Non-Bayesian Time-Varying Parameter Regression Models: Estimation and Applications

By

Mikio Ito

A DISSERTATION SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DERGEE OF

DOCTOR OF PHILOSOPHY

(ECONOMICS)

at

Keio University

2015

Non-Bayesian Time-Varying Parameter Regression Models: Estimation and Applications

Mikio Ito

At Keio University

This dissertation presents empirical results for financial systems using a new method originally developed in the author's own paper, Ito (2007). The method employed in this dissertation, an extended one from the paper's, sheds new light on estimation of a state space model. In the context of econometrics, one can regard the model as a linear regression model with time-varying parameters. A state space model has been conventionally estimated by the Kalman filtering and smoothing; it can be done by his new method, based on an ordinary linear regression, a non-Bayesian time-varying regression. The extended method not only guarantees robust estimation of the regression but also allows ones to conduct statistical inferences based on a sequence of critical values of state variables of the state space model. It enables this dissertation to attain fruitful results when examined is structure of financial markets such as stock markets or commodity futures markets.

The first chapter provides an overview of background of this dissertation. In particular, it contains a survey of literature on market efficiency of financial markets, which is a long history of debate. The chapter elucidates the reason why the debate has lasted for goods, discussing the analytical tools employed by the literature. It also shows a new approach of research paying attention to unceasing structural changes in financial markets such as stock markets, commodity futures markets, etc. Accordingly, required is such a new approach that allows us to deal with the markets under unceasing structural changes. The first

chapter asserts that his non-Bayesian time-varying regression provides a powerful tool to analyze the financial markets.

In the second chapter, a non-Bayesian time-varying model is developed by introducing the concept of the degree of market efficiency that varies over time. With new methodologies and a new measure of the *degree of market efficiency*, the author examines whether the U.S. stock market evolves over time. In particular, a time-varying autoregressive (AR) model is employed. His main findings are: (i) the U.S. stock market has evolved over time and the degree of market efficiency has cyclical fluctuations with a considerably long periodicity, from 30 to 40 years; and (ii) the U.S. stock market has been efficient with the exception of four times in his sample period: during the long-recession of 1873-1879; the recession of 1902-1904; the New Deal era; and the recession of 1957-1958 and soon after it.

The third chapter develops a non-Bayesian methodology to analyze the time-varying structure of international linkages and market efficiency in G7 countries. The author considers a non-Bayesian time-varying vector autoregressive (TV-VAR) model, and apply it to estimate the joint degree of market efficiency in the sense of Fama (1970, 1991). His empirical results provide a new perspective that the international linkages and market efficiency change over time and that their behaviors correspond well to historical events of the international financial system.

The fourth chapter examines how the Tokyo and Osaka rice futures markets in prewar Japan were evolving in view of market efficiency. Applying a non-Bayesian time-varying model approach to analyze the famous equation for the futures premium, the author finds that the market efficiency of the two major rice futures markets varied with time. Such time-varying structure of the rice futures markets in prewar Japan corresponds well to historical changes in the Japanese colonial policy and domestic development of railroad system and port facilities.

The last chapter summarizes the empirical results of this dissertation and concludes. First, it asserts the significance of paying attention to unceasing structural changes in real financial markets. In other words, a test of parameter constancy is inevitable and a model with time-varying parameters is preferable when a parametric model is employed to analyze financial markets, say, stock markets or commodity futures markets. Second, market efficiency of such financial markets varies over time with high probability. The findings not only correspond well to historical events in the financial markets but also reconciles the never-ending debates over market efficiency of financial systems.

Acknowledgments

The research underlying this dissertation was conducted between 2009 and 2014. Foremost, I would like to express my deepest gratitude to the dean of the graduate school, professor Shinsuke Nakamura, for his continuous support and helpful advice to my Ph.D. study and research. His guidance helped me in all the time of writing of this dissertation. I could not have imagined receiving a better advice for my Ph.D. study. Besides, I would like to thank my dissertation committee: professors, Takuji Arai, Shinichi Suda, and Taisuke Otsu, for their encouragement, insightful comments, and hard questions. My gratitude also goes to professors, Kazusuke Tsujimura and Shuhei Shiozawa, both of who gave me encouragement and support.

My sincere appreciation also goes to Tatsuma Wada and Akihiko Noda, for their long-time great help for research and writing joint papers. I owe several chapters of this dissertation to them. I could not imagine any words that express my true thanks to them. They kindly gave me their permissions of reprint of the joint works. In particular, Akihiko Noda has always helped me in typesetting papers, surveying literature, fruitful discussions and giving useful advice etc. Also I would like to express my thanks to Kiyotaka Maeda for his permission of reprint of a joint paper with him and Akihiko Noda, to whom I owe a chapter of this dissertation. I had a good time writing the paper with this historian.

I would like to show my greatest appreciation to professors Denzo Kamiya and Hiroaki Osana for mentoring me during my youth. My appreciation also goes to professors Genshiro Kitagawa for enlightening me the first glance of research and Taisuke Otsu for his thoughtful suggestions. I would like to express my deep thanks to my colleagues and long-time companions in research, Yukihiro Ikeda and Masahiro Kawamata, for their unceasing encouragement. Moreover I thank my friend, Shunsuke Sugiyama, for giving me an unforgettable trigger of research.

Last but not the least, I would like to thank my family: my parents Michihiro Ito and Keiko Ito, for giving birth to me at the first place and my wife Eriko, daughters Yuki and Mizuki.

February 2015

Mikio Ito

To my parents, wife and daughters

Abstract

Chapter 1. The first chapter provides an overview of background of this dissertation. In particular, it contains a survey of literature on market efficiency of financial markets, which is a long history of debate. The chapter elucidates the reason why the debate has lasted for goods, discussing the analytical tools employed by the literature. It also shows a new approach of research paying attention to unceasing structural changes in financial markets such as stock markets, commodity futures markets, etc. Accordingly, required is such a new approach that allows us to deal with the markets under unceasing structural changes. The first chapter asserts that his non-Bayesian time-varying regression provides a powerful tool to analyze the financial markets.

Chapter 2. A Non-Bayesian time-varying model is developed by introducing the concept of the degree of market efficiency that varies over time. With new methodologies and a new measure of the *degree of market efficiency*, we examine whether the U.S. stock market evolves over time. In particular, a time-varying autoregressive (AR) model is employed. Using a simulation based technique, we obtain a sequence of critical values of the time varying coefficients under the efficient market hypothesis. Then we detect the period in which the hypothesis does not hold by comparing the estimated coefficient and the critical value for each period. Our main findings are: (i) the U.S. stock market has evolved over

time and the degree of market efficiency has cyclical fluctuations with a considerably long periodicity, from 30 to 40 years; and (ii) the U.S. stock market has been efficient with the exception of four times in our sample period: during the long-recession of 1873-1879; the recession of 1902-1904; the New Deal era; and the recession of 1957-1958 and soon after it.

Chapter 3. This chapter develops a non-Bayesian methodology to analyze the time-varying structure of international linkages and market efficiency in G7 countries. We consider a non-Bayesian time-varying vector autoregressive (TV-VAR) model, and apply it to estimate the joint degree of market efficiency in the sense of Fama (1970, 1991). Our empirical results provide a new perspective that the international linkages and market efficiency change over time and that their behaviors correspond well to historical events of the international financial system.

Chapter 4. This chapter examines how the Tokyo and Osaka rice futures markets in prewar Japan were evolving in view of market efficiency. Applying a non-Bayesian time-varying model approach to analyze the famous equation for the futures premium, we find that the market efficiency of the two major rice futures markets varied with time. Such time-varying structure of the rice futures markets in prewar Japan corresponds well to historical changes in the Japanese colonial policy and domestic development of railroad system and port facilities.

Chapter 5. The last chapter summarizes the empirical results of this dissertation and concludes. First, it asserts the significance of paying attention to unceasing structural changes in real financial markets. In other words, a test of parameter constancy is inevitable and a model with time-varying parameters is preferable when a parametric model is employed to analyze financial markets, say, stock markets or commodity futures markets.

Second, market efficiency of such financial markets varies over time with high probability. The findings not only correspond well to historical events in the financial markets but also reconciles the never-ending debates over market efficiency of financial systems.

Contents

List of Figures

List of Tables

Chapter 1

The Efficient Market Hypothesis and Non-Bayesian Time-Varying Parameter Models

1.1 Introduction[1]

This chapter has two purposes: literature review about market efficiency in financial systems, which this dissertation addresses, and a suggestion for a new approach of study of market efficiency, which it adopts. Although the Efficient Market Hypothesis (EMH) has been a source of a number of papers since Samuelson (1965) introduced the concept of fair game to financial economics, the EMH is still in controversy. The EMH concerns return predictability from past price changes; it mathematically says that the log price of a financial commodity is a martingale if its market is efficient. Since the so-called random walk hypothesis of price of a financial commodity is its sufficient condition, the literature

[1]This chapter is based on Ito (2007) and Ito and Sugiyama (2009).

contains a tremendous number of papers that study and test the latter hypothesis statistically. However many papers have studied and tested the hypothesis, researchers' views on the EMH have changed during the past 50 years. Fama (1970) summarized that, in early researches, many papers supported the EMH; Fama (1991) himself slightly modified his view because of a lot of anomalies reported after Fama (1970). Even in more recent years, Malkiel (2004, ch.11) sticks to his stance that the EMH is almost true; Shiller (2005, ch.10) is skeptical of the hypothesis. The two Nobel Prize winners take opposite stances! This fact represents the current researches sinking into a morass. Today nobody can summarize the debate over the EMH in a couple of lines. The battle between proponents of the EMH and advocates of behavioral finance is still ongoing and, the author believes, will never end.

The author asserts that any effort to verify the EMH by the classical hypothesis testing is in vain and that it is productive to establish a measure of relative market inefficiency as Campbell, Lo, and Mackinlay (1997, ch.1.5.2) point out. One can find recent literature which compares relative performance of market efficiency while there has been little literature that studies the relative market efficiency through time based on time series data. The possibility of time-varying market efficiency has attracted attention recently. Among such studies, there is one that concerns empirical evidences of evolving stock return predictability can be rationalized within a framework called Lo's (2004) adaptive markets hypothesis. This dissertation addresses the similar approach; the author realizes an empirical study measuring a gradually time-varying structure of market inefficiency of a single or multiple financial commodities, such as stock or commodity futures.

This dissertation deals with the EMH in the weak sense of Fama (1970); it implies one can never predict returns of an asset by analyzing past price history. Inefficiency can be interpreted as implying existence of exploitable opportunities. While some degree of serial correlation implies the predictability, it may not imply inefficiency if the predictability is insufficient to overcome transaction costs. The author also notes that delayed trading of some stocks after a shock may impart specious autocorrelation to an index. He considers some derivative of an autocorrelation of stock returns as a good proxy of market inefficiency when one deals with the EMH in the weak sense Fama (1970).

The author's task is to measure the degree of market efficiency derived from the MA or VMA process that represents returns of a financial commodities, supposed to be time-varying. The author adopts a time-varying AR or VAR model, in which the AR or VAR coefficients can vary over time. Applying a state space model, he obtains the estimate the time-varying AR or VAR coefficients by some regression techniques. Then he derives the degrees. It is a new approach that the author presents in this dissertation.

This chapter is organized as follows. Section 1.2 presents a standard definition of the EMH and its significance. Section 1.3 reviews the standard literature on the EMH. It also discusses a limit of studies in the literature and asserts needs of a new approach. Then Section 1.4 suggests a new approach by time-varying parameter models and the advantages of the approach. At the same time, it shows a way to detect in what periods financial markets are efficient in the weak form. Finally, the Section 1.5 is allocated for conclusion.

1.2 The Efficient Market Hypothesis

This chapter briefly presents the theoretical framework of the efficient market hypothesis that is founded by Fama (1970). He focuses his attention on stochastic properties of financial commodities when he deals with efficiency of financial markets. Fama considers that it is significant to regard such uncertainty in financial systems that no arbitrage could exist if the market is efficient. In other words, nobody can predict future price in such a way that he gains profit without risks. From the view point of information, the market efficiency of a financial market says that new information is so quickly and correctly reflected in its current security price that nobody can enjoy arbitrage. In what follows, the author shows this famous idea formally to provide some convenience when he presents his own idea about the degree of market efficiency.

Let r_t denote the return of a financial commodity. That is, it is a log difference $\ln P_t - \ln P_{t-1}$ of the price P_t.[2] What to focus on this chapter is reduced to the condition

$$E\left[r_t \mid \mathcal{I}_{t-1}\right] = 0 \tag{1.1}$$

where $\mathcal{I}_{t-1}$ is an information set available at $t-1$. In other words, the time-t expected capital gain given the information set $\mathcal{I}_{t-1}$ is zero. This condition is equivalent to the following familiar form:

$$\ln P_{t-1} = E\left[\ln P_t \mid \mathcal{I}_{t-1}\right] \tag{1.2}$$

on condition that the information set $\mathcal{I}_{t-1}$ contains the past price P_{t-1}, which is quite natural. As for this form, mathematically the log of the price is a martingale. Economically, both the two conditions assert that there exists no arbitrage, which can be regarded as positive capital gain with certainty. Equations (1.1) and/or (1.2) is called the efficient market hypothesis. This dissertation focuses on the case when the available information set contains the past price history – the weak form of the EMH. [3]

---

[2]The expression $\ln P_t - \ln P_{t-1}$ is a convenient alternative for the discrete form of the security return

$$\frac{P_t - P_{t-1}}{P_{t-1}}.$$

[3]Fama (1970) classifies his market efficiency into three categories considering the available information set: (i) weak form, (ii) semi-strong form and (iii) strong form. The last two forms concern cases when their underlying information sets contain not only the past price history of the market but also information about the so-called fundamentals of the economy.

1.3 Review of Standard Literature on the EMH

There is much literaturxe about the EMH since Fama (1970) published his famous paper (see Andreou, Pittis, and Spanos (2001), Park and Irwin (2007), Yen and Lee (2008) and Lim and Brooks (2011) for recent survey papers on the EMH). Its amount remains huge even if its subject is restricted to the weak-form category. While it could be broken down the literature into some sub-categories, the author provides a brief literature review paying his attention to statistical tools employed to investigate market efficiency empirically. The tools are categorized into some groups: (i) estimating serial correlations, (ii) testing variance ratios (VR) (iii) testing unit roots, (iv) investigating non-linear dependence such as chaos and (v) estimating long memory using an autoregressive fractionally integrated moving average (ARFIMA) model. Considering sizes of literature for each tools, the author reviews (i) through (iii) with concentration. In fact, the literature underlying the tools of (iv) and (v) is much smaller than (i) through (iii).

As for the existence of correlations of security returns, one can recently find less studies employing conventional autocorrelation tests than before. However, the VR test by Lo and MacKinlay (1988) has been accepted as a major tool to detect serial correlation of financial time series because of its easy calculation. The VR test statistic is basically designed to be unity under the null hypothesis of a random walk while many students have improved it theoretically (see Wright (2000) for example). The extensive survey paper by Charles and Darné (2009) provides a review on the many studies relying on this test. While the efficiency of developed markets has attracted less attention than before, found are more works on markets in developing countries and areas such as China and the Middle East (see Kim and Shamsuddin (2008) for instance). Such works provide mixed findings on market efficiency; some are favorable and others are negative. This suggests that simple tests again serial correlation or random walk have a limit in detecting whether a financial market is efficient or not.

The unit root test, say, the augmented Dickey-Fuller (ADF) test, is another primary tools for researchers to examine the weak-form EMH. Employing this test in the earlier literature, most of them have concluded that the log-levels of stock prices are non-stationary and that most financial markets are relatively efficient. Considering the low power of the ADF test and the presence of structural breaks in most financial time series, several researchers investigate efficiency of stock markets. For example, Lean and Smyth (2007) show some random walk evidences using the Lagrange multiplier (LM) unit root tests with one and two structural breaks. A nice literature review by Narayan and Smyth (2007) examines the price indices for stock markets in G7 countries containing a unit root accounting for the structural breaks. However, as Rahman and Saadi (2008) stressed, an existence of a unit root is a necessary condition for the random walk hypothesis but not a sufficient condition. That is, the presence of a unit root in financial time series does not always provide useful information when the market efficiency is concerned.

There is more literature about the EMH focussing another points: nonlinear serial dependence in time series, long memory and chaos etc. For a limited survey for studies about the nonlinear serial dependence using a non-linear model, say, smooth transition autoregressive (STAR) models, see Tsay (2010). One can find a good review on studies on long memory using ARFIMA models in Baillie (1996), which does not always concern the EMH. For tests of nonlinear serial dependence, Hommes and Wagener (2009) provide reviews from the view point of behavioral finance. The author dare say that the studies in these lines contribute little to the research on the EMH; the empirical results vary too much over times, countries and securities etc.

Summing up, the vast literature employing the approaches reviewed above has not let us attain some consensus on the debate over whether financial markets hold the EMH or not. Even Fama's own confidence in the EMH seemed to surge in 1990. In fact, Fama (1970) concludes that stock markets are efficient in the weak form as far as considering the studies before 1970; Fama (1991) keeps his stance while he cannot ignore many anomalies suggesting deviation from the EMH, which are observed in 1990s. In the following subsections, the author shows the way to reconcile such tangled situation

1.4 New Approach to the EMH

What we learn from the vast literature about the EMH is summarized as follows: (i) conventional statistical approaches, parametric or non-parametric, have been unable to detect whether a financial market is efficient or not, (ii) to what extent efficient is a financial market depends on times, countries and financial commodities etc. and (iii) new approaches such as Lo (2004) are emerging recently.

This dissertation seeks a new approach to the EMH considering the above three points. Since Equation (1.2) is a necessary condition for the random walk hypothesis, there is vast literature conducting tests of the hypothesis as already pointed out. However, as will be discussed in detail in Chapter 2, any test regarding Equation (1.2) –including unit-root tests of the stock price– is essentially a test of a joint null hypothesis of market efficiency and the model of the market equilibrium (See Fama (1970, p.386)). Thus, when Equation (1.2) or (1.1) does not hold, there are some possibilities: (i) risk averters dominate the underlying market and risk premium always exits; (ii) the market is not efficient (not in equilibrium); (iii) exogenous positive dividend exists.

The author intends to investigate whether Equation (1.1) or (1.2) holds, he focuses his attention on the time periods for which the efficient market hypothesis is true, rather than estimating auotcorrelation or testing a unit root in $\ln P_t$ for the entire sample fixed. He allows the log price $\ln P_t$ to follow a very flexible process (namely, the time-varying autoregressive process) that permits the autoregressive coefficients to change from time to time, whatever univariate or multivariate series. The reason why the author chooses this approach is that he is more interested in the time-varying nature of the market, or the evolution of the market, which possibly comes from shifts or fluctuations in the parameters or variables that pertain to Equation (1.1) or (1.2). He embarked on investigating time-varying nature of the U.S. stock market by simply estimating time-varying AR coefficients of its returns (See Ito and Sugiyama (2009)).[4]

The idea of this dissertation consists of three parts: (i) any financial time series supposed non-stationary is approximated by a state space model, (ii) the state space model can be estimated by a non-Bayesian econometric method rather than the conventional Kalman filtering method and (iii) the estimates of time-varying parameters suggest in what periods a market is efficient. As will be formally discussed in the technical appendix of Chapter 2, the method of this method has several advantages. First, the time-varying estimates can be regarded as the ones obtained by the rolling window method with the optimal window widths[5]for period by period. Second, thanks to an integrated notation, it avoids such numerical iterative method as the Kalman filter by employing several modern numerical methods using large sparse matrix techniques. Third, the method of this dissertation allows easy calculation of Monte Carlo and bootstrap statistics for time-varying estimates as will be presented in the following Chapters. Fourth, this dissertation proposes a novel degree of

[4]Some researcher have the same interest as the author has (see Kim, Shamsuddin, and Lim (2011) and Lim, Luo, and Kim (2013)). They practically take a rolling method, which shifts a fixed subsample (window width) of data, while the method of this dissertation take a random parameter model, which allows flexible window width.

[5]The word “optimal” reflects the fact that the window widths for each periods are determined through minimizing the sum of squared residuals

market efficiency based on impulse responses. Combining the degree with the non-Bayesian parameter models allows us to show the time-varying nature of efficiency of a market.

1.5 Concluding Remarks

This chapter discusses the limit of preceding works on the EMH by reviewing standard literature on the subject and shows a new approach by using a non-Bayesian time-varying models and a newly proposed degree of market efficiency. Whatever conventional econometric techniques are employed, say, autocorrelation analyses, unit root tests and non-linearity tests etc., most statistical inferences based on a fixed sample fail to investigate whether a financial market is efficient or not. In fact, as the brief literature review here showed, there has never been any consensus on the EMH; one study concludes the efficiency of a market and another study denies it with other sample periods. Such situations would be more tangled by variation of markets and countries etc.

Thus, a new approach is needed to investigate the EMH under the tangled situation. The author proposes a new approach: (i) estimating a non-Bayesian time-varying parameter model, (ii) calculating a time-varying degree of market efficiency based on the time-varying estimates and (iii) conducting statistical inference on the time-varying degree. The method provides in what periods a market to be investigated is efficient in the weak form. Furthermore, the method has several advantages in comparing with the conventional Kalman filtering method or rolling window method.

Chapter 2

The Evolution of Stock Market Efficiency in the U.S.: A Non-Bayesian Time-Varying Model Approach

2.1 Introduction[1]

Over the past forty years, research on the efficient market hypothesis (EMH) has been advanced by a large number of researchers. Utilizing autocorrelation tests and reviewing the preceding studies, Fama's (1970) seminal paper concludes that stock markets are almost always efficient.[2] Two decades later, Fama (1991) addresses the same issue shedding lights on slightly different aspects such as anomalies arising from return seasonality (Ariel (1987,

[1]This chapter is based on Ito, Noda, and Wada (2012).

[2]His survey includes Fama (1965), Fama and Blume (1966), Fama, Fisher, Jensen, and Roll (1969), Jensen (1968), and Blume (1970) among as others.

1990), Harris (1986), and Keim (1983, 1989)) and the firm size effect (Banz (1981), Conrad and Kaul (1988), Chan and Chen (1991), and Fama and French (1992)). After evaluating the vast literature concerning market efficiency, Fama (1991) presents mounting evidence for the predictability of stock returns, indicating that markets are inefficient. Further, Fama (1998) again reviews the empirical work on event studies concerning several long-term return anomalies.

With rigorous statistics, Malkiel (2003) and Schwert (2003) critically examine the validity of the predictability reported in stock returns. While the former generally agrees with Fama's (1970) conclusion, the latter reports a number of anomalies, concurring with Fama's (1991) conclusion. More recently, Yen and Lee (2008) suggest that whether or not the EMH is supported depends upon sample periods: 1960s data are generally affirmative. Yet in the 1990s, this idea receives attacks from the school of behavioral finance.

As Malkiel, Mullainathan, and Stangle (2005) clearly point out, the controversy over the EMH between proponents of the EMH and advocates of behavioral finance is ongoing. The lack of consensus on the EMH is partly attributable to the following: Previous studies have primarily focused on whether the returns of stock follow a random walk process. While the random walk returns support the EMH, returns following a non-random walk process do not necessarily rule out the EMH. In fact, as Nyblom (1989) shows, a process with a single break point can also be a martingale process. Therefore, a rejection of the random walk hypothesis as null may give little information about the efficiency of the stock market.

However, there is an attempt to reconcile the opposing sides. Considering an evolutionary alternative to market efficiency, Lo (2004, 2005) proposes an hypothesis, which he calls the adaptive market hypothesis (AMH). In this new hypothesis, reconciliation is reached in a consistent manner between the proponents of the EMH and behavioral finance: market efficiency is not an object that is statistically tested, but a framework from which most researchers can discuss various perspectives. This framework allows us to explore the possibility that the stock market evolves over time and market efficiency also varies with time. In line with this approach, recent papers, such as Ito and Sugiyama (2009), Kim et al. (2011) and Lim et al. (2013), conclude that market efficiency (or the degree thereof) varies with time. To our knowledge, however, there is no study that directly (employing appropriate measures and methodologies) examines whether or not the stock market evolves over time.

The contribution of this chapter is twofold. First, new convenient methodologies are developed to examine the evolution of the U.S. stock market. The proposed technique is found to be very easy to implement; and applicable to a variety of models with time-varying coefficients. Second, with the concept and measure of the *degree of market efficiency*, it is shown that the U.S. stock market evolves over time and its degree of efficiency changes accordingly. In particular, we consider a non-Bayesian time-varying autoregressive (TV-AR) model, and apply it to the time-varying moving average (TV-MA) model. Then, our main results demonstrate the following: (i) the U.S. stock market evolves over time and its market efficiency changes accordingly. More specifically, the degree of market efficiency in the U.S. stock market has a cyclical fluctuation with a very long periodicity, from 30 to 40 years. (ii) the U.S. stock market is mostly efficient, except for four times (three recessions and the New Deal era) in our sample period.

This chapter is organized as follows. In Section 2.2, the model and new methodologies for our non-Bayesian TV-AR model are presented. The data on the U.S. stock market, together with preliminary unit root test results, are described in Section 2.3. In Section 2.4, we show empirical results that the market efficiency in the U.S. stock market periodically varies over time and that the amplitude of the cycle decreases in the long-run. Section 2.5 concludes. The Appendix provides mathematical and statistical discussions about our new estimation methodologies.

2.2 Model and Methodology

2.2.1 Preliminaries

For a representative household, the first order conditions for its utility maximization problem result in the following Euler equation:

$$p_t = E_t \left[m_{t+1} \left(p_{t+1} + \kappa_{t+1} \right) \right], \tag{2.1}$$

where p_t is the stock price; m_{t+1} is the stochastic discount factor;[3] κ_{t+1} is the dividend; and $E_t[\cdot]$ represents the conditional expectation given the information available at t. Provided m_{t+1} is constant over time and close to 1,[4] together with the condition $E_t[\kappa_{t+1}] = 0$, the stock price should follow a random walk process (or a martingale). In such a case, the expected return is unpredictable since the expected (gross) return is always constant and one, or $E_t[p_{t+1}/p_t] = 1$. As Fama (1970) states, any test regarding Equation (2.1) –including unit-root tests of the stock price– is essentially a test of a joint null hypothesis of market efficiency and the model of the market equilibrium.

Therefore, it is possible that Equation (2.1) does not hold in times when: i) the house-

[3] The stochastic discount factor is defined as

$$m_{t+1} = \delta \frac{u'(C_{t+1})}{u'(C_t)},$$

where δ is the subjective discount factor; $u'(\cdot)$ is the first derivative of a utility function; and C_t is consumption at t.

[4] If the household has a risk-neutral preference,

$$m_{t+1} = \delta$$

for all t.

hold is risk averse and its stochastic discount factor (the marginal rate of substitution of two periods) varies; ii) $E_t[\kappa_{t+1}] \neq 0$; or iii) the market is not efficient (not in equilibrium). While this chapter investigates whether Equation (2.1) holds, our approach here is to find the time periods for which equation (2.1) is true, rather than to test a unit root in p_t for the entire sample period. In doing so, we allow p_t to follow a very flexible process (namely, the time-varying autoregressive process) that permits the autoregressive coefficients to change from time to time. The reason why we choose this approach is that we are more interested in the time-varying nature of the market, or the evolution of the market, which is perhaps due to shifts or fluctuations in the parameters or variables that pertain to Equation (2.1).

Our main idea stems from the observation in Figure 2.1, which exhibits returns on the S&P 500 from January 1871 through June 2012. It is quite evident that the Wall Street Crash in October 1929 and subsequent years, known as the Great Depression, caused an irregular pattern in returns. Due to the fact that outliers often lead to incorrect conclusions in statistical testing (see for example Perron (1989)), one way to carefully investigate whether the efficiency condition Equation (2.1) is satisfied is to exclude the Great Depression period from our sample. Still, a practical question remains: Which observation(s) should be removed from our sample as a result of the Great Depression? Without getting into data mining, we employ a model that allows flexible specifications for the process of the variable, namely, a time-varying coefficients model.

2.2.2 Impulse-Responses and Long-Run Multipliers

From the argument in the previous subsection, our main focus is reduced to the condition

$$E[x_t \mid \mathcal{I}_{t-1}] = 0 \tag{2.2}$$

where $x_t = \ln p_t - \ln p_{t-1}$. In other words, the time-t expected captal gain given the information set available at $t-1$ is zero.

Assuming that x_t is stationary, by the Wold decomposition we write the time-series process of x_t as

$$x_t = \phi(L) u_t,$$

where $\phi(L) = \phi_0 + \phi_1 L + \phi_2 L^2 + \cdots$,[5] with $\phi_0 = 1$; L is the lag operator; $\{u_t\}$ is an i.i.d. process with the mean of zero, and a variance of σ^2. Note that Equation (2.2) holds if and only if $\phi(L) = 1$. Put differently, under the assumption that the discount factor is constant and close (or equal) to 1, EMH is equivalent to $x_t = u_t$. Since the capital gain is random and serially uncorrelated, an unanticipated shock at t affects the capital gain at t, but not in any subsequent periods. This argument leads us to two important features of the capital gain under EMH: First, the long-run multiplier of a shock, i.e.,

$$\phi_\infty \equiv \phi(1) = \phi_0 + \phi_1 + \phi_2 + \cdots,$$

is always 1. Second, impulse-responses quickly disappear after impact. In other words, the effect of an unanticipated shock on the capital gain is short-lived. Therefore, in this chapter, we shall compute both the long-run multipliers and the impulse-response functions of a shock in order to investigate whether the EMH holds.

[5] We also assume $\sum_{i=0}^{\infty} \phi_i^2 < \infty$.

2.2.3 The Non-Bayesian Time-Varying AR Model

It is well known that an invertible MA model of order p is equivalently written as an AR(∞) model, which can be approximated by an AR(q) model:

$$x_t = \alpha_0 + \alpha_1 x_{t-1} + \alpha_2 x_{t-2} + \cdots + \alpha_q x_{t-q} + \varepsilon_t$$

where ε_t is an error term with $E[\varepsilon_t] = 0$, $E[\varepsilon_t^2] = \sigma^2$ and $E[\varepsilon_t \varepsilon_{t-m}] = 0$ for all $m \neq 0$. Unlike the common time-series analysis where α_l's are parameters that are assumed to be constant over time, we assume in our time-varying AR model that they vary with time. Thus, our model to be estimated is

$$x_t = \alpha_{0,t} + \alpha_{1,t} x_{t-1} + \alpha_{2,t} x_{t-2} + \cdots + \alpha_{q,t} x_{t-q} + \varepsilon_t. \tag{2.3}$$

The martingale formulation – introduced by Nyblom (1989) and Hansen (1992) – has substantial flexibility and covers a wide range of parameter dynamics. For instance, it allows for a stochastic process with a single structural break (as Nyblom (1989) points out) and a random walk (as specified in many articles). Nyblom (1989) and Hansen (1992) provide several test statistics for the null hypothesis of the constant parameters against the alternative hypothesis of at least one martingale parameter, in both linear and non-linear models.

We use Hansen's (1992) statistic in this chapter.[6] Among several alternatives of Hansen's test (i.e., martingale formulation), we adopt the assumption that all the AR coefficients, except for the one that corresponds to the intercept term, follow independent random walk

[6]For the detail of test statistics, see Hansen (1992).

processes. Specifically,

$$\alpha_{l,t} = \alpha_{l,t-1} + v_{l,t} \tag{2.4}$$

where $\{v_{l,t}\}$ satisfies $E\left[v_{l,t}\right] = 0,\ \ E\left[v_{l,t}^2\right] = \sigma^2$ and $E\left[v_{l,t} v_{l,t-m}\right] = 0$ for all t, l and $m \neq 0$. This is because, as we will see in Section 2.4, the data in the U.S. stock market are in favor of unstable autoregressive coefficients.

If, instead, the AR coefficients were constant over time (i.e., an ordinary AR model), one could estimate them using the following linear regression model:[7]

$$\boldsymbol{x} = \begin{pmatrix} 1 & {}^t\boldsymbol{z}_1 \\ 1 & {}^t\boldsymbol{z}_2 \\ \vdots & \vdots \\ 1 & {}^t\boldsymbol{z}_T \end{pmatrix} \boldsymbol{\alpha} + \boldsymbol{u} \tag{2.5}$$

where $\boldsymbol{x} = {}^t(x_1\ x_2\ \cdots\ x_T)$, $\boldsymbol{\alpha} = {}^t(\alpha_0\ \alpha_1\ \cdots\ \alpha_q)$, $\boldsymbol{u} = {}^t(u_1\ u_2\ \cdots\ u_T)$, $\boldsymbol{z}_t = {}^t(x_{t-1}\ x_{t-2}\ \cdots\ x_{t-q})$; superscript t denotes a transpose of a vector.

Alternatively, when a non-Bayesian time-varying model with AR coefficients that evolve by random walk processes (specified in the above martingale formulation; Equations (2.3) and (2.4)) is estimated, (2.5) should be modified; and our model can be set in a state space form:

$$\begin{aligned} x_t &= \alpha_0 + {}^t\boldsymbol{z}_t \widetilde{\boldsymbol{\alpha}}_t + u_t \quad (t = 1, 2, \ldots, T) & (2.6) \\ \widetilde{\boldsymbol{\alpha}}_t &= \widetilde{\boldsymbol{\alpha}}_{t-1} + v_t \quad (t = 1, 2, \ldots, T) & (2.7) \end{aligned}$$

where $\widetilde{\boldsymbol{\alpha}}_t = {}^t(\alpha_{1,t}\ \alpha_{2,t}\ \cdots\ \alpha_{q,t})$. Our model is non-Bayesian because it does not necessitate

[7] One could estimate this model by the ordinary least squares (OLS), which is used in this chapter. Because of this, our estimation method can be seen as conditional maximum likelihood estimation, see Hamilton (1994).

the prior distributions of parameters.

Equations (2.6) and (2.7) are called the observation equation and the state equation, respectively. Notice that we assume the intercept α_0 is time-invariant. In order to avoid over-fitting that would occur if the time-varying intercept were employed.

Utilizing the method, of which idea is essentially identical to Maddala and Kim (1998, pp.469–470), we regard equations (2.6) and (2.7) as a system of simultaneous equations:

$$\boldsymbol{x} = Z\boldsymbol{\beta} + \boldsymbol{u} \tag{2.8}$$

$$\boldsymbol{\gamma} = W\boldsymbol{\beta} + \boldsymbol{v} \tag{2.9}$$

where

$$\underbrace{Z}_{T\times(1+qT)} = \begin{pmatrix} 1 & {}^t\boldsymbol{z}_1 & & O \\ \vdots & & \ddots & \\ 1 & O & & {}^t\boldsymbol{z}_T \end{pmatrix}; \quad \underbrace{W}_{qT\times(1+qT)} = \begin{pmatrix} 0 & -I & & & O \\ 0 & I & -I & & \\ \vdots & & \ddots & \ddots & \\ 0 & O & & I & -I \end{pmatrix};$$

I is a $q \times q$ identity matrix; $\boldsymbol{\beta} = {}^t(\alpha_0 \ {}^t\widetilde{\boldsymbol{\alpha}}_1 \ \cdots \ {}^t\widetilde{\boldsymbol{\alpha}}_T)$; $\boldsymbol{\gamma} = {}^t(-{}^t\widetilde{\boldsymbol{\alpha}}_0 \ {}^t\mathbf{0} \ \cdots \ {}^t\mathbf{0})$; and $\boldsymbol{v} = {}^t(v_{1,1} \ \cdots \ v_{q,1} \mid v_{1,2} \ \cdots \ v_{q,2} \mid \cdots \ \cdots \ \cdots \mid v_{1,T} \ \cdots \ v_{q,T})$.

Note that one can regard $\widetilde{\boldsymbol{\alpha}}_0 = {}^t(\alpha_{1,0} \ \alpha_{2,0} \ \cdots \ \alpha_{q,0})$ as a prior vector of non-Bayesian TV-AR coefficients, while α_0 in the observation equation (2.6) represents the constant intercept term. For convenience, we stack equations (2.8) and (2.9), in the system of simultaneous equations.

$$\begin{bmatrix} \boldsymbol{x} \\ \cdots \\ \boldsymbol{\gamma} \end{bmatrix} = \begin{bmatrix} Z \\ \cdots \\ W \end{bmatrix} \boldsymbol{\beta} + \begin{bmatrix} \boldsymbol{u} \\ \cdots \\ \boldsymbol{v} \end{bmatrix} \tag{2.10}$$

We estimate the non-Bayesian TV-AR coefficients by applying least squares techniques, OLS or generalized least squares (GLS), to equation (2.10).[8][9]

Three major advantages of our method over the conventional Kalman smoothing (e.g., Hamilton (1994)) are as follows.

First, our method is quite simple and fast. Unlike the conventional Kalman filtering and following smoothing, no iteration is required. Second, a wide variety of models can be easily dealt with even when the state equations of such models are not represented by simple stochastic difference equations. This is especially beneficial for models with random parameter variations because stochastic constraints or moment conditions are simply put in the state equation. Third, it is evident that our estimation is based on classical regression. Non-Normal errors are allowed for both the observation errors $\boldsymbol{u}$ *and* the state equation errors $\boldsymbol{v}$: not only heteroskedasticity in errors ($\boldsymbol{u}$, $\boldsymbol{v}$, or both), but also correlation between observation errors and state equation errors (correlation between $\boldsymbol{u}$ and $\boldsymbol{v}$) is permitted. The least squares method for Equation (2.10) gives us:

[8]In the Appendix, we present an least square estimator of a state space model that involves a projection technique. Also, the Appendix formally discusses why the least squares technique can be used to estimate the non-Bayesian TV-AR coefficients.

[9]While the regressor of our system in equation (2.10) may be an extremely large matrix of the data exceeding one thousand, this does not pose much of a problem for most recent computers. Furthermore, recent numerical methods for large sparse matrices help our calculation. Indeed, in our case, only several minutes are necessary for estimation of the parameters, impulse responses, and the long-run multipliers. Paige and Saunders (1977) show a numerically efficient method for this framework.

$$\widehat{\boldsymbol{\beta}} = \underset{(1+qT)\times(T+qT)}{Q} \begin{bmatrix} \boldsymbol{x} \\ \cdots \\ \boldsymbol{\gamma} \end{bmatrix}$$

where Q is a $(1 + qT) \times (T + qT)$ matrix.

2.2.4 The Relation to the Moving-Window Method

One of the interesting features of our method is that the smoothed estimate of the time-varying AR coefficient vector α_t can be represented by the weighted average of the observed data and a constant.[10] To highlight the difference between our method and the moving-window method (for example, Kim et al. (2011) and Lim et al. (2013)), let us rewrite the estimated p-th coefficient of the time-varying AR(q) model at time-t is as follows:

$$\hat{\alpha}_{p,t} = \sum_{\tau=1}^{T} \omega_{p,\tau,t} x_\tau + \omega_{p,0,t}, \quad (p = 1, \cdots, q);$$

where $\omega_{p,\tau,t}$'s are weights. Our method allows us to readily compute the weights by the following matrices: By a $(1 + qT) \times (1 + qT)$ permutation matrix P, we have a modified coefficient vector

$$\begin{aligned} \underbrace{\boldsymbol{b}}_{(1+qT)\times 1} &\equiv {}^{t}(\alpha_0 \mid \alpha_{1,1}\ \alpha_{1,2}\ \cdots\ \alpha_{1,T} \mid \alpha_{2,1}\ \cdots\ \alpha_{2,T} \mid \cdots\ \cdots\ \cdots \mid \alpha_{q,1}\ \cdots\ \alpha_{q,T}) \\ &= P\boldsymbol{\beta}. \end{aligned}$$

[10] This is due to the fact that the smoothed estimator is, in its probability limit, the projection of the state vector onto the space spanned by the data and the constant. See Hamilton (1994, ch.13) and Sargent (1987, ch.10), among others.

Then, defining a $(1+qT) \times (T+qT)$ matrix Ω such that:

$$\begin{aligned}
\underbrace{\Omega}_{(1+qT)\times(T+qT)} &= \underbrace{P}_{(1+qT)\times(1+qT)} \underbrace{Q}_{(1+qT)\times(T+qT)} \\
&= \begin{bmatrix} \boldsymbol{r_x} & \boldsymbol{r}_\gamma \\ \Omega_{1,\boldsymbol{x}} & \Omega_{1,\gamma} \\ \Omega_{2,\boldsymbol{x}} & \Omega_{2,\gamma} \\ \vdots & \vdots \\ \Omega_{q,\boldsymbol{x}} & \Omega_{q,\gamma} \end{bmatrix},
\end{aligned}$$

we arrive at the least squares estimator for $\boldsymbol{b}$:

$$\widehat{\boldsymbol{b}} = \Omega \begin{bmatrix} \boldsymbol{x} \\ \dots \\ \gamma \end{bmatrix}.$$

Thus, $\omega_{p,\tau,t}$ is the (t,τ)-th element of $\Omega_{p,\boldsymbol{x}}$; and $\omega_{p,0,t}$ is a linear combination of t-th row of $\Omega_{p,\gamma}$.

It is important to point out that in the case of the moving-window method, the bandwidth (size) of the window for t is fixed and smaller than the whole sample size, T. In contrast, our method is to find the orthogonal projection onto the space spanned by all the information, $(x_1, \dots, x_T)$, as shown in Equation (2.10). Because of this, in fact, our estimator is the minimized mean squared error estimator (MMSE).

(Figure 2.2 here)

Figure 2.2 exhibits estimated weights $\omega_{\tau,t}$ following Equation (2.10). The smoothed estimate for α_{1942}, for example, requires approximately 25 years of observations x_t, from

1930 to 1954, with variant weights on each observation. Another way to interpret this finding is that the band-width for α_{1942} is about 25 years. Since these weights are the result of the least squares estimation, our implicitly-defined band-width can be seen as the "optimally-chosen band-width," in the sense of the MMSE, as opposed to an arbitrarily selected band-width in the moving-window method.

2.2.5 Time-Varying Impulse Responses and Time-Varying Long-Run Multipliers

In this subsection, we present the method that provides time-varying impulse responses and time-varying long-run multipliers. They are calculated from the non-Bayesian TV-AR coefficients in each period utilizing the method described in previous subsections. Statistical inference on our estimates can be conducted by the Monte Carlo technique under the hypothesis that all of the coefficients are zero.

While this idea is quite simple, the following two caveats need special attention: (1) the non-Bayesian TV-AR model that we estimate is only an approximation of the real data generating process, which may be a very complex process; and (2) we consider the estimated stationary AR(q) model index by each period t as a local approximation of the underlying complex process.

First, let us consider a time-varying AR(q) model. To find the appropriate order of q, we employ the Schwartz Bayesian information criterion (SBIC). After the order, q, is selected, we estimate the non-Bayesian TV-AR(q) by the method presented in Section 2.2.3.

Second, from the non-Bayesian TV-AR(q) model, the TV-MA(∞) model is derived:

$$x_t = u_t + \phi_{1,t} u_{t-1} + \phi_{2,t} u_{t-2} + \cdots .$$

The coefficients of the TV-MA(∞) can be readily computed from the estimated TV-AR(q) in the following manner: With the estimates ${}^t(\widehat{\alpha}_{1,t}, \ldots, \widehat{\alpha}_{q,t})$ and $\widehat{\alpha}_0$, and Proposition 2.4 of Lütkepohl (2005, Section 2.3.2), the TV-MA(∞) coefficients are found by using:

$$\widehat{\phi}_{0,t} = 1, \quad \widehat{\phi}_{k,t} = \sum_{j=1}^{k} \widehat{\phi}_{k-j,t}\widehat{\alpha}_{j,t}.$$

The time-varying long-run multiplier associated with the time-varying coefficients can be computed by, (see Equation (2.3.26) in Lütkepohl (2005, p.56)),

$$\widehat{\phi}_{\infty,t} = \frac{1}{1 - \widehat{\alpha}_{1,t} - \widehat{\alpha}_{2,t} - \cdots - \widehat{\alpha}_{q,t}}. \tag{2.11}$$

We pay special attention to the time-varying long-run multiplier because it measures the deviation from efficient market. Note that in the case of efficient market where $\phi_1 = \phi_2 = \cdots = 0$ and $\alpha_1 = \alpha_2 = \cdots = \alpha_q = 0$, the long-run multiplier $\phi_{\infty,t}$ becomes one; otherwise, $\phi_{\infty,t}$ deviates from one. Hence, we consider $\phi_{\infty,t}$ to be a measure of the degree of market efficiency.

In order to conduct statistical inference on our time-varying impulse response, we build a set of Monte Carlo samples of the time-varying AR estimates under the hypothesis that all of the TV-AR coefficients are zero. That is, we derive a simulated distribution of the time-varying AR coefficients estimated if the stock return process were generated under the efficient market hypothesis. In addition, we can compute the corresponding distributions of the impulse response and long-run multipliers based on the simulated one. Finally, we conduct statistical inference on our estimates by using a sequence of critical values of time varying coefficients under the efficient market hypothesis derived from such simulated distributions.

2.3 Data

Monthly returns for the S&P500 stock price index from January 1871 through December 2012 (obtained from Robert Shiller's website) are utilized. In practice, we compute the first-difference of the logarithm of the S&P500 stock price index as the returns. Figure 2.1 presents time series plots of the returns for the S&P500.

(Figure 2.1 here)

We check whether the variables satisfy the stationarity condition, we apply the ADF-GLS test of Elliott, Rothenberg, and Stock (1996). Together with the procedure proposed by Ng and Perron (2001), this unit root test is robust against size-distortions. The results of the ADF-GLS test along with descriptive statistics of the data are presented in Table 2.1: The ADF-GLS test rejects the null hypothesis that the variable contains a unit root at conventional significance levels.[11]

(Table 2.1 here)

2.4 Empirical Results

In this section, we report three sets of results. They are: i) our preliminary estimation together with Hansen's (1992) test which confirm that the parameters in the standard AR model are not constant over the sample period; ii) estimation of the non-Bayesian TV-AR model which reveals the validity of our model; and iii) the impulse-responses and the long-run multiplier that suggest market efficiency for a limited period of time.

[11] For selecting the optimal lag length, we employ the Modified Bayesian Information Criterion (MBIC) instead of the Modified Akaike Information Criterion (MAIC). This is because, from the estimated coefficient of the detrended series, $\widehat{\psi}$, we do not find the possibility of size-distortions (see Elliott et al. (1996); Ng and Perron (2001)).

2.4.1 Preliminary Estimation and Parameter Constancy Test

Assuming a standard AR(q) model with constant parameters, we utilize the SBIC of Schwarz (1978) to select the lag-order, q. As a result, $q = 2$ – the second order autoregressive model – is obtained.[12]

(Table 2.2 here)

Our estimation result for an AR(2) model with the whole sample is summarized in Table 2.2: all AR estimates are statistically significant at the 1% level. It is notable that the first-order autoregressive estimate is about 0.31 (and the second one is about -0.08). This implies that approximately 10% of an unanticipated shock to the average stock return during any month will remain in the average stock returns for two months later.

Are the AR coefficients constant over the sample period? One approach that we consider useful is to apply a test of the parameter constancy. As presented in Table 2.2 (the entry below L_C), Hansen's (1992) test reject the null hypothesis of constant parameters at the 1% significance level (The asymptotic critical value at the 1% significance level is 1.60). Having found non-constant parameters in the AR(q) model, we move forward to focus on the time-varying AR model in order to see whether gradual changes occur in the U.S. stock market.

2.4.2 Non-Bayesian TV-AR Estimation

Given the fact that the test of the parameter constancy rejects the null hypothesis against the alternative hypothesis of the AR parameters following the random walk process, our non-Bayesian time-varying estimation method is carried out to estimate our TV-AR(2)

[12] The heteroskedasticity and autocorrelation consistent (HAC) covariance matrix estimator of Newey and West (1987, 1994) is used.

model. Because of the properties we discussed in Section 2.2.4, their method is shown to have a particular advantage over the simple moving window method that assumes a fixed width to compute AR coefficients or the correlation coefficient.[13] With optimally selected window widths, the coefficients of the TV-AR(2) model are computed.[14]

(Figure 2.2 here)

Figure 2.2 presents the weight for $t = 851$ (February, 1942). As is discussed in Section 2.2.4, the estimate utilizes a wide range of observations. In this case, the smoothed estimate for $t = 851$ requires the data points of 170 months.

(Figure 2.3 here)

Demonstrated in Figure 2.3 (solid lines), the estimated AR coefficients are very unstable over time.[15] As a statistical inference, we provide significance bands in Figure 2.3. Due to the fact that our method is based on least squares for subsamples (i.e., widths), note that our estimates may suffer from a downward bias (see for example, Andrews (1993a)). Taking into account such a possibility, we construct significance bands for the estimates as follows. For the data generating process assuming $\alpha_{1,t} = \alpha_{2,t} = 0$ for all t, the TV-AR(2) model is estimated. Repeating this process 5000 times and tabulating $\widehat{\alpha}_{1,t}$ and $\widehat{\alpha}_{2,t}$ for each $t = 1, \ldots, T$, we plot the 99% upper and lower limits of the estimates. Therefore, Figure 2.3 shows that the AR(1) coefficient is significant most of the time, while the AR(2) coefficient is not.

From the estimated AR(1) coefficient, remarkably, the market crash and the following Great Depression did not cause a great deal of deviation of the AR(1) coefficient from its

[13] This point distinguishes our work from previous studies, such as Kim et al. (2011) and Lim et al. (2013).

[14] We confirm that our model is locally stationary – the stationary AR(2) process for all t – by checking the roots of the equation $1 - \alpha_{1,t}z - \alpha_{2,t}z^2 = 0$. All of the roots lie outside the unit circle.

[15] Yet, Ljung and Box's (1978) test for the residuals do not reject the null hypothesis of autocorrelation. To estimate the TV-AR model, the HAC estimator of Newey and West (1987, 1994) is employed.

historical average: In fact, in the late 1980s, a slightly larger magnitude of deviation can be seen. However, care must be taken in interpreting the estimates of the AR coefficients. Our ultimate goal in this chapter is to compute the long-run multipliers, defined in Equation (2.11), and the time-varying impulse-responses to see whether there is evidence of market efficiency. The next subsection presents these two measures.

2.4.3 Time-Varying Impulse Responses and Time-Varying Long-Run Multipliers

Figure 2.4 exhibits the time-varying impulse-responses.

(Figure 2.4 here)

Notably, the time path of an exogenous shock's effect on return varies widely with time. For example, in December 1919, when the estimated time-varying AR(1) coefficient reaches its whole sample minimum, only less than 20% of the shock to the average stock return remains two months after the shock (Figure 2.4, bottom left); whereas in November 1987, when the estimated AR(1) coefficient reaches its whole sample maximum, more than 40% of the shock is preserved two months after impact (Figure 2.4, bottom right).

The other measure of market efficiency, the time-varying long-run multiplier – the degree of market efficiency – is demonstrated in Figure 2.5.

(Figure 2.5 here)

There are two important observations to be pointed out.

First, despite the extraordinary magnitudes of the Great Crash of 1929 and Financial Crisis of 2008 (as displayed in Figure 2.1), the degrees of market efficiency for such periods are not shown to be outliers.

Second, after the 1930s, the degree of market efficiency tends to rise during periods of expansion, and tends to decline during periods of contraction.[16] Since the degree of market efficiency that is higher than 1 means "more inefficient," expansions are associated with more inefficient markets and vice versa. Put differently, shocks that affect the return on stocks linger for quite while; but for depressed markets (which are often led by a crash in the stock market), shocks quickly disappear. One possible interpretation of this phenomenon is that individuals become more irrational during expansions while they become more rational during contractions. This argument is, however, not very convincing because not all recessions/expansions exhibit the pattern described above.

With that being said, the degree of market efficiency is greater than 1 for the entire sample period. Does this make our argument invalid? For the same reason we compute them in the previous subsection, significance bands for the degree of market efficiency are provided. By doing so, we arrive at the conclusion that the stock market is efficient for the most of our sample period at the 1% significance level. With such bands, we are able to find at least four clearly inefficient markets in our sample. Out of four inefficient markets, theree of them have their highest degree of market efficiency (i.e., the largest deviation from market efficiency) during recessions. The first one appears during the long-depression (1873-1879), following the financial panic of 1873. Note that this is the longest recession NBER has ever recorded (65 months). The second inefficiency take place during the recession of 1902-1904, which follows the panic of 1901. The third peak is seen in August 1958 when a short, but very severe recession just passes its "trough." A study by Perron and Wada (2009) reveals that the cyclical component of the U.S. post war real GDP, after taking into account the structural break in the slope of the trend in 1973, reaches its

[16]To see the relationship between overall economic conditions and the degree of market efficiency, we provide the NBER business cycle dates in Figure 2.5 (the shaded areas are recessions).

lowest point in the 1957-1958 recession. This result is in accordance with the trend-cycle decomposition by the Beveridge-Nelson decomposition (Beveridge and Nelson (1981); see Morley, Nelson, and Zivot (2003)). One exception took place in the expansion between 1933 and 1937, the aftermath of the Great Depression; and the economy was recovering only due to the aggressive fiscal policy called the New Deal.

One finding stands out: Not all recessions (or expansions) create inefficient markets, although some turning points of business cycles and those of the degree of market efficiency seem to be related. Therefore, we can conclude that the degree of market efficiency does not fluctuate as often as macrovariables such as GDP and consumption. From the view point of the spectral analysis, Figure 2.6 confirms this conclusion.[17]

(Figure 2.6 here)

The power spectrum of the estimated degree of efficiency has a peak and its most power in lower frequencies than in standard business cycle frequencies (periodicity corresponding to 72- 384 months; shaded in Figure 2.6. See also Baxter and King (1999).), indicating that the degree of market efficiency has a very long periodicity. Thus, market inefficiency emerges only infrequently.

However, note that the evidence against efficient markets is found for extraordinary times, but not for times of financial panics or bubbles. What is unclear is that the deviations from efficient market are attributed to either irrational behaviors of market participants or drastic changes in the individuals' stochastic discount factor, which is the function of: the consumption growth rate, the degree of risk aversion, and the subjective discount factor.

[17]The power spectrum is estimated by utilizing the Bartlett window and the Quadratic window in order for the estimate to be consistent (See, for example, Brockwell and Davis (1991)). To select the bandwidth, Andrews's (1991) method that are designed to consistently estimate the spectral density function at frequency zero is employed. We also examine the consistent estimator proposed by Newey and West (1994) that is also desinged to estimate the zero-frequency spectral density consistently. From Figure 2.6, all estimators of the spectral density exhibit qualitative similarity.

The latter is as likely as the former, because all of the factors that affect the stochastic discount factor may have been altered during such times.

2.5 Concluding Remarks

Focusing on market efficiency that may vary with time, we develop a non-Bayesian time-varying model to examine whether or not the U.S. stock market has evolved over time. In particular, the non-Bayesian time-varying AR (TV-AR) model is applied by taking into account various possibilities, namely, structural changes, regime shifts, and gradual changes. In addition, a new measure of the degree of market efficiency is introduced and estimated. With a new and convenient technique, it is found that the U.S. stock market evolves slowly over time: Our estimated power spectrum indicates the periodicity of the degree of market efficiency is 30 to 40 years. After careful consideration based on statistical inferences, the degree of market efficiency is found to be in favor of efficient markets for the vast majority of our sample period. This is in line with our impulse-response analysis which uncovers that any shock to stock return quickly disappears most of the time – indicating that market is generally efficient over the sample period. However, a little evidence for inefficient markets is also discovered. They are, during: (i) the longest recession defined by NBER (1873-1879); (ii) the 1902-1904 recession; (iii) the New Deal era; and (iv) just after the very severe 1957-1958 recession. These results suggest that the deviation from efficient markets occurs in extraordinary times, but not during times of panic (or bubble). This, in turn, raises another question of whether the market is inefficient due to irrationality or is efficient but the individuals' stochastic discount factor changes dramatically. While this unanswered question opens up new avenues of research for market efficiency, we believe that our approach – allowing the possibility of the evolving market and introducing the concept of the degree of market efficiency – will provide researchers with a more in-depth view of market efficiency.

Figure 2.1: The Returns on S&P500

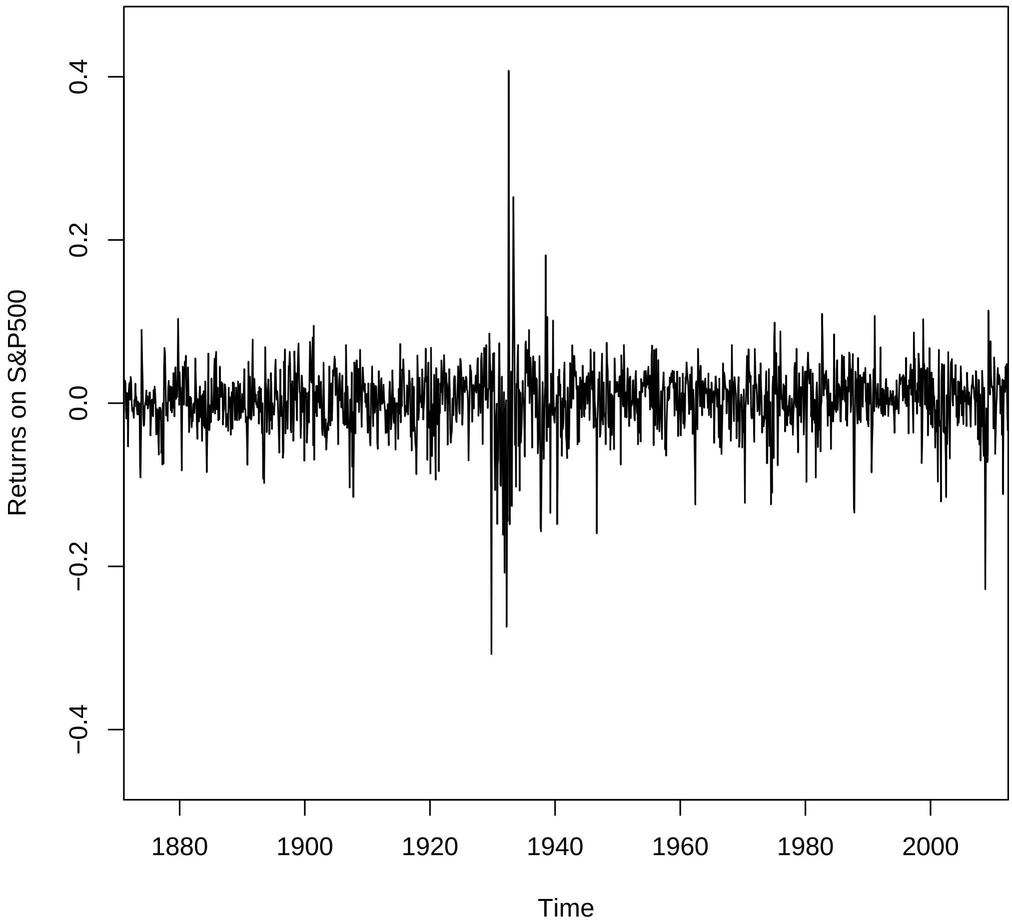


Figure 2.2: Optimal Weights for the Smoother

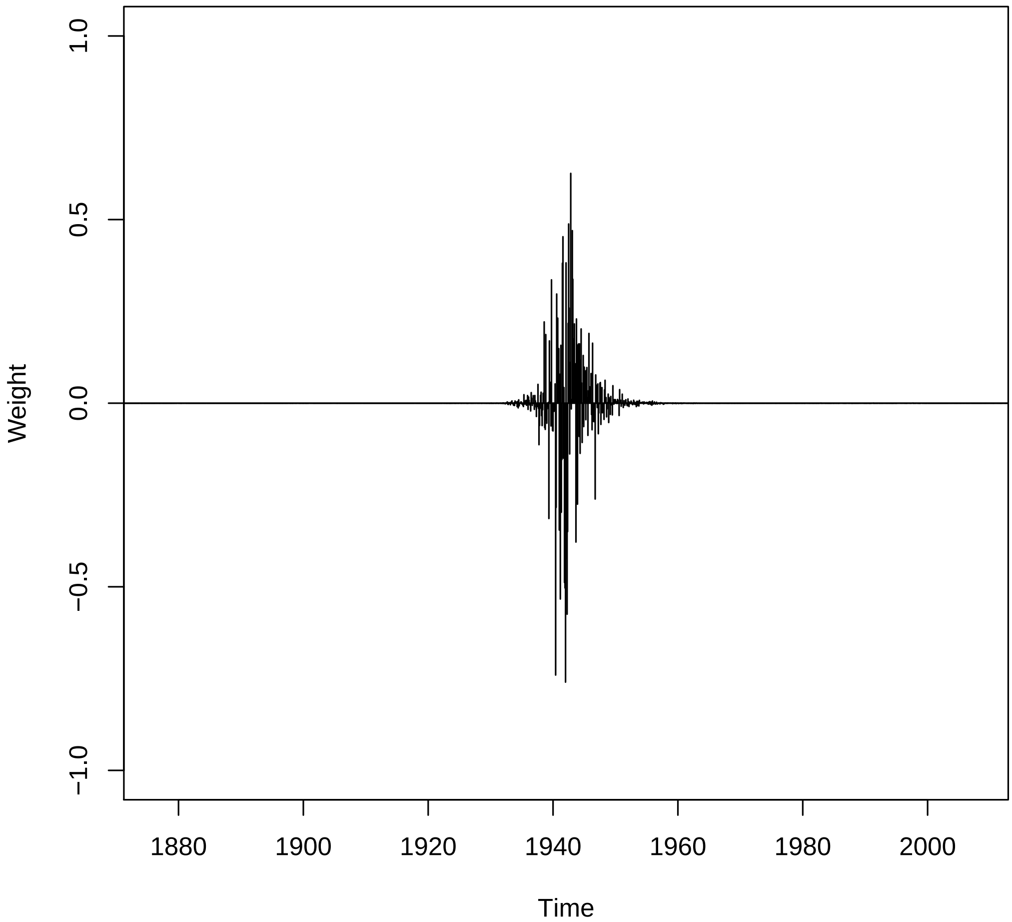


Table 2.1: Descriptive Statistics and Unit Root Test

Mean	SD	Min	Max	ADF-GLS	Lag	$\hat{\psi}$	$\mathcal{N}$
0.0034	0.0410	-0.3075	0.4075	-30.1543	0	0.3033	1703

Notes:

(1) "ADF-GLS" denotes the ADF-GLS test statistics, "Lag" denotes the lag order selected by the MBIC, and "$\hat{\psi}$" denotes the coefficients vector in the GLS detrended series (see equation (6) in Ng and Perron (2001)).

(2) In computing the ADF-GLS test, a model with a time trend and a constant is assumed. The critical value at the 1% significance level for the ADF-GLS test is "-3.42."

(3) "$\mathcal{N}$" denotes the number of observations.

(4) R version 3.0.2 was used to compute the statistics.

Table 2.2: Preliminary Estimation and Parameter Constancy Test

Constant	R_{t-1}	R_{t-2}	$\bar{R}^2$	L_C
0.0026	0.3089	-0.0808	0.0860	53.4309
[0.0010]	[0.0281]	[0.0308]		

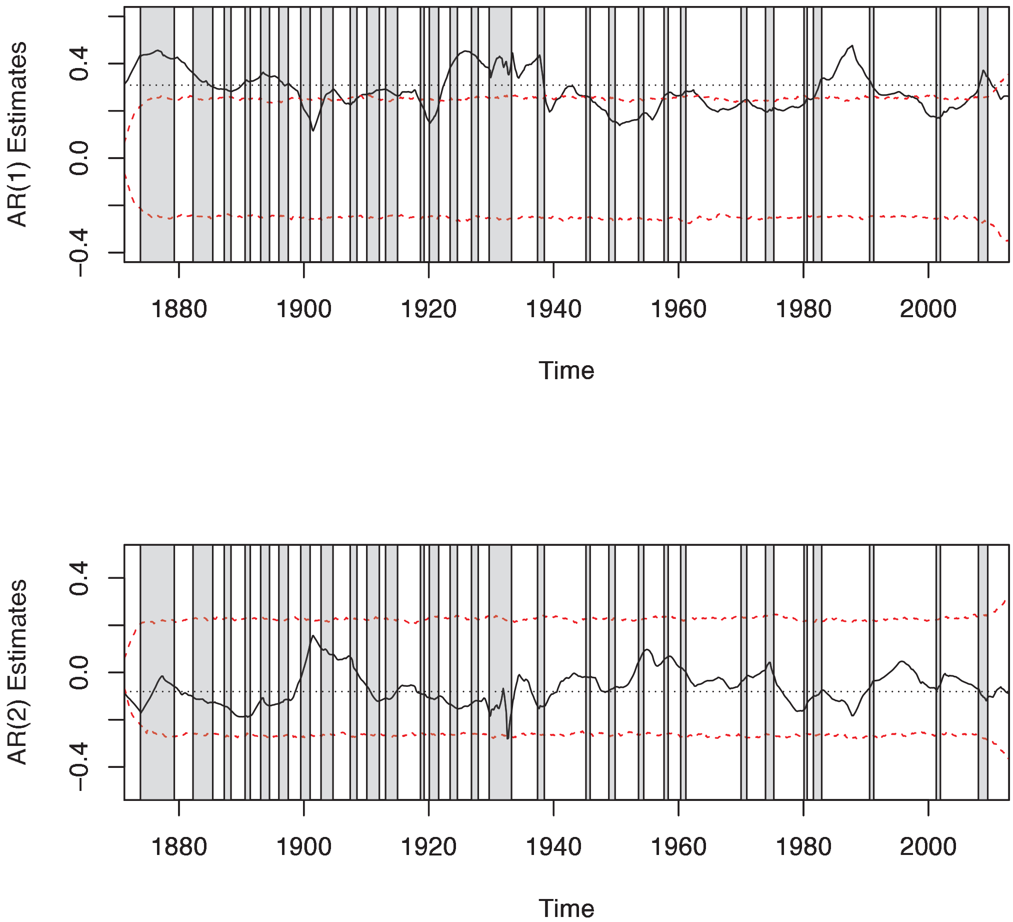
Notes:

(1) "R_{t-1}," "R_{t-2}," "$\bar{R}^2$," and "L_C" denote the AR(1) estimate, the AR(2) estimate, the adjusted R^2, and Hansen's (1992) joint L statistic with variance, respectively.

(2) Newey and West's (1987) robust standard errors are in brackets.

(3) R version 3.0.2 was used to compute the estimates.

Figure 2.3: Non-Bayesian TV-AR Estimation



Notes:

(1) The dashed red lines represent the 99% critical values of the TV-AR estimates in case of efficient market.

(2) We run 5000 times Monte Carlo sampling to calculate the critical values.

(3) The dotted lines also represent the TV-AR estimates with whole sample, respectively.

(4) The shade areas are recessions reported by the NBER business cycle dates.

(5) R version 3.0.2 was used to compute the estimates.

Figure 2.4: Time-Varying Impulse Responses

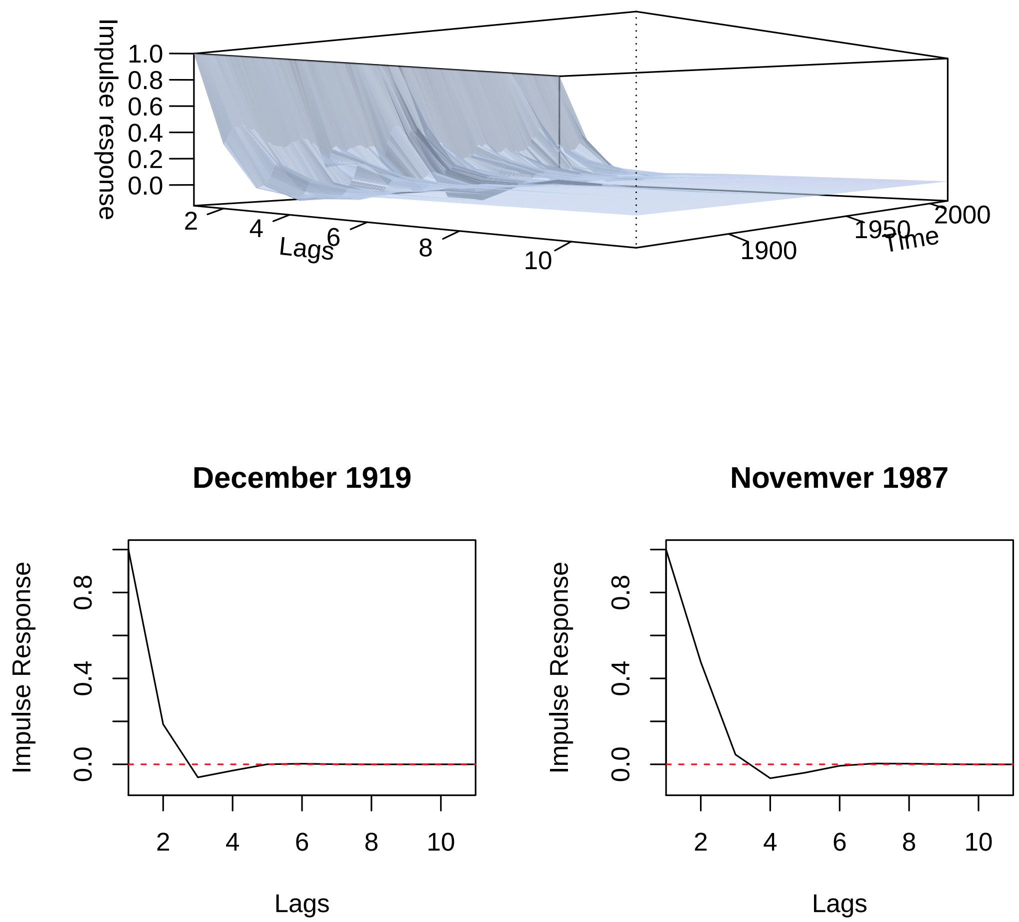
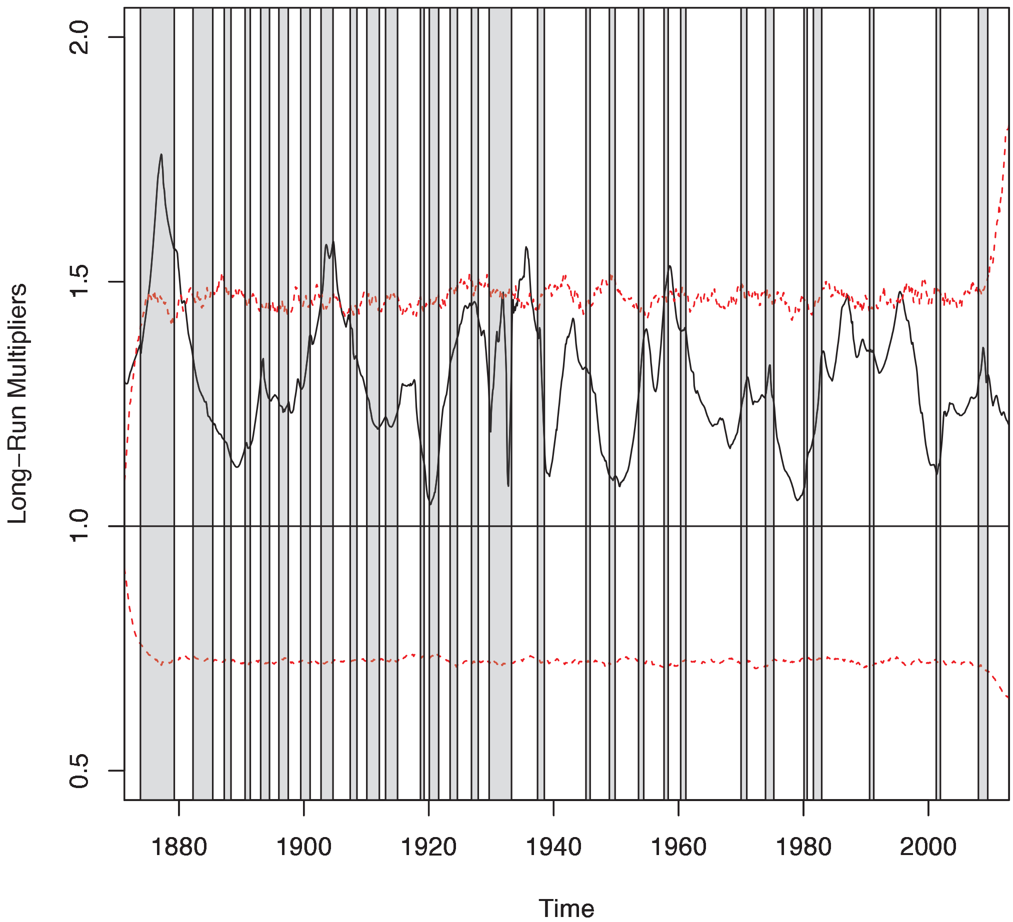


Figure 2.5: Time-Varying Long-Run Multipliers



Notes:

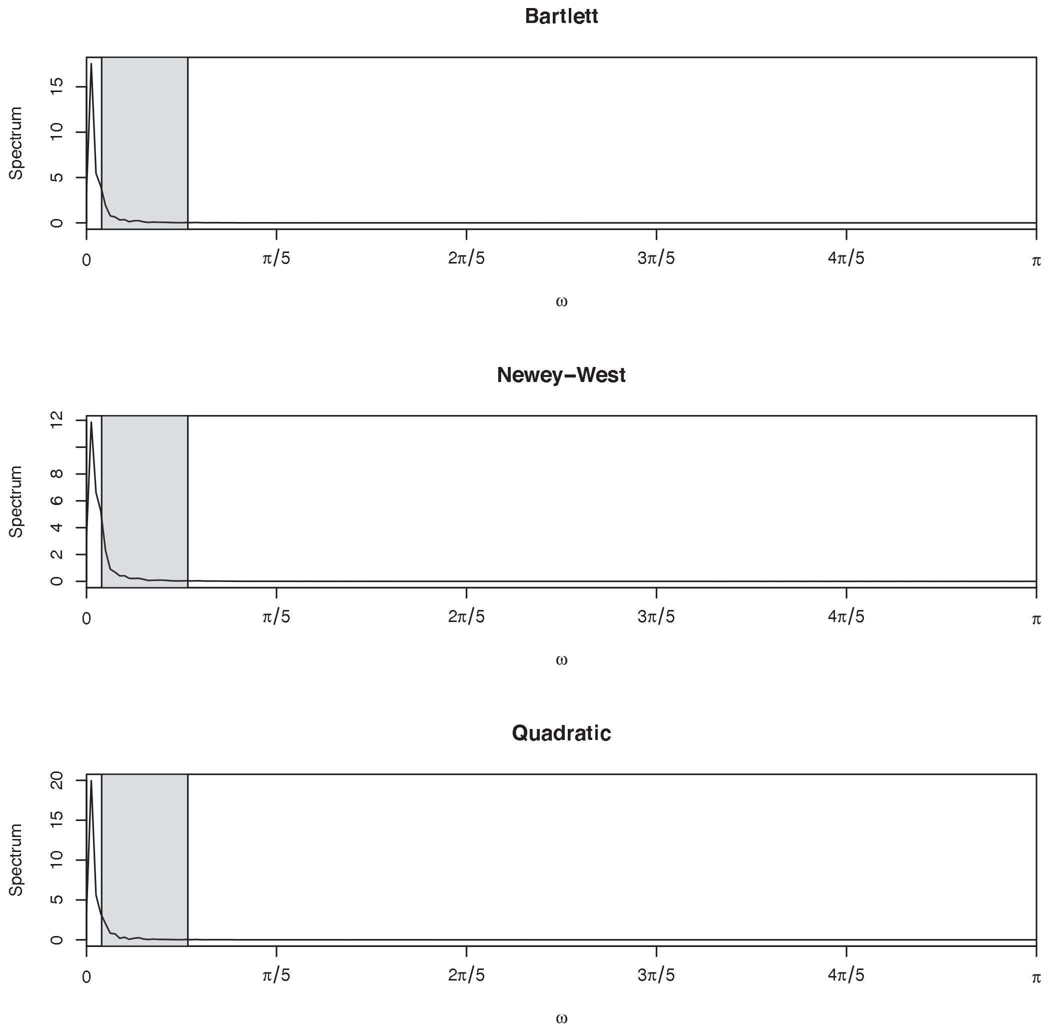
(1) The dashed red lines represent the 99% critcal values of the time-varying long-run multipliers in case of efficient market.

(2) We run 5000 times Monte Carlo sampling to calculate the critical values.

(3) The shade areas are recessions reported by the NBER business cycle dates.

(4) R version 3.0.2 was used to compute the estimates.

Figure 2.6: Power Spectrum Analysis



Notes:

(1) "Bartlett," "Newey-West," and "Quadratic" denote the type of the window kernel to estimate the power spectrums, respectively.

(2) The shade area is the standard business cycle frequencies.

(3) R version 3.0.2 was used to compute the statistics.

2.A Technical Appendix

2.A.1 State Space Regression Model

This section shows a framework allowing us to deal with a state space models as a linear regression model with time-varying parameters. Such a framework provides an alternative of the Kalman smoother. The alternative estimator takes a crucial role in this chapter. The idea is so simple that we reduce a state space model to more general regression models, usually called random parameter regression models. Our framework covers wide range of state space models, which have random parameters with occasional jumps.

We first reexamine the structure of the state space model as a linear regression model, Equations (2.8) and (2.9), which we call state space regression model. We regard Equation (2.9) as an ordinary linear regression model: we regard Equation (2.9) as the one specifying a time-varying structure of parameters. Equation (2.9) is considered as an difference equation with disturbance terms.

On the other hand, when we stand on the regression theory, Equation (2.9) illustrates how the parameters are randomized. Given some prior vector $\boldsymbol{\gamma}$, we rewrite Equation (2.9) as follows:

$$\boldsymbol{\gamma} = W\boldsymbol{\beta} + \boldsymbol{v}.$$

To simplify the discussion, we temporally suppose that the matrix W is non-singular and that we have the following representation of Equation (2.9):

$$\boldsymbol{\beta} = W^{-1}\boldsymbol{\gamma} - W^{-1}\boldsymbol{v}. \tag{2.A.1}$$

We regard the first term of RHS in Equation (2.A.1) as the expected values of the *randomized* parameters and the second one the random effects of the disturbance terms. Notice that the vital supposition in our discussion is not the invertibility of W but the existence of such a decomposition of randomized parameters.

Considering the case where some parameters might not be random or the one where the disturbance terms affect the parameters in degenerated ways, we generalize Equation (2.9) in the following simple form:

$$\boldsymbol{\beta} = \bar{\boldsymbol{\beta}} + D\boldsymbol{w}, \tag{2.A.2}$$

where D is a matrix with the size of $N \times \ell, \ell \leq N, N := nT$ and

$$\boldsymbol{w} \overset{iid}{\sim} \mathcal{WS}(\mathbf{0}, \Sigma_{\boldsymbol{w}}),$$

where $\boldsymbol{w}$ might have smaller dimension than that of $\boldsymbol{v}$ and $\mathcal{WS}(\mathbf{0}, \Sigma_{\boldsymbol{w}})$ is any distribution depending only on mean $\mathbf{0}$ and variance $\Sigma_{\boldsymbol{w}}$ such as the normal distribution. We will use this notation to represent a distribution with wider sense assumptions.

We assumes that the rank of D is $\ell > 0$ and that $\bar{\boldsymbol{\beta}}$ is known. Note that when $\text{rank}D = \ell = 0$, Equation (2.A.2) has no parameter to estimate. Notice that the matrix D reflects the time-varying structure in place of W.

This class of random parameter regression models covers a conventional state space model:

$$\boldsymbol{y}_t = X_t\boldsymbol{\beta}_t + \boldsymbol{u}_t, \quad \boldsymbol{u}_t \overset{iid}{\sim} \mathcal{WS}(\mathbf{0}, R_t)$$

$$\boldsymbol{\beta}_{t+1} = \Phi_{t+1,t}\boldsymbol{\beta}_t + G_t\boldsymbol{w}_t, \quad \boldsymbol{w}_t \overset{iid}{\sim} \mathcal{WS}(\mathbf{0}, \Sigma_{\boldsymbol{w}}), \tag{2.A.3}$$

where the matrix G_t in Equation (2.A.3) has the size $m \times \ell(\ell \leq m)$ and its rank is ℓ. It is natural to assume that $\Sigma_{\boldsymbol{w}} = I$ because this assumption implies $\text{cov}(G_t\boldsymbol{w}_t) = G_tG_t'$ and thus G_t has all information about the covariance matrix of the state equation. We make additional remark that the class of the random parameter regression model defined above covers quite wide range of linear models such as linear models for panel data.

2.A.2 Random Parameter Regression

In this section, we demonstrate that OLS and GLS assure the MMSLE estimator in the random parameter regression model defined above. To simplify our notations and discussion, we present the random parameter regression model as follows. Let M denote the number of unknown parameters and N the dimension of observation vector $\boldsymbol{y}$ is N.

$$\boldsymbol{y} = X\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \tag{2.A.4}$$

and

$$\bar{\boldsymbol{\beta}} = \boldsymbol{\beta} - D\boldsymbol{w}, \tag{2.A.5}$$

where D is a matrix known with the size of $M \times \ell,\ \ell \leq M$,

$$\boldsymbol{\varepsilon} \overset{iid}{\sim} \mathcal{WS}(\mathbf{0}, \Sigma_{\boldsymbol{\varepsilon}}) \text{ and } \boldsymbol{w} \overset{iid}{\sim} \mathcal{WS}(\mathbf{0}, \Sigma_{\boldsymbol{w}}),$$

and $\bar{\boldsymbol{\beta}}$ is known.

We supose some regularity conditions for the least square estimation.

Assumption 1

$$\textit{rank}\ D = \ell > 0,$$

and there is a generalized inverse D^{-} of D such that

$$D^{-}D = I.$$

We stack Equations (2.A.4) and (2.A.5) multiplied by D^-.

$$\mathcal{Y} = \mathcal{X}\boldsymbol{\beta} + \boldsymbol{\xi}, \tag{2.A.6}$$

for

$$\begin{bmatrix} D^-\bar{\boldsymbol{\beta}} \\ \boldsymbol{y} \end{bmatrix} = \begin{bmatrix} D^- \\ X \end{bmatrix} \boldsymbol{\beta} + \begin{bmatrix} -\boldsymbol{w} \\ \boldsymbol{\varepsilon} \end{bmatrix}.$$

Equation (2.A.6) can be written as

$$(\mathcal{Y} - \mathcal{X}\boldsymbol{\beta}) \sim \mathcal{WS}(\mathbf{0}, \Sigma_{\boldsymbol{\xi}}), \tag{2.A.7}$$

where

$$\Sigma_{\boldsymbol{\xi}} = \begin{pmatrix} \Sigma_{\boldsymbol{w}} & \boldsymbol{O} \\ \boldsymbol{O} & \Sigma_{\boldsymbol{\varepsilon}} \end{pmatrix}.$$

Table 2.A.1 summarizes the dimensions of the vectors and matrices.

Table 2.A.1: Dimensions of Random Parameter Regression Model

vector		matrix	
$\boldsymbol{y}$	N	$\mathcal{X}$	$(N+\ell) \times M$
$\boldsymbol{\beta}$	M	$\Sigma_{\boldsymbol{w}}$	$\ell \times \ell$
$\boldsymbol{\varepsilon}$	N	$\Sigma_{\boldsymbol{\varepsilon}}$	$N \times N$
$\boldsymbol{w}$	ℓ	$\Sigma_{\boldsymbol{\xi}}$	$(N+\ell) \times (N+\ell)$
$\mathcal{Y}$	$N+\ell$	D	$M \times \ell$

The above framework enables us to deal with the random parameter regression theory as we do with the familiar regression theory. In the usual regression theory, an important estimator is the weighted least squares estimator (WLSE) of $\boldsymbol{\beta}$.

Definition 1 (WLSE (weighted least squares estimator)) $\boldsymbol{b}$ *is the WLSE of* $\boldsymbol{\beta}$ *for Equation (2.A.7) if and only if*

$$(\mathcal{Y} - \mathcal{X}\hat{\boldsymbol{\beta}})'\Sigma_{\boldsymbol{\xi}}^{-1}(\mathcal{Y} - \mathcal{X}\hat{\boldsymbol{\beta}}),$$

is minimized when $\boldsymbol{b} = \hat{\boldsymbol{\beta}}$*, that is,*

$$\boldsymbol{b} = \mathrm{argmin}\{(\mathcal{Y} - \mathcal{X}\hat{\boldsymbol{\beta}})'\Sigma_{\boldsymbol{\xi}}^{-1}(\mathcal{Y} - \mathcal{X}\hat{\boldsymbol{\beta}}) : \hat{\boldsymbol{\beta}} \in \mathbf{R}^M\}$$

where $\mathcal{Y}$ *and* $\mathcal{X}$ *are given and* $\Sigma_{\boldsymbol{\xi}}$ *is known.*

Note that WLSE $\boldsymbol{b}$ does not always hold unbiasedness and consistency when $\mathcal{Y}$ and any column of $\mathcal{X}$ are statistically correlated.

Chapter 3

International Stock Market Efficiency: A Non-Bayesian Time-Varying Model Approach

3.1 Introduction[1]

How are the world's stock markets linked to each other? Is it possible that a boom in one country's stock market could be accompanied by a boom in another country's stock market, while a drop in equity prices occurs simultaneously in many countries' stock markets? Recent technological progress in the financial sector has enabled information and funds to be rapidly transmitted, thereby providing investors with many opportunities in world stock markets, rather than in just a local stock market. Also, economic activities of firms have become more international, mostly thanks to low transportation costs stemming from technological progress. The answer to the aforementioned (second) question seems to be

[1]This chapter is based on Ito, Noda, and Wada (2014d).

affirmative in such an environment.

If so, what are implications of the international linkage for Fama's (1970) market efficiency, which requires zero unexploited excess profit given the information available to the public? For example, consider the case where one nation's stock market is not efficient, but that stock market is jointly efficient with another country's stock market. One possible interpretation for this phenomenon would be that investors have opportunities to invest in two markets – as a result of portfolio diversification – and arbitrages occur across these two markets. Hence, it is conceivable that the "joint efficiency" among several markets appears when those markets are highly integrated.[2]

As we shall see in the next section, which reviews the literature, there are mainly two types of papers written in the 1990s and early 2000s. The first approach employs a vector autoregressive (VAR) model to determine whether there is any international linkage of stock prices, especially in short-run relationships among stock price indices (see, for example, Jeon and Von Furstenberg (1990) and Tsutsui and Hirayama (2004)). Studies in the second category, on the other hand, shed more light on the long-run equilibrium relationship among returns on stocks by using a vector error correction (VEC) model and cointegration tests (see, for example, Kasa (1992), Chan, Gup, and Pan (1997)). Further, some of the studies examine short-run deviations of returns from long-run relationships by evaluating variance decomposition and impulse response functions.

What is common in the two approaches is that they assume constant parameters in their VAR or VEC with few exceptions that consider structural changes (Narayan and Smyth (2006), for example). In other words, the relationships appearing in the models are characterized by the parameters that are invariant over time. This assumption seems too strong or

[2]The idea of joint efficiency is not new; for example, MacDonald and Taylor (1989) focus on the relationship between cointegration and joint efficiency in foreign exchange markets.

even implausible for international financial markets where regime switchings or structural changes occur perhaps due to policy changes and the increasing presence of emerging markets. Furthermore, given the fact that technological changes have been connecting global equity markets more and more strongly – by facilitating instant and international trades –, the assumption of the stable relationships among international markets is not appealing.

In order to take into account regime shifts or structural changes, traditionally, researchers have often used subsamples of the data to exclude such aberrant events. By splitting the whole sample into several subsamples so that each subsample does not contain a break or shift, one can easily estimate the relationship among variables within the subsamples. The biggest challenge to this method is to find the cutoff points that determine subsamples.

Another approach that is preferred by some economists is the so-called rolling window method. Instead of using subsamples defined by breaks, this method allows a researcher to slide the fixed-width window (i.e., subsample) by a given increment. For example, to find the correlation coefficient of two variables at time-t, one can utilize data from $t-h$ to $t+h$ as a subsample; likewise, the correlation coefficient at $t+1$ needs data from $(t+1)-h$ to $(t+1)+h$. Here, the width of the window is $2h+1$. In effect, to estimate the correlation coefficient for all t, one must use overlapping subsamples. The remaining problem for this approach is that a researcher has to determine the width of the window (or h).

The proposed approach in this chapter is free from choices of cutoff points or the width of windows. In fact, we are able to demonstrate that our method, a non-Bayesian time-varying VAR (TV-VAR) model, can quite easily be implemented. This result is in line with that of Ito et al. (2012), who propose the TV-VAR model for a univariate case. Furthermore, by defining and computing the time-varying degree of the market efficiency from the estimated TV-VAR, we are able to discover the time-varying structure of world stock markets. This time-varying degree of market efficiency is used in conjunction with its statistical inference – which is calculated by our new bootstrap method – to determine whether or not the stock market is efficient.

To sum up, this chapter's contribution is two-fold. First, we look at the linkage of international equity markets through the novel non-Bayesian TV-VAR that possesses statistically more preferable properties than alternatives. This is a clear deviation from previous studies that employ standard VARs and VECs. As the second contribution, we propose the "time-varying degree of market efficiency" associated with the TV-VAR model and its statistical inference. This new measure and its statistical inference provide information as to whether the linkage is so strong that investors regard the markets for which efficiency is jointly confirmed as a single market. Our method of evaluating (joint) market efficiency is an alternative to the previous studies that employ the VEC model, where cointegration may imply market inefficiency. This is because (i) we incorporate the time-varying nature of market efficiency and (ii) the evaluation is based on estimated TV-VAR coefficients.

Our empirical results affirm the time-varying structure in the linkages and in the market efficiency of international stock markets. We also find that the time-varying nature of the market structure corresponds well to historical events in the international financial system.

This chapter is organized as follows. In Section 3.2, the model and new methodologies for our non-Bayesian TV-VAR model are presented. We also introduce our new measure, the degree of market efficiency in Section 3.2. The data on G7 stock markets, together with preliminary unit root test results, are described in Section 3.3. In Section 3.4, our empirical results show that international stock markets are inefficient in the late 1980s; and the European markets become jointly inefficient after the late 1990s. Section 3.5 concludes. The Technical Appendix provides mathematical and statistical discussions about the new estimation methodologies that we have developed.

3.2 The Model

3.2.1 Preliminaries

First, we consider a simple VAR model to analyze the stock market linkage in the context of market efficiency. Let $\boldsymbol{y}_t$ denote a k-dimensional vector representing the rates of return for k stocks at t. More specifically, $\boldsymbol{y}_t$ consists of k countries' stock market returns. In the literature, a number of previous studies employ VAR(p) models to analyze the linkage of stock markets:

$$\boldsymbol{y}_t = \boldsymbol{\nu} + A_1\boldsymbol{y}_{t-1} + \cdots + A_p\boldsymbol{y}_{t-p} + \boldsymbol{u}_t; \quad t = 1, 2, \ldots, T, \tag{3.1}$$

where $\boldsymbol{\nu}$ is a vector of intercepts; $\boldsymbol{u}_t$ is a vector of error terms that are serially uncorrelated. Note that the unexpected, excess returns on stocks at t are solely driven by the error term $\boldsymbol{u}_t$ because $\boldsymbol{y}_t - E\left[\boldsymbol{y}_t \mid \boldsymbol{y}_{t-1}, \boldsymbol{y}_{t-2}, \cdots\right] = \boldsymbol{u}_t$, where $E\left[\boldsymbol{y}_t \mid \boldsymbol{y}_{t-1}, \boldsymbol{y}_{t-2}, \cdots\right]$ represents the conditional expectation of the k countries' stock returns at t given the returns at any previous periods $t-1, t-2, \cdots$. This set-up is in accordance with the view of the efficiency market hypothesis (EMH) in the sense that there is no unexploited excess profit given the information available to the public (see Fama (1970, 1991), for example).

Oftentimes, it is understood that the VAR(p) model is a reduced form of a data generating process that is a VMA(∞) model:

$$\begin{aligned} \boldsymbol{y}_t &= \boldsymbol{\mu} + \Phi_0\boldsymbol{u}_t + \Phi_1\boldsymbol{u}_{t-1} + \Phi_2\boldsymbol{u}_{t-2} + \cdots \\ &= \boldsymbol{\mu} + \Phi\left(L\right)\boldsymbol{u}_t, \end{aligned} \tag{3.2}$$

where $\Phi\left(L\right)$ is a matrix lag polynomial of a lag operator L, i.e., $\Phi\left(L\right) = \Phi_0 + \Phi_1 L + \Phi_2 L^2 +$

$\cdots$, with all eignvalues of $\Phi(1)$ being outside the unit circle. One clear assumption here is that coefficient matrices $\{\Phi_i\}_{i=0}^{\infty}$ are time-invariant or parameter matrices of $k \times k$. With such matrices, one can compute the impulse-response functions along with the identification assumptions such as $\Phi_0 = I$ (an identity matrix). Note that the long-run effect of $\boldsymbol{u}_t$ on $\boldsymbol{y}$ is given by $\Phi(1)$; the vector of expected excess returns, $E\left[\boldsymbol{y}_t \mid \boldsymbol{y}_{t-1}, \boldsymbol{y}_{t-2}, \cdots\right] - \boldsymbol{\mu}$, is zero when $\Phi(1) = I$.

In this chapter, we view Equation (3.2) as the model and Equation (3.1) serves as a device for estimation. Our focus as to whether EMH holds hinges on whether $\Phi(1) = I$, or equivalently, whether $[I - A(1)]^{-1} = [I - A_1 - A_2 - \cdots - A_p]^{-1} = I$. The following subsection explains the extended version of our estimation model, Equation (3.1).

3.2.2 Non-Bayesian Time-Varying VAR Model

When the parameters in our VAR coefficient matrices $\{A_i\}_{i=0}^{p}$ do not seem to be constant over time, more precisely, when a parameter consistency test such as Hansen (1992) rejects the null hypothesis of constant parameters,[3] it is more appropriate to use the time-varying (TV-) VAR model that allows the parameters in $\{A_i\}_{i=0}^{p}$ to vary with time:

$$\boldsymbol{y}_t = \boldsymbol{\nu} + A_{1,t}\boldsymbol{y}_{t-1} + \cdots + A_{p,t}\boldsymbol{y}_{t-p} + \boldsymbol{u}_t; \quad t = 1, 2, \ldots, T, \tag{3.3}$$

with

$$A_{i,t} = A_{i,t-1} + V_{i,t} \tag{3.4}$$

for $i = 1, \ldots, p$, and $t = 1, 2, \ldots, T$. The assumption for the process, Equation (3.4), stems from the fact that the alternative hypothesis of Hansen's (1992) test is the multivariate random walk process. For this time-varying specification, we consider the underlying time-

[3] See Technical Appendix 3.A.1.

varying VMA model:

$$\boldsymbol{y}_t = \boldsymbol{\mu}_t + \Phi_{0,t}\boldsymbol{u}_t + \Phi_{1,t}\boldsymbol{u}_{t-1} + \Phi_{2,t}\boldsymbol{u}_{t-2} + \cdots \tag{3.5}$$

with $\Phi_{0,t} = I$ for all t.

Taken together, the following is the extended version of Ito et al. (2012), who consider the time-varying model within a univariate context. In this chapter we propose the Non-Bayesian time-varying model with a vector process. First of all, it is convenient for us to set our model in a state-space form, especially for estimation purposes. To this end, the observation Equation (3.3) is written as

$$\boldsymbol{y}_t = \boldsymbol{\nu} + A_t Z_{t-1} + \boldsymbol{u}_t; \quad t = 1, 2, \ldots, T,$$

where

$$A_t = \begin{bmatrix} A_{1,t} & \cdots & A_{p,t} \end{bmatrix} \text{ and } Z_{t-1} = \begin{bmatrix} \boldsymbol{y}'_{t-1} & \cdots & \boldsymbol{y}'_{t-p} \end{bmatrix}';$$

and the state Equations (3.4) becomes

$$vec\,(A_t) = vec\,(A_{t-1}) + \boldsymbol{v}_t,$$

where $\boldsymbol{v}_t = vec\,(V_t) = vec\left(\begin{bmatrix} V_{1,t} & \cdots & V_{p,t} \end{bmatrix}\right)$ is a k^2p-vector.

Furthermore, to see the advantages of our estimation method, we stack equations (3.3) from $t = 1$ through T,

$$\boldsymbol{y} = D\boldsymbol{\beta} + \boldsymbol{u}$$

where $\boldsymbol{y} = \begin{pmatrix} \boldsymbol{y}'_1 & \cdots & \boldsymbol{y}'_T \end{pmatrix}'$, $\boldsymbol{u} = \begin{pmatrix} \boldsymbol{u}'_1 & \cdots & \boldsymbol{u}'_T \end{pmatrix}'$, D is a matrix consisting of $Z_0, \quad \cdots \quad Z_{T-1}$, and $\boldsymbol{\beta}$ is a vector that contains $A_1 \quad \cdots \quad A_T$. For more details, see Technical Appendix

3.A.2.

The time-varying nature of the VAR coefficients, more specifically Equation (3.4), is also stacked from $t = 1$ to T, and written in a matrix form:

$$\boldsymbol{\gamma} = W\boldsymbol{\beta} + \boldsymbol{v}$$

where a vector $\boldsymbol{\gamma}$ includes $vec\,(A_0)$ and zeros; W is a matrix consisting of identity matrices and zeros; and $\boldsymbol{v} = \begin{pmatrix} \boldsymbol{v}_1' & \cdots & \boldsymbol{v}_T' \end{pmatrix}'$. Noticing that both systems of equations share the same vector $\boldsymbol{\beta}$, we combine them together and our entire system now becomes

$$\begin{bmatrix} \boldsymbol{y} \\ \gamma \end{bmatrix} = \begin{bmatrix} D \\ W \end{bmatrix} \boldsymbol{\beta} + \begin{bmatrix} \boldsymbol{u} \\ \boldsymbol{v} \end{bmatrix}. \tag{3.6}$$

One of the very convenient features of our method is that the system (3.6) allows us to estimate the unknown parameters $\boldsymbol{\beta}$ (time-varying coefficients) and the variance-covariance of the error term by OLS.[4] Unlike the traditional method or the Bayesian estimation, the time-varying model does not necessitate the Kalman filtering which requires iterations from $t = 1$ to T, to find the likelihood value. Indeed, this OLS based method estimates unknown parameters at one time. This method can also handle cases such as heteroscedasticity, serial correlations, and mutual correlations within the vector $(\boldsymbol{u}, \boldsymbol{v})$ by utilizing the Generalized Least Squares (GLS) method. Note that, when the Kalman filter is applied to this model, mutual correlations in $\boldsymbol{u}$ and $\boldsymbol{v}$ need special treatments, such as what Anderson and Moore (1979) describe. Other desirable properties of our method are more thoroughly explained in Ito et al. (2012).

The next subsection describes a different way to examine whether EMH holds em-

[4]See Technical Appendix 3.A.3 for more details.

pirically, under our assumption that the structure or autoregressive coefficients on stock returns vary over time.

3.2.3 Time-Varying Degrees of Market Efficiency

While our main focus is whether the vector of expected excess returns, $E\left[\boldsymbol{y}_t \mid \boldsymbol{y}_{t-1}, \boldsymbol{y}_{t-2}, \cdots\right] - \boldsymbol{\mu}$, is zero in Equation (3.2), the time-varying VAR model (3.3) modifies this condition as $E\left[\boldsymbol{y}_t \mid \boldsymbol{y}_{t-1}, \boldsymbol{y}_{t-2}, \cdots\right] - \boldsymbol{\mu}_t = 0$ where $\boldsymbol{\mu}_t = \left(I - A_{1,t} - \cdots - A_{p,t}\right)^{-1} \boldsymbol{u}_t$. Still, with the estimated coefficient matrices obtained by (3.6) via OLS (or GLS), we are able to determine if EMH holds by observing the time-varying matrices $A_{1,t}, \ldots, A_{p,t}$ or $\Phi_{1,t}, \Phi_{2,t}, \ldots$. This is because the TV-VAR model (3.3) can be seen as a locally stationary model. To implement this strategy, let us consider a matrix

$$\Phi_t(1) = \sum_{j=1}^{\infty} \Phi_{j,t},$$

which is a cumulative sum of the time-varying VMA coefficient matrices which appears in Equation (3.5). Since this matrix measures the long-run effect of shocks $\{\boldsymbol{u}_\tau\}_{\tau=t}^{-\infty}$ on stock return $\{\boldsymbol{y}_t\}$, we call it the long-run multiplier that is in line with Ito et al. (2012). This multiplier, $\Phi_t(1)$, is, in fact, regarded as a metric measure for market efficiency in the sense that $\Phi_t(1) = I$ suggests EMH; while its deviations from an identity matrix measure the degree of market (in)efficiency. Once again, we do not impose the identification assumptions about shocks. In other words, we do not label which shock originated from country $j \in \{1, \ldots, k\}$. Therefore, it is important for us to investigate whether or not k countries' stock markets are all efficient without specifying the roles of the shocks; and hence, measuring the distance between $\Phi_t(1)$ and I provides us with a good idea as to how k countries' stock markets are closer to or farther from efficiency.

The distance between $\Phi_t(1)$ and I is measured by the spectral norm:

$$\zeta_t = \sqrt{\max \lambda \left[(\Phi_t(1) - I)' (\Phi_t(1) - I) \right]}, \tag{3.7}$$

i.e., the square root of the largest eigenvalue of the square matrix $(\Phi_t(1) - I)'(\Phi_t(1) - I)$ for each t. Clearly, the distance becomes zero ($\zeta_t = 0$) when $\Phi_t(1) = I$. Yet, the distance becomes a large (positive) number as the two matrices deviate from each other, in the sense of the spectral norm. We call ζ_t the degree of market efficiency that measures how close to or far from the efficient markets the actual markets are.[5] By computing this for all the sample points, t, we are able to know when the markets are efficient (or close enough to indicate that the markets are efficient) and when the markets are obviously inefficient.

From the correspondence between Equations (3.3) and (3.5), provided $(I - A_{1,t} - \cdots - A_{p,t})$ is non-singular, we have

$$\Phi_t(1) = (I - A_{1,t} - \cdots - A_{p,t})^{-1}.$$

In practice, we estimate the TV-VAR Equation (3.3) to obtain $A_{1,t}, \ldots, A_{p,t}$, and then, we compute the degree of market efficiency, Equation (3.7).

3.3 Data

We utilize the monthly data on the Morgan Stanley capital index for the Group of 7 (G7) countries (United States, Canada, United Kingdom, Japan, Germany, France, and Italy) from December 1969 through March 2013, as obtained from the Thomson Reuters

[5] In case of univariate data, this norm is essentially same as the degree defined by Ito et al. (2012). That is, the degree in this chapter is a natural extension of the one in Ito et al. (2012).

Datastream. To compute the ex-post stock return series, we take the first difference of the natural log of the stock price index.

(Table 3.1 around here)

Provided in Table 3.1, descriptive statistics show that monthly returns on the stock market range from Italy's 0.34% (4.08% per annum) to the U.K.'s 0.57% (6.84% per annum). The average return on the largest stock market, U.S. is 0.52% (6.24% annum) and the third highest after the U.K. and Canada (0.54% or 6.48% annum). Volatility for the U.S. is the smallest among the seven countries, 3.96%, while that for Italy is 6.17%, which is the largest of the seven countries. Except for the U.K. and France, the largest month-to-month change is negative, not positive, i.e., crashes have larger magnitudes than booms.

Table 3.1 also shows the results of a unit root test. It is important for our estimation that the variables in the TV-VAR model are all stationary. Thus, we examine this by employing the ADF-GLS test of Elliott et al. (1996). It is then confirmed that for all the variables, the ADF-GLS test uniformly rejects the null hypothesis of the variable's having a unit root at the 1% significance level.

3.4 Empirical Results

3.4.1 The Time-Invariant VAR Model

Having presented the time-varying model, we must verify that the TV-VAR model is more appropriate to use than the traditional, time-invariant VAR model. To do so, we first estimate the time-invariant VAR model, and then apply the parameter constancy test of Hansen (1992) to investigate whether the time-invariant model is a better fit for our data.

Here we consider the parameter stability for five groups using five VAR models. They are: (i) North American markets (U.S. and Canada); (ii) two largest stock markets (U.S. and U.K.); (iii) three largest stock markets (U.S., U.K. and Japan); (iv) European markets (U.K., Germany, France and Italy); and (v) all seven markets.

(Table 3.2 around here)

To select the length of the lags, we adopt the Bayesian information criterion (BIC; Schwarz (1978)). Table 3.2 presents both estimated coefficients and Hansen's (1992) joint parameter constancy test statistics that are shown in the last column, under "L_C." For all the five models, the parameter constancy test rejects the null hypothesis of constancy at the 1% level, against the alternative hypothesis stating that the parameter variation follows a random walk process. These results suggest that the time-invariant VAR model does not fit our data; rather, we should use the TV-VAR model for the G7 data.[6]

3.4.2 TV-VAR Model and the Degree of Market Efficiency

Now, let us focus on the TV-VAR model and the degree of market efficiency associated with the model. Once again, using the spectral norm, we measure the stock markets' deviation from the efficient condition, by Equation (3.7). For example, considering U.S. and Canadian stock markets, the degree of market efficiency tells us how the two markets are different from the efficient markets. If $\zeta_t = 0$ for time-t, the two stock markets are jointly efficient at that time.

Note that the degree of market efficiency is computed from the estimates of $\Phi_t(1)$, therefore, ζ_t is subject to sampling errors. Because of this, we provide a sequence of the critical values of ζ_t under the null hypothesis of market efficiency.[7] We do not find

[6]Table 3.A.1 in Technical Appendix provides the results of a univariate AR process for each country. From "L_C" statistics, we can confirm that the TV model is more appropriate for all countries.

[7]While we use the bootstrap method (i.e., using estimated residuals) to compute the critical values,

evidence of inefficient markets whenever estimated ζ_t is less than the upper critical value; inefficient markets are detected with an estimated ζ_t that is larger than the upper limit. As detailed in Technical Appendix 3.A.4, Monte Carlo simulations construct a sequence of the critical values of the estimates by (i) generating multi-variate i.i.d. processes for $\boldsymbol{y}_t$, then, (ii) applying and estimating the TV-VAR model for those processes, and finally, (iii) computing ζ_t.

(Figures 3.1 to 3.5 around here)

Looking at North American markets, as demonstrated in Figure 3.1, the markets become inefficient in the late 1980s and in the late 2000s; and almost completely inefficient in the early 2000s. Interestingly, these correspond well with Black Monday in 1987 and the Savings and Loan (S&L) crisis in the late 1980s, the world financial crisis of 2008-2009, and the dot-com bubble burst in 2001. To provide a better understanding of the dynamics of the international market linkages, Figure 3.6 presents the degree of market efficiency for each of the seven countries. While the degrees for the U.S. and Canada have similar fluctuation patterns, inefficiency in the early 2000s seems to attribute to the Canadian market inefficiency.

(Figure 3.6 around here)

The two largest stock markets become inefficient in the early to middle 1970s and the late 1980s. Yet, the U.K. market's inefficiency in the 1970s stands out. In addition to the period when the two largest markets are inefficient, the three largest stock markets, including Japan, become inefficient in the middle 2000s and after 2010.

the Monte Carlo method (i.e., using random draws from the normal distribution) can generate pretty much the same critical values. Therefore, our empirical results presented in this chapter do not depend on which method is used to compute the statistical inference for the degree of market efficiency.

Like other markets, the middle 1970s and 1980s are times for inefficient markets in Europe, the degree of market efficiency in the European markets is nevertheless increasing since the late 1990s. Why is the European markets' efficiency decreasing (the degree of market efficiency increasing) after the late 1990s? From individual degrees of efficiency presented in Figure 3.6, it is hard to understand why. In addition, even when we compute the degree of market efficiency for each country individually (See Figure 3.6), none of the European countries indicates such tendency. Remarkably, however, the introduction of the common currency euro in 1999 somewhat coincides the increase of the degree of market efficiency.

Finally, the efficiency of all seven countries changes over time. But importantly, in the early-to-middle 1970s, between the late 1980s and early 1990s, and after 2000 are three periods when the world (G7) markets indicate their inefficiency.

3.4.3 Why Do Stock Markets Become Inefficient?

There are, possibly, several reasons why stock markets become inefficient. First, as pointed out by Ito et al. (2012) who focus solely on the US stock market, (i) people become irrational when they face extraordinary events such as severe recessions or financial crises[8]; (ii) people are rational, but their stochastic discount factor changes for those time periods.[9] It is also possible that the stock markets in our sample were indeed efficient, in combination with other stock markets that are not included in our sample. To fully understand why markets become inefficient, further investigation, perhaps with larger samples, is necessary.

[8]Ito et al. (2012) utilize a much longer sample period for the US, and find that the world financial crisis in 2008 did not cause an inefficient market in the US. According to their study, inefficient markets emerged during: (i) the longest recession defined by NBER (1873-1879); (ii) the 1902-1904 recession; (iii) the New Deal era; and (iv) just after the very severe 1957-1958 recession.

[9]In such a case, the efficiency condition does not hold even with people's rational behavior.

3.5 Concluding Remarks

Stock market efficiency, especially the international linkage of stock markets, is studied. Paying attention to the stability of VAR coefficients, we determine that the time-varying (TV) VAR model fits better with the stock market data, with the help of Hansen's (1992) parameter constancy test. As an extension of Ito et al. (2012), we develop a new metric for market efficiency, namely the degree of market efficiency that is the spectral norm between the estimated time-varying parameter matrix and an identity matrix.

Our empirical results show that the degree of the market efficiency is, in fact, time-varying; and there are times when international markets are jointly efficient and inefficient. After considering five groups of countries, we interestingly find that (i) international stock markets are inefficient in the late 1980s; and (ii) European markets become jointly inefficient – despite the fact that each country's stock market indicates otherwise – after the late 1990s, which is almost corresponding to the commencement of their common currency.

However, caution should be taken in interpreting our results. For example, failing to find evidence of market efficiency in our sample does not necessarily mean that the markets are inefficient. This is because those markets are, possibly, jointly efficient with other markets that are not included in our sample. Also, further investigation should reveal the mechanism of how the degree of market efficiency fluctuates from time to time – especially why the degree of market efficiency exhibits inefficient markets during extraordinary times – by utilizing larger samples of data and by formal statistical tests.

Table 3.1: Descriptive Statistics and Unit Root Tests

	Mean	SD	Min	Max	ADF-GLS	Lags	$\hat{\phi}$	$\mathcal{N}$
R^{US}	0.0052	0.0396	−0.2324	0.1596	−3.9408	4	0.5766	519
R^{CA}	0.0054	0.0430	−0.2365	0.1496	−17.4826	0	0.2569	519
R^{GB}	0.0057	0.0470	−0.2468	0.3685	−15.3813	1	0.2828	519
R^{JP}	0.0036	0.0477	−0.2515	0.1597	−15.9591	0	0.3391	519
R^{DE}	0.0038	0.0498	−0.2924	0.1386	−8.7929	2	0.4023	519
R^{FR}	0.0050	0.0525	−0.2004	0.2040	−17.0041	0	0.2823	519
R^{IT}	0.0034	0.0617	−0.2290	0.2061	−17.5358	0	0.2540	519

Notes:

(1) "R^{US}", "R^{CA}", "R^{GB}", "R^{JP}", "R^{DE}", "R^{FR}", and "R^{IT}" denote the returns of the Morgan Stanley capital index for the G7 countries (United States, Canada, United Kingdom, Japan, Germany, France, and Italy), respectively.

(2) "ADF-GLS" denotes the ADF-GLS test statistics, "Lags" denotes the lag order selected by the MBIC, and "$\hat{\phi}$" denotes the coefficients vector in the GLS detrended series (see Equation (6) in Ng and Perron (2001)).

(3) In computing the ADF-GLS test, a model with a time trend and a constant is assumed. The critical value at the 1% significance level for the ADF-GLS test is "−3.42".

(4) "$\mathcal{N}$" denotes the number of observations.

(5) R version 3.1.0 was used to compute the statistics.

Table 3.2: Standard VAR Estimations

	$Constant$	R_{t-1}^{US}	R_{t-1}^{CA}	R_{t-1}^{GB}	R_{t-1}^{JP}	R_{t-1}^{DE}	R_{t-1}^{FR}	R_{t-1}^{IT}	$\bar{R}^2$	L_C
(1) North America										
R_t^{US}	0.0044	0.2230	−0.0306	–	–	–	–	–	0.0361	68.1322
	[0.0018]	[0.0686]	[0.0732]							
R_t^{CA}	0.0038	0.1532	0.1277	–	–	–	–	–	0.0603	
	[0.0020]	[0.0904]	[0.0808]							
(2) U.S. and U.K.										
R_t^{US}	0.0041	0.1274	–	0.1017	–	–	–	–	0.0453	52.6841
	[0.0019]	[0.0730]		[0.0414]						
R_t^{GB}	0.0035	0.2242	–	0.1726	–	–	–	–	0.1000	
	[0.0019]	[0.0772]		[0.0636]						
(3) (2) plus Japan										
R_t^{US}	0.0041	0.1411	–	0.1103	−0.0343	–	–	–	0.0447	
	[0.0019]	[0.0820]		[0.0392]	[0.0481]					
R_t^{GB}	0.0035	0.2256	–	0.1735	−0.0036	–	–	–	0.0983	85.8165
	[0.0019]	[0.0823]		[0.0659]	[0.0490]					
R_t^{JP}	0.0020	0.1320	–	0.0373	0.2055	–	–	–	0.0829	
	[0.0021]	[0.0807]		[0.0499]	[0.0590]					
(4) EU Countries										
R_t^{GB}	0.0039	–	–	0.2475	–	−0.0603	0.1476	−0.0347	0.0854	
	[0.0018]			[0.0586]		[0.0564]	[0.0833]	[0.0522]		
R_t^{DE}	0.0026	–	–	0.0600	–	0.1692	0.0646	0.0097	0.0630	117.8807
	[0.0022]			[0.0485]		[0.0596]	[0.0721]	[0.0444]		
R_t^{FR}	0.0035	–	–	0.0747	–	0.0613	0.1507	0.0049	0.0520	
	[0.0023]			[0.0746]		[0.0704]	[0.0764]	[0.0635]		
R_t^{IT}	0.0021	–	–	−0.0590	–	−0.0145	0.2442	0.1221	0.0619	
	[0.0026]			[0.0649]		[0.0895]	[0.1051]	[0.0742]		
(5) G7 Countries										
R_t^{US}	0.0042	0.01814	−0.0459	0.1131	−0.0285	−0.0336	0.0064	0.0179	0.0393	
	[0.0019]	[0.0964]	[0.0793]	[0.0453]	[0.0502]	[0.0508]	[0.0597]	[0.0370]		
R_t^{CA}	0.0039	0.1448	0.1241	0.0065	0.0653	0.0024	−0.0576	0.0194	0.0570	
	[0.0020]	[0.0983]	[0.0819]	[0.0617]	[0.0531]	[0.0601]	[0.0669]	[0.0381]		
R_t^{GB}	0.0034	0.2481	−0.0142	0.1850	0.0110	−0.1094	0.1027	−0.0390	0.1018	
	[0.0018]	[0.1051]	[0.0891]	[0.0661]	[0.0515]	[0.0543]	[0.0729]	[0.0523]		
R_t^{JP}	0.0020	0.1836	−0.0397	0.0395	0.2175	−0.1273	0.0397	0.0514	0.0853	160.5454
	[0.0021]	[0.0972]	[0.0771]	[0.0490]	[0.0615]	[0.0666]	[0.0642]	[0.0364]		
R_t^{DE}	0.0022	0.2668	−0.0211	0.0081	−0.0521	0.1308	0.0248	0.0094	0.0785	
	[0.0022]	[0.0997]	[0.0812]	[0.0491]	[0.0496]	[0.0616]	[0.0604]	[0.0470]		
R_t^{FR}	0.0033	0.2438	−0.1069	0.0441	−0.0330	0.0359	0.1333	0.0067	0.0581	
	[0.0024]	[0.1082]	[0.0955]	[0.0797]	[0.0578]	[0.0772]	[0.0782]	[0.0642]		
R_t^{IT}	0.0017	0.2024	−0.0312	−0.1029	−0.0031	−0.0491	0.2136	0.1201	0.0636	
	[0.0026]	[0.1087]	[0.1107]	[0.0644]	[0.0602]	[0.0807]	[0.1103]	[0.0709]		

Notes:

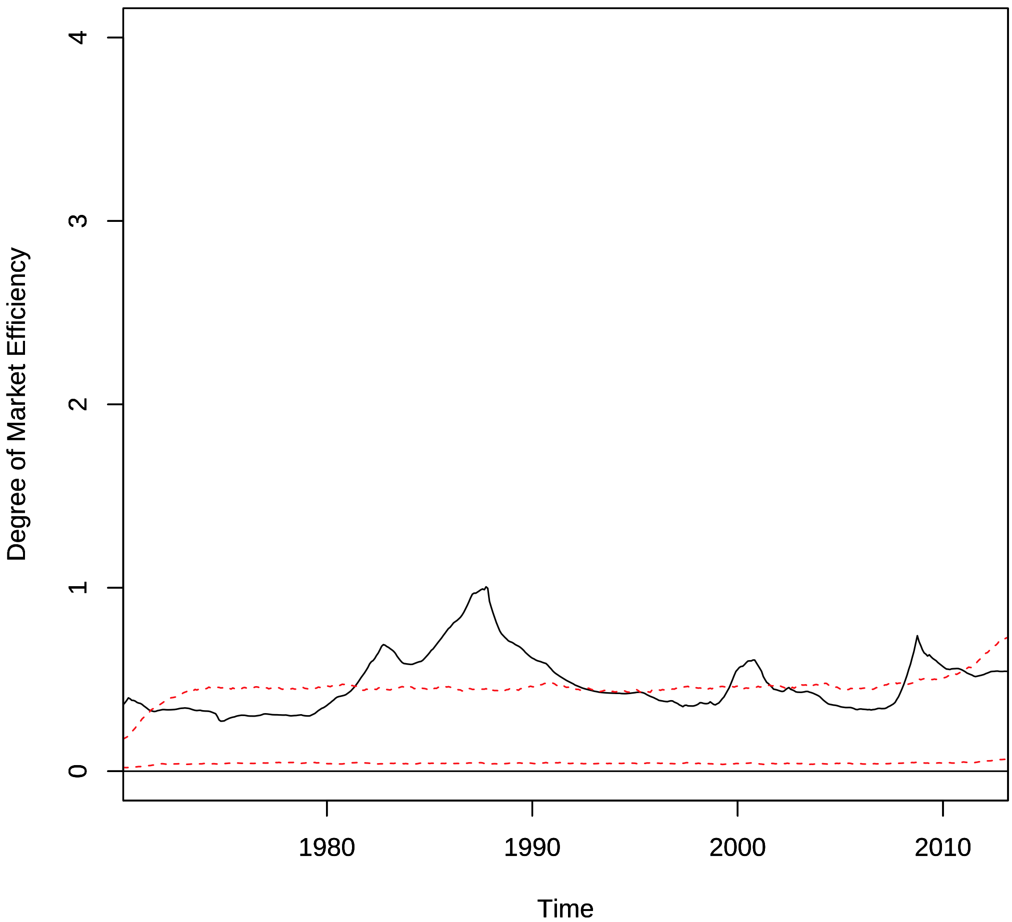
(1) "R_{t-1}^{US}", "R_{t-1}^{CA}", "R_{t-1}^{GB}", "R_{t-1}^{JP}", "R_{t-1}^{DE}", "R_{t-1}^{FR}", and "R_{t-1}^{IT}" denote the lagged returns of the Morgan Stanley capital index for the G7 countries (United States, Canada, United Kingdom, Japan, Germany, France, and Italy), respectively.

(2) "$\bar{R}^2$" denotes the adjusted R^2, and "L_C" denotes Hansen's (1992) joint L statistic with variance.

(3) Newey and West's (1987) robust standard errors are in brackets.

(4) R version 3.1.0 was used to compute the estimates and the statistics.

Figure 3.1: Time-Varying Degree of Market Efficiency: North America



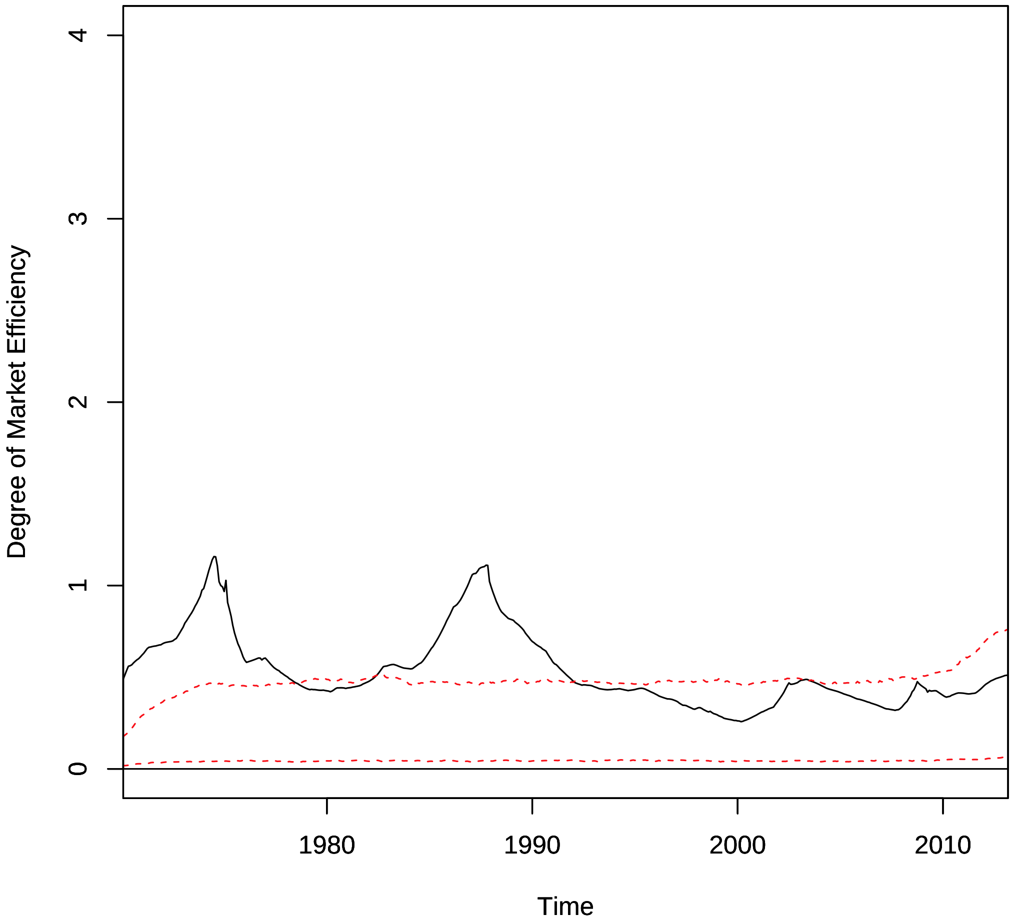
Notes:

(1) The dashed red lines represent the 99% critical values of the time-varying spectral norms in case of efficient market.

(2) We run 5000 times Monte Carlo sampling to calculate the critical values.

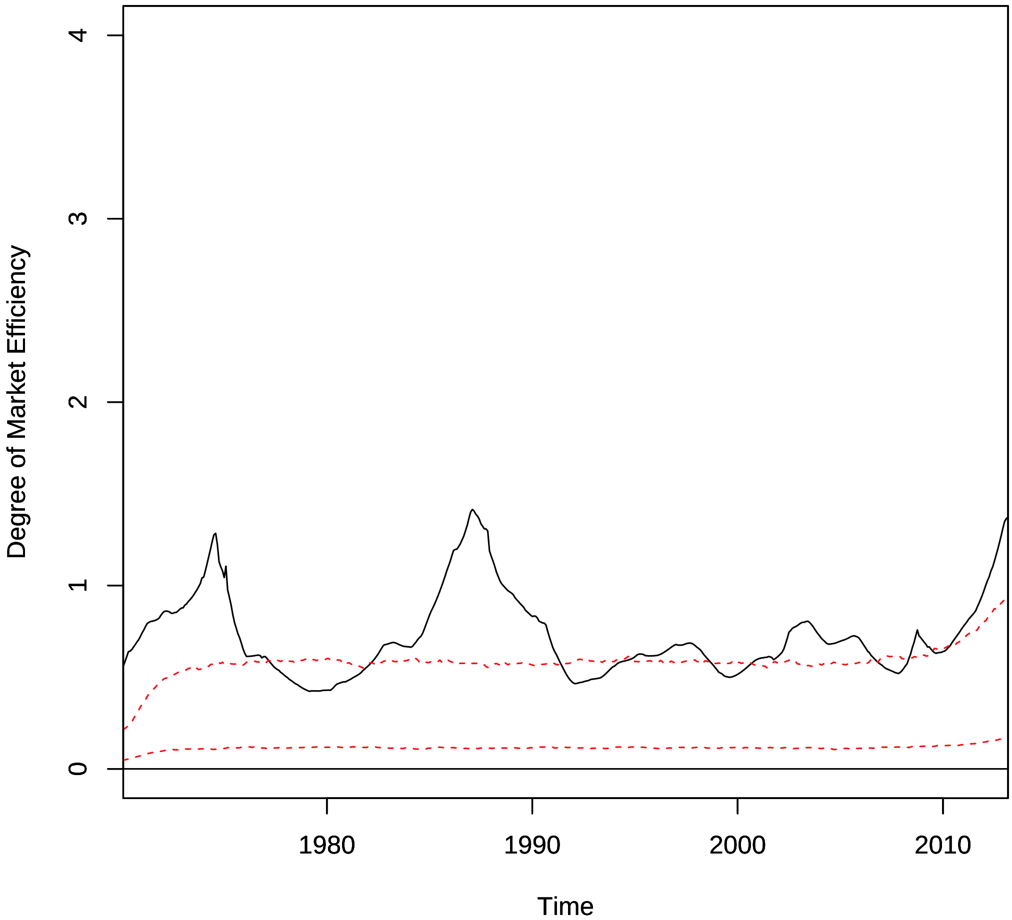
(3) R version 3.1.0 was used to compute the estimates.

Figure 3.2: Time-Varying Degree of Market Efficiency: U.S. and U.K.



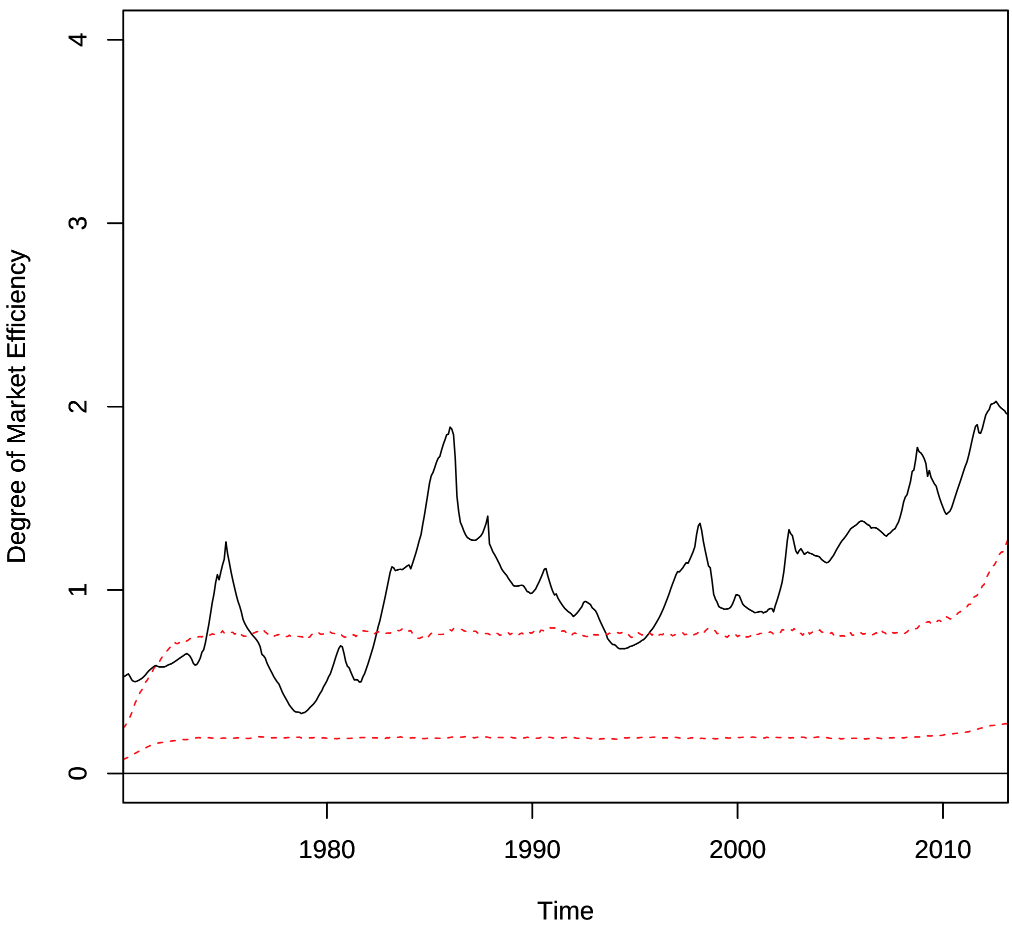
Note: As for Figure 3.1.

Figure 3.3: Time-Varying Degree of Market Efficiency: U.S., U.K. and Japan



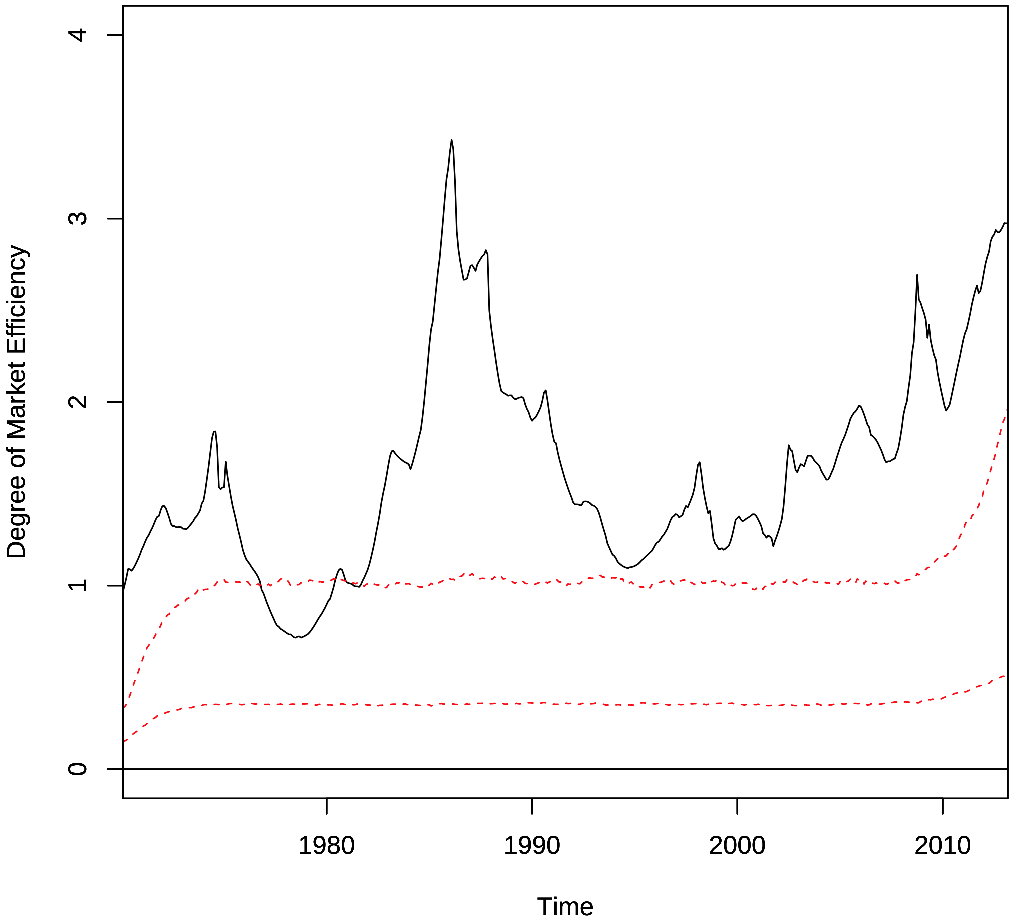
Note: As for Figure 3.1.

Figure 3.4: Time-Varying Degree of Market Efficiency: European Countries



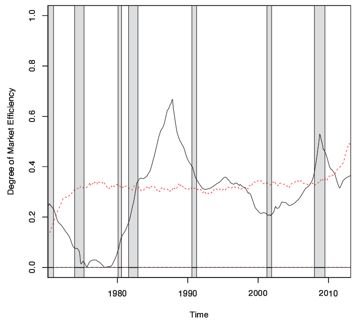
Note: As for Figure 3.1.

Figure 3.5: Time-Varying Degree of Market Efficiency: G7 Countries

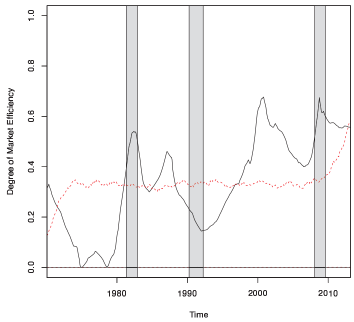


Note: As for Figure 3.1.

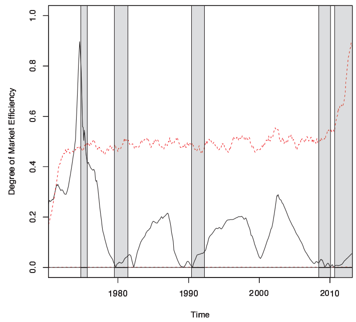
Figure 3.6: Time-Varying Degree of Market Efficiency: Individual Countries



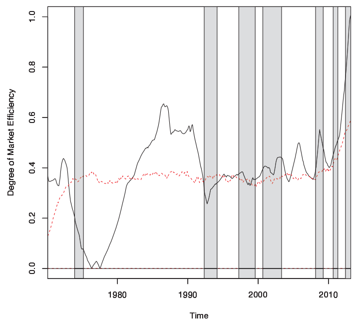
(a) U.S.



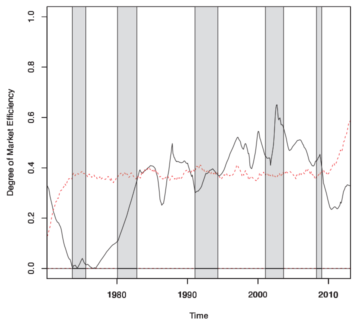
(b) Canada



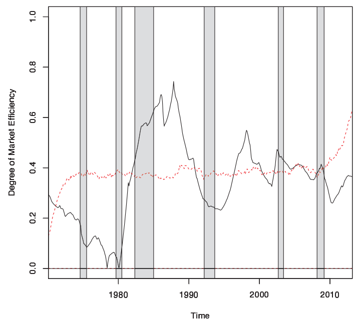
(c) U.K.



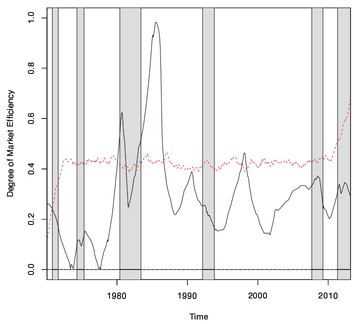
(d) Japan



(e) Germany



(f) France



(g) Italy

<u>Notes</u>:

(1) The dashed red lines represent the 99% critical values of the time-varying spectral norm in case of efficient market.

(2) We run 5000 times Monte Carlo sampling to calculate the critical values.

(3) The shade areas represent recessions reported by the NBER business cycle dates for the U.S. and the Economic Cycle Research Institute business cycle peak and trough dates for the other countries.

(4) R version 3.1.0 was used to compute the estimates.

3.A Technical Appendix

3.A.1 Parameter Constancy Test for Time-Invariant VAR Model

There are some parameter constancy tests, for example Andrews (1993b) and Nyblom (1989). Hansen (1992) develops a parameter constancy test for linear and non-linear models. The test is made under the null and alternative hypotheses: (H_0) the parameters are constant over time and (H_1) they follow a martingale process. In practice, we reformulate Equation (3.1) to extend Hansen's test for a VAR(p) model.

Define the $(k \times T)$ data matrix as

$$Y := (\boldsymbol{y}_1, \boldsymbol{y}_2, \cdots, \boldsymbol{y}_T),$$

where T is the number of time series observations. To reformulate the VAR(p) model with a intercept term into a regression formula, we introduce the following auxiliary vector and matrix:

$$Z_t := \begin{pmatrix} 1 \\ \boldsymbol{y}_t \\ \vdots \\ \boldsymbol{y}_{t-p} \end{pmatrix},$$

and

$$Z := (Z_0, Z_1, \cdots, Z_{T-1}).$$

Note that Z_0 is regarded as a prior of estimation. Defining the $(k \times kT)$ disturbance matrix as

$$U := (\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_T),$$

we obtain a matrix form of equation (1):

$$Y = BZ + U,$$

where a $k \times (1 + kp)$ matrix B is $[\boldsymbol{\nu}, A_1, \cdots, A_p]$. Using the *vec* operator, which transforms a $(k \times T)$ matrix into an (kT) vector by stacking the columns, we have the following regression form of Equation (3.1):

$$vec(Y) = (Z' \otimes I_k)vec(B) + vec(U), \tag{3.A.1}$$

where $\otimes$ is the Kronecker product and I_k denotes k identity matrix. In case $k = 2$ and $p = 1$, Equation (3.A.1) is specifically exhibited as

$$\begin{pmatrix} y_{11} \\ y_{21} \\ y_{12} \\ y_{22} \\ \vdots \\ y_{1T} \\ y_{2T} \end{pmatrix} = \begin{pmatrix} 1 & 0 & y_{10} & 0 & y_{20} & 0 \\ 0 & 1 & 0 & y_{10} & 0 & y_{20} \\ 1 & 0 & y_{11} & 0 & y_{21} & 0 \\ 0 & 1 & 0 & y_{11} & 0 & y_{21} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & y_{1,T-1} & 0 & y_{2,T-1} & 0 \\ 0 & 1 & 0 & y_{1,T-1} & 0 & y_{2,T-1} \end{pmatrix} \begin{pmatrix} \nu_1 \\ \nu_2 \\ \alpha_{11} \\ \alpha_{21} \\ \alpha_{12} \\ \alpha_{22} \end{pmatrix} + \begin{pmatrix} u_{11} \\ u_{21} \\ u_{12} \\ u_{22} \\ \vdots \\ u_{1T} \\ u_{2T} \end{pmatrix}.$$

In order to extend the procedure of Hansen (1992), we rewrite Equation (3.A.1) as

$$\boldsymbol{y} = X\boldsymbol{\beta} + \boldsymbol{u},$$

where $\boldsymbol{y} = vec(Y)$, $X = Z' \otimes I_k$ and $\boldsymbol{u} = vec(U)$. Given $\boldsymbol{y}$, we obtain the OLS estimate $\hat{\boldsymbol{\beta}} = (X'X)^{-1}X'\boldsymbol{y}$ and the estimate of the covariance matrix $\hat{\Sigma}_u$ as UU'/T or $UU'/(T - kp - 1)$

(see Lütkepohl (2005, ch.3) for details).

Furthermore, to extend Hansen's (1992) procedure to our system, let us organize the above equation as follows:

$$\boldsymbol{y}_t = X_t\boldsymbol{\beta} + \boldsymbol{u}_t,\ t = 1, 2, \cdots, T. \tag{3.A.2}$$

For simplicity, we consider an OLS estimation of Equation (3.A.2) with the i.i.d. assumption, $E[\boldsymbol{u}_t\boldsymbol{u}_t'] = \Sigma_u, t = 1, 2, \cdots$. The OLS estimates $\hat{\boldsymbol{\beta}}$ as well as $\hat{\Sigma}_u$ should hold the following equations.

$$\mathbf{0} = \sum_{\tau=1}^{T} X_\tau' \boldsymbol{e}_\tau$$

and

$$\mathbf{0} = \sum_{\tau=1}^{T} (vech(\boldsymbol{e}_\tau\boldsymbol{e}_\tau') - vech(\hat{\Sigma}_u)),$$

where $\boldsymbol{e}_t = \boldsymbol{y}_t - X_t\hat{\boldsymbol{\beta}}$ and the *vech* operator is closely related to *vec*, which only stacks the elements on and below the main diagonal of a symmetric matrix. To construct Hansen (1992) test statistic for a simultaneous linear equation system, define $(k(1+kp)+k(k+1)/2)$ vector,

$$f_\tau(\hat{\boldsymbol{\beta}}, \hat{\Sigma}_u) = \begin{bmatrix} vec(X_\tau'\boldsymbol{e}_\tau) \\ vech(\boldsymbol{e}_\tau\boldsymbol{e}_\tau') - vech(\hat{\Sigma}_u) \end{bmatrix},\ \tau = 1, 2, \cdots, T.$$

Let $f_{i\tau}$ denotes i-th component of $f_\tau(\hat{\boldsymbol{\beta}}, \hat{\Sigma}_u)$, $S_s = \sum_{\tau=1}^{s} f_\tau$ and $S_{is} = \sum_{\tau=1}^{s} f_{i\tau}$ for $i = 1, 2, \cdots, k(1+kp)+k(k+1)/2,\ s = 1, 2, \cdots, T$. Note that we abbreviate $\hat{\boldsymbol{\beta}}, \hat{\Sigma}_u$ when there is no confusion. Hansen's (1992) individual statistics of parameter constancy test are

$$L_i = \frac{1}{TV_i}\sum_{\tau=1}^{T} S_{i\tau}^2,\ i = 1, 2, \cdots, (k(1+kp)+k(k+1)/2),$$

where $V_i = \sum_{\tau=1}^{T} f_{i\tau}^2$. His joint statistic is defined:

$$L_c = \frac{1}{T}\sum_{\tau=1}^{T} S_\tau' V^{-1} S_\tau,$$

where $V = \sum_{\tau=1}^{T} f_\tau f_\tau'$. These test statistics follow singular distributions represented by the Brownian motions and bridges. See Hansen (1990) for the tables of the statistics.

3.A.2 Regressor of State Space Regression Models

This section shows the reader a significant property of the method that we present for dealing with a TV-VAR(p) model in Section 3.2. The method enables us to handle the model as a conventional econometric model by identifying the corresponding state space model with a linear regression model. Recall that the observation equation and the state one appeared in the state space model in Section 3.2: Equations (3.3) and (3.4).

We prove that estimating method of the smoother of a state space model guarantees the exact estimate whatever data is applied. Mathematically we assert that the regressor matrix of our underlying linear system has full rank. This implies that our estimation never suffers from multicolinearity and that our model is always identifiable in the sense of Rothenberg (1971). The following results do not depend on whatever time-varying model is assumed; they cover our TV-VAR model, a typical state space model.

The model consists of two equations as follows:

$$\boldsymbol{y}_t = X_t \boldsymbol{\beta}_t + \boldsymbol{u}_t, \quad \boldsymbol{u}_t \overset{iid}{\sim} \mathcal{WS}(\mathbf{0}, \boldsymbol{\Sigma}_{\boldsymbol{u}}), \quad t = 1, 2, \cdots, T, \tag{3.A.3}$$

and

$$\boldsymbol{\beta}_{t+1} = B_t \boldsymbol{\beta}_t + \boldsymbol{v}_t, \quad \boldsymbol{v}_t \overset{iid}{\sim} \mathcal{WS}(\mathbf{0}, \boldsymbol{\Sigma}_{\boldsymbol{v}}), \quad t = 1, 2, \cdots, T, \tag{3.A.4}$$

where $\mathcal{WS}(\mathbf{0}, \mathbf{\Sigma}_{\boldsymbol{u}})$ and $\mathcal{WS}(\mathbf{0}, \mathbf{\Sigma}_{\boldsymbol{v}})$ are any distributions depending only on their means and variances such as the normal distributions. Note that $\boldsymbol{y}_t$'s are $k-$dimensional vectors and $\boldsymbol{\beta}_t$'s are $m-$dimensional vectors as parameters. X_t and B_t are $k \times m$ and $m \times m$ matrices respectively. Notice that two disturbances, $\boldsymbol{u}_t$ and $\boldsymbol{v}_t$, follow *iid* normal distributions with zero means and covariance matrices, $\mathbf{\Sigma}_{\boldsymbol{u}}$ and $\mathbf{\Sigma}_{\boldsymbol{v}}$, respectively.

By transforming Equation (3.A.4) to the following form, we can regard the state equation as a linear regression model.

$$\mathbf{0} = B_t \boldsymbol{\beta}_t - I \boldsymbol{\beta}_{t+1} + \boldsymbol{v}_t, \quad \boldsymbol{v}_t \overset{iid}{\sim} \mathcal{WS}(\mathbf{0}, \mathbf{\Sigma}_{\boldsymbol{v}}), \quad t = 1, 2, \cdots, T. \tag{3.A.5}$$

Combining Equations (3.A.3) with (3.A.5) into the following linear system, we can employ conventional econometric techniques to obtain the Kalman smoother.

$$\hat{\boldsymbol{y}} = \begin{bmatrix} \boldsymbol{y} \\ \hdashline \gamma \end{bmatrix} = \begin{bmatrix} \boldsymbol{X} \\ \hdashline \boldsymbol{W} \end{bmatrix} \boldsymbol{\beta} + \begin{bmatrix} \boldsymbol{u} \\ \hdashline \boldsymbol{v} \end{bmatrix}$$

where

$$\boldsymbol{y} = \begin{pmatrix} \boldsymbol{y}_1 \\ \boldsymbol{y}_2 \\ \vdots \\ \boldsymbol{y}_T \end{pmatrix}, \quad \gamma = \begin{pmatrix} -B_0 \boldsymbol{\beta}_0 \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{pmatrix}, \quad \boldsymbol{u} = \begin{pmatrix} \boldsymbol{u}_1 \\ \boldsymbol{u}_2 \\ \vdots \\ \boldsymbol{u}_T \end{pmatrix}, \quad \boldsymbol{v} = \begin{pmatrix} \boldsymbol{v}_1 \\ \boldsymbol{v}_2 \\ \vdots \\ \boldsymbol{v}_T \end{pmatrix},$$

and

$$\boldsymbol{X} = \begin{pmatrix} X_1 & & & \boldsymbol{O} \\ & X_2 & & \\ & & \ddots & \\ \boldsymbol{O} & & & X_T \end{pmatrix}, \quad \boldsymbol{W} = \begin{pmatrix} -I & & & \boldsymbol{O} \\ B_1 & -I & & \\ & \ddots & \ddots & \\ \boldsymbol{O} & & B_{T-1} & -I \end{pmatrix}.$$

Our main theorem of this appendix is as follows:

Theorem 1 *The regressor matrix has full rank. That is*

$$\operatorname{rank}\left[\begin{array}{c}\boldsymbol{X}\\ \hdashline \boldsymbol{W}\end{array}\right] = mT.$$

Proof: Let $\hat{\boldsymbol{X}}$ denote the regressor matrix. It is of size $(k+m)T \times mT$. Thus, $\operatorname{rank}\hat{\boldsymbol{X}} \leq mT$. Clearly, each ranks of the two submatrices, $\boldsymbol{W}$ and $\boldsymbol{X}$, are fewer than or equal to mT. Formally,

$$\operatorname{rank}\boldsymbol{X} \leq mT,$$

and

$$\operatorname{rank}\boldsymbol{W} \leq mT.$$

On the other hand, $\operatorname{rank}\boldsymbol{W} = \operatorname{rank}\boldsymbol{W}'$. Since W' has the following reduced row echelon form:

$$\begin{pmatrix} -I & B_1 & & \boldsymbol{O} \\ & -I & \ddots & \\ & & \ddots & B_{T-1} \\ \boldsymbol{O} & & & -I \end{pmatrix},$$

its rank is clearly mT. Finally, $mT \leq \operatorname{rank}\hat{\boldsymbol{X}} \leq mT$. This implies $\operatorname{rank}\hat{\boldsymbol{X}} = mT$. $\square$

The reader should notice that the rank of $\hat{\boldsymbol{X}}$ depends neither on the data matrices, $X_1, X_2, \cdots, X_T$, nor on the state transition matrices, $B_1, B_2, \cdots, B_{T-1}$. This signifies that classical least square techniques such as OLS or GLS are always applicable to estimate linear time-varying models that can be represented as state space models.[10]

[10] The reader also should note that the "identification" simply means invertibility of the matrix $\hat{\boldsymbol{X}}'\hat{\boldsymbol{X}}$. This remark is applicable to the next theorem.

In this chapter, our TV-VAR(1) model with constant drift. Such a model has two types of parameters to be estimated: time-varying and time-invariant. In the rest of this appendix, we confirm that the above discussion holds even if linear time-varying model such as our model has time-invariant parameters. In the case, we should modify equation (3.A.3) as follows:

$$\boldsymbol{y}_t = Z_t \boldsymbol{\alpha} + X_t \boldsymbol{\beta}_t + \boldsymbol{u}_t, \quad \boldsymbol{u}_t \overset{iid}{\sim} \mathcal{WS}(\mathbf{0}, \boldsymbol{\Sigma}_{\boldsymbol{u}}), \quad t = 1, 2, \cdots, T,$$

where $\boldsymbol{\alpha}$ is a $\ell-$vector of time-invariant parameters and Z_t is a $k \times \ell$ matrix of data. The regressor matrix turns to be

$$\left[\begin{array}{cc} \boldsymbol{Z} & \boldsymbol{X} \\ \hdashline \boldsymbol{O} & \boldsymbol{W} \end{array}\right],$$

where the above $\boldsymbol{O}$ is $mT \times \ell$ matrix of zero and

$$\boldsymbol{Z} = \begin{pmatrix} Z_1 \\ Z_2 \\ \vdots \\ Z_T \end{pmatrix}.$$

The following theorem holds.

Theorem 2

$$\operatorname{rank} \left[\begin{array}{cc} \boldsymbol{Z} & \boldsymbol{X} \\ \hdashline \boldsymbol{O} & \boldsymbol{W} \end{array}\right] = \operatorname{rank} \boldsymbol{Z} + mT.$$

Proof: Let $\tilde{X}$ denote the regressor matrix in the case. According to Theorem 1, rank$\hat{\boldsymbol{X}}$ is mT. Thus, using the reduced row echelon form of $\boldsymbol{W}$, say, $\boldsymbol{W}_0$, we can transform $\hat{\boldsymbol{X}}$ to the following form:

$$\left[\begin{array}{c} \boldsymbol{O} \\ \hdashline \boldsymbol{W}_0 \end{array}\right].$$

Then there are two non-singular matrices, P of size $((k+m)T) \times ((k+m)T)$ and Q of size $(\ell + mT) \times (\ell + mT)$, such that

$$P\tilde{\boldsymbol{X}}Q = \left[\begin{array}{cc} \boldsymbol{Z} & \boldsymbol{O} \\ \hdashline \boldsymbol{O} & \boldsymbol{W}_0 \end{array}\right].$$

Considering $\mathrm{rank}\boldsymbol{Z} = mT$ and

$$\mathrm{rank}\left[\begin{array}{c} \boldsymbol{Z} \\ \hdashline \boldsymbol{O} \end{array}\right] = \mathrm{rank}\boldsymbol{Z},$$

we have $\mathrm{rank}\tilde{\boldsymbol{X}} = \mathrm{rank}\boldsymbol{Z} + mT$. $\square$

Note that we can deal with the case of the TV-VAR(1) with time-invariant drift term such as our model in this chapter. As appeared in section 3.2.2, one regards $\boldsymbol{Z}$ as

$$\begin{pmatrix} I \\ I \\ \vdots \\ I \end{pmatrix},$$

where each I denotes $k \times k$ identity matrix. Then, $\mathrm{rank}\boldsymbol{Z} = k$. Considering the number of parameters to be estimated is $k + k^2T$ and $m = k^2$, our TV-VAR model has no problem of identification.

3.A.3 Non-Bayesian Time-Varying VAR Model

If the parameter constancy test exhibits instability of VAR coefficients over time, we choose an alternative model that holds the alternative hypothesis on the parameter dynamics and then estimate the model. Focusing our attention on linkage of stock markets which is supposed to vary over time, we suppose that only VAR coefficients vary over time while any components of the intercept term are invariant. In what follows, we present how to estimate the non-Bayesian TV-VAR model. Our approach has several good points in comparison with the Bayesian one (see, for example, Cogley and Sargent (2001, 2005) and Primiceri (2005)).

We start by setting VAR coefficients varying over time for an ordinary VAR(p) model as follows:

$$\boldsymbol{y}_t = \boldsymbol{\nu} + A_{1,t}\boldsymbol{y}_{t-1} + \cdots + A_{p,t}\boldsymbol{y}_{t-p} + \boldsymbol{u}_t,\ t = 1, 2, \cdots, T.$$

This equation can be represented as follows:

$$\boldsymbol{y}_t = \boldsymbol{\nu} + A_t Z_{t-1} + \boldsymbol{u}_t,\ t = 1, 2, \cdots, T, \tag{3.A.6}$$

where

$$A_t = [A_{1,t} \ \cdots \ A_{p,t}] \text{ and } Z_{t-1} = \begin{bmatrix} \boldsymbol{y}_{t-1} \\ \vdots \\ \boldsymbol{y}_{t-p} \end{bmatrix}.$$

Equation (3.A.6) corresponds to an observation equation when we regard our TV-VAR model as a state space model. As to the corresponding state equation, let us assume the following random walk process:

$$A_{i,t} = A_{i,t-1} + V_{i,t},\ i = 1, 2, \cdots, p \text{ and } t = 1, 2, \cdots, T,$$

where each $V_{i,t}$ is a $k \times k$ matrix of random variables, say, following normal distributions. Let V_t denote $= [V_{1,t} \ \cdots \ V_{p,t}]$ or equivalently

$$vec(A_t) = vec(A_{t-1}) + \boldsymbol{v}_t, \ t = 1, 2, \cdots, T, \tag{3.A.7}$$

where each $\boldsymbol{v}_t$ is a k^2p-vector of random variables, $vec(V_t) = vec([V_{1,t} \ \cdots \ V_{p,t}])$.

In place of the Kalman smoothing, we estimate $\boldsymbol{\nu}, A_1, \cdots, A_T$ by considering together Equations (3.A.6) and (3.A.7) as a simultaneous system of linear equations. Equation (3.A.6) turns out to be

$$\boldsymbol{y} = \begin{pmatrix} \boldsymbol{y}_1 \\ \boldsymbol{y}_2 \\ \vdots \\ \boldsymbol{y}_T \end{pmatrix} = \begin{pmatrix} I_k & Z_0' \otimes I_k & & \boldsymbol{O} \\ I_k & & Z_1' \otimes I_k & \\ \vdots & & \ddots & \\ I_k & \boldsymbol{O} & & Z_{T-1}' \otimes I_k \end{pmatrix} \begin{pmatrix} \boldsymbol{\nu} \\ vec(A_1) \\ vec(A_2) \\ \vdots \\ vec(A_T) \end{pmatrix} + \begin{pmatrix} \boldsymbol{u}_1 \\ \boldsymbol{u}_2 \\ \vdots \\ \boldsymbol{u}_T \end{pmatrix}.$$

This is equivalent to the following equation.

$$\boldsymbol{y} = [\mathbf{1} \otimes I_k \ \ \mathrm{diag}(Z_0' \otimes I_k, Z_1' \otimes I_k, \cdots, Z_{T-1}' \otimes I_k,)]$$
$$vec([\boldsymbol{\nu}, A_1, A_2, \cdots, A_{T-1}]) + vec(U), \tag{3.A.8}$$

where $\mathbf{1} = (1 \ 1 \ \cdots \ 1)' \in \boldsymbol{R}^T$. On the other hand, the state equation can be represented as:

$$\begin{pmatrix} -vec(A_0) \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{pmatrix} = \begin{pmatrix} \mathbf{0} & -I_{k^2p} & & & \\ \mathbf{0} & I_{k^2p} & -I_{k^2p} & & \\ \vdots & & \ddots & \ddots & \\ \mathbf{0} & & & I_{k^2p} & -I_{k^2p} \end{pmatrix} \begin{pmatrix} \boldsymbol{\nu} \\ vec(A_1) \\ vec(A_2) \\ \vdots \\ vec(A_T) \end{pmatrix} + \begin{pmatrix} \boldsymbol{v}_1 \\ \boldsymbol{v}_2 \\ \vdots \\ \boldsymbol{v}_T \end{pmatrix},$$

where $vec(A_0)$ is a prior vector of VAR coefficients, of whichever level is set, do not influence our results except starting burn-in periods.

Considering the following linear system, we estimate the coefficients of our time-varying VAR by using OLS.

$$\begin{bmatrix} \boldsymbol{y} \\ \boldsymbol{\gamma} \end{bmatrix} = \begin{bmatrix} D \\ W \end{bmatrix} \boldsymbol{\beta} + \boldsymbol{\varepsilon},$$

where $D = \left[\mathbf{1} \otimes I_k \,|\, \mathrm{diag}(Z_0' \otimes I_k, Z_1' \otimes I_k, \cdots, Z_{T-1}' \otimes I_k)\right]$,

$$\boldsymbol{\gamma} = \begin{bmatrix} -vec(A_0) \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{bmatrix},$$

$$W = \begin{pmatrix} \mathbf{0} & -I_{k^2p} & & & \\ \mathbf{0} & I_{k^2p} & -I_{k^2p} & & \\ \vdots & & \ddots & \ddots & \\ \mathbf{0} & & & I_{k^2p} & -I_{k^2p} \end{pmatrix},$$

$$\boldsymbol{\beta} = vec([\boldsymbol{\nu}, vec(A_1), \cdots, vec(A_T)])$$

and

$$\boldsymbol{\varepsilon} = \begin{bmatrix} \boldsymbol{u} \\ \boldsymbol{v} \end{bmatrix}.$$

The OLS estimate is:

$$\hat{\boldsymbol{\beta}} = \left(\begin{bmatrix} D \\ W \end{bmatrix}' \begin{bmatrix} D \\ W \end{bmatrix} \right)^{-1} \begin{bmatrix} D \\ W \end{bmatrix}' \begin{bmatrix} \boldsymbol{y} \\ \gamma \end{bmatrix}.$$

As Ito et al. (2012) show, the OLS estimate, $\hat{\boldsymbol{\beta}}$, of the above linear model provides the same estimation, $\hat{A}_t$, $(t = 1, 2, \cdots, T)$ as those of the Kalman smoothing for the original state space model. Not using iterative procedure of the traditional Kalman smoothing, our approach allows us to use familiar econometric techniques such as heteroskedastic autoeregressive consistent (HAC) estimations (see, for example, Newey and West (1987, 1994)). Such a HAC estimate enables us to obtain the covariance estimate of $\Sigma_{vec(A_t)}, (t = 1, 2, \cdots, T)$ from the covariance estimate, $\Sigma_{\boldsymbol{\beta}}$. It is possible to construct the critical values of the VAR coefficients for each period under the efficient market hypothesis by using the squared diagonal components of $\Sigma_{\hat{\boldsymbol{\beta}}}$. Furthermore, as shown in Technical Appendix 3.A.2, our approach never confronts with the problem of identifiability, involving the rank condition, unlike usual econometric analyses using least square techniques. In comparison with Bayesian approaches, our method is so simple that we are free from sometimes agonizing choice of prior distribution of parameters (for a typical example of the Bayesian TV-VAR model approach, see Cogley and Sargent (2001, 2005) and Primiceri (2005)).

3.A.4 Monte Carlo Method for TV-VAR Estimations

This subsection provides our method of statistical inferences on the TV-VAR estimates and their derived statistics, the degree of market efficiency. The idea is so simple that the Monte Carlo technique brings about their critical values under the hypothesis that any markets are efficient at any periods.

The practical procedure is as follows. We first estimate the means and standard deviations of time series of stock market returns over the periods, $\hat{\boldsymbol{\mu}} = (\hat{\mu}_1 \cdots \hat{\mu}_k)'$ and $\hat{\boldsymbol{\sigma}} = (\hat{\sigma}_1 \cdots \hat{\sigma}_k)'$ using the original data. Then, we derive N time series samples with length T by a Monte Carlo method from an i.i.d. normal distribution with the following means and variance structure:

$$\boldsymbol{y}_t^{(n)} = \boldsymbol{u}_t^{(n)}, \quad \boldsymbol{u}_t^{(n)} \overset{iid}{\sim} \mathcal{WS}\left(\begin{pmatrix} \hat{\mu}_1 \\ \vdots \\ \hat{\mu}_k \end{pmatrix}, \begin{pmatrix} \hat{\sigma}_1^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \hat{\sigma}_k^2 \end{pmatrix}\right), \ (n = 1, \cdots, N, \ t = 1, \cdots, T).$$

Note that we consdier as the efficient market hypothesis in the semi-strong sense as the null of our infernece. Thus, we suppose that the above uncorrelated disturbance vector $\boldsymbol{u}_t^{(n)}$, follows the above normal distribution. In other words, they are generated by a VAR process with time-invariant zero VAR coefficients.

Secondly we estimate the TV-VAR coefficients and residuals for our system of equations (3.A.7) and (3.A.7) in Section 3.2.2 from the above N artificial samples:

$$S_M := \{(\hat{A}_{1,t}^{(1)}, \cdots, \hat{A}_{p,t}^{(1)}, \hat{\boldsymbol{u}}_t^{(1)}, \hat{\boldsymbol{v}}_t^{(1)}), \cdots, (\hat{A}_{1,t}^{(N)}, \cdots, \hat{A}_{p,t}^{(N)}, \hat{\boldsymbol{u}}_t^{(N)}, \hat{\boldsymbol{v}}_t^{(N)})\}.$$

We derive N Monte Carlo samples of TV-VAR(p) model from S_M; it is natural to consider a distribution of the statistics with respect to each Monte Carlo samples. A set of N samples is available through the way shown in section 3.3 using the set of samples, S_M. Finally, we construct the critical values under the hypothesis from the N Monte Carlo samples.[11]

3.A.5 Bootstrap Method for TV-VAR Estimations

We can adopt another simulation method, bootstrap one, to attain the same goal as the Monte Carlo method does. The idea itself is very similar to that of the Monte Carlo technique in the previous section; it only differs by resampling process from the Monte Carlo technique.

The practical procedure is as follows. First we identify the stock returns data $\{\boldsymbol{y}_1, \cdots, \boldsymbol{y}_T\}$ with the residuals $D^0 = \{\hat{\boldsymbol{u}}_1, \cdots, \hat{\boldsymbol{u}}_T\}$ of VAR estimation under the hypothesis of all zero coefficients. Then we extract N samples $\{\tilde{\boldsymbol{y}}_1^{(i)}, \cdots, \tilde{\boldsymbol{y}}_T^{(i)}\}, i = 1, 2, \cdots, N$ with replace from D^0 regarding it as an empirical distribution of the residuals.

Secondly we estimate the TV-VAR coefficients and residuals for our system of equations (3.A.7) and (3.A.8) in Section 3.2.2 from the above N artificial samples:

$$S_b := \{(\tilde{A}_{1,t}^{(1)}, \cdots, \tilde{A}_{p,t}^{(1)}, \tilde{\boldsymbol{u}}_t^{(1)}, \tilde{\boldsymbol{v}}_t^{(1)}), \cdots, (\tilde{A}_{1,t}^{(N)}, \cdots, \tilde{A}_{p,t}^{(N)}, \tilde{\boldsymbol{u}}_t^{(N)}, \tilde{\boldsymbol{v}}_t^{(N)})\}.$$

We derive N bootstrap samples of TV-VAR(p) model from S_b; it is natural to consider a distribution of the statistics with respect to each bootstrap samples. A set of N samples is available through the way shown in Section 3.2.3 using the set of samples, S_b. Finally, we construct a sequece of the critical values from the N bootstrap samples in the same way as the Monte Carlo technique.[12]

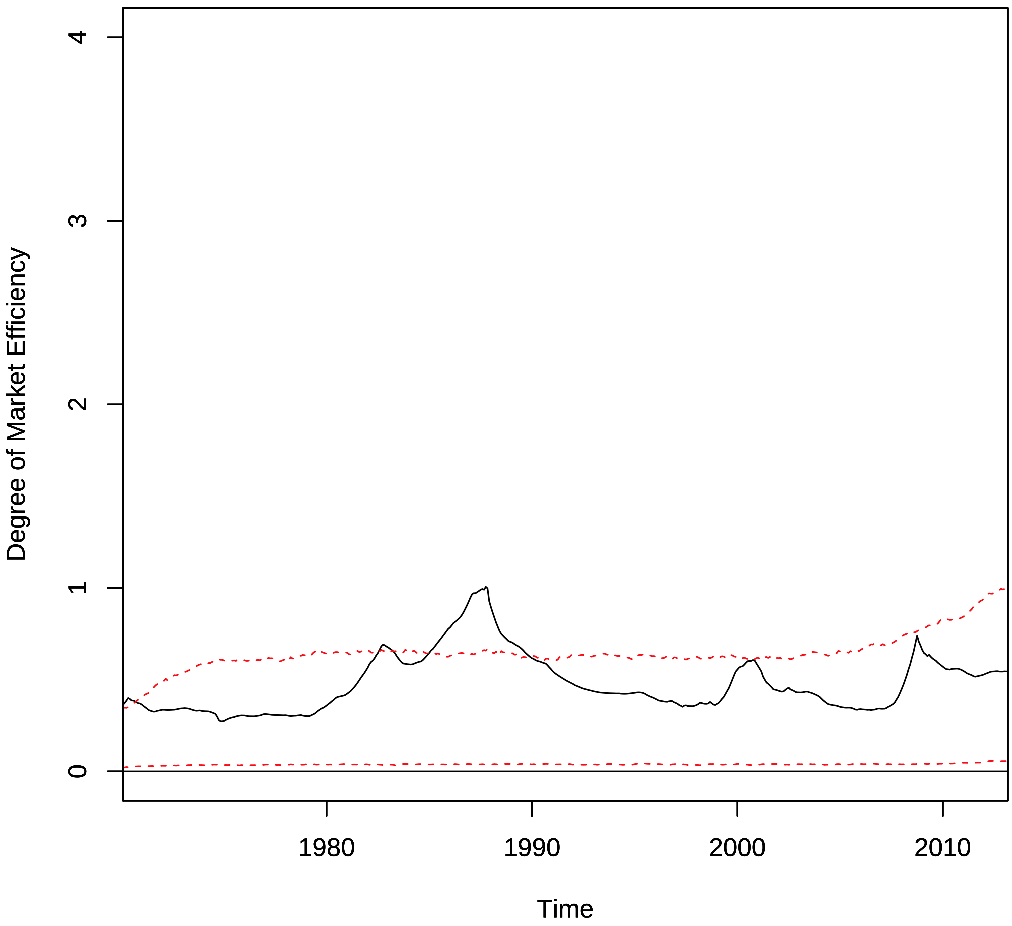
[11] Our Monte Carlo method is applicable to an AR(p) model for univarite data.
[12] Our bootstrap method is also applicable to an AR(p) model for univarite data.

Table 3.A.1: Time-Invariant AR Estimations

	$Constant$	R_{t-1}	R_{t-2}	$\bar{R}^2$	L_C
R_t^{US}	0.0043	0.1975	–	0.0356	28.3517
	[0.0018]	[0.0528]			
R_t^{CA}	0.0041	0.2361	–	0.0521	41.0607
	[0.0020]	[0.0558]			
R_t^{GB}	0.0047	0.3172	−0.1222	0.0886	39.3477
	[0.0021]	[0.0628]	[0.0526]		
R_t^{JP}	0.0027	0.2738	–	0.0711	37.5094
	[0.0021]	[0.0453]			
R_t^{DE}	0.0030	0.2512	–	0.0596	27.2011
	[0.0022]	[0.0454]			
R_t^{FR}	0.0038	0.2299	–	0.0492	47.9710
	[0.0023]	[0.0442]			
R_t^{IT}	0.0026	0.2119	–	0.0412	44.5418
	[0.0027]	[0.0496]			

Note: As for table 3.2.

Figure 3.A.1: Time-Varying Degree of Market Efficiency: North America



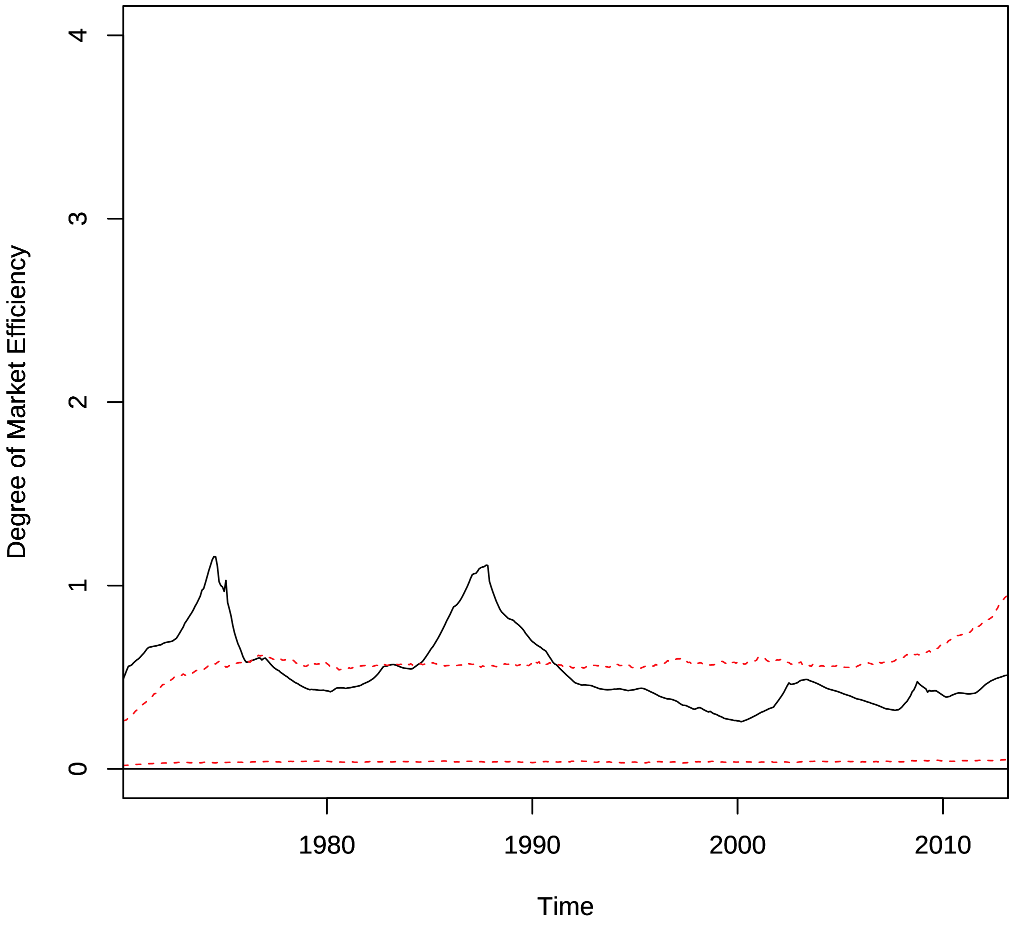
Notes:

(1) The dashed red lines represent the 99% critical values of the time-varying spectral norm in case of efficient market.

(2) We run 5000 times bootstrap sampling to calculate the critical values.

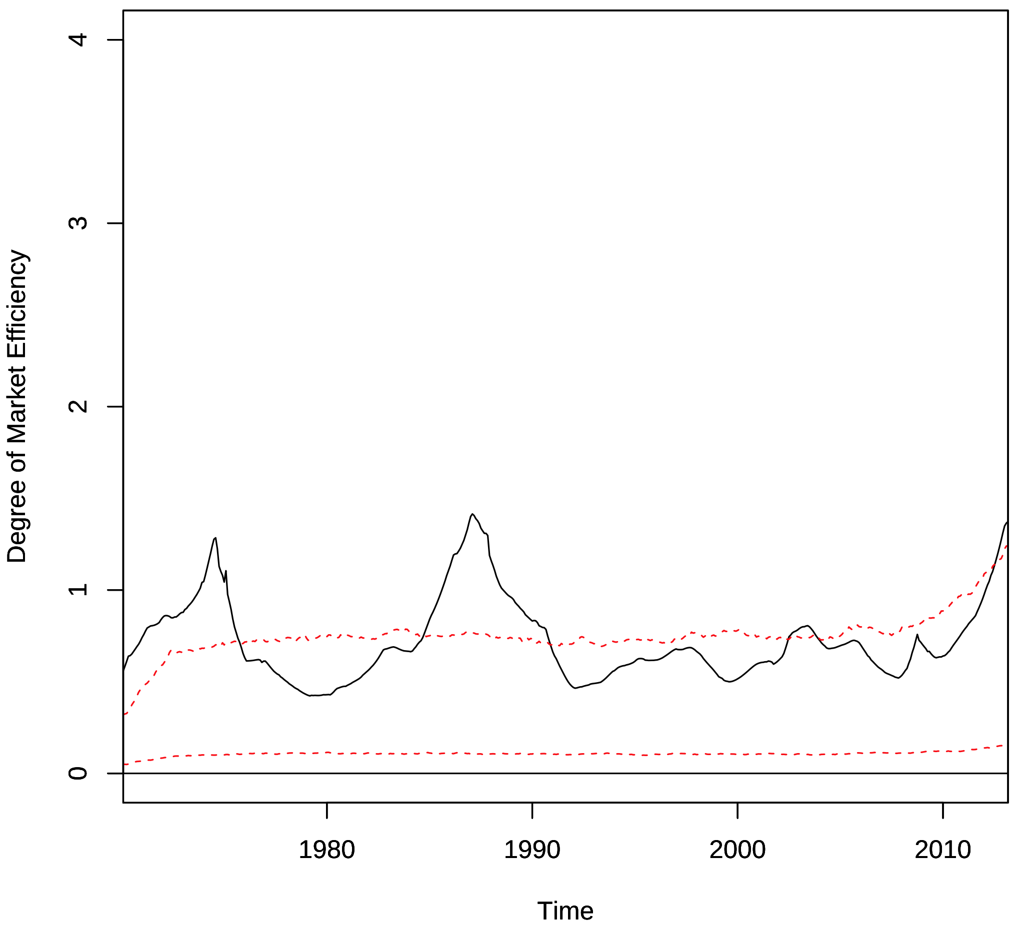
(3) R version 3.1.0 was used to compute the estimates.

Figure 3.A.2: Time-Varying Degree of Market Efficiency: U.S. and U.K.



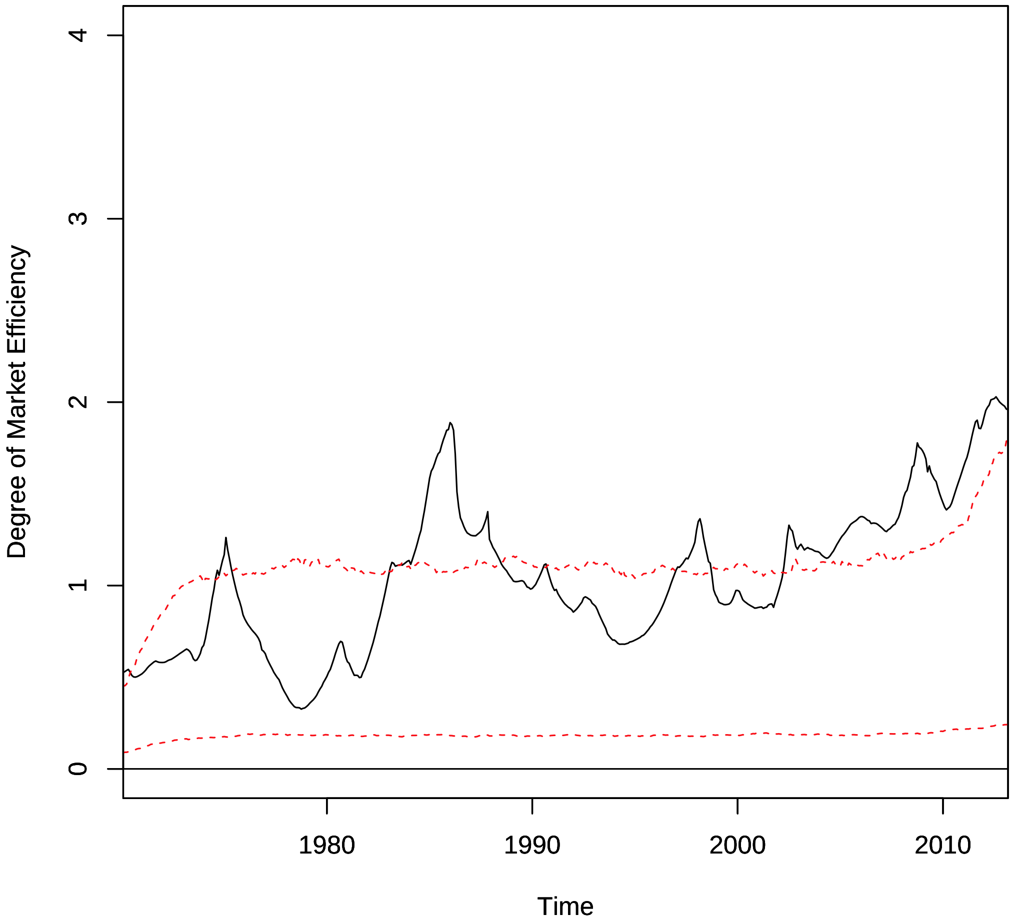
Note: As for figure 3.A.1.

Figure 3.A.3: Time-Varying Degree of Market Efficiency: U.S., U.K. and Japan



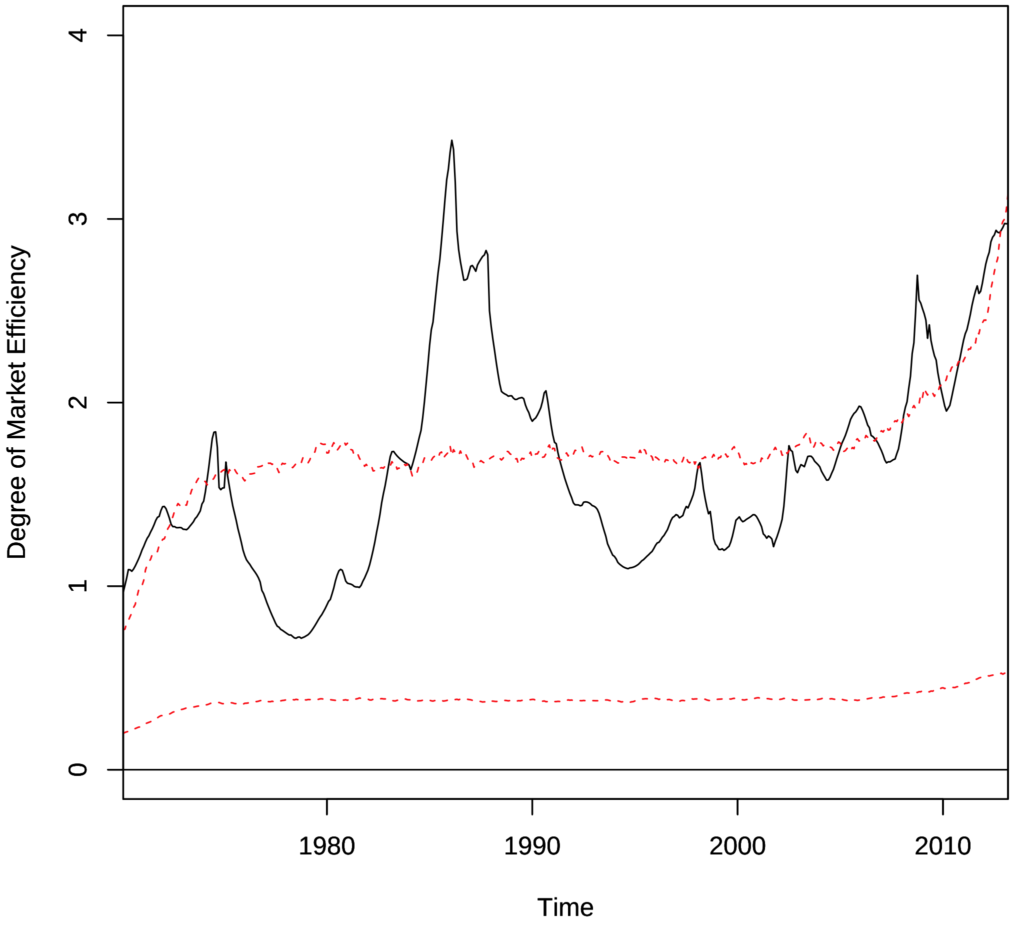
Note: As for figure 3.A.1.

Figure 3.A.4: Time-Varying Degree of Market Efficiency: European Countries



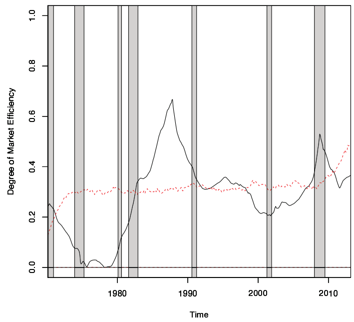
Note: As for figure 3.A.1.

Figure 3.A.5: Time-Varying Degree of Market Efficiency: G7 Countries

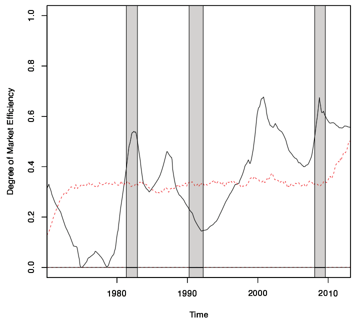


Note: As for figure 3.A.1.

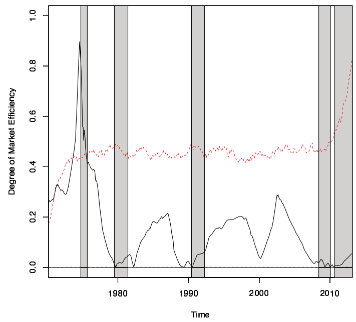
Figure 3.A.6: Time-Varying Degree of Market Efficiency: Individual Countries



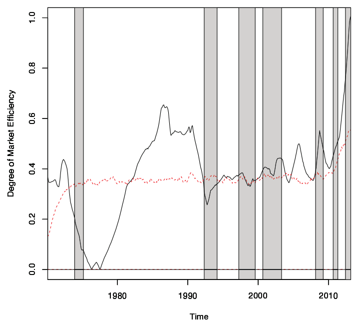
(a) U.S.



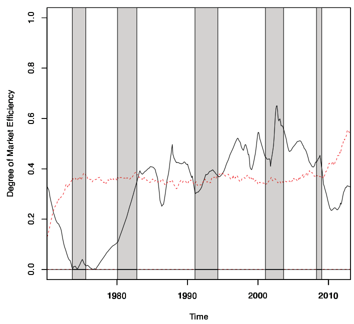
(b) Canada



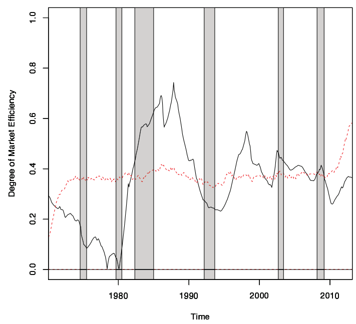
(c) U.K.



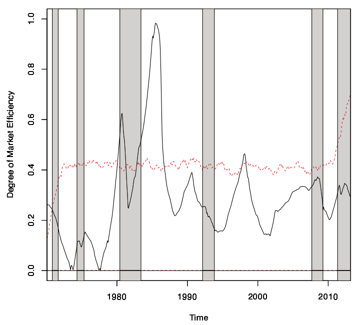
(d) Japan



(e) Germany



(f) France



(g) Italy

Notes:

(1) The dashed red lines represent the 99% critical values of the time-varying spectral norm in case of efficient market.

(2) We run 5000 times bootstrap sampling to calculate the critical values.

(3) The shade areas represent recessions reported by the NBER business cycle dates for the U.S. and the Economic Cycle Research Institute business cycle peak and trough dates for the other countries.

(4) R version 3.1.0 was used to compute the estimates.

Chapter 4

Futures Premium and Efficiency of the Rice Futures Markets in Prewar Japan

4.1 Introduction[1]

People have always confronted with price risks caused by some reasons, lean crop, wars, economic crises and etc. Historically, futures markets have been established in the world for hedging such risks; futures markets are characterized as well organized markets where forward contracts of a certain standardized commodity are transacted. The Osaka-Dojima rice futures market in the 18th century of Japan and the Chicago grain market in the 19th century of U.S. are famous historical examples. In particular, data of Japanese historical rice markets is famous for its richness. Accordingly, there is much literature about the markets. Miyamoto (1988), Kato (2001) and Takatsuki (2012) historically study the Osaka-

[1] This chapter is based on Ito, Maeda, and Noda (2014c).

Dojima rice futures market in the Tokugawa-period (1603-1868). They are interested in the hedging performance of the historical rice futures markets. As for the data of the Osaka-Dojima rice futures market, Ito (1993), Wakita (2001) and Takatsuki (2008) examine the famous equations of rice futures premium while Hamori, Hamori, and Anderson (2001) applies an error correction model to the rice futures premium. Kakizaka (2012) conducts some unit root tests for spot and futures prices of rice.

We also find literature about the rice markets in prewar Japan (1868-1932) when its economy was in a stage of takeoff. At this time, Japan had two major rice futures markets in Tokyo and Osaka. Taketoshi (1999) examines market efficiency of the Tokyo rice futures market by using Fama and French's (1987) model. Shizume (2011) studies rational expectation formation of rice futures in the Osaka rice futures market by using Hamilton's (1987) model. Nakanishi (2002) and Koiwa (2003) discuss evolution of the two major rice futures markets in Tokyo and Osaka considering the fact that correlation of their prices and shares of trade changed with time. Recently Ito, Maeda, and Noda (2014b) show that each market efficiency of Tokyo and Osaka rice futures markets varied with time. And they conclude that each time-varying structure of market efficiency strongly depended on often government interventions the two rice exchanges. However, they do not discuss interrelation among the two major rice markets in prices and distributions of rice.

In this chapter, we further focus our attention on the interrelation among the two major rice markets in prewar Japan in the view of market efficiency, considering they were under continual structural changes. Particularly we pose the following two questions. When were the futures premium supposed to be an unbiased predictor of the ex-post spot return under the rational expectations hypothesis? And to what extent did the above relation hold? To answer these questions, we apply a linear regression model of the spot returns on the futures premium to the two major rice markets. Nevertheless, our equation is not the one usually used for foreign exchange markets, but the one with time-varying parameters. That is, we estimate the parameters of the usual regression model at period by period supposing the two major rice markets have confronted with continual structural changes. A non-Bayesian time-varying model approach allows us to estimates them and to conduct a statistical inference by a residual-based bootstrap technique (see Ito et al. (2012, 2014d)). Our empirical results affirm the time-varying structure in the relation between the futures premium and the spot returns in prewar Japanese rice futures markets. It is found that such time-varying structure of the rice futures markets in prewar Japan corresponds well to historical changes in the Japanese colonial policy and domestic development of railroad system and port facilities.

This chapter is organized as follows. Section 4.2 provides a short historical review of rice markets in prewar Japan. Section 4.3 presents methodologies with respect to our non-Bayesian technique for a linear regression equation with time-varying parameters. This section also provides a new statistical inference for the above time-varying parameters based on a residual-based bootstrap method. The data on the rice futures markets in prewar Japan are described in Section 4.4. In Section 4.5, our empirical results show that the market efficiency in rice markets in prewar Japan varied over time and that its time-varying structure depended on the Japanese colonial policy for Taiwan and Korea. Section 4.6 concludes.

4.2 A Historical Review of the Rice Markets in Prewar Japan

Tokyo and Osaka, which this chapter mainly addresses, have been the two biggest cities in Japan since the Tokugawa-period (1603-1868). They also have been two centers of consumption and distribution of rice, a staple in Japan. Furthermore, Osaka was a cross-point of coastal routes (see Miyamoto (1988, pp.131-162) and Ishii (1986, pp.26-30)) in the Tokugawa-period. Thus the biggest market of rice was formed there because of its facility of transportation of rice as agricultural tax at the time. As well known, rice futures (*choaimai*) had been developed in Osaka to hedge price risks of rice since the early eighteenth century (see Suzuki (1940, pp.17-54)).

Following Osaka, a new market in rice futures was opened in Tokyo in 1870s. As population of Tokyo increased after the Meiji Restoration in 1868, the city had consumed more rice. This trend enhanced development of a new rice futures exchange and of a physical rice market for dealing with large volume of rice arrived from remote places (see Omameuda (2000, p.195)). Before 1870s, most of rice transported to Tokyo was from the neighboring region of Tokyo. As consumption of rice increased in Tokyo, an area named Fukagawa next to estuary of Sumida river, running through the center of Tokyo, attracted many merchants of rice. The merchants such as Mitsui Bussan were mainly wholesale dealers of rice cropped in remote places from Tokyo. They established Tokyo Rice Wholesalers Association in 1883 and then opened the Tokyo-Fukagawa spot rice market to deal with physical rice (see Omameuda (2004, pp.309-316)).

In the period after 1880s, which mainly concerns us, facilities of transportation and infrastructure of transaction were developed as for the rice market in Tokyo. Each of the rice exchanges in Osaka and Tokyo formed a major rice market respectively; nevertheless, they are not always integrated. After 1880s, because of population increase and general rise in living standards all over Japan, consumption of rice continuously increased. Japan in 1890s suffered from chronic shortage of rice. Then import of rice from Southeast Asia such as Vietnam and Thai began at the time. After the Russo-Japanese War (1904-1905), Japan imported more rice from Taiwan and Korea, both of which were Japanese colonies at the time. Note that such imported rice had a different variety than domestic rice. The former is indica rice and the latter japonica; they have difference in texture and taste.

Although price of imported rice was lower than that of domestic rice, most consumers in Japan did not always accept the rice as a staple. Thus even poor people in urban areas preferred uneaten food in barracks, boarding houses and restaurants, which were sold for them (see Omameuda (2007, pp.58-62)). As for imported rice until early 1910s, in general, poor people in Tohoku area, northern Japan and coal miners in Kyusyu area had consumed it. That is, ordinary consumers did not regard the imported rice as the same diet as domestic rice that they regularly ate (see Mochida (1970, pp.51-62)). However, the contents of imported rice began to change in 1910s; inland Japan imported more japonica rice from its colonies, Korea and Taiwan. The historical facts resulted from promotion of japonica rice cropping in the colonies. The Government-General of Korea first promoted the policy in 1910, and then Taiwan followed it in mid 1920s (see Tohata and Okawa (1939, pp.438-439) and Government–General of Taiwan (1945, p.244)). In fact, inland Japan imported more and more rice from Korea and Taiwan from 1920s. However, the imported rice at that time was not always consumed uniformly in all over Japan. In particular, most rice cropped in Korea was consumed in western Japan, including Osaka, because the shipping cost for the region was lower than that for eastern Japan (see Mochida (1970, p.140)). Because more japonica rice was cropped in Korea and Taiwan and more such rice was imported, the shortage of rice in inland Japan from 1890s and 1910 resolved by degrees.

4.3 Model and Empirical Method

This section presents our econometric method to examine the time-varying structure of the rice futures market in the Meiji, Taisho and early Showa periods Japan (1868-1945). It is partly based on Ito et al. (2014d), who analyze dynamic linkages of stock prices among G7 countries and examine efficiency of the international stock market. Unlike their method, ours focuses on the dynamic relation between the rice futures premium and the returns on spot rice in prewar Japan; we exmine the time-varying estimates of the famous equation for the futures premium. Employing a non-Bayesian time-varying model approach, we conduct statistical inference for the time-varying estimates by using a residual-based bootstrap technique.

4.3.1 Preliminaries

This chapter pays attention on the relation between between the rice futures premium and the returns on spot rice. Following many earlier studies (e.g., Ito (1993) and Wakita (2001)), we adopt the famous equation for the futures premium as our starting point:

$$(\log S_{t+k} - \log S_t) = \alpha + \beta(\log F_{t+k|t} - \log S_t) + u_t, \quad (t = 1, 2, \cdots, T-k), \tag{4.1}$$

S_t is the spot price at time t, $F_{t+k|t}$ is the futures price at time t for delivery of rice at time $t+k$ and u_t follows an independent and identically distributed process. That is, we consider the k-period forward contract. This type of regression equations is usually employed to test the unbiasedness hypothesis in the context of financial market efficiency. In practice, the null hypothesis is that $(\alpha, \beta) = (0, 1)$ and $\{u_t\}$ is serially uncorrelated. Note that the hypothesis refers to the joint hypothesis of risk neutrality (or no-risk premium) and rationality. In other words, the hypothesis says that no well informed speculators can expect to make excess returns (see Brenner and Kroner (1995) for details).

The vast literature on the exchange futures reports that the estimated β of Equation (4.1) is closer to or less than zero. It is well known as the futures premium puzzle, implying strong rejections of the unbiasedness hypothesis. Unlike most exchange futures data, our data of the rice futures market in prewar Japan provides diferrent features as will be shown in the later section.

4.3.2 Non-Bayesian Time-Varying Regression Models

In this chapter, we discuss time-varying structure of market efficiency of rice market in prewar Japan based on Equation (4.2). In particular, we are using the following equation:

$$(\log S_{t+k} - \log S_t) = \alpha + \beta_t(\log F_{t+k|t} - \log S_t) + u_t, \quad (t = 1, 2, \cdots, T-k). \tag{4.2}$$

This is a linear regression of the returns on spot rice on the rice futures premium. We concentrate our attention on whether β varies with time or not; we regard α as time-invariant considering its insignificance in the preceding works. Equation (4.2) itself cannot be estimated because it is not identifiable. Thus many preceding works estimating a model with time-varying parameters usually suppose a dynamic equation with respect to such parameters, say, a random walk. Thus we assume that β_t follows a random walk. It is represented as follows:

$$\beta_{t+1} = \beta_t + v_t, \quad (t = 1, 2, \cdots, T-k), \tag{4.3}$$

where each v_t follows an independent and identically distribution. We can regard Equations (4.2) and (4.3) together as a state space model. Following Ito et al.'s (2014d), we employ a non-Bayesian technique to estimate the parameters in the state space model using time series of S_t and $F_{t+k|t}$ as data.[2] We consider Equations (4.2) and (4.3) together to estimate the parameters, $\alpha, \beta_1, \cdots, \beta_{T-k}$ by using the following matrix form:

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_{T-k} \\ \beta_0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} 1 & x_1 & & & \boldsymbol{O} \\ 1 & & x_2 & & \\ \vdots & & & \ddots & \\ 1 & \boldsymbol{O} & & & x_{T-k} \\ 0 & 1 & 0 & & \boldsymbol{O} \\ 0 & -1 & 1 & & \\ \vdots & & \ddots & \ddots & \\ 0 & \boldsymbol{O} & & -1 & 1 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_{T-k} \end{pmatrix} + \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_{T-k} \\ v_1 \\ v_2 \\ \vdots \\ v_{T-k} \end{pmatrix}, \tag{4.4}$$

where $x_t = \log S_{t+k} - \log S_t$ and $y_t = \log F_{t+k|t} - \log S_t$ for $t = 1, 2, \cdots, T-k$. It helps us not only to estimate the coefficients but also to conduct statistical inference using a bootstrap technique as will be shown in the next subsection. One can estimates the state vectors $(\alpha\ \beta_1\ \beta_2 \cdots \beta_{T-k})'$ at one time based on observations $x_1, x_2 \cdots, x_{T-k}, y_1, y_2 \cdots, y_{T-k}$ and a prior β_0 given by OLS or GLS, familiar with economists. Note that such a vector estimated are identical with the the Kalman smoother with a fixed interval of the corresponding state space model, Equations (4.2) and (4.3).[3]

[2]See Technical Appendix 2.A.1 for details.
[3]See Technical Appendix 3.A.1 and Maddala and Kim (1998, ch.15) for details.

4.3.3 Statistical Inference for Time-Varying Parameters

This subsection presents our methodology of statistical inference on the time-varying estimates of β_t's in Equation (4.2). The idea is so simple that we examine the *estimates* with the joint distribution of our time-varying *estimators* $\hat{\beta}_1, \hat{\beta}_2, \cdots, \hat{\beta}_{T-k}$ under the unbiasedness hypothesis that $\beta_t = 1$ for any t and $\alpha = 0$. In particular, after setting a significance level, say, 95 percent, we construct upper and lower bounds based on such a distribution as two time series: $\beta_1^U, \beta_2^U, \cdots, \beta_{T-k}^U$ and $\beta_1^L, \beta_2^L, \cdots, \beta_{T-k}^L$. Then, using such a critical values forming the two series of bounds, we examine in what peridod the estimates $\hat{\beta}_t$'s are inside the band to identify the periods when the market is efficient.

However, one can hardly employ any asymptotic theory to construct the critical values above on time-varying coefficients supposed to follow a random walk. Such difficulty stems from their asymptotes involved in Browinian motion or Brownian bridge. It is quite hard to derive them theoretically; their distributions can only be represented with some Brownian motion or Brownian bridge if it were successfull. That is, such distributions need Monte Carlo technique when we use them to conduct some statistical inference. In fact, this obliges us to adopt a residual-based bootstrap technique to construct critical values of parameters of a state space model under the assumption of the unbiasedness hypothesis: $(\alpha, \beta_1, \cdots, \beta_{T-k}) = (0, 1, \cdots, 1)$ for our time-varying model. We are considering the case where each β_t's are estimated by the above method while the data were generated by Equation (4.1) when $(\alpha, \beta) = (0, 1)$.

In practice, our residual-based bootstrap techinique consists of the following steps.[4] First we fit Equation (4.1) to the observed data. Then we obtain a residual sequence $D_0 = (\hat{u}_1, \hat{u}_2, \cdots, \hat{u}_{T-k})$. Secondly, we extract N bootstrap samples $D_i = (\hat{u}_1^{(i)}, \cdots, \hat{u}_{T-k}^{(i)}), i = 1, 2, \cdots, N$ with replace from D_0 regarding it as an empirical distribution of the residuals.

[4]See Technical Appendix 3.A.3, of which base idea is found in Lütkepohl's (2005) Appendix D.3.

Thirdly, we generate N sets of $(\hat{u}_1^{(i)}, \cdots, \hat{u}_{T-k}^{(i)})$ for $i = 1, 2, \cdots, N$ by using D_i for $i = 1, 2, \cdots, N$. Fourth, we estimate N set of the time-varying coefficients and residuals by applying our time-varying model (4.4) to N bootstrap data, $(x_1, x_2, \cdots, x_{T-k}, y_1^{(i)}, y_2^{(i)}, \cdots, y_{T-k}^{(i)})$ and a prior, β_0 given, for $i = 1, 2, \cdots, N$. Finally, we construct critical values of $(\alpha, \beta_1, \cdots, \beta_{T-k})$.

4.4 Data

We utilize the monthly data on the rice spot and futures prices of the Tokyo and Osaka rice exchanges in prewar Japan. As for the rice futures data, three contract months are available: nearby contract (one month), second nearest contract (two months), and deferred contract (three months). These datasets consist of the following two statistics: (i) Nakazawa (1933) for the Tokyo rice exchange from October 1880 through November 1932, and (ii) Miyamoto, Uekawa, Uemura, Shiba, Hiroyama, Chakepaichayon, Mori, and Sakata (1979), the *Osaka City Statistics* and the Ministry of Commerce and Industry (1931, p.342) for the Osaka rice exchange from April 1881 through November 1932.[5] There are a few missing values found in both statistics. Therefore, we fill the missing values by using seasonal Kalman filter. We take the first difference of natural log of the rice spot prices to obtain the returns on spot rice. And we subtract the natural log of the rice spot prices from the natural log of the rice futures prices to calculate the rice futures premium.

(Table 4.1 around here)

From Table 4.1, we can confirm that the more the contract month is long, the more the futures premium is volatile. The table also shows the results of the unit root test along

[5] Note that the rice futures data of nearby and nearest contract months in Tokyo have different starting date from that of deferred contract month. In particular, the data of nearby contract month is available from January 1888, and the data of nearest contract month is available from January 1898.

with descriptive statistics of the data. For our estimations, all variables that appear in the moment conditions should be stationary. To confirm whether the variables satisfy the stationarity condition, we apply the ADF-GLS test of Elliott et al. (1996). We employ the Modified Bayesian Information Criterion (MBIC) instead of the Modified Akaike Information Criterion (MAIC) to select the optimal lag length. This is because, from the estimated coefficient of the detrended series, $\hat{\psi}$, we do not find the possibility of size-distortions (see Elliott et al. (1996); Ng and Perron (2001)).

4.5 Empirical Results

4.5.1 Time-Invariant Regression Model

We assume a time-invariant regression model with constant parameters for our preliminary estimations. Our estimation results for a time-invariant regression model with whole sample is summarized in Table 4.2: all estimates of α are almost zero, and as the contract month of the rice futures is longer, the corresponding estimates of β are increasing. These results suggest that the more the contract month is long, the more the rice market in prewar Japan is efficient.[6] Taketoshi (1999) obtains different estimates of β for each subsamples when he splits the whole sample into four subsamples, but he does not examine whether the rice market in prewar Japan is efficient using conventional statistical inferences. We therefore use Ito et al.'s (2012) non-Bayesian time-varying regression model to estimate the β because the rice market in prewar Japan may be not always efficient through the time. However, we should verify if the time-varying regression model is more appropriate than the time-invariant regression model before we adopt a new approach. Then we apply

[6]Note that our estimates of β are little larger than Taketoshi's (1999) ones because our dataset is seasonally unadjusted.

Hansen's (1992) parameter constancy test to investigate whether the time-invariant model is a better fit for our data.

(Table 4.2 around here)

Table 4.2 presents the results of preliminary estimations: the estimates of time-invariant regression models (α and β), and their corresponding Hansen's (1992) joint parameter constancy test statistics (L_C). For our time-invariant regression model, the joint parameter constancy test rejects the null hypothesis of constancy at the 5% level, against the alternative hypothesis stating that the parameter variation follows a random walk process. These results suggest that the time-invariant regression model does not accommodate our data; rather, we should use the time-varying regression model for the data of prewar Japanese rice markets.

4.5.2 Time-Varying Regression Model and Market Efficiency

Figures 4.1 and 4.2 show that the farthest contract month transaction is more efficient than those of the nearest and next-nearest ones in both the Tokyo and Osaka rice exchanges.

(Figures 4.1 and 4.2 around here)

That is, the longer the contract month (maturity) in futures, the more successful for traders to hedge price risks in the rice market. It resulted from the large amount of the farthest contract month transaction in futures trading: 72.1 percent of whole futures in Tokyo and 62.7 percent in Osaka (see Kamibayashi (1935)). The fact suggests that such transaction included a large part of required information to hedge price risks in the rice market in prewar Japan. However, examining Figures 4.1 and 4.1 carefully, one can find some periods when even the futures of farthest contract month failed to hedge price risks. Specifically, Figure

4.1 exhibits three periods as for the Tokyo rice exchange: late 1880s, late 1890s and from mid 1900s to 1910s; Figure 4.2 exhibits two periods as for the Osaka rice exchange: late 1890s and from mid 1900 to mid 1920s. In what follows, we discuss what made even the farthest month transaction sometimes fail to hedge price risks in the rice market: first, Tokyo and second Osaka in order.

First, we discuss the Tokyo rice exchange in late 1880s. From 1884 to 1889, rice export shipped from the Port of Kobe to Europe noticeably grew. The growing export resulted from upturn of terms of trade, caused by stable price of silver, and decrease in price of domestic rice at the time. The rice export from Kobe amounted to 1.39 million *koku* in 1888, which is more than the amount of physical rice arrived in the Tokyo-Fukagawa rice spot market, 1.12 million *koku* in the same year. Such large amount of rice export deterred further decrease in rice price. In particular, rice price in Tokyo subsequently remained at stable levels with a support line under influence of the movement in rice market in the Osaka-Kobe area (see Omameuda (1993, pp.18-21)). This situation resulted from both the spot and futures prices of rice depending on those of the Osaka rice exchange in late 1880 (see Ito et al. (2014b)). These observations provide an evidence that the rapid increase in rice export from the Port of Kobe in late 1880s triggered the strong influence of rice price in western Japan, covering Osaka and Kobe on rice price in eastern Japan, including Tokyo. Both the prices in spot and futures of rice in Tokyo essentially depended on those in Osaka and the latter deteriorated primary relation between spot and futures prices of rice at that time. However, population increase at a high growth rate of big cities such as Tokyo caused growing consumption of rice after Japan experienced a serious rice failure in 1890. Then, rice export showed rapid decrease and the above link between rice prices in Tokyo and Osaka deteriorated showed recovery. In early 1890s, it became possible to hedge price risks of rice in the Tokyo rice exchange.

Second, we discuss the market inefficiency in late 1890s. On the contrary to the case of late 1880s, import of rice rapidly grew in the period.

(Figures 4.3 and 4.4 around here)

Figure 4.3 exhibits amounts of physical rice arrived in the Tokyo-Fukagawa spot rice market for each place of production in the Tokyo-Fukagawa rice spot market. Figure 4.4 exhibits the rate of imported rice to all physical rice arrived in the market. They show that imported rice explosively increased; an amount of the imported rice had 43 percent of all physical rice in the market. This rapid increase in import of rice resulted from lean crop of rice in 1897 due to bad weather (see Ota (1938, p.184)). The imported rice from Indochina, cropped in Yangon, Ho Chi Minh and Taiwan are categorized as indica rice while domestic rice is categorized as japonica rice. Thus because of difference in texture and taste, consumers in Japan regarded such imported rice as alternative grain. In practice, they mixed domestic rice and the imported one when they cooked (see Omameuda (1993, pp.37-40)). Notice that both the Tokyo rice exchange and Osaka-Dojima rice exchange used to deal with only two brands of rice, Musashi and Settsu, of a standard variety as domestic rice. This situation disrupted the rice market at that time. In fact, the rice market in Japan at that time dealt with different grain since only domestic rice was listed in futures market and both the imported and domestic rice was traded in spot market. Thus, the futures price failed to be a fine index of the expected price of rice. Such a situation reappeared when rice export grew again after mid 1900s.

Third, we consider market inefficiency from mid 1900s to 1910s. Figure 4.4 shows that imported rice rose its rate to all physical rice arrived in the Tokyo-Fukagawa rice spot market again after mid 1900s. At the same time, futures in the Tokyo rice exchange failed to hedge price risks of physical rice. In 1910s, the imported rice often amounted to more than 50 percent of all physical rice arrived in the market; rice futures failed to hedge effectively price risks of in the rice market. However, in 1920s, rice futures appeared to work well again. In order to explain this phenomenon, we need further discussion about drastic change in substance of rice cropped in Korea and Taiwan in what follows.

Japan experienced *kome-soudou* (the Rice Riots) in 1918. It caused large unrest in Japanese economy. Turmoil in rice market was settled in 1920; spot price of rice lowered and an amount of imported rice decreased (see Figure 4.4). At the same time, the rate of imported rice to all physical rice arrived lowered to around 30 percent. In 1920s, amounts of physical rice arrived for each place of production were changing. From 1917 to 1919, their rates to all physical rice arrived in the Tokyo-Fukagawa rice spot market are as follows. The imported rice exhibited 60.7 percent, rice cropped in Taiwan 24.5 percent and rice cropped in Korea 14.7 percent respectively. From 1921 to 1923, each rates exhibits 27.7 percent, 34.1 percent and 38.2 percent in the same order. The amount of foreign rice rapidly decreased (see Sasaki (1937, pp.278-279,286)); rice cropped in Korea shipped to Japan was japonica rice in the period. After 1910s, the Government-General of Korea promoted rice cropping of the same varaiety as domestic rice. In fact, as for the rate of such rice to whole rice production in Korea, it exhibited only 5 percent in 1912, 69 percent in 1921 and 79 percent in 1932 (see Tohata and Okawa (1939, pp.438-439)). Since japonica rice cropping was introduced to Korea, rice cropped in Korea had occupied the central position in physical rice arrived in the Tokyo-Fukagawa rice spot market. In effect, it amounted to 74.3 percent of the imported rice from 1927 to 1929 (see Sasaki (1937, pp.278-279,286)). Furthermore, following Korea, the Government-General of Taiwan introduced japonica rice cropping in 1924 (Government–General of Taiwan (1945, p.244)). Such promotion of japonica rice cropping in Korea as well as Taiwan made rice cropped in Korea and rice cropped in Taiwan amount to 94.4 percent of whole imported rice from 1930 to 1932 (see

Sasaki (1937, p.286))). This means that the fraction of indica rice of the imported rice arrived in rice spot market rapidly became very small. Thus, the imported rice had hardly affected pricing in rice market in futures as well as spot even under the situation of a large amount imported rice. In summary, the change in variety of cropping rice in Korea and Taiwan let people inland consume more imported rice; it also let the rice exchanges work better in hedging price risk.

In the following part, we discuss the Osaka-Dojima rice exchange. First, we study the market inefficiency in late 1890s. In this period, distributions of imported rice in Osaka increased in the same way as Tokyo.

(Figures 4.5 and 4.6 around here)

Figure 4.5 shows amounts of rice in year's end stock from 1891 to 1914 and an amount of physical rice in circulation for each year from 1909 to 1932 respectively. Figure 4.6 shows rates of imported rice to rice in year's end stock from 1891 to 1914 and to rice in circulation from 1909 to 1932. Because there is no statistics presenting volume of transaction of rice before 1908, we use rice in year's end stock as a substitute of volume of transaction of rice for such periods. Figure 4.6 demonstrates that rate of imported rice to whole rice in year's end stock rose from 1896 to 1898. Figures 4.4 and 4.6 show that the rate kept at relatively higher level than that of Tokyo while it was less than 10 percent. This resulted from the two facts: large portion of rice cropped in Korea was landed at the Port of Kobe and the Port of Osaka and it was consumed in Kinki and Chugoku regions (see Okada (1911, p.38)). In 1890s rice cropped in Korea was dealt more in the Osaka rice market than in the Tokyo market; rice cropped in Korea at that time was indica rice before introduction of japonica rice cropping as described above. Thus futures trade in the Osaka-Dojima rice exchange failed to hedge price risk of physical rice for exactly the same reason as in the Tokyo rice exchange.

Second, we discuss the market inefficiency from mid 1900s to mid 1920s. As already pointed out, futures in the Tokyo rice exchange in 1910s failed to provide a fine index of the expected price of spot rice. In the Osaka-Dojima rice exchange had experienced market inefficiency for 20 years, longer than Tokyo. Such longer inefficiency resulted from two factors. The first factor is strong dependence of the Osaka rice exchange on the Tokyo rice exchange. Futures price in the Osaka-Dojima rice exchange had become dependent on that of the Tokyo rice exchange since mid 1900s. More specifically, the former depended on the latter and not vice versa from mid 1900s to early 1920s as Ito, Maeda, and Noda (2014a) pointed. According to Figure 4.3, we find that more rice cropped in western Japan was shipped to the Tokyo-Fukagawa rice spot market at the period from mid 1900s to early 1920s. Large portion of such rice came from Kyushu region. In particular, rice cropped in Kyusyu accounted for more than 90 percent of the rice cropped in western Japan shipped to the Tokyo-Fukagawa rice spot market in 1910s. Such rice cropped in Kyusyu had been transported to regions on the shore of the Setouchi Sea, say, Osaka, up until mid 1900. Then, as railroad system and port facilities in Kyusyu area were developed, in turn, rice cropped in Kyusyu was transported to Tokyo (see Nakamura (2003, pp.68-72)). This change of distribution of rice cropped in Kyusyu let the Tokyo rice market deal with rice cropped in all over Japan while the exchange had mostly dealt with rice from eastern Japan, including Tohoku area. Finally, in this period, the Osaka-Dojima rice market, which mostly dealt with rice cropped in western Japan, became to take the second position in rice market in Japan. Consequently, price of rice futures in the Osaka-Dojima rice exchange strongly depended on that in the Tokyo rice exchange. This means that significance of the former deteriorated. The second factor is the fact that more imported rice had been traded in Osaka since 1910s. Figure 4.6 demonstrated that the rate of imported rice to physical rice in circulation increased from 1910s. These two factors let the function to price futures effectively deteriorate in the Osaka-Dojima rice exchange.

In mid 1920s, in turn, the function then ameliorate. Such change of market efficiency comes from the following two points. The first point is that an amount of rice cropped in western Japan transported to Tokyo decreased. Figure 4.3 proves that the amount began to decrease in the Tokyo-Fukagawa rice spot market in mid 1910s. This phenomenon resulted from soaring transportation cost due to World War I and increasing consumption of rice in Kyusyu area caused by industrialization of northern Kyusyu area (see Sasaki (1937, p.277-278) and Mochida (1970, pp.186-187)). The second point is that structure of distributions of physical rice in Tokyo and Osaka had changed since mid 1920s. Figure 4.4 shows that imported rice amounted to about 60 percent of all physical rice arrived in the Tokyo-Fukagawa rice spot market in mid 1920s. At the same time, imported rice accounted for from 30 to 40 percent of total physical rice in circulation of Tokyo (see Sasaki (1937, p.273-274,281-282)). In Osaka, imported rice amounted to about 80 percent of all physical rice in circulation in mid 1920s as Figure 4.6 exhibits. That is, domestic rice was dealt in Tokyo while imported rice was dealt in Osaka. Note that about 95 percent of imported rice consisted of rice cropped in Korea of the same variety as domestic rice because of promotion of japonica rice cropping (see Osaka City Government (1931, ch.5,p.42)). Thus, in the same situation as the Tokyo rice exchange, rice price in futures provided an efficient index of the expected spot price of rice.

To sum up, we assert that the Osaka-Dojima rice exchange attained a significant position in rice futures again because the amount of rice cropped in western Japan decreased in rice market in Tokyo from mid 1910s and the Osaka rice market had different structure than the Tokyo one from mid 1920s. Finally, futures price in the Osaka-Dojima rice exchange became again an effective index of expected price of physical rice in Osaka.

4.6 Concluding Remarks

This chapter discusses each market efficiency of the Tokyo and Osaka rice futures markets in prewar Japan using a non-Bayesian time-varying model approach. An extent to which rice futures markets hedges price risks depend on time; our estimates of β in the famous equation for the futures premium vary with time. The estimates of time-varying β for Tokyo and Osaka suggest how each markets deviated from the efficient market respectively. Considering the historical data of rice distribution in prewar Japan, we elucidate that increasing import of japonica rice after 1900s affected each market efficiency of the Tokyo and Osaka rice futures markets. Such increase in imported rice was caused by Japanese colonial policy of rice cropping in Korea and Taiwan.

There are three periods when the Tokyo rice futures market was inefficient: late 1880s, late 1890s and from mid 1900s to 1910s. At first, rice export grew since the goverment promote rice export to improve low price of domestic rice in late 1880s. The most of rice was shipped to overseas from the Port of Kobe. The rapid increase in exported rice influenced the rice futures market in Osaka. At that time, rice price in Tokyo strongly depended on that in Osaka since the Osaka futures market dominated the Tokyo's. Thus, the Tokyo futures market experienced market inefficiency. In late 1890s, Tokyo confronted with population increase at a high rate in late 1890s. It caused rapid increase in demand of rice, Japanese staple. Accordingly, more imported rice was required while it was indica rice. Such increasing import of rice caused disorder of rice distribution; different varieties of rice was distributed in Tokyo. Then the disorder let the Tokyo rice futures market inefficient. In turn, from mid 1900s to 1910s, imported rice increased at a high rate whereas most of the rice became japonica variety because of Japanese colonial policy of rice cropping in Korea and Taiwan. This rapid increase in imported rice from Korea and Taiwan caused temporary inefficiency in the Tokyo rice futures market from mid 1900s to 1910s. Since the imported rice was of the same variety as domestic one, the Tokyo rice futures market was efficient even in the period including the Rice Riots of 1918.

The Osaka rice futures market experienced two periods of inefficiency: late 1890s and from mid 1900 to mid 1920. The inefficiency in the Osaka rice futures market in late 1890s resulted from the increased in imported rice from Korea and Taiwan in the same way as Tokyo. In contrast, the inefficiency from mid 1900s to mid 1920s resulted from different reason than that of 1890s case. The Osaka rice futures market depended on the Tokyo rice futures market. At the same time, as railroad system and port facilities in Kyusyu were developed, more rice became shipped from western Japan, especially Kyusyu, to Tokyo in eastern Japan than before. Thus, such change in rice distribution caused the inefficiency from mid 1900s to mid 1920s.

We conclude that the two major rice futures markets in Tokyo and Osaka in prewar Japan confronted with the change of market efficiency. The historical facts resulted from imported rice from Korea and Taiwan, where the two Government-Generals promoted japonica rice cropping as a colonial policy. The time-varying structures of market efficiency in the Tokyo and Osaka rice futures markets differed from each other because of their facilities of transportation of imported rice.

Table 4.1: Descriptive Statistics and Unit Root Tests

	Tokyo						Osaka					
	One Month		Two Month		Three Month		One Month		Two Month		Three Month	
	SR	FP	SR	FP	SR	FP	SR	FP	SR	FP	SR	FP
Mean	0.0025	−0.0172	0.0015	−0.0172	0.0024	−0.0176	0.0013	−0.0227	0.0027	−0.0212	0.0042	−0.0201
SD	0.0536	0.0688	0.0908	0.0822	0.1145	0.0902	0.0619	0.0789	0.0923	0.0920	0.1171	0.1045
Min	−0.3670	−0.3623	−0.4601	−0.3600	−0.5360	−0.3549	−0.2939	−0.3780	−0.3821	−0.5279	−0.4582	−0.5991
Max	0.2791	0.1769	0.3116	0.2037	0.4752	0.1618	0.6569	0.2190	0.7260	0.3132	0.8131	0.2494
ADF-GLS	−16.1633	−6.9630	−4.2744	−5.8867	−6.6164	−7.9606	−13.8955	−8.8637	−5.7255	−7.9995	−4.8899	−7.5746
Lags	0	0	8	0	10	1	1	0	8	1	2	18
$\hat{\phi}$	0.3442	0.8317	0.7294	0.8444	0.7843	0.8629	0.1968	0.7739	0.6053	0.7220	0.7153	0.7657
$\mathcal{N}$	538		417		624		619		618		617	

Notes:

(1) "ADF-GLS" denotes the ADF-GLS test statistics, "Lags" denotes the lag order selected by the MBIC, and "$\hat{\phi}$" denotes the coefficients vector in the GLS detrended series (see equation (6) in Ng and Perron (2001)).

(2) In computing the ADF-GLS test, a model with a time trend and a constant is assumed. The critical value at the 1% significance level for the ADF-GLS test is "−3.42".

(3) "$\mathcal{N}$" denotes the number of observations.

(4) R version 3.1.0 was used to compute the statistics.

Table 4.2: Time-Invarant Estimations

	Tokyo			Osaka		
	One Month	Two Month	Three Month	One Month	Two Month	Three Month
α	0.0056	0.0066	0.0118	0.0076	0.0100	0.0119
	[0.0028]	[0.0096]	[0.0085]	[0.0030]	[0.0058]	[0.0106]
β	0.1771	0.2927	0.5298	0.2771	0.3424	0.3823
	[0.0743]	[0.1682]	[0.1572]	[0.0435]	[0.0693]	[0.1023]
$\bar{R}^2$	0.0500	0.0679	0.1728	0.1232	0.1150	0.1149
L_C	1.0740	1.0510	1.6171	1.5787	1.8331	2.4301

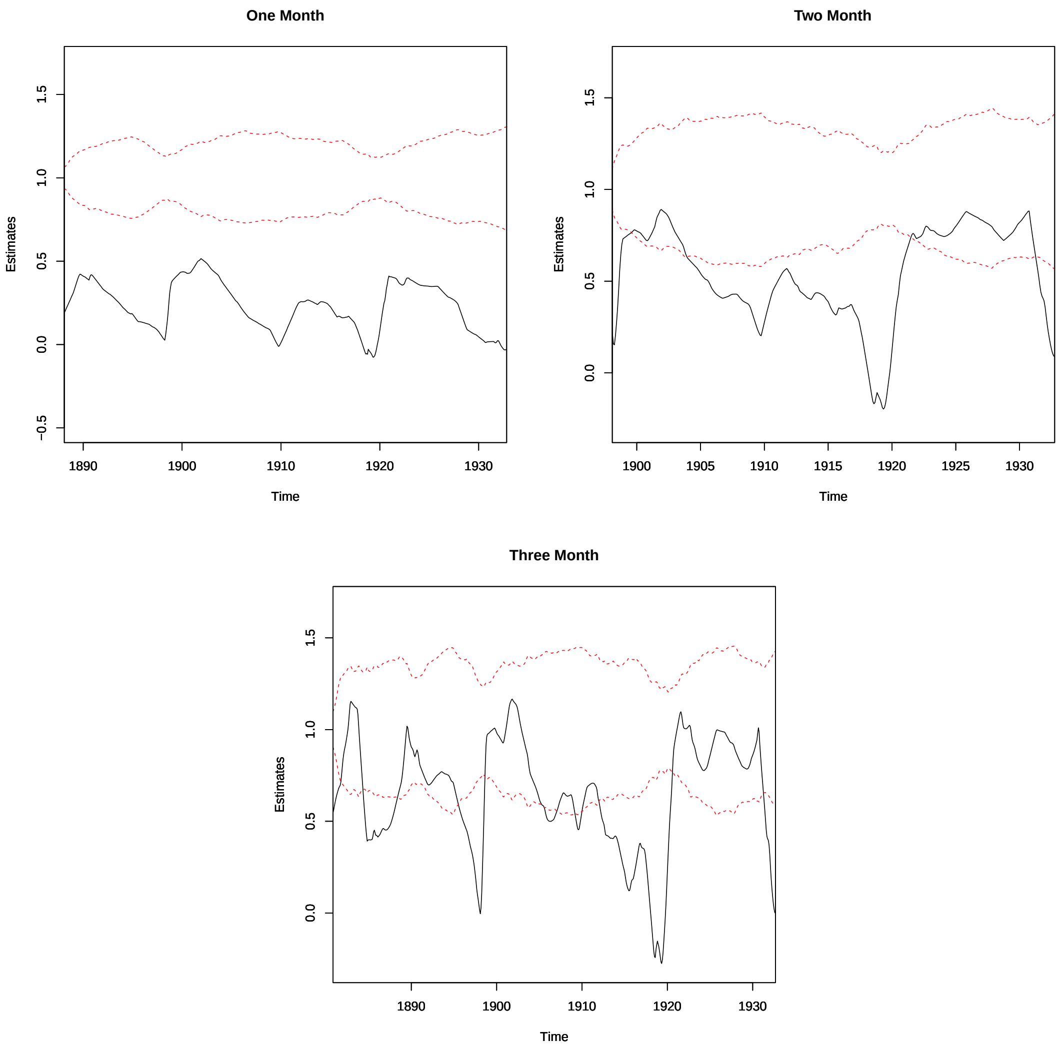
Notes:

(1) "$\bar{R}^2$" denotes the adjusted R^2, and "L_C" denotes Hansen's (1992) joint L statistic with variance.

(2) Newey and West's (1987) robust standard errors are in brackets.

(3) R version 3.1.0 was used to compute the estimates and the statistics.

Figure 4.1: Time-Varying Estimates of β: The Case of Tokyo Rice Market



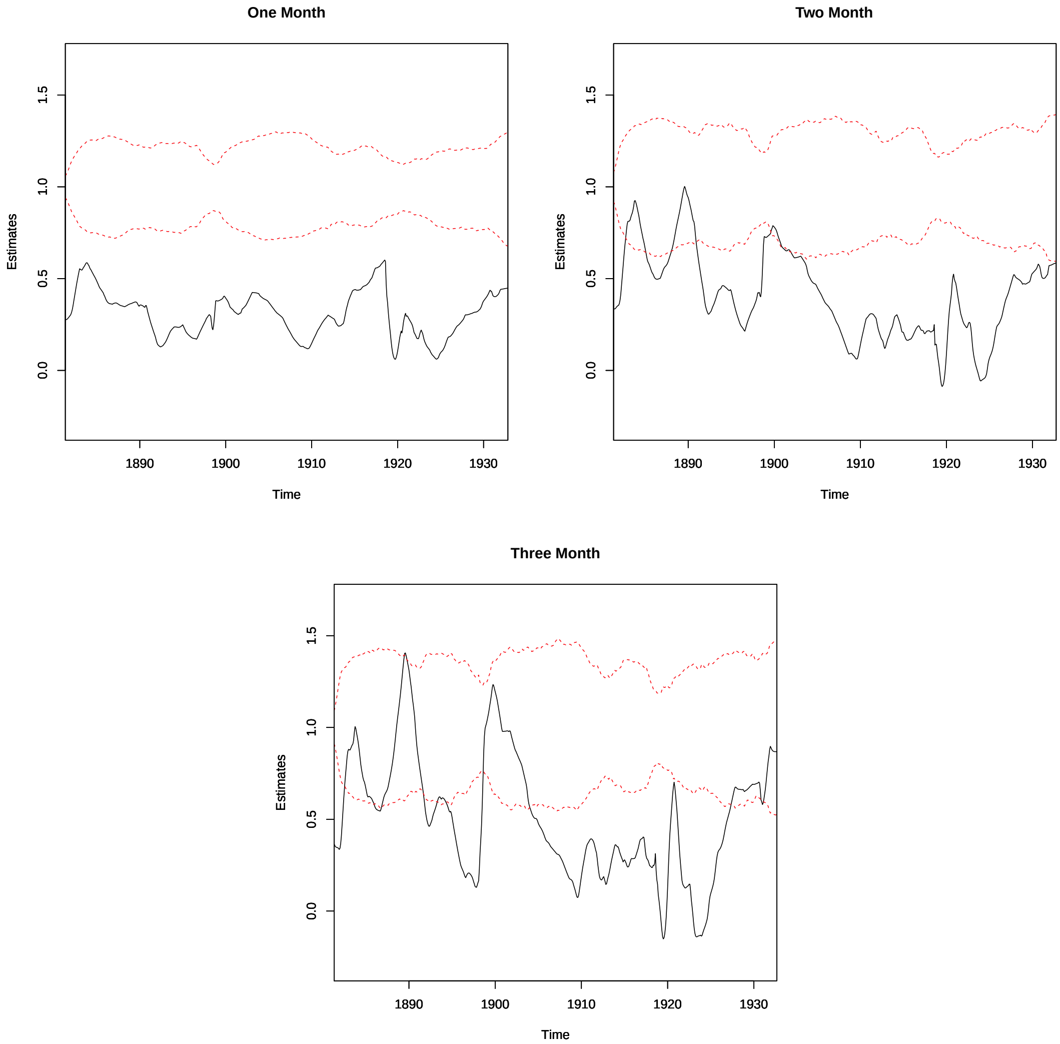
Notes:

(1) The dashed red lines represent the 95% ccritical values of the estimates in case of efficient market.

(2) We run 5000 times bootstrap sampling to calculate the critical values.

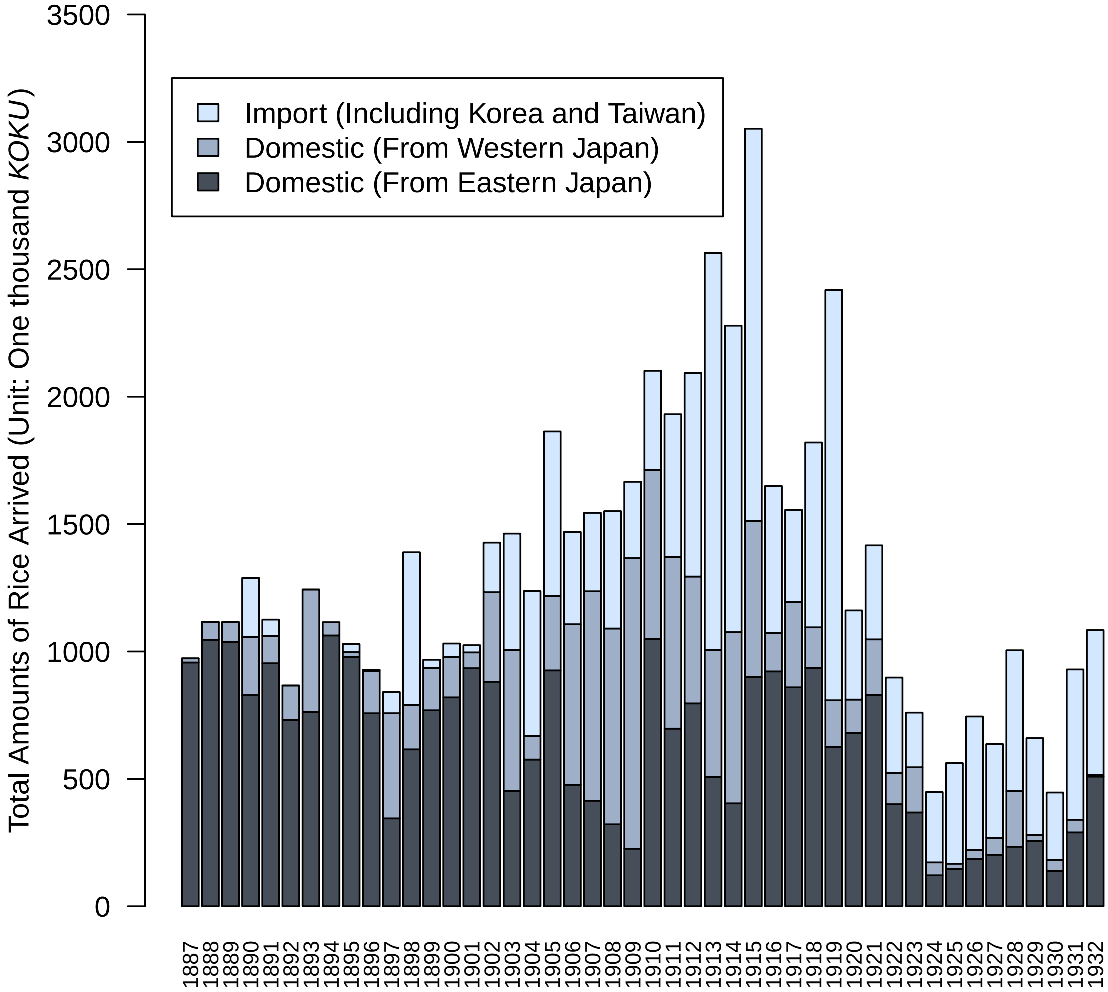
(3) R version 3.1.0 was used to compute the estimates.

Figure 4.2: Time-Varying Estimates of β: The Case of Osaka Rice Market



Note: As for Figure 4.1.

Figure 4.3: Total Amounts of Rice Arrived in the Tokyo-Fukagawa Spot Rice Market



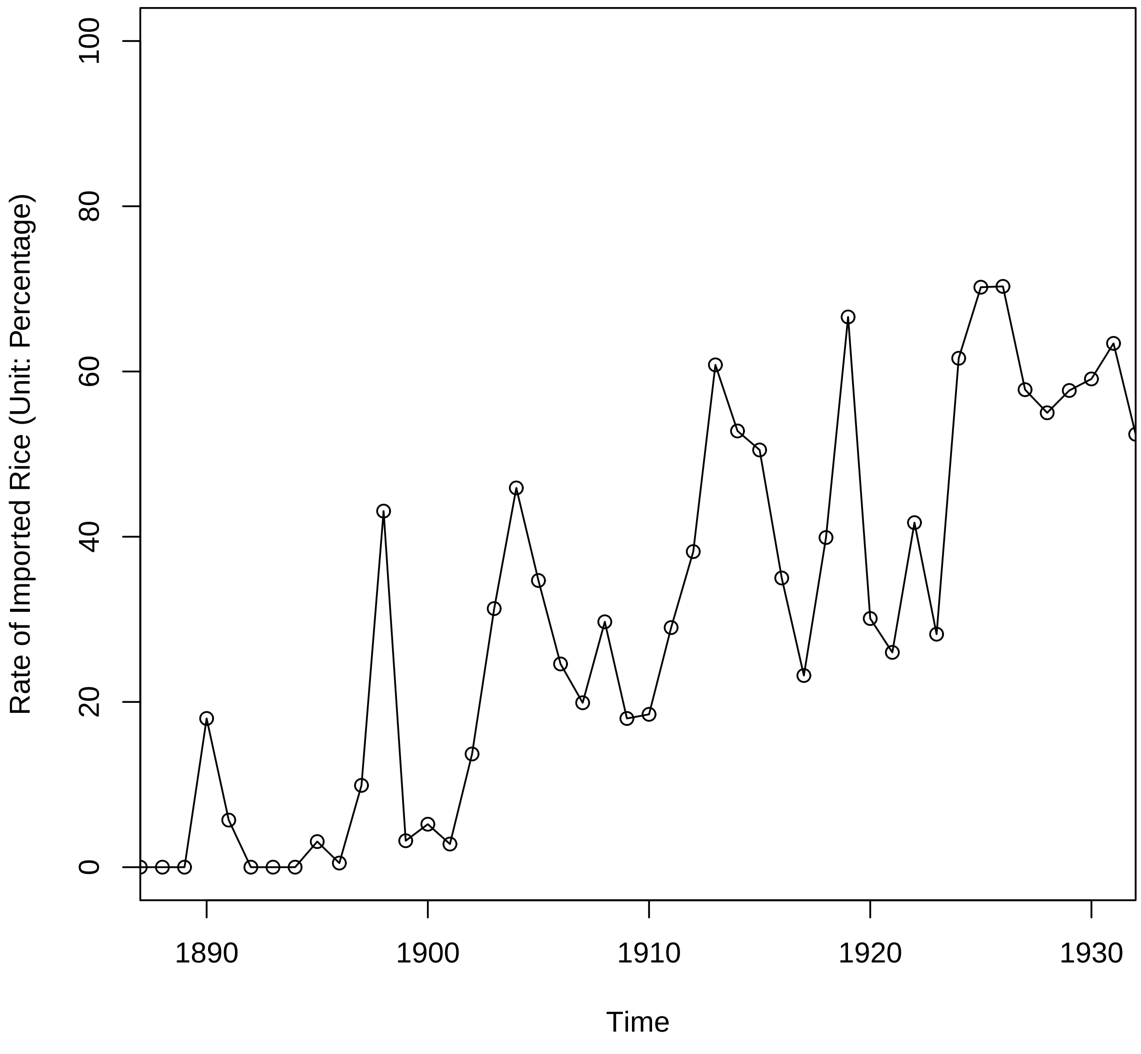
Data Sources:

(1) Tokyo Chamber of Commerce (1927, pp.6–7).

(2) Tokyo Chamber of Commerce and Industory (1934, pp.132–133).

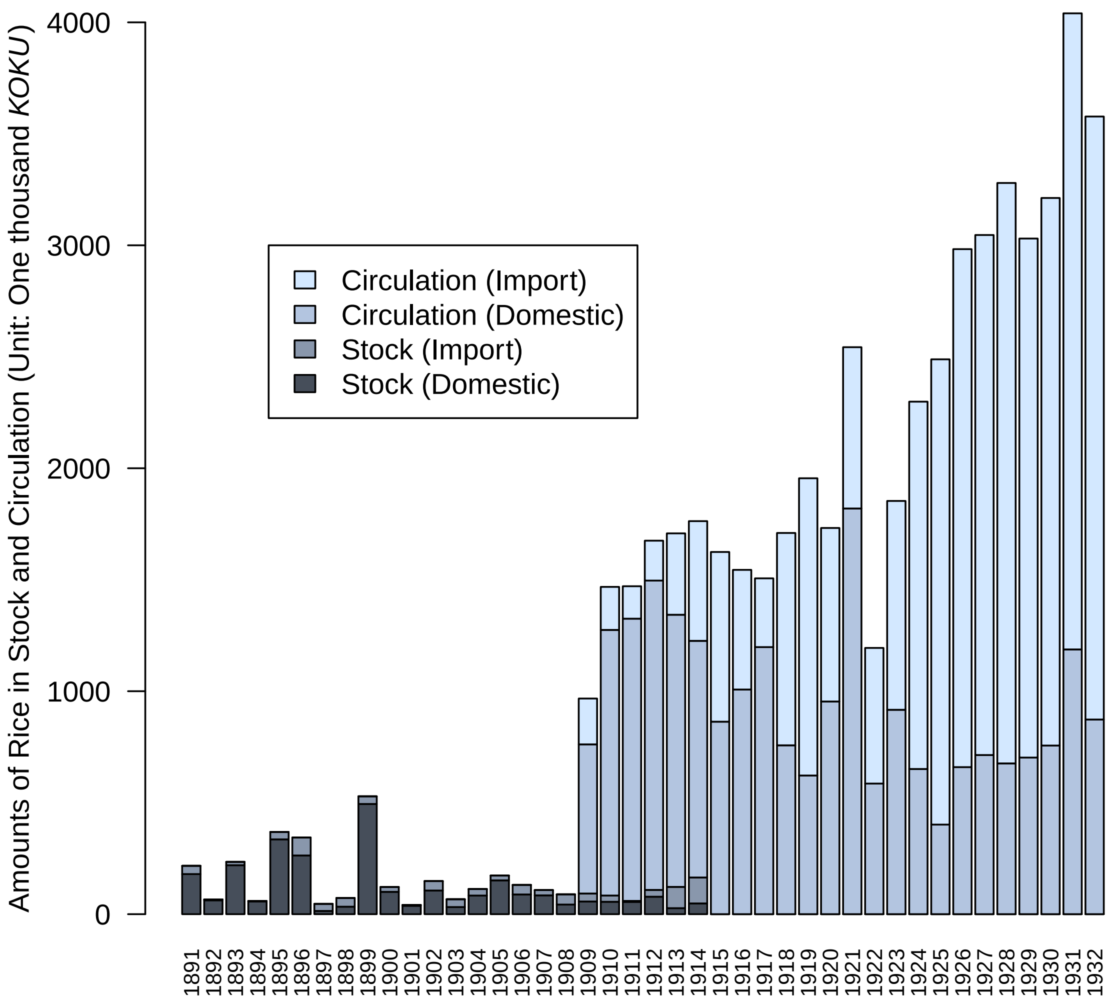
(3) Sasaki (1937, pp.268–270).

Figure 4.4: The Rate of Imported Rice in the Tokyo-Fukagawa Rice Spot Market



Note: As for Figure 4.3.

Figure 4.5: Total Amounts of Rice in Stock and Circulation in Osaka



Data Sources:

(1) Osaka City Government (1913, pp.253–254).

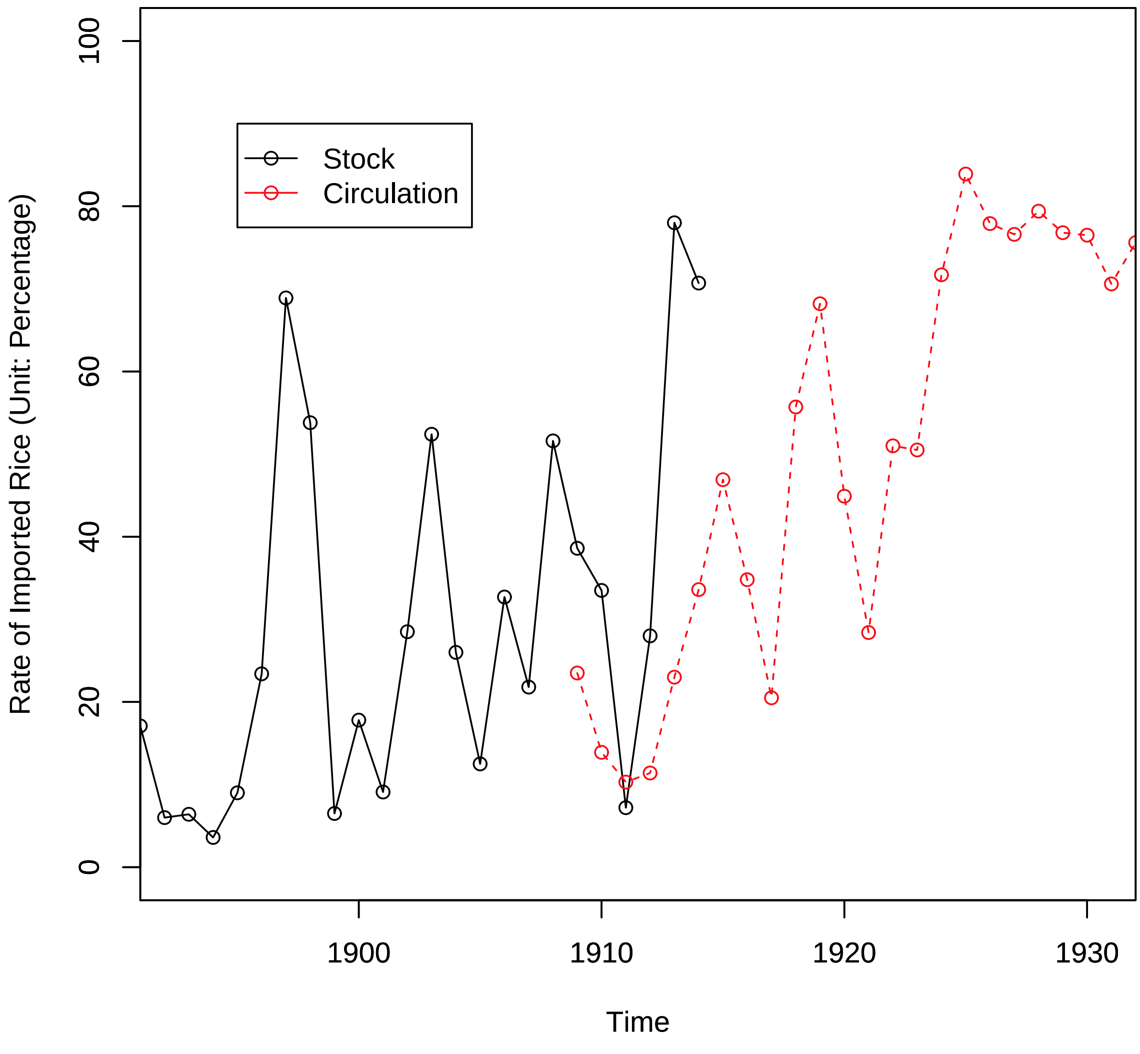
(2) Osaka-Dojima Rice Exchange (1915, pp.100–102).

(3) Osaka City Government (1920, ch.7, pp.54–55).

(3) Osaka City Government (1926, ch.7, pp.54–55).

(3) Osaka City Government (1933, ch.5, pp.34–35).

Figure 4.6: The Rate of Imported Rice to All Physical Rice Arrived in Osaka



Note: As for Figure 4.5.

Chapter 5

Conclusion

The dissertation aimed at examining the contemporary financial system from the view of market efficiency by using non-Bayesian time-varying models: TV-AR, TV-VAR and a linear regression with time-varying parameters. As the author discussed in Chapter 2, such an approach was designed to deal with most financial systems, which have confronted with unceasing structural changes. The approach successfully attained the goal if the author dare say in short. The rest of this chapter summarize the arguments of Chapter 1 and the empirical results of Chapters 2, 3 and 4 proving the validity of the approach. Then it concludes.

Chapter 1 asserts the significance of paying attention to unceasing structural changes in financial markets in the real world. In other words, a test of parameter constancy is inevitable and a model with time-varying parameters is preferable when a parametric model is employed to analyze financial markets, say, stock markets or commodity futures markets. The chapter surveyed the literature about market efficiency since Fama (1970) is published. The literature is regarded as a long history of debate over whether financial markets are efficient or not. Chapter 1 also discussed the reason why the literature is never-ending debate and argued that analytical tools or statistical method employed by literature failed to detect if financial markets are efficient or not based on available data. Under the supposition that there is no structural change, any time-invariant parametric models could never deal with the most financial systems that confront with unceasing structural changes such as financial crises like the recent Lehman bankruptcy. Methodologically saying, the author argued that such a model necessary fails since it assumes some fundamental uniform structure. Since most real financial systems confront with unceasing structural changes in financial markets such as stock markets, commodity futures markets, etc., required is such a new method that allows us to deal with the markets under unceasing structural changes. The first chapter asserted that the non-Bayesian time-varying approach in this dissertation provides a powerful tool to analyze the financial markets.

Chapters 2, 3, and 4 analyzed some financial systems, stock markets and commodity futures markets, from the view of time-varying market efficiency by employing a non-Bayesian time-varying model as well as by testing parameter instability for conventional time series models. They elucidated that market efficiency of such financial markets vary over time with high probability. The findings not only correspond well to historical events in the financial markets but also reconciles the never-ending debates over market efficiency of financial systems. In particular, Chapter 2 addressed the long-run data of U.S. stock market to examine how efficiency of the market evolved over time; a non-Bayesian time-varying model is developed by introducing the concept of the degree of market efficiency that varies over time. With the new methodology and a measure of the degree of market efficiency, the author examined whether the U.S. stock market evolves over time. Employing a time-varying autoregressive (AR) model and measuring the degree, he argued that (i) the U.S. stock market had evolved over time and the degree of market efficiency has cyclical fluctuations with a considerably long periodicity, from 30 to 40 years; and (ii) the U.S. stock market had been efficient with the exception of four times in my sample period: during the long-recession of 1873-1879; the recession of 1902-1904; the New Deal era; and the recession of 1957-1958 and soon after it. It is a novel finding that the market efficiency in U.S. stock market has such long periodicity is absolutely new although there are no more persuasive reason of the periodicity than the historical events well corresponded to the periodicity.

As for the Chapter 3, developed was a non-Bayesian methodology to analyze the time-varying structure of international linkages and market efficiency in G7 countries. The author considered a non-Bayesian time-varying vector autoregressive (TV-VAR) model, and applied it to estimate the joint degree of market efficiency for G7 countries in the sense of Fama (1970, 1991). The empirical results in the chapter provides a new perspective that the international linkages and market efficiency vary over time and that their behaviors correspond well to historical events of the international financial system in the same way as Chapter 2 did. Chapter 3 concludes that the contemporary international stock markets have dynamic linkages in the view of market efficiency and that exogenous shocks might perturb the efficiency.

Chapter 4 examined how the Tokyo and Osaka rice futures markets in prewar Japan were evolving in view of market efficiency. Applying a non-Bayesian time-varying model approach to analyze the famous equation for the futures premium, the author finds that the market efficiency of the two major rice futures markets varied with time. In contrast to Chapter 3, he employed the famous equation regarding the futures premium and spot returns of rice futures. Then he examined the time-varying structure of its key parameter that could signify market efficiency of the two rice futures market in Tokyo and Osaka. As the empirical result, the author demonstrated that the two markets have a little different time-varying patterns of the key parameter. It reflected that the two markets experienced different time-varying market structures respectively. Such time-varying structure of the rice futures markets in prewar Japan corresponds well to historical changes in the Japanese colonial policy and domestic development of railroad system and port facilities. This finding provides a new perspective about commodity futures.

Bibliography

Anderson, B. D. O. and Moore, J. B. (1979), *Optimal Filtering*, Prentice-Hall, Englewood Cliffs, NJ.

Andreou, E., Pittis, N., and Spanos, A. (2001), "On Modelling Speculative Prices: The Empirical Literature," *Journal of Economic Surveys*, 15, 187–220.

Andrews, D. W. K. (1991), "Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation," *Econometrica*, 59, 817–858.

— (1993a), "Exactly Median-Unbiased Estimation of First Order Autoregressive/Unit Root Models," *Econometrica*, 61, 139–165.

— (1993b), "Tests for Parameter Instability and Structural Change With Unknown Change Point," *Econometrica*, 61, 821–856.

Ariel, R. A. (1987), "A Monthly Effect in Stock Returns," *Journal of Financial Economics*, 18, 161–174.

— (1990), "High Stock Returns Before Holidays: Existence and Evidence on Possible Causes," *Journal of Finance*, 45, 1611–1626.

Baillie, R. T. (1996), "Long Memory Processes and Fractional Integration in Econometrics," *Journal of Econometrics*, 73, 5–59.

Banz, R. W. (1981), "The Relationship between Return and Market Value of Common Stocks," *Journal of Financial Economics*, 9, 3–18.

Baxter, M. and King, R. G. (1999), "Measuring Business Cycles: Approximate Band-Pass Filters for Economic Time Series," *Review of Economics and Statistics*, 81, 575–593.

Beveridge, S. and Nelson, C. R. (1981), "A New Approach to Decomposition of Economic Time Series Into Permanent and Transitory Components with Particular Attention to Measurement of the 'Business Cycle'," *Journal of Monetary Economics*, 7, 151–174.

Blume, M. E. (1970), "Portfolio Theory: A Step Toward Its Practical Application," *Journal of Business*, 43, 152–173.

Brenner, R. J. and Kroner, K. F. (1995), "Arbitrage, Cointegration, and Testing the Unbiasedness Hypothesis in Financial Markets," *Journal of Financial and Quantitative Analysis*, 30, 23–42.

Brockwell, P. J. and Davis, R. A. (1991), *Time Series: Theory and Methods*, Springer, 2nd ed.

Campbell, J. Y., Lo, A. W., and Mackinlay, A. C. (1997), *Econometrics of Financial Markets*, Princeton University Press.

Chan, K. C. and Chen, N. (1991), "Structural and Return Characteristics of Small and Large Firms," *Journal of Finance*, 46, 1467–1484.

Chan, K. C., Gup, B. E., and Pan, M. (1997), "International Stock Market Efficiency and Integration: A Study of Eighteen Nations," *Journal of Business Finance & Accounting*, 24, 803–813.

Charles, A. and Darné, O. (2009), "Variance-Ratio Tests of Random Walk: An Overview," *Journal of Economic Surveys*, 23, 503–527.

Cogley, T. and Sargent, T. J. (2001), "Evolving Post-World War II U.S. Inflation Dynamics," *NBER Macroeconomics Annual*, 16, 331–373.

— (2005), "Drifts and Volatilities: Monetary Policies and Outcomes in the Post WWII US," *Review of Economic Dynamics*, 8, 262–302.

Conrad, J. and Kaul, G. (1988), "Time-Variation in Expected Returns," *Journal of Business*, 61, 409–425.

Elliott, G., Rothenberg, T. J., and Stock, J. H. (1996), "Efficient Tests for an Autoregressive Unit Root," *Econometrica*, 64, 813–836.

Fama, E. F. (1965), "The Behavior of Stock-Market Prices," *Journal of Business*, 38, 34–105.

— (1970), "Efficient Capital Markets: A Review of Theory and Empirical Work," *Journal of Finance*, 25, 383–417.

— (1991), "Efficient Capital Markets: II," *Journal of Finance*, 46, 1575–1617.

— (1998), "Market Eifficiency, Long-Term Returns, and Behavioral Finance," *Journal of Financial Economics*, 49, 283–306.

Fama, E. F. and Blume, M. E. (1966), "Filter Rules and Stock-Market Trading," *Journal of Business*, 39, 226–241.

Fama, E. F., Fisher, L., Jensen, M. C., and Roll, R. (1969), "The Adjustment of Stock Prices to New Information," *International Economic Review*, 10, 1–21.

Fama, E. F. and French, K. R. (1987), "Commodity Futures Prices: Some Evidence on Forecast Power, Premiums, and the Theory of Storage," *Journal of Business*, 60, 55–73.

— (1992), "The Cross Section of Expected Stock Returns," *Journal of Finance*, 47, 427–465.

Government–General of Taiwan (1945), "Taiwan Touchi Gaiyou [Administrative Report of Taiwan]," Taipei, Taiwan.

Hamilton, J. D. (1987), "Monetary Factors in the Great Depression," *Journal of Monetary Economics*, 19, 145–169.

— (1994), *Time Series Analysis*, Princeton University Press.

Hamori, S., Hamori, N., and Anderson, D. A. (2001), "An Empirical Analysis of the Efficiency of the Osaka Rice Market during Japan's Tokugawa Era," *Journal of Futures Markets*, 21, 861–874.

Hansen, B. E. (1990), "Lagrange Multiplier Tests for Parameter Instability in Non-linear Models," University of Rochester.

— (1992), "Testing for Parameter Instability in Linear Models," *Journal of Policy Modeling*, 14, 517–533.

Harris, L. (1986), "A Transaction Data Study of Weekly and Intradaily Patterns in Stock Returns," *Journal of Financial Economics*, 16, 99–117.

Hommes, C. and Wagener, F. (2009), "Complex Evolutionary Systems in Behavioral Finance," in *Handbook of Financial Markets: Dynamics and Evolution*, eds. Hens, T. and Schenk-Hoppé, K. R., North Holland, pp. 217–276.

Ishii, K. (1986), "Kokunai Shijo no Keisei to Tenkai [The Rise and Development of Domestic Market in Japan]," in *Kindai Nihon no Shohin Ryutsu* [The Circulation of Commodities in Modern Japan], eds. Yamaguchi, K. and Ishii, K., University of Tokyo Press, pp. 26–35.

Ito, M. (2007), "A New Method for Estimating Economic Models with General Time-Varying Structures," Keio Economic Society Discussion Paper Series, KESDP 07-8.

Ito, M., Maeda, K., and Noda, A. (2014a), "Did Rice Markets Integrate in Prewar Japan?" Mimeo.

— (2014b), "Dynamic Linkages between Tokyo and Osaka Rice Futures Markets in Prewar Japan," [Preprint: arXiv:1404.1164], Available at http://arxiv.org/pdf/1404.1164.pdf.

— (2014c), "Futures Premium and Efficiency of the Rice Futures Markets in Prewar Japan," [Preprint: arXiv:1404.5381], Available at http://arxiv.org/pdf/1404.5381.pdf.

Ito, M., Noda, A., and Wada, T. (2012), "The Evolution of Market Efficiency and Its Periodicity: A Non-Bayesian Time-Varying Model Approach," [Preprint: arXiv:1202.0100], Available at http://arxiv.org/pdf/1202.0100.pdf.

— (2014d), "International Stock Market Efficiency: A Non-Bayesian Time-Varying Model Approach," *Applied Economics*, 46, 2744–2754.

Ito, M. and Sugiyama, S. (2009), "Measuring the Degree of Time Varying Market Inefficiency," *Economics Letters*, 103, 62–64.

Ito, T. (1993), "Efficiency of the Rice Futures Market in Osaka in the 18th Century (in Japanese)," *Economic Review*, 44, 339–350.

Jensen, M. C. (1968), "The Performance of Mutual Funds in the Period 1945-1964," *Journal of Finance*, 23, 389–416.

Jeon, B. N. and Von Furstenberg, G. M. (1990), "Growing International Co-Movement in Stock Price Indexes," *Quarterly Review of Economics and Business*, 30, 15–30.

Kakizaka, M. (2012), "A Study on the Efficiency of the Dojima Rice Market in the 19th Century (in Japanese)," *JCER Economic Journal*, 66, 72–87.

Kamibayashi, M. (1935), *Shohin Torihikijo Ron* [Commodity Exchanges], Iwanami Shoten.

Kasa, K. (1992), "Common Stochastic Trends in International Stock Markets," *Journal of Monetary Economics*, 29, 95–124.

Kato, K. (2001), *Kinsei Koki Keizai Hatten no Kozo* [Structure of Economic Development in Late Modern Japan], Seibundo.

Keim, D. B. (1983), "Size-Related Anomalies and Stock Return Seasonality: Further Empirical Evidence," *Journal of Financial Economics*, 12, 13–32.

— (1989), "Trading Patterns, bid-ask Spreads, and Estimated Security Returns: The Case of Common Stocks at Calendar Turning Points," *Journal of Financial Economics*, 25, 75–97.

Kim, J. H. and Shamsuddin, A. (2008), "Are Asian Stock Markets Efficient? Evidence from new Multiple Variance Ratio Tests," *Journal of Empirical Finance*, 15, 518–532.

Kim, J. H., Shamsuddin, A., and Lim, K. P. (2011), "Stock Return Predictability and the Adaptive Markets Hypothesis: Evidence from Century-Long U.S. Data," *Journal of Empirical Finance*, 18, 868–879.

Koiwa, N. (2003), *Kindai Nihon no Beikoku Shijo* [The Rice Market in Modern Japan], Association of Agriculture & Forestry Statistics.

Lean, H. H. and Smyth, R. (2007), "Do Asian Stock Markets Follow a Random Walk? Evidence from LM Unit Root Tests with One and Two Structural Breaks," *Review of Pacic Basin Financial Markets and Policies*, 10, 15–31.

Lim, K. P. and Brooks, R. (2011), "The Evolution of Stock Market Efficiency Over Time: a Survey of the Empirical Literature," *Journal of Economic Surveys*, 25, 69–108.

Lim, K. P., Luo, W., and Kim, J. H. (2013), "Are US Stock Index Returns Predictable? Evidence from Automatic Autocorrelation-Based Tests," *Applied Economics*, 45, 953–962.

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

Lo, A. W. (2004), "The Adaptive Markets Hypothesis: Market Efficiency from an Evolutionary Perspective," *Journal of Portfolio Management*, 30, 15–29.

— (2005), "Reconciling Efficient Markets with Behavioral Finance: The Adaptive Markets Hypothesis," *Journal of Investment Consulting*, 7, 21–44.

Lo, A. W. and MacKinlay, A. C. (1988), "Stock Market Prices do not Follow Random Walks: Evidence from a Simple Specification Test," *Review of Financial Studies*, 1, 44–66.

Lütkepohl, H. (2005), *New Introduction to Multiple Time Series Analysis*, Springer.

MacDonald, R. and Taylor, M. P. (1989), "Foreign Exchange Market Efficiency and Cointegration: Some Evidence from the Recent Float," *Economics Letters*, 29, 63–68.

Maddala, G. S. and Kim, I. (1998), *Unit Roots, Cointegration, and Structural Change*, Cambridge University Press.

Malkiel, B. G. (2003), "The Efficient Market Hypothesis and its Critics," *Journal of Economic Perspectives*, 17, 59–82.

— (2004), *A Random Walk Down Wall Street: The Time-Tested Strategy for Successful Investing*, W. W. Norton & Company, Inc.

Malkiel, B. G., Mullainathan, S., and Stangle, B. E. (2005), "Market Efficiency versus Behavioral Finance," *Journal of Applied Corporate Finance*, 17, 124–136.

Ministry of Commerce and Industry (1931), "Statistics of Wholesale Prices," Tokyo, Japan.

Miyamoto, M. (1988), *Kinsei Nihon no Shijo Keizai* [The Market Economy of Pre-Modern Japan], Yuhikaku.

Miyamoto, M., Uekawa, Y., Uemura, M., Shiba, T., Hiroyama, K., Chakepaichayon, V., Mori, N., and Sakata, Y. (1979), "Osaka Wholesale Prices (1) 1881–1912: Rice, Sake, Coal and Cotton," *Osaka Economic Papers*, 29, 140–190.

Mochida, K. (1970), *Beikoku Shijo no Tenkai Katei* [The Development of Rice Market], University of Tokyo Press.

Morley, J. C., Nelson, C. R., and Zivot, E. (2003), "Why are the Beveridge-Nelson and Unobserved-Components Decompositions of GDP so Different?" *Review of Economics and Statistics*, 85, 235–243.

Nakamura, N. (2003), "Hokubu-Kyushu ni Okeru Kindaiteki Kotsu Kikan to Shohin Ryutsu [Transportation Facilities and Distribution of Commodities in Northern Kyushu]," in

Shohin Ryutsu no Kindaishi [A History of Distribution in Modern Japan], eds. Nakanishi, S. and Nakamura, N., Nihon Keizai Hyouron Sha, pp. 49–84.

Nakanishi, S. (2002), "Kindai no Shohin Shijo [Commodity Market in Modern Japan]," in *Ryutsu Keizaishi* [Economic History of Distribution], eds. Sakurai, E. and Nakanishi, S., Yamakawa Shuppansha, pp. 275–328.

Nakazawa, B. (1933), *Nihon Beika Hendo Shi* [History of Rice Price Fluctuations in Japan], Kashiwa Shobo.

Narayan, P. K. and Smyth, R. (2006), "Random Walk versus Multiple Trend Breaks in Stock Prices: Evidence from 15 European Markets," *Applied Financial Economics Letters*, 2, 1–7.

— (2007), "Mean Reversion versus Random Walk in G7 Stock Prices Evidence from Multiple Trend Break Unit Root Tests," *Journal of International Financial Markets, Institutions and Money*, 17, 152–166.

Newey, W. K. and West, K. D. (1987), "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, 55, 703–708.

— (1994), "Automatic lag Selection in Covariance Matrix Estimation," *Review of Economic Studies*, 61, 631–653.

Ng, S. and Perron, P. (2001), "Lag Length Selection and the Construction of Unit Root Tests with Good Size and Power," *Econometrica*, 69, 1519–1554.

Nyblom, J. (1989), "Testing for the Constancy of Parameters Over Time," *Journal of the American Statistical Association*, 84, 223–230.

Okada, S. (1911), *Chosen Yushutsumai Jijyo* [Export Rice from Korea], Doubunkan.

Omameuda, M. (1993), *Kindai Nihon no Shokuryo Seisaku* [Food Policy in Modern Japan], Minerva Shobo.

— (2000), "Tokyo Shijo wo Meguru Jimawari Mai to Ongoku Mai [Regional Rice and Extra-Regional Rice in Tokyo Rice Market]," in *Shohin Ryutsu to Tokyo Shijo* [Commodity Distribution and the Tokyo Market], eds. Oikawa, Y. and Omameuda, M., Nihon Keizai Hyoron Sha, pp. 193–221.

— (2004), "Beikoku Ryutsu no Saihen to Shonin [Restructuring of the Rice Distribution and Merchants]," in *Meiji Zenki no Nihon Keizai* [Japanese Economy in Early Meiji-Period], ed. Takamura, N., Nihon Keizai Hyoron Sha, pp. 289–322.

— (2007), *Okome to Shoku no Kindai Shi* [History of Rice and Food in Modern Japan], Yoshikawa Kobunsha.

Osaka City Government (1913), "The 11th Osaka City Statistics," Osaka, Japan.

— (1920), "The 17th Osaka City Statistics," Osaka, Japan.

— (1926), "The 23th Osaka City Statistics," Osaka, Japan.

— (1931), "The 29th Osaka City Statistics," Osaka, Japan.

— (1933), "The 31th Osaka City Statistics," Osaka, Japan.

Osaka-Dojima Rice Exchange (1915), "Kabushiki Gaisha Osaka Dojima Beikoku Torihikijo Enkaku [History of Osaka-Dojima Rice Exchange]," Osaka, Japan.

Ota, K. (1938), *Meiji Taisho Showa Beika Seisaku Shi* [Rice Price Policy in the Meiji, Taisho and Showa Eras], Maruyamasha Shoten.

Paige, C. C. and Saunders, M. A. (1977), "Least Squares Estimation of Discrete Linear Dynamic Systems Using Orthogonal Transformations," *SIAM Journal on Numerical Analysis*, 14, 180–193.

Park, C. H. and Irwin, S. H. (2007), "What Do We Know About the Profitability of Technical Analysis," *Journal of Economic Surveys*, 21, 786–826.

Perron, P. (1989), "The Great Crash, the Oil Price Shock, and the Unit Root Hypothesis," *Econometrica*, 57, 1361–1401.

Perron, P. and Wada, T. (2009), "Let's Take a Break: Trends and Cycles in US Real GDP," *Journal of Monetary Economics*, 56, 749–765.

Primiceri, G. E. (2005), "Time Varying Structural Vector Autoregressions and Monetary Policy," *Review of Economic Studies*, 72, 821–852.

Rahman, A. and Saadi, S. (2008), "Random Walk and Breaking Trend in Financial Series: An Econometric Critique of Unit Root Tests," *Review of Financial Economics*, 17, 204–212.

Rothenberg, T. J. (1971), "Identification in Parametric Models," *Econometrica*, 39, 577–591.

Samuelson, P. A. (1965), "Proof that Properly Anticipated Prices Fluctuate Randomly," *Industrial Management Review*, 6, 41–49.

Sargent, T. (1987), *Macroeconomic Theory*, Academic Press, 2nd ed.

Sasaki, N. (1937), "Tokyo Kaimai Donya Kumiai Fukagawa Shomai Shijo 50-nen Shi [Fifty Year's History of Tokyo Rice Wholesalers Association and Fukagawa Rice Spot Market]," Tokyo, Japan.

Schwarz, G. (1978), "Estimating the Dimension of a Model," *Annals of Statistics*, 6, 461–464.

Schwert, G. W. (2003), "Anomalies and Market Efficiency," in *Handbook of the Economics of Finance*, eds. Constantinides, G. M., Harris, M., and Stulz, R. M., North-Holland, chap. 15, pp. 937–972.

Shiller, R. J. (2005), *Irrational Exuberance: Second Edition*, Princeton University Press.

Shizume, M. (2011), "Formulation of Expectations in Price Fluctuation in Japan during the Interwar Period: An Analysis Using Commodity Futures Price Data (in Japanese)," *Socio-Economic History*, 77, 25–47.

Suzuki, S. (1940), *Dojima Kome Shijo Shi* [History of the Dojima Rice Market], Nippon Hyoron Sha.

Takatsuki, Y. (2008), "The Center of Rice Trade in Tokugawa Japan: Efficiency of the Dojima Rice Market (in Japanese)," *Socio-Economic History*, 74, 3–22.

— (2012), *Kinsei Komeshijo no Keisei to Tenkai* [The Rise and Development of Rice Exchange Market in Tokugawa Japan], University of Nagoya Press.

Taketoshi, K. (1999), "Meiji kara Showa Shoki ni Okeru Kome Sakimono Kakaku ni Kansuru Keiryo Bunseki [An Econometric Analysis of Rice Futures Prices in Prewar Japan]," *Sakimono Torihiki Kenkyu* [Japan Futures Studies], 4, 125–145.

Tohata, S. and Okawa, K. (1939), "Chosen Beikoku Keizairon [Rice Economy of Korea]," in *Beikoku Keizai no Kenkyu* [Study of Rice Economy], ed. Japan Society for the Promotion of Science, Yuhikaku, vol. 1, pp. 297–446.

Tokyo Chamber of Commerce (1927), “The Annual Statistical Report of the Tokyo Chamber of Commerce 1924,” Tokyo, Japan.

Tokyo Chamber of Commerce and Industory (1934), “The Annual Statistical Report of the Tokyo Chamber of Commerce and Industry 1932,” Tokyo, Japan.

Tsay, R. S. (2010), *Analysis of Financial Time Series*, John Wiley & Sons, 3rd ed.

Tsutsui, Y. and Hirayama, K. (2004), “Appropriate Lag Specification for Daily Responses of International Stock Markets,” *Applied Financial Economics*, 14, 1017–1025.

Wakita, S. (2001), “Efficiency of the Dojima Rice Futures Market in Tokugawa-Period Japan,” *Journal of Banking and Finance*, 25, 535–554.

Wright, J. H. (2000), “Alternative Variance-Ratio Tests Using Ranks and Signs,” *Journal of Business & Economic Statistics*, 18, 1–9.

Yen, G. and Lee, C. (2008), “Efficient Market Hypothesis (EMH): Past, Present and Future,” *Review of Pacific Basin Financial Markets and Policies*, 11, 305–329.