การศึกษาเชิงประจักษ์ของประสิทธิภาพของตลาดหุ้นไทยภายใต้ช่วงความเชื่อมั่นของดัชนีเฮิร์สทโดยวิธีดีเอฟเอ

นางสาวสิรภัทร สุกใส

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต
สาขาวิชาคณิตศาสตร์ประยุกต์และวิทยาการคณนา
ภาควิชาคณิตศาสตร์และวิทยาการคอมพิวเตอร์
คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2559

EMPIRICAL STUDY ON EFFICIENCY OF THAILAND STOCK MARKET BASED ON CONFIDENCE INTERVAL OF HURST INDEX USING DFA METHOD

Miss Sirapat Suksai

A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Science Program in Applied Mathematics and Computational Science
Department of Mathematics and Computer Science
Faculty of Science
Chulalongkorn University
Academic Year 2016

Thesis Title	EMPIRICAL STUDY ON EFFICIENCY OF THAILAND STOCK MARKET BASED ON CONFIDENCE INTERVAL OF HURST INDEX USING DFA METHOD
By	Miss Sirapat Suksai
Field of Study	Applied Mathematics and Computational Science
Thesis Advisor	Assistant Professor Khamron Mekchay, Ph.D.

Accepted by the Faculty of Science, Chulalongkorn University in Partial Fulfillment of the Requirements for the Master's Degree

. . . . Dean of the Faculty of Science
(Associate Professor Polkit Sangvanich, Ph.D.)

THESIS COMMITTEE

. . . . Chairman
(Assistant Professor Krung Sinapiromsaran, Ph.D)

. . . . Thesis Advisor
(Assistant Professor Khamron Mekchay, Ph.D.)

. . . . Examiner
(Jiraphan Suntornchost, Ph.D.)

. . . . External Examiner
(Thidaporn Supapakorn, Ph.D.)

สิรภัทร สุกใส : การศึกษาเชิงประจักษ์ของประสิทธิภาพของตลาดหุ้นไทยภายใต้ช่วงความเชื่อมั่นของดัชนีเฮิร์สทโดยวิธีดีเอฟเอ. (EMPIRICAL STUDY ON EFFICIENCY OF THAILAND STOCK MARKET BASED ON CONFIDENCE INTERVAL OF HURST INDEX USING DFA METHOD) อ.ที่ปรึกษาวิทยานิพนธ์หลัก : ผศ.ดร. คำรณ เมฆฉาย, 54 หน้า.

ในงานวิจัยนี้สนใจที่จะศึกษาความมีประสิทธิภาพของตลาดหลักทรัพย์ในประเทศไทยทั้งตลาดหลักทรัพย์แห่งประเทศไทย (SET) และตลาดหลักทรัพย์เอ็มเอไอ (MAI) โดยศึกษาผ่านดัชนีเฮิร์สทที่ประมาณค่าด้วยวิธี Detrended Fluctuation Analysis (DFA) การศึกษาครั้งนี้เป็นการขยายงานวิจัยของ เจษฎา สุขพิทักษ์ และ วรากร เฮ้งปัญญา (2016) โดยได้สร้างช่วงความเชื่อมั่นของค่าดัชนีเฮิร์สทสำหรับตลาดที่มีประสิทธิภาพ ซึ่งสร้างโดยวิธีมอนติคาร์โลและวิธีการทางสถิติ ภายใต้สมมุติฐานที่ว่าราคาสินทรัพย์ในตลาดมีประสิทธิภาพสามารถจำลองได้ด้วยการเคลื่อนที่แบบบราวน์ เพื่อนำมาเป็นตัวชี้วัดในการอธิบายความมีประสิทธิภาพของตลาดหลักทรัพย์ในประเทศไทย ผลการศึกษาพบว่าพฤติกรรมของดัชนีเฮิร์สทในการเลื่อนหน้าต่างของเวลาที่แตกต่างกันนั้นสามารถอธิบายความมีประสิทธิภาพของตลาดหลักทรัพย์ของประเทศไทยได้สอดคล้องกันในทุกระดับความเชื่อมั่น ในการศึกษานี้พบว่าทั้งตลาดหลักทรัพย์แห่งประเทศไทยและตลาดหลักทรัพย์เอ็มเอไอมีบางช่วงเวลาที่พบความไม่มีประสิทธิภาพ แต่พบว่าตลาดอาจจะมีประสิทธิภาพในช่วงหลัง เพื่อที่จะให้ข้อสรุปเกี่ยวกับความมีประสิทธิภาพของประเทศไทยถูกต้องมากยิ่งขึ้น ควรใช้เครื่องมือและวิธีการอื่นเข้ามาช่วยสรุปเพิ่มเติม

ภาควิชา	คณิตศาสตร์และ	ลายมือชื่อนิสิต	....
	วิทยาการคอมพิวเตอร์	ลายมือชื่อ อ.ที่ปรึกษาหลัก	....
สาขาวิชา	คณิตศาสตร์ประยุกต์		
	และวิทยาการคณนา		
ปีการศึกษา	2559		

5872071523 : MAJOR APPLIED MATHEMATICS AND COMPUTATIONAL SCIENCE, KEYWORDS : THAILAND STOCK MARKET / EFFICIENCY OF STOCK MARKET / DETRENDED FLUCTUATION ANALYSIS / CONFIDENCE INTERVAL

SIRAPAT SUKSAI : EMPIRICAL STUDY ON EFFICIENCY OF THAILAND STOCK MARKET BASED ON CONFIDENCE INTERVAL OF HURST INDEX USING DFA METHOD. ADVISOR : ASSISTANT PROFESSOR KHAMRON MEKCHAY, Ph.D., 54 pp.

We are interested in investigating the efficiency of Thailand stock markets (SET and MAI markets) by using the Hurst index based on the Detrended Fluctuation Analysis (DFA). This study is an extension of the work of Sukpitak and Hengpunya (2016) by using Monte Carlo simulation and statistical analysis to construct a confidence interval of Hurst index for the efficient market based on the assumption that the sample of Brownian motion represents assets prices in efficient markets, then applying the interval as an indicator for the efficiency of Thailand stock markets. We found that there is a consistency in the behavior of Hurst index among different time windows sizes based on a constructed confidence interval. Based on this study, the result shows that both SET and MAI markets have some inefficient periods and we fail to reject that the markets are efficient in the recent years. To get more accurate conclusion on the efficiency of Thailand markets, other tools or techniques are required.

Department	: Mathematics and Computer Science	Student's Signature	....
Field of Study	: Applied Mathematics and Computational Science	Advisor's Signature	....
Academic Year	: 2016		

ACKNOWLEDGEMENTS

First, I would like to sincerely thank my advisor Assistant Professor Dr. Khamron Mekchay, for his continuous support, guidance, motivation along with useful comments and encouragement.

I would like to thank Assistant Professor Dr. Krung Sinapiromsaran, Dr. Jiraphan Suntornchost, and Dr. Thidaporn Supapakorn, for being my thesis committee and giving useful suggestions and comments to improve this thesis.

I appreciate the His Royal Highness Crown Prince Maha Vajiralongkorn Scholarship from the Graduate school, Chulalongkorn University to commemorate the 72^{nd} Anniversary of his Majesty King Bhumibol Adulyadej for supporting all tuition fee during my master degree program.

I thank all AMCS students and staffs for all their sincerely help and all valuable time we shared together.

Finally, I am grateful to my family for their support and always being by my side throughout writing this thesis and all my life.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

There is a cornerstone assumption in the financial market which is efficiency market hypothesis (EMH) (Fama, 1970) stating that asset price totally reflects all available information such as the past prices or the news about that stocks. EMH is developed under the assumption that stocks are always traded at their fair values because everyone knows all information about stocks, therefore, it is impossible to buy the stock at a lower price or sell at a higher price than the market fair price. Moreover, the future price depends on the random future information that no one knows, thus, we cannot consistently earn an excess profit by predicting the movement of the future prices using historical price. Hence, the technical analysis based on the historical price will not work in the efficient market. Mathematically, EMH implies that the returns or the increment of stocks prices should follow the process that has independent and uncorrelated increments, which agree with the properties of the Brownian motion. Hence, one can use the Brownian motion, which is the Gaussian process that has independent and uncorrelated increments, to describe some behavior of the asset price in the efficient market in order to estimate or predict the behavior of the prices, which will be useful in financial investment.

There is a long discussion in EMH. Some research (Lo, 1991) found that in some stock markets, share prices may have the different property from the assumption of EMH such as long-range dependence (LRD) (Beran, 1994), where the current prices depend on prices in the past. In this case, the market price cannot be described by the Brownian motion, then one required the fractional Brownian motion (Mandelbrot and Ness, 1968) instead. In literature, one often refers to the parameter H, known as the Hurst index, in order to investigate the efficiency of the markets; theoretically, if H is close or equal to 0.5, one often say that the market is efficient, while H is quite different from 0.5, one expects the existence of long-range dependence in the market (inefficient).

Capturing the behavior of the Hurst index is one way to investigate the efficiency of the stock market. There are many methods to estimate the Hurst index from the data series. The detrended fluctuation analysis (DFA) is one of the well-known methods that have been used in investigating the long memory behavior of data in various fields such as DNA sequences (Peng et al., 1994a), heartbeat signals (Peng et al., 1994b), traffic data (Shang et al., 2008), and financial time series (Cajueiro and Tabak, 2004; Costa and Vasconcelos, 2003; Sukpitak and Hengpunya, 2016). The idea of the DFA method is constructed based on the power-law scaling of the fluctuation functions in order to obtain the Hurst index.

Study the efficiency of Thailand stock markets is one interesting topic in general so that one can model financial assets using appropriate models. For example, if one knows that the market is efficient, then one will consider a model that relies on the Brownian motion. Otherwise, if the market is not efficient, one would consider other processes such as the fractional Brownian motion. There is few investigation of the efficiency of Thailand stock markets in the past. Cajueiro and Tabak (2004) used R/S method to investigate the efficiency of emerging market including Thailand SET market for the period of 1992-2002 and found no conclusion about the efficiency of Thailand market in general, but they found that the efficiency of all emerging markets is increasing over time. Sukpitak and Hengpunya (2016) had investigated the efficiency of Thailand stock market by using the DFA method to obtain the Hurst index for the Thailand SET and MAI indices as numbers to check for efficiency. They showed that the SET market is efficient in recent years after 2012 by implying from DFA result that the Hurst indices look close to 0.5 in that period, while the MAI market shows inefficiency at that time (2002-2015).

In this thesis, we are interested in investigating the efficiency of Thailand stock markets from both SET and MAI markets by extending the work of Sukpitak and Hengpunya (2016). Instead of obtaining the Hurst index for the data series to check for efficiency, we propose the investigation based on the confidence interval of Hurst index for the efficient market, as suggested in general by Weron (2002). This concept is more important for investigation of the efficiency of a market than using the single Hurst index that represents the market efficiency because the Hurst index usually varies in time, and it is not clear how close to 0.5 it should be for concluding that the market is efficient. The confidence intervals of estimated Hurst index for the efficient market are produced based on the assumption that the sample of Brownian motion represents assets prices in efficient markets, which are constructed using Monte Carlo simulation to simulate data and estimated its Hurst index, then applying statistics to obtain properties, and finally using data fitting techniques to get a suitable function for the confidence interval.

The rest of the thesis is organized as follows. The next chapter describes the background knowledge of EMH, fractional Brownian motion (FBM), the DFA and other background knowledge that are used in this research. Chapter 3 provides all required steps used in this thesis for estimating the time-varying Hurst index and constructing the empirical confidence interval. The empirical results and analysis of the constructed empirical confidence interval, the efficiency of Thailand stock market and the comparison between the efficiency of the SET market and the MAI market are presented in Chapter 4. Finally, Chapter 5 provides the comments and conclusion of the result. The R code for the work is shown in the appendix.

CHAPTER II

BACKGROUND KNOWLEDGE

2.1 Efficient Market Hypothesis (EMH)

The efficient market hypothesis (EMH) is introduced by Fama (1970) stating that the asset price totally reflects its relevant information. EMH is developed under the assumption that stocks are always traded at their fair values, because everyone knows all information about stocks. Therefore, it is impossible to buy the stock at the lower price or sell at the higher price than the market fair price. Moreover, the future price depends on the random future information that no one knows, thus, we cannot gain the excess profit by predicting the movement of the future prices.

This knowledge of EMH is classified as the "weak-form EMH", stating that the assets prices reflect their past prices. Thus, the investment strategies that based on the historical prices do not work in the case of that market is weak-form efficient.

We say that the market is weak-form efficient (efficient, for short) when the asset price reflects the information about the past prices, and the changes of the price (returns) are independent in this hypothesis. Thus, one cannot predict the movement of the future prices using the historical price. Hence, technical analysis and other strategies based on the historical information will not be able to make an abnormal profit in the weak-form efficient market.

Samuelson (1973) (Samuelson, 1973) suggested that the stock prices follow a martingale that is the expectation of the stock price at time $t+1$, y_{t+1}, containing all available information is equal to the stock price at time t

$$\mathbf{E}(y_{t+1}|\Phi_t) = y_t,$$

where Φ_t is the information available at time t which includes all past and the present price of that stock $\ldots, y_{t-2}, y_{t-1}, y_t$.

He defined the return or price change as $r_t = y_{t+1} - y_t$ for consecutive times t and $t+1$. Then

$$\mathbf{E}(r_t|\Phi_t) = \mathbf{E}(y_{t+1} - y_t|\Phi_t) = 0,$$

that is the expected return based on information at time t or the average of market profits and losses is zero and he then proved that the return is uncorrelated.

In conclusion, according to the EMH and Samuelson's proof, stock prices follow a martingale which is described in the next section and its returns are uncorrelated and independent. Then the process that can use for representing a stock price in the efficient market should be a martingale and have independent and uncorrelated increments.

2.2 Brownian Motion

Definition 2.1 (Brownian motion)**.** (see Olofsson and Andersson (2012), for example) Brownian motion $B(t)$ is a stochastic process in real time $t \geq 0$ with the following properties;

1. (Independent increments) For all $t_0 < t_1 < \ldots < t_m$, the increments

$$B(t_1) - B(t_0), B(t_2) - B(t_1), \ldots, B(t_m) - B(t_{m-1})$$

are independent.

2. (Stationary increments) $B(t) - B(s)$ follows a normal distribution with zero mean and variance $t - s$ where $t \geq s$.

3. (Continuous paths) $B(t), t \geq 0$ is a continuous function of t.

4. $B(0) = 0$.

The followings are some basic properties of the Brownian motion

1. From the Definition 2.1 (2), the increment $B(t) - B(0)$ follows a normal distribution then

$$B(t) = B(t) - B(0) \sim N(0, t),$$

$B(t)$ has a normal distribution with mean $\mathbf{E}[B(t)] = 0$ and variance $\mathbf{Var}[B(t)] = t$.

2. Covariance

For the covariance function, suppose $t \geq s$. Since $B(t)$ and $B(t) - B(s)$ are independent follow from the assumption of independent increment, we have

$$\begin{aligned}
\mathbf{Cov}(B(t), B(s)) &= \mathbf{Cov}(B(s) + (B(t) - B(s)), B(s)) \\
&= \mathbf{Var}[B(s)] + \mathbf{Cov}(B(s), B(t) - B(s)) \\
&= s + 0 \\
&= s.
\end{aligned}$$

3. Correlation

$$\mathbf{Corr}(B(s), B(t)) = \frac{\mathbf{Cov}(B(s), B(t))}{\sqrt{\mathbf{Var}(B(s))}\sqrt{\mathbf{Var}(B(t))}} = \frac{s}{\sqrt{s}\sqrt{t}} \qquad , where \quad t \geq s.$$

Definition 2.2 (Martingale)**.** (see Olofsson and Andersson (2012), for example) A stochastic process $X_t, t \geq 0$ is a martingale if for any t, $\mathbf{E}[X_t]$ exists, and for any $s > 0$

$$\mathbf{E}[X_{t+s}|\Phi_t] = X_t,$$

where Φ_t is the information about the process available at time t.

Theorem 2.3**.** *Let $B(t)$ be a Brownian motion. Then $B(t)$ is a martingale.*

Proof. Let Φ_t is the information about the process B_t. By Definition 2.1 (2), $B(t) \sim N(0, t)$, then $\mathbf{E}[B(t)] = 0$. Then, for $t \geq s$,

$$\begin{aligned}\mathbf{E}[B(t)|\Phi_s] &= \mathbf{E}[B(s) + (B(t) - B(s))|\Phi_s] \\ &= \mathbf{E}[B(s)|\Phi_s] + \mathbf{E}[B(t) - B(s)|\Phi_s] \\ &= B(s) + \mathbf{E}[B(t) - B(s)],\end{aligned}$$

Since $t \geq s$, then $B(t) - B(s)$ follows the assumption of stationary increment which implies that $\mathbf{E}[B(t) - B(s)] = 0$. Hence $\mathbf{E}[B(t)|\Phi_s] = B(s) + \mathbf{E}[B(t) - B(s)] = B(s)$. So, $B(t)$ is a martingale. □

By weak-form EMH, one can use the Brownian motion to represent the movement of stock prices in the efficient market since it has properties mentioned in Section 2.1; i.e. the martingale and the uncorrelated and independent increments properties.

2.3 Hurst Index

In the case that the market is different from the efficient market, i.e. the inefficient market, there are some studies found that the returns of stock prices of some markets exhibit the behavior called "long-range dependence" or "long memory". If there exists the long-range persistence in the returns of the financial assets, it means that the return is not random likes we assumed from the market efficiency hypothesis. The existence of long-range dependence not only found in the financial market but also found in many natural phenomena as well.

In 1951, Hurst (Hurst, 1956) observed 800 years of record of Nile river. He found that the flow of Nile river is not random but have patterned. He describes this phenomenon with the constant which later called the Hurst exponent or the Hurst index, H follows his name. In 1968, Mandelbrot (Mandelbrot and Ness, 1968) proposed the method that could capture this phenomenon in financial data which is R/S method.

Generally, the parameter H represents 3 cases of the long-memory property of the increments of time series;

1. If $H \in (0, 0.5)$, the increments of the series are negatively correlated and the series exhibits antipersistent.
2. If $H = 0.5$, the increments are uncorrelated and the process corresponds to the Brownian motion.
3. If $H \in (0.5, 1)$, the increments of series are positively correlated and the series is said to have long-memory or long range dependence.

2.4 Long-range Dependence

Definition 2.4. (Beran, 1994) A stationary process X_t for time $t \geq 0$ is said to have the long-range dependence, if its autocorrelation function $\rho(k) = \dfrac{\mathbf{Cov}(X_t, X_{t+k})}{\mathbf{Var}(X_t)}$ where k is the lag or time difference at any time t, holds:

$$\sum_{k=-\infty}^{\infty} \rho(k) = \infty. \tag{2.1}$$

That is, the autocorrelations decay slowly to zero that their sum does not converge.

Beran (1994) showed that if the autocorrelation $\rho(k)$ is approximately equal to

$$c|k|^{-\alpha}, \tag{2.2}$$

with a constant c and a parameter $\alpha \in (0, 1)$ then Equation (2.1) holds.

2.5 Fractional Brownian Motion

To describe behavior of a process with the long-range dependence, we introduce a stochastic process generalized from the Brownian motion that can exhibit the long-range dependence, namely the fractional Brownian motion (FBM) (Mandelbrot and Ness, 1968).

Definition 2.5. Let $B_H(t)$ for $t \geq 0$ be a stochastic process with continuous sample paths and such that

1. $B_H(t)$ is Gaussian.
2. $B_H(0) = 0$.

3. $\mathbf{E}[B_H(t) - B_H(s)] = 0$ where $t \geq s$.

4. $\mathbf{Cov}[B_H(t), B_H(s)] = \frac{1}{2}\left(t^{2H} + s^{2H} - (t-s)^{2H}\right)$ where $t \geq s$.

for any $H \in (0,1)$, called the Hurst index. Then $B_H(t)$ is called the fractional Brownian motion (FBM).

The difference between FBM and Brownian motion is the increments $B_H(t) - B_H(s)$ of FBM may be dependent. For a special case of FBM when $H = 0.5$, FBM is the same as the Brownian motion.

We called the increment $B_H(t) - B_H(s)$ of FBM as the fractional Gaussian noise (FGN)

The covariance at lag k of FGN is defined as follows

$$\begin{aligned}\gamma(k) &= \mathbf{Cov}[B_H(t) - B_H(t-1), B_H(t+k) - B_H(t+k-1)] \\ &= \frac{1}{2}\left(|k+1|^{2H} - 2|k|^{2H} + |k-1|^{2H}\right).\end{aligned}$$

Beran (1994) shows that $\rho(k)$ is approximately equal to $H(2H-1)k^{2H-2}$ for $k \to \infty$. Then for $H > 0.5$, the autocorrelation function $\rho(k)$ follows the property of long-range dependence in Equation (2.1) with $\alpha = 2 - 2H \in (0,1)$. Hence, FGN has the long-range dependence when $H > 0.5$ and other cases as described in Section 2.3.

2.6 Detrended Fluctuation Analysis (DFA)

The DFA method (Peng et al., 1994a) is well-known and often used in general for obtaining the Hurst index of data series. The idea of the DFA method is to subtract linear trends from the time series and then calculate the fluctuation of the detrended data. Then we can investigate the long-range correlation in the data using the power-law relationship (Taqqu et al., 1995) between fluctuation functions and its time lag to produces the power-law scale, which is used to estimate the Hurst index.

We apply the DFA method with the returns of the stock with the following steps:

1. Calculating the returns of the stock when the return r_t at time t is given as the difference of logarithm of the consecutive daily closed prices y_t, (y_t represents daily closed price at time t)

$$r_t = \ln(y_{t+1}) - \ln(y_t) = \ln\left(\frac{y_{t+1}}{y_t}\right), \qquad \text{for} \quad t = 1, 2, \ldots, T,$$

2. Cumulating the difference from overall mean of the original time series $\{r_t\}_{t=1,2,\ldots,T}$ to obtain a new time series,

$$X(t) := \sum_{i=1}^{t}(r_i - \bar{r}) \qquad \text{for} \quad t = 1, 2, \ldots, T$$

, where $\bar{r} = \frac{1}{T}\sum_{i=1}^{T} r_i$.

3. Dividing $\{X(t)\}$ into $N = \left\lfloor \frac{T}{\tau} \right\rfloor$ non-overlapping subseries I_n, for $n = 0, 1, \ldots, N-1$, of length τ where τ is the positive integer.

4. Finding the local linear trend $Y_\tau(t)$ that fits the data in each subseries I_n for $n = 0, 1, \ldots, N-1$.

5. Computing the fluctuation function $F(\tau)$ defined as the root mean square

error of $X(t)$ with respect to the linear trend function $Y_\tau(t)$,

$$F(\tau) = \sqrt{\frac{1}{T}\sum_{t=1}^{T}[X(t) - Y_\tau(t)]^2}.$$

6. Repeating all previous steps for different values of τ to get the relationship between $F(\tau)$ and τ. In this work, we use τ from 5 to $\left\lfloor \frac{N}{5} \right\rfloor$ as recommended by Peng et al. (1994a) and Sukpitak and Hengpunya (2016).

7. Since $F(\tau)$ increases with the size τ, the relationship between $F(\tau)$ and τ follows the power-law with the power law scale H, the Hurst index. The value of the Hurst index is estimated from the slope of the log-log plot between $F(\tau)$ and τ.

2.7 Normal Distribution

In this section, we provide the knowledge about the normal distribution which related to the Brownian motion.

Definition 2.6. (Olofsson and Andersson, 2012) If a random variable X has a probability density function (pdf)

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}e^{-(x-\mu)^2/2\sigma^2} \qquad , x \in \mathbb{R} \tag{2.3}$$

It is said to have the normal or Gaussian distribution with parameters μ and σ^2, written as $X \sim N(\mu, \sigma^2)$ which has mean $\mathbf{E}[X] = \mu$ and variance $\mathbf{Var}[X] = \sigma^2$.

Theorem 2.7. *(Olofsson and Andersson, 2012) Suppose that $X \sim N(\mu, \sigma^2)$ and let*

$$Z = \frac{X - \mu}{\sigma}. \tag{2.4}$$

Then $Z \sim N(0,1)$ *which is called the standard normal distribution.*

If the random variable X is normally distributed, then the probability that X assumes any value in the interval $a < X < b$ is

$$P(a < X < b) = \frac{1}{\sigma\sqrt{2\pi}} \int_a^b e^{-(x-\mu)^2/2\sigma^2} dx. \tag{2.5}$$

The Equation (2.5) can be evaluated by the integral

$$P(Z \leq z) = \Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-t^2/2} dt$$

and using (2.4) to substitution and then get

$$P(a < X < b) = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right).$$

The curve of probability density function of the normal distribution is called the bell-shaped curve. It is symmetric with respect to the center $x = \mu$.

2.8 Test for Normality

2.8.1 Skewness

The skewness is a statistical measure of the asymmetry of the probability distribution of a random variable about its mean. In other words, skewness tells the amount and direction of departure from the horizontal symmetry. The skewness of a pdf can be measured in terms of its third moment about the mean follows

$$\frac{\mathbf{E}[(X-\mu)^3]}{\sigma^3}.$$

The skewness of the normal distribution which is symmetric is equal to 0 (Larsen and Marx, 2001). The sample skewness can be calculated as follows

$$\frac{\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^3}{(\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2)^{3/2}},$$

where x_i is the data from the sample and x̄ is the sample mean.

2.8.2 **Kurtosis**

The kurtosis is a statistical measure of the height of the peak of the probability distribution of a random variable. The kurtosis of a pdf can be measured in terms of its fourth moment about the mean follows

$$\frac{\mathbf{E}[(X-\mu)^4]}{\sigma^4}.$$

The kurtosis of the normal distribution is equal to 3 (Larsen and Marx, 2001). The sample skewness can be calculated as follows

$$\frac{\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^4}{(\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2)^{2}},$$

where x_i is the data from the sample and x̄ is the sample mean.

2.8.3 **Histogram Plots**

The histogram is a graphical representation of the data. The histogram is obtained by dividing the range of the data into equal-sized bins, then count the number of points from the data in each range and plot the quantity into a graph. The shape of the histogram is similar to the shape of the pdf of the distribution. If the histogram of the sample data indicates a symmetric and no skew which resembles the bell shape of the pdf of the normal distribution, then the normal distribution might be a model of the data.

2.8.4 Normal Q-Q Plots

The normal quantile-quantile (Q-Q) plots are plots of the empirical quantile from the sample against the theoretical quantile from a standard normal distribution. The normal Q-Q plot is another graphical representation of how well the data are modeled by a normal distribution. The normal Q-Q plot from the normal distributed data point should lies approximately on a straight line.

2.8.5 Kolmogorov-Smirnov Test

The Kolmogorov-Smirnov (K-S) (Massey, 1951) test is used for testing the assumption that the sample data are from a reference-distributed population. The null hypothesis H_0 that the data come from a certain distribution function $F(x)$ and the alternative hypothesis H_1 is that a certain function F_x is not the distribution of a population. Hence, if the p-value from the test is less than 0.05 then the null hypothesis is rejected which means rejecting the assumption.

Let $\bar{F}(x)$ be the corresponding sample distribution of an approximation of $F(x)$. The Kolmogorov-Smirnov test statistic, D_n, is defined by

$$D_n = \sup_x |\bar{F}(x) - F(x)|.$$

If the distribution of the sample is similar with the distribution $F(x)$, then D_n is closed to zero.

2.9 Confidence Intervals

The usual way to quantify the amount of uncertainty in an estimator is to construct a confidence interval defined as the point estimator plus the margin of error with a specified confidence level. The confidence intervals are ranges of numbers that have a high probability of containing the unknown parameter within the ranges.

Random intervals can be constructed based on a confidence level. In practice, the significance level α is usually chosen from 0.10, 0.05, and 0.01 which gives the corresponding confidence intervals $100(1-\alpha)\%$ which are 90%, 95% and 99%, when α is equal to 0.10, 0.05 and 0.01, respectively.

From the property of the standard normal distribution, the confidence interval is constructed based on $P(Z \geq z_{\alpha/2}) = \alpha/2$. For example, $z_{\alpha/2} = z_{0.025} = 1.96$. A $100(1-\alpha)\%$ confidence interval for standard normal distribution is the range of numbers

$$(\mu - z_{\alpha/2}\sigma, \mu + z_{\alpha/2}\sigma).$$

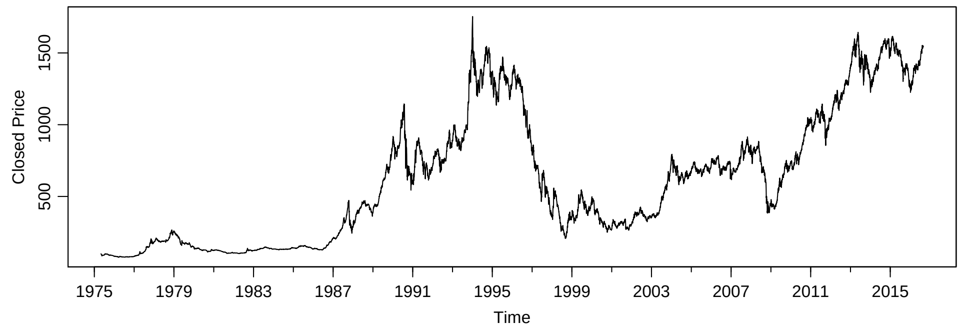
2.10 Data of Thailand Stock Markets

The stock index is a statistical indicator representing the performance of the given stock market. Investors use the stock index as a benchmark for tracking the changes in a stock market, thus studying behavior of the stock index is useful in helping investors to see overall behavior of the market. In this thesis, we study on Thailand stock market established in 1975 through the index of the largest and well known market : Stock Exchange of Thailand (SET) index which is calculated from all stocks in the market that has big market capitalization and market of alternative investment (MAI) index which is the market for small and medium businesses that started from 2002.

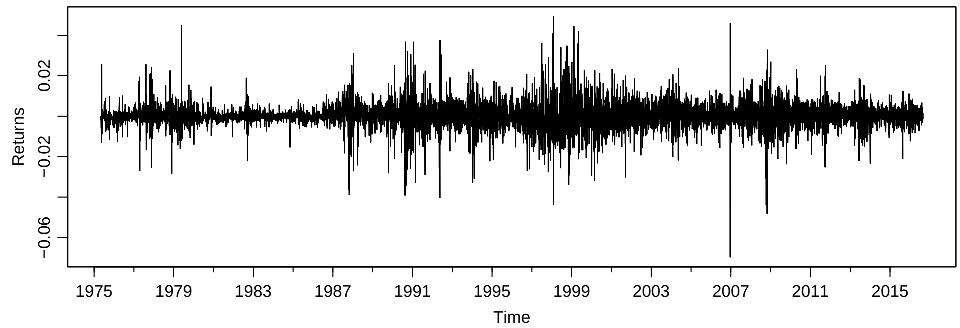
The data is provided by `http://siamchart.com/stock/` (Kurayami Corp., 2016) in the period from 2 May 1975 to 29 August 2016 for the SET index and from 3 September 2002 to 29 August 2016 for the MAI index. The closed prices and the returns of the SET and MAI markets are shown in Figures 2.1 and 2.2, respectively. The other market information is in Table 2.1.

Index	Number of listed companies	Market capitalization	Average trading volume
SET	522	15,079,272.11	10,943.80
MAI	134	425,364.18	1,056.38

Table 2.1: Market information at 31 December 2016

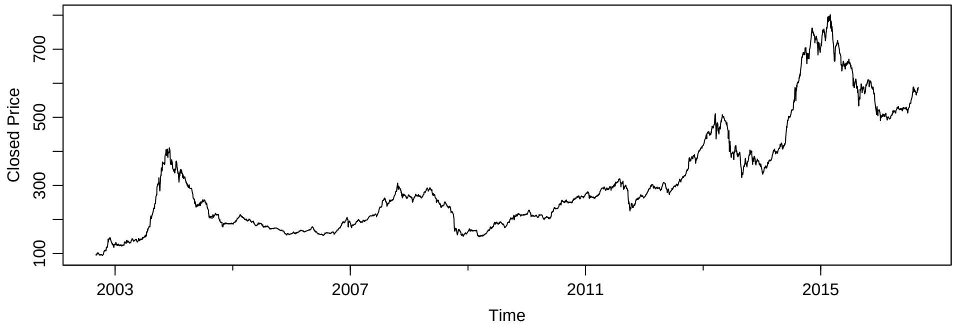


(a) The closed prices of SET index

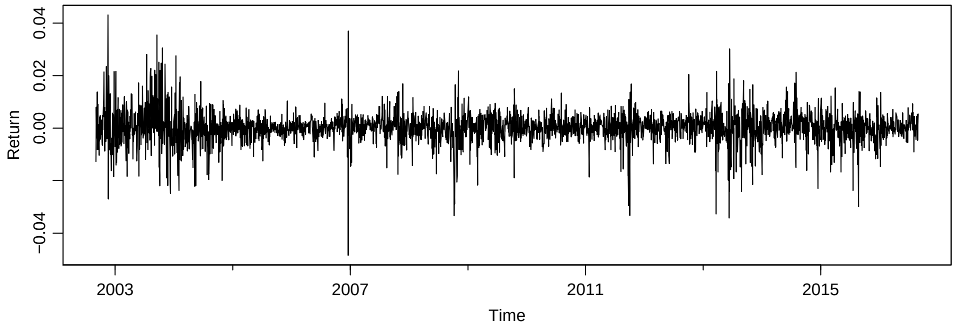


(b) The returns of SET index

Figure 2.1: Plots of the closed prices and the returns of SET index



(a) The closed prices of MAI index



(b) The returns of MAI index

Figure 2.2**:** Plots of the closed prices and the returns of MAI index

CHAPTER III

METHODOLOGY

In this chapter, we provide the methodology using in this thesis to accomplish the two main objectives of the thesis;

1. To construct empirical confidence interval for the Hurst index based on the sample of an efficient market to describe behavior of the efficiency of the market.
2. To investigate behavior of the efficiency of Thailand stock markets using the Hurst index with the DFA method.

We perform the same core method, DFA, for both objectives as described in Section 2.6. For the first objective, we construct the empirical confidence intervals using the method provided in Section 3.1. Then, we consider the Hurst index from return data of Thailand stock markets as the time-varying Hurst index as shown in Section 3.2. Finally in Section 3.3, to complete the second objective, we apply the constructed confidence interval from Section 3.1 with the data of the time-varying Hurst index of Thailand stock markets from Section 3.2 for investigating the behavior of the efficiency of Thailand stock markets.

3.1 Construction of Empirical Confidence Intervals

Theoretically, when one considers the Brownian motion as a representing of efficient market prices, one would expect that the Hurst index is $H_{eff} = 0.5$. However, in the empirical study, it hardly gets the Hurst value exactly 0.5 by estimation. Therefore in this work, we construct the empirical confidence intervals of the estimated of Hurst index ($\hat{H}_{eff}$) when the market is efficient by simulating the sample from efficient markets of the population of the efficient market. Then we simulate a large number of paths of returns as increments of Brownian motion considered as independent and $N(0,1)$ distributed for each different data length as the sample from efficient markets and estimate the Hurst index by DFA method. In this work, we check the property of the normal distribution, rather than quantile in Weron (2002), and then we construct the confidence interval function, as follows;

1. Simulate a sample from efficient matket of the population of the efficient market. Then, sample the returns from the efficient market with 10,000 paths of return of Brownian motion where each point in a path is independent and has $N(0,1)$ distributed, which is randomly generated by R programming (Team, 2013) with the various data lengths in the forms of $N = 2^k$, where $k = 8, 9, ..., 18$.

2. Estimate the Hurst index using DFA method for each sample path to have 10,000 data of the Hurst index of each data length representing the estimated Hurst index for each sample from efficient markets with different data length.

3. Obtain the statistical properties of the distributions of the estimated H index for each data length. In this case, check for the normal distribution using the histogram, normal Q-Q plot and Kolmogorov-Smirnov test. We believe that the Hurst index from the sample of efficient markets has the normal distribution $H_{eff} \sim N(\mu, \sigma^2)$.

4. Build the confidence interval which follows the properties of normal distribution. In this case, we consider the 90%, 95% and 99% confidence interval as follow;

   For the standard normal distribution $Z \sim N(0,1)$ and $0 < \alpha < 1$, we consider $(1-\alpha)100\%$ confidence interval from

$$P(-z_{\alpha/2} < Z < z_{\alpha/2}) = 1-\alpha.$$

Following the results from Step 3, we assume that the Hurst index for the efficient market has the normal distribution $H_{eff} \sim N(\mu,\sigma^2)$ where $Z = \dfrac{H_{eff}-\mu}{\sigma}$ then we consider a $(1-\alpha)100\%$ confidence interval of the estimated Hurst index when the market is efficient as follow;

$$P\left(-z_{\alpha/2} < \frac{H_{eff}-\mu}{\sigma} < z_{\alpha/2}\right) = 1-\alpha,$$

$$P\left(\mu - z_{\alpha/2}\sigma < H_{eff} < \mu + z_{\alpha/2}\sigma\right) = 1-\alpha,$$

The $(1-\alpha)100\%$ confidence interval of the estimated Hurst index for the sample of efficient markets $(\hat{H}_{eff})$ is

$$\left(\bar{H}_{eff} - z_{\alpha/2}s_{eff}, \bar{H}_{eff} + z_{\alpha/2}s_{eff}\right), \tag{3.1}$$

where $\bar{H}_{eff}$ is the sample mean of H_{eff} and s_{eff} is the sample standard deviation of H_{eff}.

5. To obtain confidence intervals for other data lengths different from the lengths simulated in Step 1 ($N = 2^8, 2^9, \ldots, 2^{18}$), we construct the functions representing the lower and upper bounds based on the confidence intervals of the estimated Hurst index obtained in Step 4 as functions of data length N. The construction of the functions is performed based on the knowledge of data fitting using the linear regression for suitable form of functions. In this case, we have 11 data points (lower/upper bounds of the confidence intervals) for $N = 2^8, 2^9, \ldots, 2^{18}$ to produce the best-fit functions. The functions will be constructed for 90%, 95% and 99% confidence intervals.

Note The functions for the lower and upper bounds will be used later to obtain the confidence intervals when the measurement of the Hurst indices for various time windows (data lengths).

3.2 Time-varying Hurst Index

In this work, we extend the idea of Sukpitak and Hengpunya (2016) by computing the Hurst index as time-varying Hurst index using various time window size. In general, the value of H is not constant over time, one can consider using sliding time windows to get a local value of H. There is still no conclusion which size of time windows is the best. The previous work just suggests using time window size that is neither too big nor too small, the period around 3-4 years is often used as in (Costa and Vasconcelos, 2003; Sukpitak and Hengpunya, 2016).

In this work, we estimate the value of the time-varying Hurst index of Thailand Stock market returns (SET and MAI markets) by using different sliding time windows sizes 300, 512, 700, 1024 and 1500 for SET market and 300, 512, 700, 1024 for MAI market. For example, time window of size 512, we estimate the Hurst index using DFA method of the first 512 data and refer to it as the local value of the Hurst index at time 512. Next, we advance the time series to the next day and estimate the local Hurst index at time 513 using data from day 2 to day 513 and repeat this process until the last day of data. The obtained time-varying H-index of the markets will be investigated for efficiency based on the confidence intervals.

3.3 Efficiency of Thailand Markets

To investigate the efficiency of Thailand markets, we use the data provided by `http://siamchart.com/stock/` from 2 May 1975 to 29 August 2016 for SET index and 3 September 2002 to 29 August 2016 for MAI index. We do the following for each data series (SET and MAI indices).

1. Compute the confidence interval of the estimated Hurst index for efficient markets ($\hat{H}_{eff}$) for each time windows sizes using formula obtained from step 5 in Section 3.1.

2. Fit the time-varying Hurst index obtained from Section 3.2 for each windows size with the corresponding confidence interval.

3. Investigate the efficiency for each windows size at each 90%, 95% and 99% confidence intervals.

CHAPTER IV

RESULTS

The study of efficiency of Thailand stock markets in this thesis consists of two main steps;

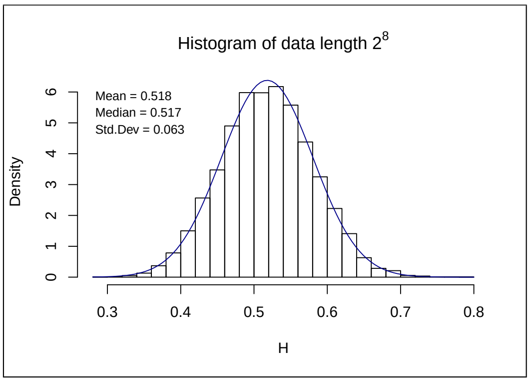
1. Construct confidence interval of the estimated Hurst index for efficient markets ($\hat{H}_{eff}$) at 90%, 95%, 99% confidence intervals based on the Monte Carlo simulation of paths of the returns of the Brownian motion in different lengths, then applied DFA method to get the Hurst index, and study statistical behavior to obtain the confidence interval as presented in Section 4.1.
2. Apply the confidence intervals of $\hat{H}_{eff}$ from step 1 with the time-varying Hurst index of the markets (the SET and MAI markets) using different time windows (data lengths). The results are discussed in Section 4.2.

4.1 Empirical Confidence Intervals

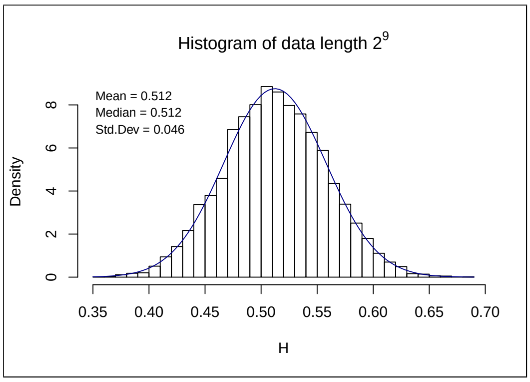
Based on the construction of confidence intervals of the estimated Hurst index when the market is efficient observed in Section 3.1, we simulate 10,000 paths of returns of the Brownian motions as a sample returns from the efficient markets with different data lengths in the form of 2^k, where $k = 8, 9, ..., 18$, and apply the DFA method to obtain the Hurst index for each path. Table 4.1 shows the statistic of the Hurst indices $\hat{H}_{eff}$ for each data length. As results, the skewness that is close to zero and the kurtosis that is close to three, which are similar to the properties of the normal distribution. To confirm that the obtained distributions behave like normal distributions, we check the histograms (see Figure 4.1), the normal Quantile-Quantile plots (see Figure 4.2) and the Kolmogorov-Smirnov test (see Table 4.2) as discussed in Section 2.8.

Data length	Mean	Median	Std.Dev.	Skewness	Kurtosis
2^8	0.518	0.517	0.063	0.090	2.982
2^9	0.512	0.512	0.046	0.037	2.912
2^{10}	0.509	0.508	0.034	-0.007	2.953
2^{11}	0.508	0.508	0.027	-0.047	2.932
2^{12}	0.506	0.506	0.021	-0.026	3.012
2^{13}	0.505	0.505	0.018	-0.041	2.988
2^{14}	0.504	0.504	0.015	-0.026	3.012
2^{15}	0.504	0.503	0.013	0.010	3.000
2^{16}	0.503	0.503	0.011	0.011	2.972
2^{17}	0.503	0.503	0.010	0.021	3.029
2^{18}	0.502	0.502	0.009	-0.008	2.941

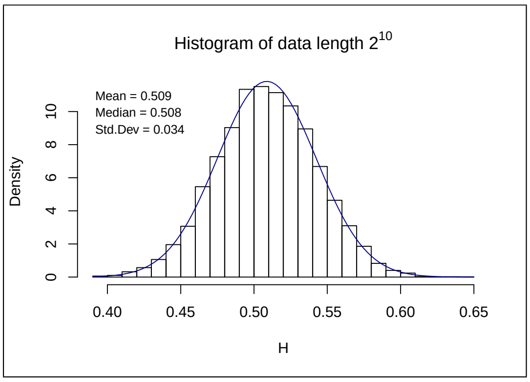
Table 4.1: Table of statistical summary shows that mean and median are approximately equal to 0.5 for each data lengths



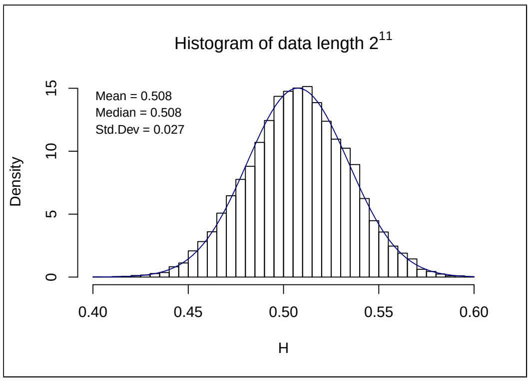
(a) Length 2^8



(b) Length 2^9

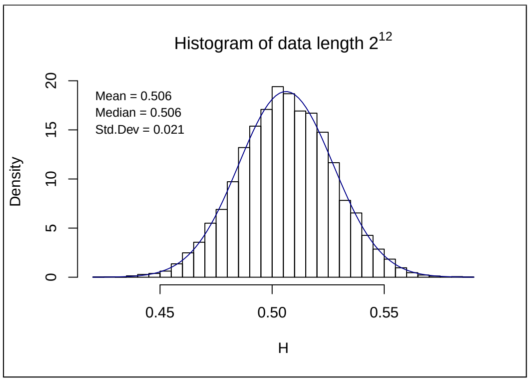


(c) Length 2^{10}

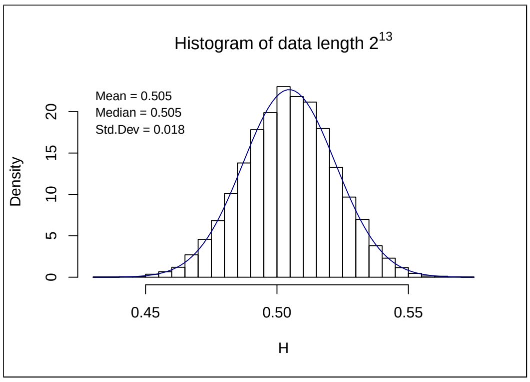


(d) Length 2^{11}

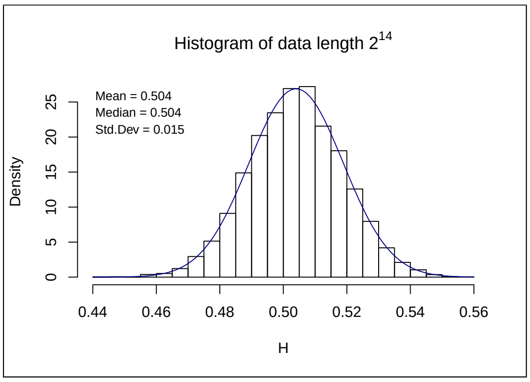
Figure 4.1: Histogram plots of the simulated Hurst indices



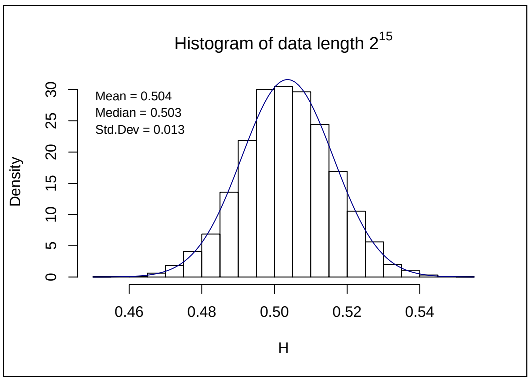
(e) Length 2^{12}



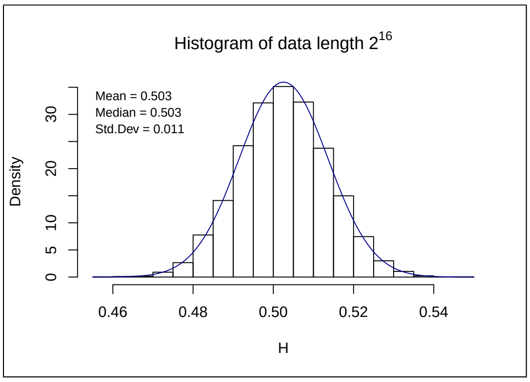
(f) Length 2^{13}



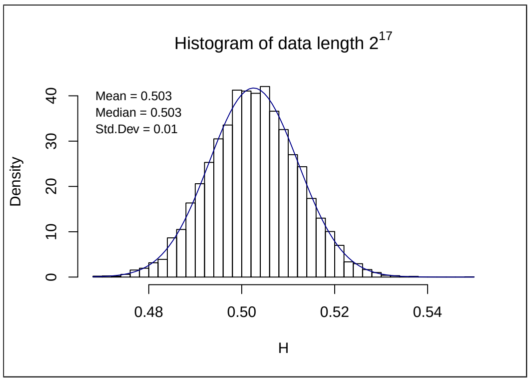
(g) Length 2^{14}



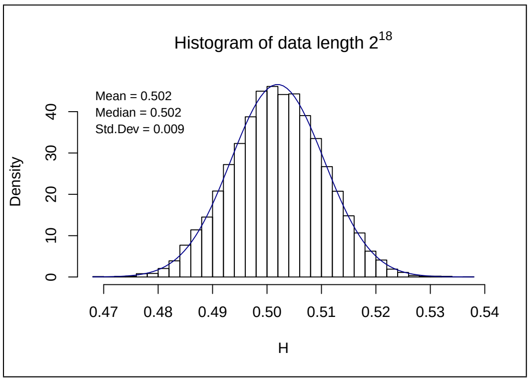
(h) Length 2^{15}



(i) Length 2^{16}

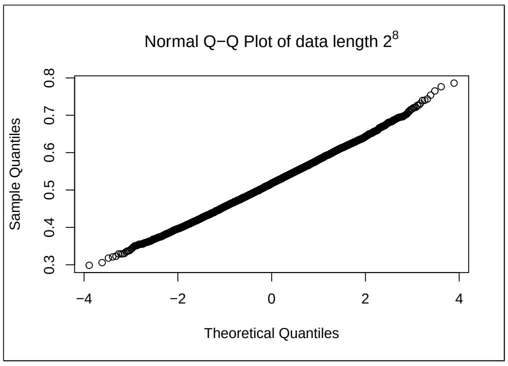


(j) Length 2^{17}

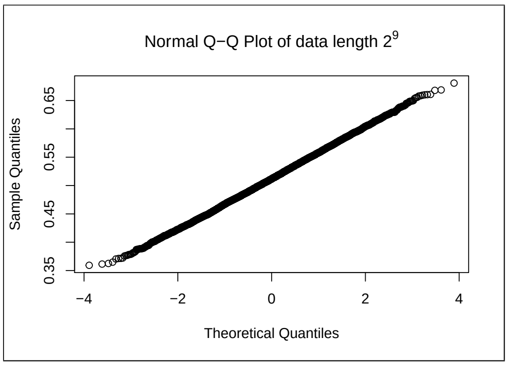


(k) Length 2^{18}

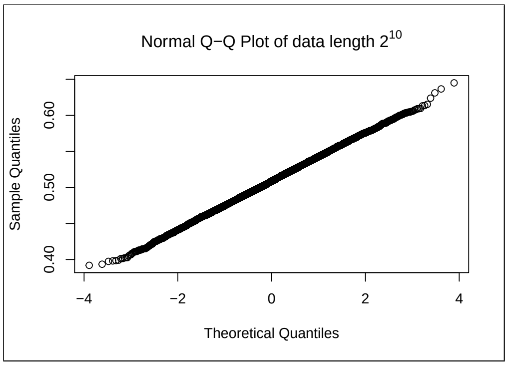
Figure 4.1: Histogram plots of the simulated Hurst indices



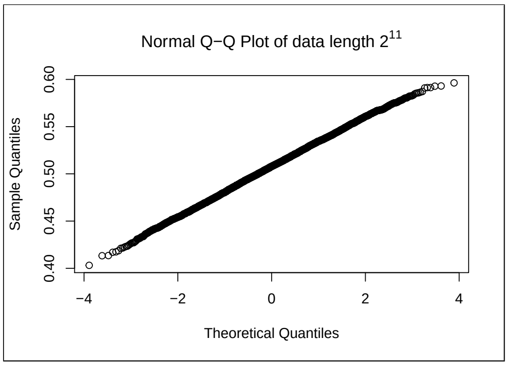
(a) Length 2^8



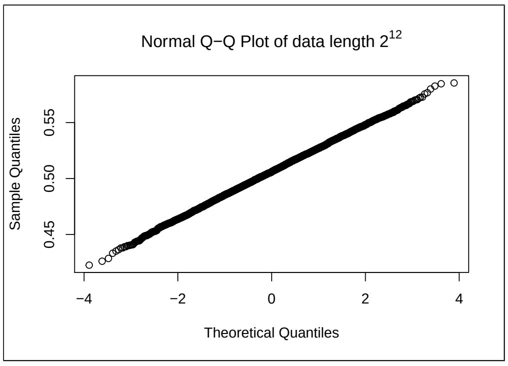
(b) Length 2^9



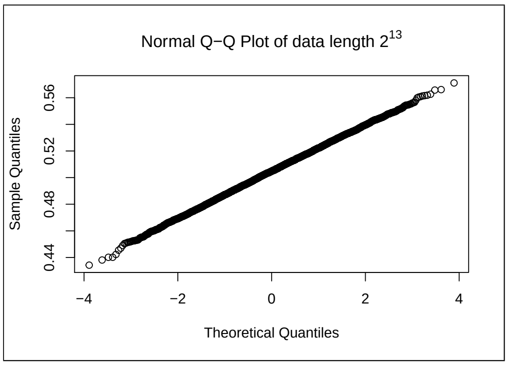
(c) Length 2^{10}



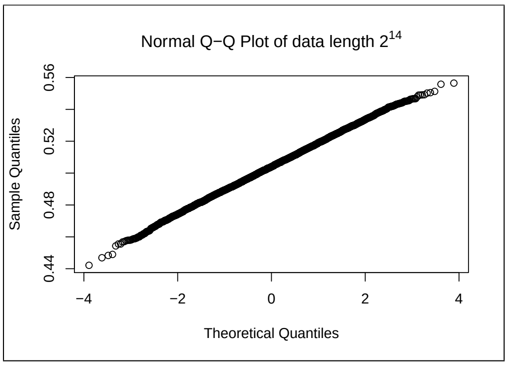
(d) Length 2^{11}



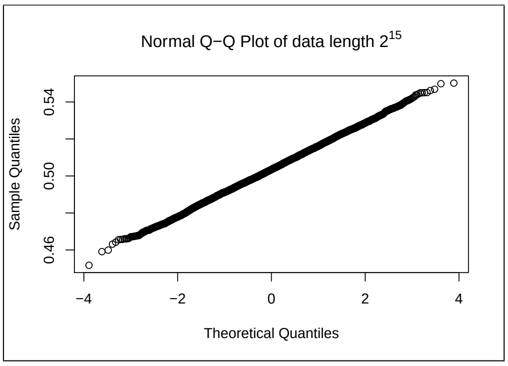
(e) Length 2^{12}



(f) Length 2^{13}

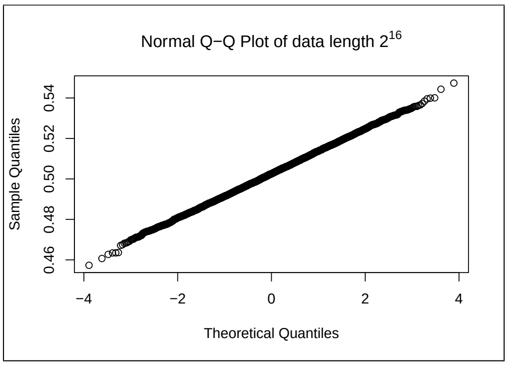


(g) Length 2^{14}

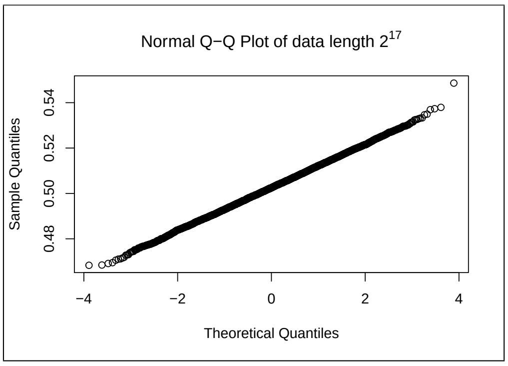


(h) Length 2^{15}

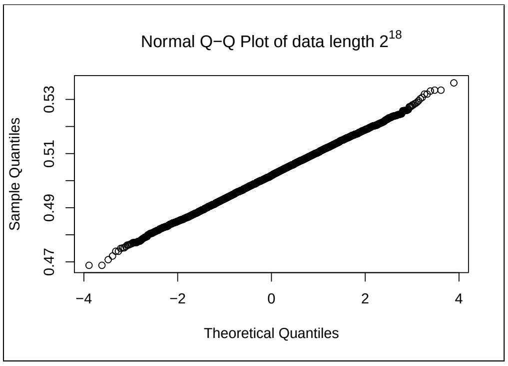
Figure 4.2**:** Normal Q-Q plots of the simulated Hurst indices



(i) Length 2^{16}



(j) Length 2^{17}



(k) Length 2^{18}

Figure 4.2**:** Normal Q-Q plots of the simulated Hurst indices

	Kolmogorov-Smirnov Distance	Kolmogorov-Smirnov p-value
2^8	0.010	0.319
2^9	0.006	0.852
2^{10}	0.006	0.830
2^{11}	0.006	0.858
2^{12}	0.006	0.892
2^{13}	0.009	0.440
2^{14}	0.005	0.967
2^{15}	0.008	0.622
2^{16}	0.006	0.814
2^{17}	0.004	0.998
2^{18}	0.007	0.780

Table 4.2**:** Table of statistic from Kolmogorov-Smirnov test

In Figure 4.1, each histogram of $\hat{H}_{eff}$ fits quite well with the corresponding normal distribution curve created from the obtained mean and standard deviation of $\hat{H}_{eff}$ (Table 4.1) for all data lengths. To confirm with other statistical analysis, the normal Quantile-Quantile plots in Figure 4.2, shows the relation between the sample quantiles and the theoretical quantiles from the normal distribution, which are almost linear implying that the sample distributions behave like the normal distributions. For Kolmogorov-Smirnov test, shown in Table 4.2, the distance is very small and the p-value is above 0.05, which fail to reject the null hypothesis that the distributions of $\hat{H}_{eff}$ from simulated paths are normally distributed at a significance level of 0.05 for all data lengths.

This statistical analysis of the Hurst index when the market is efficient obtained from Monte Carlo simulations of 10,000 sample paths for each data length gives the result that the $\hat{H}_{eff}$ is normally distributed with the corresponding mean and variance given in Table 4.1. Therefore, the construction of confidence interval of $\hat{H}_{eff}$ for each data length is obtained via the fitted normal curve shown in Figure 4.1, where the result is described in the next section.

4.1.1 **Confidence Intervals**

In this work, we are interested in obtaining the confidence intervals of the Hurst index for efficient markets $\hat{H}_{eff}$ at 90%, 95%, and 99% confidence intervals for each data length. The intervals are obtained using the normal curves with means and standard deviation are given in Table 4.1.

For example, for data length 2^8, the mean and standard deviation of the sample are $\bar{H}_{eff} = 0.518$ and $s_{eff} = 0.063$, so the 95% level confidence interval which is calculated as describe is Section 3.1 is $(\bar{H}_{eff} - z_{0.025}s_{eff}, \bar{H}_{eff} + z_{0.025}s_{eff})$ which is (0.395,0.641) as shown in Table 4.3. The confidence intervals at all levels for all data lengths are given in Table 4.3, with the plot given in Figure 4.3. As expected in general that the intervals approach 0.5 as the data lengths approach infinity. We also observed that the higher percents of levels have the bigger intervals will be see Table 4.3 and Figure 4.3, as expected from the formula

(3.1) from Section 3.1.

Data length	Lower 90%	Upper 90%	Lower 95%	Upper 95%	Lower 99%	Upper 99%
2^8	0.415	0.621	0.395	0.641	0.357	0.679
2^9	0.437	0.587	0.423	0.602	0.395	0.630
2^{10}	0.453	0.564	0.443	0.575	0.422	0.596
2^{11}	0.464	0.551	0.455	0.560	0.439	0.576
2^{12}	0.471	0.541	0.465	0.548	0.452	0.561
2^{13}	0.476	0.534	0.470	0.539	0.459	0.550
2^{14}	0.480	0.528	0.475	0.533	0.466	0.542
2^{15}	0.483	0.524	0.479	0.528	0.471	0.536
2^{16}	0.484	0.521	0.481	0.524	0.474	0.531
2^{17}	0.487	0.518	0.484	0.521	0.478	0.527
2^{18}	0.488	0.516	0.485	0.519	0.480	0.524

Table 4.3: The lower and upper bound of the confidence intervals from generated Hurst index

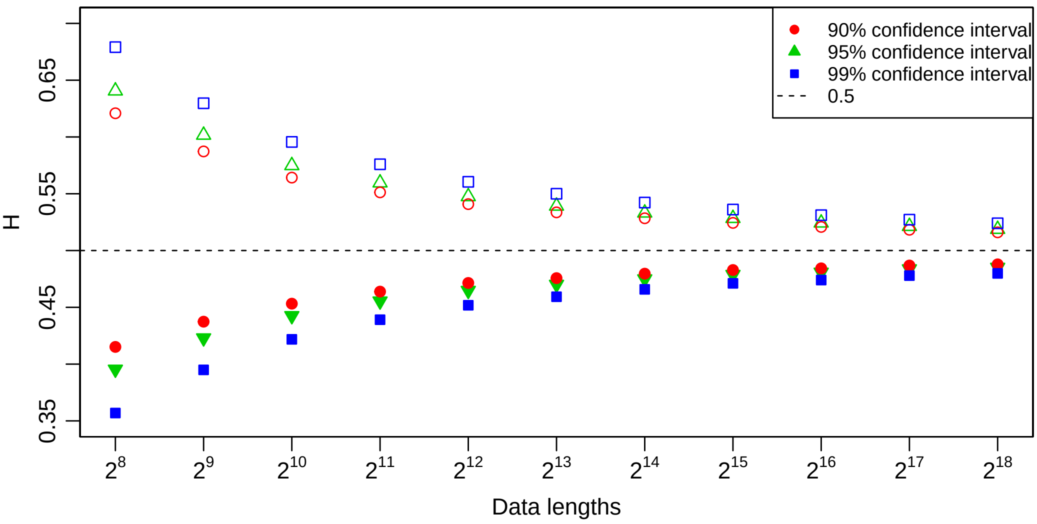


Figure 4.3: The upper and lower bounds of Hurst indices when the market is efficient for 90%, 95%, 99% confidence intervals at various data lengths. The filled shapes represent the lower bounds and the unfilled shapes represent the upper bounds when the red, green and blue shapes belong to 90%, 95% and 99% confidence level, respectively.

4.1.2 **Functions of Confidence Intervals**

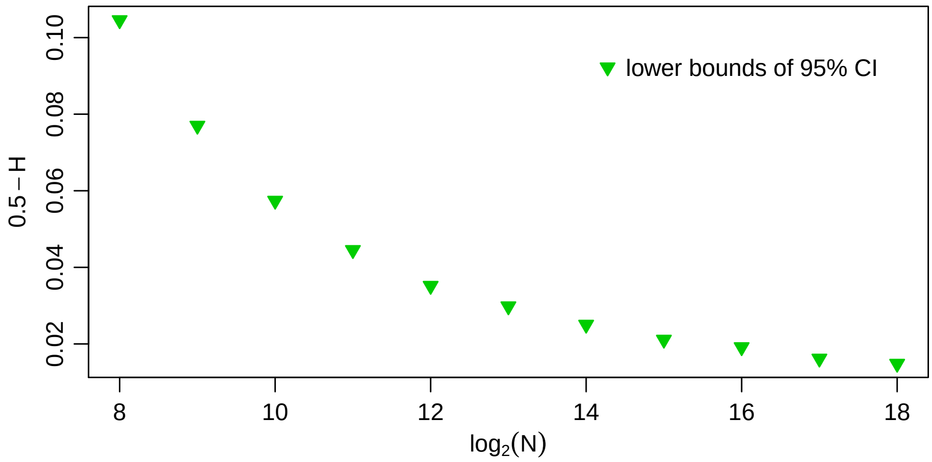
The confidence intervals when the market is efficient at 90%, 95%, 99% confidence intervals for 11 data lengths (from Table 4.3) are used to generate functions for the lower and upper bounds of confidence intervals as functions of the data length N. The functions are obtained via data fitting techniques as described as follows.

For example, at 95% level for the lower bound, the plot of $0.5-H$ and $\log_2(N)$ is shown in Figure 4.4(a) where H is the lower bound of 95% confidence interval presented in Table 4.3. The data-point behaves like the exponential decay. After scaling by log-log plot (see Figure 4.4(b)), the data looks almost linear, and by linear fitting we obtain $\log_2(0.5 - H) = 3.92819 - 3.490524\ \ln(\log_2(N))$ only if $H < 0.5$ in the case of the lower bound. To obtain the function, we solve for H to get, in this case $f_L(N) = H = 0.5 - 2^{-3.490524\ \ln(\log_2(N))+3.92819}$ as in Table 4.4. When applying for time windows in Section 3.2, we compute the confidence interval using the equation in Table 2, for example, when N=1500,

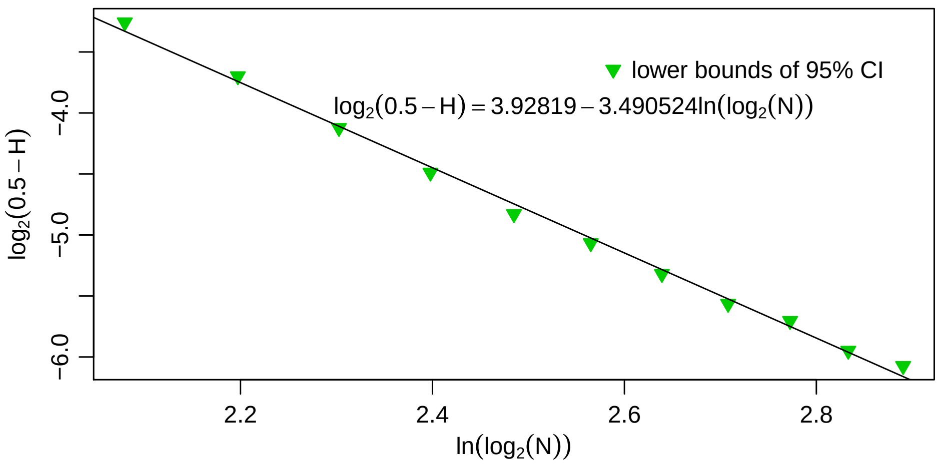
$$\text{the lower bound of 95\% CI} : 0.5-2^{-3.490524\ \ln(\log_2(1500))+3.928190} = 0.4490993,$$

$$\text{the upper bound of 95\% CI} : 0.5+2^{-3.558221\ \ln(\log_2(1500))+4.495110} = 0.5675098.$$

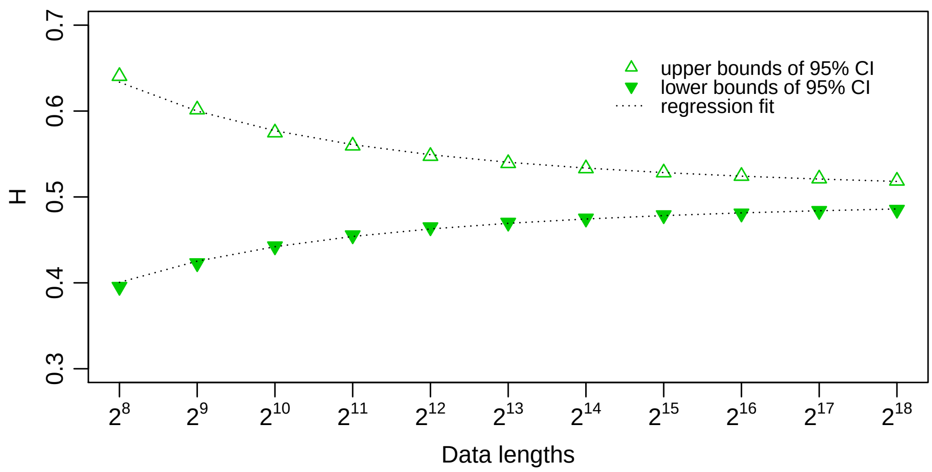
We perform similar technique for constructing the upper bound function, using the data of $H-0.5$ instead of $0.5-H$ in constructing the lower bound function and perform the same for 90% and 99% confidence intervals and finally get the functions of upper bounds and lower bounds as shown in Table 4.4.



(a) Plot of $0.5 - H$ vs $\log_2(N)$ for different sample sizes N



(b) The linear fitting between $\log(0.5 - H)$ and $\ln(\log_2(N))$



(c) The fitting for the lower and upper bounds of 95% confidence interval

Figure 4.4**:** Example of construction of the lower bound function of 95% confidence interval

Level	Lower bound functions	Upper bound functions
90	$0.5 - 2^{-3.481834\ \ln(\log_2(N))+3.609013}$	$0.5 + 2^{-3.563098\ \ln(\log_2(N))+4.287324}$
95	$0.5 - 2^{-3.490524\ \ln(\log_2(N))+3.928190}$	$0.5 + 2^{-3.558221\ \ln(\log_2(N))+4.495110}$
99	$0.5 - 2^{-3.500689\ \ln(\log_2(N))+4.401579}$	$0.5 + 2^{-3.551823\ \ln(\log_2(N))+4.831190}$

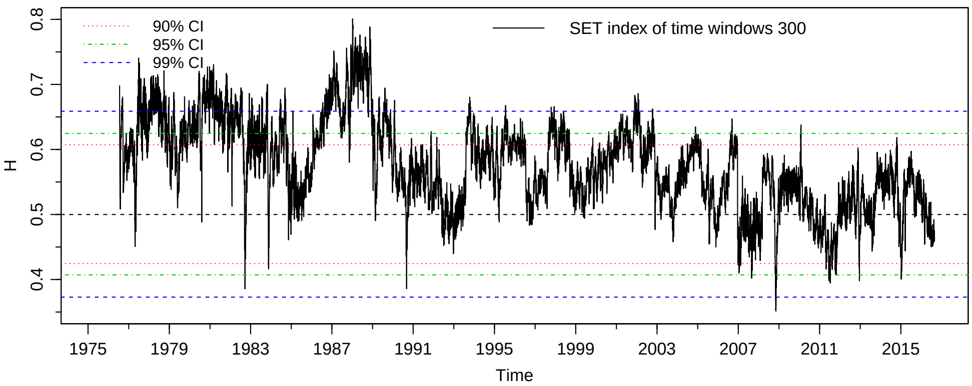
Table 4.4**:** Empirical 90%, 95% and 99% confidence intervals for data length N

4.2 Efficiency of Thailand Stock Market

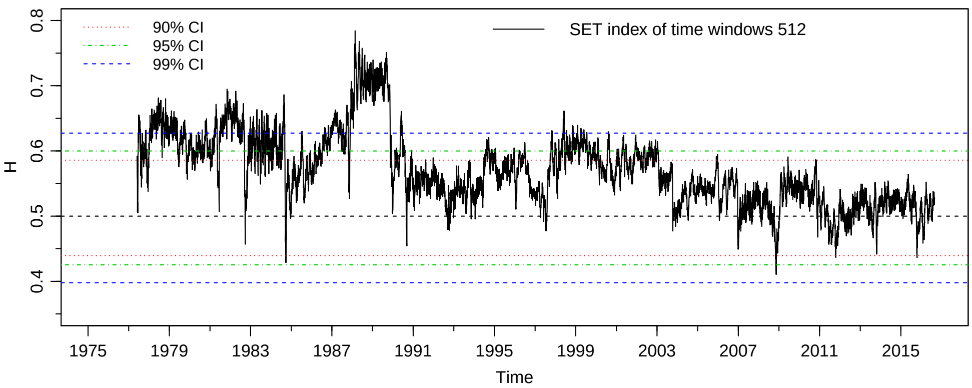
To study the behavior of the efficiency of Thailand stock market, we apply our constructed confidence intervals at 90%, 95%, 99% with the data of the Hurst index estimated by the DFA method of Thailand stock market based on the SET index with 5 different time windows sizes of 300, 512, 700, 1024 and 1500 and also apply with the MAI index (we do not consider at time windows size of 1500 which is too large for the MAI index), we represent time-varying Hurst index of the MAI market as the comparison of time-varying Hurst index of the SET index and MAI index with same time windows sizes as mentioned above.

4.2.1 Time-varying Hurst Index of the SET Index

The results of the time-varying Hurst index of the SET index for various time windows sizes (data lengths) of 300, 512, 700, 1024 and 1500 are presented in Figure 4.5 which is also indicated the confidence intervals when the market is efficient (bounds) at 90%, 95%, 99% levels by dash lines, where outer bound is for 99% and the inner bound is for 90%. Note also that the intervals are smaller when the time windows are bigger. We observed that the results of the H-indices from these time windows behave quite similar (Figure 4.5). They show that from the start of the SET market to 1991, the Hurst indices are often outside the confidence intervals of $\hat{H}_{eff}$ at 90% level, which can conclude that the SET market is inefficient during this period, when considered at 90% level. This conclusion of inefficiency at this period agrees with Sukpitak and Hengpunya (2016). However, when considered at 99%, the H-indices are both within and outside the interval, which gives no conclusion on the inefficient of the SET market during this time. When look closely at the period 1988-1990, all figures gives the H-indices that are outside the 99% confidence interval, which imply that the SET market is inefficient during this time, and move downward the confidence intervals quite rapidly after that. We also can conclude from the results of these figures that we fail to reject that the SET market is efficient after 2004 where the H-indices stay within the 90%,95% and 99% confidence intervals for all time windows, which again agree with the conclusion obtained in Sukpitak and Hengpunya (2016) and Cajueiro and Tabak (2004) that the emerging markets become more efficient over time.

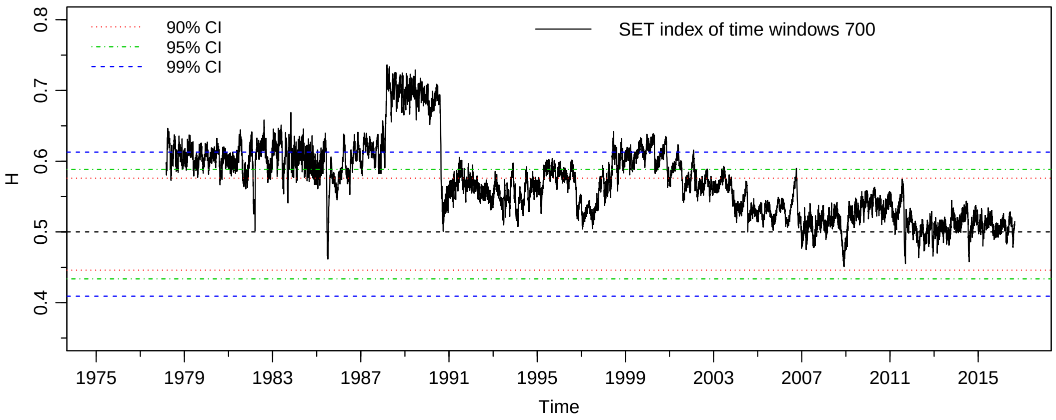


(a) Time-varying Hurst index of the SET market with time window size 300

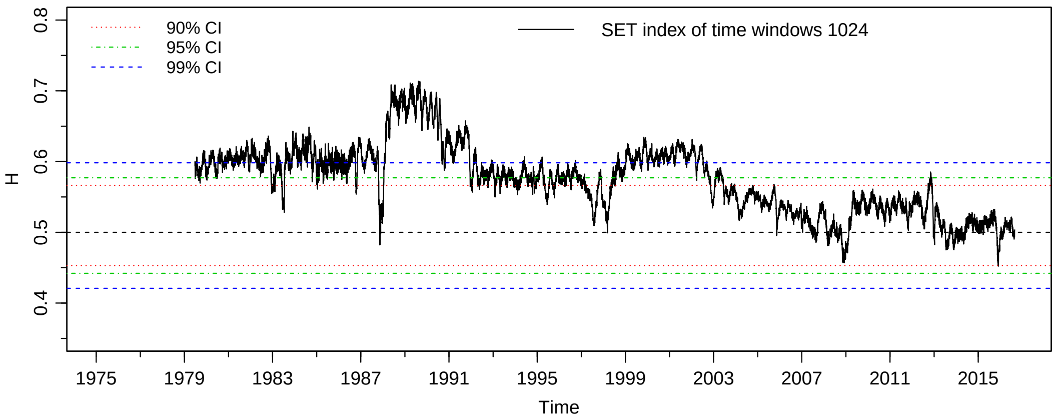


(b) Time-varying Hurst index of the SET market with time window size 512

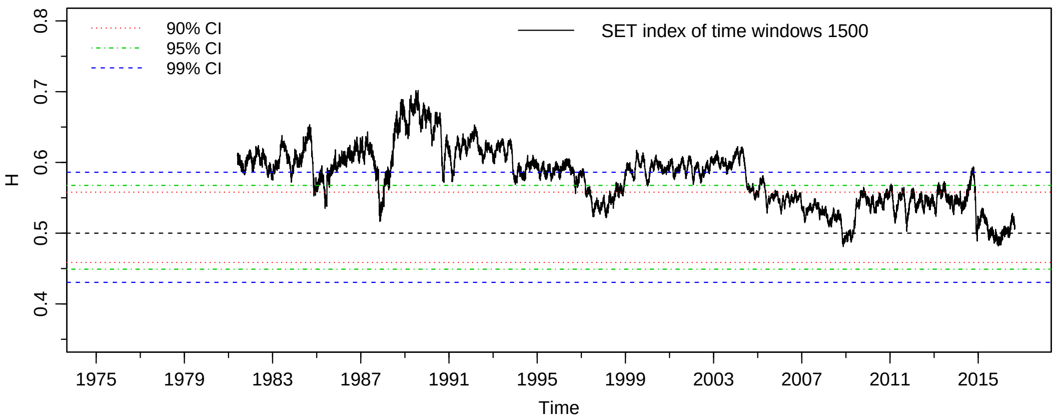
Figure 4.5: Time-varying Hurst index of the SET market



(c) Time-varying Hurst index of the SET market with time window size 700



(d) Time-varying Hurst index of the SET market with time window size 1024

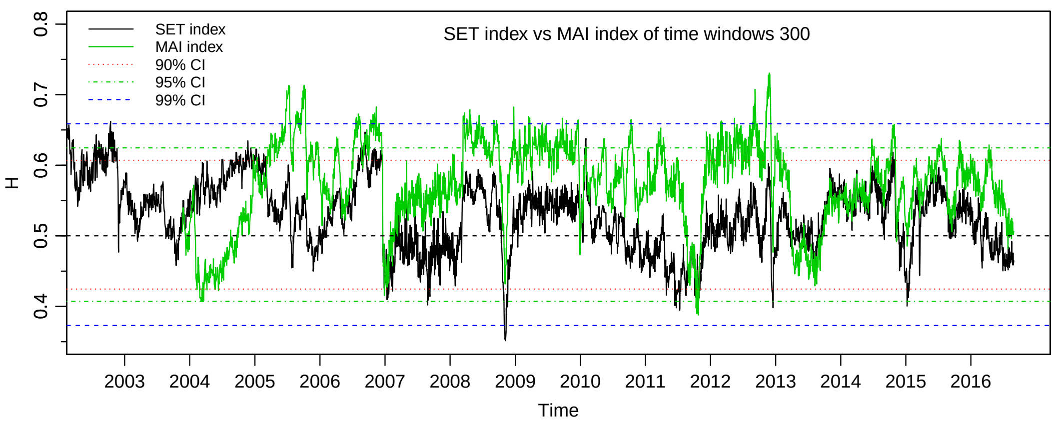


(e) Time-varying Hurst index of the SET market with time window size 1500

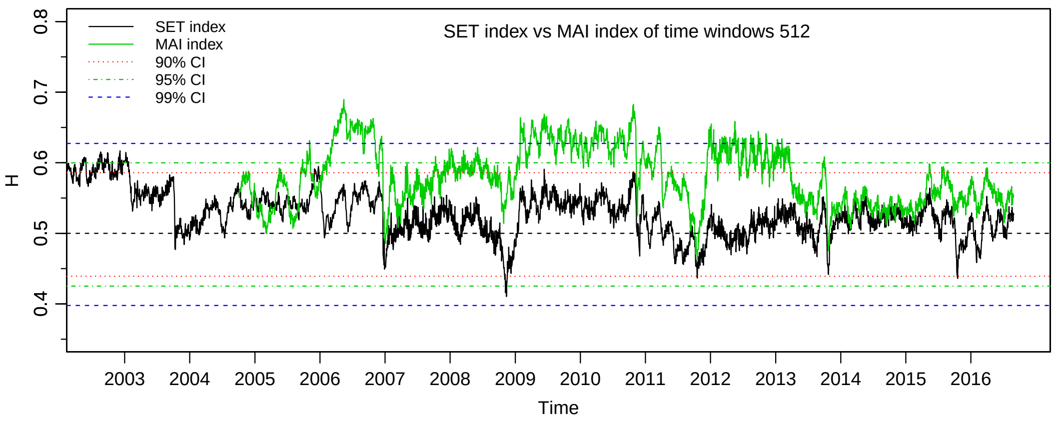
Figure 4.5**:** Time-varying Hurst index of the SET market

4.2.2 Comparison between the Efficiency of the SET Market and MAI Markets

The results of time-varying H-indices of the MAI market for time window sizes 300, 512, 700 and 1024 compared with the SET market are shown in Figure 4.6. The 90%, 95%, 99% levels of confidence intervals are indicated in Figure 4.6 to investigate the efficiency of the market. Figure 4.6 shows that the Hurst indices of the MAI market move almost in the direction parallel to the SET market, but have higher values than those from the SET market from the beginning of the market. The similarity of the behavior of the H-indices of these two markets confirms that both markets are driven under the same environments such as the laws and regulations in Thailand, agreeing with Sukpitak and Hengpunya (2016). In the work of Sukpitak and Hengpunya (2016) that studied the efficiency of the MAI market from 2002 to 2016 with time window size 1024 concluded that the MAI market is not efficient from the beginning to 2016 because the values of H-indices of the MAI market are around 0.6. However, in this study we observed from Figure 4.6 that we fail to reject that the MAI market is quite efficient within 99% confidence interval for the periods 2007-2009 and 2013-2016. Moreover, we also observed the raising of Hurst indices of the SET and MAI markets after 2009 might be the effect of Hamburger crisis in late 2008 which agree with Lim et al. (2008) that the financial crisis makes the increasing of the inefficiency of the market (H-indices). Therefore, this result of the investigation of the efficient market under confidence intervals can give more details of the behavior of the time-varying Hurst index, when compared with the single value of the Hurst index.

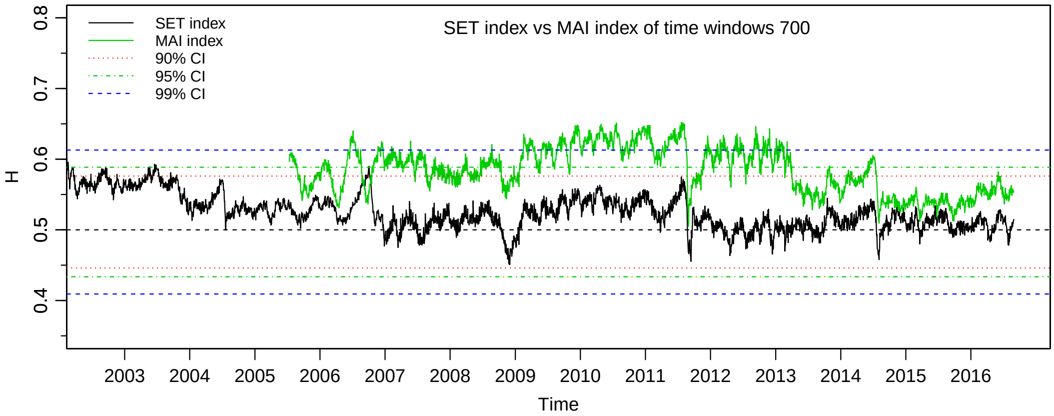


(a) The comparison between the time-varying Hurst index of the SET and MAI markets with time window size 300

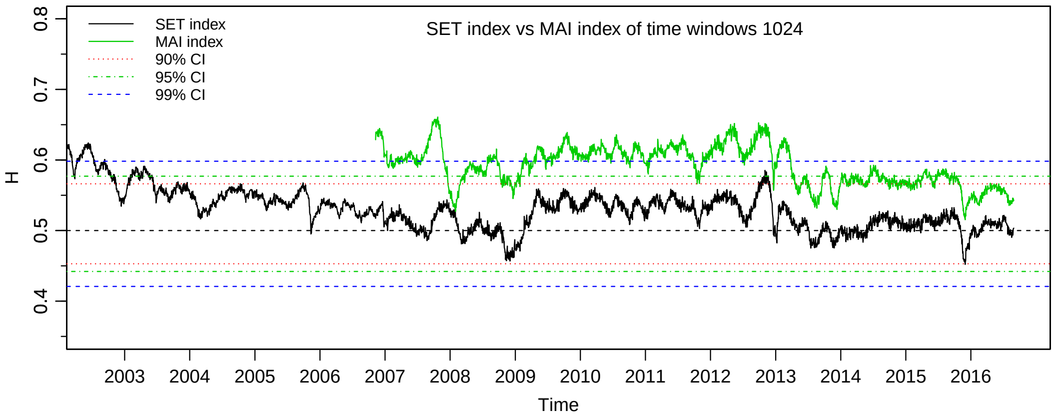


(b) The comparison between the time-varying Hurst index of the SET and MAI markets with time window size 512

Figure 4.6**:** The comparison between the time-varying Hurst index of the SET and MAI markets



(c) The comparison between the time-varying Hurst index of the SET and MAI markets with time window size 700



(d) The comparison between the time-varying Hurst index of the SET and MAI markets with time window size 1024

Figure 4.6**:** The comparison between the time-varying Hurst index of the SET and MAI markets

CHAPTER V

CONCLUSIONS AND DISCUSSIONS

In this work, we construct confidence intervals as indicators in measuring the efficiency of markets instead of using the value of Hurst index at 0.5 as the threshold from previous studies. The construction of confidence intervals is performed based on statistical simulation, where we assumed that the prices in the efficient market follows the Brownian motion.

The obtained confidence intervals is applied to investigated the efficiency of Thailand markets (SET and MAI markets). As explained in the previous sections, the results show that the SET market is quite efficient after 2004, based on the time-varying Hurst indices using various time windows and 90%, 95%, 99% confidence intervals of the estimated Hurst index when the market is efficient. The results for the MAI market are also similar to Sukpitak and Hengpunya (2016) that the MAI market is inefficient from starting in 2002 to 2013 but H-indices move in the similar way as the SET market; however, we might conclude that the MAI market is efficient from 2007-2009 and after 2013.

In this work, we have showed a way to investigate and obtained the result that the efficiency of Thailand markets correspond with the previous result in Sukpitak and Hengpunya (2016). However, the way we looked at the efficiency is performed via the confidence interval that is constructed using statistical simulation of Brownian motion to represent the efficient market.

The conclusion from this study suggests that for efficient markets one can use Brownian motion in a model to describe financial prices. However, when the markets are not efficient, one need to be careful in order to model the asset prices using the Brownian motion, one may need to use the fractional Brownian motion instead.

5.1 Discussions

The result of this study is limited with some assumptions of the sample when constructing the confidence interval that is based on the sample of efficient market. Thus, when apply these intervals to observed data, we can only conclude the inefficiency when the H-indices are outside the confidence interval. One can make this study more complete to investigate the efficiency of the market by following suggestions;

1. One can improve the confidence intervals to measure the efficiency by using the simulations of samples for both efficient and inefficient markets.

2. Since it is not known in general about the range of estimated Hurst index for efficient market, one may need to apply additional hypothesis testing or other techniques to investigate the market efficiency.

3. Another way to construct confidence interval is analyzing the DFA method, which may involved some mathematical knowledge to investigate the bound of the noise (error) of the method, which can lead to more accurate in constructed confidence interval.

REFERENCES

Beran, J. 1994. Statistics for Long-Memory Processes. Chapman and Hall/CRC.

Cajueiro, D. O. and Tabak, B. M. 2004. The hurst exponent over time: testing the assertion that emerging markets are becoming more effcient. Physica A: Statistical Mechanics and its Applications 336.3-4 (2004): 521–537.

Costa, R. L. and Vasconcelos, G. 2003. Long-range correlations and nonstationarity in the brazilian stock market. Physica A: Statistical Mechanics and its Applications 329.1-2 (November 2003): 231–248.

Fama, E. F. 1970. Efficient capital markets: a review of theory and empirical work. The Journal of Finance 25.2 (1970): 383–417.

Hurst, H. E. 1956. The problem of long-term storage in reservoirs. International Association of Scientific Hydrology 1.3 (1956): 13–27.

Kurayami Corp. 2016. Set index historical [Online]. Available from: http://siamchart.com/stock/ [2016,August].

Larsen, R. J. and Marx, M. L. 2001. An Introduction to Mathematical Statistics and Its applications. Prentice-Hall,Inc.

Lim, K.-P., D.Brooks, R., and H.Kim, J. 2008. Financial crisis and stock market efficiency: empirical evidence from asian countries. International Review of Financial Analysis 17 (2008): 571–591.

Lo, A. W. 1991. Long-term memory in stock market prices. Econometrica 59.5 (1991): 1279–1314.

Mandelbrot, B. B. and Ness, J. W. V. 1968. Fractional brownian motions, fractional noises and applications. SIAM Review 10.4 (1968): 422–437.

Massey, F. J. 1951. The kolmogorov-smirnov test for goodness of fit. Journal of the American Statistical Association 46.253 (1951): 68–78.

Olofsson, P. and Andersson, M. 2012. Probability, Statistics, and Stochastic Processes. John Wiley & Sons, Inc.

Peng, C.-K., Buldyrev, S. V., Havlin, S., Simons, M., Stanley, H. E., and Goldberger, A. L. 1994a. Mosaic organization of DNA nucleotides. Physical Review E 49.2 (1994): 1685–1689.

Peng, C.-K., Havlin, S., Stanley, H. E., and Goldberger, A. L. 1994b. Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series. Chaos: An Interdisciplinary Journal of Nonlinear Science 5.1 (1994): 82–87.

Samuelson, P. A. 1973. Proof that properly anticipated prices fluctuate randomly. The Bell Journal of Economics and Management Science 4.2 (1973): 369–374.

Shang, P., Lu, Y., and Kamae, S. 2008. Detecting long-range correlations of traffic time series with multifractal detrended fluctuation analysis. Chaos, Solitons & Fractals 36.1 (2008): 82–90.

Sukpitak, J. and Hengpunya, V. 2016. Efficiency of thai stock markets: detrended fluctuation analysis. Physica A: Statistical Mechanics and its Applications 458 (2016): 204–209.

Taqqu, M. S., Teverovsky, V., and Willinger, W. 1995. Estimators for long-range dependence: an empirical study. Fractals 3 (1995): 785–798.

Team, R. C. 2013. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. Available from: http://www.R-project.org/ .

Weron, R. 2002. Estimating long-range dependence: finite sample properties and confidence intervals. Physica A: Statistical Mechanics and its Applications 312.1-2 (2002): 285–299.

APPENDIX

APPENDIX A

SOURCE CODE

A.1 Construction of Confidence Intervals

A.1.1 Monte Carlo Simulation

```
library(nonlinearTseries)
for (k in 8:18)
{nrSamples <- 10000
data_len <- 2^k
e <- list(mode="vector",length=nrSamples)
for (i in 1:nrSamples) {
  e[[i]] <- rnorm(data_len)
}

white.estimation = c()
for (i in 1:nrSamples)
{
  dfa.analysis = dfa(time.series = e[[i]],
                     window.size.range=c(5,floor(data_len/5)), do.plot=FALSE)
  white.estimation[i] = estimate(dfa.analysis,do.plot=FALSE)
}
save(white.estimation, file=paste("rnorm_",nrSamples,".rda",sep=""))
}

#save(white.estimation, file="whitenoise.rda")
#load("whitenoise.rda")
```

A.1.2 Statistical Summary of Simulated Hurst Indices

```
library(xtable)
library(normtest)
```

```
##############import monticarlo data#######
load("DATA/rnorm_10000sample_len256.rda")
r2_8 <- white.estimation
load("DATA/rnorm_10000sample_len512.rda")
r2_9 <- white.estimation
load("DATA/rnorm_10000sample_len1024.rda")
r2_10 <- white.estimation
load("DATA/rnorm_10000sample_len2048.rda")
r2_11 <- white.estimation
load("DATA/rnorm_10000sample_len4096.rda")
r2_12 <- white.estimation
load("DATA/rnorm_10000sample_len8192.rda")
r2_13 <- white.estimation
load("DATA/rnorm_10000sample_len16384.rda")
r2_14 <- white.estimation
load("DATA/rnorm_10000sample_len32768.rda")
r2_15 <- white.estimation
load("DATA/rnorm_10000sample_len65536.rda")
r2_16 <- rnorm_10000sample_len65536
load("DATA/rnorm_10000sample_len131072.rda")
r2_17 <- rnorm_10000sample_len131072
load("DATA/rnorm_10000sample_len262144.rda")
r2_18 <- rnorm_10000sample_len262144

##########histogram,qq plot##########
for (i in 8:18)
{data <-get(paste("r2_",i,collapse ="",sep=""))
MyMean <- mean(data)
MyMedian <- median(data)
MySd <- sd(data)

hist(data, breaks=30,prob=TRUE,xlab="H",font.main = 2,main=substitute(paste("Histogram of data length ", 2^i),list(i=i)))
     #,xlim=c(0.3,0.8),ylim=c(0,1500))
curve(dnorm(x, mean=MyMean, sd=MySd),
      col="darkblue", lwd=2, add=TRUE, yaxt="n")
legend("topleft", legend = c(paste("Mean =", round(MyMean, 3)),
                             paste("Median =",round(MyMedian, 3)),
                             paste("Std.Dev =", round(MySd, 3))),
       bty = "n",cex=0.85)
#qqnorm(data,main=substitute(paste("Normal Q-Q Plot of data length ", 2^i),list(i=i)))
}
```

```

#########stat summary##############
for (i in 8:18)
{data <-get(paste("r2_",i,collapse ="",sep=""))
MyMean <- mean(data)
MyMedian <- median(data)
MySd <- sd(data)

}

r_df <- data.frame(r2_8,r2_9,r2_10,r2_11,r2_12,r2_13,r2_14,r2_15,r2_16,r2_17,r2_18)
summ_r_df <- data.frame(apply(r_df, 2, mean),apply(r_df, 2, median)
                   ,apply(r_df, 2, sd),apply(r_df, 2, skewness),apply(r_df, 2,
                       kurtosis))
colnames(summ_r_df) <- c("Mean", "Median", "Std.Dev.", "Skewness", "Kurtosis")
xtable(summ_r_df,digits=rep(3,6))

############# normality test ###############
library(normtest)

norm_df <- data.frame()
for (i in 8:18)
{data <-get(paste("r2_",i,collapse ="",sep=""))
MyMean <- mean(data)
MyMedian <- median(data)
MySd <- sd(data)
kstest <-ks.test(data, "pnorm",MyMean,MySd)
}
colnames(norm_df) <- c("ks D","ks pvalue")
rownames(norm_df) <- c("$2^8$","$2^9$","$2^10$","$2^11$","$2^12$","$2^13$","$2^14$","$
    2^15$","$2^16$","$2^17$","$2^18$")
xtable(norm_df,digits=c(0,3,3))

#######CI#############
r_mean <- colMeans(r_df)
r_sd <- apply(r_df, 2, sd)

df_CI <- c()
for (CI in c(90,95,99))
{up <- 1-(1-CI/100)/2
r_err <- qnorm(up)*r_sd #0.975 for 95 CI
r_lower <- r_mean-r_err
```

```
r_upper <- r_mean+r_err

df_CI <- cbind(df_CI,r_lower,r_upper)
}

rownames(df_CI) <- c("$2^8$","$2^9$","$2^10$","$2^11$","$2^12$","$2^13$","$2^14$","$
    2^15$","$2^16$","$2^17$","$2^18$")
colnames(df_CI) <- c("lower 90%","upper 90%","lower 95%","upper 95%","lower 99%","
    upper 99%")
xtable(df_CI,digits=c(0,3,3,3,3,3,3))

df_CI <- data.frame(df_CI)
```

A.1.3 Regression Fit for Constructing Function of Confidence Interval

```
load("DATA/rnorm_10000sample_len256.rda")
r2_8 <- white.estimation
load("DATA/rnorm_10000sample_len512.rda")
r2_9 <- white.estimation
load("DATA/rnorm_10000sample_len1024.rda")
r2_10 <- white.estimation
load("DATA/rnorm_10000sample_len2048.rda")
r2_11 <- white.estimation
load("DATA/rnorm_10000sample_len4096.rda")
r2_12 <- white.estimation
load("DATA/rnorm_10000sample_len8192.rda")
r2_13 <- white.estimation
load("DATA/rnorm_10000sample_len16384.rda")
r2_14 <- white.estimation
load("DATA/rnorm_10000sample_len32768.rda")
r2_15 <- white.estimation
load("DATA/rnorm_10000sample_len65536.rda")
r2_16 <- rnorm_10000sample_len65536
load("DATA/rnorm_10000sample_len131072.rda")
r2_17 <- rnorm_10000sample_len131072
load("DATA/rnorm_10000sample_len262144.rda")
r2_18 <- rnorm_10000sample_len262144

r_df <- data.frame(r2_8,r2_9,r2_10,r2_11,r2_12,r2_13,r2_14,r2_15,r2_16,r2_17,r2_18)

```

```
r_mean <- colMeans(r_df)
r_sd <- apply(r_df, 2, sd)

#Can change CI in range of 90,95,99
CI <- 95
up <- 1-(1-CI/100)/2
r_err <- qnorm(up)*r_sd #0.975 for 95 CI
r_lower <- r_mean-r_err
r_upper <- r_mean+r_err

#Regression fit for lower bound
k=8:18
x=2^k

plot(log(log2(x)),log2(0.5-r_lower))
abline(lsfit(log(log2(x)),log2(0.5-r_lower)))

lowerfit_new <- lsfit(log(log2(x)),log2(0.5-r_lower))
# x is log2(N)
lowerfn_new <- function(x) 0.5-2^(lowerfit_new$coefficients[[2]]*log(x)+lowerfit_new$coefficients[[1]])
c(lowerfit_new$coefficients[[2]],lowerfit_new$coefficients[[1]])
# [1] -3.490524  3.928190 for 95
# [1] -3.481834  3.609013 for 90
# [1] -3.500689  4.401579 for 99
##lowerfn_new <- function(x) 0.5-2^((-3.490524)*log(x)+3.928190)

#Regression fit for upper bound
plot(log(x),log2(r_upper-0.5))
plot(log(log2(x)),log2(r_upper-0.5))
abline(lsfit(log(log2(x)),log2(r_upper-0.5)))

upperfit_new <- lsfit(log(log2(x)),log2(r_upper-0.5))
upperfit_new
c(upperfit_new$coefficients[[2]],upperfit_new$coefficients[[1]])
upperfn_new <- function(x) 0.5+2^(upperfit_new$coefficients[[2]]*log(x)+upperfit_new$coefficients[[1]])

plot(8:18,r_lower,type="p",col=3,pch=25,bg=3,xaxt="n",ylim=c(0.3,0.7),ylab="H",xlab="Data lengths")
par(new=T)
```

```
plot(8:18,lowerfn_new(8:18),type="l",col="black",ylim=c(0.3,0.7),lty=3,axes=FALSE,ann=
    FALSE)
par(new=T)
plot(8:18,r_upper,type="p",col=3,pch=24,bg="white",ylim=c(0.3,0.7),axes=FALSE,ann=
    FALSE)
par(new=T)
plot(8:18,upperfn_new(8:18),type="l",col="black",ylim=c(0.3,0.7),lty=3,axes=FALSE,ann=
    FALSE)
par(new=F)
ticks <- 8:18
labels <- sapply(ticks, function(i) as.expression(bquote(2^ .(i))))
axis(1, at=log2(datasize(8:18)), labels=labels)

legend("topright",inset=0.1, legend = c("upper bounds of 95% CI","lower bounds of 95%
    CI","regression fit"),
       pch=c(2,25,NA), col=c(3,3,1),pt.bg =c(0,3,NA),lty=c(NA,NA,3),cex=0.8,bty="n")
```

A.2 Time-varying Hurst Indices

A.2.1 Construction of time-varying Hurst indices

```
df <-read.csv("C:/Users/Sirapat/Google Drive/Thesis - AMCS/data_MAI_prepared.csv")
mai_df <- df[2:length(df[,2]),c(2,4)]
plot(df[,c(2,3)],main="MAI price")
plot(mai_df,main="MAI return")

w <- 300 #moving size
L <- length(set_df[,2]) #3407 for mai not include the first day


library(nonlinearTseries)
H = c()
for (k1 in 2:(L-w+1))
{
  dfa.analysis = dfa(time.series = set_df[k1:(k1+w-1),2],
                     window.size.range=c(5,floor(w/5)), do.plot=FALSE)
  H[k1] = estimate(dfa.analysis,do.plot=FALSE)
  k1=k1+1
}
```

```
k=2:(L-w+1);
plot(k,H[2:length(H)],type="l")
save(H,file="H_mai_300.RDA")
```

A.2.2 Time-varying Hurst Indices of SET Index

```
dfs <-read.csv("C:/Users/Sirapat/Google Drive/Thesis/Master/data_SET_prepared.csv")
set_df <- dfs[2:length(dfs[,2]),c(2,4)]

set_date <- as.Date(set_df[,1],format="%d-%b-%y")
set_return <- set_df[,2]
set_price <- dfs[2:length(dfs[,2]),c(2,3)][,2]


library(zoo)
set <- zoo(set_return,set_date)
plot(set,xlab="Time",ylab="Returns",xaxt="n")
ticks <- seq(as.Date("1975/1/1"), as.Date("2016/1/1"), "2 years")
axis(1, at = ticks, labels =F, tcl = -0.25)
ticksm <- seq(as.Date("1975/1/1"), as.Date("2016/1/1"), "4 years")
axis(1, at = ticksm, labels =format(ticksm, "%Y"), tcl = -0.5)


setprice <- zoo(set_price,set_date)
  plot(setprice,xlab="Time",ylab="Closed Price",xaxt="n")
ticks <- seq(as.Date("1975/1/1"), as.Date("2016/1/1"), "2 years")
axis(1, at = ticks, labels =F, tcl = -0.25)
ticksm <- seq(as.Date("1975/1/1"), as.Date("2016/1/1"), "4 years")
axis(1, at = ticksm, labels =format(ticksm, "%Y"), tcl = -0.5)

##### Set limit date for plot ###########
x_lim_set=c(as.Date("30-Apr-75", format="%d-%b-%y"),
            as.Date("29-Aug-16", format="%d-%b-%y"))

for(win_length in c(300,512,700,1024,1500,2048))
{
  x <- log(win_length,2)
  lowerfn_95 <- function(x) 0.5-2^((-3.490524)*log(x)+3.928190)
  lowerfn_90 <- function(x) 0.5-2^((-3.481834)*log(x)+3.609013)
```

```
lowerfn_99 <- function(x) 0.5-2^((-3.500689)*log(x)+4.401579)
#upper
upperfn_95 <- function(x) 0.5+2^(-3.558221*log(x)+4.495110)
upperfn_90 <- function(x) 0.5+2^(-3.563098*log(x)+4.287324)
upperfn_99 <- function(x) 0.5+2^(-3.551823*log(x)+4.831190)

Hsetdata<- paste("H_SET_",win_length,".RDA",sep="")
load(Hsetdata)
H.set <- H

par(mar=c(3.5, 4.1, 1, 1.5),mgp=c(2,0.7,0))
H_set <- NULL
H_set <- H.set[2:length(H.set)]
Hset <- zoo(H_set,set_date[(win_length+1):length(set_date)])
plot(Hset,ylim=c(0.35,0.8),xlim=x_lim_set,xlab="Time",ylab="H",col=1,xaxt="n")

  ticks <- seq(as.Date("1975/1/1"), as.Date("2016/1/1"), "2 years")
axis(1, at = ticks, labels =F, tcl = -0.25)
ticksm <- seq(as.Date("1975/1/1"), as.Date("2016/1/1"), "4 years")
axis(1, at = ticksm, labels =format(ticksm, "%Y"), tcl = -0.5)
Hticks <- seq(0.35,0.8,0.05)
axis(2, at = Hticks, labels =F, tcl = -0.25)

abline(h=0.5,col="Black",lty=2)
abline(h=lowerfn_90(x),col=2,lty=3)
abline(h=upperfn_90(x),col=2,lty=3)
abline(h=lowerfn_95(x),col=3,lty=4)
abline(h=upperfn_95(x),col=3,lty=4)
abline(h=lowerfn_99(x),col=4,lty=2)
abline(h=upperfn_99(x),col=4,lty=2)
legend("topright", legend = paste("SET index of time windows ",win_length,sep=""),
       lty=c(1,1), col=c(1,3),cex = 1,bty = "n")
legend("topleft", legend = c("90% CI","95% CI","99% CI" ),
       lty=c(3,4,2), col=c(2,3,4),bty = "n",cex = 0.9)
}
```

A.2.3 Comparison between Time-varying of SET and MAI Index

```
library(zoo)

```

```
df <-read.csv("C:/Users/Sirapat/Google Drive/Thesis/Master/data_MAI_prepared.csv")
dfs <-read.csv("C:/Users/Sirapat/Google Drive/Thesis/Master/data_SET_prepared.csv")
mai_df <- df[2:length(df[,2]),c(2,4)]
#mai_df[,1] date
set_df <- dfs[2:length(dfs[,2]),c(2,4)]

mai_date <- as.Date(mai_df[,1],format="%d-%b-%y")
mai_return <- mai_df[,2]
mai_price <- df[2:length(df[,2]),c(2,3)][,2]
set_date <- as.Date(set_df[,1],format="%d-%b-%y")
set_return <- set_df[,2]

maiprice <- zoo(mai_price,mai_date)
plot(maiprice,xlab="Time",ylab="Closed Price",xaxt="n")
ticks <- seq(as.Date("1975/1/1"), as.Date("2016/1/1"), "2 years")
axis(1, at = ticks, labels =F, tcl = -0.25)
ticksm <- seq(as.Date("1975/1/1"), as.Date("2016/1/1"), "4 years")
axis(1, at = ticksm, labels =format(ticksm, "%Y"), tcl = -0.5)

mairetrun <- zoo(mai_return,mai_date)
plot(maireturn,xlab="Time",ylab="Return",xaxt="n")
ticks <- seq(as.Date("1975/1/1"), as.Date("2016/1/1"), "2 years")
axis(1, at = ticks, labels =F, tcl = -0.25)
ticksm <- seq(as.Date("1975/1/1"), as.Date("2016/1/1"), "4 years")
axis(1, at = ticksm, labels =format(ticksm, "%Y"), tcl = -0.5)


##### Set limit date for plot ###########
x_lim_set=c(as.Date("30-Apr-75", format="%d-%b-%y"),
            as.Date("29-Aug-16", format="%d-%b-%y"))
x_lim_mai=c(as.Date("03-Sep-02", format="%d-%b-%y"),
            as.Date("29-Aug-16", format="%d-%b-%y"))

############# plot for both SET and MAI ##################
for(win_length in c(300,512,700,1024,1500,2048))
{
  x <- log(win_length,2)
  lowerfn_95 <- function(x) 0.5-2^((-3.490524)*log(x)+3.928190)
  lowerfn_90 <- function(x) 0.5-2^((-3.481834)*log(x)+3.609013)
  lowerfn_99 <- function(x) 0.5-2^((-3.500689)*log(x)+4.401579)
  #upper
  upperfn_95 <- function(x) 0.5+2^(-3.558221*log(x)+(4.495110))
```

```
  upperfn_90 <- function(x) 0.5+2^(-3.563098*log(x)+(4.287324))
  upperfn_99 <- function(x) 0.5+2^(-3.551823*log(x)+(4.831190))

  Hmaidata<- paste("H_MAI_",win_length,".RDA",sep="")
  load(Hmaidata)
  H.mai <- H

  Hsetdata<- paste("H_SET_",win_length,".RDA",sep="")
  load(Hsetdata)
  H.set <- H

  H_set <- NULL
  H_set <- H.set[2:length(H.set)]
  Hset <- zoo(H_set,set_date[(win_length+1):length(set_date)])

  H_mai <- NULL
  H_mai <- H.mai[2:length(H.mai)]
  Hmai <- zoo(H_mai,mai_date[(win_length+1):length(mai_date)])

  par(mgp=c(2,0.7,0))
  plot(Hset,ylim=c(0.35,0.8),xlim=x_lim_mai,xlab="Time",ylab="H",col=1,xaxt="n")
  ticks <- seq(as.Date("2002/1/1"), as.Date("2016/1/1"), "years")
  axis(1, at = ticks, labels =format(ticks, "%Y"))
  Hticks <- seq(0.35,0.8,0.05)
  axis(2, at = Hticks, labels =F, tcl = -0.25)

  par(new=T)
  plot(Hmai,ylim=c(0.35,0.8),ann=F,xaxt="n",ylab=F,col=3,xlim=x_lim_mai)
  par(new=F)
  abline(h=0.5,col="Black",lty=2)
  abline(h=lowerfn_90(x),col=2,lty=3)
  abline(h=upperfn_90(x),col=2,lty=3)
  abline(h=lowerfn_95(x),col=3,lty=4)
  abline(h=upperfn_95(x),col=3,lty=4)
  abline(h=lowerfn_99(x),col=4,lty=2)
  abline(h=upperfn_99(x),col=4,lty=2)
  legend("topleft", legend = c("SET index","MAI index","90% CI","95% CI","99% CI"),
         lty=c(1,1,3,4,2), col=c(1,3,2,3,4),cex = 0.8,bty = "n")
  legend("topright", legend = paste("SET index vs MAI index of time windows ",win_
      length,sep=""),
        ,cex = 1,bty = "n")
}
```

BIOGRAPHY

Name Ms. Sirapat Suksai

Date of Birth 12 February 1993

Place of Birth Nakhon Ratchasima, Thailand

Education B.Sc. (Mathematics) (Second Class Honours), Chulalongkorn University, 2014

- Suksai, S. and Mekchay, K. 2017. Efficiency of thailand stock market based on empirical confidence interval of hurst index using dfa method. In - Proceedings of Annual Pure and Applied Mathematics Conference 2017, pp. 164–176.