Physica A 390 (2011) 1991–2001

Contents lists available at ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

Identifying the structure of group correlation in the Korean financial market

Sanghyun Ahn [a], Jaewon Choi [a], Gyuchang Lim [a], Kil Young Cha [a], Sooyong Kim [a], Kyungsik Kim [b,*]

[a] Department of Physics, Korea Advanced Institute of Science and Technology, Daejeon 305-701, Republic of Korea

[b] Department of Physics, Pukyong National University, Pusan 608-737, Republic of Korea

ARTICLE INFO

Article history:
Received 12 February 2010
Received in revised form 17 December 2010
Available online 31 January 2011

Keywords:
Random matrix theory
Cross-correlation
Group correlation matrix
Inverse participation ratio
Korean stock market
Ensemble

ABSTRACT

We investigate the structure of the cross-correlation in the Korean stock market. We analyze daily cross-correlations between price fluctuations of 586 different Korean stock entities for the 6-year time period from 2003 to 2008. The main purpose is to investigate the structure of group correlation and its stability by undressing the market-wide effect using the Markowitz multi-factor model and the network-based approach. We find the explicit list of significant firms in the few largest eigenvectors from the undressed correlation matrix. We also observe that each contributor is involved in the same business sectors. The structure of group correlation can not remain constant during each 1-year time period with different starting points, whereas only two largest eigenvectors are stable for 6 years 8–9 eigenvectors remain stable for half-year. The structure of group correlation in the Korean financial market is disturbed during a sufficiently short time period even though the group correlation exists as an ensemble for the 6-year time period in the evolution of the system. We verify the structure of group correlation by applying a network-based approach. In addition, we examine relations between market capitalization and businesses. The Korean stock market shows a different behavior compared to mature markets, implying that the KOSPI is a target for short-positioned investors.

1. Introduction

The correlation-based approach used in the study of financial markets is highly relevant among the various methodologies in the econophysics field [1]. Cross-correlation between the stock entities of financial markets provides not only a physical interpretation of the market but also empirical or real information about the correlation structure between market entities [2–14]. Basically, financial market fluctuations result from the correlated decision making between buy and sell orders of various stock entities participating in the market. Because the conceptions of the market are reflected in price fluctuations, we can obtain the structure of a given financial market by investigating financial cross-correlation data. Random matrix theory has provided highly meaningful information about financial market structures, especially that of the U.S. market [15–22]. As reported in previous works, the seemingly complicated structure of stock market cross-correlation can be divided into three categories: a bulk random part, a market-wide part and the group correlation between firms in the same business sector [17,22]. This indicates that investors engaged in the stock market manage their stock portfolios to be optimized in relation to the structure of a particular business cluster. Clearly decomposed group correlation is one of the key features of a mature stock market. Does this hold true for the Korean stock market? Korea has a distinctive economic

* Corresponding author. Tel.: +82 51 629 5562; fax: +82 51 629 5549.
E-mail address: kskim@pknu.ac.kr (K. Kim).

0378-4371/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.physa.2010.12.027

structure and background compared with other typical mature markets [23]. Although several papers studying the cross-correlation of the Korean stock market have been published recently, no work has investigated the explicit structure of group correlation [24,25]. Genuine information about group correlation was obtained by undressing market-wide effects using the Markowitz multifactor models. In the first part of this paper, we found the structure of business clusters in the Korean stock market by applying multifactor models. We verified the business structures by a network approach. Time stability of the market structure is another major issue in this research. We investigated the time stability of the obtained group correlation by separating the time series into several parts. We ultimately tested whether the structure of group correlation in the Korean stock market was maintained consistently during a sufficiently long time period or whether it showed a different pattern in the evolution of the system.

2. Method: random matrix theory

We first obtained a correlation coefficient between a pair of stocks. In order to calculate the correlation coefficient between stock entities from the logarithmic returns of a financial time series, we defined the logarithmic return, $R_i(t)$, as follows,

$$R_i(t, \Delta t) = \log P_i(t + \Delta t) - \log P_i(t) \tag{1}$$

where $P_i(t)$ denotes the daily price of stock entity i at time t and Δt represents one day. In order to remove the effect of the magnitude of prices on the correlation coefficient, we used the normalized returns, $r_i(t)$, as follows:

$$r_i(t) = \frac{R_i(t) - \langle R_i \rangle}{\sigma_i}, \qquad \sigma_i = \sqrt{\langle R_i^2 \rangle - \langle R_i \rangle^2} \tag{2}$$

where σ_j is the standard deviation of the return and $\langle \ldots \rangle$ indicates a time average during L steps. The correlation coefficient can be derived by following relation

$$C \equiv \langle r_i(t) r_j(t) \rangle = \begin{pmatrix} 1 & \ldots & c_{1N} \\ \vdots & \vdots & \vdots \\ c_{N1} & \ldots & 1 \end{pmatrix}. \tag{3}$$

The elements of the correlation matrix C_{ij} have a range of $-1 \le C_{ij} \le 1$. When the stocks i and j are perfectly-correlated, the value of C_{ij} is 1. On the other hand, $C_{ij} = -1$ if two stocks are perfectly anti-correlated. $C_{ij} = 0$ corresponds to the case where both stocks are uncorrelated. As mentioned above, the cross-correlation between the stock entities provides us with significant and meaningful information about the intrinsic structure of the stock market. The empirical time series of financial data always contains many effects, and therefore different levels of information exist in the market in the form of different levels of the cross-correlation coefficient. Previous work published in recent years [17] has suggested that the financial cross-correlation matrix C is separated into three groups: (a) a random part; (b) a market-wide effect part; and (c) the group correlation between entities included in the same industry, i.e. the so-called business sector. Therefore, for the latter, apart from the largest eigenvalue, genuine information about the business sector among stocks was reconstructed by several next-largest eigenvalues according to

$$C = \sum_{i=1}^{N_r} \lambda_i |\lambda_i\rangle\langle\lambda_i| + \sum_{j=1}^{N_g} \lambda_j |\lambda_j\rangle\langle\lambda_j| + \lambda_{largest} |\lambda_{largest}\rangle\langle\lambda_{largest}| \tag{4}$$

where N_r and N_g represent the number of eigenvalues attributed to the random part and the correlations based on the business sector, respectively.

3. Financial data

3.1. Data analyzed

We analyzed the financial time-series data of Korean stock market securities from the Korean Composite Stock Price Index (KOSPI). We extracted the 586 largest stocks from the entire KOSPI that were exchanged on the Korean stock market during the period 2003–2008, forming $L = 1500$ records of daily returns for $N = 586$ Korean stocks. The mean value of the return of the Korean stock market is $\langle R_{ij}(t + \Delta t) - R_{ij}(t) \rangle = 5.149$ and that of normalized log return is $\langle R_{ij} \rangle = 0.000283$. This result indicated that the average stock price increased in the time period from 2003 to 2008.

3.2. Brief statistics of correlation coefficient

We analyzed the distribution $P(C_{ij})$ of the elements $\{C_{ij}; i \neq j\}$ in the cross-correlation matrix C. We first examined $P(C_{ij})$ for daily returns from KOSPI data for the 6-year period. The $P(C_{ij})$ of the Korean stock market has an asymmetric coefficient distribution and each coefficient was positively correlated with the others; in other words, $\langle C_{ij} \rangle > 0$. We calculated the mean value of the empirical correlation coefficient $\langle C_{ij} \rangle = 0.12$. In the case of random time-series, the shape of the distribution $P(R_{ij})$ was similar to the Gaussian distribution. $P(C_{ij})$ corresponded well to the Gaussian distribution