Available online at www.sciencedirect.com
ScienceDirect
Procedia Computer Science 60 (2015) 1247 – 1254
Procedia Computer Science

19th International Conference in Knowledge Based and Intelligent Information and Engineering Systems

Verification of the Relationship Between the Stock Performance and the Randomness of Price Fluctuation

Mieko Tanaka-Yamawaki[a]*, XinYang[b], Yuuta Mikamori[c]

[a]*Department of Information.& Knowledge Eng., Graduate School of Engineering, Tottori University, 101-4, Koyamacho-Minami, Tottori,680-8662 Japan*
[b]*Taizhou Vocational & Technical College, Qianjiang Rd, Jiaojiang, Taizhou, Zhejiang, China*
[c]*Probizmo., Ltd., Eiko-Ekinan Bldg., 1-3-2 Minamimachi, Izumo, Shimane, 693-0008 Japan*

Abstract

The authors examine the validity of the empirical rule connecting the randomness of high frequency stock prices to its future performance in a bully market conditions. For this purpose, the U.S. market in the period of 1993-1997 is chosen for investigation. The rule was first discovered in a bear market of 2007-2009 in Tokyo market, as one of the useful applications of the RMT-Test which is a new tool to measure the randomness of given time series based on the random matrix theory, showing that the stock of the highest randomness is more profitable than the Nikkei Average Price throughout the following year. The previous analysis was limited to the period of bear market, and inclusive analysis for a wider market conditions are necessary in order to establish the validity of the rule.

Peer-review under responsibility of KES International

Keywords: Randomness, RMT-test, tick-wise stock price, bear market, bull market

1. Introduction

The authors have proposed to measure the randomness of a given long time series based on RMT[1], named the RMT-test[68], and have demonstrated the effectiveness of this method by measuring the randomness of the physical random numbers and the pseudo-random numbers, and applied the same tool to measure various real-world time series including price fluctuation this process, a new empirical rule has been discovered, stating that the stocks *having higher randomness in the previous year tend to perform better than the stocks of lower randomness*[5] Although it is a striking discovery, the analysis was limited only the period 2007-2009 in the Tokyo market, which was a typical example of the weak market condition. In order to establish the validity of the rule, it is necessary to examine

* Corresponding author. Tel.: +81-0857-31-5223; fax: +81-0857-31-0879.
E-mail address: mieko@eecs.tottori-u.ac.jp.

Peer-review under responsibility of KES International
doi:10.1016/j.procs.2015.08.190

whether the same rule holds in a strong market. For this purpose, the U.S. market in the period of 1993 - 1997 is employed and thoroughly examined.

2. The RMT test

2.1. Mathematical Preparation[2,3]

The method of the RMT-test is outlined as follows. We aim to test the randomness of a long 1-dimensional sequence of numerical data, S.

At the first step, we cut S into N pieces of equal length T, then shape them in an N×T matrix, $S_{i,j}$, by placing the first T elements of S in the first row of the matrix $S_{i,j}$, and the next T elements in the 2nd row, etc., by discarding the remainder if the length of S is not divisible by T. Each piece, $S_i = (S_{i,1}, S_{i,2}, \dots, S_{i,T})$, is converted to a normalized vector $x_i = (x_{i,1}, x_{i,2}, \dots, x_{i,T})$ by means of

$$x_{i,t} = \frac{S_{i,t} - \langle S_i \rangle}{\sigma_i} \quad (1)$$

where,

$$\langle S_i \rangle = \frac{1}{T} \sum_{t=1}^{T} S_{i,t} \quad (2)$$

and

$$\sigma_i = \sqrt{\langle S_i^2 \rangle - \langle S_i \rangle^2} \quad (3)$$

such that every row in the new matrix x has mean=0, variance =1. Since the original sequence S is random, in general all the rows are independent, i.e., no pair of rows is identical.

The cross correlation matrix $C_{i,j}$ between two stocks, i and j, is constructed by the inner product of the two time series, $x_{i,t}$ and $x_{j,t}$,

$$C_{i,j} = \frac{1}{T} \sum_{t=1}^{T} x_{i,t} x_{j,t} \quad (4)$$

thus the matrix $C_{i,j}$ is symmetric under the interchange of i and j.

A real symmetric matrix C can be diagonalized by a similarity transformation $V^{-1}CV$ by an orthogonal matrix V satisfying $V^t=V^{-1}$, each column of which consists of the eigenvectors of C. Such that

$$C v_k = \lambda_k v_k \ (k=1,\dots,N) \quad (5)$$

where the coefficient λ_k is the k-th eigenvalue and v_k is the k-th eigenvector.

According to the RMT, the eigenvalue distibution spectrum of the cross correlation matrix C of random series is given by the following formula called Marcenko-Pastur distribution[4]

$$P_{RMT}(\lambda) = \frac{Q}{2\pi} \frac{\sqrt{(\lambda_+ - \lambda)(\lambda - \lambda_-)}}{\lambda} \quad (6)$$

$$\lambda_\pm = (1 \pm Q^{-1/2})^2 \quad (7)$$

valid at the limit of N and T going to infinity, keeping

$$Q = K/N \quad (8)$$

as a constant.

2.2. Quantitative Version of the RMT-Test

The quantitative version of the RMT-Test[5] is used to discriminate subtle difference of randomness among good random numbers. The k-th moment of the eigenvalues

$$m_k^{EXP} = \frac{1}{N}\sum_{i=1}^{N}(\lambda_i)^k \tag{9}$$

is compared to the corresponding theoretical moments derived from P_{RMT} in Eqs. (1)-(3) of the random matrix theory.

$$m_k^{RMT} = \int_{\lambda_-}^{\lambda_+} \lambda^k P_{RMT}(\lambda) d\lambda \tag{10}$$

The theoretical moments up to k=6 can be explicitly obtained as a function of Q=L/N as follows[11].

$$m_1^{RMT} = 1 \tag{11}$$

$$m_2^{RMT} = 1 + \frac{1}{Q} \tag{12}$$

$$m_3^{RMT} = 1 + \frac{3}{Q} + \frac{1}{Q^2} \tag{13}$$

$$m_4^{RMT} = 1 + \frac{6}{Q} + \frac{6}{Q^2} + \frac{1}{Q^3} \tag{14}$$

$$m_5^{RMT} = 1 + \frac{10}{Q} + \frac{20}{Q^2} + \frac{10}{Q^3} + \frac{1}{Q^4} \tag{15}$$

$$m_6^{RMT} = 1 + \frac{15}{Q} + \frac{50}{Q^2} + \frac{50}{Q^3} + \frac{15}{Q^4} + \frac{1}{Q^5} \tag{16}$$

The quantified criterion of randomness is identified as the deviation of Eq.(9) from Eq.(10). The authors chose to use the 6th moment (i.e., k=6) in order to define the level of randomness in the RMT-test [2,3,5] as the inverse of

$$|\mathrm{Error}| = |m_6^{EXP} - m_6^{RMT}| \tag{17}$$

Note that we do not need to solve the eigenvalue problem in Eq. (5) directly but simply obtain the trace (sum of all the diagonal elements)

$$\sum_{i=1}^{N}(C^k)_{i,i} = \sum_{i=1}^{N}(\lambda_i)^k \tag{18}$$

of the matrix of the k-th power C^k of the cross correlation matrix to obtain the k-th moment.

3. Randomness and Stock Performance ('07-'09)

The RMT-test is now applied to measure the randomness of the stock time series. It is customary to convert the price time series $p_1, p_2, ..., p_T$ to the log-return time series $r_1, r_2, ..., r_T$

$$r_i = \log(p_i/p_{i-1}) \tag{19}$$

in the financial analysis, in order to eliminate the unit/size dependence of different stock prices. Minute data of stocks in TOPIX 500 from 2007 to 2009 per minute satisfying the appropriate conditions are selected and used for analysis. Minute data mean the time series stamped in seconds or minutes which record the information of traded or quoted prices. Since trades or quotes may not occur at every time period, the lengths of the minute data are not fixed. For this reason, some work is required to prepare the fixed-length time series to serve for analysis. The blanks are filled by copying the previous data as long as the added part is less than the 20 percent of the total length in order to calculate equal time correlation of each stock price. As a result, the data length (T) and the number of stocks (N) are different at each year. The values of T and N for each year used in this analysis was N=211, T=66338 for 2007, N=240, T=66338 for 2008, and N229, T=65945 for 2009. The free parameter Q=K/N of the RMT-test was chosen to be Q=4, to give an empirical rule:"*having higher randomness in the previous year tend to perform better than the stocks of lower randomness"*

Note that this period was a falling market throughout, as shown in Fig.1. Moreover, the Nikkei average price index fell from 13000 yen to 8000 yen in a month from October to November, 2008, due to the huge earthquake known as the Lehman shock. The ranking of randomness measured by using the log-return time series in Eq.(19) based on the RMT-test is listed in Table 1, in the descending order of randomness. The prices of the stocks 9504, 6460, 9506, 9508 which are the top 4 in this ranking of 2007 are compared to the performance of 7201 which had the lowest randomness in 2007 are shown to outperform throughout the year of 2008, in Fig.2.

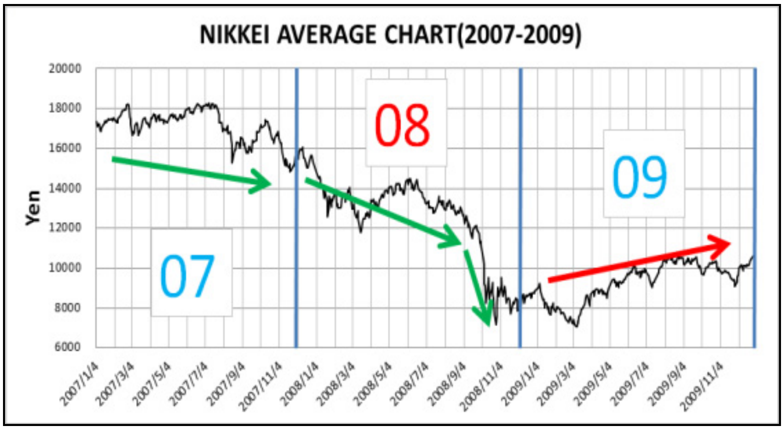

Fig.1 NIKKEI average chart from 2007 to 2009(Fig.5 in Ref. 5

Table 1 The ranking of randomness by using the minute data of 2007(Table 3 in Ref.5)

Rank	Sector	Code	|Error|
1	Electric/ Gas	9504	26.4
2	Machinery	6460	37.6
3	Electric/ Gas	9506	38.2
4	Electric/ Gas	9508	43.3
5	Information & Communication	4676	44.9
	...		
207	Electric Equipment	6506	740.9
208	Non-ferrous Metal	5802	797.3
209	Chemistry	4043	799.8
210	Iron & Steel	5541	1001.5
211	Car	7201	1209.6

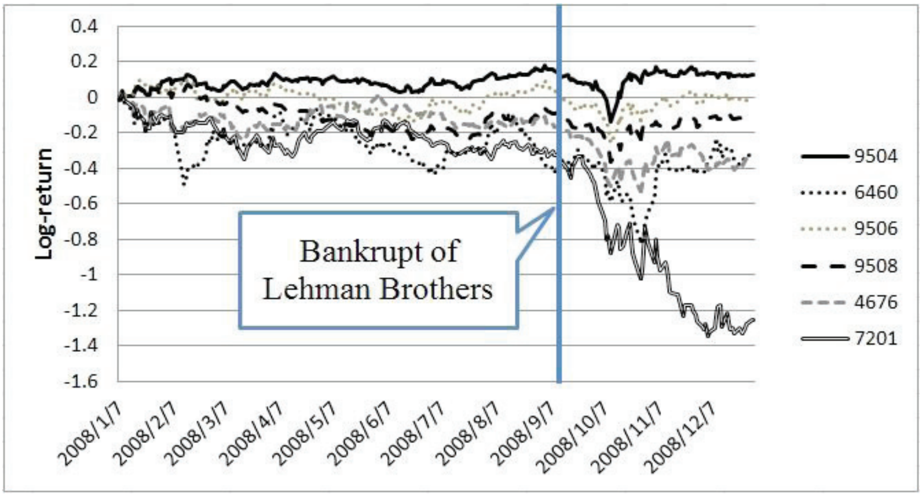

Fig.2 The Top5 stocks of the highest randomness are safer than the stock of the lowest randomness, 7201;Fig.6 in Ref.5)

4. Rule on the randomness and Stock Performance tested in '93-'97: the period of rising market

Although the empirical rule seems to work in the period 2007-2009 in Tokyo market, this period was a weak market throughout. Thus it is necessary to check the rule in a different type of market conditions. For this reason, the authors investigate the period 1993-1997 in U.S. market, as an example of strong markets. As shown in Fig. 3, the S&P500 index draws a rising curve in the period of 1993-1997.

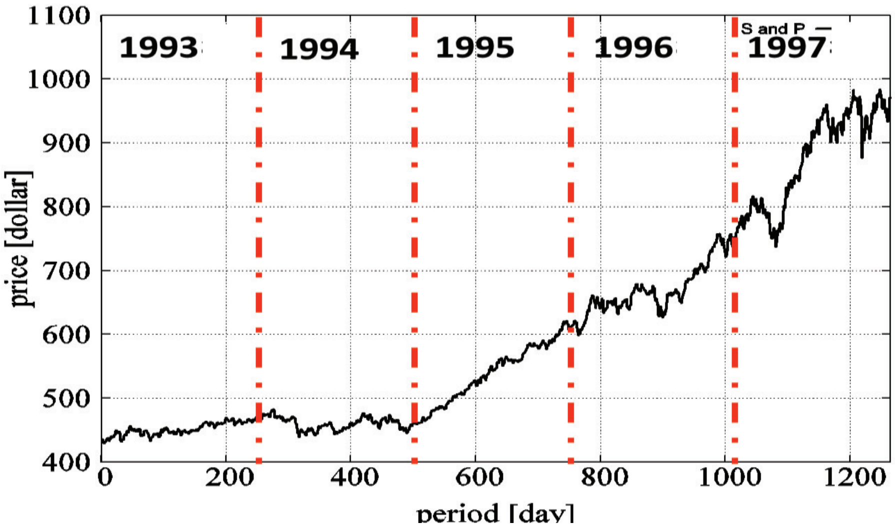

Fig.3 S&P500 index in the period of 1993-1997

In order to make a comparison to the case of falling market investigated in Ref.5, the original tick-wise prices are converted to equal interval time series per 1 minute. Substitution is made by the previous price so that the price table for every minute is filled. However, the condition applied in the previous section for 2007-2009 TOPIX data "as long as the substitution does not exceed 20% of the entire string" is not applied. Because such a condition would reduce N (number of stocks) to 15-50 for each year, not sufficient to apply RMT-Test. The data set thus chosen is 222, 244, 250, 263 stocks for 1993, 1994, 1995, 1996, respectively.

The result of quantifying the randomness in 1993-1997 US Stock Market in terms of |Error| defined in Eq.(17) are summarized in Tables 2-5. Based on the rankings obtained, the next years performance of H (average of the top 5) and L (average of the bottom 5) are compared with S&P500 index in the following years, in Fig. 4-7.

Table 2 Randomness in 1993 (The highest 5 and the lowest 5)

Order	Stock	Sector	|Error|%
1	AMP	Finance	3.0
2	BLL	Material	7.2
3	IFF	Material	7.3
4	CSC	IT	7.6
5	WHR	Consumer/Service	9.5
…			
218	PEP	Daily Necessities	2935
219	T	Telecommunication	2973
220	MO	Oil	3166
221	WMT	Daily Necessities	3205
222	BMY	Healthcare	3436

Table 3 Randomness in 1994 (The highest 5 and the lowest 5)

Order	Stock	Sector	|Error|%
1	ETN	Capital goods	9.1
2	DHR	Capital goods	18
3	MAT	Consumer/Service	21
4	KMB	Daily Necessities	23
5	BDX	Healthcare	24
…			
240	PEP	Daily Necessities	2200
241	GE	Capital goods	2201
242	MRK	Healthcare	2814
243	T	Telecommunication	3367
244	RHI	Capital goods	4576

Table 4 Randomness in 1995(The highest 5 and the lowest 5)

Order	Stock	Sector	|Error|%
1	AFL	Finance	1.2
2	TE	Public service	1.3
3	HON	Capital goods/Service	3.7
4	PH	Capital goods/Service	4.7
5	HP	Energy	10
…			
246	WMT	Daily Necessities	2100
247	MRK	Healthcare	2113
248	F	Car	2393
249	T	Telecommunication	2435
250	GE	Capital goods/Service	2794

Table 5 Randomness in 1996 (The highest 5 and the lowest 5)

Order	Stock	Sector	|Error|%
1	MAR	Consumer/Service	0.24
2	VMC	Material	0.68
3	DE	Capital goods/Service	1.7
4	CTL	Telecommunication	2.9
5	CPB	Daily Necessities	4.3
…			
259	ADM	Daily Necessities	1615
260	KO	Daily Necessities	1925
261	F	Consumer/Service	1963
262	PEP	Daily Necessities	2414
263	T	Telecommunication	30.91

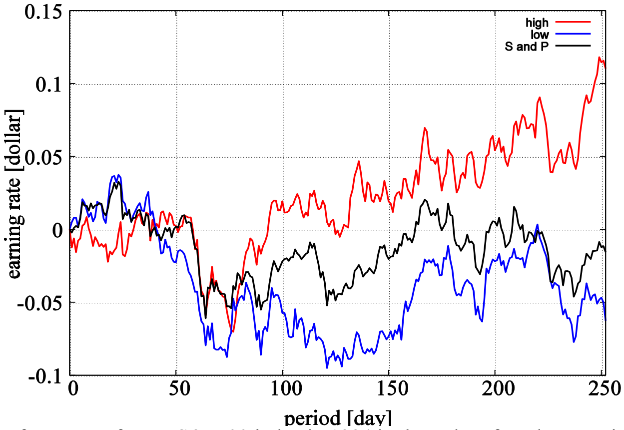

Fig. 4 Performance of H, L, S&P500 index in 1994 in the order of randomness in 1993

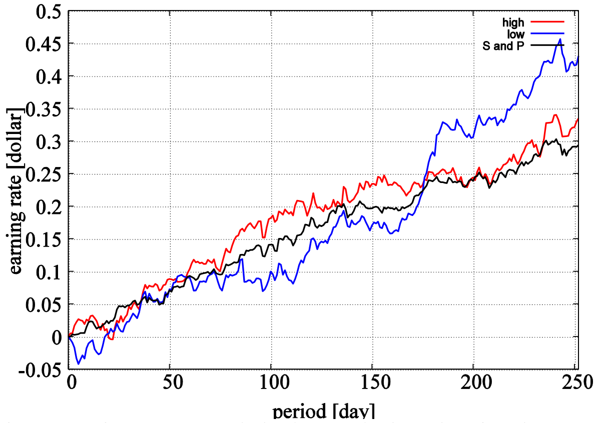

Fig.5 Performance of H, L, S&P500 index in 1995 in the order of randomness in 1994

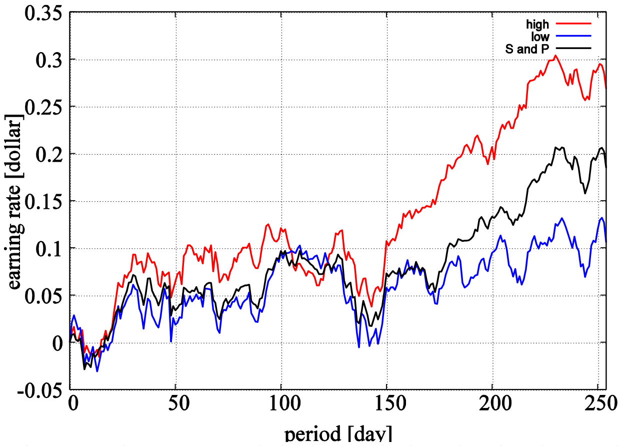

Fig. 6 Performance of H, L, S&P500 index in 1996 in the order of randomness in 1995

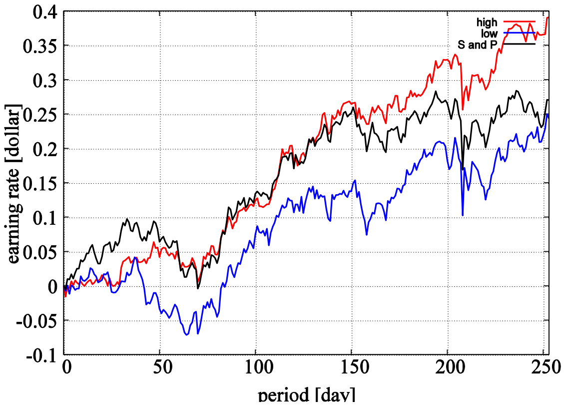

Fig. 7 Performance of H, L, S&P500 index in 1997 in the order of randomness in 1996

5. Summary

In this paper**,** the authors have investigated whether the empirical rule "stocks of higher randomness perform better in the next year" holds in wider cases than the example studied in Ref.5, since the rule was originally found for 2007-2009 Japanese market, a typical bear market. The same line of analysis based on the RMT-test is applied on 1993-1997 US market, which is an example of bull market.

It is found that this rule holds for three periods out of four, except the period 1994-1995, meaning that the randomness is computed by using the minute-wise log-return and the performance is tested in 1995. A possible reason for this can be attributed to the sharp rise of stocks in 1995, due to the recovery of economy under Clinton administration (1993-) backed by the IT revolution. As shown in Fig.3, S&P500 index climbed rapidly from $460 to $620 from the beginning to the end of 1995, after almost two years of relatively flat period. Similar to the case of the Lehman Brothers bankruptcy in 2008 in the period 2007-2010 discussed in Ref.5, the rule did not hold. In other words, the empirical rule 'higher randomness implies better performance" has proven to hold as long as the market stays in the normal condition.

References

1. Mehta, M.: Random Matrices, 3rd edn. Academic Press, 2004.
2. Plerou, V., Gopikrishnan, P., Rosenow, B., Amaral, L. A. N. and Stanley, H. E.: Universal and Non-Universal Properties of CrossCorrelations in Financial Time Series, Physical Review Letter 83, pp.1471-1474, 1999.
3. Plerou, V., Gopikrishnan, P., Rosenow, B., Amaral, L.A.N., and Stanley, H.E.,"Random matrix approach to cross correlation in financial data", Physical Review E 65, 066126, 2002.
4. Marcenko, *V*. and Pastur, L., Distribution of eigenvalues for some sets of random matrices, Mathematics of the USSR-Sbornik **1**, pp.457-483, 1994.
5. Xin Yang, Yuuta Mikamori, Mieko Tanaka-Yamawaki: Predicting the security levels of stock investment by using the RMT-test, Procedia Computer Science,22 pp.1172-1181,2013.
6. Yang X., Itoi R., and Tanaka-Yamawaki M.: Testing Randomness by Means of RMT Formula, Intelligent Decision Technologies, SIST 10: pp.589-596, 2011.
7. Yang, X., Itoi, R. and Tanaka-Yamawaki, M.: Testing randomness by means of Random Matrix Theory, Progress of Theoretical Physics Supplement 194,pp. 73-83, 2012.
8. Tanaka-Yamawaki, M., Yang, X. and Itoi, R.: Moment Approach for quantitative evaluation of randomness based on RMT formula, Intelligent Decision Thechnologies, SIST 16, pp. 423-432, 2012.