

Full Length Research Paper

A newly application of the random matrix theory on the exploration of the efficiency of investment portfolios in the Taiwan Stock Market

Chen, Ying-Chi^{1*} and Goo, Yeong-Jia²

¹IBM Taiwan Corporation Consultant 2F, No. 70, Lane 190, Min-Chiuan E Road, Taipei 11490, Taiwan, R.O.C.

²Department of Business Administration, National Taipei University, Taiwan.

Accepted 2 March, 2011

In the past, the investment portfolio theory regarding stock price deviation was considered to be individual risk that occurred in the market at random. However, the outcome efficiency of portfolios could be influenced by other factors that were found to be interrelated. In this study, we attempted to filter out those random parts of correlation matrix by the walk forward approach (WFA), and applied the Random matrix theory (RMT) that was developed from nuclear physics. We constructed a portfolio and then estimate the accumulation returns and Sharpe ratio. Based on the daily trading records from the Taiwan Stock Exchange (TWSE), the component indexes were grouped into 19 categories from Jan. 2, 2007 to Jan. 29, 2010, and there were a total of 767 data set entries as input information to this study. We employed variance tests, ANOVA, and least squares mean (LSM) to test the results. This study finds that, in general, the random part of stock return increases the portfolio risk, and consequently decreases the efficiency of the portfolio. In addition, this study adapted the concept of principal component analysis to analyze the information eigenvectors, which provided the information sequence for detecting the presence of unusual information that might affect the portfolio. As a result, the new approach could be helpful in forecasting the direction of price fluctuation.

Key words: Stock index, Random matrix theory, investment portfolio theory, principal component analysis, interrelated factor.

INTRODUCTION

Bachelier (1900) first introduced the random concept into the economic system, arguing that the reason for price changes of stocks and bonds was intangible random variation. Samuelson (1965) and Mandelbrot (1966) used mathematical methods to prove that stock price change is one kind of random motion, and they regarded the physical particle of Brownian movement as a rational economic entity, and then they produced the Random walk theory (RWT). Fama (1965) assumed that the economic entity is rational and proved the stock price was

stated as the model of random walk. He thought that the market would respond to all messages, and therefore he proposed the Efficient Market Hypothesis (Fama, 1970). However, the scholars of behavioral finance did not agree with him —Shiller (1981) and De Bondt and Thaler (1985) were opposed to Fama and thought the stock price existed because of over-reaction phenomena. Lakonishok et al. (1994) argued that both excessively optimistic and pessimistic investors cause stock price fluctuation, and even over-reaction phenomena. Cutler et al. (1991), and Bernard and Thomas (1990) argued that short-term return showed a positive autoregressive phenomenon, indicating a delay reaction due to insufficient information in the short run, which then led to over-reaction in the long term. Hong

*Corresponding author. E-mail: robinchen0925@yahoo.com.tw.

et al. (2000) believe that the different communication speeds of good and bad information cause delayed reaction to short-term information. All of these researches have proven that the change of stock price was not like RWT, and they deny the existence of efficient markets.

Fama (1998) argued that the presences of both price and insufficient information resulted in dual reaction, so stock price might occur randomly. Meanwhile, the estimation of abnormal return is very sensitive to statistical method variation, and so it could not be used to deny the existence of an efficient market by abnormal return researches (Fama, 1998). Shefrin (2000), who pointed out that the phenomenon of insufficient response occurs in the short term, and over-response in the long term, had opposed Fama (1998) and thought that both are the random view. Abeysekera (2001) and Sharma (2009), based on an empirical study of the stock market, denied the existence of the efficient market hypothesis. However, Milionis and Moschos (2000), Malkiel (2005), as well as Ozdemir (2008), pointed out firmly after empirical research that the weak form efficient market exists. Samuelson (1965) assumed that investors are similar to the physical particle of random motion. However, the physical particle is not independent, whereas humans are independent animals. Therefore, when messages appear in the market, some people will possibly make the rational response and stroll randomly in the market. Some people certainly also have an incomplete rational reaction also because of the speculative motive, and because of excessive optimism or pessimism. Therefore, we might need to do a rational analysis and avoid the two extremes to explain the market.

Because the efficient market assumes that the investor is a rational doer, it may be regarded as a real information carrier, and its responded price is the basic price. The incomplete rational doer's decision making will disturb the response of the market and cause the performance of the short-term market to deviate from the basic. If this decision is wrong, the performance of the long-term market will tend to adjust to the basic price. This argument was confirmed by many researchers (Shiller, 1981; De Bondt and Thaler, 1985; Jagadeesh and Titman, 1993; Lakonishok, Shleifer, and Vishny, 1994; Rouwenhorst, 1998).

Markowitz (1952), based on the relative degree of stock price returns, proposed the investment portfolio theory. In the estimation of the relevance of stock price of an economic system, both the change of real value of individual stock prices and the price of deviation were considered at the same time. The portion of stock price deviation was usually considered as individual risk and regarded as random. If high correlations among stock returns were observed and were unable to disperse, it would affect the efficiency of the investment portfolio. The key issue is decided by the "correlation matrix" of stock returns. If we could find a way to separate the random part

and keep the substantive, relevant portion of basic price, then we could possibly gain better investment portfolio achievements. The RMT is a technology developed by nuclear physics that observes the highly excited atomic spectrum (Wigner, 1951) and it has proven to be a good analytical tool in analyzing the U.S. stock market. Laloux et al. (1999) investigated the S&P 500 by using RMT and discovered that the random matrix was truly a disturbance of the investment portfolio. Therefore, it is arguable that we can use the correlation coefficients to analyze a portfolio directly. Plerou et al. (2002) studied 1000 American corporation's stocks from 1994 to 1995 and found that most eigenvalues are within the boundary of RMT, with only a few maximal eigenvalues deviating out of boundary. It means that using RMT to describe the real world between the correlation of the investment portfolio or stock price returns is more suitable. The random matrix in RMT which does not possess persistence, is similar in character to the randomness (Sharifi et al., 2004). In addition, the RMT might extract part of the basic price and reveals latent important information in the market. This would allow us to catch the message of timing in terms of price rise and drop.

The contributions of this study are as follows. (1) We adopt principal component analysis to analyze the information eigenvectors, trying to capture the turning points from bull/bear market transfer to bear/bull market. (2) We empirically test the portfolio efficiency of the Taiwan stock market based on the RMT methodology.

Therefore, the objectives of this study are: (1) to examine whether the filtration randomness could enhance investment portfolio efficiency, and (2) to decompose the portion of information derived from the RMT and use the concept of principal component analysis to match up trading volume for market exploration. We could expect to find out the switch time message of bull and bear markets. The paper is organized as follows. The first is the introduction. Second we explain the RMT and its connection with investment portfolio theory, and then explain the research design of this study. The fourth part presents the research results, and the final part is the conclusion of the study.

RANDOM MATRIX THEORY AND THE INVESTMENT PORTFOLIO

Nuclear physicists in the 1950s observed the interaction of atoms and found that very complex interactions of atoms could not be described and forecasted effectively, and they so could be regarded as a random acts. Wigner (1951) analyzed the eigenvalue of the random matrix, and found that it is suitable to describe the interactions of atoms using randomness. Suppose, there are K Gaussian (Gauss) sequences, the length of each sequence is N , the mean value is 0, and the variance as σ . If we demonstrate

K random sequences as a matrix, the correlation matrix is as below:

$$C = \frac{1}{N} GG', \quad (1)$$

where C; random sequence correlation matrix of $K \times K$; G: random sequence matrix of $K \times K$; $'$: transpose symbol.

Let $q \equiv N/K$, ($q > 1$), be a constant. If $K \rightarrow \infty$, $N \rightarrow \infty$, the limit distribution of the eigenvalue of correlation matrix C will have the following type (Marcenko and Pastur, 1967; Stein, 1969):

$$p_r(\lambda) = \begin{cases} \frac{q}{2\pi} \frac{\sqrt{(\lambda_{\max} - \lambda)(\lambda - \lambda_{\min})}}{\lambda}, & \lambda_{\min} \leq \lambda \leq \lambda_{\max} \\ 0, & \text{elsewhere} \end{cases} \quad (2)$$

$$\lambda_{\max} = \left(1 + \frac{1}{\sqrt{q}}\right)^2 \quad (3)$$

$$\lambda_{\min} = \left(1 - \frac{1}{\sqrt{q}}\right)^2, \quad (4)$$

where $p_r(\lambda)$: the probability density function of random matrix eigenvalue; λ : the eigenvalue of random matrix; λ_{\max} : the maximal eigenvalue of random matrix; λ_{\min} : the minimal eigenvalue of random matrix; λ_{\max} and λ_{\min} were described within random matrix theory as the maximum and minimum eigenvalues that will apply to define the boundary of a random correlation matrix. Through using the RMT to infer the boundary conditions, we can easily extract the non-random part in the correlation matrix.

The random matrix in the stock market cannot be directly observed—we could only calculate the correlation matrix that includes the random matrix. However, Bouchaud and Potters (2000) proposed a method by which we do not need to calculate the random matrix but instead remove it from the correlation matrix. The filtering steps are as follows:

1. Calculate the correlation matrix of stock returns and find out the eigenvalue: $\det(c - \lambda I) = 0$; \det is a determinant, c is the correlation coefficient matrix, I is a matrix. Find out K eigenvalues (λ) by the simultaneous equations.

2. To sort the eigenvalue from large to small order, and then apply the formulas (3) and (4) random matrix calculations of upper and lower bounds.

3. Calculate the average of eigenvalue within the boundaries of the random matrix: $\sum \lambda / s, s \leq K$.

4. Replace the original eigenvalue that is within the boundaries of the random matrix by an average.

5. Use both the original eigenvector and the new eigenvalue matrix to reconstruct the correlation matrix:

$$C_{new} = \vec{V} \times \Lambda \times \vec{V}'.$$

Therefore, it can remove the random matrix from the correlation matrix. Although Bouchaud and Potters's (2000) filter method (2000) has poor persistence, only some of the larger eigenvalues possess persistence nature (Sharifi et al., 2004), in other words, the part that is larger than the upper bound of random matrix possess persistence, vice versa. However, in the theory of factor analysis, the explanation ability would be worse if the eigenvalue gets smaller (Kaiser, 1960). Although the lower bound, which is smaller than the random matrix, is the basic price, its explanation ability is very weak. Therefore, this study suggests not to consider its influence ability. Stock returns variation has the characteristic of mean return (French and Roll, 1986), which indicates that the price deviation will disappear as time goes by. Therefore, the randomness that is generated by the high correlation cannot be persistent within the market. With regard to persistence study, Sharifi et al. (2004) also pointed out that the random matrix possesses poor persistence, which corresponds to the features that interfere with price, and this is an important reason that random matrix is suitable for this study. We should define some terms of RMT to meet the theme of this study. The study has defined the random matrix as a correlation matrix that was generated by randomness, also known as a random matrix. In accordance with the filtering step and reconstructing the correlation matrix that was defined as the correlation matrix after interference filter, it is known as the information matrix. The original matrix—that is, the original correlation matrix—is defined as the original matrix.

RESEARCH DESIGN

This study is based on the daily trading records from the TWSE, and the component indexes were grouped into 19 categories which included cement, food, plastics, textiles, electrical machinery, electrical cables, chemical industry, glass, ceramics, paper, steel, rubber, automotive, electronics, construction, transport, tourism, finance total merchandise trade, and other types from Jan. 2, 2007 to Jan. 29, 2010. There were totally 767 data set entries as input information to this study and the study used programs developed by

Matlab (2009a). First of all, we transfer the stock index into stock index return. As this research is following the efficient market hypothesis, we use the continuous rate of return, calculated as follows:

$$r_t = \ln\left(\frac{p_t}{p_{t-1}}\right) \quad (5)$$

where r_t : the index return of stocks of t stage; p_t : the index of stocks of t stage; p_{t-1} : the index of stocks of $t-1$ stage; $\ln(\bullet)$: logarithmic function.

After getting the stock index return, we take 30 days as the formation phase and 15 days as the forecast phase for WFA to process the filtering of the random matrix and calculation of the portfolio performance. In each phase of WFA, we calculated the correlation coefficient matrix (original matrix) of the first 30 days for 19 stocks, then filtered out the impact of random matrices and built an information matrix using the Bouchaud and Potters (2000) steps. Then we used the information concept of minimum variance portfolio weighting that was proposed by Markowitz (1952), to put the original matrix and the information matrix into the concept. We then calculated the optimal weights of two portfolios derived from the original matrix, known as the original portfolio. The portfolio that was calculated by the information matrix was called the information portfolio. Then, according to the optimal weight of the two portfolios and assigned to 19 kinds of stock, we calculated cumulative return and the Sharpe ratio for the next 15 days.

After finishing the process of the WFA phase, we obtained 722 data sets of cumulative returns and Sharpe ratios, followed by statistical analysis of these data sets. First, we calculated the variance of the information portfolio and the original portfolio by cumulative return rate, and we tested the variance of the two portfolios by F-test to confirm which risk is smaller. Then, we used the ANOVA approach to test the Sharpe ratio for the two portfolios to find the differences, and we processed a paired test by the least square mean, comparing the two portfolios to find out which performed better (the portfolio performance test process is shown in Figure1). We chose Sharpe ratio stands for the portfolio performance index for the following reasons. (1) The Sharpe ratio uses risk as a unit standard, with a higher Sharpe ratio standing for higher expected return for a portfolio under the same measure of risk (Sharpe, 1966, 1975), to be suitable for applying to expression of the portfolio performance with a corresponding relationship between risk and reward. (2) The Sharpe ratio was calculated in accordance with the efficient frontier of the portfolio theory of Markowitz (Sharpe, 1994), and the minimum variance portfolio that was developed by Markowitz was also used for this study. So the Sharpe ratio is very suitable to be the performance indicator of the original and information portfolios.

From Sharifi et al.'s (2004) research it can be observed that several larger eigenvalues possess a persistence nature and powerful influence, and we argued that these eigenvalues were information eigenvalues, with the eigenvalue of their corresponding eigenvector being called the information eigenvector. This study also took the WFA of the 30-day formation phase, calculated the information eigenvector, obtained 737 sets of information eigenvector and used the concept of principal component analysis to discuss the relationship between stocks and elements of information eigenvector. By the concept of principal component analysis, the greater the eigenvalue, the larger is the interpretation ability of overall change. Similarly, the greater the coefficient of the principal component, the greater the influence of such stock on the

principal component. Therefore, we used the analysis of information eigenvector to discover the message of stock variation.

This study was analyzed using eigenvector, by which corresponding to maximum eigenvalue, we obtained 737 sets of eigenvector after running WFA, and each set of eigenvector has 19 elements (coefficient) corresponding to 19 kinds of stock. And since the eigenvector represents the impact extent of each kind of stock, therefore, absolute value of eigenvector which was used to measure the size of the impact extent was taken.

When time is taken as axis, 19 sets of standard deviations to show the various extent of each kind of stock of information carrying capacity in this study, with the greater the variability indicating greater variation of information that was carried by such stock were obtained. Similarly, If space is taken as axis, 737 sets of standard deviations, which are known as the spatial variability sequence and which show the 19 stocks' daily information on the carrying capacity of the variation of the 19 stocks will be obtained. If the variation degree is bigger, then it expresses that the amount of information carried in the more inconsistent of that day's various stocks, and the difference between changes in stocks, may be larger. Thus, this study was analyzed by spatial variability analysis.

As the size of the information-carrying capacity was defined by the interaction of investors and the market, whether by the efficient market hypothesis or the behavioral finance theory, once an investor makes an investment, the phenomenon of rational judgments or over-optimism of information will react on the market. Karpoff (1986) argues that the relevant information of price changes is one source of the trading volume, and that the trading volume can provide some information that could not be provided by price (Blume et al., 1994). Furthermore, the trading volume can disclose information other than stock returns, and it also can be taken as the future direction of stock prices further amended (Choi and Kim, 2001). Therefore, this study suggests that the stock market trading volume reflects the information-carrying capacity of the market. The greater the volume, the greater the capacity of information carried. In contrast, the smaller the volume, the smaller the carrying capacity. When the transaction volume is big, the various kinds of stock will also carry information simultaneously and spatial variability would be reduced. Otherwise, if the volume is small, only part of the stock will be carrying information, and then spatial variation will be amplified. So when the trading volume and spatial variability are not symmetrical, the market may have unusual information. This research processed normalization for trading volume and spatial variability series, and multiplied the two series to obtain the information sequence. Under normal circumstances, the information sequence is relatively stable, while under the abnormal circumstances there will be a large change in sequence. In this study, we use the average of the information sequence as a basis for judging. When the data are larger than the average, that means the phenomenon of abnormal information, and catch the sign of market variation.

RESULTS

In this study, we classified stock index in accordance with equation (1), transfer to a continuous rate of return, then take a descriptive statistical result of return as shown in Table 1. The main purpose of the study is to compare the information and original portfolios, to identify which kind of construction method of a portfolio could improve the efficiency of the portfolio. With access to the index return of stocks, we began to establish the 30-day formation phase, and the 15-day forecast phase for WFA. We applied the Bouchaud and Potters (2000) steps to filter

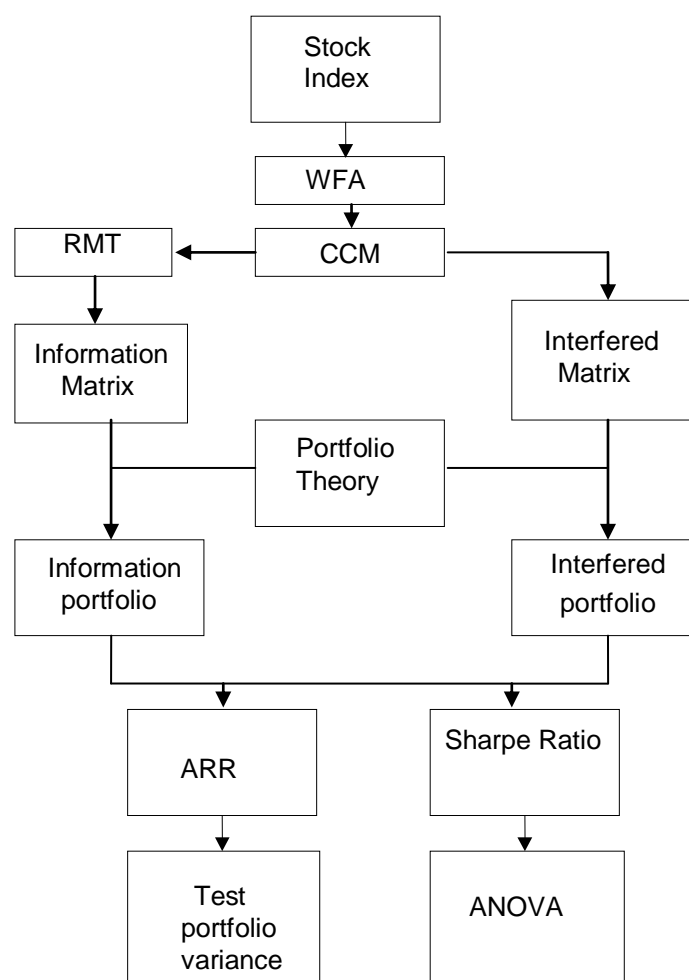


Figure 1. Portfolio performance test flowchart. We obtain 722 data sets of cumulative returns and Sharpe ratios after the WFA process, followed by statistical analysis of these data sets. First, we calculated the variance of information and the original portfolio by cumulative return rate, then we tested the variance of two portfolios by F test to confirm which risk is smaller. Then we used the ANOVA approach to test the Sharpe ratio for two portfolios to find the differences, and the process paired test by the least square mean, comparing the two portfolios to discover which performance is better.

out the impact of random matrices and build an information matrix, and we used the minimum variance portfolio weights that were proposed by Markowitz (1952) to calculate the forecast performance of the original portfolio and the information portfolio to obtain 722 groups of forecast performances. We carried out the variance test during the return of the holding period, and the test results are listed in Table 3, which shows which risk is smaller, and in summary in Table 2, which shows the statistical result of the return rate and the Sharpe ratio during the holding period. Next, we used the ANOVA approach to test the difference of the Sharpe ratio between the original and the information portfolios and then summarized the

results as shown in Table 4.

In Table 4, the means of the two Sharpe ratio indexes are significant, but we could not find out which one is higher than the other. We also calculated and compare the means of least square with those paired test and determine which performance is better, shown in Table 5.

Table 5 shows that the Sharpe ratio of information, and the original portfolio is significantly different, with a significance level of 5%. Therefore, the mean of the Sharpe ratio of the information portfolio is more significant than that of the original portfolio. The portfolio after filtered random matrix is indeed more efficient.

After analyzing the efficiency of the portfolio, the part of

Table 1. Described statistical tables of index returns¹.

Stock index	Mean (%)	Std. deviation	Skewness	kurtosis
Cement ind.	0.006	0.026	0.033	3.333
Food ind.	0.037	0.022	-0.208	3.860
Plastics ind.	0.019	0.018	0.092	4.872
Textile ind.	0.015	0.022	-0.142	3.507
Electric machinery	-0.021	0.016	-0.537	4.925
Cable ind.	-0.040	0.022	-0.305	3.850
Chemical ind.	0.011	0.020	-0.462	3.939
Glass-ceramic ind.	0.010	0.025	-0.136	3.784
Papermaking ind.	-0.011	0.022	-0.241	3.910
Steel ind.	0.002	0.020	-0.005	4.072
Rubber ind.	0.070	0.023	-0.248	3.547
Motor vehicle ind.	-0.039	0.020	0.069	4.206
Electron ind.	-0.007	0.018	-0.284	4.574
Construction ind.	-0.006	0.027	-0.189	3.096
Transportation ind.	-0.010	0.022	-0.160	3.654
Tourism ind.	0.013	0.027	-0.012	3.481
Finance ind.	-0.035	0.023	-0.099	4.258
General merchandise	0.033	0.019	0.041	4.089
Others	0.005	0.015	-0.284	4.684

Table 2. Described Statistical Result of Returns and Sharpe Ratio During Holding Period².

	Portfolio	Mean (%)	Std. Deviation	Skewness	Kurtosis
ARR	Original	0.0096	0.1486	0.6402	1.7756
	Information	0.0022	0.0861	-0.0055	1.0495
SR	Original	0.0065	0.2887	0.0043	0.0231
	Information	0.0409	0.3024	0.2117	0.3735

Table 3. The difference test table of return variance of information³ and original portfolio during holding period.

H0 of portfolio	F value	P value
Information / Original	0.3349**	< 0.001

* = 5 significance level, ** = 1 significance level.

Table 4. Sharpe ratio's variance analysis of information and original portfolio⁴.

Variance source	Degree of freedom	Sum squares	Mean squares	F-value	P-value
Between variance	1	0.4266	0.4266	4.88*	0.027
Within variance	1442	126.0222	0.0874		
Sum of variance	1443	126.4488			
Samples	1444				

* = 5 significance level, ** = 1 significance level.

¹ 766 samples

² 722 samples

³ There are 722 samples and used a two-tailed test in this study in Table 3, the variance of the information portfolio is more significant than that of the original portfolio, and the F value is less than 1, indicating that the variance of the information portfolio is smaller than that of the original portfolio. It can be seen, after filtering the random matrix, that the risk with the structured portfolio is less than with the filtered random matrix.

⁴In Table 4, the means of the two Sharpe ratio indexes are significant, but we could not find out which one is higher than the other. To calculate the means of least square with paired test and determine which performance is better, we can compare the results with those shown in Table 5.

Table 5. Sharpe Ratio's Compare Table of Information and Original Portfolio⁵

Portfolio	Original (%)	Information (%)	Mean difference (%)	T-value	P-value
Result	0.0065	0.0409	- 0.034	-2.21*	0.0273

* = 5 significance level, ** = 1 significance level.

unusual information is discussed. Random matrix theory described that the information eigenvalue has a strong persistence nature, which means its information-carrying capacity is larger. The principal component analysis demonstrates that the first principal component is the greatest degree of overall data interpretation. Therefore, this study will analyze eigenvector, which corresponded to the largest eigenvalue. First, we adopt the WFA of the 30-day formation phase, then we calculate the eigenvalue and eigenvector of each phase, and select the eigenvector that corresponded to the largest eigenvalue (hereinafter referred to as eigenvectors). First, we take time as the axis to calculate the eigenvector of stocks of standard deviations throughout the study period (a total of 19 standard deviations). Then we take space as the axis and calculate the standard deviation of eigenvector of 19 stocks within each time phase to obtain the spatial variability sequence (a total of 737 standard deviations). This study suggests that the stock market trading volume, just to reflect the information-carrying capacity of the market, has the symmetry with spatial variation. First, to normalize the standard sequence with which we calculated the base on the spatial axis and stock trading volume, which were corresponding to the standard sequence, we make two data sets limited to between 0 and 1. Then, we multiply the two data sets to obtain a new information sequence. Figure 2 shows the weighted index, spatial variability, stock market trading volume, and information sequence after normalization, respectively. There are several peaks in the information sequence corresponding to the transition time point of market ups and downs. Figure 3 shows the results of time variation, with the horizontal axis being the eigenvectors corresponding to the 19 kinds of stocks, and the vertical axis being time variation.

By the time variation, we can find the size of the information-carrying capacity variation for the various kinds of stocks in this study. The stock of number 16 (tourism) has the greatest variation, which points to the deregulation by the Taiwan government in July 2008 to let

Mainland China tourists visit Taiwan, with all of the related impacts this caused.

Figure 4, in the lower figure, shows the information sequence, horizontal representative of mean value, and a comparison of the information sequence with upper figure of the market index, and we find that information sequences will exceed the mean before the trend of stock market decline. This study suggests that we take the mean to be a filter criterion, define the portion of the information sequence that is greater than the mean as abnormal information; then there are large amounts of abnormal information around on 110 (peak 1), 160 (peak 2), 300 (peak 3), and 730 (peak 4) days (about 2007/7/31, 2007/10/31, 2008/5/14, and 2010/1/4) before a decline in the markets.

However, there is little abnormal information that appears before the upward trend, around 460, 540, and 690 days (about 2008/11/27, 2009/4/17, and 2009/11/20), which shows that the proposed method of this study is more suitable for detecting the situation of bull market change to bear market.

Conclusion

Through the method of filtering out interrelated factors by the random matrix theory, this study provides a set of correlation coefficients based on total randomness for the improvement of portfolio efficiency. The test results of variance showed that the variance of accumulated return of the information portfolio was significantly lower than that of the original portfolio, which means that the information portfolio has lower risk. We found that the mean of the information portfolio Sharpe ratio was significantly different and greater than the original portfolio after we analyzed the Sharpe ratio by ANOVA. These results indicate that, if we are able to filter out the impact of the randomness, it can effectively reduce portfolio risk and improve portfolio performance. By the test results, there is sufficient evidence in this study to claim that the randomness has a significant impact on portfolio theory. Such randomness will increase the risk of the portfolio, reducing the efficiency of the portfolio, if we use the Markowitz portfolio theory hastily, and will cause a higher risk with poor performance during the constructing of the portfolio.

The second part of this study, using the concept of principal component to analyze the information eigenvectors, hoped to capture the conversion time point of bull market transfer to bear market. Using both spatial variability extent and stock market trading volume of information eigenvector, we measured at various points the information-carrying capacity so as to inspect the signals of abnormal information carrying. With the mean of the information sequence as a judgment to filter out some value less than the mean, the signals of abnormal

⁵Number of sample 1444(722 * 2) in Table 5 and shows the Sharpe ratios of the information, the original portfolios are significantly different with a significance level of 5. Therefore, this test showed a significance level at 5, and so the mean of the Sharpe ratio of the information portfolio is more significant than that of original portfolio. The portfolio after filtered random matrix is indeed more efficient. Mean difference = Original – Information.

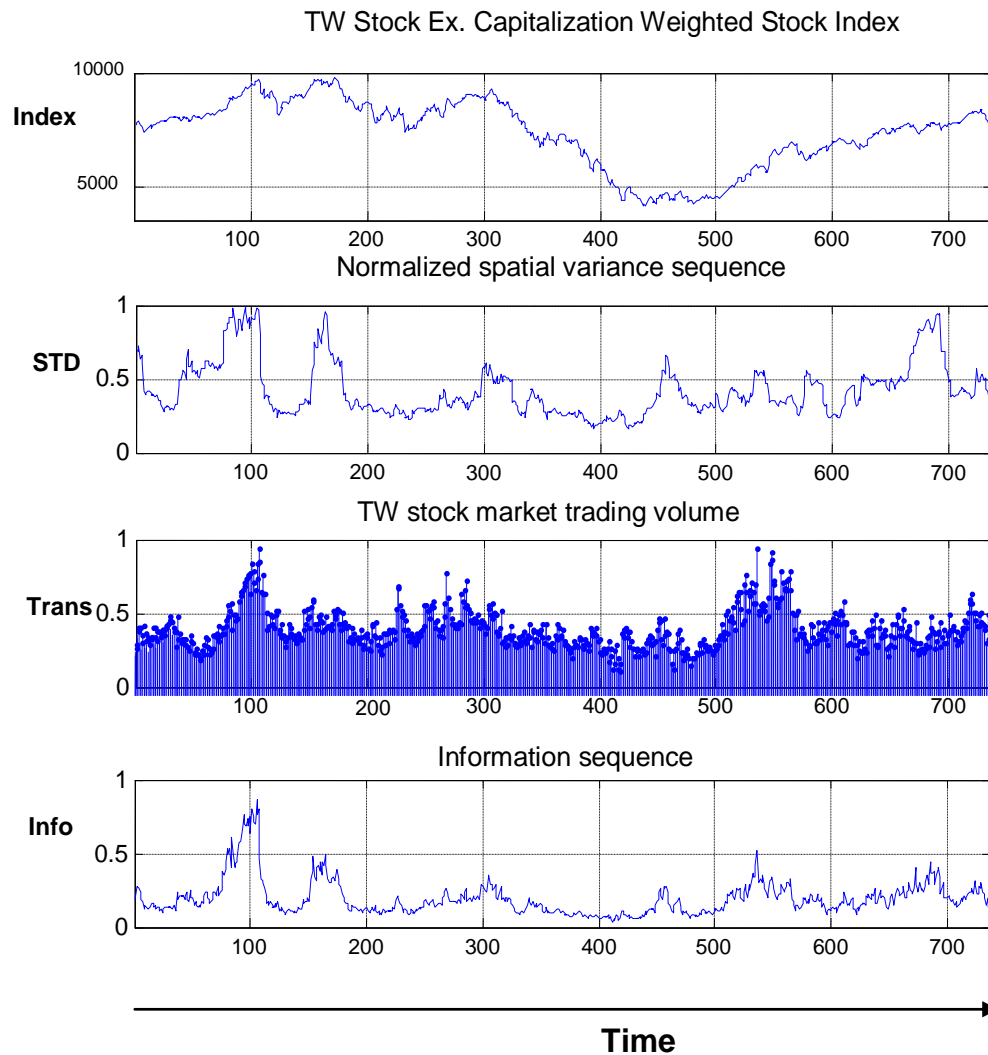


Figure 2. Information sequence and relevant figure. **Index** for the stock market index, **STD** for the spatial variability sequence after normalization, **Trans** for the trading volume after the normalization, **info** for the information sequence. For all charts, the x-axis is time, with the time period being from 2007/2/13, to 2010/1/29.

information carrying can be clearly seen. However, after corresponding with the weighted stock index, we only found at first two turning points during the sub-prime mortgage crisis (the first peak close to the weighted stock index of 10,000) and the economy cooling period (the second peak close to weighted stock index of 10,000), while the period of the global financial crisis, although detectable, is not significant when compared with the first two peaks. The second half of study period, due to the fact that the market was flooded by abnormal information signals, showed that this method cannot focus on the detection of highlights of bull market and bear market conversion. We argued that we cannot effectively detect the conversions for two reasons. (1) This study only used

the first principal component analysis, and we may add the other main component in to get better results. (2) In this study, only 19 stocks were analyzed, which is sufficient to describe the stock market, but it is too small a sample to calculate spatial variance, which may have led to the poor results.

In this study, we use the stock index of the Taiwan Stock Exchange (TWSE), and so because of the regulation the findings may be impacted by price fluctuation limits. This study confirmed that the randomness does, through the correlation matrix, impact the portfolio, but how to form and what is randomness? There are no answers in the research literature right now, and these could be considered as future research directions. In this

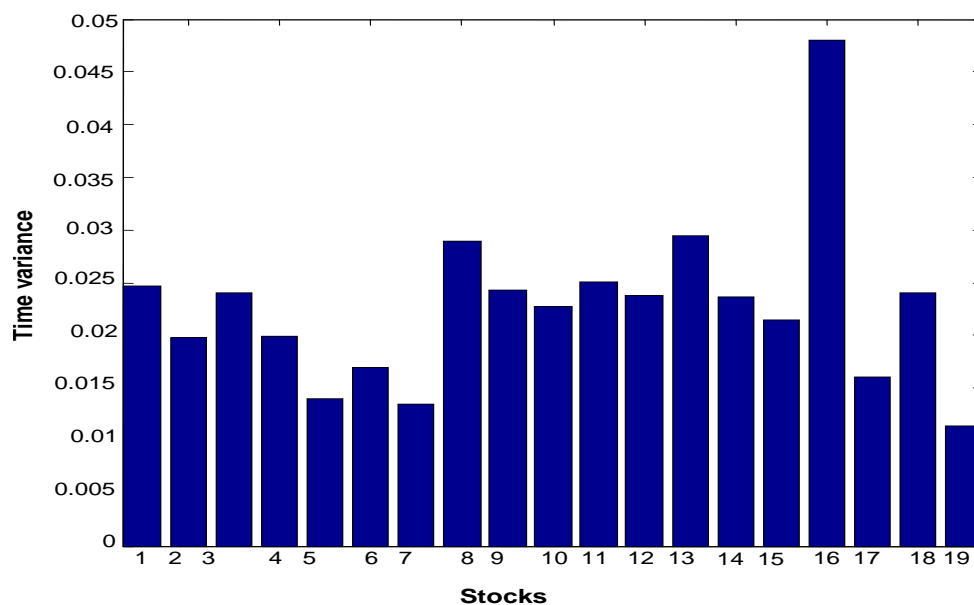


Figure 3. Bar chart of time variance.

Remarks: The statistic of time variation of eigenvector for 19 kinds of stock.

Arabic numerals stand for: 1: cement industry, 2: food industry, 3: Plastics industry, 4: Textile industry, 5: Electric machinery industry, 6: Cable industry, 7: Chemical Industry, 8: Glass-ceramic industry, 9: Papermaking industry, 10: Steel industry, 11: Rubber industry, 12: motor vehicle industry, 13: Electronics industry, 14: Construction industry, 15: Transportation industry, 16: Tourism industry, 17: Finance industry, 18: General merchandise industry, and 19: others.

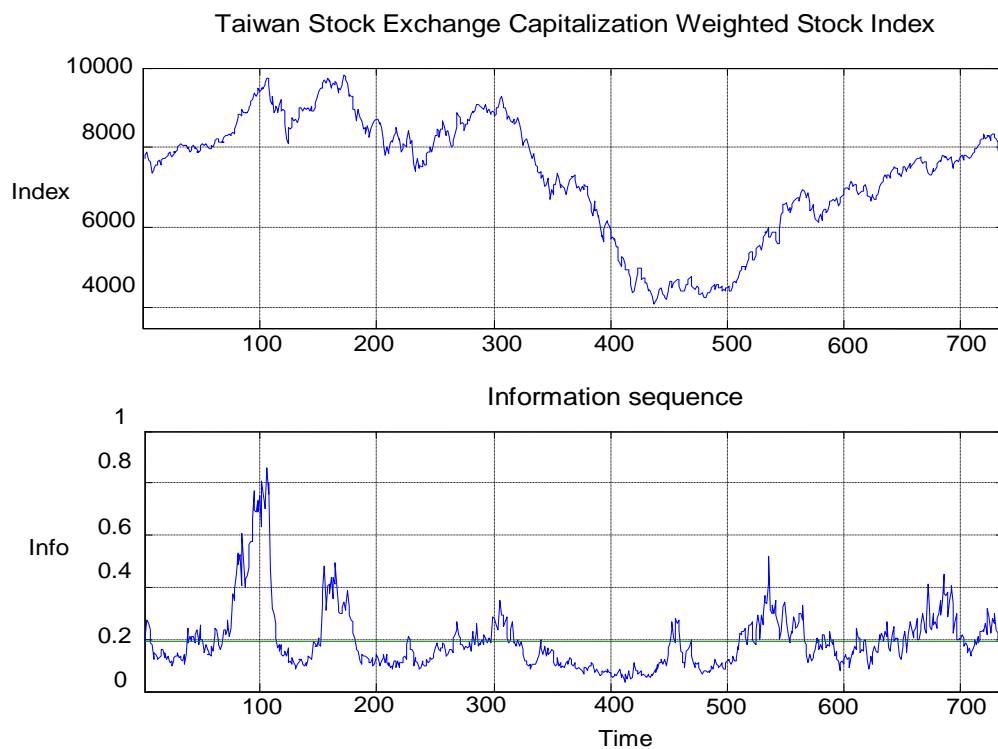


Figure 4. Information sequence

Index for the market index, info for the information row. All charts, x-axis is time, with time to start 2007/2/13, termination date 2010/1/29.

study, we use stock index for the information eigenvector analysis, but we may also try to use industry chain for further research, in which we may get better information about the interaction mechanism.

ACKNOWLEDGMENTS

We thank Professor Gregory C.C. Niu and Ray Hardesty for English modification. Shih-Chang Shu, and Sze-Hsun Chang for helpful discussions.

REFERENCES

- Abeysydera SP (2001). Efficient Markets Hypothesis and the Emerging Capital Market in Sri Lanka: evidence from the Colombo Stock Exchange – a note. *J. Bus. Financ. Account.* 28(1/2):249-261.
- Bachelier L (1900). *Theorie de la speculation*. Annales scientifiques de l'École Normale Supérieure 17:21-86
- Bernard VL, Thomas JK (1990). "Evidence that stock prices do not fully react the Implications of current earnings for future earnings". *J. Account. Econ.* 13(4):305-340.
- Blume L, Easley D, O'Hara M (1994). Market statistics and technical analysis: the role of volume. *J. Financ.* 49: (153-182).
- Bouchaud J, Potters M (2000). *Theory of financial risk-from statistical physics to risk management*. Cambridge University Press, Cambridge, UK.
- Choi W, Kim JW (2001). Underreaction, trading volume, and post-announcement drift. *Social Science Research Network* pp.1-51.
- Cutler D, Porterba J, Summers L (1991). *Speculative Dynamics*. *Rev. Econ. Stud.* 58(195):529-546.
- De Bondt WFM, Thaler RH (1985). Does the stock market overreact? *J. Financ.* 40(3):793-808.
- Fama EF (1965). The behavior of stock market prices. *J. Bus.* 38(1):34-105.
- Fama EF (1970). Efficient capital market: a review of theory and empirical work. *J. Financ.* 25(2):383-417.
- Fama EF (1998). Market efficiency, long-term returns and behavioral finance. *J. Financ. Econ.* 49(3):283-306.
- French K, Richard R (1986). Stock return variance: the arrival of information and the reaction of traders. *J. Financ. Econ.* 17(1):5-26.
- Hong H, Terence L, Jeremy CS (2000). Bad news travels slowly: size, analyst coverage, and the profitability of momentum strategies. *J. Financ.* 55(1):265-295.
- Jegadeesh N, Titman S (1993). Returns to buying winners and selling losers: implications for stock market efficiency. *J. Finan.* 48(1):65-91.
- Kaiser HF (1960). The application of electronic computers to factor analysis. *Educ. Psychol. Meas.* 20(1):141-151.
- Karpoff JM (1986). A theory of trading volume. *J. Financ.* 41:1069-1087.
- Lakonishok J, Shleifer A, Vishny R (1994). Contrarian investment, extrapolation, and risk. *J. Financ.* 49(5):1541-1578.
- Laloux L, Cizeau P, Bouchaud JP, Potters M (1999). Noise dressing of financial correlation matrices. *Phys. Rev. Lett.* 83(7):1467-1470.
- Malkiel BG (2005). Reflections on the efficient market hypothesis: 30 Years Later. *Financ. Rev.* 40(1):1-9.
- Mandelbrot B (1966). Forecasts of future prices, unbiased markets, and martingale models. *J. Bus.* 39(special supplement):394-419.
- Marcenko VA, Pastur LA (1967). Distributions of eigenvalues of some sets of random matrices. *Mathematics of the USSR-Sbornik* 1(4): 507-536.
- Markowitz HM (1952). Portfolio selection. *J. Financ.* 7(1):77-91.
- Milonis AE, Moschos D (2000). On the validity of the weak-form efficient markets hypothesis applied to the London stock exchange: comment. *Appl. Econ. Lett.* 7(7):419-421.
- Ozdemir ZA (2008). Efficient market hypothesis: evidence from a small open-economy. *Appl. Econ.* 40(5):633-641.
- Plerou VP, Gopikishnan P, Rosenow B, Amaral LAN, Guhr T, Stanley HE (2002). Random matrix approach to cross correlations in financial data. *Phys. Rev. E* 65(6):1-18.
- Rouwenhorst KG (1998). International momentum strategies. *J. Financ.* 53(1):267-284.
- Samuelson PA (1965). Proof that properly anticipated prices fluctuate randomly. *Ind. Manag. Rev.* 6(2):41-49.
- Sharifi S, Crane M, Shamaie A, Ruskin H (2004). Random matrix theory for portfolio optimization: a stability approach. *Physica A: Stat. Mech. Appl.* 335(3/4):629-643.
- Sharma A (2009). Impact of public announcement of open offer on shareholders return: an empirical test for efficient market hypothesis. *IUP J. Appl. Financ.* 15(11):38-51.
- Sharpe WF (1966). Mutual fund performance. *J. Bus.* 39(1):119-138.
- Sharpe WF (1975). Adjusting for risk in portfolio performance measurement. *J. Portf. Manag.* 1(2): 29-34.
- Sharpe WF (1994). The Sharpe ratio. *J. Portf. Manag.* 21(1):49-58.
- Shefrin H (2000). *Beyond Greed and Fear: understanding behavioral finance and the psychology of investing*, Boston, MA: Harvard Business School Press.
- Shiller RJ (1981). Do Stock Prices move too much to be justified by subsequent changes in dividends? *Am. Econ. Rev.* 71(3):421-498.
- Stein C (1969). *Multivariate analysis I*, Technical Report, Stanford University, Department of Statistics. Notes prepared by ML Eaton in 1966:79-81.
- Wigner EP (1951). On a class of analytic functions from the quantum Theory of Collisions. *Ann. Math.* 53(1):36-67.