Empirical Study on Information Arrival in Market

——Evidences from China's Markets

Fenghua Wen and Xiaoguang Yang*

Academy of Mathematics and Systems Science, Chinese Academy of Sciences

Abstract

That information arrives in the market in a random way underlies modern financial theories. In this paper, we use information mines as indicators of market information, and explore the probability distribution of information arrival by analyzing the distribution of time interval between 2 pieces of adjacent information mines. Using data from China's security markets, we find that the way of information arrival does not exhibit a normal distribution, but a fractal distribution, instead. This implies that information arrives in the security market in a biased way.

JEL Classification: F3

Keywords: EMH hypothesis, Information Arrival, Distribution Test, R/S Analysis, Jarque-Beta Test

Acknowledgement

This project was supported by Natural Science Foundation of China(No.70425004). The authors are indebted to many for helpful comments including Linjie He, Qiujun Lan, Xiaoyong Zhang, and Chaoqun Ma.

* Corresponding author: Academy of Mathematics and Systems Science, Chinese Academy of Sciences. No.55, Zhongguancun East Road, Haidian District, Beijing, P. R. China, Postcode:100080. Tel: +86 1062651380.
Email Address: xgyang@iss.ac.cn.

1. Introduction

Over the past 30 years, the efficient market hypothesis (EMH) has been the cornerstone of modern financial theories. Osberne and Fama have been the two influential researcher in the development of EMH. Osberne prosposes that the stock price follows a random walk, and investors value a stock according to their expected values or their return rate estimates; the expected value is the weighted average of the possibilities of the return rates. In Osberne's point of view of investor's rationality, investors set their subjective probabilities in a unbiased way. On that basis, Fama constructed EMH. EMH consists of three gradually weakened hypotheses: when the investors are rational, they can value the stock rationally, and thus the market is efficient; even if some of the investors are irrational, there won't be price deviations as the transactions are randomly conducted; even if investors' irrational behaviors doesn't occur randomly, they will meet with rational arbitragers, and this will keep the price reverting to the fundamental value; at last, even if irrational investors trade at prices away from the fundamental value, their wealth will gradually diminish and the investors will be edged out of the market in the end.

Fama (1970), based on the information in relation to asset pricing, further categorized three kinds of efficient markets, and argued that the price reflects all the publicly available information and the consensus of investors forms the reasonable price. Therefore, EMH implies that known information cannot be used to gain excess profit on the market[1]. According to EMH, the current price of a financial product is the result of the current information, and the current return rate is not related with past return rates. Return rates are independent from each other, and follow a random walk. When the return rates are collected, they will exhibit a normal distribution. This argument is based on the first 2 hypotheses. The first basic hypothesis of EMH is that of investors' rationality—investors have included all the information in determining the price of the security, therefore the price is independent, and moves only when new information arrives in the market. The second basic hypothesis is that information arrives in the market in a random way, since information is generated from the economic system which is a system with great degrees of freedom and a myriad of factors. This is a reasonable hypothesis in terms of the linear approach[2]. Under the 2 basic hypotheses, the normal distribution of the return rate and the random walk of the stock price are natural corollaries of EMH.

However, empirical evidences from the security market show that the return rate is not a normal distribution[3] (Fig. 1), and there are many abnormalities that the standard financial theory cannot explain. With these in background, some researchers develop theory of behavioral finance. Behavioral finance proposes that investors' bounded rationality is a part of the reasons for the abnormalities and the non-normality of the return rate distribution, thus leading to the release of the first of the three basic hypotheses. Meanwhile, as modern science develops, many researchers believe that the security market is a complex system, and the non-normality of the return rate distribution is a reflection of the complexity of the market[4]. That viewpoint is to see the security market as a complex system with random disturbance, or specifically, thinks that information arrives in the market in a random way[5].

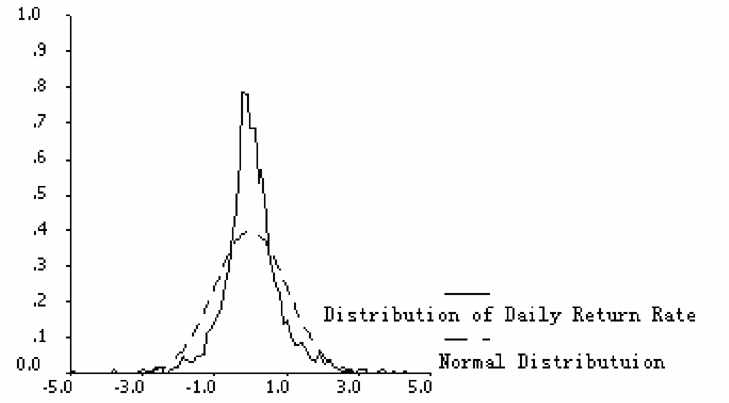

Figure 1 the Distribution of the Daily Return Rate of Shanghai Composite Index and the Normal Distribution

As to the second basic hypothesis, some researchers have long explained that if information arrives in the market cluster by cluster rather than in a linear way, the return rate distribution will be leptokurtic. Furthermore, if information arrives in the market in a biased manner, the abnormalities on the market and the non-normality of the return rate distribution can be well explained. This proposition will change the premises for the behavioral finance and the security market researches from the perspective of complex science, and have great impact on the two approaches to the research of the security market. Due to the lack of an effective measure of information arrival in the market, the research on information in the market has been confined to qualitative analysis in want of support from empirical evidences. This paper attempts to conduct an empirical research on the manner in which information arrives in the market.

It is a new adventure to explore empirically whether information arrives in the market in a random way. This paper takes the Chinese security market as an example and uses the information mine as the representative of market information. We define t as the variable of the time interval of two pieces of adjacent information arriving in the market, and transform the empirical study of whether information arrives in the market in a random walk into a test of the distribution of the new variable t.

2. Hypothesis and Data

The multitude of information in the security market makes it difficult to test the distribution of all information. It is critical to sample representative information and design a workable measure for the manner in which information arrives in the market.

In terms of the range in which information exercises influence, we can roughly categorize information as follows: common market information— macro-economic information, relevant information of policies affecting certain industries, or sectors; stock-specific information— information that only affects the return rate of an individual stock, for example, changes in the operation, adjustments of the strategy, and the financial condition of a listed company. When common market information arrives in the security market, it will have influence on the entire market, and the return rate of the market will move; while the stock-specific information arrives in the market, the influence on the involved individual stock is obvious and that on the entire market is very little.

It is a universally acknowledged fact that investors have bounded rationality. “Bounded rationality”, to a larger extent, denotes “being subjectively rational but unable to consummate the rationality in practice”. This, in fact, is consistent with the notion that people have rational consciousness, but do not have sufficient ability to behave with perfect rationality[6]. Individual investors with bounded rationality cannot browse all the newspapers and webpages to search for the relevant information. Even if they can do that, they may neglect certain important information, as their understanding of the information is subjected to a variety of factors. To meet the needs of grasping the representative information in the market and making effective decisions, security market information advisory companies that cater to security investors come into being. Information advisory companies develop and sell information mines, i.e. information sets which may affect the security market or the prices of individual securities—information affecting individual stocks are called stock mines, and those involving the entire market called market mines.

In information advisory companies, professionals spend much more time analyzing and sorting out relevant information with more rationality than individual investors. In addition, an information advisory company can only survive through the mechanism of market selection after the validity of their information mines are tested in the market. The accuracy of information mines are thus warranted. Therefore we believe that information mines are reasonable substitutes for information.

This paper takes as the research object the market information mines developed by two well recognized investment information advisory companies in China: Wanguo Securities co. Ltd., and Wind Info. The data points are the entire market information mines of Shanghai Composite Index from January 1, 1997 to October 30, 2005, since China’s securities market has a history of less than 20 years. Before 1997, it operates with poor efficiency, transparency of information or market responsiveness to information.

Now we will define a new variable *t* as the time interval between two adjacent information mines:

$$t_n = T_n - T_{n-1}$$

where T_n is the time the n*th* (in the sequence of arriving the market) information mine arrives the market. Thus, the distribution of the random variable *t* represents the distribution of information arriving in the market. If information arrives in the market in a random way, the distribution of *t* variable will be normal. Through devising the variable *t*, the empirical study of whether information arrives in the market is transformed into a test of the distribution of variable *t*.

3. Empirical Evidence

3.1 Methodology

There are a number of procedures to test whether the data is a normal distribution. In this paper, we first apply Jarque-Bera (J-B) test, since it is an effective measure to test the normality which well incorporates the information on the skewness and the kurtosis of the sample[7]. The statistic of the J-B test is:

$$L = n[\frac{\hat{S}^2}{6} + \frac{(\hat{K}-3)^2}{24}]$$

where *n* is the number of observations, $\hat{S}$ is the estimate of the overall skewness of the sample data, and $\hat{K}$ is the estimate of the overall kurtosis of the sample data. Under the assumption that the observations are independent and is a normal distribution, the J-B statistic *L* is a chi-square distribution with 2 degrees of freedom. a chi-square distribution with 2 degrees of freedom.

This paper adopts the Rescaled Range Analysis (R/S Analysis) to test the distribution of *t*, which was first developed by the eminent British hydraulic engineer H.E.Hurst in his Long-Term Storage Capacity of Reservoirs in 1951. This is a kind of robust non-parametric test.

As we know, the Brownian Motion is the primary mode of the random walk. Einstein discovered that the distance that random particles roam will increase with the square root of time in his research in the process of the Brownian Motion:

$$R = T^{0.5}$$

Where *R* is the distance and *T* is the time. This is the well known the $T^{0.5}$ rule. The $T^{0.5}$ rule applies only to the cases where the time series in the form of an ordinary Brownian motion with an average of 0 and a standard deviation of 1. Inspired by the $T^{0.5}$ rule, Hurst derived the generalized form,

$$(R/S)_n = C * nH$$

Where *R* denotes range，*S* standard deviation，*H* Hurst Exponent，*C* a constant，*n* time. The above equation is the Rescaled Range Analysis equation.

The virtue of the *R/S* Analysis rests with that it is a non-parametric approach, that is, the *R/S* Analysis does not assume in advance that the subject being researched is of a particular distribution, thus making it extensively applicable to both Gaussian distribution and non-Gaussian distributions such as *t* distribution and Gamma distribution etc.

In *R/S* Analysis, Hurst Exponent has different meanings. Generally speaking, when the Hurst Exponent of the security price time series $H = 0.5$, it means the security price does follow a Brownian Motion, i.e. for any time *t* the correlation between the increments of the past price and the current price are zero, and thus exhibiting an independent random process; when $H \neq 0.5$, it shows that the security price on the capital market follows a fractional Brownian Motion, that is to say, the increments of the past price and the current price are correlated for any time *t*; specifically, in the case that $H > 0.5$, the correlation continues, and exhibits persistence.

The *R/S* Analysis can be computed step by step as follows:

(1) Divide the sample data $\{X_i\}^N_{i=1}$ into $Int(N/W)$ independent time series with the interval *W*, and compute the averages $(EX)_w$ and the standard deviations $S(W)$ of the independent time series separately.

$$(EX)_W = \frac{1}{W}\sum_{i=1}^{w} x_i \ , \quad S(W) = \{\frac{1}{W}\sum_{i=1}^{w}[X_i - (EX)]_W\ ]^2\}^{1/2}$$

(2) Calculate the cumulative mean deviations $X(i,W)$.

$$X(i,W) = \sum_{t=1}^{i}[X_t - (EX)_W] = \sum_{t=1}^{i} X_t - i * (EX)_W \ , \quad (1 \le i \le W)$$

(3) Work out the range $R(W)$.

$$R(W) = \max_{1 \le i \le W} X(i,W) - \min_{1 \le i \le W} X(i,W)$$

(4) Calculate $R(W)/S(W)$ and take the average $[E(R/S)]_w$ of the $Int(N/W)$ *R/S* values as the *R/S* estimate of the time series with the interval of *W* weeks.

(5) In the paired log coordinate system, utilize the points $(\log W, \log[E(R/S)]_w)$ to calculate the slope *H* of the best fitting line with the least squares procedure:

$\log[E(R/S)_w] = H * \log(W) + \log c$, where *c* is a constant.

3.3 Empirical Evidence

3.3.1 Empirical Evidence (A)

From January 1, 1997 to October 30, 2005, the entire market information mine technology of Wanguo Securities Co. Ltd. warned 216 information mines, and therefore provided 215 *t* indexes. From January 4, 2001 to October 30, 2005, Wind Info reported 530 information mines for the entire market, and thus derived 529 *t* indexes. We tested the normality of the 2 sets of sample data, and obtained the results as shown in Figures 2 and 3.

On the basis of the above statistics, it can be found that *t* is not normally distributed, i.e., the research findings support that information does not arrive in the securities market in a random manner, which is contradictory to the assumption of EMH. Therefore, we will perform a further *R/S* Analysis of the sample data below.

3.3.2 Empirical Evidence (B)

First, we conducted a *R/S* Analysis of the weekly return rate of Shanghai Composite Index from January 1, 1997 to October 10, 2005. The result is shown in Figures 4 and 5.

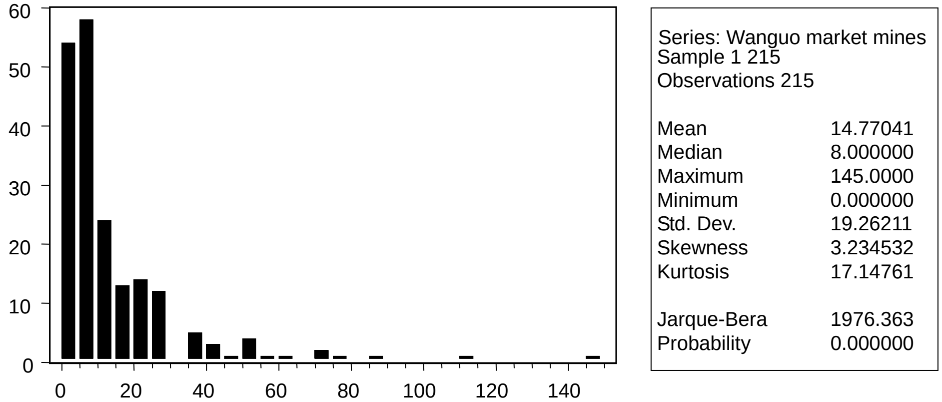

Figure 2 The Distribution Characteristics of Information Mines' Index *t* of Wanguo

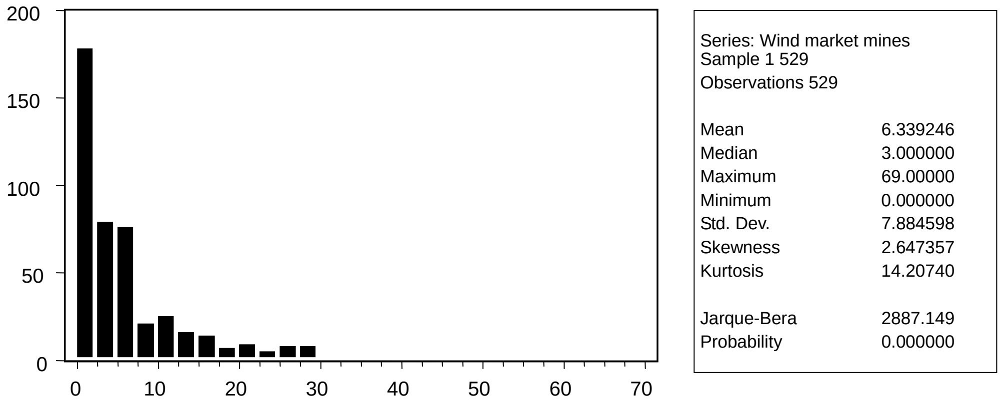

Figure 3 The Distribution Characteristics of Information Mines' Index *t* of Wind

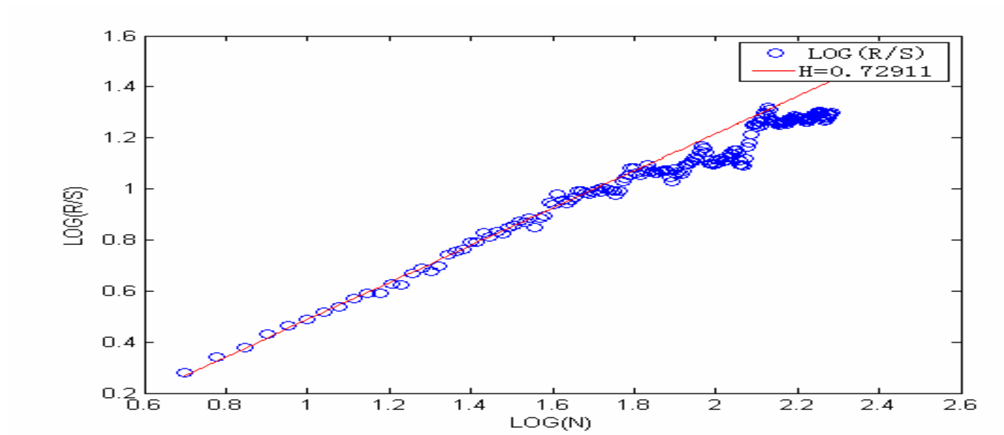

Figure 4 *R/S* Empirical Result of the Shanghai Composite Index

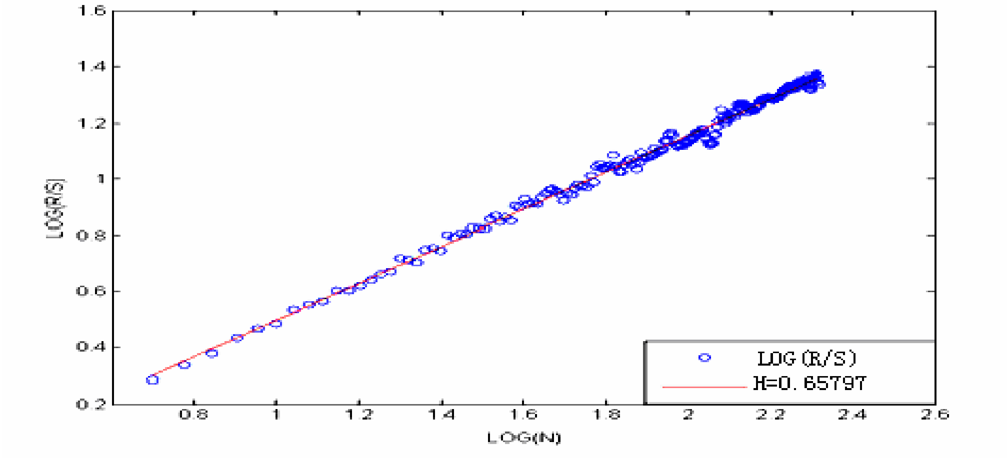

Figure 5 ***R/S*** **Empirical Result of the Shenzhen Component Index**

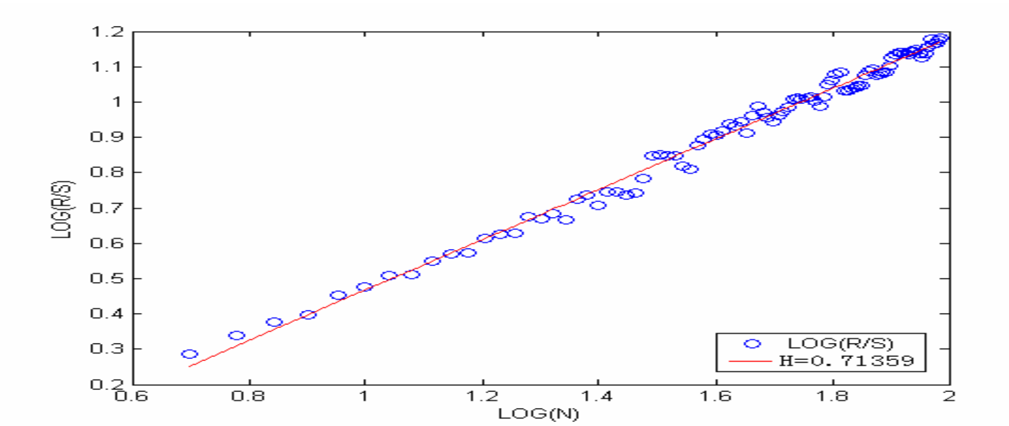

Figure 6 ***R/S*** **Empirical Result of Wind Info Market Information Mine**

From Figures 4 and 5, it can be observed that the Hurst Exponents of the Return Rate of Shanghai Composite Index and those of the Return Rate of Shenzhen Component Index are 0.72911 and 0.65797 respectively. Meanwhile, *R/S* analyses were also conducted of the 2 sets of *t* indexes for the entire market information mines. The results are graphed in Figures 6 and 7.

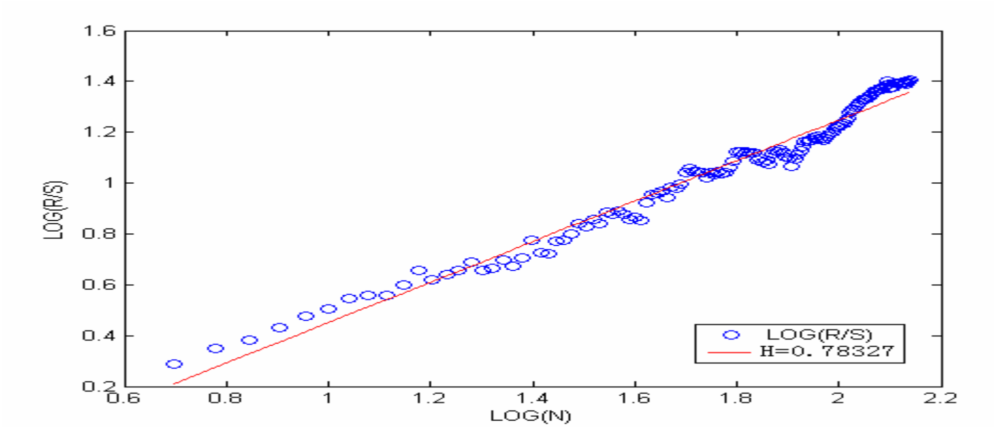

Figure 7 ***R/S*** **Empirical Result of Dazhihui Market Information Mine**

According to the empirical results derived from the *R/S* analyses of figures 6 and 7, the Hurst Exponents of the Return Rate of Shanghai Composite Index and those of Shenzhen Component

Index are respectively 0.71359 and 0.78327, which indicates that *t* accords with the fractal distribution and the information mine arrives in the market in a biased manner.

Furthermore, compared with the results from Figure 5, the Hurst Exponents of the return rate of the market and those of the information distribution are both greater than 0.5. It is of comparatively higher consistence, which implies that, to a certain extent, the distribution characteristics of the return rate of the securities market and the biased information arrival are positively correlative. A problem deserved to be explored further is that whether the Hurst Exponents of the information distribution are higher than those of the return rate of the market. It implies that there is a phenomenon of excess kurtosis in the distribution of information arriving in the market represented by information mines, or that the investors react to the information mines inadequately.

4. Conclusion

Based on a basic hypothesis that information arrives in the securities market in a random manner, this research defines a new variable *t* as the interval two adjacent pieces of information mines arrive in the securities market, and then transforms the empirical research whether information arrives in the securities market in a random manner into a normality test of *t*. Utilizing Jarque-Bera test and *R/S* Analysis of the samples, we find that *t* is not a normal distribution, i.e., the interval in which information arrives in the securities market is of a fractal distribution. This finding shows that information arrives in the securities market in a biased manner.

The securities market information is quite complex, which makes it a challenge to study the ways of the information's arriving in the market. There's much to be researched, for example, looking for a better market information representative index than the information mine so as to test the distribution of the market information arrival in a better way; constructing stock return rate distribution models in which the distribution of information (including the byte size of the information) is incorporated; perfecting the theory of behavioral finance under the assumption that the information arrives in the market in a biased manner, etc. It is wished hopefully that this research would bring in more fruitful and profound development of the research of this kind.

Reference

[1] Fama, E. F. (1970), Efficient Capital Markets: A Review of Empirical Work[J], Journal of Finance, Vol.25:383-417

[2] Zvi Bodie,Alex Kane,Alan J.Marcus(1999),Investment[M],McGraw-Hill,Inc,4th:280

[3] Shiller, R. J.(1981)，Do Stock Prices Move Too Much to be Justified by Subsequent Changes in Dividends?[J]，American Economic Review, Vol.71,No.3:421-436

[4] Jin Wulun, Ge Yuanlin（2004），Science of Complexity and Its Evolution[J]，Complex System and Science of Complexity，Vol.1,No.1：1－5. (Chinese)

[5] Wu Haihua（2002），On Complexity of the Securities Market：A Case Study of Shanghai Securities Market [J]，Journal of System Simulation，Vol.14 No11：1486－1490. (Chinese)

[6] Wang Dingding, Meditations Reaped From Between Economics and Philosophy [M], Beijing:

China Social Sciences Press, 1996, P:76. (Chinese)
[7] John Y. Campbell，Andrew W. Lo, A. Craig Mackinlay，(Zhu Pingfang,et al Translated)，The Econometrics of Financial Markets[M]，the Press of Shanghai University of Finance and Economics，2003，p:25. (Chinese)