Impacts of Social Networks in an Agent-based Artificial Stock Market

Xiaoting Dai[a], Jie Zhang[b], Victor Chang[c,1]

[a]*School of Business, Jiangnan University, Wuxi 214122, China, Email address:Xiaoting.Dai@alumni.xjtlu.edu.cn and 452608555@qq.com*

[b]*International Business School (Suzhou), Xi'an Jiaotong-Liverpool University, Suzhou 215123, China, Email address:Jie.Zhang@xjtlu.edu.cn*

[c]*Artificial Intelligence and Information Systems Research Group, School of Computing, Engineering and Digital Technologies, Teesside University, Middlesbrough, UK, Email address: victorchang.research@gmail.com and V.Chang@tees.ac.uk*

Abstract

We propose an agent-based artificial stock market to investigate the influences of social networks on the financial market. The artificial stock market contains four types of traders whose information sets and trading strategies are different. Genetic Programming is employed in informed and uninformed traders' learning behavior and heterogeneity with the application of artificial intelligence. When information is exogenous, social networks result in higher market volatility and trading volume, and decrease price distortion and bid-ask spread. When information is endogenous, the influences of social networks on the financial market are reversed, which indicates that social networks harm market efficiency, decreases the trading volume and increases bid-ask spread. The reason is that social networks harm information production after traders tend to rely on information from communication, instead of spending a cost on it.

Keywords: social networks, artificial intelligence, genetic programming, volatility, price distortion, trading volume, bid-ask spread

[1]Correspondence to: Victor Chang, Teesside University, UK Emails address: victorchang.research@gmail.com and V.Chang@tees.ac.uk

Funding information: VC Research (VCR 0000146)

1 Introduction

The literature has long acknowledged that information moves stock prices. In financial markets, information transmission among market participants is one of the most pervasive features. Traders rationally aggregate information and incorporate it into investment decisions. Studying how information spreads among traders in financial markets and into stock prices becomes a popular topic. A lot of direct channels, such as communication among neighbors and friends or other social interactions, can be documented as information sharing through social networks. Recently, more people migrate their activities and communication to the web with the popularity of web-based applications. A lot of literature has investigated the influence of communication through social networks on financial markets, which concludes that they have significant implications for traders' decision making on stock participation and the whole financial market (e.g. Shiller and Pound, 1989; Cohen et al., 2008; Brown et al., 2008; Shive, 2010; Colla and Mele, 2010; Ozsoylev and Walden, 2011; Han and Yang, 2013 and Hvide and Ostberg, 2015). These papers suggest that the dissemination of information through social networks is crucial for many market outcomes, such as price efficiency, liquidity, trading profits, etc. This paper employs an application of artificial intelligence to investigate the exact influences of social networks on the financial stock market.

Many models are introduced to explore the complex network theory. Colla and Mele (2010) propose a cyclical network and introduce that social networks help improve market efficiency by considering the trading volume and price informativeness. They conclude that an investor will have positively correlated trades when he is close in the network, such as neighbors, friends and etc. Thus, social networks help improves the market depth and trading volume. Ozsoylev and Walden (2011)'s theoretical framework proves that social networks help improve market efficiency after introducing a rational expectations equilibrium model with general large-scale social networks. They conclude that trading volume increases with network connectedness, while the price volatility and network connectedness have a non-monotonic relationship. Moreover, they also emphasize that an increase in network connectedness makes the price reveal more information resulting in higher market efficiency. In addition, Han and Yang (2013) complement those studies by investigating when there is endogenous information, social communication can oppositely influence market qualities, which is exogenous information. They give evidence that social communication help improves market efficiency, trading volume and reduces the cost of capital if there is exogenous information. However, those results are opposite when information is endogenous. In recent years, social communication has developed rapidly with increasing online social media platforms. Bu et al. (2021) use the text mining methods and find that social media can significantly influence the relationship between customers and hosting firms. They confirm the importance of social networks in analyzing the customer perceived value of products. Liu et al. (2021) first investigate opportunistic behavior in supply chain nance by drawing on a social media perspective. They collect social media data from the biggest Chinese micro-blog platform and find that information governance by social networks could moderate the opportunistic behavior, which had negative effects on all the supply chain nance participants. Ram and Zhang (2021) use Nvivo coding and matrix queries to analyze the data from social media, manufacturing, IT, service industries and telecommunication. They conclude that social media analytics help expands competitive intelligence to businesses beyond the known scope of competitor analysis before. The social network analysis method is used to identify the factors influencing the seasonality effect in the stock market by Kajol et al. (2020). They show that volatility is the most important factor of the seasonality effect. They suggest the policymakers should hold training to traders through an awareness campaign. Sohaib (2021) employs partial least squares structural equation modeling methods to provide an empirical analysis of the effects of social networking services on social trust towards social commerce intention.

There is rare literature contributing to empirical or agent-based intelligent analysis of social network influences. We fill the literature gap by proposing artificial intelligence instead of theoretical models into this line of research. In this paper, we follow the settings introduced by Han and Yang (2013) and use an artificial stock market (ASM), which can be seen as an application of artificial intelligence, to demonstrate the effects of social communication on market outcomes. We compare these influences by proposing artificial intelligence with those generated by the model of Han and Yang (2013). It is interesting to see whether similar settings between ASMs and theoretical models will bring similar conclusions. In this regard, we provide another method to follow the research with complex networks.

This paper contributes to working on the advanced artificial intelligence for complex networks and can be seen as one of artificial intelligence-powered approaches in simplifying complex networks. It might be challenging for empirical studies to identify the dissemination of information through complex social networks since direct communications among traders are not easily observed and documented. Specifically, more people migrate their activities and communication to the web with the popularity of web-based applications. The range and volume of social networks become unprecedentedly large, which are difficult to generate efficiently. There are still studies providing evidence that indirect proxies of social communication could do the same, such as common schooling (Cohen et al., 2008), geographic proximity (Brown et al., 2008) and coworkers (Hvide and Ostberg, 2015). Nevertheless, Ahern (2017) suggests that although these proxies might indicate social communication, they could also reflect the homophily among these traders because they have the same background and might act alike in decision-making. After suffering from the data collecting, Cecconi and Campenni (2019) point out that agent-based artificial finance makes it possible to see how macro-outcomes will appear and how an equilibrium state will be reached through endogenous interactions from large numbers of autonomous and heterogeneous agents, rather than from a typical isolated individual behavior. Moreover, the application of artificial intelligence in stock market trading develops and popularizes accelerated recently. For example, stock price prediction is one popular topic by using artificial intelligence (e.g. Fischer and Krauss, 2018;Long et al., 2019;Zhong and Enke, 2019 and Nabipour et al., 2020). Thus, we employ artificial intelligence, the ASM with Genetic Programming (GP), to overcome the difficulty in the lack of data and extend the line of research by demonstrating the effects of social communication on the financial markets. Artificial intelligence has been introduced as one way to extract valuable data from complex networks and further analyze network performances. Artificial intelligence has received great concern from the research field in recent years. Especially, more studies increase the attention on artificial intelligence after the event of COVID because of decreasing human elements in society and increasing automation(Jaklic et al. 2019 and Collins et al. 2021). These artificial intelligence algorithms are able to capture the complex micro characteristics of investors sufficiently, including heterogeneous beliefs and trading strategies with intelligence. Following the artificial intelligence-powered approach, we distinguish different types of traders at the beginning and form the complex social networks ex-ante. Information is then independent and severally received or transferred by relative traders. Complex networks are then can be simplified in this way.

In this paper, we first aim to build the ASM reasonable and reliable after the model is calibrated. Following that, we seek to investigate the implications of social networks for market qualities that can be explained by market volatility, price distortion, trading volume and bid-ask spread by using artificial intelligence. We also examine the comparison of the influences of social networks on financial markets when information is exogenously given versus when information has endogenously acquired a cost.

We contribute to constructing the agent-based limit order ASM by setting two assets in the market, a risk-free asset with a constant interest rate per period and a risky asset, based on the framework presented in Yeh and Yang (2010). The ASM contains four types of traders: informed, social, uninformed and noise traders. We call informed, social and uninformed traders are rational traders in the market. Rational traders employ CARA utility with a risk-averse coefficient and they can choose whether to spend a cost on diverse private signals regarding the stock payoff before making investment decisions on the risky asset. At first, informed traders receive private and public information, while uninformed traders only can receive public information from the market. Informed and uninformed traders use the technique of GP, as one application of artificial intelligence with complex networks, to evolve their learning behaviors and then update their trading strategies to improve their investment decisions. Informed traders can communicate with friends (social traders) in their group. Social traders will then also receive the information, however, with a noisy signal. The financial market then opens and all traders make their own investment decisions in the market. Following that, trade information with the bid or ask orders is updated on the board and observable for all traders.

Our paper examines that social networks help increase market volatility and trading volume and decrease bid-ask spread when information is exogenous. The reason might be that sharing information with group members enlarges traders' information set. They can estimate more precisely the stock, thereby increasing their trading aggressiveness and leading the market more liquid. Social networks also improve market efficiency by decreasing price distortion when information is exogenous, since there is more information being impounded into the market after sharing information via social networks. When information is endogenous acquired at a cost, the fraction of informed traders decreases as two free-riding channels through social communication (Han and Yang, 2013). The first channel is "free-riding on friends", in which communication among friends will influence traders' ex-ante incentive to acquire more information. They attempt to rely on the information via social communication, rather than spending a cost on that. The second is "free-riding on price". Traders prefer to communicate information and learn from market price and directly generate social communication benefits on price informativeness. Thus, we investigate that social communication harms market efficiency, decreases the trading volume and increases bid-ask spread under endogenous information while increasing market volatility.

The rest of this paper is organized as follows. In Section 2, we introduce the design of our ASM and detailed settings. We compare statistical features of our artificial stock market with real stock markets in the world in this section. Section 3 gives the results and explanations. In section 4, we perform the sensitivity analysis to test how the simulation results are sensitive to the parameters of the ASM. Finally, we make some concluding remarks in Section 5.

2 Experimental Design

2.1 Methodology

The application of artificial intelligence of modeling an ASM with interacting agents from the bottom up has three advantages. The first advantage is that ASM allows us to set up experiments (simulation) in a controlled environment to ensure the accuracy of the conclusion. ASM is also able to capture the complex micro characteristics of investors sufficiently, including heterogeneous beliefs and trading strategies with intelligence. Due to the lack of data, empirical studies cannot easily be set up to trace the activities of social networks from traders. An artificial stock market would no longer have limitations of data availability. Our ASM model is based on the framework of Yeh and Yang (2010), which is followed the model of Santa Fe ASM of Arthur et al. (1997) and Lebaron et al. (1999), and the studies of Brock and H.Hommes (1998).

2.2 Market structure

2.2.1 Assets

Two assets are introduced in the market. One is a risk-free asset (cash) paying interests at a constant rate. The gross return of the risk-free asset is $R = 1 + r_f$ for one period, where r_f is risk-free interest rate. The second asset is a risky asset (stock), in which paying a stochastic dividend which is assumed to be a AR(1) (the first-order auto-regressive) process:

$$D_{t+1} = D + \rho(D_t - D) + \mu_{t+1} \tag{1}$$

where D is the average dividend over a long period. ρ is a coefficient to indicate how fast the dividend value approaches the average value. μ is positive white noise, $\mu_t = N(0, \sigma_u^2)$. The setting of dividends is similar to that used in Lebaron et al. (1999). This process helps to avoid the dividend process getting too close to non-stationary dividend processes.

2.2.2 Wealth

The trader i's wealth at t, W_t is given by

$$W_t = RW_{i,t-1} + (P_{t+1} + D_{t+1} - RP_t)h_{i,t} \tag{2}$$

where $W_{i,t-1}$ is the trader i's wealth at $t-1$, P_t is the current stock price per share and $h_{i,t}$ indicates the shares of the stock held by trader i at time t.

2.2.3 Traders

In this market, we consider four types of traders.

- *Informed traders (I)*: I traders can receive private and informative signals about the dividend and fundamental value of the risky asset. They are rational investors in the market. I traders generate an expectation of future stock prices via updated trading strategies through a learning system (namely GP), which will be further discussed later). We assume that there is one I trader in each group.
- *Social traders (S)*: S traders do not acquire any private information. They only update trading strategies through social communication within the group. We assume each S trader receives a noisy version of the result through social communication.
- *Uninformed traders (U)*: U traders are also rational agents. Like I traders, they update trading strategies through GP by using public information, such as recent dividends and stock prices.
- *Noise traders (N)*: N traders trade on a spurious signal that they believe is informative, but in fact, the signal is a purely noise. They do not have a learning system but have a biased belief about next period stock price as the previous period clearing price, given by a noise term.

2.2.4 The learning system and traders' beliefs

GP is defined as a technique by running computer programs that are encoded as a set of genes that evolved using an evolutionary algorithm. It is popular in the application of artificial intelligence. Langdon and Poli (2002) provided a detailed explanation of the technical issues of GP. In our model, informed traders studied the market and updated their trading strategies via GP. The function set and terminal set are displayed in table 1 . Traders have heterogeneous learning frequencies between 5 to 95 periods. They form their believes about future stock price and dividends as following:

$$E_{i,t}(P_{t+1}+D_{t+1}) = \begin{cases} (P_t+D_t)\left[1+\theta_0 \tanh(\frac{\ln(1+f_{i,t})}{\omega})\right] & if\ f_{i,t} \geq 0 \\ (P_t+D_t)\left[1-\theta_0 \tanh(\frac{\ln(|-1+f_{i,t}|)}{\omega})\right] & if\ f_{i,t} < 0 \end{cases} \tag{3}$$

where $f_{i,t}$ is determined as a forecasting accuracy indicator. θ and ω are constants.

We employ a setting for noise traders similar as Banerjee and Green (2015). Noise traders form their beliefs without learning, but purely utilizing a biased belief about the previous period clearing prices:

$$E_{i,t}(P_{t+1}+D_{t+1}) = P_t + D_t + \varepsilon_{i,t} \tag{4}$$

where $\varepsilon_{i,t} = N(0, \sigma_\varepsilon^2)$ is biased belief of trader i.

The reservation price (P_i^R) at time period t is derived based on the expectation of each trader, in which it follows Yeh and Yang (2010).

$$P_i^R = \frac{E_{i,t}(P_{t+1}+D_{t+1}) - \lambda h_{i,t} V_{i,t}(R_{t+1})}{R} \tag{5}$$

where R_{t+1} is the excess return at time $t+1$, i.e. $P_{t+1} + D_{t+1} - RP_t$, and $V_{i,t}(R_{t+1})$ is the forecast of trader i regarding the conditional variance at time $t+1$ given his/her information up to t.

2.2.5 Price determination

After traders collect information and update their reservation price P_i^R with their own trading believes, each trading round starts with a random trader entering the stock market. The random trader determines his order of bid or ask with following other random traders. The rules of how they make orders work as follows.

The price determination of this paper is realized by a simplified continuous double auction process. The process is similar to the framework of Yeh and Yang (2010), which is important in the application of artificial intelligence. There are N rounds of the continuous double auction process in each period, which determine trading information.

The highest price for buying (the best bid B_b) and the lowest price for selling (the best ask B_a) are observable for traders. Traders make orders, i.e. accept a bid (an ask) or submit an order based on their own reservation prices of the risky asset. The process works as follows. Traders may come across four scenarios: (1) both B_b and B_a exist; (2) only B_a exists; (3) only B_b exists; and (4) neither bid nor ask exists. In the first scenario, the trader will either post a market buy (sell) order at $B_a(B_b)$ when$P_i^R > B_a(P_i^R < B_b)$; or post a limit buy (sell) order at his reservation price when $B_b \leq P_i^R \leq B_a$ and $P_i^R \geq (B_a + B_b)/2$ $(P_i^R < (B_a + B_b)/2)$. In the second scenario, the trader will post a market buy order at B_a when $P_i^R > B_a$; or he will post a limit buy order at his reservation price when $P_i^R \leq B_a$. Under the third condition, the trader will post a market order and sell at B_b when $P_i^R < B_b$;or he will post a limit sell order at his reservation price when $P_i^R \geq B_b$. Under the fourth condition, the trader will post a limit buy or a limit sell order at his reservation price in equal chance.

(1) Both B_b and B_a exist,

- If $P_i^R > B_a$, the trader will post a market buy order at B_a.
- If $P_i^R < B_b$, the trader will post a market sell order at B_b.
- If $B_b \leq P_i^R \leq B_a$and $P_i^R \geq (B_a + B_b)/2$, the trader will post a limit buy order at his reservation price P_i^R.
- If $B_b \leq P_i^R \leq B_a$and $P_i^R < (B_a + B_b)/2$, the trader will post a limit sell order at his reservation price P_i^R.

(2) Only B_a exists,

- If $P_i^R > B_a$, the trader will post a market buy order at B_a.
- If $P_i^R \leq B_a$, the trader will post a limit buy order at his reservation price P_i^R.

(3) Only B_b exists,

- If $P_i^R < B_b$, the trader will post a market order and sell at B_b.
- If $P_i^R \geq B_b$, the trader will post a limit sell order at his reservation price P_i^R.

(4) Neither bid nor ask exists,

- The trader will post a limit buy or a limit sell order at his reservation price P_i^R in equal chance.

2.3 Experimental designs

Table 1 lists important parameters used in the experiments by employing artificial intelligence, which is the ASM, after model calibration in section 2.4. Each trader has an initial wealth of 200 bonds and 1 stock. The return for bonds is a 0.8% interest rate for each period. The average dividend is 0.2. We collect a total of 20,000 periods for each simulation run. We tried to collect data for more than 20,000 periods and found that there is no significant difference among results. We simulate 20 different runs for each market with different network connectedness and measure the average results of these runs. Thus, the following results are averaged by all simulation runs.

Table 1: Settings of the simulated system

The artificial stock market	
Initial cash (M_0)	200
The initial number of stocks	1
Stock initial price	25
Interest rate (r)	0.008
Dividend for each period	$0.2 + 0.95(D_t - 0.2) + N(0, 0.02)$
Number of periods	20,000
Number of strategies of each informed trader	2
Evolutionary cycle	2
θ_0	0.2
ω	15
λ	0.5
Function set	$\{ifelse, +, -, \times, \div, sqrt, sin, con, abs\}$
Terminal set of informed traders	$\{P_{t-1}, ..., P_{t-5}, D_t, D_{t-1}, ..., D_{t-5}, P_f\}$
Terminal set of uninformed traders	$\{P_{t-1}, ..., P_{t-5}, D_{t-1}, ..., D_{t-5}, P_{f-1}\}$
Probability of immigration P_i	0.1
Probability of crossover P_c	0.7
Probability of mutation P_m	0.2

Based on the empirical properties presented in Table 3 and the first q-q plot in Figure 1, the model specification of our simulation is calibrated to mimic these stylized facts in the real financial stock market. The important parameters in the system are then presented in this table. Each trader in the market has initial wealth of 200 cash and 1 stock. The return for cash is a 0.8% interest rate during each period. The initial stock price in the system is 25. The risky asset (stock), paying a stochastic dividend which is assumed to be an AR(1) process. 20,000 periods for each simulation run is collected. The second and third parts of Table 4 present important parameters in the learning process, which is GP.

Table 2 shows the numbers of different traders in the market. We set 50% of total traders are noise, which aims to provide liquidity in the market. The numbers of total and noise traders are constant among different experiments, which are 200 and 100. Only informed and social traders are in groups and evenly distributed among groups. The remaining participants are uninformed traders in the market. The network is formed within each group, which can be interpreted as social media, community, or friendships. Traders in each group communicate and share their information with each other. However, there are no connections between any two groups. This method we used to simulate the network is the island-connection network explained by Jackson (2008). According to Ozsoylev and Walden (2011), traders are defined as being linked when they share their information within a group. They further define traders' connectedness (N) as the number of traders' links in the network. In our experiment, we assume each trader shares their private information with others within one network, which was also introduced by Han and Yang (2013). Therefore, N is the expression of the number of traders, also each trader's

Table 2: Settings of traders in markets

Information	Exogenous					Endogenous				
No. of groups (G)	10	10	10	10	10	10	10	10	10	10
No. of traders in the market	200	200	200	200	200	200	200	200	200	200
No. of traders in each group(N)	2	4	6	8	10	2	4	6	8	10
Informed traders	10	20	30	40	50	10	10	10	10	10
Social traders	10	20	30	40	50	10	30	50	70	90
Uninformed traders	80	60	40	20	0	80	60	40	20	0
Noise traders	100	100	100	100	100	100	100	100	100	100

connectedness and network connectedness.

In the market with exogenous information, we follow Han and Yang (2013) and assume the fraction of informed traders is fixed in each group. We fix the fraction at 50%. For example, 3 informed traders and 3 social traders are in each group when there are 6 traders in each group. Therefore, the total numbers of informed and social traders are 30, respectively, with all 10 groups. There are also 100 noise and 40 uninformed traders in this market.

In the market with endogenous information, Han and Yang (2013) summarize that communication with informed traders could reduce traders' ex-ante incentive to spend a cost on private information. They conclude the action of "free-riding on friends" results in decreasing the fraction of informed traders in each group. Therefore, we follow them and assume that the proportion of informed traders reduces as social connectedness increases. Thus, we set the number of informed traders is fixed at 1 in each group among all experiments with endogenous information. By this assumption, the proportion of informed traders is 50% (1 informed trader) in each group when the network connectedness (N) is 2, while it becomes 10% (also 1 informed trader) in each group when the network connectedness (N) is 10. Therefore, the total numbers of informed, social, uninformed and noise traders are 10, 10, 80 and 100, respectively, with all 10 groups when the network connectedness (N) is 2.

2.4 Model calibration and statistics of stock price

Table 3 presents several statistical properties found in financial markets, and we further provide these features, which are also observed in our experiments. Panel A in Table 3 presents the basic statistical properties of the Nasdaq Composite Index (Nasdaq), the S&P 500 in the U.S., and DJI Index (Dow Johns Industria

Average), Hang Seng Index (HSI) in Hong Kong, and Nikkei 225 in Japan. The first column of Table 3 presents the name of these indexes and the second column shows the time periods that the stock indices are considered for analysis. The next two columns describe the minimum returns and maximum logarithmic returns in percentage. It is clear that the largest absolute daily returns range from 10.51% to 40.54%. We simply use the average of absolute returns to measure the market volatility shown in the fifth column. It ranges from 0.74% to 1.07%. The skewness of raw returns, which presents in the sixth column, are all negative. The kurtosis of raw returns of all financial markets is larger than three, indicating fat tails. We present the results of ASM in panel B. It illustrates those statistic properties of stock prices in our ASM from 160 basic simulation runs (20 runs for each experiment based on different network connectedness (N)). The stock price shows negative skewness and fat tail on average. Comparing with the empirical results summarized in Panel A, our model matches several financial markets within a reasonable range.

Figures 1 and 2 show prices and logarithmic returns of Nasdaq and the ASM (one random example from simulations). The stock price of Nasdaq contains the time during 1972-2021 where bubbles and crashes exist. Figure 3 shows two quantile-quantile plots (q-q plots) of the S&P 500 and the artificial stock market. These two q-q plots help to test whether two sets of the sample come from the same distribution. The first figure presents how the return of S&P 500 to the standard normal distribution, and the second one compares the return of our ASM to the standard normal distribution. It is obvious that the points of return of S&P 500 and our ASM all fall approximately along the 45-degree reference line. It suggests that two sets of sample data in one q-q plot come from the same distributions.

Table 3: Stylized facts of market returns
Panel A: Financial Markets

Series	Period	r_{min}	r_{max}	$\|r\|$	Skewness	Kurtosis
Nasdaq	1971-2019	-12.04	13.25	0.81	-0.30	9.64
S&P 500	1980-2019	-22.90	10.96	0.74	-1.15	26.53
DJI	1985-2019	-25.63	10.51	0.73	-1.66	41.66
HSI	1986-2019	-40.54	17.25	1.07	-2.35	58.58
Nikkei	1984-2019	-16.14	13.23	0.99	-0.36	8.15

Panel B: Artificial Stock Market (ASM)

Network connectedness (N)	r_{min}	r_{max}	$\|r\|$	Skewness	Kurtosis
2	-13.53	12.86	0.73	-0.15	13.53
4	-15.20	13.72	0.74	-0.24	15.54
6	-13.68	13.44	0.75	-0.09	13.28
8	-14.63	13.16	0.77	-0.28	13.93
10	-13.82	12.37	0.78	-0.21	12.78
12	-15.32	13.85	0.78	-0.23	16.48
14	-14.80	12.18	0.77	-0.27	13.60
16	-15.02	12.82	0.78	-0.25	13.90

Based on the stylized facts of daily data in the financial markets presented in Table 3 and the calibrated approaches displayed in Figure 1-3, the ASM is calibrated to mimic those facts of financial markets. The ASM specifications with control parameters are then produced and displayed in Table 1.

Figure 1: Price and return series of Nasdaq

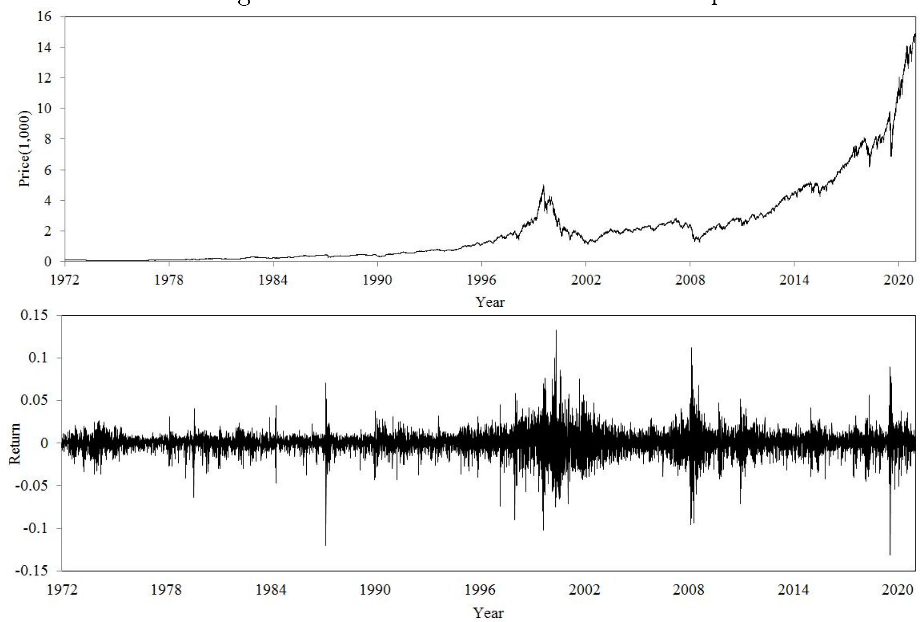


Figure 2: Price and return series of ASM

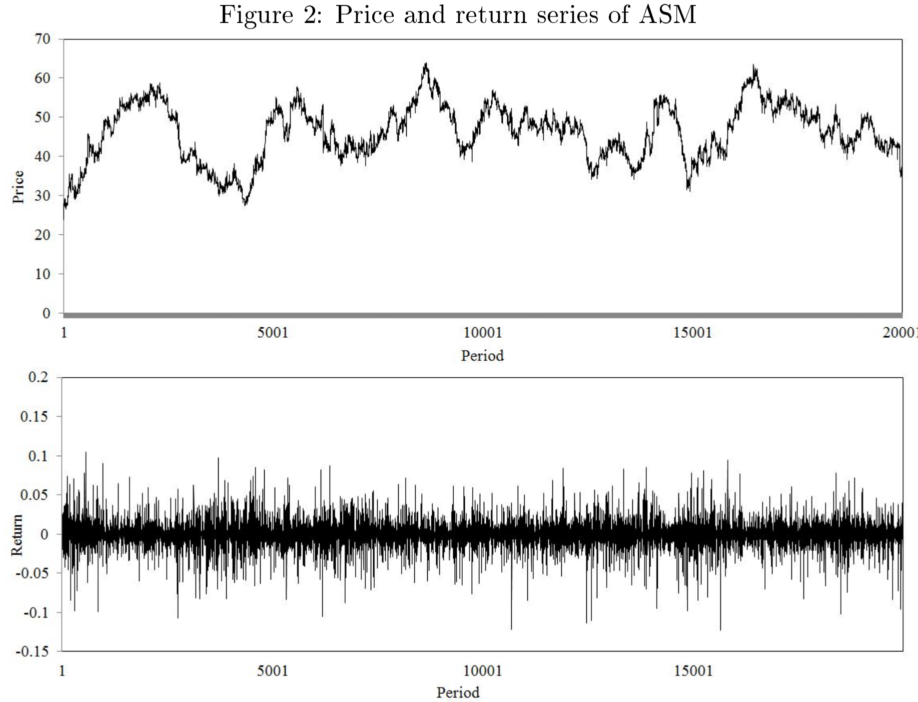
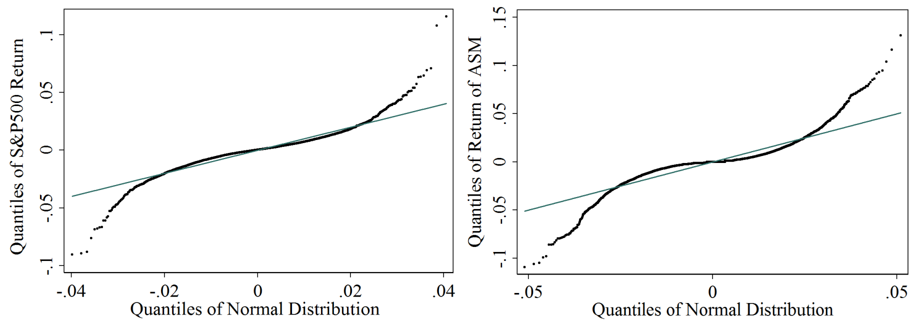


Figure 3: Quantile-quantile plots (q-q plots) of financial market and ASM



3 Results

3.1 Markets with exogenous information

We first analyze the influence of network connectedness on financial markets when the information is exogenous given, which means we do not consider the information acquired cost. Following Han and Yang (2013), we assume a consistent proportion of informed traders in each group in the system., which is 50%.

3.1.1 Volatility and price distortion

In order to examine the influence of social network, we first investigate how the market volatility change when group network connectedness (N) increase. We apply the measurement of market volatility (P_v) which is introduced by Westerhoff (2003),

$$P_v = \frac{100}{N_{T-1}} \Sigma_{\mathrm{t}=1}^{N_T} \left| \frac{P_t - P_{t-1}}{P_{t-1}} \right| \tag{6}$$

where P_t is the price per share of the stock, P_{t-1} is the price of the last period and N_T is the number of periods.

The left panel of Figure 4 shows the change of volatility as network connectedness increases for all 100 simulation runs. The two lines, solid and dashed, present the numerical values of mean and standard deviation among 20 runs with relative network connectedness, respectively. The figure also presents the results of all these 20 runs. It shows that network connectedness increases market volatility. It has the lowest value when the network connectedness is 2, while it reaches the highest point when the network connectedness is 10. It is clear that market volatility increases as network connectedness larger. As network connectedness increases, traders are better informed about the return of risky assets. They are able to estimate the value of risky assets more precisely. Thus, these traders' demands become more aggressive, making market volatility increased. Nevertheless, there is no monotonic relationship between the standard deviation of all these runs and network connectedness.

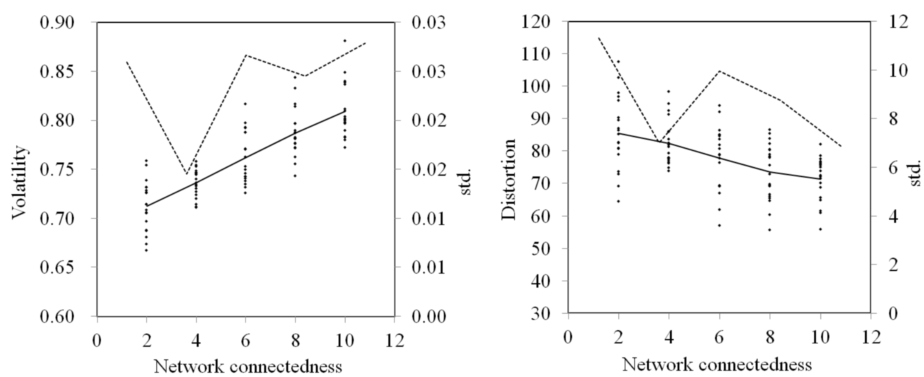
According to Westerhoff (2003), we use the following measurement to analyze the change of market price distortion (P_D),

$$P_D = \frac{100}{N_T}\Sigma_{\mathrm{t=1}}^{N_T}\left|\frac{P_t - P_f}{P_f}\right| \tag{7}$$

where P_f refers to the fundamental price according to $P_f = D_t/r_f$ Gordon (1962).

The right panel of Figure 4 shows that how network connectedness (N) influences price distortion (P_D) from the experiments by employing artificial intelligence. The two lines, solid and dashed, present the numerical values of mean and standard deviation among 20 runs with relative network connectedness, respectively. As we expected, price distortion decreases as network connectedness becomes larger. It is obvious that price distortion is the lowest, which reveals more information in the market when network connectedness reaches 10. This finding is consistent with Colla and Mele (2010), Ozsoylev and Walden (2011) and Han and Yang (2013). More information is impounded into the stock price by friends' communication and sharing, decreasing the price distortion and improving market efficiency. In addition, sharing information within a group may lead one trader to lose part of his own biased judgments on the stock, thereby improving market efficiency.

Figure 4: Volatility and distortion effects of network connectedness with exogenous information

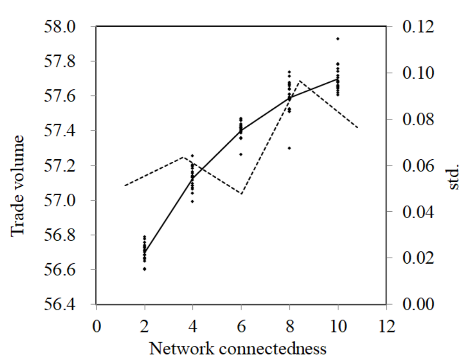


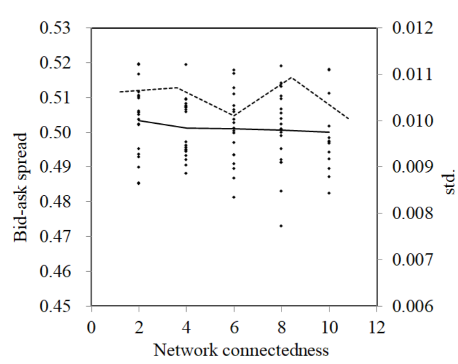
3.1.2 Trading volume and bid-ask spread

We also investigate the impact of network connectedness on trading volume with exogenous information. We use the average value of trading volume overall 20,000 periods as a measurement. The results are also straightforward and obvious in the left panel of Figure 5: network connectedness increases trading volume with exogenous information. The reason might be that more traders receive information through communication, and then they can estimate the risky asset more precisely. In addition, the same information linkages in a group may lead one trader to lose his own monopolistic information power because other group members receive the same signal. This may encourage each trader to increase their aggressiveness to anticipate his group members. As a result, they may be more aggressive in trading and thus, market trading volume increases. The results generate a convex line between trading volume and network connectedness, which is similar to the results provided by Han and Yang (2013) when information is exogenous in the market. At the beginning level, traders communicate with informed friends and communication enlarges their information sets, and thereby they trade aggressively. As network connectedness increases at a certain level, the group members with those demands are more and cannot be all satisfied in the market. Thus, the trading volume provides a convex line with network connectedness.

In the simplified continuous double auction process of the application of artificial intelligence, which is the ASM, we have the highest price for buying (the best bid B_b) and the lowest price for selling (the best ask B_a) on the board. We measure the bid-ask spread by the difference value of the best ask B_a and the best bid B_b($B_b - B_a$) for each period and then average all these 20,000 periods. From the right panel of Figure 5, we find that bid-ask spread slightly reduces as network connectedness (N) increases. Although the highest point (0.503 when N is 2) is only 0.003 higher than the lowest point (0.500 when N is 10) in the picture, it still provides a steady increasing trend generally. In the market with exogenous information by giving a fixed fraction of informed traders, social communication helps enlarge more traders' information set and corresponding impounding more information to the market, thereby helping lower bid-ask spread and make the market more liquid.

Figure 5: Trading volume and bid-ask spread effects of network connectedness with exogenous information





3.2 Markets with endogenous information

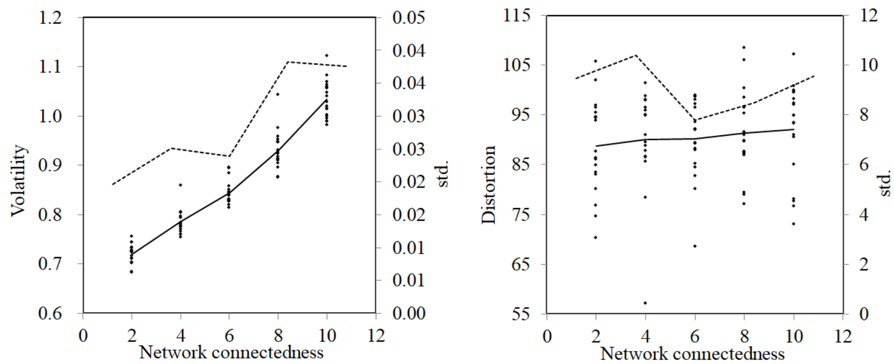
Han and Yang (2013) conclude that the proportion of informed traders reduces as network connectedness increases when considering the cost of information acquisition. They suggest that the reason might be two free-riding channels through social communication. The first one could be “free-riding on friends”. Social communication among friends will influence informed traders’ incentives spend a cost on private information. They will attempt to rely on the information via social communication, rather than spending a cost on that. The second one could be “free-riding on price”. Traders will prefer to communicate information and learn from market price, rather than acquire private information by themselves. They will directly generate the benefits of social communication on price informativeness. Thus, we follow their conclusion, which is decreasing the proportion of informed traders as network connectedness increasing in a group when information is endogenously in the market. In our system, the fractions of informed traders are 50%, 25%, 16.67%, 12.5% and 10% when the network connectedness is 2,4,6,8,and 10 respectively.

3.2.1 Volatility and price distortion

When private information is endogenously acquired at a cost, the market volatility also increases as network connectedness becomes larger. Nevertheless, volatility is a concave function with network connectedness in Figure 6. Comparing with the results by given exogenous information, the market provides higher volatility obviously. Especially when the network connectedness is 10, market volatility (P_v) increases by 0.22% (1.03% minus 0.81%). Han and Yang (2013) suggests that social communication in groups has a negative effect on information production when information is endogenous. This might be the reason that makes the market more volatile. The standard deviation of 20 runs also increases as network connectedness is raised, which presents in a dashed line in the left panel of Figure 6. This also reveals that social communication increases more market volatility when information is endogenous, which is acquired at a cost.

The right panel of Figure 6 shows that how social communication affects price distortion when information is endogenous. We demonstrate that price distortion averagely increases from 88.81% to 92.02% as network connectedness rises from 2 to 10, which is consistent with the conclusion drawn by Han and Yang (2013). The finding is different from the result with exogenous information. Han and Yang (2013) suggests that social communication deters the production of information and then reduces the market efficiency when information is endogenous. In other words, when the proportion of informed traders is endogenously determined, gross information content in the market reduces when increasing network connectedness (N). It then deteriorates market efficiency, which provides an opposite result of exogenous information.

Figure 6: Volatility and price distortion effects of network connectedness with endogenous information

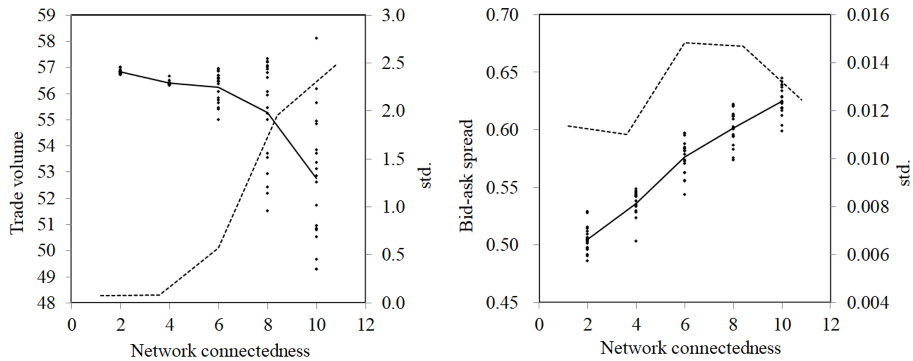


3.2.2 Trading volume and bid-ask spread

We also demonstrate a different conclusion of trading volume effects of social communication when information is endogenously acquired at a cost. The left panel of Figure 7 presents that trading volume decreases as there is more social communication in the market. Under endogenous information, the fraction of informed traders decreases as social communication increases, which will deteriorate the market's information production and further increase the overall volatility and risk in the market. Therefore, it results in those traders in groups trade less aggressively, thereby causing trading volume to be decreasing as network connectedness (N) rises. As is shown in the figure, trading volume especially decreases sharply from network connectedness (N) 6 to 10. It is also evident that the deterioration of information production by social communication significantly harms market liquid under endogenous information. The standard deviation of trading volume, presented by the dashed line in the figure, has an increasing trend as network connectedness (N) increases. It is also evident that the market is more volatile and risky when there is more social communication.

The second panel of Figure 7 demonstrates that bid-ask spread increases as there is more social communication under endogenous information. It suggests that reducing the fraction of informed traders harms information production and then deteriorates market liquidity. It is almost a straight solid line in the picture. Compared to the results observed under exogenous information, which is a slight and steady increasing trend, the results are exactly the opposite.

Figure 7: Trading volume and bid-ask spread effects of network connectedness with endogenous information



4 Sensitivity analysis

We perform the sensitivity analysis of Figures 8-11 to investigate how the findings are sensitive to the parameters of the simulated model. Following Yeh and Yang (2010), different pre-specified parameters are performed in this part. The first parameter is each trader's initial money (M_0). We perform the simulation with $M_0 = 300$, which is different from the $M_0 = 200$ in the calibrated model. The second parameter is regarding the learning behavior of each rational trader, which is the probability set for mutation and crossover used in GP. In the sensitivity analysis, we simulate 5 runs for each parameter set.

Figure 8 and Figure 10 show the market outcomes under the condition with 300 initial cash when information is exogenously given and when information is endogenously acquired at a cost, respectively. Figure 9 and Figure 11 present the market outcomes under (P_c, P_m) is (0.6,0.3) when information is exogenously given and when information is endogenously acquired at a cost respectively. We examine the effects of social network connectedness on the financial market, including market volatility, price distortion, trading volume and the bid-ask spread, when information is exogenous or endogenous. We summarize the results to examine whether or not the effects of social networks are influenced by those changed parameters. Overall, our results illustrate that the simulations under those chosen parameters do not present significantly different results to those observed from the original model. It is obvious that Figure 8-11 are quite similar to those obtained from Section 3.

Figure 8: Market outcomes with exogenous information under initial cash = 300

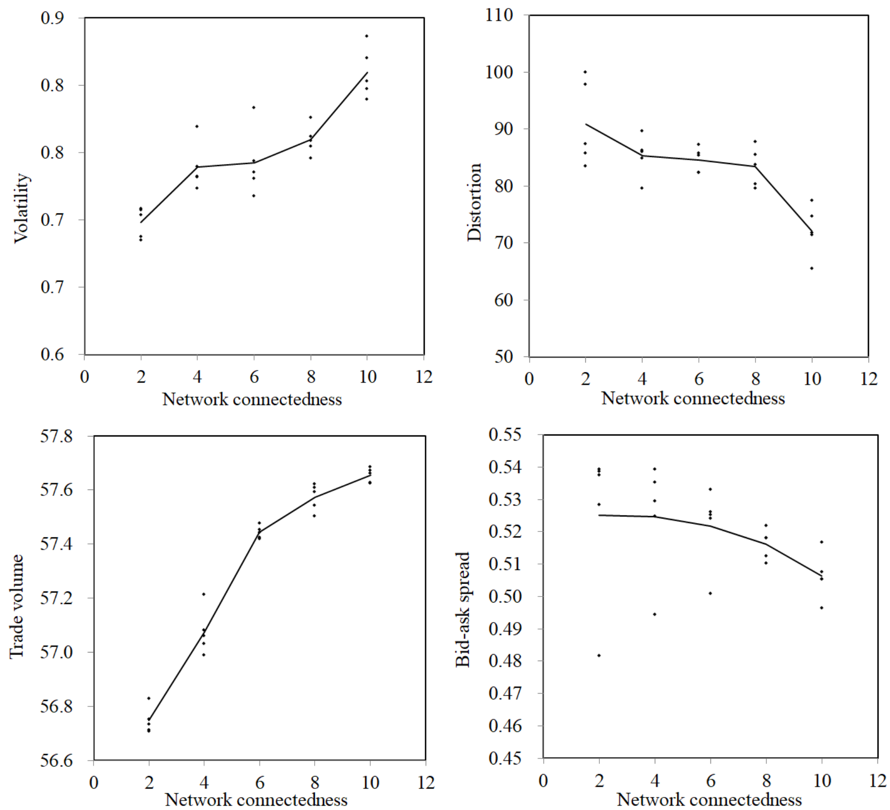


Figure 9: Market outcomes with exogenous information under $P_c = 0.6$,$P_m = 0.3$

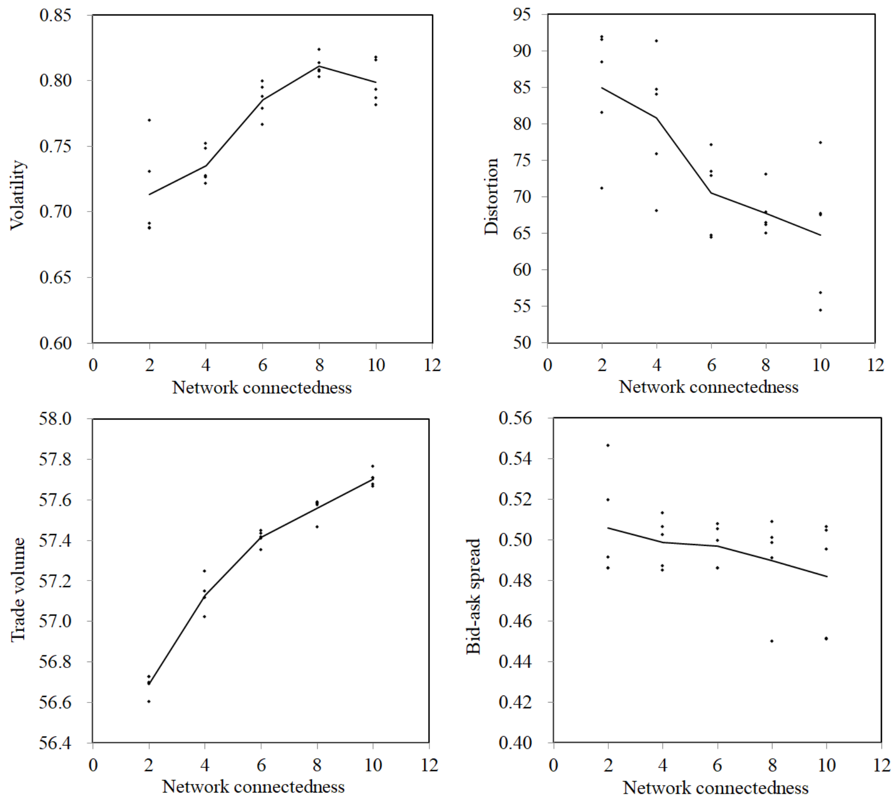


Figure 10: Market outcomes with endogenous information under initial cash = 300

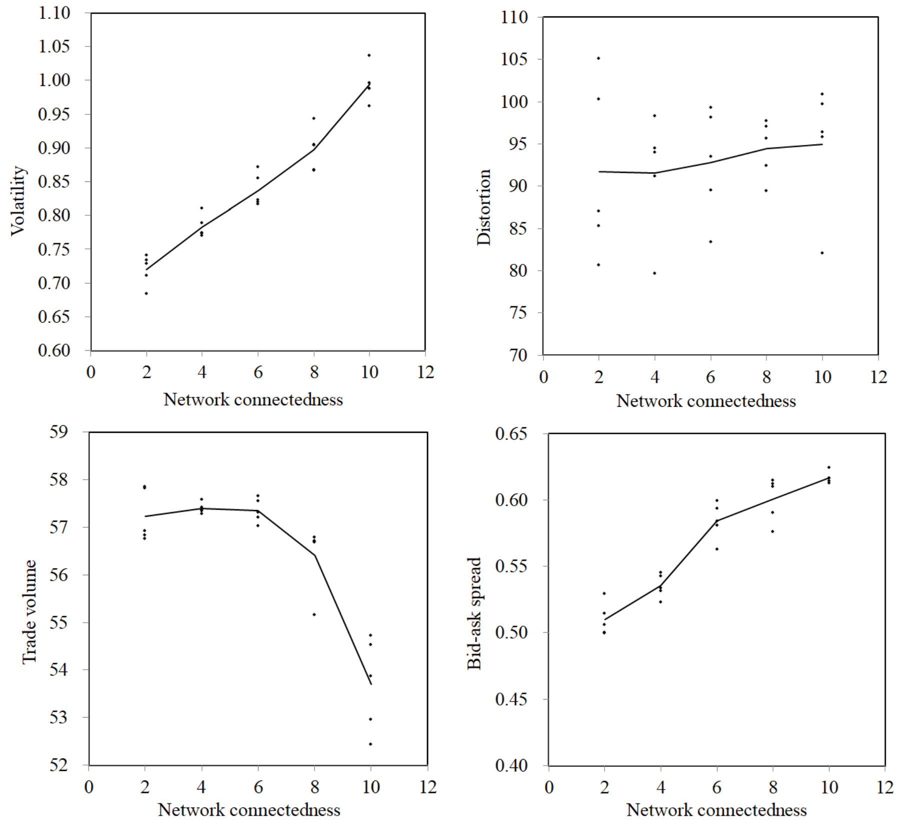
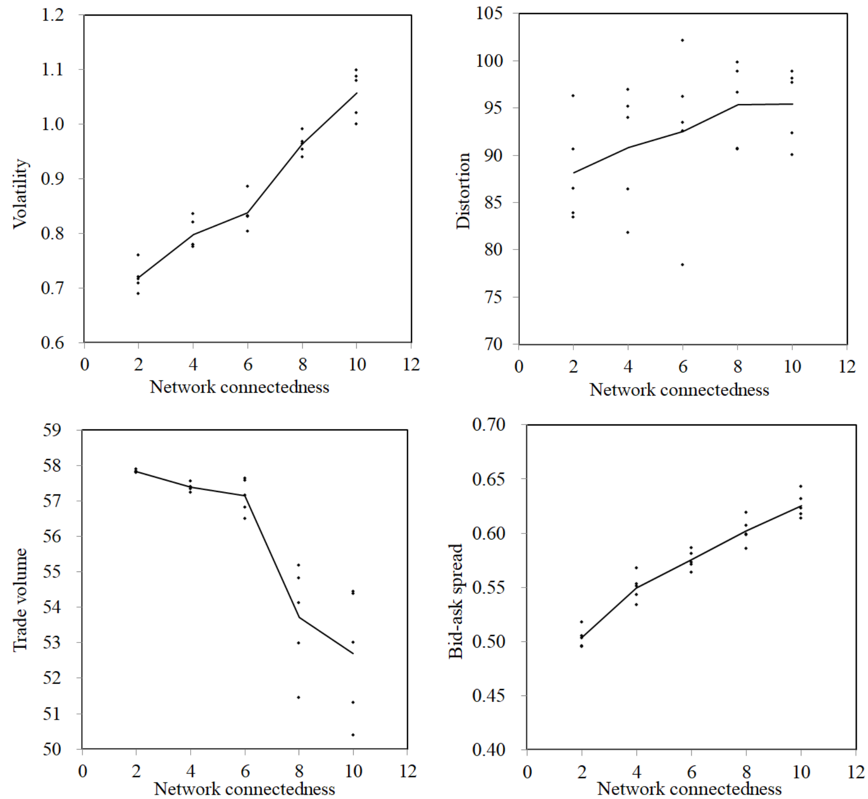


Figure 11: Market outcomes with endogenous information under $P_c = 0.6, P_m = 0.3$



5 Conclusion

This paper examines the effects of social networks on financial markets by constructing an agent-based limit order artificial stock market(ASM), an application of artificial intelligence based on the framework presented in Yeh and Yang (2010). We further divide market participators into four types: informed, social, uninformed and noise traders in the ASM. Different traders receive different kinds of information from the market and have different learning behaviors and trading strategies. GP is used to evolve rational traders' learning behaviors and their investment decisions then can be heterogeneous. Social and informed traders form communication networks in the experiments by using artificial intelligence.

Based on the proposal of Han and Yang (2013), we investigate how social networks influence financial markets under two conditions: information is exogenously given versus information is endogenously acquired at a cost. Our results, which are proven by an application of artificial intelligence, are consistent with those presented by Han and Yang (2013). We also find that the effects of social networks on the market are quite different when considering the change of information production. When the information is exogenous, which means the proportion of informed traders is fixed in each group, social networks increase market volatility and trading volume and decrease price distortion and bid-ask spread. When information acquires at a cost, which means information is endogenously and results in reducing the fraction of informed traders in each group, the influences of social networks on the financial market are reversed. Social communication harms market efficiency, decreases the trading volume and increases bid-ask spread under endogenous information because it harms information production. Thus, it indicates that information production plays an important role in financial markets because the results are different depends on whether it takes into account.

In future work, the model in this paper could be improved to calibrated with more sensible empirical data to make it more reliable and reasonable. The use of artificial intelligence is possible to combine with empirical data to predict financial facts. For example, a survey could be employed to collect data. In the survey, we first make sure whether the investor is a social investor or not and collect information and return of investments from him. At the same time, we could collect the returns of social investors in our ASM to compare and calibrate with the data from the survey. If there are no significant differences, we assert that the model is reliable and able to predict more financial facts.

Acknowledgments

The authors would like to thank Professor Victor Chang for the research support from VC Research (VCR 0000146).

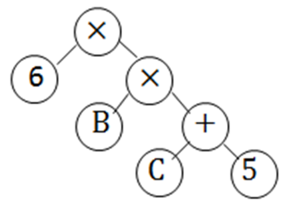
Appendix

Genetic Programming (GP) is a technique by running computer programs that are encoded as a set of genes that evolved using an evolutionary algorithm. As an important part of artificial intelligence, GP is determined by the structure of a parse tree, which contains functions and terminals. Generally, the terminal set is composed of those dependent variables, such as stock prices or dividends. The elements of the function set can be explained as functions used to combine the terminals with building up the functions. For example,

$$A = 6B(C + 5) \tag{8}$$

The GP tree structure is described as follows:

Figure 12: GP tree strcuture



The first function model is generated randomly according to the pre-specified terminal and function set. The performance of each function model is evaluated by the resulting fitness. GP generates new function models in three ways: immigration, crossover and mutation. They help form new functions in the following generation. Immigration is the process that gives an existed function, and thus immigration shows no innovation in creating forecasting models. The crossover procedure randomly selects one point from each of two GP trees named parents here. Then exchange the whole parts below that point of the parents to generate a new function model. The procedure shows that combining two kinds of knowledge into one idea. A mutation is used to randomly change a part of the sub-tree of the original function tree. It can be regarded as an innovation relying on current knowledge.

References

Ahern, K. R. (2017). Information networks: Evidence from illegal insider trading tips. *Journal of Financial Economics*, 125:26–47.

Arthur, W. B., Holland, J. H., Palmer, B. L. R. G., and Tayler, P. (1997). Asset pricing under endogenous expectations in an artificial stock market. *Social Science Electronic Publishing*, 23:1487–1516.

Banerjee, S. and Green, B. (2015). Signal or noise? uncertainty and learning about whether other traders are informed. *Journal of Financial Economics*, 117:398–423.

Brock, W. A. and H.Hommes, C. (1998). Heterogeneous beliefs and routes to chaos in a simple asset pricing model. *Journal of Economic Dynamics and Control*, 22:1235–1274.

Brown, J. R., IvkoviÄ, Z., Smith, P. A., and Weisbenner, S. (2008). Neighbors matter: Causal community effects and stock market participation. *Journal of Finance*, 63(3):1509–1531.

Bu, X., Huang, Z., and Zhao, Q. (2021). Mining analysis of customer perceived value of online customisation experience under social commerce. *Enterprise Information Systems*, pages 1–25.

Cecconi, F. and Campenni, M., editors (2019). *Information and Communication Technologies (ICT) in Economic Modeling*. Computational Social Sciences.

Cohen, L., Frazzini, A., and Malloy, C. (2008). The small world of investing: Board connections and mutual fund returns. *Journal of Political economy*, 116(5):951–979.

Colla, P. and Mele, A. (2010). Information linkages and correlated trading. *The Review of Financial Studies*, 23(1):203–246.

Collins, C., Dennehy, D., Conboy, K., and Mikalef, P. (2021). Atificial intelligence in information systems research: A systematic literature review and research agenda. *International Journal of Information Management*, 60:1–17.

Fischer, T. and Krauss, C. (2018). Deep learning with long short-term memory networks for financial market predictions. *European Journal of Operational Reseach*, 270(2):654–669.

Gordon, M. J. (1962). The investment, financing, and valuation of the corporation. *Homewood, IL: R. D. Irwin*.

Han, B. and Yang, L. (2013). Social networks information acquision and asset prices. *Management Science*, 59(6):1444–1457.

Hvide, H. K. and Ostberg, P. (2015). Social interaction at work. *Journal of Financial Economics*, 117(3):628–652.

Jackson, M. O. (2008). *Social and Economics Networks*. Princeton University.

Jaklic, J., Grubljesic, T., and Coelho, P. S. (2019). The shift to socio-organizational drivers of business intelligence and analytics acceptance. *Journal of Organizational and End User Computing (JOEUC)*, 31(2):37–64.

Kajol, K., Nath, M., Singh, R., Singh, R., and Das, A. K. (2020). Factors affecting seasonality in the stock market: a social network analysis approach. *International Journal of Accounting & Finance Review*, 5(4):39–59.

Langdon, W. B. and Poli, R. (2002). *Foundations of Genetic Programming*. SpringerVerlag.

Lebaron, B., Arthur, W. B., and Palmer, R. (1999). Time series properties of an artificial stock market. *Journal of Economic Dynamics and Control*, 23:1487–1516.

Liu, X., Wang, S., Yao, K., and Sun, R. (2021). Opportunistic behaviour in supply chain finance: a social media perspective on the 'noah event'. *Enterprise Information Systems*, pages 1–28.

Long, W., Lu, Z., and Cui, L. (2019). Deep learning-based feature engineering for stock price movement prediction. *Knowledge-Based Systems*, 164:163–173.

Nabipour, M., Nayyeri, P., Jabani, H., and Mosavi, A. (2020). Deep learning for stock market prediction. *Entropy*, 22(8):840.

Ozsoylev, H. N. and Walden, J. (2011). Asset pricing in large information networks. *Journal of Economic Theory*, 146(6):2252–2280.

Ram, J. and Zhang, C. (2021). Examining the role of social media analytics in providing competitive intelligence: The impacts and limitations. *Journal of Global Information Management (JGIM)*, 29(6):1–18.

Shiller, R. J. and Pound, J. (1989). Survey evidence on diffusion of interest and information among investors â. *Journal of Economic Behavior & Organization*, 12(1):47–66.

Shive, S. (2010). An epidemic model of investor behavior. *Journal of Financial & Quantitative Analysis*, 45(1):169–198.

Sohaib, O. (2021). Social networking services and social trust in social commerce: A pls-sem approach. *Journal of Global Information Management*, 29(2):23–44.

Westerhoff, F. (2003). Speculative markets and the effectiveness of price limits. *Journal of Economic Dynamics and Control*, 28(3):493–508.

Yeh, C.-H. and Yang, C.-Y. (2010). Examining the effectiveness of price limits in an artificial stock market. *Journal of Economic Dynamic and Control*, 34:2089–2108.

Zhong, X. and Enke, D. (2019). Predicting the daily return direction of the stock market using hybrid machine learning algorithms. *Financial Innovation*, 5(1):1–20.