A LETTERS JOURNAL EXPLORING THE FRONTIERS OF PHYSICS

LETTER

A minority game with expected returns for modeling stock correlations

To cite this article: M.-Y. Yang *et al* 2018 *EPL* **123** 18001

View the article online for updates and enhancements.

A LETTERS JOURNAL EXPLORING THE FRONTIERS OF PHYSICS

July 2018

EPL, **123** (2018) 18001
doi: 10.1209/0295-5075/123/18001

www.epljournal.org

A minority game with expected returns for modeling stock correlations

M.-Y. YANG[1], S.-P. LI[2], L.-X. ZHONG[3] and F. REN[1,4,5(a)]

[1] *School of Business, East China University of Science and Technology - Shanghai 200237, China*
[2] *Institute of Physics, Academia Sinica - Taipei 115 Taiwan*
[3] *School of Finance, Zhejiang University of Finance and Economics - Hangzhou 310018, China*
[4] *School of Science, East China University of Science and Technology - Shanghai 200237, China*
[5] *Research Center for Econophysics, East China University of Science and Technology - Shanghai 200237, China*

received 25 March 2018; accepted in final form 13 July 2018
published online 3 August 2018

PACS 89.65.Gh – Economics; econophysics, financial markets, business and management
PACS 07.05.Tp – Computer modeling and simulation
PACS 05.45.Tp – Time series analysis

Abstract – Financial systems are complex systems which have been widely studied in recent years. We here propose a model to study stock correlations in financial markets, in which an agent's expected return for one stock is influenced by the historical return of the other stock. Each agent makes a decision based on his expected return with reference to information dissemination and the historical return of the stock. We find that the returns of the stocks are positively (negatively) correlated when agents' expected returns for one stock are positively (negatively) correlated with the historical return of the other. We provide both numerical and analytical studies and give explanations to stock correlations for cases with agents having either homogeneous or heterogeneous expected returns. The result still holds when other factors such as holding decisions and external events are included which broadens the practicability of the model.

Introduction. – Complex systems are common in geology, biology as well as social and economic areas. There has been interest to understand the complex cooperative behavior of complex systems as a result of simple rules of interactions between microscopic constituents among researchers. The correlation in a complex system and how the information flows within the system is an interesting and important topic of current research. Understanding the underlying mechanism and the cooperative behavior of one complex system would be helpful in the understanding of other complex systems. Among them, financial systems are complex systems which have been widely studied in recent years. Understanding stock correlation is academically and practically useful. Stock correlations are high during financial crises, which is a phenomenon in stock markets worldwide [1,2]. Only making clear the cause of this phenomenon, can one better understand the mechanism underlying the stock correlation, and apply it more accurately in asset pricing, investor decision-making, and financial risk regulations.

Later on, scholars have explained stock correlations from the macroscopic perspective, for instance the market [3,4], industry [5,6], and firm [7,8] level. Morck *et al.* [3] and Dang *et al.* [4] found that stock correlations tended to be higher in poor and emerging markets with poor institutions and property rights. Kallberg and Pasquariello [5], Antón and Polk [6] showed that firms in the same industry had correlated earnings and therefore returns. Qiao *et al.* [7] and Zhang *et al.* [8] have shown that a firm's specific information is the primary cause of stock correlations.

Recently, scholars have also studied stock correlations from the perspective of investors' actions [9–11]. Some empirical studies show that investors' sentiment and irrational actions will cause stock correlations [9]. Other studies focus on information dissemination through investors [10,11]. To date, only few studies are devoted to explain the microcosmic reason of stock correlations from the perspective of agents' actions, most of which are based on theoretical analysis. Our study is motivated by research on stock correlations with agent-based models and reveals its microcosmic mechanism by modeling the trading behavior of individual agents.

(a) E-mail: fren@ecust.edu.cn

Minority game is a highly successful agent-based model that can be used to explain most stylized facts in financial markets [12]. The standard minority game was first proposed by Challet and Zhang based on the El Farol bar problem [13]. It describes a system in which heterogeneous agents adaptively compete for scarce resource, and it captures some key features of a generic market mechanism and the basic interactions between agents and public information. As a successful agent-based model to simulate the stock market, minority game has been under rapid development and largely used in memory size [14,15], evolution mechanism [16], strategy selection [17,18], stock market simulations [19], market impact [20] and agent behavior [21]. One of the most important applications of the minority game is its modeling for multi-assets [22,23]. Inspired by these works, we here introduce a new model based on minority game in which agents may trade two stocks, and focus on its explanation to stock correlations. This extension makes the model closer to the real stock markets, and can be widely applied to asset pricing, investor decision-making and financial risk regulations.

We propose that an agent's expected return on one stock is influenced by its own and other stock's historical return. This expected return is modeled with reference to studies on the theory of information dissemination and the facts that exist in real stock markets. Some studies on information dissemination theory imply that stock price movements are influenced by the spread of information among firms [24,25]. In addition, some other studies demonstrated that information from different firms could be obtained and spread by investors, which, in turn, could affect the price movements of stocks [26,27]. In a real stock market, we can see that investors trade stocks by referring to the performance of other stocks in the same industry, sometimes even the same industry chain and the competitive industry. These investors' behaviors will ultimately have an impact on the stock correlations. The information dissemination theory proposed in the literature and the above phenomena exist in real stock markets which provide the theoretical and realistic basis for the establishment of the expected return in our model.

We also propose that an agent makes a decision based on his/her expected return and historical returns of the stock in the minority game. The simulation results of our model suggest that the stock returns are positively (negatively) correlated when the expected returns are positively (negatively) correlated, and the correlation of stock returns is proportional to the correlation of the expected returns of the stock. Compared to previous studies, this paper makes two contributions. First, we propose a new model based on the MG model to explain the stock correlations from the perspective of agent-based modeling. Second, we theoretically demonstrate that the investors' expectation of one stock influenced by another stock essentially leads to the correlation between the two by both numerical simulation and analytical analysis.

Table 1: A strategy of agent i for stock j. Here $r_j(t-m)$ denotes the historical return of the stock j at time $t-m$, and $r^e_{j,i}(t)$ denotes the expected return of agent i for stock j at time t. $+$ $(-)$ means the return is positive (negative). $\sigma^i_j(t)$ denotes the decision that agent i makes for stock j according to the history and expectation. 1 (-1) means buying (selling) 1 unit of stock.

History			Expectation	Decision
$r_j(t-m)$	...	$r_j(t-1)$	$r^e_{j,i}(t)$	$\sigma^i_j(t)$
$+$	...	$+$	$+$	1
$-$	...	$-$	$-$	-1
...	...	...	...	...
$+$	...	$+$	$-$	1
$+$	...	$-$	$-$	-1

Model and assumptions. – To study the correlation between individual stocks, we assume the simple case when there are only two stocks traded by agents. Our model takes the form of a repeated game with an odd number N of agents who must choose an independent decision of buying or selling actions according to their strategies. A strategy is a way of decision-making in which an investor makes a decision according to the historic returns in the last m time steps and the expected return for a specific stock. Table 1 illustrates an agent's strategy for the transaction of stock j in our model. The history of a stock is the stock return series in the last m steps, and the expectation is an agent's expected return. The decision is denoted by 1 and -1, respectively in the table, representing buying or selling 1 unit of stock. Since there are 2^m possible bit strings of historical returns, two possible options (up or down) for each agent's expectation, and two decisions of buying or selling actions, hence there are a total of $2^{2^{m+1}}$ strategies. We also consider the case when the agent can take a holding position, making a decision about not buying or selling [19]. We find that the introduction of the holding position does not change the main results, hence this feature will not be included in our model.

The expected return is defined as follows:

$$r^e_{j,i}(t) = a_j r_j(t-1) + b_{j,i} r_{j'}(t-1), \qquad (1)$$

where $r^e_{j,i}(t)$ $(j = 1, 2;\ i = 1, \ldots, N)$ is the expected return of agent i for stock j at time t $(t = 0, 1, \ldots, T)$, and $r_j(t-1)$ is the return of stock j at time $t-1$, $j' = 1$ (2) when $j = 2$ (1). a_j is the first-order autocorrelation coefficient of the stock j and $b_{j,i}$ denotes the impact of the first-order lag return of the other stock on stock j for agent i.

We propose the expected return with reference to the studies on the information dissemination theory. Zhang *et al.* [8] and Chuang [28] both demonstrated that the dissemination of information between firms caused stock price co-movement. Mondria [11] found that changes in one asset affected both asset prices when investors could choose linear combinations of asset payoffs to update information about the assets. Based on these studies, we propose that the expected return of agents to be in the form of a linear combination of stock returns in our model.

We can also see that investors trade stocks by considering the performance of other stocks in real stock markets, and this provides a realistic basis for the assumption of the above expectation. The underlying reason for the investors' behaviors that focus on the performances of multi-assets may lie in the fundamental correlations between these stocks, which have business relationships between each other. For example, investors will refer to the performance of other stocks in the same industry, sometimes even the same industry chain and the competitive industry when they trade stocks.

To begin with, each agent i randomly picks S strategies from the full strategy space, sticks with them throughout the game, and keeps track of the cumulative performance of his/her trading strategy s ($s = 1, \ldots, S$) by assigning a score $U_j^{i,s}(t)$ to each of them. The initial scores of the strategies are set to be zero. Initially, agent i randomly selects a strategy among his/her S own strategies, and the trading decision is also randomly selected since there are no historical and expected returns at $t = 0$. At time step $t > 0$, each agent i adopts a strategy s with the highest score. If $U_j^{i,s}(t)$ is the highest at time t, the decision of agent i to trade stock j is $\sigma_j^i(t) = \sigma_j^{i,s}(t)$ corresponding to the current historical and expected returns of stock j.

After all agents have made their decisions of actions, the excess demand of stock j at time t is calculated as

$$A_j(t) = \sum_{i=1}^{N} \sigma_j^i(t). \tag{2}$$

$A_j(t)$ measures the imbalance between buyers (demand) and sellers (supply), which is commonly used to update the price of the stock [16,19]. The price of stock j at time t is updated according to

$$P_j(t) = P_j(t-1) + \mathrm{sgn}[A_j(t)]|A_j(t)|^{0.5}, \tag{3}$$

where the square root in the formula is commonly used in the price dynamics of the MG model [19], which is also supported by the evidence of empirical studies [29]. The return of stock j at time t is

$$r_j(t) = \log(P_j(t)) - \log(P_j(t-1)). \tag{4}$$

The score of strategy s held by the agent i is then updated as below:

$$U_j^{i,s}(t) = U_j^{i,s}(t-1) + g_j^{i,s}(t), \tag{5}$$

where the payoff $g_j^{i,s}(t)$ to strategy s is

$$g_j^{i,s}(t) = -\mathrm{sgn}[A_j(t)]\sigma_j^{i,s}(t). \tag{6}$$

The payoff is calculated with the excess demand, which is in line with the price update.

The agent then repeatedly chooses the best strategy from his/her fixed S strategies according to their updated scores, and makes the trading decision according to the newly updated historical returns and expected return. The model evolves by repeating the steps listed above.

Simulation results and analysis. – For simplicity and the computational efficiency of the model, we choose the case where $N = 1001$, m and S are relatively small. Though the correlation becomes weaker as m increases, the pattern of the correlation still remains for $m \leq 4$. Besides, considering the finite memory size of the investors in a real stock market, we let m take a small value, and set $m = 1$ as an illustrative example. The choice of S strategies, selected from the whole pool of $2^{2^{m+1}}$ possible strategies, can also affect the results of the correlations. For large S, the agents have many repeated strategies, resulting in the same decisions of buying or selling actions, which will cause problems in the correlation between the stocks. An extreme case is to use the entire strategy pool. In fact, the investors generally have only a small number of strategies to use in a real market. Therefore, S should not be large, and we set $S = 2$ in our model. The simulation result becomes stable when $T \sim 1000$. Therefore, we choose T to be 1000. The initial price of our model is set to be 2000, large enough to ensure that the price is positive throughout the entire evolution period. We set the values of a_j according to the first-order autoregressive coefficients of the 15 A-share stocks trade on the Shanghai and Shenzhen Stock Exchanges from January 4, 2011 to December 31, 2015, which are found to be larger than 0 and smaller than 1. Therefore, we take the value of a_j to be between 0 and 1, under which the first-order autoregressive coefficients of the simulated data are consistent with those of the empirical data. We consider four combinations of different values of a_j, namely $a_1 = \{0.1, 1\}$ and $a_2 = \{0.1, 1\}$ in our simulation, and the values of $b_{j,i}$ in two cases: homogeneous agents with the same expectation and heterogeneous agents with different expectations.

To discuss the correlation between stocks, we calculate the Pearson's correlation coefficient, which is defined as

$$\rho_r = \frac{\mathrm{Cov}(r_1, r_2)}{\sqrt{\mathrm{Var}(r_1)}\sqrt{\mathrm{Var}(r_2)}}, \tag{7}$$

where $\mathrm{Cov}(r_1, r_2)$ is the covariance of the two stock return series, and $\mathrm{Var}(r_1)$ and $\mathrm{Var}(r_2)$ are the variances of the two stock return series, respectively. We perform 50 runs and take the mean correlation coefficient as the final simulation result.

Simulation results and analysis for homogeneous agents with the same expectation. When all agents have the same expected return, we have $b_{j,i} = b_j$. We here assume that the values of b_j are between -1 and 1, and the interval is 0.1. The simulation results are shown in fig. 1.

From the simulation result in fig. 1(a), we can see that the correlation coefficients between the stock returns are

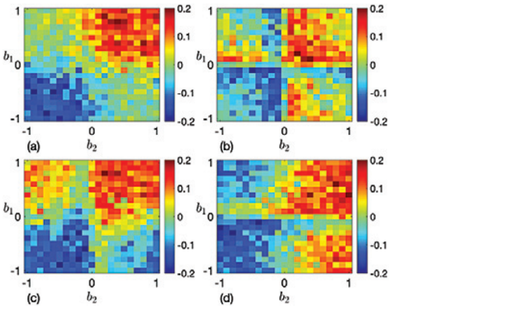

Fig. 1: (Color online) The correlation coefficients between the stock returns with parameters $m = 1$ and $S = 2$ for (a) $a_1 = a_2 = 1$, (b) $a_1 = a_2 = 0.1$, (c) $a_1 = 1, a_2 = 0.1$, (d) $a_1 = 0.1, a_2 = 1$.

positive (negative) when $b_1 > 0$ and $b_2 > 0$ ($b_1 < 0$ and $b_2 < 0$), and the absolute values of the correlation coefficients are large when b_1 and b_2 are near 1 or -1. In fig. 1(b), the correlation coefficients between the stock returns are small in most cases, and the absolute values of the correlation coefficients are large when b_1 and b_2 are near 0.1 or -0.1. For $a_1 = 1$ and $a_2 = 0.1$ in fig. 1(c), the expected return of stock 2 is almost only affected by stock 1 and the expected return of stock 1 is affected by both stocks, hence the correlation coefficients between two stocks' returns are positive (negative) when $b_1 > 0$ ($b_1 < 0$). In addition, the expected return of stock 2 is almost unaffected by either stock when b_2 is near 0, hence the correlation coefficients between the stock returns are very small. The results in fig. 1(d) are similar to the results in fig. 1(c). Therefore, we will take the simulation results in fig. 1(a) as a representative to analyze.

It is worth studying how much investors' expectations of stock returns account for the stock return correlation when other factors are included in their trading decisions. For instance the exogenous factor [30], which considers a case when the market is impacted by external events. The internal and external contributions to the overall excess demand are introduced as

$$A_j(t) = A_j^{int}(t) + A_j^{ext}(t), \tag{8}$$

where $A_j^{int}(t)$ is the excess demand defined in eq. (2), and $A_j^{ext}(t)$ represents the contribution of external events, which has a general formula

$$A_j^{ext}(t) = (-1)^{\theta(t)} \tilde{A}_j E_j(t). \tag{9}$$

To make $A_j^{ext}(t)$ more realistic, we use the empirical data of the news of A-share stocks trade on Shanghai Stock Exchange during the period from December 16, 2013 to November 22, 2016 to calculate the average probability of the occurrence of an external event per minute, and use it to determine the probability $p(E_j(t) = 1) = 0.0082$. The basic results remain to be the same if $p(E_j(t) = 1)$ is not too large. $\theta(t)$ is a randomly generated integer, which determines the sign of the impact. $\tilde{A}_j$ reflects the impact

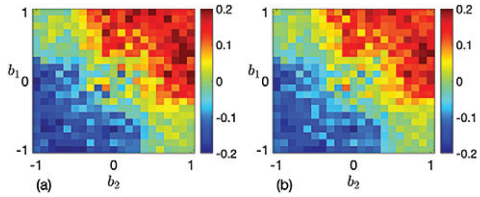

Fig. 2: (Color online) The correlation coefficients between the stock returns for the case in which external events are included into investors' trading decisions with parameters $m = 1$, $S = 2$ and $a_1 = a_2 = 1$ for (a) $k = 1$ and (b) $k = 4$.

strength of the external event, which is measured in units of the standard deviation s_j of $A_j^{int}(t)$, *i.e.*, $\tilde{A}_j = ks_j$. The simulation results for the case when $k = 1$ and 4 are shown in fig. 2. As the impact strength of external events increases, the correlation between the stocks becomes relatively smaller. However, the main results are essentially similar to the ones shown in fig. 1(a). The stock returns are still correlated under the interference of exogenous factor, and this shows that the stock return correlations are largely determined by the mechanism of the expected returns proposed in our model.

From the simulation results, regularities in stock return correlations can be seen. However, what causes the regularities, and why the correlation coefficient is very close to zero in some cases and greater (less) than zero in other cases? To explain these phenomena, it is necessary to analyze the source of the stock correlation. From the assumptions of the model, we know that the correlation between the stock returns is mainly determined by the agent's expected returns of the stocks. Since the expected return of the agent on one stock takes into account the influence of the other stock, the correlation between the expected returns is bound to affect the correlation between the stock returns. To verify this, we should analyze the relationship between the stock return and its expected return.

First, we draw the scatter plots of both stocks when $m = 1$, $S = 2$ and $a_1 = a_2 = 1$ in fig. 3. It can be seen that generally there is a linear relationship between the return and its expected return, though the scatter points cluster in three regions, which may be related to the crowd effects. In the minority game, agents show crowd effects when $2^m \ll N$, and the synchronization of their actions induce periodicity in the return distribution [31]. We thus observe that the scatter points cluster in fig. 3. We further perform a linear regression analysis between the return and its expected return, and find the p-value for the t-test, a significance test of the regression coefficient, is very close to 0 while the R-square is about 0.75. Hence, there is a significant positive linear correlation between the return and its expected return. We therefore analyze the correlation between the two stocks' expected returns in order to explain the correlation between two stocks' returns.

Since the correlation between the stock returns is mainly determined by the values of b_j, we concentrate on how the

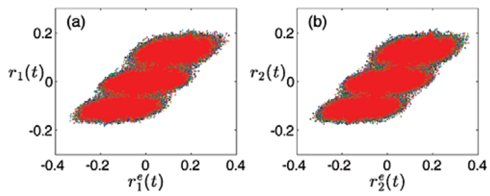

Fig. 3: (Color online) Scatter plots of the return *vs.* its expected return for (a) stock 1 and (b) stock 2 when $m = 1$, $S = 2$ and $a_1 = a_2 = 1$. Colors correspond to the returns of different runs.

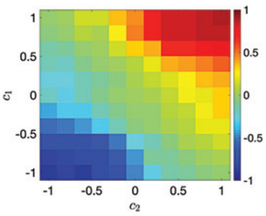

Fig. 4: (Color online) The correlation coefficients between the stock returns when $m = 1$, $S = 2$, $a_1 = a_2 = 1$ and $\delta_1 = \delta_2 = 1$.

values of b_j affect the correlation between the expected returns of the stocks. The correlation between the stocks' expected returns will be explored using the variances of their expected returns. From eq. (1), when $a_1 = a_2 = 1$, the variance of the expected return for stock j is

$$\Delta r_j^e(t) = \Delta r_j(t-1) + b_j \Delta r_{j'}(t-1), \tag{10}$$

where the variance of stock j is defined as $\Delta r_j(t-1) = r_j(t-1) - r_j(t-2)$. Since the agents are homogeneous, the variance of the expected return for different agents is the same. We further study the positive (negative) correlation between the expected returns by calculating the probability of having the same (opposite) sign of their variances under different combinations of the parameters b_1 and b_2. See appendix for details of the analysis.

From the results of fig. 3 and the regression analysis, we have concluded that there is a significant positive linear correlation between the returns and expected returns. We then summarize the results of simulated returns based on the correlation of expected returns in the following. The correlation between the stock returns is positive (negative) when $b_1 > 0$ and $b_2 > 0$ ($b_1 < 0$ and $b_2 < 0$), and gets stronger as b_j ($j = 1, 2$) increases (decreases). The correlation is weak when b_1 and b_2 have opposite signs, and changes from negative to positive as b_j ($j = 1, 2$) increases. We can also observe that the results in fig. 1(a) and the analytic results in the above thus agree with each other.

Simulation results and analysis for heterogeneous agents with different expectations. In a real stock market, agents are heterogeneous and their expected returns are not exactly the same. Indeed, we can consider the heterogeneous agents that have different values for both a_j and $b_{j,i}$. The parameter $b_{j,i}$ can reflect the information dissemination through stocks, which plays a core role in explaining the correlation between stocks. Besides, we have also studied the case with a uniform distribution of $a_j \sim U(0, 1)$, and found that the results remain similarly. Hence, we mainly focus on the parameter $b_{j,i}$ in the following discussions. We will make an additional assumption that each agent i in the model matches a unique $b_{j,i}$, which is subject to a uniform distribution, *i.e.*, $b_{j,i} \sim U(c_j - \delta_j, c_j + \delta_j)$. The center of the distribution is c_j, and the distribution range is $2\delta_j$.

It is worth noting that the results do not depend on the specific formula of the distribution, but only the symmetry feature of the distribution significantly affects the results. As one will see in the following discussions, the stock correlation is positive (negative) when the distribution center is bias to positive (negative), which may be driven by the forces of news or big events in real stock markets.

We first consider the case of a uniform distribution with varying distribution center and fixed distribution range. The simulation result for $a_1 = a_2 = 1$ and $\delta_1 = \delta_2 = 1$ is shown in fig. 4. Centers c_1 and c_2 are from -1 to 1, and the intervals are 0.2. Similar to the analysis for homogeneous agents with the same expectation, we draw the scatter plots of the returns *vs.* expected returns of the two stocks for four combinations of c_1 and c_2 in fig. 5. We also perform a regression analysis for the stocks, where the p-value is very close to 0 and the R-square is within 0.73–0.88. Hence, there is a significant positive linear correlation between the return and its expected return when b_1 and b_2 are uniformly distributed.

From the linear relationship between the return and its expected return, one can interpret the correlation between the returns based on the correlation between the expected returns. For heterogeneous agents who have different expected returns, the variances of their expected returns should be averaged over different agents

$$\overline{\Delta r_j^e}(t) = \frac{1}{N}\sum_{i=1}^{N} a_j \Delta r_j(t-1) + \frac{1}{N}\sum_{i=1}^{N} b_{j,i} \Delta r_{j'}(t-1). \tag{11}$$

Since $b_{j,i}$ are subject to the uniform distribution, we have $\frac{1}{N}\sum_{i=1}^{N} b_{j,i} = c_j$. When $a_j = 1$, the average variances in eq. (11) become

$$\overline{\Delta r_j^e}(t) = \Delta r_j(t-1) + c_j \Delta r_{j'}(t-1). \tag{12}$$

We now study the correlation of the expected returns from eq. (12), which takes the same form as eq. (10), with c_j replacing b_j. Similar to the analysis for fixed values of b_j, we can obtain the correlation between the returns, which is summarized below. The correlation between the stock returns is positive (negative) when $c_1 > 0$ and $c_2 > 0$ ($c_1 < 0$ and $c_2 < 0$), and becomes stronger as c_j ($j = 1, 2$) increases. The correlation is weak when c_1 and c_2 have opposite signs, and the correlation changes from negative to positive as c_j increases, in agreement with fig. 4.

We then consider the case of a uniform distribution with fixed distribution center and varying distribution range.

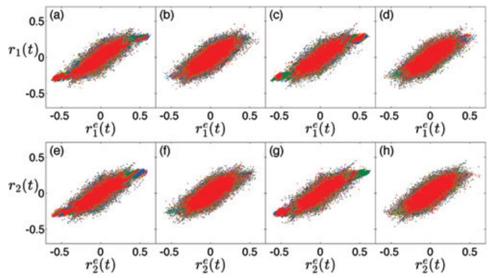

Fig. 5: (Color online) Scatter plots of the return *vs.* its expected return for (a) stock 1 and (e) stock 2 with $0 < c_1 < 1$ and $0 < c_2 < 1$, (b) stock 1 and (f) stock 2 with $-1 < c_1 < 0$ and $-1 < c_2 < 0$, (c) stock 1 and (g) stock 2 with $0 < c_1 < 1$ and $-1 < c_2 < 0$, (d) stock 1 and (h) stock 2 with $-1 < c_1 < 0$ and $0 < c_2 < 1$. Colors correspond to the returns of different runs, with parameters $m = 1$, $S = 2$ and $a_1 = a_2 = 1$.

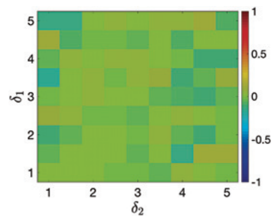

Fig. 6: (Color online) The correlation coefficients between the stock returns for $m = 1$, $S = 2$, $a_1 = a_2 = 1$ and $c_1 = c_2 = 0$.

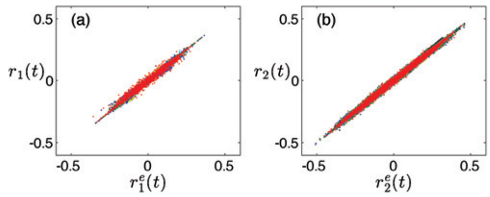

Fig. 7: (Color online) Scatter plots of the return *vs.* its expected return for (a) stock 1 and (b) stock 2 when $m = 1$, $S = 2$ and $a_1 = a_2 = 1$. Colors correspond to the returns of different runs.

The simulation results for $a_1 = a_2 = 1$ and $c_1 = c_2 = 0$ are shown in fig. 6. The ranges of δ_1 and δ_2 are from 1 to 5, and the intervals are 0.5. From fig. 6, we can see that the correlation coefficients are small, and the correlation between the stock returns is weak.

Scatter plots of the returns *vs.* the expected returns of the stocks when $c_1 = c_2 = 0$ are shown in fig. 7, where a linear relationship between the return and its expected return can be clearly observed. We therefore carry out a regression analysis of the stocks separately, where the p-value is close to 0 and the R-square is about 0.99. Hence, there is a significant positive linear correlation between the return and its expected return, based upon which one can interpret the correlation between the returns by analyzing the correlation between the expected returns.

From eq. (11), when $a_j = 1$ and $c_j = 0$, we have $\overline{\Delta r_j^e}(t) = \Delta r_j(t-1)$. In this case, we can easily see that the mean value of the expected return is only relevant to the stock return itself. Hence, the expected returns of the stocks are uncorrelated. From fig. 7 and the regression analysis results, we see that the return and its expected return are strongly correlated, implying that stock returns are not correlated. This is confirmed by the simulation results in fig. 6.

Conclusion. – In this paper, we study stock correlations by using a model based on the minority game, in which we propose that an agent makes a decision with the historical return of the stock and his/her expected return, which is influenced by the historical return of the other stock. We model the stock correlation for homogeneous agents with the same expectation, and find that the investor's expectation regarding the stock return is an important factor for the stock correlation. We also find that stock returns are positively (negatively) correlated when the expectations of returns are positively (negatively) correlated, and the correlation of stock returns is proportional to the correlation of stock expected returns. We then model the stock correlation for heterogeneous agents with different expectations, whose parameters $b_{j,i}$ obey a uniform distribution. The simulation results suggest that the center of the distribution has a significant influence on the stock correlation but the range of the distribution has no influence on the stock correlation. These results remain to be true when we consider other factors such as external events in financial markets.

To the best of our knowledge, this is the first time that one models the correlations between stocks from the perspective of agent-based modeling. Our model is derived from the standard MG model, in which the agents are allowed to trade multi-assets simultaneously. The expected returns of the agents are modeled with reference to information dissemination in financial markets. To improve the practicability of the model, we further introduce variables that model external news and events in financial markets and fix their values from the analysis of real data. These features make our model simulate real stock markets more closely, and help to expand its practical implications in many issues dealing with multi-assets or systematic problems. The model not only can reveal the microscopic mechanism underlying the stock correlations, but also can be applied to asset pricing, investor decision-making and financial risk regulations.

* * *

This work was partially supported by the National Natural Science Foundation (Nos. 10905023, 71131007 and 71371165), the Humanities and Social Sciences Fund sponsored by Ministry of Education of the People's Republic of China (No. 17YJAZH067), the Collegial Laboratory Project of Zhejiang Province (No. YB201628), and the Fundamental Research Funds for the Central Universities (2015).

Appendix

Analysis for the correlation between the stocks' expected returns under four combinations of the parameters b_1 and b_2: I) $0 < b_1 < 1$ and $0 < b_2 < 1$, II) $-1 < b_1 < 0$ and $-1 < b_2 < 0$, III) $0 < b_1 < 1$ and $-1 < b_2 < 0$, IV) $-1 < b_1 < 0$ and $0 < b_2 < 1$.

I) $0 < b_1 < 1$ and $0 < b_2 < 1$.

a) If $\Delta r_1(t-1) > 0$ and $\Delta r_2(t-1) > 0$, then

$$\Delta r_j^e(t) = \Delta r_j(t-1) + b_j \Delta r_{j'}(t-1) > 0. \qquad \text{(A.1)}$$

Here, the expected returns of both stocks increase at the same time, and the correlation between the expected returns is positive.

b) If $\Delta r_1(t-1) < 0$ and $\Delta r_2(t-1) < 0$, the analysis is similar to I) a). Here, the correlation between the expected returns is positive.

c) If $\Delta r_1(t-1) > 0$ and $\Delta r_2(t-1) < 0$, we discuss the possibilities for four cases of combinations of $\Delta r_1^e(t)$ and $\Delta r_2^e(t)$.

According to eq. (10), the condition satisfying $\Delta r_1^e(t) < 0$ and $\Delta r_2^e(t) > 0$ is

$$-\frac{\Delta r_2(t-1)}{b_2} < \Delta r_1(t-1) < -b_1 \Delta r_2(t-1). \qquad \text{(A.2)}$$

Since $-\Delta r_2(t-1)/b_2 > -b_1 \Delta r_2(t-1)$ when $\Delta r_2(t-1) < 0$, eq. (A.2) cannot be satisfied, therefore the case when $\Delta r_1^e(t) < 0$ and $\Delta r_2^e(t) > 0$ is impossible.

For the case when $\Delta r_1^e(t) > 0$ and $\Delta r_2^e(t) < 0$ are satisfied, according to eq. (10), the condition will be

$$-b_1 \Delta r_2(t-1) < \Delta r_1(t-1) < -\frac{\Delta r_2(t-1)}{b_2}. \qquad \text{(A.3)}$$

The value range of $\Delta r_1(t-1)$ becomes smaller when b_1 and b_2 approach 1. Therefore, the negative correlation between the expected returns becomes weaker.

If $\Delta r_1^e(t) > 0$ and $\Delta r_2^e(t) > 0$ are satisfied, according to eq. (10), the condition will be $\Delta r_1(t-1) > -\Delta r_2(t-1)/b_2$. The value range of $\Delta r_1(t-1)$ becomes larger when b_2 approaches 1. Therefore, the positive correlation between the expected returns becomes stronger.

Similarly, if $\Delta r_1^e(t) < 0$ and $\Delta r_2^e(t) < 0$ are satisfied, the positive correlation of the expected returns is stronger when b_1 approaches 1.

Among the four cases above, three of them are valid, which include two cases when the expected returns are positively correlated with large probability. Therefore, the correlation is positive on average, and becomes stronger when b_j approaches 1.

d) If $\Delta r_1(t-1) < 0$ and $\Delta r_2(t-1) > 0$, the analysis is similar to I) c). Here, the positive correlation between the expected returns becomes stronger when b_j approaches 1.

The analysis of II), III) and IV) are similar to I), and we will not go into details here.

References

[1] Aste T., Shaw W. and Di Matteo T., *New J. Phys.*, **12** (2010) 085009.
[2] Liu X. F. and Tse C. K., *Int. J. Bifurcat. Chaos*, **22** (2012) 1250142.
[3] Morck R., Yeung B. and Wu W., *J. Financ. Econ.*, **58** (2000) 215.
[4] Dang T. L., Moshirian F. and Zhang B., *J. Financ. Econ.*, **116** (2015) 82.
[5] Kallberg J. and Pasquariello J., *J. Empir. Finance*, **15** (2008) 481.
[6] Antón M. and Polk C., *J. Finance*, **69** (2014) 1099.
[7] Qiao H.-S., Xia Y. and Li Y., *PLoS ONE*, **11** (2016) e0156784.
[8] Zhang W.-P., Li C.-P., Ye Y.-M., Li W.-J. and Ngai E. W. T., *IEEE Intell. Syst.*, **30** (2015) 26.
[9] Eun C. S., Wang L. L. and Xiao S. C., *J. Financ. Econ.*, **115** (2015) 283.
[10] Filzen J. J. and Schutte M. M., *N. Am. J. Econ. Finance*, **39** (2017) 19.
[11] Mondria J., *J. Econ. Theory*, **145** (2010) 1837.
[12] Challet D., Marsili M. and Zhang Y.-C., *Physica A*, **294** (2001) 514.
[13] Challet D. and Zhang Y.-C., *Physica A*, **246** (1997) 407.
[14] Johnson N. F., Hui P.-M., Zheng D.-F. and Hart M., *J. Phys. A: Math. Theor.*, **32** (1999) L427.
[15] Lee C.-Y., *Phys. Rev. E*, **64** (2001) 015102.
[16] Challet D., Marsili M. and Zhang Y.-C., *Quant. Finance*, **1** (2001) 168.
[17] Ren F., Zheng B., Qiu T. and Trimper S., *Phys. Rev. E*, **74** (2006) 041111.
[18] Zhong L. X., Zheng D. F., Zheng B. and Hui P. M., *Phys. Rev. E*, **72** (2005) 026134.
[19] Yeung C. H., Wong K. Y. M. and Zhang Y.-C., *Phys. Rev. E*, **77** (2008) 26107.
[20] Barato A. C., Mastomatteo I., Bardoscia M. and Marsili M., *Quant. Finance*, **13** (2013) 1343.
[21] Zhang J. Q., Huang Z. G., Wu Z. X., Su R. and Lai Y. C., *Sci. Rep.*, **6** (2016) 20925.
[22] Bianconi G., De Martino A., Ferreira F. F. and Marsili M., *Quant. Finance*, **8** (2008) 225.
[23] De Martino A., Castillo I. P. and Sherrington D., *J. Stat. Mech.*, **2007** (2007) P01006.
[24] Bordino I., Battiston S., Caldarelli G., Cristelli M., Ukkonen A. and Weber I., *PLoS ONE*, **7** (2012) e40014.
[25] Heiberger R. H., *PLoS ONE*, **10** (2015) e0135311.
[26] Chen J.-J., Tan L. and Zheng B., *Sci. Rep.*, **5** (2015) e8399.
[27] Marcet F., *Emerg. Mark. Rev.*, **32** (2017) 1.
[28] Chuang H.-W., *N. Am. J. Econ. Finance*, **36** (2016) 172.
[29] Zhang Y.-C., *Physica A*, **269** (1999) 30.
[30] Papadopoulos P. and Coolen A. C. C., *J. Phys. A: Math. Theor.*, **41** (2008) 365002.
[31] Challet D., Marsili M. and Zhang Y.-C., *Minority Games: Interacting Agents in Financial Markets* (Oxford University Press, Oxford) 2005.