Testing the Weak-Form Market Efficiency Hypothesis for Canadian and Chinese Stock

Markets

by

Bin Hou

and

Mingyang Sun

PROJECT SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE IN FINANCE

In the Master of Science in Finance Program
of the
Faculty
of
Business Administration

© 2014

SIMON FRASER UNIVERSITY

Term Spring Year 2014

Approval

Name: **Bin Hou; Mingyang Sun**

Degree: **Master of Science in Finance**

Title of Project: **Testing the Weak-Form Market Efficiency Hypothesis for Chinese and Canadian Stock Markets**

Supervisory Committee:

_____

Title and Name
Andrey D. Pavlov
Professor

_____

Title and Name
Jijun Niu
Associate Professor

Date Approved: _____

Abstract

The main empirical test methods for Weak-form efficiency market hypothesis can be divided into two categories: one is to test the randomness of stock prices; the other is to test the invalidity of technical analysis, which testing the predictability of earnings. This study mainly focused on the first category.

To examining the hypothesis whether Canadian and Chinese stock markets are efficient in the weak form, two types of test are conducted. They are parametric and non-parametric tests. For Non-parametric test, we implement the Runs test and Kolmogrov–Smirnov goodness of fit test. For parametric test, autocorrelation (LBQ test), variance ratio and ARMA model have been chosen. The empirical analysis in this study uses daily closing prices of indices from Shanghai Stock Exchange (SSE) and Toronto Stock Exchange (TSX). To avoiding the biases of choosing testing period, we implemented the same tests among different sample periods for each market.

The overall testing results are mixed from sample period to sample period for both markets. In general, for the early testing period, almost all testing techniques generate un-favoured results against the weak-form efficient market hypothesis (EMH) for both TSX and SSE. Several testing results based on more recent sample periods align with the assumption under the EMH, but it is still early to claim that either the Canadian or the Chinese stock market has become the weak form efficient. More comprehensive testing results and analysis can be found under section 5 and 6.

Key Words: weak form market efficiency; runs test; autocorrelation test; ARMA model

Table of Contents

1. Introduction

As one of the cornerstone theories of traditional finance literature, the efficient market hypothesis states that an efficient market is one in which asset prices fully reflect all available information and instantaneously adjust to and fully incorporate new information. There are three types of efficient markets, which are the weak form, the semi-strong and the strong-form efficient market (Fama, 1970).

In a weak-form efficient market, stock prices reflect information contained in market trading data, such as past prices and trading volume. In other words, future stock prices cannot be anticipated by using past stock price information, as past price information is already incorporated into the current stock price. Therefore, the prices of stock behave randomly, or without any identifiable pattern, and investors cannot generate abnormal returns by using technical analysis.

If by increasing the information set to include publicly available information (e.g. information on companies annual earning, announcement of dividends, stock splits, money supply, exchange rate, interest rates, etc.), it is still not possible for a market participant to make abnormal profits, then the market is said to be semi-strong efficient. In other word, both technical and fundamental analyses are not useful in a market like that.

In a strong form market efficiency, stock prices reflect all publicly available and private information, which means that there is no way to make unusual or excess returns by using any types of information.

In this paper, we analyse two financial markets: one is a well-known developed market (Canada); the other is a typical emerging market (China). It is initiative to claim that the longer the development history of the market, the more efficient it could be. The contributions of our studies are not only limited to testing the weak form efficiency for each market, but also providing evidence that there is no clear relationship between the maturity level of the market and the level of its efficiency. The following paragraph briefly introduces some background information of each stock market.

The Shanghai Stock Exchange (SSE) was founded on Nov 26th, 1990 and in operation on Dec 19th the same year. As of March 7, 2014, there are 959 listed companies in this exchange. The total market capitalization of SSE hit RMB 14,511 billion (SSE website). Founded on October 25, 1861, Toronto Stock Exchange (TSX) has become the largest stock exchange in Canada. As of January 31, 2014, there are 1525 listed issuers with total market capitalization of 2,342 billion of Canadian dollars (TMX website). From size wise, SSE is relatively larger than TSX; however, TSX has much longer development history than SSE.

2. Literature Review

Over the past decade, many empirical studies have focused on the efficiency of China's stock market. However, from those studies, it is difficult to draw a definitive conclusion regarding whether China's stock market is efficient or not. The evidence on weak form efficiency is controversial especially in developing and less developed countries market.

Song and Jin (1995) determine that the Shanghai Stock Exchange has already been weakly efficient. At the same time, S. N. Wu (1996), S. X. Song and W. G. Jin (1995), X. Y. Chen et al. (1997), their empirical studies have concluded that China's stock market has reached or come close to the weak form efficient. However, Zhang and Zhou (2001) conclude that China's stock market has not achieved weak-form efficiency. In the late 1990s and the early 2000s, an increasing number of studies found that China's stock market has been weak form efficient (e.g. Li et al., 2001; Chen, 2002). On the contrary, Max Chen and Y. M. Hong (2003) and D. J. Yuan and W. J. Li (2002) both drawn the same conclusion that although the effectiveness of the China's stock market has increased at a later stage as a result of market development and market-oriented trend, it is still too early to draw the conclusion that the China's stock market reached the level of weak form efficiency.

As for the developed market, Empirical evidence on the weak form efficiency indicates mixed results as well. Cooper (1982) studied world stock markets using monthly, weekly and daily data for 36 countries. He examined the validity of the random walk hypothesis by employing correlation analysis, run tests and spectral analysis. With respect to the US and UK, the evidence supported the random walk hypothesis. Seiler and Rom (1997) examined the behaviour of daily stock returns of the US market from February 1885 to July 1962, partitioned annually. Using Box-Jenkins analysis for each of the 77 years, they indicated that changes in historical stock prices were completely random. On the other hand, Fama (1965), using 30 US companies of the Dow Jones, found evidence of dependence in the price change. The first order autocorrelation of daily returns were positive for 23 of the 30 companies and they were significant for 11 of the 30 companies. Lanne and Saikkonen (2004) analysed monthly excess US stock returns from January 1946 to December 2002. The results indicate the presence of conditional skewness in stock returns. The evidence seems to suggest that there is informational inefficiency and stock prices can be predicted with a fair degree of reliability.

Because there are conflicting evidences presented by these studies on the efficiency of China's stock market and some developed markets, it is necessary to examine China's stock market and Canadian stock market, which is a representative of the developed market, by using more recent market data and more comprehensive techniques to determine whether they are efficient or not.

To be more specific, unlike the most of the studies mentioned above, we apply two different methods to perform our testing for each testing target. For example, in order to test the randomness of the raw data, we adopt both the runs test and variance ratio test. The benefit of doing that is to increase the credibility of our testing results. If two different techniques come to the same conclusion, we will have more solid evidence to support our arguments. Moreover, another unique feature of this study is that we conduct each test on both Canadian and Chinese market in parallel. By doing so, we will have more direct statistic comparison between these two markets. Last but not the least, the raw data we used in this study is very up-to-data, which will lead to more accurate results. The following section describes the raw data in more details.

3. Data

In this study, we use stock index from each stock market as the raw data. For Chinese market, we choose SSE Composite Index, which includes all listed stocks (A shares and B shares) at Shanghai Stock Exchange. The Base Day for SSE Composite Index is December 19, 1990. The index was launched on July 15, 1991 (SSE website).

For Canadian market, the S&P/TSX Composite Index is used. This index launched in 1977 with approximately 95% coverage of the Canadian equities market, it is the primary gauge for Canadian-based, Toronto Stock Exchange listed companies (TMX website).

The daily closing prices for each index are retrieved from the Yahoo! Finance for three time periods, which are from 1997-01-01 to 2006-12-31, 2007-01-01 to 2009-12-13 and 2010-01-01 to 2013-12-31. The reason for splitting the entire period from 1997 to 2013 into three sub-periods is due to the high volatility of stock prices during the time of financial crisis happened around 2008. By doing so, not only can we compare our testing results across different sample periods, but also isolate the impact of the financial crisis on our study.

The reason why the daily closing price has been choosen instead of monthly index value is beacasue the transcation costs for active trading has decreased signigicantly in recent years. Thus, it is more reasonable to use daily historcial information to conduct the weak form efficient market testing. For example, if the testing results indicate that the target market is not weak form efficient, certain trading strategy can be formed to exploit this inefficiency. In this case, it makes more sence to use the daily returns as the performance measurement if this is a very active tradeing strategy. Therefore, we decided to use the daily historical index values as the raw data to perform the testing.

The daily market log returns are used as an individual time series variable for most of the testing techniques. Daily market returns (R_t) are calculated from the daily index value such as follows:

$$R_t = \text{Ln} \left(\frac{PI_t}{PI_{t-1}} \right)$$

Where, R_t = market return in period t; PI_t = Index value at time t; PI_{t-1}= the index value at time (t - 1) and Ln = natural log. The plots of raw data and transformed data are available in the Appendix A.

The reasons to take logarithm returns are justified both theoretically and empirically. Theoretically, logarithmic returns are analytically more tractable when linking together sub-period returns to form returns over longer intervals. Moreover, empirically logarithmic returns are more likely to be normally distributed which is a prior condition of standard statistical techniques (A. Mobarek, A. S. Mollah and R. Bhuiyan 2005, N. Strong 1992).

Before performing any time series analysis by using different techniques, we calculated some key statistic information for each time series in four different periods. Period 1 stands for 1997-01-01 to 2006-12-31, Period 2 covers the time from 2007-01-01 to 2009-12-31, Period 3 represents 2010-01-01 to 2013-12-31, and whole period is the summation of these three sub-periods. The detailed statistic information can be found in Table 1.

	Log Returns_TSX			
	Period 1	Period 2	Period 3	Whole Period
Mean	0.0311%	-0.0127%	0.01375%	0.0196%
Standard Deviation	1.0233%	1.7968%	0.85807%	1.1655%
Skewness	-0.6519	-0.5653	-0.33019	-0.6778
Kurtosis	5.39	5.30	2.32	8.55
Minimum	-8.4656%	-9.7879%	-4.12256%	-9.7879%
Maximum	4.6833%	9.3702%	3.94092%	9.3702%
Count	2,511	754	1,003	4,268

	Log Returns_SSE			
	Period 1	Period 2	Period 3	Whole Period
Mean	0.0411%	0.0250%	-0.0432%	0.0192%
Standard Deviation	1.4558%	2.3533%	1.2021%	1.5989%
Skewness	0.0048	-0.3569	-0.2938	-0.2127
Kurtosis	6.29	1.50	1.91	4.81
Minimum	-9.3343%	-9.2562%	-5.4449%	-9.3342%
Maximum	9.4008%	9.0343%	4.2332%	9.4008%
Count	2,607	752	989	4348

Table 1: Statistics Summary

From the Table 1, it can be seen that the volatility of SSE is significantly higher than TSX across all sample periods. In other words, the Chinese stock market is more volatile in the past 15 years compared with Canadian market. Skewness provides an insight about the asymmetry in the distribution of the stock returns. For normal distribution, the skewness should be zero. The above table shows that both markets have negative skewness for all sample periods except for the first period of SSE, which has a slightly positive skewness of 0.0048. Thus, in general, negative skewness in both markets indicate that the distribution of market returns has fat left tail, which means that there is a greater chance for investors to get extremely negative outcomes in both markets. Kurtosis is a measure of the "peakedness" of the probability distribution of a random variable. For standard normal distribution, the kurtosis should be 3. In general, the kurtosis of Canadian market returns is higher than the value in the Chinese stock market. Higher kurtosis means the distribution has fatter tails on both sides. Specifically, in both period 2 and 3, investors have larger chance to get extreme outcomes, both positive and negative, from the Canadian market due to the high peakness in its return's distribution. From the entire testing period perspective, both markets return series have kurtosis greater than 3, 8.55 and 4.81 for TSX and SSE, respectively, which implies a leptokurtic distribution.

Based on these descriptive statistics, we can draw a preliminary conclusion that both markets' log returns crossing all sample periods are not normally distributed and are characterised as leptokurtic or platykurtic and skewed. In other words, the non-normal frequency distributions of the return series deviate from the prior condition of random walk model. To confirm the distribution pattern for these return series, Kolmogrov－Smirnov Goodness of Fit test is also used in the following test, which provides further evidence whether the distribution confirms to a normal distribution or not.

4. Methodologies for Weak Form Market Efficiency Testing

The empirical test in our study uses multi-approach or triangulation of statistical techniques rather than individual approach, which makes it easier to compare the results from different findings and increases reliability of this research. This study includes both Non-parametric (e.g. Runs test and K-S normality test) and Parametric tests (e.g. Variance Ratio test, Auto-correlation test, and ARMA model).

4.1 Non-Parametric Tests

Runs Test

As a well-known and widely used non-parametric test approach, the run test is able to test and investigate serial dependence in share price movements (randomness), which may not be detected by the parametric auto-correlation test. It is a strong test for proving the random walk model because it is independent of the normality and constant variance of data and ignores the properties of distribution.

A run can be defined as a series of price changes of the same sign preceded and followed by the price changes of different sign. The numbers of runs are computed as a sequence of the price changes of the same sign (such as $+ +, - -, 0\ 0$) (Siegel 1956). Null hypothesis of the test is that successive price changes are independent and random. The null hypothesis of randomness of the daily return series is rejected if the expected number of run is dramatically different from the observed number of runs.

The null hypothesis for the run test is as follows:

H0: The observed series is a random series.

H1: The observed series is not a random series.

The test statistics for the number of runs is computed as follows:

$$E(R) = \frac{2N_1 N_2}{N} + 1 \qquad (1)$$

$$\text{Variance: } \sigma_R^2 = \frac{2N_1N_2(2N_1N_2-N)}{(N)^2(N-1)} \tag{2}$$

Where, N = total number of observations; N_1 = number of '+' symbols; N_2 = number of '-' symbols; R = number of runs

The Z statistics tests the significance of the difference between observed and expected number of runs and it is able to give the probability of difference between the actual and expected number of runs. If the Z value is greater than or equal to ± 1.96, we can reject the null hypothesis at 5% level of significance (J. L. Sharma and R. E. Kennedy 1977).

$$\text{The test statistics } Z = \frac{R-E(R)}{\sqrt{\text{var}(R)}} \tag{3}$$

Kolmogorov-Smirnov (K-S) Test

As one of the best-known and most widely used goodness-of-fit tests, the K-S test was originally proposed in the 1930s (Kanji 1999) .The K-S test is of significant importance to provide evidence whether a random sample of data fits a particular distribution (Poisson, uniform and normal) or not. The procedure of one-sample K-S test is to compare the cumulative distribution function (CDF) of the uniform or normal distribution with the empirical distribution function (ECDF) of the sample and test whether the distributions are homogeneous.

4.2 Parametric Tests

In addition to non-parametric tests, this study investigates the parametric tests to examine if the findings by using non-parametric tests confirmed by the findings in parametric tests.

Variance ratio test

The single variance ratio test, proposed by Lo and Mackinlay (1988), demonstrates that the variance ratio test is more powerful than the well-known Dickey–Fuller unit root or the Box–Pierce Q-tests because it uses the fact that if a series of stock prices follows a random walk, then the increments are said to be serially uncorrelated and that the variance of those increments should increase linearly in the sampling intervals. The first null hypothesis is stated as follows:

Let Y_t denote a time series generated from the following equation:

$$Y_t = u + Y_{t-1} + \varepsilon_t \tag{4}$$

where u is an arbitrary drift parameter. The key properties of a random walk that we would like to test are $E(\varepsilon_t) = 0$, for all t, and $E(\varepsilon_t \varepsilon_{t-1}) = 0$.

We may denote estimators for the mean of first difference and variance as:

$$\bar{u} = \frac{1}{T}(Y_t - Y_0) \tag{5}$$

$$\overline{\sigma}^2 = \frac{1}{T}(Y_t - Y_{t-1} - \bar{u})^2 \tag{6}$$

Based on above statistics, we can estimate the variance of the q-period difference to be:

$$\bar{\sigma}^2(q) = \frac{1}{Tq}\sum_{t=1}^{T}(Y_t - Y_{t-1} - q\bar{u})^2 \tag{7}$$

The variance ratio developed by Lo and MacKinlay (1988) is then given by the following:

$$\text{VR}(q) = \frac{\hat{\sigma}^2(q)}{\hat{\sigma}^2(1)} \tag{8}$$

Lo and MacKinlay (1988) also generated the variance ratio z-statistics:

$$Z(q) = \frac{\text{VR}(q)-1}{\varphi(q)} \sim N(0, 1) \tag{9}$$

where $\varphi(q) = [\frac{2(2q-1)(q-1)}{3q(nq)}]^{1/2}$

Serial Correlation Test

Serial Correlation coefficient is the most direct and widely used parametric test for testing EMH in weak form. It provides a measure of the relationship between returns in the current period and those in the previous period. Autocorrelation tests determine whether the correlation coefficients are significantly different from zero.

Auto-correlation coefficient denoted as $\rho(k)$ is defined as

$$\rho(k) = \frac{\text{Cov}(r_t, r_{t-k})}{\sqrt{\text{Var}(r_t)}\sqrt{\text{Var}(r_{t-k})}} = \frac{\text{Cov}(r_t, r_{t-k})}{\text{Var}(r_t)} \tag{10}$$

10

Where $\rho(k)$ is the autocorrelation coefficient on a security at time t, k is the lag of the period. $Cov(r_t, r_{t-k})$ denotes the covariance between the return of an index over time period (t) and its lagged return t-k periods earlier, and $Var(r_t)$ denotes the variance on the return of a security over time period.

The Q-statistic is also used to test the joint hypothesis that all autocorrelation up to 20th lag are simultaneously equal to zero. This study uses the Ljung-Box (LB) statistics by Ljung and Box (1978) which is defined as

$$Q_{LB} = n\ (n+2)\sum_n^m \frac{\rho^2(k)}{(n-k)}, \sim X_m^2 \tag{11}$$

Where n is the sample size and m is the lag length.

ARMA model

In addition to above statistics, this paper also runs ARMA model which consists of two processes: Auto-Regression (AR) and Moving average (MA). Developed by Box and Jenkins (1978), ARMA model is used for describing stationary time series and forecasting future values in time series analysis. This study applies ARMA model to evidence whether the current returns y_t might not only depend on the past return series, but also on the past and current forecast errors. This study follows a method similar to the Box-Jenkins approach (1978) for ARMA building to understand the structure of market index returns and to determine if lagged regressors have predictive powers.

ARMA Model is defined as:

$$y_t = \alpha_0 + \sum_{i=1}^{p} \alpha i y_{t-1} + \sum_{i=1}^{q} \beta i \varepsilon_{t-1} + \varepsilon_t \tag{12}$$

Where y_t is the market returns at time t; α_0 is the expected price change of drift, p denotes the lag values of y_t and q, the lag values of error term ε_t which are independently and identically distributed with the mean of zero and variance σ^2. This implies that current market returns are determined by autoregressive elements of the returns AR (p) as well as moving average of error terms MA (q).

5. Empirical Testing Results Analysis

Since MATLAB has built in functions for all the parametric and non-parametric methods mentioned above. We tested the time series, either the index value itself or the transformed series, e.g. log prices and log returns, in MATLAB by using corresponding MATLAB built-in function calls. The following is outputs of our testing results for both markets with detailed analysis.

5.1 Non-Parametric Test Results Analysis

Runs test (up and down test)

Function 'runstest' in MATLAB performs the runs test on a sequence of observations in the time series, which are the daily index values in this case. The null hypothesis of this test is that the values in the time series come in random order, against the alternative that they do not. In particular, our runs test performs a test for the number of runs up or down. Too few runs indicate a trend, and too many runs indicate an oscillation. Index values exactly equal to the preceding value are discarded. The following two tables summarize some key statistic outputs for the runs test:

	TSX_Index Value			
	Period 1	Period 2	Period 3	Whole Period
Number of runs	1148	394	509	2051
NO. of '+' symbol	1337	414	543	2296
NO. of '-' symbol	1164	338	459	1961
Z – stats	-24.6253	-9.2736	-11.9054	-28.6045

	SSE_Index Value			
	Period 1	Period 2	Period 3	Whole Period
Number of runs	1191	350	483	2025
NO. of '+' symbol	1257	412	482	2152
NO. of '-' symbol	1161	319	485	1966
Z – stats	-20.3009	-12.0391	-12.3226	-26.6191

Table 2: Runs test results (up and down)

From the table 2, it can be seen that all the Z statistics for both indices across different sample periods excess the critical value of -1.96 significantly, which means that the null

hypothesis of the time series following random order can be rejected. The further the Z statistic exceeding its critical value, the less likely the time series follow random walk process. In both markets, the Z statistic shows significant improvement in more recent testing periods. For example, in the first testing period, the Z statistics for TSX and SSE are -24.6253 and -20.3009, respectively. On the contrary, the Z statistics in both markets have improved to the level of -10 for the second and third sample periods. Based on this observation, it is safely to conclude that the randomness of both market indices have enhanced dramatically in recent years; however, there is still a large gap to fill before we can claim that TSX and SSE stock indices are following random walk process.

In addition, by doing cross-markets comparison, it is hard to distinguish the statistics difference of the test results between these two markets. However, from the more recent period perspective, the Z statistics of Canadian market (-9.2736 and -11.9054 for period 2 and 3, respectively) are slightly better than its counterparty (-12.0319 and -12.3226) in this study.

Overall, based on the runs test, on one hand, we can conclude that the stock indices of both TSX and SSE are not following random process across all sample periods, which is against the weak-form market efficiency hypothesis; on the other hand, the randomness of index values show improvements for both markets from the statistic perspective.

Runs test (log returns test)

As indicated by above runs test, the index value itself does not follow random order. Are the log returns of the index also a non-random time series? In this section, we reapply the runs test to the index log returns in each market. In this case, the test is based on the number of runs of consecutive values above or below the mean of the tested time series. If the Z statistic is below -1.96 or above +1.96, we can reject the null hypothesis that the time series follows a random order. The testing results for both markets are showed in the following two tables:

	TSX_Lognorm Returns			
	Period 1	Period 2	Period 3	Whole Period
Number of runs	1154	394	515	2059
Above mean	1302	419	534	2259
Below mean	1209	335	469	2011
Z – stats	-4.0086	1.4892	0.895	-2.1284

13

	SSE_Log Returns			
	Period 1	Period 2	Period 3	Whole Period
Number of runs	1235	358	481	2073
Above mean	1207	409	523	2110
Below mean	1400	343	466	2240
Z – stats	-2.4368	-1.1476	-0.7889	-3.0524

Table 3: Runs test results (Log returns)

Based on testing output in the Table 3, in the case of TSX, the Z statistics for period 2 and 3 are 1.4892 and 0.895, respectively, which are within the range of two critical values, -1.96 and +1.96; thus, we cannot reject the null hypothesis. In other words, the log returns of TSX index follow random walk process. On the contrary, for the first and the entire sample periods, the Z statistics of -4.0086 and -2.1284 are exceeding the critical value of -1.96. Therefore, the log returns of TSX index show a non-random process based on the data from these two sample periods. Similarly, the log returns of SSE index follow random walk in period 2 and 3 with Z statistics of -1.1476 and -0.7889, respectively. For the first and entire sample periods, Z statistics of -2.4368 and -3.0524 provide the evidence of non-random process for the log returns of SSE index.

Interestingly, we observed a similar trend in the Z statistics as the runs test, which conducted for the index values. Specifically, this study shows better Z statistic values, which align with randomness hypothesis, for more recent sample periods. Since the randomness feature is an important characteristic of the weak form efficient market, we can conclude that the market efficiency in both markets has been improved in their latest stages. Regarding the runs test on log returns for each index, it is reasonable to conclude that they are following random process for more recent periods, which is an evidence to support the weak form market efficiency hypothesis.

Kolmogorov-Smirnov test

MATLAB function 'kstest' performs the Kolmogorov-Smirnov (K-S) test to determine if a random sample X could have come from a standard normal distribution, N(0,1). The decision to reject the null hypothesis is based on comparing the p-value with significant level, which is 5% in this testing. In order to accommodate the raw data for the usage of this MATLAB function, the log index return for each market has been standardized by subtracting its mean and dividing its standard deviation. The following two tables include the key statistic results for this test:

	TSX			
	Period 1	Period 2	Period 3	Whole Period
p-value	7.3255e-08	5.0019e-07	2.1852e-04	3.9614e-24
KS Stats	0.0583	0.1001	0.0672	0.0798
Critical value	0.0270	0.0492	0.0427	0.0207

	SSE			
	Period 1	Period 2	Period 3	Whole Period
p-value	4.0773e-15	5.0438e-05	0.0019	6.5808e-25
KS Stats	0.0804	0.0836	0.0591	0.0804
Critical value	0.0265	0.0493	0.0430	0.0206

Table 4: K-S test results

Based on results from the Table 4, all the p-values are significantly smaller than the 5% of the significant level across different sample periods. Thus, the null hypothesis, which states that the time series is a standard normal distribution, can be rejected. In other words, the log returns of each index do not have the characteristic of normal distribution, which confirms our observation from the statistic summary in pervious section.

By comparing the p-values among different testing periods in each market, we can see that for more recent sample period, the p-values are becoming larger in both markets, from 7.3255e-08 and 4.0773e-15 in the first period to 2.1852e-04 and 0.0019 in the third sample period, respectively. Therefore, the distributions of log returns for both indices are getting closer to normal distribution. Overall, however, the results of the K-S test are still against the weak form market efficiency hypothesis, which requires the time series to be a normal distribution.

5.2 Parametric Test Results Analysis

Variance Ratio Test for Randomness

The variance ratio test assesses the null hypothesis that a univariate time series, which is the log price of each index in this case, is a random process. The MATLAB function to perform this test is 'vratiotest'. The variance ratio test statistics in MATLAB are based on a ratio of variance estimates of returns $r(t) = y(t) - y(t-1)$, where $y(t)$ is the log price of the index and $r(t)$ is the log return. For the q period in the equation (7), we use the default value of 2 in the MATLAB. The following two tables listed the testing results for each market:

	TSX			
	Period 1	Period 2	Period 3	Whole Period
p-value	0.002	0.1672	0.1298	0.9108
stats	3.0958	-1.3812	1.5149	0.1121
Critical value	1.96	-1.96	1.96	1.96

	SSE			
	Period 1	Period 2	Period 3	Whole Period
p-value	0.6749	0.888	0.8263	0.6782
stats	-0.4194	-0.1409	-0.2194	-0.4149
Critical value	-1.96	-1.96	-1.96	-1.96

Table 5: Variance Ratio test results

Based on statistic results in the Table 5, it can be seen that all the p-values for SSE index are greater than the significant level of 5%, which leads to a conclusion that the SSE index is a random walk across all sample periods. The testing result for TSX index shows that the index also follows random process for all testing periods, except for the first period, where the p-value is smaller than 5%; thus, we can reject the null hypothesis that the underlying time series is a random walk.

The results of randomness testing by using variance ratio method are similar to the results from runs test on index log returns. In particularly, these two techniques draw the same conclusion for TSX index in the sample periods of 1, 2 and 3 but not for the entire period sample. In the case of SSE index, both methods generate the same conclusion for testing period 2 and 3.

Autocorrelation test (LBQ tests)

One way to determine if there is autocorrelation in the time series is to plot the ACF with different number of lags and check if the correlation coefficient for each lag is significantly different from 0. The ACF and PACF plots for each index among each sample period are available in the Appendix A. The more convenient method to test series correlation is to use the LBQ test. The built in function for this test is 'lbqtest' in MATLAB. The following two tables are the testing results for the log return time series among different sample periods in both markets:

	TSX			
	Period 1	Period 2	Period 3	Whole Period
p-value	2.4255e-06	0.0022	0.0559	3.5822e-04
stats	62.9950	42.7611	30.945	48.5306
Critical value	31.4104	31.4104	31.4104	31.4104

	SSE			
	Period 1	Period 2	Period 3	Whole Period
p-value	0.0093	0.4011	0.6225	1.1650e-05
stats	37.8232	20.9328	17.4673	58.6127
Critical value	31.4104	31.4104	31.4104	31.4104

Table 6: LBQ test results

The testing results from LBQ test are mixed. From the Table 6, it can be seen that the log returns for TSX index have serial correlation in sample periods of 1, 2 and whole period. For its third testing period, the p-value is 0.0559, which is greater than the 5% significant level. Thus, we cannot reject the null hypothesis of non-existence of autocorrelation. In the case of SSE index, the p-values in the second and third sample periods are 0.4011 and 0.6225, respectively, which are greater than the significant level of 5%. Therefore, there is no serial correlation in these two testing periods for SSE. On the contrary, the statistic shows that there are significant autocorrelations in both period 1 and whole period for the Chinese stock market.

In conclusion, there is serial correlation in both markets if we based our testing on the entire sample period. For each individual sample period, however, the results are mixed in each market.

ARMA Model

Based on the conclusion from the LBQ test, both indices show autocorrelation in the whole sample period. Thus, we fit an ARMA model for each time series in that period. The way to determine the best-fitted model is by trial and error. In other words, we keep increasing the orders of the ARMA model until its residuals become s white noise (expect for the constant variance requirement).

For the TSX index, we find ARMA (4, 2) is the best-fitted model for the whole sample period; for the SSE log return time series, the ARMA (2, 5) is the best-fitted model. For choosing the best-fitted model in the ARMA model, we considered the diagnostic checking, which is ACF

(Autocorrelation function) of the residual errors accept the null hypothesis that the residuals are white noise.

According to the weak-form EMH, future market returns will not depend on past (lag) values of share returns or on the past error terms. The coefficients of auto-regression (AR) or moving average (MA) are significantly different from zero suggest dependency of the series, which violates the assumption of the random walk model and weak form efficiency hypothesis. From the Table 7, it can be seen that all the orders for AR and MA model have statistical significant coefficients. This implies that the first four-lag orders of returns in the Canadian market have a predictive power for current returns. Specifically, significant positive autocorrelations in lag 1 and lag 3 indicate the momentum or trend in the market. In addition, the mean reversion characteristic in the market return series can be claimed due to the significant negative autocorrelations in lag 2 and lag 4. The significant MA (1) and MA (2) coefficients indicate that the forecasted past errors have the ability to predict the future market returns as well. Therefore, when the index returns can be predicted on the basis of data on past returns and on forecasted past errors together, this gives rise to ARMA model. Thus, it suggests that the TSX market is not weak form efficient based on the data from whole sample period.

Parameter	Value	Standard Error	T Statistic
AR(1)	1.0998	0.10892	10.0978
AR(2)	-0.57734	0.10538	-5.4787
AR(3)	0.074347	0.012174	6.1072
AR(4)	-0.074335	0.0087236	-8.5211
MA(1)	-1.0972	0.11106	-9.8792
MA(2)	0.53125	0.10644	4.9913
K	0.00013504	1.3393e-06	100.8288

Table 7: ARMA (4, 2) model parameters for TSX in whole sample period

Similar, based on the results in the Table 8, the coefficient of AR (1) and AR (2) have t-statistics of -74.8218 and -54.9023, respectively. The significant coefficient values for the first two-lags indicate the predicative power of historical price information for the SSE. Also, the significant negative autocorrelations on lag 1 and lag 2 reflect that there is a strong mean-reversion characteristic existing in the Chinese stock market. In addition, the significant coefficient values for MA(1), MA(2), MA(3) and MA(5) also provide evidence that the forecasted past errors have the ability to predict the future market returns, which is against the

weak form EMH. In conclusion, SSE market does not satisfy the requirements of weak-form market efficiency hypothesis based on the whole sample period data.

Parameter	Value	Standard Error	T Statistic
AR(1)	-1.2989	0.017359	-74.8218
AR(2)	-0.93865	0.017097	-54.9023
MA(1)	1.2891	0.020182	63.8761
MA(2)	0.90822	0.024989	36.3447
MA(3)	0.0034947	0.020963	0.1667
MA(4)	0.055291	0.0184	3.0050
MA(5)	0.059248	0.011561	5.1248
K	0.00025356	3.0792e-06	82.3477

Table 8: ARMA (2, 5) model parameters for SSE in whole sample period

6. Conclusion

The robustness of the results of this study is assessed using different statistical tests and finding that the results are mixed for different sample periods. By using both parametric and non-parametric tests, we get a more comprehensive coverage and generate results that are more reliable. The following two tables summarize all the testing results across different sample periods including all testing techniques for each market:

TSX Index				
Testing techniques	Period 1	Period 2	Period 3	Overall
Runs test (randomness test for index value)				
Runs test (randomness test for index log-return)		✓	✓	
VR test (randomness test for index log-return)		✓	✓	✓
Description statistics (normality test for log-return)				
K-S test (normality test for index log-return)				
LBQ test (autocorrelation test for log-return)			✓	
ARMA	n/a	n/a	n/a	

Table 9 & 10: Testing results summary

SSE Index				
Testing techniques	Period 1	Period 2	Period 3	Overall
Runs test (randomness test for index value)				
Runs test (randomness test for index log-return)		✓	✓	
VR test (randomness test for index log-return)	✓	✓	✓	✓
Description statistics (normality test for log-return)				
K-S test (normality test for index log-return)				
LBQ test (autocorrelation test for log-return)		✓	✓	
ARMA	n/a	n/a	n/a	

Note: '✓' represents that the result based on certain technique supports the weak-form EMH in that sample period; otherwise, the cell will be left as blank.

In conclusion, based on the testing summary in the Table 9 and 10, it can be seen that the index value in each market does not follow the random walk process across all sample periods. From the statistic point of view, this result is rather expected due to the nature of index price (or stock price), which generally has trends and non-constant mean embedded. Thus, the focus of this study is on the log return time series in each market. In general, both parametric and non-parametric tests support the randomness feature in both TSX index and SSE index, especially for the more recent sample periods, such as the second and third testing periods. As a result, the randomness test is in favour of the weak-form efficient market hypothesis. In addition, the autocorrelation test supports the weak-form EMH for the third sample period in Canadian market and both second and third testing periods in Chinese market. On the contrary, the normality tests did not pass for any single sample period in both markets.

Moreover, by comparing the testing results between the Chinese and Canadian stock markets, we cannot draw a clear conclusion regarding which market is more efficient than the other, even though the Canadian market is a more developed market with much longer development history. Therefore, there is no necessary positive relationship between the development time frame and its level of efficiency.

Overall, among all the different testing periods, we have more faith to claim that the third sample period, which is from 2010-01-01 to 2013-12-31, provides more solid evidence to support the weak form efficient market hypothesis compared with other testing periods in both Canadian and Chinese stock markets.

Since it is still early to claim that either TSX or SSE has already become the weak form efficient market, there are opportunities for investors to form certain trading strategies to exploit the market inefficiency. To be more specific, the following actively traded market instruments, which have very low transaction costs, can be used as the tools to make potential abnormal returns. For TSX, the iShares S&P/TSX Capped Composite Index ETF (ticker XIC) is a suitable candidate. Similar, iShares CSI 300 A-Share Index ETF tracks the SSE index most closely. Therefore, it can be used as the underlying instrument for certain trading strategies.

Bibliography

Works Cited

A. Mobarek, A. S. Mollah and R. Bhuiyan, Market efficiency and stock returns behaviour in emerging market: Evidence from the Bangladesh stock market, Published in Advances in Research in Business and Finance, Vol. 1 as International Conference on Business & Finance – 2005 Proceedings (2005).

Box, G. P. E and G. M. Jenkins (1978), Time Series Analysis: Forecasting and Control, Holden Day, San Francisco.

Chen, Lixin, 2002, "An empirical study on the efficiency of Shanghai stock market," Zhongguo Ruankexue (China Soft Science), No. 5, pp. 23–9.

Cooper, J C B (1982): 'World Stock Markets: Some Random Walk Tests', Applied Economics, 14, pp. 515-31.

Dickey, D A and W A Fuller (1979): 'Distribution of the Estimators for Autoregressive Time Series with a Unit Root', Journal of the American Statistical Association, 74, pp 427-31.

Dickey, D A and W A Fuller (1981): 'Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root', Econometrica, 49, pp 1057-72.

D. J. Yuan and W. J. Li. "China's Capital Market Efficiency of the Theoretical and Empirical Analysis (I)". Journal of Henan Institute of Financial Management, No. 5, 2002, pp. 1-4.

Fama, E F (1965): 'The Behaviour of Stock Market Prices', Journal of Business, 38, pp 34-105.

Fama, E F (1991): 'Efficient Capital-Market: II', Journal of Finance, 46, pp 1575-1617.

J. L. Sharma and R. E. Kennedy, A comparative analysis of stock price behavior on the Bombay, London and New York stock exchanges, Journal of Financial and Quantitative Analysis (1977) 391–413.

Kanji, G K (1999): 100 Statistical Test. Sage Publications, New Delhi.

Ljung, G. M. and G. E. P. Box (1978), 'On a Measure of Lack of Fit in Time Series Models', Biometrika, Vol. 65, pp. 297–303.

Lo, A.W. & MacKinlay, A.C. (1988). Stock market prices do not follow random walks: Evidence from a simple specification test. Review of Financial Studies, 1(1), 41–66.

Li, Xue, Jianmin Liu and Yunhui Jin, 2001, "A run test on the efficiency of China's securities market," Tongji Yanjiu (Statistical Research), No. 12, pp. 43–6.

Lanne, M and PSaikkonen (2004):'A Skewed Garch-in-Mean Model: An Application to US Stock Returns', unpublished paper, University of Jyvaskyla, Finland

Max Chen and Y. M. Hong. "Has Chinese Stock Market Become Efficient—Evidence from a New Approach". China Economic Quarterly, vol.3, No. 1, Oct. 2003, pp. 68-78.

N. Strong, Modelling abnormal returns: A review article, Journal of Business Finance and Accounting 19(4) (1992) 533–553.

Nwosu.E.O, Orji. A and Anagwu.O (2013): ' African Emerging Equity Markets Re-examined: Testing the Weak Form Efficiency Theory', African Development Review, Vol. 25, No. 4, 2013, 485–498

Siegel, Sidney (1956): Non-parametric Statistics for Behavioural Sciences, McGraw-Hill Company, New York.

Song, Songxin and Weigen Jin, 1995, "An empirical study on the efficiency of Shanghai stock market," Jingji Xuejia (The Economists), No. 4, pp. 56–64.

S. N. Wu. "Analysis of China's Securities Market Efficiency". Economic Research Journal, No. 4, 1996, pp. 13-19.

Seiler, Michael J and Walter Rom (1997): 'A Historical Analysis of Market Efficiency: Do Historical Returns Follow a Random Walk', Journal of Financial and Strategic Decisions, 10.

S. X. Song and W. G. Jin. "Empirical Research of Shanghai Stock Market's efficiency". Economist, No. 4, 1995, pp. 107-113.

SSE website: Retrieved from http://www.sse.com.cn/ on March 7th 2014

TMX website: Retrieved from http://www.tmx.com/en/mig/archives.html

Wu, Shinong, 1996, "An analysis of the efficiency of China's securities market," Jingji Yanjiu (Economic Research Journal), Vol. 31, No. 4, pp. 92–101.

X. Y. Chen and X. Chen, B. Gu. "The Empirical Research of China's Stock Market's weak-form efficiency ". Accounting Research, No. 9, 1997, pp. 13-17.

Yahoo! Finance: Retrieved from http://ca.finance.yahoo.com/

Zhang, Yichun and Yinggang Zhou, 2001, "Are China's stock market weak-form efficient?" Jinrong Yanjiu (Journal of Financial Research), No. 3, pp. 15–26.

Appendix A: Plots
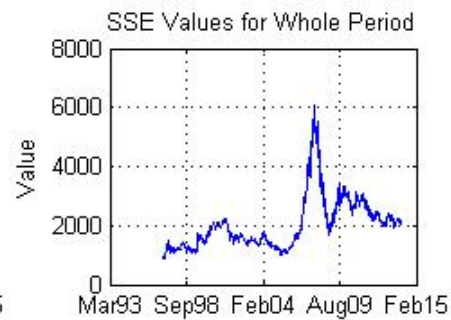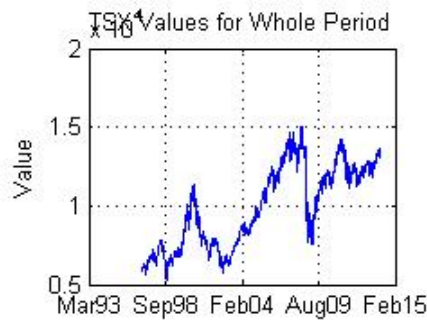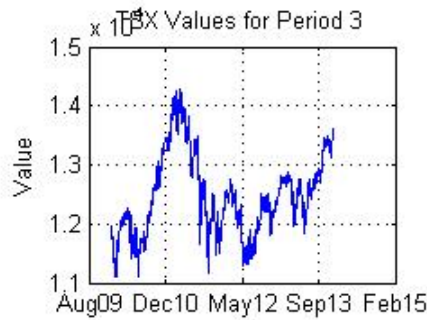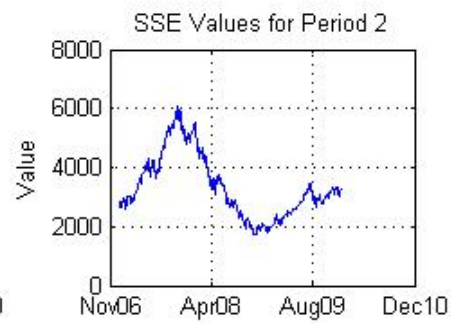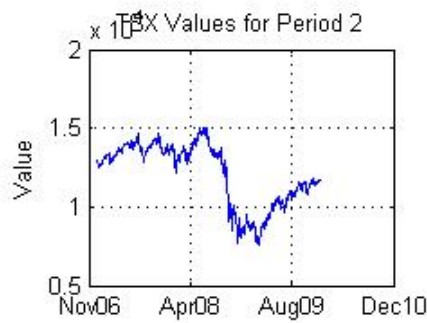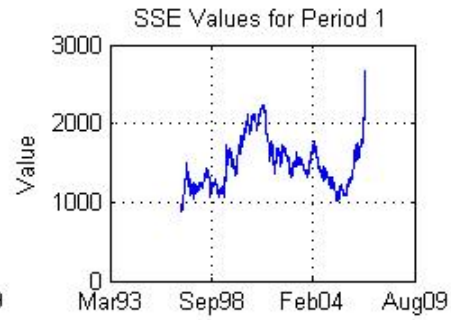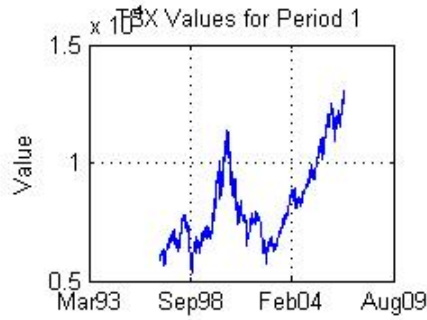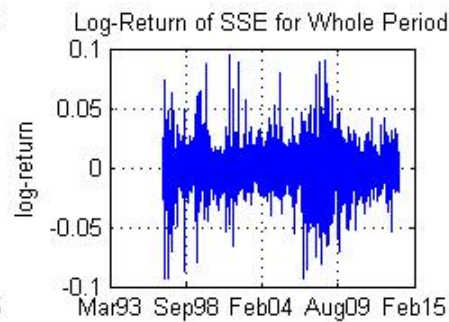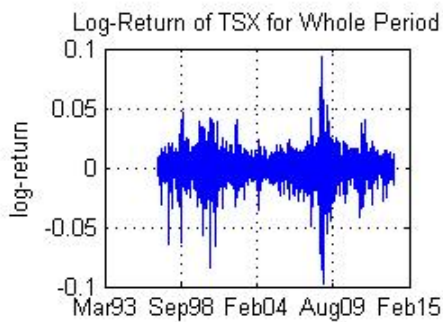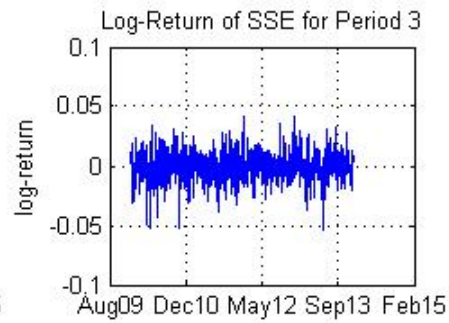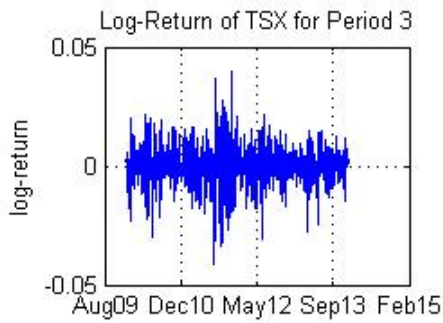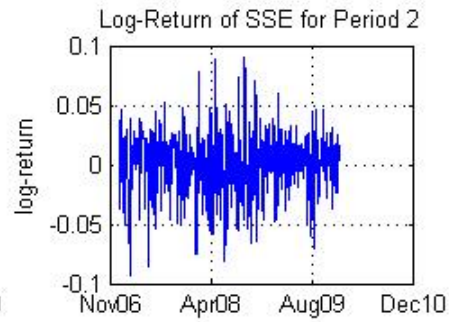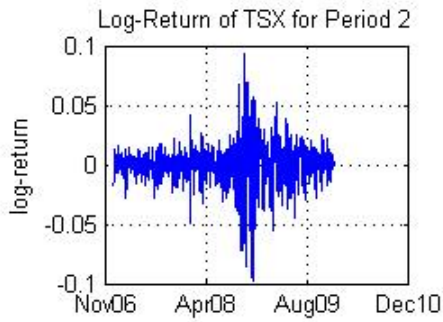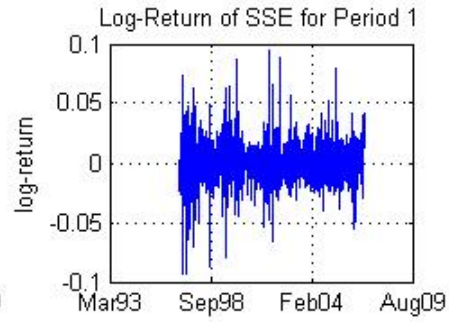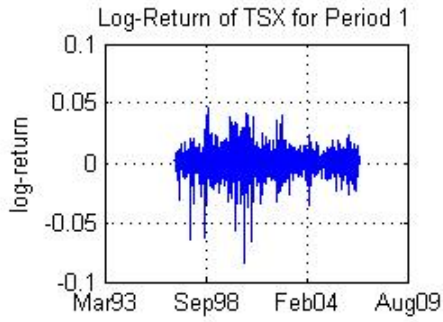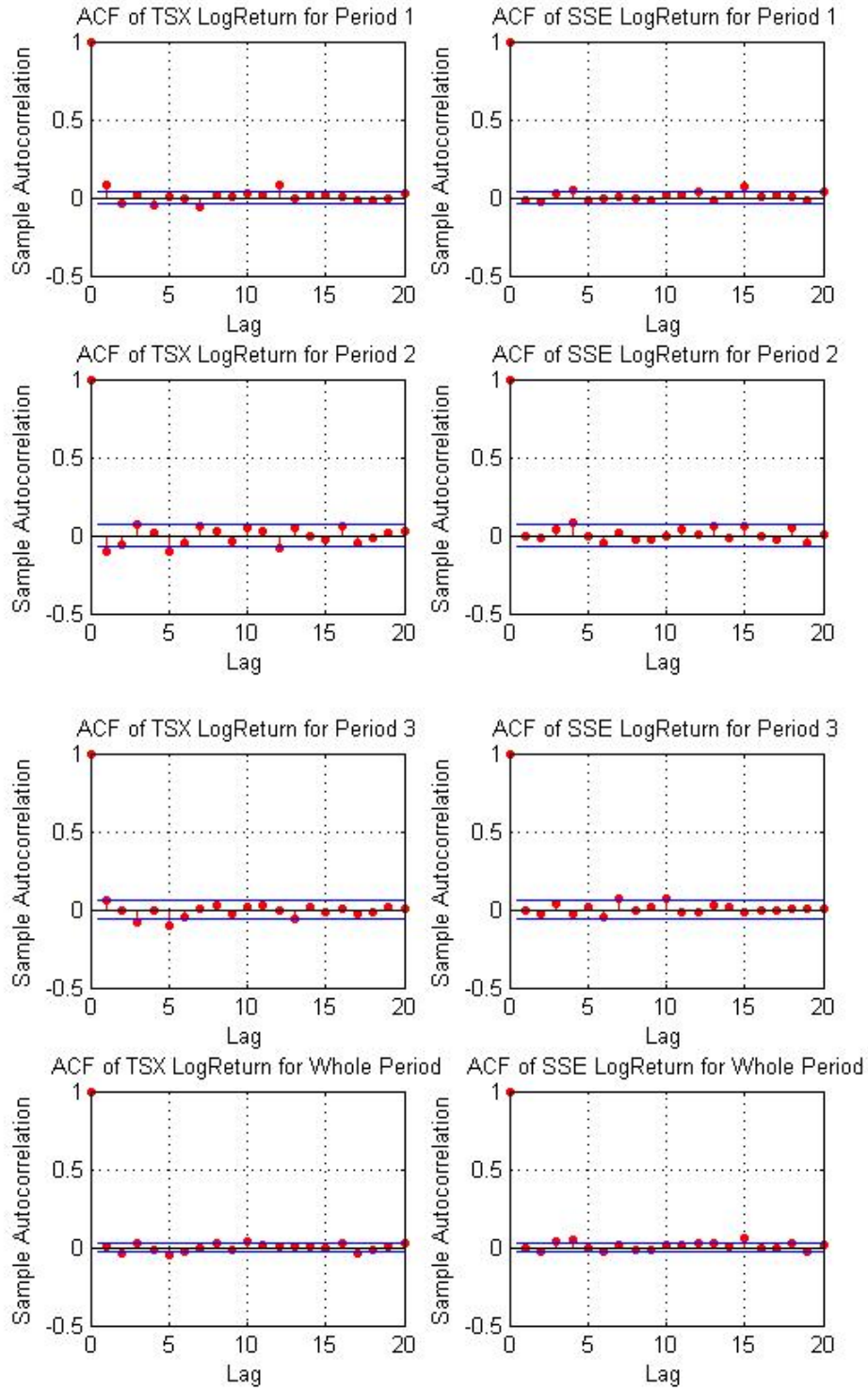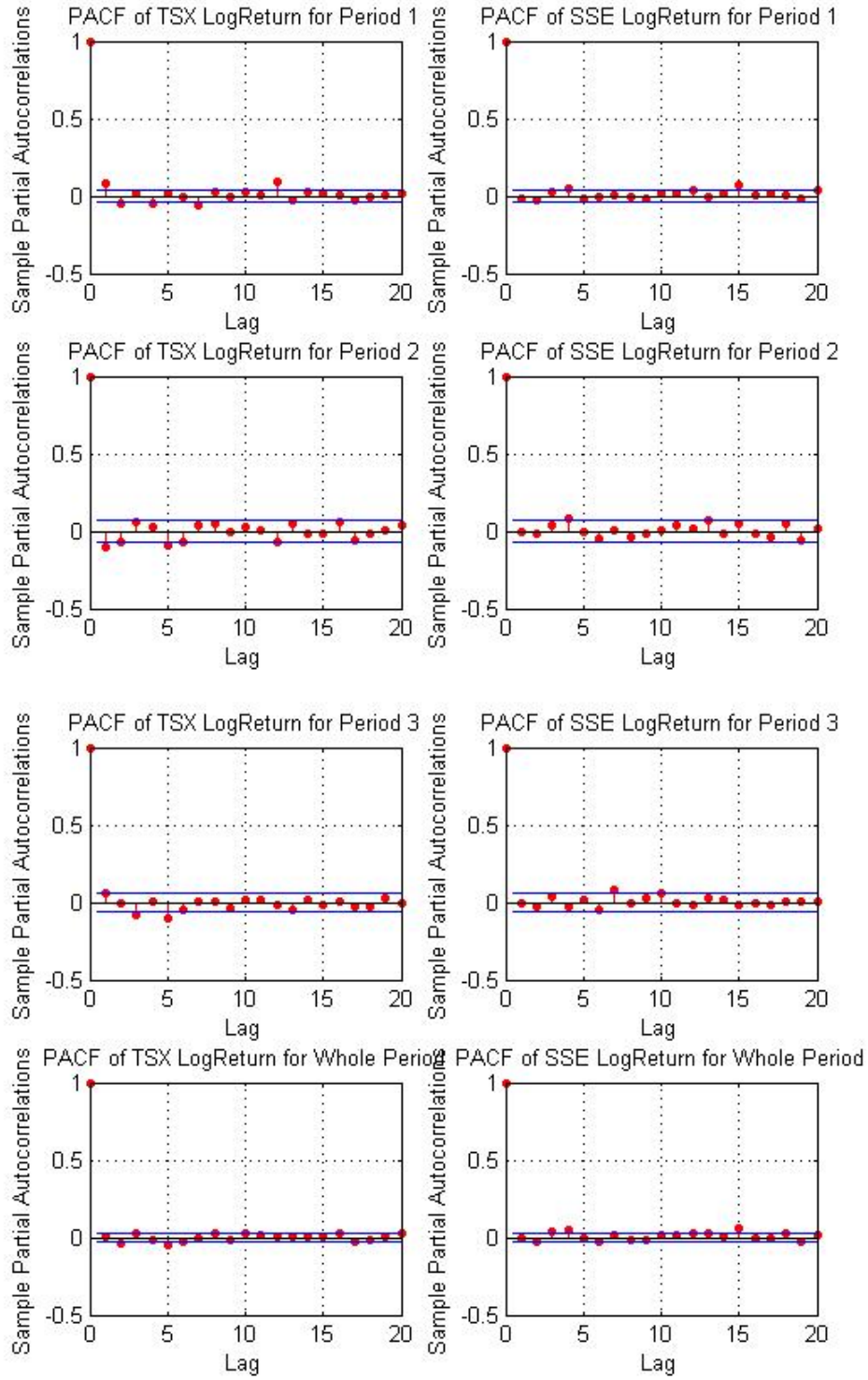
Appendix B: MATLAB Scripts

```matlab
% Formatting
clear
clc
format compact

% To set startDate, endDate and ticker
startDate_1 = '01012010';
endDate_1 = '31122013';

startDate_2 = '01011997';
endDate_2 = '31122013';

% Collect two periods of price data from Yahoo finance for each index
TSX_P1 = hist_stock_data(startDate_1, endDate_1,
'^GSPTSE','frequency','d');
SSE_P1 = hist_stock_data(startDate_1, endDate_1,
'000001.ss','frequency','d');

TSX_P2 = hist_stock_data(startDate_2, endDate_2,
'^GSPTSE','frequency','d');
SSE_P2 = hist_stock_data(startDate_2, endDate_2,
'000001.ss','frequency','d');

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
% Reverse order of both dates and prices, since yahoo imports these
starting with most recent date and price
% Convert date character strings to internal serial numbers using
datenum
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
newDate_TSX_P1 = datenum(TSX_P1.Date(end:-1:1));
newPrice_TSX_P1 = TSX_P1.AdjClose(end:-1:1);

newDate_SSE_P1 = datenum(SSE_P1.Date(end:-1:1));
newPrice_SSE_P1 = SSE_P1.AdjClose(end:-1:1);

newDate_TSX_P2 = datenum(TSX_P2.Date(end:-1:1));
newPrice_TSX_P2 = TSX_P2.AdjClose(end:-1:1);

newDate_SSE_P2 = datenum(SSE_P2.Date(end:-1:1));
newPrice_SSE_P2 = SSE_P2.AdjClose(end:-1:1);

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
% Plot raw data (Indice values)
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
figure('Name','Raw Data','NumberTitle','off')
subplot(2,2,1)
plot(newDate_TSX_P1, newPrice_TSX_P1)
axis auto;
datetick('x','mmmyy', 'keepticks')
```

```matlab
ylabel('Value')
title('TSX Values for Period 3')
grid on

subplot(2,2,2)
plot(newDate_SSE_P1, newPrice_SSE_P1)
axis auto;
datetick('x','mmmyy', 'keepticks')
ylabel('Value')
title('SSE Values for Period 3')
grid on

subplot(2,2,3)
plot(newDate_TSX_P2, newPrice_TSX_P2)
axis auto;
datetick('x','mmmyy', 'keepticks')
ylabel('Value')
title('TSX Values for Whole Period')
grid on

subplot(2,2,4)
plot(newDate_SSE_P2, newPrice_SSE_P2)
axis auto;
datetick('x','mmmyy', 'keepticks')
ylabel('Value')
title('SSE Values for Whole Period')
grid on


%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
% Transform the raw data to log normal returns
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
LogReturn_TSX_P1 = diff(log(newPrice_TSX_P1));
LogReturn_SSE_P1 = diff(log(newPrice_SSE_P1));

LogReturn_TSX_P2 = diff(log(newPrice_TSX_P2));
LogReturn_SSE_P2 = diff(log(newPrice_SSE_P2));


%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
% Plot the transformed data (logreturns)
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
figure('Name','Transformed Data (LogReturn)','NumberTitle','off')
subplot(2,2,1)
plot (newDate_TSX_P1(1:end-1), LogReturn_TSX_P1)
axis auto;
datetick('x','mmmyy','keepticks')
ylabel('log-return')
title('Log-Return of TSX for Period 3')
grid on

subplot(2,2,2)
plot (newDate_SSE_P1(1:end-1), LogReturn_SSE_P1)
axis auto;
datetick('x','mmmyy','keepticks')
```

```matlab
ylabel('log-return')
title('Log-Return of SSE for Period 3')
grid on

subplot(2,2,3)
plot (newDate_TSX_P2(1:end-1), LogReturn_TSX_P2)
axis auto;
datetick('x','mmmyy','keepticks')
ylabel('log-return')
title('Log-Return of TSX for Whole Period')
grid on

subplot(2,2,4)
plot (newDate_SSE_P2(1:end-1), LogReturn_SSE_P2)
axis auto;
datetick('x','mmmyy','keepticks')
ylabel('log-return')
title('Log-Return of SSE for Whole Period')
grid on


%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
% Plot ACF of TSX and SSE using built-in function autocorr
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
figure('Name','ACF Plots','NumberTitle','off')
subplot(2,2,1)
autocorr(LogReturn_TSX_P1,20)
title('ACF of TSX LogReturn for Period 3')

subplot(2,2,2)
autocorr(LogReturn_SSE_P1,20)
title('ACF of SSE LogReturn for Period 3')

subplot(2,2,3)
autocorr(LogReturn_TSX_P2,20)
title('ACF of TSX LogReturn for Whole Period')

subplot(2,2,4)
autocorr(LogReturn_SSE_P2,20)
title('ACF of SSE LogReturn for Whole Period')


%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
% Plot PACF of TSX and SSE using built in function parcorr
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
figure('Name','PACF Plots','NumberTitle','off')
subplot(2,2,1)
parcorr(LogReturn_TSX_P1,20)
title('PACF of TSX LogReturn for Period 3')

subplot(2,2,2)
parcorr(LogReturn_SSE_P1,20)
title('PACF of SSE LogReturn for Period 3')

subplot(2,2,3)
parcorr(LogReturn_TSX_P2,20)
```

```matlab
title('PACF of TSX LogReturn for Whole Period')

subplot(2,2,4)
parcorr(LogReturn_SSE_P2,20)
title('PACF of SSE LogReturn for Whole Period')


%%%%%%%%%%%%%%%%%%%%%%%%%  Non-Parametric Testings
%%%%%%%%%%%%%%%%%%%%%%%%

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
% Perform the runs test for TSX ans SSE
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
runs_TSX_P1 = runstest(newPrice_TSX_P1,'ud');
if runs_TSX_P1 == 1
fprintf ('The TSX index does not follow random walk process for the
first testing period based on runs test.\n')
else
fprintf ('The TSX index follows random walk process for the first
testing period based on runs test.\n')
end

runs_TSX_P2 = runstest(newPrice_TSX_P2,'ud');
if runs_TSX_P2 == 1
fprintf ('The TSX index does not follow random walk process for the
second testing period based on runs test.\n')
else
fprintf ('The TSX index follows random walk process for the second
testing period based on runs test.\n')
end
disp('----------------------------------------------------------------
--')

runs_SSE_P1 = runstest(newPrice_SSE_P1,'ud');
if runs_SSE_P1 == 1
fprintf ('The SSE index does not follow random walk process for the
first testing period based on runs test.\n')
else
fprintf ('The SSE index follows random walk process for the first
testing period based on runs test.\n')
end

runs_SSE_P2 = runstest(newPrice_SSE_P2,'ud');
if runs_SSE_P2 == 1
fprintf ('The SSE index does not follow random walk process for the
second testing period based on runs test.\n')
else
fprintf ('The SSE index follows random walk process for the second
testing period based on runs test.\n')
end
disp('----------------------------------------------------------------
--')


%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
% Kolmogorov-Smirnov (K-S) Test
```

```matlab
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
KS_TSX_P1 = kstest((LogReturn_TSX_P1-
mean(LogReturn_TSX_P1))/std(LogReturn_TSX_P1));
if KS_TSX_P1 == 1
fprintf ('For testing period 1, the log return of TSX index does not
follow normal distribution.\n')
else
fprintf ('For testing period 1, the log return of TSX index follows
normal distribution.\n')
end

KS_TSX_P2 = kstest((LogReturn_TSX_P2-
mean(LogReturn_TSX_P2))/std(LogReturn_TSX_P2));
if KS_TSX_P2 == 1
fprintf ('For testing period 2, the log return of TSX index does not
follow normal distribution.\n')
else
fprintf ('For testing period 2, the log return of TSX index follows
normal distribution.\n')
end
disp('----------------------------------------------------------------
--')

KS_SSE_P1 = kstest((LogReturn_SSE_P1-
mean(LogReturn_SSE_P1))/std(LogReturn_SSE_P1));
if KS_SSE_P1 == 1
fprintf ('For testing period 1, the log return of SSE index does not
follow normal distribution.\n')
else
fprintf ('For testing period 1, the log return of SSE index follows
normal distribution.\n')
end

KS_SSE_P2 = kstest((LogReturn_SSE_P2-
mean(LogReturn_SSE_P2))/std(LogReturn_SSE_P2));
if KS_SSE_P2 == 1
fprintf ('For testing period 2, the log return of SSE index does not
follow normal distribution.\n')
else
fprintf ('For testing period 2, the log return of SSE index follows
normal distribution.\n')
end
disp('----------------------------------------------------------------
--')


%%%%%%%%%%%%%%%%%%%%%%%%%%%% Parametric Testings
%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
% Unit Root Test by using the Augmented Dickey - Fuller (ADF) Test
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
ADF_TSX_P1 = adftest(newPrice_TSX_P1);
if ADF_TSX_P1 == 1
```

```matlab
fprintf ('The TSX index is an random process for the first testing
period.\n')
else
fprintf ('The TSX index is not an random process for the first testing
period.\n')
end

ADF_TSX_P2 = adftest(newPrice_TSX_P2);
if ADF_TSX_P2 == 1
fprintf ('The TSX index is an random process for the second testing
period.\n')
else
fprintf ('The TSX index is not an random process for the second testing
period.\n')
end
disp('----------------------------------------------------------------
--')

ADF_SSE_P1 = adftest(newPrice_SSE_P1);
if ADF_SSE_P1 == 1
fprintf ('The SSE index is an random process for the first testing
period.\n')
else
fprintf ('The SSE index is not an random process for the first testing
period.\n')
end

ADF_SSE_P2 = adftest(newPrice_SSE_P2);
if ADF_SSE_P2 == 1
fprintf ('The SSE index is an random process for the second testing
period.\n')
else
fprintf ('The SSE index is not an random process for the second testing
period.\n')
end
disp('----------------------------------------------------------------
--')


%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
% Testing the autocorrelation of the transformed data by using LBQ test
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
[~,LBQ_TSX_P1] = lbqtest(LogReturn_TSX_P1);
if LBQ_TSX_P1 <= 0.05
fprintf ('There is serial correlation among the log-return of TSX index
for the first period.\n')
else
fprintf ('There is no serial correlation among the log-return of TSX
index for the first period.\n')
end

[~,LBQ_TSX_P2] = lbqtest(LogReturn_TSX_P2);
if LBQ_TSX_P2 <= 0.05
fprintf ('There is serial correlation among the log-return of TSX index
for the second period.\n')
else
```

```matlab
fprintf ('There is no serial correlation among the log-return of TSX
index for the second period.\n')
end
disp('-----------------------------------------------------------------
--')

[~,LBQ_SSE_P1] = lbqtest(LogReturn_SSE_P1);
if LBQ_SSE_P1 <= 0.05
fprintf ('There is serial correlation among the log-return of SSE index
for the first period.\n')
else
fprintf ('There is no serial correlation among the log-return of SSE
index for the first period.\n')
end

[~,LBQ_SSE_P2] = lbqtest(LogReturn_SSE_P2);
if LBQ_SSE_P2 <= 0.05
fprintf ('There is serial correlation among the log-return of SSE index
for the second period.\n')
else
fprintf ('There is no serial correlation among the log-return of SSE
index for the second period.\n')
end
disp('-----------------------------------------------------------------
--')


%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
% Variance Ratio Test for Randomness
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%
VR_TSX_P1 = vratiotest(log(newPrice_TSX_P1));
if VR_TSX_P1 == 1
fprintf ('The log normal return of TSX index does not follow random-
walk process for the first testing period.\n')
else
fprintf ('The log normal return of TSX index follows random-walk
process for the first testing period.\n')
end

VR_TSX_P2 = vratiotest(log(newPrice_TSX_P2));
if VR_TSX_P2 == 1
fprintf ('The lognormal price of TSX index does not follow random-walk
process for the second testing period.\n')
else
fprintf ('The log normal price of TSX index follows random-walk process
for the second testing period.\n')
end
disp('-----------------------------------------------------------------
--')

VR_SSE_P1 = vratiotest(log(newPrice_SSE_P1));
if VR_SSE_P1 == 1
fprintf ('The lognormal price of SSE index does not follow random-walk
process for the first testing period.\n')
else
```

```matlab
fprintf ('The lognormal price of SSE index follows random-walk process
for the first testing period.\n')
end

VR_SSE_P2 = vratiotest(log(newPrice_SSE_P2));
if VR_SSE_P2 == 1
fprintf ('The lognormal price of SSE index does not follow random-walk
process for the second testing period.\n')
else
fprintf ('The lognormal price of SSE index follows random-walk process
for the second testing period.\n')
end
disp('----------------------------------------------------------------
--')


% Fit an ARMA model using garchset and garchfit
% The best fit model parameters are determined by trail and error

% For the log return of TSX in the first period, model ARMA(6,6) has
been used
spec_TSX_P2 = garchset('R',4,'M',2,'C',nan,'Display','off');
[coeff_TSX_P2,errors_TSX_P2,LLF_TSX_P2,Innovations_TSX_P2,Sigmas_TSX_P2
] = garchfit(spec_TSX_P2,LogReturn_TSX_P2);
garchdisp(coeff_TSX_P2,errors_TSX_P2);

% For the log return of SSE in the first period, model ARMA(6,2) has
been used
spec_SSE_P2 = garchset('R',2,'M',5,'C',nan,'Display','off');
[coeff_SSE_P2,errors_SSE_P2,LLF_SSE_P2,Innovations_SSE_P2,Sigmas_SSE_P2
] = garchfit(spec_SSE_P2,LogReturn_SSE_P2);
garchdisp(coeff_SSE_P2,errors_SSE_P2);
```