharpe ratio is measured by:

$$\text{Sharpe ratio} = \frac{E[R(t)]}{\sigma(R(t))} \tag{20}$$

As stated in the previous part, the spoofing agent profits by establishing the position when it enters the market, placing a limit order in the opposite direction, and manipulating the price to the level where the limit order is traded. To calculate its Sharpe ratio, when the spoofing agent arrivals at the market, we will assume it places a unit size order with market price and unit size limit order with the price it intends to manipulate.

Figure 10 presents the boxplot of the average Sharpe ratio in different agent types for 50 independent simulations. The beginning of the evaluation window is set to the arrival time of the spoofing agent, and the spoofing end time follows the definition in Equation (13). From the plot, it is clear that the market manipulation activity of the spoofing agent would give it an unfair advantage and result in a significantly higher Sharpe ratio than the other two types of agents. Compared to the fundamentalist agents, the Chartist agents would have more benefit opportunities and be able to benefit not only from the price recovery process but also from the price increase process. Thus, they have a higher Sharpe ratios than the fundamentalist agents.

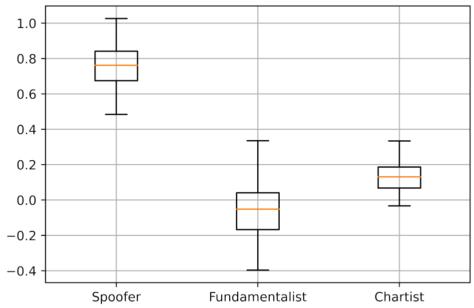


Fig. 10: Boxplot for Sharpe ratios of different agent types

VI. Conclusions

In this paper, we introduce an agent-based simulation model to replicate the spoofing market manipulation activity. The model consists of four types of agents: fundamentalist, chartist, zero-intelligence agent, and spoofing agent. The agents decide their trading actions with their heuristic function, the parameters of agents are sampled from uniform distributions to ensure their heterogeneity. Without the presence of the spoofing agents, the market can form an equilibrium price series, and its statistical property is verified with the stylized facts, such as the absence of autocorrelation, aggregational gaussianity, heavy tail distribution of returns, gain/loss asymmetry, and leverage effect. During the spoofing simulation, the spoofing agents manipulate the market by supplying the spoofing orders, and the model can simulate both upward and downward price movement of spoofing activity.

In addition, the market dynamics, volatility, and agents' profit are analyzed. From the simulation, the spoofing orders of the spoofing agent would exert the order book imbalance and coordinate the chartist agents' trading actions to manipulate the price level of the market. After the spoofing agent obtained its profit and exited the market, the market would return to its original price level with the interactions of chartist and fundamentalist agents. The presence of the spoofing agent and its market manipulation activity would increase the market volatility and exacerbate the price variations. The fundamentalist agents would suffer a loss during the spoofing time but would be able to make a profit during the price recovery process. The chartist agents would suffer a loss when the spoofing agent realized its profit and the price recovery process started, at which they falsely believed the price movement trend would continue. The Sharpe ratio analysis also indicates the market manipulation activity of the spoofing agent would give it an unfair advantage and result in a significantly higher Sharpe ratio than the other two types of agents.

This simulation model can be further extended by adding more types of agents. The current model only consists of four types of agents, which may be limited in representing the role of various investors in the real financial market. The model fidelity to the real market can be further improved by incorporating more agent types such as market makers, liquidity suppliers, and arbitrageurs. Furthermore, adding the agents with learning ability would also be a possible direction. In the existing model framework, the agents' behaviors are well defined by their heuristic functions, which fails to capture the learning characteristics of investors and makes the agents always act in the same way no matter how the market condition changes.

VII. Acknowledgment

We want to thank Dr. Dan Pirjol who served as the reader for this research project. He provided technical guidance in measuring agents' performance and suggestion for agent behaviors. The authors also want to thank the staff who supported the Master thesis work.

References

[1] T. Hendershott, C. M. Jones, and A. J. Menkveld, "Does algorithmic trading improve liquidity?" *The Journal of finance*, vol. 66, no. 1, pp. 1–33, 2011.

[2] T. Hendershott, R. Riordan *et al.*, "Algorithmic trading and information," *Manuscript, University of California, Berkeley*, 2009.

[3] J. D. Montgomery, "Spoofing, market manipulation, and the limit-order book," *Available at SSRN 2780579*, 2016.

[4] E. M. Aldrich, J. Grundfest, and G. Laughlin, "The flash crash: A new deconstruction," *Available at SSRN 2721922*, 2017.

[5] CFTC and SEC, "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," 2010.

[6] J. C. Jackson, D. Rand, K. Lewis, M. I. Norton, and K. Gray, "Agent-based modeling: A guide for social psychologists," *Social Psychological and Personality Science*, vol. 8, no. 4, pp. 387–395, 2017.

[7] T. Mizuta, S. Kosugi, T. Kusumoto, W. Matsumoto, K. Izumi, I. Yagi, and S. Yoshimura, "Effects of price regulations and dark pools on financial market stability: an investigation by multiagent simulations," *Intelligent Systems in Accounting, Finance and Management*, vol. 23, no. 1-2, pp. 97–120, 2016.

[8] S. Nagumo, T. Shimada, N. Yoshioka, and N. Ito, "The effect of tick size on trading volume share in two competing stock markets," *Journal of the Physical Society of Japan*, vol. 86, no. 1, p. 014801, 2017.

[9] X. Yang, J. Zhang, and Q. Ye, "Tick size and market quality: Simulations based on agent-based artificial stock markets," *Intelligent Systems in Accounting, Finance and Management*, vol. 27, no. 3, pp. 125–141, 2020.

[10] R. M. Bookstaber and M. E. Paddrik, "An agent-based model for crisis liquidity dynamics," *Office of Financial Research Working Paper*, no. 15-18, 2015.

[11] S. Y. Yang, A. Liu, and S. Y. K. Mo, "Twitter financial community modeling using agent based simulation," in *2014 IEEE Conference on Computational Intelligence for Financial Engineering & Economics (CIFEr)*. IEEE, 2014, pp. 63–70.

[12] M. Paddrik, R. Hayes, A. Todd, S. Yang, P. Beling, and W. Scherer, "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)*. IEEE, 2012, pp. 1–8.

[13] R. Hayes, A. Todd, N. Chaidarun, S. Tepsuporn, P. Beling, and W. Scherer, "An agent-based financial simulation for use by researchers," in *Proceedings of the Winter Simulation Conference 2014*. IEEE, 2014, pp. 300–309.

[14] F. H. Westerhoff, "The use of agent-based financial market models to test the effectiveness of regulatory policies," in *Agent Based Models for Economic Policy Advice*. De Gruyter Oldenbourg, 2016, pp. 195–227.

[15] S. J. Leal, M. Napoletano, A. Roventini, and G. Fagiolo, "Rock around the clock: An agent-based model of low- and high-frequency trading," *Journal of Evolutionary Economics*, vol. 26, no. 1, pp. 49–76, 2016.

[16] S. J. Leal and M. Napoletano, "Market stability vs. market resilience: Regulatory policies experiments in an agent-based model with low-and high-frequency trading," *Journal of Economic Behavior & Organization*, vol. 157, pp. 15–41, 2019.

[17] X. Wang, C. Hoang, Y. Vorobeychik, and M. P. Wellman, "Spoofing the limit order book: A strategic agent-based analysis," *Games*, vol. 12, no. 2, p. 46, 2021.

[18] S. Maslov and M. Mills, "Price fluctuations from the order book perspective-empirical facts and a simple model," *Physica A: Statistical Mechanics and its Applications*, vol. 299, no. 1-2, pp. 234–246, 2001.

[19] R. Cont, "Empirical properties of asset returns: stylized facts and statistical issues," *Quantitative finance*, vol. 1, no. 2, p. 223, 2001.