Masteroppgave

Master of Science in Business

Finansiering og investering

Julie Marie Aksdal and Ulrikke Grønberg

Trading strategies based on the lead-lag relationship between the spot and the futures prices for the Nikkei 225 Stock Average Index

Trondheim, mai 2015

Høgskolen i Sør-Trøndelag
Handelshøyskolen i Trondheim

Julie Marie Aksdal and Ulrikke Grønberg

Trading strategies based on the lead-lag relationship between the spot and the futures prices for the Nikkei 225 Stock Average Index

Masteroppgave, Master of Science in Business Administration
Trondheim, mai 2015

Spesialiseringsretning:	Finance and investments
Veileder:	Khine Kyaw

Høgskolen i Sør-Trøndelag
Handelshøyskolen i Trondheim

Preface

This master thesis is written in the spring semester of 2015 as a completion of our Master of Science in Business Administration at Trondheim Business School (Handelshøyskolen i Trondheim, HiST).

In the thesis we take a closer look at the Japanese financial markets, and the analysis aims at distinguishing the effects that the worldwide financial crisis beginning in 2008 have had on this market in terms of market efficiency. The reason for choosing this topic is both the major impact that this crisis have had on global financial markets, and also that the theory on market efficiency forms a central basis for other important theories on the behaviour of prices of financial assets.

We would like to express our gratitude to our supervisor, Associate Professor Khine Kyaw, who has provided us with valuable advice and feedback in all stages of the process.

The contents and opinions given in this thesis are the sole responsibility of the authors. The institution and the supervisor are not responsible for the theories and methods used, nor the results and conclusions drawn, through the approval of this thesis.

Trondheim, May 21st 2015

Julie Marie Aksdal   Ulrikke Grønberg

Abstract

The main objective for this master thesis was to investigate the relationship between the spot and futures market for the Nikkei 225 Stock Average Index, in order to give an indication about the efficiency of the Japanese financial markets. Furthermore, the sample period from July 2004 to July 2014 provided an opportunity to compare the nature of the relationship before and after the financial crisis of 2008. The focus of the study was the following research question:

Is it possible to earn an abnormal return in the Japanese market, as represented by the Nikkei 225 Stock Average Index, by trading in the spot and futures market based on an error correction model?

The efficient market hypothesis and arguments concerning arbitrage formed the theoretical foundation of the study. The relationship between the spot and the futures market for the Nikkei 225 Stock Average Index was investigated using unit root tests, the Engle and Granger two-step co-integration method, error correction models and ARIMA frameworks. The study showed that the two prices were co-integrated and shared a long-run relationship. A Granger causality test suggested that the relationship was bi-directional, with a strong feedback effect. Furthermore, the study uncovered a shift in the relationship from the period before September 15th 2008 to the period after. The results implied that the flow of information had slowed down, and that corrections back to the equilibrium state were slower, and therefore a deterioration of market efficiency. The models were tested further using forecasting techniques on out-of-sample periods. The best models from the post financial crisis period were applied in a test of different trading strategies. The selection of models was based on RMSE (root mean squared error), MAE (mean average error) and percentage of correct prediction of direction. None of the active strategies provided abnormal profits after the deduction of transaction costs. This was the case for both trade of futures contracts and trade of the spot price. Therefore, the market seemed to be efficient in terms of the arbitrage argument, even after the financial crisis of 2008.

Sammendrag

Hovedformålet med denne masteravhandlingen var å undersøke forholdet mellom spot – og futuresmarkedet for Nikkei 225 Stock Average-indeksen for å kunne si noe om graden av effisiens i de japanske finansmarkedene. Utvalgsperioden, som varte fra juli 2004 til juli 2014, gjorde det mulig å sammenlikne relasjonens natur før og etter finanskrisen i 2008. Studien fokuserte på følgende forskningsspørsmål:

Er det mulig å oppnå en meravkastning i det japanske markedet, representert av Nikkei 225 Stock Average-indeksen, ved å handle i spot – og futuresmarkedet basert på en feilkorreksjonsmodell?

Hypotesen om markedseffisiens og arbitrasjeargumenter dannet det teoretiske grunnlaget for studien. Relasjonen mellom spot – og futuresmarkedet for Nikkei 225 Stock Average-indeksen ble undersøkt ved å bruke enhetsrot-tester, Engle og Granger to-stegsmetode for kointegrasjon, feilkorreksjonsmodeller og ARIMA-rammeverk. Studien viste at prisene var kointegrerte og hadde en langsiktig forbindelse. En Granger-kausalitetstest antydet at denne relasjonen var tosidig med en sterk feedback-effekt. Studien avdekket også en endring i relasjonen fra perioden før 15. september 2008 til perioden etter. Resultatene indikerte at strømmen av informasjon gikk langsommere og at korrigeringer tilbake til likevektstilstanden gikk saktere, og dermed at markedet var blitt mindre effisient etter finanskrisen. Modellene ble testet videre ved bruk av prediksjonsteknikker i perioder utenfor utvalgsperioden. De beste modellene fra perioden etter finanskrisen ble brukt til å teste ulike handelsstrategier. Valget av modeller baserte seg på mål som RMSE (root mean squared error/kvadratroten av gjennomsnittlig kvadrert feilledd), MAE (mean average error/gjennomsnittlig feilledd) og prosentandel av riktig predikert retning. Ingen av de aktive strategiene ga meravkastning etter at transaksjonskostnader var trukket fra. Dette gjaldt både handel basert på futureskontrakter og handel basert på spot-prisen. Markedet syntes derfor å være effisient i henhold til arbitrasjeargumentet også etter finanskrisen i 2008.

Table of contents

List of tables

List of figures

1 Introduction

1.1 Context

The main focus of this thesis is described by the following title:

Trading strategies based on the lead-lag relationship between the spot and futures prices for the Nikkei 225 Stock Average Index.

Japan is a large and important economy with great implications for the overall world economy. Therefore, the choice to investigate a potential lead-lag relationship in the Japanese financial markets was not a random one. Market analysts and great economies such as the US and large European countries have taken interest in how Japan has dealt with financial instability in the past years (Irwin, 2013).

For example, in 1945, after Japan was defeated by China in the Pacific War, the country managed to achieve rapid growth and industrialization due to global economic conditions, which allowed Japan to export high-quality manufactured goods (Ohno, 2006). From the mid 1950s and through the 1970s the Japanese economy grew to become the second largest in the world (Central Intelligence Agency, 2014). However, following the Japanese asset price bubble in the late 1980s, the country entered into what has become known as “the lost decade”. During the 1990s the Japanese economy practically stopped growing and deflation became a reality.

Several reforms were introduced by the Koizumi government, and in spite of more difficulties related to the IT recession in 2001, the economy seemed to pick up speed by 2005 (Ohno, 2006). However, in the second half of 2008, and especially in the fourth quarter of 2008, the Japanese economy was hit hard by the financial crisis (Fukao and Yuan, 2009). The GDP contraction was almost twice that of the US, and the economy experienced a steep fall in external demand and an appreciation of the Yen.

After the 2012 election, Shinzō Abe, who then took office as the Prime Minister of Japan, pledged to do “whatever it takes” to turn a deflating economy into inflation (Irwin, 2013). The new era of "the Abenomics" involved taking on Keynesian remedies in combination with structural reforms to end the more than two decades long period of Japanese recession. Abe seems to be succeeding (Trading Economics, 2015). Despite still suffering from the aftermaths of "the lost decade" and the financial crisis, with huge government debt (Central Intelligence Agency, 2014) and massive unemployment (Trading Economics, 2015), Japan is the third largest economy in the world in 2015 (Bergmann, 2015). The country now has a higher GDP growth rate than the US (Trading Economics, 2015).

The use of the Nikkei 225 Stock Average Index as the underlying asset for this analysis makes it possible to monitor the pulse of the Japanese economy (Riley, 2015). A 10-year sample period with a time span from July 2004 to June 2014 is analysed in this thesis. Thus, the sample contains data from before, during and after the latest financial crisis triggered by the bankruptcy of the American Lehman Brothers Holdings Inc. in 2008. The reasoning behind this is to look at how the Japanese financial markets were affected in terms of maintaining market efficiency after the latest financial crisis.

1.2 Research question

Although there have been many studies on market efficiency and error correction models related to the early years of the futures market for the Nikkei 225 Stock Average Index and the financial crisis of the late 1980s, there seem to be rather few that concern the later years and longer time spans. This, despite the fact that the Japanese economy has undergone many changes related to "the lost decade" and changes related to the worldwide economic situation, such as increasing international trade and cross-border challenges and crises. Following Tse (1995) and Brooks et al. (2001), co-integration techniques are used to develop error correction models that form a basis for forecasting and trading strategies. Thus, the research question of this thesis reads as follows:

Is it possible to earn an abnormal return in the Japanese market, as represented by the Nikkei 225 Stock Average Index, by trading in the spot and futures market based on an error correction model?

The purpose of this research question is to both describe a potential lead-lag relationship and to investigate the potential for profitable trading given transaction costs, thus testing for semi-strong market efficiency in the chosen market. The question also gives other interesting implications because of the time span of the investigated period. For example, the effects of the financial crisis of 2008 can be investigated by dividing the period into two sub periods, one before and one after the crisis.

The rest of this thesis paper is organized in the following manner: First, the literature review chapter aims at presenting the theoretical foundation of the efficient market hypothesis and arbitrage, as well as a selection of previous research and findings on the lead-lag relationship between spot and futures prices. Here, the focus will be on research based on co-integration methodology, error correction models and Granger causality, and there will be a separate section for research related to the Japanese market and the Nikkei 225 Stock Average Index. Secondly, there will be a chapter presenting the essential methodology, namely stationarity and unit root tests, co-integration tests, error correction models, Granger causality and ARIMA modelling. Thirdly, the data used in the study is presented and the limitations of the data are discussed, before the results and interpretations of the econometric analysis are outlined in the following section. Next, the effectiveness of different trading strategies is investigated, before a conclusion is drawn. Finally, there is a critique of the thesis and suggestions for further research are presented.

2 Literature review

2.1 The random walk model and the efficient market hypothesis

Bachelier (1900) was the first to describe the determination of successive stock price movements by the stochastic process called Brownian motion. His PhD is considered the beginning of modern finance (Walter, 2003). However, it was not until the mid 1900s that the ideas of Bachelier caught widespread interest in the field of mathematical and empirical finance. Kendall and Hill (1953) were among the first to adopt his theories. Studying British stock indices, they concluded that there was no point in trying to predict future stock prices based solely on the past because stock price changes appeared to follow a random walk. Later, Roberts (1959) found similar results while analysing American stock prices. Also, when searching for a significant serial correlation coefficient in American stock price time series, Fama (1965) was unable to reject the hypothesis of the random walk phenomenon in stock prices. Neither was he able to use mechanical trading strategies to create a greater profit than that of a passive buy-and-hold strategy.

Later, Fama (1970) formulated the Efficient Market Hypothesis (EMH), a hypothesis that relies on the notion of the random walk of stock prices. If successive stock prices truly are independent, it is reasonable to suggest that financial securities markets, where stock prices are the underlying asset, are efficient in the sense that they absorb and reflect available information as it reaches the markets. Hence, in an efficient market, the current price of an asset is no indicator of the future price and information efficiency will ensure that the security price is an unbiased estimate of the true value of the underlying asset.

Furthermore, the EMH implies that financial derivatives must be traded at a fair price in an efficient market, giving no opportunity for long-term arbitrage and making investors unable to outperform the market without taking on excess risk (McDonald et al., 2006). The no-arbitrage rule of an efficient market also gives rise to the law of one price, which states that securities with the same future cash flows must be identically priced today.

Fama (1970) also categorized the EMH into three levels on which a market can be said to be efficient. Each level corresponds to an assumption about what type of information is available and therefore can be reflected in the prices. Under the weak form of market efficiency, it is assumed that all historical information is reflected in current market prices. The semi-strong form of market efficiency implies that all publically available information also should be reflected in current prices. Under the strong form of market efficiency private information, as well as historical and public information, is assumed to be reflected in the market prices. Fama pointed out that there was no significant evidence against the weak form of market efficiency (e.g. Bachelier (1900), Kendall and Hill (1953), Roberts (1959), Osborne (1959), Fama (1965)). However, Fama also emphasized that there is still much work to be done.

To test the semi-strong form of the EMH essentially means testing the speed of adjustment of a market to newly arrived information. Cowles (1933) was the first to empirically test the EMH in its semi-strong form. He examined the results from trading according to the advice of 45 American professional financial analysts, but found that no professional trading strategy was able to outperform a passive buy-and-hold strategy. Fama et al. (1969) later performed event studies on how stock prices adjusted to new information such as stock splits and dividend announcements on the New York Stock Exchange (NYSE). They concluded that the market adjusted very quickly to new information, implying that investors were more likely to make profits by chance than because of their knowledge and skills. Similar findings have been made in other studies, for example by Jensen in 1968 (Yen and Lee, 2008).

In terms of tests of strong market efficiency, Niederhoffer and Osborne (1966) found that two successive price changes in a positive direction had a larger probability of being followed by a new positive price change than the probability of a successive negative change in the price. Their conclusion implied that a specialist at the NYSE, e.g. a stock market investor with private information, could be able to anticipate non-random stock price movements and make a profit by using their monopolistic access to information. This result is in disfavour of the strong form of the EMH. However, an efficient market in the strong form requires transaction costs and the costs of getting information to be zero (Grossman and Stiglitz, 1980). Because there are in fact information and transaction costs, the strong from of the EMH must be rejected (Fama, 1991).

The empirical research done on the EMH has been substantial since its heydays in the 1960s (Fama, 1991) even though there is a serious joint hypothesis problem associated with testing the EMH. Market efficiency per se is not testable because a test of market efficiency must involve jointly testing both market efficiency and an equilibrium model of asset price determination (Jensen, 1978). However, this does not make empirical research on market efficiency uninteresting, but improves the understanding of the behaviour of returns on financial assets (Fama, 1991). Especially tests of the semi-strong form of the EMH are intriguing because they might come closest to allowing a break in the joint hypothesis problem and thus might give the most direct evidence on market efficiency (Fama, 1991). In this context, a significant branch of empirical literature and research has been conducted in the field of lead-lag relations between different derivatives markets. This literature is reviewed in the following section, with an emphasis on the relationship between spot and futures prices with stock indices as the underlying asset.

2.2 The lead-lag relationship between spot and futures

Theoretically, there should not exist opportunities for arbitrage in a perfect and efficient market and the contemporaneous returns from spot and futures prices on the same underlying asset should be perfectly correlated (McDonald et al., 2006). A futures contract is a legally binding agreement to buy or sell a certain quantity of the underlying asset at a price set today on a given date in the future. The futures price should be an unbiased predictor of the future spot price, and both the spot price and the futures price should reflect new information simultaneously as it reaches the market. This is consistent with the efficient market hypothesis. The theoretical relationship between the futures price and the underlying spot price based on the assumptions above is known as the cost-of-carry model (Cornell and French, 1983). According to the cost-of-carry model (COC), the futures price can be written as:

$$F_t = S_t e^{[(\mathrm{r} - \mathrm{d})(\mathrm{T} - \mathrm{t})]}$$

Here, F_t is the futures price at time t, S_t is the value of the underlying spot price at time t, r is the continuously compounded risk free rate of return and d is the continuously compounded dividend yield. T is the maturity date for the futures contract.

However, the existence of for example transaction costs and private information renders the markets imperfect and give rise to the preference of one market over another. These preferences, along with several technical reasons like nonsynchronous trading hours and differences in trading frequencies, can cause the returns of one market to lead the returns of another market. If such a lead-lag relationship exists it would be possible to profit from strategic trading. Consequently, the existence of a lead-lag relationship and the possible arbitrage opportunity would contradict the efficient market hypothesis and imply joint inefficient markets.

2.3 Previous research on the lead-lag relationship between spot and futures

Zeckhauser and Niederhoffer (1983) were among the first to investigate the relationship between stock index spot prices and stock index futures prices. They found that in the short run the futures prices could to some extent predict movements in the spot price for the S&P 500 index. However, they underlined that any relationship that they found then, considering the newness of the futures market at that time, would be expected to change as the traders gained more experience and knowledge.

Still, Kawaller et al. (1987) found a simultaneous minute-to-minute relationship between S&P 500 index spot and futures prices. Furthermore, their results suggested a lead from futures to cash prices lasting from 20 to 45 minutes. On the other hand, the lead from cash to futures prices lasted for less than a minute. A three-stage least-squares regression was used in this study. Herbst et al. (1987) also studied the S&P 500 index by applying spectral analysis and correlation analysis. Using both daily and tick-by-tick intraday data form selected days, their findings supported those of Kawaller et al. (1987). However, they found that the lead from futures to spot prices only lasted up to eight minutes, with an average lead of less than a minute. They concluded that such a short lead was unlikely to provide any profitable advantage.

Further evidence of the lead from the futures price to the spot price was provided by Stoll and Whaley (1990). They applied the ARMA technique and a multiple regression framework on intraday returns from spot and futures prices for the S&P 500 index and the MMI index. They found a five-minute lead from futures prices to spot prices, but that the markets for the most part were contemporaneous. They suggested that the lead most likely was caused by noncontinuous trade of some of the stocks in the index.

Chan (1992) found that when more stocks move together the futures price leads the spot price to a greater extent. He suggested that this effect of market-wide information implied that the markets have different access to information and that the futures price reflects new information better. His findings were based on intraday returns in five-minute intervals from the MMI cash index and futures prices and the S&P 500 index spot and futures prices. Chan used methods of serial and cross correlation as well as regression to execute his study.

The Hong Kong market and specifically the Hang Seng index was investigated in the study by Tang et al. (1992). They used daily settlement prices from the spot month and divided the data into two periods, before and after the crash in the stock market in October 1987. Their study discovered that the index futures prices seemed to lead the index spot prices in the pre cash period, and that the post crash period suggested a bi-directional relationship between the prices. They suggested that the change of causality was caused by less speculative trade because of higher perceived risks of trading as a consequence of the crash. Their discussion focused on how only institutional investors remained, and how their participation in both markets in order to hedge their investments caused the feedback effect between the markets. They applied Granger causality tests, VAR techniques and a Hsiao operational test to test the relationship between the two prices.

Wahab and Lashgari (1993) also investigated the S&P 500 index spot and futures prices. Their study also considered the FTSE 100 index spot and futures prices. They found that in both cases the cash and futures prices were co-integrated and that they could employ an error correction model to each series. However, their daily data indicated that the prices were highly simultaneous and that any lag was so short that there was no predictive power, supporting the efficient market hypothesis.

Co-integration techniques and error correction models were also the methods used by Ghosh (1993) when he studied the causality relationship between spot and futures prices for the underlying S&P 500 index and the CRB index. The results of the study provided evidence inconsistent with the efficient market hypothesis, as both series seemed to be co-integrated and related in a long-term equilibrium relationship. Furthermore, he tested the predictive power of the error correction models compared to a simple univariate OLS model and found that the error correction models outperformed the naïve strategy. However, he did not consider transaction costs, so no conclusion was made in regards to market efficiency.

Abhyankar (1995) found that the relationship between the FTSE 100 index futures and cash markets was a strong and simultaneous one in the sample period of hourly data from April 1989 to March 1990. The period was divided into three sub periods; the period before major structural reforms in October 1986, the following period up until the crash of 1987, and finally, the period after the crash. The results were in line with previous studies where the futures market seemed to lead the cash market in all three periods. Furthermore, his results implied that lower transaction costs could explain part of the preference for futures markets. He also found that the lead-lag relationship was affected by the type of news that hit the market, “good”, “moderate” or “bad” news, and also by trading volume and volatility trends. High trading and “moderate” news gave the clearest lead pattern from futures to spot prices.

Other European market indices were studied by Clare and Miffre (1995). They investigated the French MATIF CAC 40 futures contract and the German DFB DAX futures contract by building ex ante models for the stock index futures contracts based on stock and bond market forecasting variables. Their findings implied that the German market was indeed efficient and that there was no opportunity to develop a profitable trading strategy based on the model. However, the French market seemed to be weakly inefficient, and their trading rule did allow for profitable strategies. They observed weekly returns in their study.

Antoniou and Holmes (1996) continued the study of the FTSE 100 index futures contract by using the Johansen co-integration method, variance-bound tests and error correction models to investigate the efficiency both in the long run and in the short run. Their findings were that futures prices seemed to be unbiased predictors of spot prices for one, two, four and five months prior to maturity, but not for three and six months prior to maturity. They concluded that the markets for the most part were efficient, but that they showed signs of weak inefficiency in the contract months.

Pizzi et al. (1998) found that the S&P 500 index spot price and the matching futures price had a bi-directional relationship for both three- and six-month contracts in the period of January 1987 to March 1987. The futures market seemed to lead the spot market by at least 20 minutes, and the spot market seemed to lead the futures market by at least four minutes. They applied co-integration and error correction methods to their intraday minute-to-minute dataset.

Information efficiency in the Australian All-Ordinaries Index (AOI) spot and futures market diverged from the results of previous studies of lead-lag relationships. Turkington and Walsh (1999) found a strong bi-directional relationship, or feedback, between the spot and futures market when they investigated intraday five-minute interval data from January 1995 to December 1995. They observed that one market caused the other market to continue reacting for several lags, up to an hour, and none of the markets seemed to adjust more quickly than the other. The determining factor was simply which market reflected the new information first. They applied vector error correction models, ARMA and VAR models, and Granger causality and co-integration methodology.

When Brooks et al. (2001) studied the lead-lag relationship between the FTSE 100 index spot and futures prices, they also tested different trading strategies and their profitability after the deduction of transaction costs. They found, like many before them, that the futures price lead the spot price, and they tested the predictive power of different methods on an out-of-sample period. These methods consisted of two error correction models; one based on the cost-of-carry model and the other based on the traditional co-integration equation, and also a VAR and an ARMA/ARIMA model. The error correction model based on the cost-of-carry model (ECM-COC) had the strongest predictive power. They used a co-integration framework for testing their 10-minute interval observations from June 1996-1997. Even though several of the trading strategies based on the ECM-COC model provided higher returns than a passive benchmark, none of them could produce higher profits after taking transaction costs into account.

2.4 Recent studies and emerging markets

The Greek market in the period from 1999 to 2001 was investigated by Floros and Vougas (2008). They found a strong long-run relationship between spot and futures prices, and in line with previous empirical findings, they found that the futures market seemed to reflect information more quickly than the spot market. They used daily data from Athens Stock Exchange and the ADEX, and applied impulse response functions and vector error correction models.

Zakaria and Shamsuddin (2012) looked at the emerging Malaysian market in the period from January 2006 to November 2011. They used daily data and used unit root tests, co-integration tests and Granger causality, and they found a long-run relationship between the spot index and futures contract. Furthermore, they discovered that the direction of causality was unidirectional with the cash market leading the futures market. This finding suggested that the cash market reflected information more quickly than the futures market, the opposite direction of that of many of the previous studies for developed markets.

Following the trend of investigating the lead-lag relationship in emerging markets, Yang et al. (2012) used intraday five-minute interval data to study the emerging Chinese market. They observed the CSI 300 Index and the CSI 300 Index futures contract from April 2010 to July 2010. They used a recursive co-integration technique and an asymmetric ECM-GARCH model alongside the augmented Dickey-Fuller test and the Phillips-Perron test. Their findings supported those of Zakaria and Shamsuddin (2012). Judge and Reancharoen (2014) provided further support of this causality direction for emerging markets as they found that the Thai cash market lead the Thai futures market. They investigated daily data of the SET 50 spot and futures prices in the period 2006 to 2012.

One of the most recent studies was performed on the Indian market by Patra and Mohapatra (2014). They observed intraday five-minute interval data of ten blue chip Sensex stocks from diverse sectors and their futures prices and found a strong bi-directional and contemporaneous relationship between the two markets. In some cases the futures market lead the spot market, and in other cases it was the opposite. They also observed a volatility spillover from the futures market to the spot market, which strengthened the implications of their results suggesting that the futures market had a leading price discovery role for most stocks. Their methods of choice were cross correlation tests, Granger causality, VAR and GARCH frameworks. The study covered the 12-month period of 2012.

2.5 Previous research on the lead-lag relationship between spot and futures: The Japanese market

There is also a good selection of literature on the Japanese market and the Nikkei 225 Stock Average Index. Sinha (1991) was among the first to study the relationship between spot and futures prices with the Nikkei 225 as the underlying asset. His study related specifically to the

1987 crash and considered the aspect that the futures traded on SIMEX instead of on a domestic market. He applied Granger causality tests, regression analysis and ARIMA frameworks on daily data from September 1987 to March 1988. The results indicated that spot prices lead futures prices. Lim (1992) further investigated the relationship using intraday data with five-minute intervals. Considering four contracts – June 1988, September 1988, June 1989 and September 1989 – Lim found that arbitrage opportunities were very few and that the Nikkei 225 spot market seemed to be weak-form efficient. Arbitrage tests, tests for autocorrelation and the cost-of-carry model were applied to examine the observations. The results showed no sign of a lead-lag relationship from spot to futures prices or from futures to spot prices.

Tse (1995) used daily observations from December 1988 through January 1993 when he explored the relationship between the Nikkei 225 index spot and futures prices. Through extensive analysis using cost-of-carry- and Engle and Granger co-integration framework, error correction models, VAR and ARIMA methodology, he found that the futures prices lead the spot prices with daily lags. In order to test the predictive power of the lead-lag relationship and the different models, out-of sample data from February 1993 to April 1993 was used for a forecast comparison. A naïve martingale forecast was used as a base, and root mean squared error, mean absolute error, and percentage of correct prediction of direction of movements was used for the comparison. The error correction models based on the OLS residuals and the cost-of-carry model got the highest score, and predicted more than 60% correct directional movements.

Further support for the hypothesis that the Nikkei 225 futures market reflects information more quickly than the spot market was provided by Hiraki et al. (1995). They applied a GARCH framework to end-of-day return observations for the Nikkei 225 index in the period September 1988 through June 1991. Iihara and Kato (1996) also observed a strong lead from the futures to the spot market, but also a weak lead from the spot to the futures market. They considered intraday data from the period March 1989 to February 1991. GMM, ARCH and GARCH models were among the chosen methods of analysis.

Chiang and Wang (2008) looked at regime switching co-integration tests for different Asian stock index futures. They studied the MSCI Taiwan, the Nikkei 225, the Hong Kong Hang Seng and the SGX Straits Times indices. The Nikkei 225 observations were daily prices from

January 1995 through December 2003. Their findings showed that all of the indices were co-integrated in terms of spot and futures prices. Furthermore, the regime switching co-integration models seemed to better capture the long-run relationship among the variables than usual co-integration techniques.

3 Methodology

From the literature review it is clear that many different econometric approaches have been applied to investigate the relationship between spot and futures prices. One of the most popular methods, especially in more recent studies, has been the co-integration approach. The co-integration approach, alongside the error correction methodology, has made it possible to empirically investigate both the long-term and the short-term relationship between the two prices. The Engle and Granger or the Johansen approach is usually chosen to check for co-integration, and the Dickey-Fuller test or the Phillips-Perron test is used to check the required conditions for co-integration. Error correction models are included in both the Engle and Granger and the Johansen method. ARIMA models are also common approaches when dealing with time series. They are used to say something about the value of a variable today based on the past values of the same variable. The Granger causality test can be applied to describe the causal relationship between the variables. In the following section, the above-mentioned approaches will be further explained.

3.1 Stationarity and unit roots

3.1.1 The Dickey-Fuller test for unit roots

The objective of the Dickey-Fuller (DF) test is to examine the null hypothesis that $\phi = 1$, the process has one unit root, in the AR(1) model expressed by the following equation:

$$x_t = \phi x_{t-1} + e_t \tag{1}$$

The alternative is $\phi < 1$. Thus, the hypotheses can be given as:

H_0: The series contains a unit root
H_1: The series is stationary (does not contain a unit root)

Usually, economic time series are not mean reverting, i.e. not stationary, in levels. This becomes apparent when inspecting the time series variables graphically. An appropriate measure to transform the series into a stationary series is differencing the variables before applying the DF test for unit roots. Transform equation (1) into the first difference by subtracting x_{t-1} from both sides:

$$x_t - x_{t-1} = \phi x_{t-1} - x_{t-1} + e_t$$

$$\Delta x_t = (\phi - 1)x_{t-1} + e_t$$

$$\Delta x_t = \psi x_{t-1} + e_t \tag{2}$$

Now it is evident that the null hypothesis outlined above is equivalent to a test of $\psi = 0$ against the alternative $\psi < 0$. If $\psi = 1$, the AR(1) process follows a pure random walk (RW) where the variable value in the last period has no impact on the current value. Note that a time series that needs to be differenced *d* times before becoming stationary is defined as integrated of order *d*, e.g. $x_t \sim I(d)$. Most financial and economic time series variables are I(1) (Brooks, 2008).

The presence of a unit root in an AR(1) process can also be tested using two alternative DF test equations (Dickey and Fuller, 1979):

$$\Delta x_t = \beta + \psi x_{t-1} + e_t \tag{3}$$

$$\Delta x_t = \beta_1 + \beta_2 t + \psi x_{t-1} + e_t \tag{4}$$

Equation (3) includes a constant and is used to test for a unit root when the time series variable is a random walk with a drift. This way, the process includes a deterministic predictable trend in addition to the stochastic unpredictable trend. This is often the case for macroeconomic variables (Asteriou and Hall, 2007).

Apply equation (4) if the AR(1) process becomes stationary when a time trend variable is included as an independent variable. Including the trend variable ensures stationarity around the trend, i.e. the stochastic part of the process will not add to the deterministic trend of the process, but will instead disappear over time.

Before performing the DF test for unit roots it is important to know which DF test equation is best suited to the time series in question. Using a wrongly specified DF test equation will cause bias in the estimates (Campbell and Perron, 1991). When a visual inspection of the time series does not reveal which specifications are correct, a sequential test using different DF test equations is recommended. If all DF test equations give approximately the same conclusion, neither has proven to be more correct. The null hypothesis is the same for the DF test using all three DF test equations.

The Dickey-Fuller test statistic is defined as:

$$\tau = \frac{\hat{\psi}}{SE(\hat{\psi})} \tag{5}$$

The DF test statistic does not follow the Student's t distribution under the null hypothesis, but follows a non-standard distribution where critical values are derived from simulation (Fuller, 1996). In absolute terms, the DF critical values are larger than the standard normal critical values. Thus, more evidence against the null hypothesis is required to reject it when testing for unit roots than under standard t tests.

3.1.2 The augmented Dickey-Fuller test for unit roots

The Augmented Dickey-Fuller (ADF) test includes lags of the time series variable that is tested. This is done to model serial correlation in the error term e_t. Thus, the ADF test equation in the case of a RW with a drift (equation (3)) becomes:

$$\Delta x_t = \beta_0 + \psi x_{t-1} + \sum_{i=1}^{p} \beta_i \Delta x_{t-i} + e_t \tag{6}$$

In equation (6) $\gamma - 1 = \psi$ and the index p indicate the number of lags included to ensure no serial correlation in the residuals. The hypotheses are formulated as before.

It is necessary to include the appropriate number of lagged differenced terms. Including more lags than necessary to remove all significant serial correlation in the dependent variable would result in more type II errors, i.e. accepting a false null hypothesis. This is due to the large number of parameters to be estimated that affect the degrees of freedom and cause the absolute value of the test statistic to decrease. If too few lags are included, more type I errors will be made and there will still be serial correlation present in the estimated model. To determine the appropriate number of lags to include, it is recommended to choose the number of lags that minimizes an information criterion such as the Akaike Information Criterion (AIC) or the Schwartz Bayesian Information Criterion (SBIC) (Brooks, 2008), as specified in appendix A.1.

An important criticism of unit root tests such as the DF and ADF test is that they are rather poor at deciding when the underlying process of the time series is stationary but has a unit root close to the non-stationary boundary (Brooks, 2008). This is particularly a problem when working with small samples. The reason for this is the way the hypotheses of the tests are formulated. The null hypothesis is never accepted, only rejected or not rejected. Therefore, if the null hypothesis is not rejected this could either be because the null hypothesis was correct or because of insufficient information in the sample to enable rejection.

3.1.3 Phillip-Perron (PP) tests for unit roots

Phillip-Perron (PP) tests for unit roots (Phillips and Perron, 1988) are similar to the ADF test. However, in PP tests, an automatic correction is added to the DF test equation in order to allow for autocorrelated residuals (Brooks, 2008). The test equation for the PP test is the AR(1) process (equation 2). By correcting the t statistic for the coefficient ψ, potential autocorrelation in the error term is taken into account (Asteriou and Hall, 2007). The conclusions drawn from PP tests are most often the same as those from ADF tests for unit roots.

3.2 Co-integration and error correction models

3.2.1 Co-integration and error correction models

If two non-stationary variables $\{x, y\}$ are integrated of order one (i.e. $\{x, y\} \sim I(1)$) and there is a linear combination of the two variables that is stationary (I(0)), the variables are said to be co-integrated (Brooks, 2008). Co-integration should in theory only exist if there is a true relationship between the two variables, as would be the case for economic structures

(Asteriou and Hall, 2007). This relationship is also referred to as a long-run equilibrium relation as there will be a common trend linking the co-integrated variables together in spite of different movements in each variable. A linear combination of x_t and y_t can be extracted from the estimation of the following regression:

$$x_t = \beta_0 + \beta_1 y_t + e_t \tag{7}$$

and taking the residuals:

$$\hat{e}_t = x_t - \hat{\beta}_0 - \hat{\beta}_1 y_t \tag{8}$$

If $\hat{e}_t$ ~ I(0), the variables, x and y, are co-integrated. Granger (1969) was the first to introduce the concept of co-integration, and Engle and Granger (1987) further developed the concept by elaborating on the linkage between co-integrated variables and error correction models. Error correction models (ECM) takes into account both the long-term relation between the co-integrated variables and the short-term relation by using both first differenced and lagged levels of the variables (Brooks, 2008). This approach was developed as a solution to the problem related to investigating the relationship between non-stationary variables. The usual treatment by differencing the non-stationary variables was not sufficient when the long-term relation between two or more variables was of interest. An example of an error correction model could be:

$$\Delta x_t = \beta_0 + \beta_1 \Delta y_t + \beta_2 (\hat{e}_{t-1}) + u_t \tag{9}$$

Here, $\hat{e}_{t-1} = x_{t-1} - \hat{b}_0 - \gamma y_{t-1}$ is the lagged version of the $\hat{e}_t$ found above, and it is known as the error correction term. The error correction term should be I(0) in spite of the x and y being I(1). γ is the co-integration coefficient and defines the long-term relationship between x and y, β_2 is the speed of adjustment back to the defined equilibrium state, and β_1 describes the short-run relationship between changes in the respective variables. The Granger Representation Theorem (Engle and Granger, 1987) states that if there is co-integration between two series of variables then there will also be a corresponding error correction representation.

Engle and Granger two-step method (Engle and Granger, 1987) and the Johansen approach (1988) are two of the most commonly applied procedures for testing co-integration. The Granger causality test (Granger, 1969) is used to determine the existence and direction of causality of the relationship between the variables.

3.2.2 Engle and Granger two-step method

The Engle and Granger two-step method is a single equation technique and is very popular because of its simplicity and intuitive nature. The first step investigates the long-term relationship by regressing the individual I(1) variables using OLS (Brooks, 2008). The residuals from this co-integration regression are then saved and tested for unit roots. The Dickey-Fuller or the Augmented Dickey-Fuller test can be applied, but the critical values are not the same as for standard unit root tests. Critical values such as those found in MacKinnon (2010) must be used. If the residuals are I(0), stationary at level, the next step is introduced. Step two presents the error correction model and the residuals are included as one of the variables, i.e.:

$$\Delta x_t = \beta_0 + \beta_1 \Delta y_t + \beta_2 (\hat{e}_{t-1}) + u_t \tag{10}$$

where $\hat{e}_{t-1}$ is defined as mentioned above.

However, the Engle and Granger method has some problems (Asteriou and Hall, 2007). The first problem concerns the order of the variables in the co-integration regression, and the choice of independent and dependent variable can in some cases affect the stationarity properties of the residuals. Furthermore, the method can only be applied in those circumstances where only two variables are relevant because it only accounts for one set of residuals. Finally, the two-step nature of the method might be an issue, as any mistake made in the first step would affect the results in the second step.

3.2.3 The Johansen approach

The Johansen approach provides a remedy for one of the problems of the Engle and Granger two-step method by allowing to include more than two variables and thus more than one co-integrating vector (Asteriou and Hall, 2007). This test is therefore a more general approach for testing co-integrating relationships. Vector autoregressive models (VARs) are the basis of the test. VARs are system regression models with traits from univariate time series models and traits from simultaneous equations models (Brooks, 2008). When investigating a set of *g* I(1) variables ($g \geq 2$), a VAR with *k* lags including these variables could be modelled like this:

$$\underset{g\times 1}{x_t} = \underset{g\times g}{\beta_1}\underset{g\times 1}{x_{t-1}} + \underset{g\times g}{\beta_2}\underset{g\times 1}{x_{t-2}} + \cdots + \underset{g\times g}{\beta_k}\underset{g\times 1}{x_{t-k}} + \underset{g\times 1}{e_t} \tag{11}$$

The VAR must be transformed into a vector error correction model (VECM) in order to be used in the Johansen test:

$$\Delta x_t = \prod x_{t-k} + \Gamma_1 \Delta x_{t-1} + \Gamma_2 \Delta x_{t-2} + \cdots + \Gamma_{k-1} \Delta x_{t-(k-1)} + e_t \tag{12}$$

where

$$\prod = \left(\sum_{i=1}^{k} \beta_i\right) - I_g \text{ and } \Gamma_i = \left(\sum_{j=1}^{i} \beta_j\right) - I_g$$

The system is estimated by maximum likelihood. On the left side there are *g* variables in first differenced form and on the right side there are *k – 1* lags of the dependent variable, also differenced. The test is based on the relationship between the rank of a matrix and its eigenvalues. The focus is on the $\prod$ matrix, also known as the long-run coefficient matrix. The co-integration test can be performed using two different test statistics; the Trace statistic and the Maximum eigenvalue statistics, formulated as:

$$\lambda_{trace}(r) = -T \sum_{i=r+1}^{g} \ln\left(1 - \lambda_i\right) \tag{13}$$

and

$$\lambda_{max}(r, r+1) = -T \ln\left(1 - \lambda_{r+1}\right) \tag{14}$$

where *r* is the number of independent co-integrating relations, λ_i is the number of estimated values of the eigenvalues, and *T* is the number of usable observations. The trace test is considering the null hypothesis that there are less than or equal to *h* co-integrating relations ($r \leq h$) as opposed to the alternative hypothesis that there are more than *h* co-integrating relations ($r > h$). This test is therefore a joint test (Brooks, 2008). On the other hand, the maximum eigenvalue test tests each eigenvalue individually. The null hypothesis is that the number of co-integrating vectors is $r \leq h$ and the alternative hypothesis is that $r = h + 1$.

3.2.4 Granger causality

The concept of Granger causality was first developed by Granger (1969). It is a common method of empirically testing the direction of causality between two variables that are assumed to be related. The causality can be unidirectional, where one variable Granger causes the other, or bi-directional, where there is a feedback effect and the two variables cause each other. In a bivariate framework, x_t is said to Granger cause y_t if the forecast for y_t improves when lagged variables of x_t are taken into account in the forecasting equation (Brooks, 2008). The test can in the case of two stationary variables generally be expressed as:

$$y_t = \beta_{0_1} + \sum_{i=1}^{n} \beta_i x_{t-i} + \sum_{j=1}^{m} \gamma_j y_{t-j} + e_{1t} \tag{15}$$

$$x_t = \beta_{0_2} + \sum_{i=1}^{n} \theta_i x_{t-i} + \sum_{j=1}^{m} \delta_j y_{t-j} + e_{2t} \tag{16}$$

where the error terms are assumed to be uncorrelated white-noise error terms (Asteriou and Hall, 2007). The test is a normal Wald test using F statistics. If none of the coefficients are statistically significant, x_t and y_t are said to be independent. Brooks (2008) points out that the meaning of the word "causality", in terms of Granger causality, does not refer to how movements of one variable might cause the movement of the other variable. In fact, it refers to a correlation between the current value of one of the variables and the past values of the other variables.

3.3 ARIMA models

An autoregressive integrated moving average (ARIMA) model, also called a Box-Jenkins model (Box and Jenkins, 1976), is a generalization of an autoregressive moving average (ARMA) model. The ARIMA model consists of three parts: the autoregressive component (AR) indicates the included number of lags of the dependent variable. The order of integration component (I) suggests the order of integration, and the moving average (MA) component captures the effect of the error term by including q lags of past innovations.

The Box-Jenkins methodology for estimating ARIMA models involves a three-step process elaborated in Brooks (2008): In the first identification step, the order of the model is determined by inspecting the sample autocorrelation function (ACF) and sample partial autocorrelation function (PACF). In the next step, the estimation step, the model parameters are estimated using maximum likelihood (ML) techniques. Finally, in the third step, residual diagnostic checks of the model are performed. This is in order to determine if the fitted model is adequately specified and estimated. A more detailed description of the ARIMA modelling method is included in appendix A.2.

3.4 Forecasting and trading strategies

After an appropriate time series model has been fitted to the variables, the model may be used to generate a forecast of future observations and the direction of future price changes. The purpose of trading according to the model forecasts by using a specified trading strategy is to invest in a way that creates a profit superior to that of a benchmark profit. The trading strategies discussed below are inspired by Brooks et al. (2001).

3.4.1 The passive buy-and-hold strategy

The passive buy-and-hold strategy involves buying the index and holding it until the final day of the predetermined investment period. When using this trading strategy, transaction costs will be minimized. This strategy is used as a benchmark for other active trading strategies.

3.4.2 Liquid trading strategy

A liquid trading strategy involves buying the index and reinvesting any profits made when the model predicts a positive price change in the next investment period. Should the prediction be a negative price change in the next period, the investor will close his position. As long as no investments are made, the investor is assumed to earn a risk free interest rate. A new position is opened when the forecasting model predicts a positive price change.

3.4.3 Filter strategy

The filter strategy involves buying the index when the forecasting model predicts higher positive returns than the average predicted positive return. Thus, the average predicted positive return has the function of a filter rule, upon which the investor is acting. Should the model predict a negative return in the next investment period, the investor will remain passive and the investment will earn the risk free interest rate. Since no negative predicted price changes are included in the average predicted return, the number of trades made by the investor is limited.

3.5 Chosen methodology

When investigating the lead-lag relationship between the Nikkei 225 Stock Average Index spot price and the futures price based on the same index, we have chosen to apply the co-integration methodology. Furthermore, we will use the ADF test to test for unit roots and the Engle and Granger two-step approach to test for co-integration. As we only include two variables and we only want to find one co-integrating vector, this approach is more suitable than the Johansen approach. However, to make sure that the results are robust we include error correction models based on the cost-of-carry model and ARIMA models, alongside the error correction models from the two-step Engle and Granger approach. In this way, we make sure that we do not just test the model, but also thoroughly test the time series data at hand. Finally, we use an out-of-sample period to check which of the models best describe the relationship between the variables. The forecast estimation will also provide the opportunity to test different trading strategies in order to see if there is a possibility to profit from the potential relationship. Also, it will allow us to gain insight into the effects of transaction costs in the correction of market anomalies.

4 Analysis

4.1 Data

The data set consists of daily observations of the Nikkei 225 Stock Average Index (Nikkei Keikin Kabuka) quoted on the Tokyo Stock Exchange (TSE) and the futures contract based on the Nikkei 225 Stock Average Index traded at the Osaka Securities Exchange (OSE). The time series variables are collected using Thomson Reuters DATASTREAM and the Nikkei Stock Average Fact Sheet published by Nikkei Inc. (Nihon Keizai Shimbun), a leading Japanese economic newspaper.

The currency is Yen (¥) for all prices, and the dividend yield and risk free interest rate is quoted in per cent. TSE and OSE have different closing hours, which could lead to problems related to nonsynchronous trading. To avoid these problems, which may increase the probability of finding a spurious lead-lag relationship (Brooks et al., 1999), all prices are opening prices. Spot and futures opening prices will be more appropriate in the analysis as both markets open for trade at the same time. Regarding a dividend yield variable for the Nikkei 225 index, DATASTREAM is not able to provide this. Furthermore, the dividend yield is not available from any other reliable sources, such as the TSE. Therefore, as an approximation, the averaged annual Nikkei 225 index dividend yield for the years 2010-2014 (Nikkei Inc., 2014)) is used to compute the daily dividend yield. Following Tse (1995), the three month Euroyen interest rate is used as a proxy for the risk free interest rate. This is also provided by DATASTREAM. More details on the variables can be found in appendix A.3.

4.1.1 Sample sub periods

The total sample period is 01.07.2004 to 30.06.2014. Missing data from certain trading days (Monday through Friday) is due to Japanese national and observed holidays, and the market holidays of January 2nd, January 3rd and December 31st. Furthermore, the data set is divided into two sub periods for model estimation, forecasting and the testing of trading strategies; the pre financial crisis period and the post financial crisis period.

The pre financial crisis period reaches from 01.07.2004 to 15.09.2008. This sub period includes 1 098 observations where 62 observations are missing for each time series variable.

The transition from the first to the second sub period is determined by the bankruptcy of the Lehman Brothers Holdings Inc. on September 15th 2008. Thus, the post financial crisis sub period reaches from 16.09.2008 to 30.06.2014 and therefore also includes the time span of the actual crisis. The second sub period includes 1 510 potential trading days, but only 1 418 observations, due to missing observations.

For forecasting, out-of-sample periods are used for the models estimated in first and second sub period. Models estimated in the pre financial crisis sub period are tested using an out-of-sample period reaching from 16.09.2008 to 27.04.2009. Thus, the first out-of-sample period consists of 160 trading days, which is reduced to 149 observations because of missing observations. For models estimated based on the post financial crisis sub period, the out-of-sample period is 01.07.2014 to 10.02.2015. The second out-of-sample period contains a similar number of trading days and missing observations as the first out-of-sample period.

QMS EViews 7 software is used to perform the analysis. EViews uses listwise exclusion when there are missing values in the times series.

4.1.2 Data limitations

There is much consensus behind the suggestion that if there exists a lead-lag relationship between spot and futures prices, this lead only lasts up to half an hour. The use of intraday data and higher frequency data could have an impact on observed returns (Stoll and Whaley, 1990). This favours the use of intraday data of intervals shorter than half an hour. However, the use of such high frequency data has several drawbacks. One of the drawbacks is the problem concerning nonsynchronous trading, which can increase the probability of finding a spurious lead-lag relationship (Brooks et al., 1999). This problem could potentially arise due to different closing hours in the two markets where the spot and futures contracts are quoted, and also the fact that the securities are traded at different intervals throughout the day. For the OSE and TSE market, the opening hours of the two markets are the same, which makes the use of daily opening prices more suitable. Therefore, to increase the probability of discovering actual lead-lag relationships, daily logarithmic returns based on the opening price for both spot and futures prices were used in this study.

Furthermore, information on the true transaction costs involved when trading the Nikkei 225 Stock Average Index at spot price and futures price were not available for this study. This is due to the fact that the total transaction cost is dependent on several components such as stock and futures brokerage commissions and securities transfer tax when stocks are sold (Chung et al., 1994). As information on all these components was not available from the TSE or the OSE, a transaction cost proxy inspired by Brooks et al. (2001) was used when computing the profitability of strategic trading of the derivatives. To account for any bias that the use of a proxy might cause, each strategy was tested using a range of transaction cost proxies.

Also, when deciding which dividend yield to use for the Nikkei 225 Stock Average Index, very little historical information was available from the market web pages and accessible databases. Therefore, to approach a realistic dividend yield, available information from the Nikkei Inc. on the most recent five years of dividend yield was averaged. Next, the average was used as a proxy throughout the sample period. Despite using an approximation to the true dividend yield, this proxy was not entirely unrealistic since it was derived from the true dividend yield in a time period with characteristics that did not deviate very much from the sample period.

Regarding the risk free interest rate that was used both to compute the returns from a risk free position and as a cost-of-carry component, the Tokyo Interbank Euroyen three month offered rate was used as a necessary approximation.

4.2 Descriptive statistics

4.2.1 Descriptive statistics

Table 1: *Descriptive statistics spot price*

	Pre financial crisis period 01.07.04 - 15.09.08	Post financial crisis period 16.09.08 – 30.06.14	Out-of-sample period 01.07.14 – 10.02.15
Observations	1036	1417	150
Mean	14389.45	10618.71	16312.72
Median	14599.99	9834.580	15922.11
Maximum	18269.38	16269.22	18004.66
Minimum	10652.37	7059.770	14796.32
Standard Deviation	2353.464	2217.979	984.8928
Skewness	-0.080960	0.995515	0.219631
Kurtosis	1.573553	2.719635	1.466643
Jarque-Bera (prob.)	88.96518 (0.0000)	238.6938 (0.0000)	15.90084 (0.000353)

Table 2: *Descriptive statistics futures price*

	Pre financial crisis period 01.07.04 - 15.09.08	Post financial crisis period 16.09.08 – 30.06.14	Out-of-sample period 01.07.14 – 10.02.15
Observations	1036	1417	150
Mean	14387.36	10606.60	16311.80
Median	14640.00	9830.00	15915.00
Maximum	18270.00	16280.00	17970.00
Minimum	10590.00	7010.00	14710.00
Standard Deviation	2357.257	2224.755	995.2222
Skewness	-0.081751	0.988314	0.222065
Kurtosis	1.574809	2.716700	1.432995
Jarque-Bera (prob.)	88.83285 (0.0000)	235.4180 (0.0000)	16.57973 (0.000251)

The descriptive statistics of both the spot price and the futures price in the pre and post financial crisis period and the out-of-sample period share some very similar features. The mean drops significantly from the pre to the post financial crisis period and increases to a higher level in the out-of-sample period. The prices also have very similar standard deviations. Furthermore, the skewness statistics show that the prices are slightly negatively skewed in the period before the financial crisis, but positively skewed and close to 1 after the crisis. In the out-of-sample period the skewness was positive, but not as close to 1. However, the skewness in the pre crash period is so weak that the series distribution is close to symmetrical. In the post crash period, the series distribution is clearly positively skewed, but in the out-of-sample period the skewness is once again less prominent. The data is more or less platykurtic in all periods, with values for kurtosis less than 3. This is somewhat surprising as financial time series usually are leptokurtic with a higher peak at the mean and fatter tails than normal distributions. Platykurtic distributions are less peaked at the mean and have thinner tails than normal distributions.

Finally, the Jarque-Bera test statistic (JB) clearly rejects the null hypothesis of residual normality. This is expected for times series based on price data. To minimize the influence of this and of potential outliers, the logarithmic prices for spot and futures are used in the following analysis. The logarithmic price says something about the change in the price, also known as the return of the underlying asset.

4.2.2 Correlations

The theoretical relationship, the graphical impression and the inspection of data suggest that there should be a strong significant correlation between the spot and the futures price. The correlation statistics confirm this and show that the two variables are close to perfectly positively correlated. This is an indication of market efficiency. The graphical impression matches the impression from the correlation analysis.

***Table 3**: Correlations*

Correlations	***Futures***		
Spot	Pre financial crisis	Post financial crisis	Out-of-sample
Pre financial crisis	0.999590		
Post financial crisis		0.999635	
Out-of-sample			0.996738

Figure 1: *Spot and futures prices from 01.07.2004 through 30.06.2014*

Furthermore, figure 1 provides a clear picture of the financial environment, showing the distinctive drop as the world was hit by the financial crisis in September 2008. Stability for both the spot and the futures market for the Nikkei 225 index characterize the post financial crisis period, with a slight increase in prices until late 2012 where the prices rose sharply before beginning to stabilize in mid 2013.

4.3 Unit root tests

4.3.1 Pre financial crisis sample

The ADF test is employed to test for unit roots in the time series variables. The findings are reported in table 4. When determining the ADF test equation, a trend variable is never found to be significant. Therefore, test equation (3), which only includes a constant, is used in the ADF test as the time series are assumed to follow a RW with drift (Asteriou and Hall, 2007). The number of lags minimizing SIC decides the number of lags included in the test equation. As expected, the ADF tests conclude that both spot and futures prices are stationary in their first differences, e.g. they are I(1).

Table 4: *Test for unit roots in price series in levels and first difference*

	Level		First difference	
	Futures	**Spot**	**Futures**	**Spot**
N° lags included	1	0	0	0
ADF statistic (τ)	-3.346	-1.392	-32.985*	-29.749*

**1 % significant*

Note that when determining the level of statistical significance for the conclusions drawn from the ADF tests, critical values deducted from simulations by Fuller (1996) are employed. The critical values are included in appendix A.8.

4.3.2 Post financial crisis sample

As for the pre financial crisis sample, the ADF test is employed to test for unit roots in the time series variables. The tests are carried out as before and the results are reported in table 5. The conclusion is that both the spot and futures prices are I(1).

Table 5: *Test for unit roots in price series in levels and first difference*

	Level		First difference	
	Futures	**Spot**	**Futures**	**Spot**
N° lags included	2	0	0	0
ADF statistic (τ)	-0.449	-0.699	-38.332*	-32.586*

**1 % significant*

A table of critical values is included in appendix A.8.

4.4 Co-integration models – Engle and Granger step one

4.4.1 Pre financial crisis sample

In the first step of the Engle and Granger approach, equation (7) is used as to estimate a co-integration model describing the long-run equilibrium relationship between spot and futures prices, and between futures and spot prices. Thus, the estimated models reported in table 6 are the co-integration models:

$$Log(Spot_t) = \beta_0 + \beta_1 Log(Futures_t) + u_t \tag{17}$$

and

$$Log(Futures_t) = \beta_0 + \beta_1 Log(Spot_t) + v_t \tag{18}$$

***Table 6**: Co-integration regressions and ADF tests for co-integration regression residuals*

Dependent variable	$\boldsymbol{Log(Spot_t)}$	$\boldsymbol{Log(Futures_t)}$
β_0 (t statistic)	0.022** (2.702)	-0.015* (-1.818)
β_1 (t statistic)	0.998** (1 176.132)	1.001** (1 176.132)
Residual	$\boldsymbol{u_t}$	$\boldsymbol{v_t}$
N^o lags	0	0
ADF statistic (τ)	-27.904**	-27.952**

***1% significant, *10 % significant*

The estimated $\beta_1$coefficients in the co-integration regressions are the long-run coefficients. Here, they are all significant, implying that a long-run equilibrium relationship between spot and futures prices, as well as between futures and spot prices, exists.

The ADF test is employed to test for unit roots in the residuals. The test is carried out as previously described and the results are reported in table 6. The number of lags included are as suggested by the Schwartz Information Criterion (SIC). The tests conclude that for both co-integration models, equation (17) and (18), the residuals are stationary in levels, i.e. they are I(0). This confirms that both co-integration models are valid and that a long-run equilibrium relationship exists between spot and futures prices, and vice versa. Note that when testing for unit roots in the residuals, critical values deducted from simulations by MacKinnon (2010) are used to determine the statistical significance of the conclusions of the tests. The critical values are included in appendix A.8.

4.4.2 Post financial crisis sample

As for the pre financial crisis sample, the co-integrating relationship between spot and futures prices and vice versa, during the post financial crisis sample is examined using equation (17) and (18). The results are presented in table 7. For both models, the long-run coefficient β_1 is significant, indicating that a long-run equilibrium relationship between spot and futures prices exists. The ADF tests for unit roots in the co-integration regression residuals confirm the long-run relationship since both residuals are I(0). According to the Schwartz Information Criterion (SIC), three lags of the residuals are included in the tests. The results from the ADF tests are reported in table 7.

***Table** 7: Co-integration regressions and ADF tests for co-integration regression residuals*

Dependent variable	$\boldsymbol{Log(Spot_t)}$	$\boldsymbol{Log(Futures_t)}$
β_0 (t statistic)	0.052* (5.498)	-0.042* (-5.239)
β_1 (t statistic)	0.994* (1 151.649)	1.004* (1 151.649)
Residual	$\boldsymbol{u_t}$	$\boldsymbol{v_t}$
N° lags	3	3
ADF statistic (τ)	-11.824*	-11.863*

**1% significant*

4.5 Error correction models – Engle and Granger step two

The residuals from the co-integration equation are included in error correction models (ECM). The basic ECM where the futures price leads the spot price can be written as:

$$\Delta \log Spot_t = \alpha_0 + \sum_{i=1}^{r} \beta_i \Delta log\ Spot_{t-i} + \sum_{j=0}^{s} \gamma_j \Delta log\ Futures_{t-j} + \delta(\hat{u}_{t-1}) + \varepsilon_t \qquad (19)$$

and the opposite for the case where the spot price leads the futures price:

$$\Delta \log Futures_t = \alpha_0 + \sum_{j=1}^{s} \gamma_j \Delta log\ Futures_{t-j} + \sum_{i=0}^{r} \beta_i \Delta log\ Spot_{t-i} + \delta(\hat{v}_{t-1}) + \varepsilon_t \qquad (20)$$

The significant ECMs that proved to be the best models in terms of minimizing SIC and after correcting for autocorrelation and heteroskedasticity were as presented in table 8 for the pre and post financial crisis sample. Newey-West robust standard errors, using five lags, were used to correct for autocorrelation and heteroskedasticity.

Table 8: *Specification of error correction models*

Error correction models	Futures prices leading spot prices	Spot prices leading futures prices
Pre financial crisis sample	$\sum_{i=1}^{r} \beta_i \Delta log\ Spot_{t-i}$=0, s =0	$\sum_{j=1}^{s} \gamma_j \Delta log\ Futures_{t-j}$=0, r =0
Post financial crisis sample	r =1, s =1	s =1, r =1

Table 9 shows the coefficients of the preferred estimated error correction models. The alternative models can be found in appendix A.4. In the pre financial crisis sample, the best model consists of the error correction term and the first differenced logarithm of the independent variable. In other words, the model validates the assumption that there is a long-term relationship between the variables, and also that there is a strong contemporaneous relation between the changes in the respective prices. In the post financial crisis sample, the best model includes the error correction term, the differenced logarithm of the independent variable and the lagged differenced logarithms of both the dependent and the independent variable. These models suggest both a long-term and a short-term relation between the variables exist. Furthermore, the speed of adjustment for the error correction term, δ, is quite high and negative for each ECM in each period, suggesting that if the independent variable is large relative to the equilibrium state at time *t-1*, then it is expected to adjust downward in the next period. Also, because the coefficients are quite high they are expected to adjust quite quickly. As an example, in the model from the post financial crisis sample where futures prices lead spot prices, the speed of adjustment is negative 67%, which means that the variables are expected to adjust back to the equilibrium relation state at a rate of 67% at time *t-1*.

The speed of adjustment coefficient is lower in the post financial crisis sample than in the pre financial crisis sample. The slower correction back to the equilibrium state implies that deviations from this state go on for longer time spans, and short-run effects from the lagged coefficients have more influence on the dependent variable. This suggests that the respective prices adjust more slowly to changes and information in the other market, implying a minor deterioration of market efficiency.

***Table 9**: Best estimated error correction models*

ECM	Pre financial crisis sample		Post financial crisis sample	
Coefficients	**Future leads** Equation (19)	**Spot leads** Equation (20)	**Future leads** Equation (19)	**Spot leads** Equation (20)
α_0	-1.84E-05	2.52E-06	-0.000118	0.000127
β_0	-	1.015743**	-	1.000914**
β_1	-	-	-0.119456*	0.163430**
γ_0	0.846836**	-	0.864242**	-
γ_1	-	-	0.181840**	-0.216751**
δ	-0.810695**	-0.892738**	-0.670174**	-0.726984**
Adjusted R^2	0.8636	0.8660	0.8681	0.8738

**5% significant*
***1% significant*

In the models from the post financial crisis sample the coefficients for the lagged variables of the logarithmic spot and futures price are of somewhat similar size, but with opposite signs. This might indicate that these lags are spurious and almost cancel each other out in the estimated model. However, they are included in the models as they are significant and alternative models without the variables have higher SIC values and lower values for adjusted R squared (Appendix A.4). Furthermore, the squared R is lower than the Durbin Watson statistics for all of the models, suggesting that the models are not spurious. Also, the lagged variables say something about the short-term relationship between the spot and the futures price for the Nikkei 225 index. For the post financial crisis sample and specifically the futures-leads-spot ECM the coefficient for the lagged difference of the logarithmic futures price is positive. This suggests that the spot price moves in the same direction as the previous movement of the futures price. However, the coefficient for the non-lagged difference of the same variable is much higher, indicating that it explains more of the movement of the spot price. Nonetheless, the significance of the lagged variable still implies that there is in fact a price discovery role of the futures market for the spot market for the Nikkei 225 index. The spot-leads-futures ECM for the same period have the same implications for the opposite direction. There seems to be a short-run relation where the spot market has a price discovery role for the futures market for the Nikkei 225 index.

The best model for each direction in both samples has a fairly high adjusted *R* squared, suggesting that the models all explain over 86% of the variance of the dependent variable. This makes the models suitable for forecasting purposes, and the results from testing the forecasting abilities of the models will be presented in a later chapter.

The difference between the pre and the post financial crisis sample seems to be that lagged differences of the logarithmic variables are significant in the post financial crisis sample. This implies that the contemporaneous traits of the relationship remains, but that past prices also have a price discovery role. Thus, the results suggest that the markets have become less efficient.

4.6 Alternative approaches

As mentioned previously, any test of market efficiency is a joint test of market efficiency and the model that the test is based on (Jensen, 1978). Therefore, as a robustness test for the results, another error correction model based on the theoretical cost-of-carry relationship between the spot and the futures price (ECM-COC), and also a univariate ARIMA model, are included. The ECM-COC model is chosen as an alternative approach because of the theoretical link between spot and futures prices that will be represented in the error correction term of such an ECM. Thus, if the ECM-COC gives a better prediction of the spot and futures prices than the traditional ECM, changes in external factors such as the risk free interest rate and the dividend yield would seem to have a significant effect on the prices. Then, for the ECM-COC, both the changes in the spot and futures prices and the changes in the interest rate and dividend yield would affect the future prices. Univariate ARIMA models are included to investigate the opportunity that any linkages between spot and futures prices discovered by the traditional ECM are in fact not of any importance for the future price.

4.6.1 Error correction models based on the cost-of-carry relationship

The relationship between the spot and the futures prices as described by the cost-of-carry model is utilized in the following models to create error correction models. The theoretical relationship is given by the before-mentioned equation:

$$F_t = S_t e^{[(\mathrm{r} - \mathrm{d})(\mathrm{T} - \mathrm{t})]} \tag{21}$$

Taking the logarithm of this equation gives the following modification:

$$Log\ F_t = Log\ S_t + [(\mathrm{r} - \mathrm{d})(\mathrm{T} - \mathrm{t})] \tag{22}$$

The Engle and Granger two-step method is used to test for co-integration and to find the appropriate error correction model based on the co-integration equation residual. To simplify the co-integration equation the expression [(r – d)(T – t)] is calculated into one variable defined as RDT in the following. The final co-integration equation for the case where the spot price leads the futures price is defined as:

$$Log(Futures_t) = \beta_0 \log(Spot_t) + \beta_1 RDT + v_t \tag{23}$$

When the futures price leads the spot price, the co-integration equation becomes:

$$Log(Spot_t) = \beta_0 Log(Futures_t) - \beta_1 RDT + u_t \tag{24}$$

Estimated coefficients for the two co-integration models are reported in table 10 for the pre financial crisis sample (panel A) and the post financial crisis sample (panel B) respectively. As the long-run coefficient, β_1, is significant for all co-integration models and for both samples, the existence of a long-run equilibrium relationship between spot and futures prices and vice versa is implied.

Table 10: Co-integration regressions and ADF tests for co-integration regression residuals

***Panel A**: Pre financial crisis sample*		
Dependent variable	$\boldsymbol{Log(Spot_t)}$	$\boldsymbol{Log(Futures_t)}$
β_0 (t statistic)	1.000*** (40 502.100)	0.999*** (40 502.100)
β_1 (t statistic)	-0.001** (-2.412)	0.001** (2.432)
Residual	$\boldsymbol{u_t}$	$\boldsymbol{v_t}$
N^o lags	0	0
ADF statistic (τ)	-27.973***	-27.973***

****1% significant, **5 % significant*

***Panel B**: Post financial crisis sample*		
Dependent variable	$\boldsymbol{Log(Spot_t)}$	$\boldsymbol{Log(Futures_t)}$
β_0 (t statistic)	0.999* (22 007.700)	1.000* (22 007.700)
β_1 (t statistic)	0.006* (5.695)	-0.006* (-5.634)
Residual	$\boldsymbol{u_t}$	$\boldsymbol{v_t}$
N^o lags	3	3
ADF statistic (τ)	-11.750*	-11.751*

**1% significant*

The co-integration regressions show a potential shift in the relationship between the spot and the futures price from the pre to the post financial crisis period. This shift is evident by the change of sign of the coefficients. In the co-integration regression where the futures price leads the spot price the coefficient for the futures price changes sign from negative to positive. For the opposite case it is the other way around, the coefficient for the spot price changes sign from positive to negative. However, the coefficients are small, approaching zero and the change of sign might be a spurious shift with no implication for the relation.

Next, the co-integration regression residuals are tested for unit roots using the ADF test. The purpose is to confirm the implied long-run equilibrium relationship between spot and futures prices, and futures and spot prices. The ADF tests conclude that for both co-integration models in both sample periods, both residuals are stationary in levels, i.e. they are I(0). This confirms that the co-integration models are valid and that a long-run equilibrium relationship exists between spot and futures prices, and vice versa. The results from the ADF tests are reported in table 10. The Schwartz Information Criterion (SIC) is used to determine the number of lags of the residual to be included.

Finally, the I(0) residuals are included in error correction models (ECM-COC). The preferred models for each direction and each sample are selected based on minimizing SIC for significant models as previously. The basic ECM models that were tested can be expressed as before as:

$$\Delta \log Spot_t = \alpha_0 + \sum_{i=1}^{r} \beta_i \Delta log\ Spot_{t-i} + \sum_{j=0}^{s} \gamma_j \Delta log\ Futures_{t-j} + \delta(\hat{u}_{t-1}) + \varepsilon_t \tag{25}$$

for the case where futures prices leads spot prices, and:

$$\Delta \log Futures_t = \alpha_0 + \sum_{j=1}^{s} \gamma_j \Delta log\ Futures_{t-j} + \sum_{i=0}^{r} \beta_i \Delta log\ Spot_{t-i} + \delta(\hat{v}_{t-1}) + \varepsilon_t \tag{26}$$

for the case where spot prices leads futures prices. When performing the regressions with Newey-West robust standard errors, the best models proved to be as those presented in table 11.

Table 11: *Specification of error correction models*

Error correction models – COC	**Futures prices leading spot prices**	**Spot prices leading futures prices**
Pre financial crisis sample	$\sum_{i=1}^{r} \beta_i \Delta log\ Spot_{t-i}$=0, s =0	$\sum_{j=1}^{s} \gamma_j \Delta log\ Futures_{t-j}$=0, r =0
Post financial crisis sample	r =1, s =1	s =1, r =1

The best models for the cost-of-carry version of the ECM approach are presented in table 11 and they are the same as for the regular ECM. Again, the models for the pre financial crisis period confirm that there seems to be a long-term relationship between the two prices and also a contemporaneous relationship. For the post financial crisis period the models imply both a long-term and a short-term relationship. Furthermore, the speed of adjustment coefficient for the error correction term of each model is yet again quite high and negative. Alternative models can be found in appendix A.5.

Table 12: *Best estimated error correction models*

ECM – COC	Pre financial crisis sample		Post financial crisis sample	
Coefficients	**Future leads** Equation (25)	**Spot leads** Equation (26)	**Future leads** Equation (25)	**Spot leads** Equation (26)
α_0	-1.56E-05	1.15E-06	-0.000102	0.000113
β_0	-	1.016814**	-	0.996901**
β_1	-	-	-0.125218*	0.164746**
γ_0	0.846855**	-	0.867287**	-
γ_1	-	-	0.187407**	-0.217851**
δ	-0.811928**	-0.895487**	-0.655735**	-0.720690**
Adjusted R^2	0.8642	0.8656	0.8672	0.8737

** 5% significant*
*** 1% significant*

However, yet again the lagged variables of the logarithmic spot and futures prices are fairly equal in size, but with opposite signs, signalling potential spurious models. This is observed in table 12 where the coefficients of the models are presented. But, as discussed above, the adjusted squared *R* is lower than the Durbin Watson statistic, which gives reason to question the signals and to keep the models (Appendix A.5). Also, as for the traditional error correction models, the best model for each direction in both samples has a fairly high adjusted *R* squared. Here, once again, the models all explain over 86% of the variance of the dependent variable, which makes the models suitable for forecasting purposes. The results from testing the forecasting abilities of the models will be presented in a later chapter.

The interpretation of the speed of adjustment coefficient is as presented above, and a similar pattern of lower speed of adjustment coefficients in the post financial crisis sample is observed. In short, the ECM-COC versions support the indications of the regular ECMs and the implied deterioration of market efficiency from the pre to the post financial crisis sample is confirmed once more.

4.6.2 ARIMA models

Using the Box-Jenkins three-step methodology for estimating univariate ARIMA models, the ACF and the PACF of the time series variables are inspected in the first step. See appendix A.6. Because the ACF gradually dies out lag by lag and the PACF cuts off after a number of lags for both the spot and futures price series in the pre financial crisis sample, the MA component is set to zero in all estimated ARIMA models (Pham, 2013). The persistent ACF indicates that the series should be differenced in order to become stationary, just as the ADF tests for unit roots concluded. Regarding the number of AR terms, different versions are tested in combination with different orders of differencing. Finally, the minimum SIC determines which is the best-fitted univariate ARIMA model with respect to explaining current spot and futures prices.

For the univariate spot model and the univariate futures model for the pre financial crisis sample, the ARIMA(1,0,0) is best fitted to the data measured by the lowest SIC. Estimated coefficients for these ARIMA models are reported in table 13, and alternative models are found in appendix A.7. The use of Newey-West robust standard errors ensures that there is no bias caused by autocorrelation and heteroskedasticity in the estimates.

Table 13: *Best fitted estimated ARIMA models, pre financial crisis sample*

Dependent variable	ARIMA	$\bar{R}^2$	Constant	AR(1)	SIC
Log(Spot$_t$)	(1,0,0)	0.995	9.517**	0.997**	-6.118521
Log(Futures$_t$)	(1,0,0)	0.995	9.525**	0.997*	-5.925810

***1 % significant, *10 % significant*

As in the pre financial crisis sample, the ACFs gradually die out for both spot and futures prices in the post financial crisis sample. See appendix A.6. Regarding the PACF of the spot and futures prices in this sample period, it cuts off rapidly after a number of lags for both time series variables. Therefore, the procedure employed to estimate the best fitted ARIMA model in the post financial crisis sample is as before, and the MA component is set to zero. The ARIMA(1,1,0) model is determined to be the best fitted model when it comes to explaining the spot price. The ARIMA(1,0,0) model is best fitted to explain the futures price. These models are reported in table 14, and alternative models are found in appendix A.7.

Table 14: *Best fitted estimated ARIMA models, post financial crisis sample*

Dependent variable	ARIMA	$\bar{R}^2$	Constant	AR(1)	SIC
Log(Spot$_t$)	(1,1,0)	0.015	-2.33E-05	0.124**	-5.512769
Log(Futures$_t$)	(1,0,0)	0.992	9.299**	0.997**	-5.236529

***1 % significant, *5 % significant*

Table 13 and 14 show that the explanatory power of the univariate spot and futures ARIMA(1,0,0) models chosen for the pre financial crisis sample is quite high. The same goes for the chosen univariate futures ARIMA(1,0,0) model chosen for the post financial crisis sample. Comparing these, the explanatory power of the post financial crisis spot ARIMA (1,1,0) model is very low. These findings are as expected since the ARIMA(1,0,0) is a relatively simple model, while the ARIMA(1,1,0) model has a less intuitive interpretation. Nevertheless, in the case where the ARIMA(1,1,0) is chosen, this has been necessary in order to get a significant model.

4.7 Granger causality

The analysis above implies that there exists a strong relationship between the spot and the futures price in both sample periods. To further investigate this relationship, a Granger causality test is performed in the following section.

As mentioned previously, the bivariate Granger causality test uses F statistics based on Wald's test. The null hypothesis is that the spot price does not Granger-cause the futures price in a regression where the futures price is the dependent variable, or alternatively that the futures price does not Granger-cause the spot price in a regression where the spot price is the dependent variable. The causality can be unidirectional, bi-directional or non-existent. Because the sample periods consist of daily data, five lags were applied in order for the test to match the number of trading days in one week. The results of the tests are shown in table 15.

***Table 15**: Pairwise Granger causality tests*

Pairwise Granger causality tests	Pre financial crisis period		Post financial crisis period	
5 lags included	Causality	F statistics	Causality	F statistics
H_0: Spot does not cause futures	H_0: Rejected	8.42929**	H_0: Rejected	13.4368**
H_0: Futures does not cause spot	H_0: Rejected	4.20431**	H_0: Rejected	2.75349*
Conclusion	**Bi-directional**		**Bi-directional**	

** 5% significant*
*** 1% significant*

The tests show that there is a significant bi-directional relationship in both sample periods. However, the rejection of the null hypothesis, that the spot prices does not Granger-cause the futures price, is more significant in the post financial crisis period than in the pre financial crisis period. On the other hand, the rejection of the null hypothesis, that the futures price does not Granger-cause the spot price, has a lower F statistic in the post financial crisis period than in the first period. Thus, the rejection is slightly less significant in the second period. The bi-directional relationship suggests that there is a relation between the two prices, where the prices cause each other. This is in line with previous findings in the analysis, where the relationship between the spot and the futures price in the pre financial crisis period seemed to have strong contemporaneous traits. These characteristics were still significant in the post financial period, which also supports the conclusion drawn from the pairwise Granger causality test.

A causality from the futures market to the spot market is somewhat expected considering previous studies of the spot and futures market for the Nikkei 225 index (i.e. Tse (1995) and Hiraki et al. (1995)). A lead from the spot market to the futures market has been observed previously as well, by Sinha (1991). Thus, the bi-directional relationship in the sample periods suggest a new turn for the markets where information travels from both markets to the other as information feedback. Following the arguments of Tang et al. (1992), this feedback effect might indicate that the markets are popular for institutional investors in terms of hedging. These investors trade futures contracts with the purpose of hedging their investments in the stock market. As they participate in both markets, there will be a flow of information going both ways. As mentioned previously, the rejection of the hypothesis that the futures price does not cause the spot price is less significant in the post financial crisis period. This means that the lead from futures to spot prices is weaker in this period. Again, following Tang et al. (1992), this might be caused by less speculative trade in the futures market as this became more risky as the underlying index fell when the financial crisis hit the market.

4.8 Preliminary conclusion of analysis

As mentioned above, error correction models based on the theoretical cost-of-carry relationship between the spot and the futures price and univariate ARIMA models are estimated in addition to the traditional error correction models. The best-fitted model from each model category is chosen based on a SIC minimizing criterion and the alternative models serve the purpose of being a robustness test of the traditional error correction models.

In the pre financial crisis sample, the traditional error correction models chosen to describe the logarithmic return on the Nikkei 225 Stock Average Index and the return on the matching futures contracts validate the assumption that there is a long-term relationship between the variables. Furthermore, a short-term relation between the spot and the futures prices is suggested and the existence of a strong contemporaneous relation between changes in the spot and futures prices is confirmed. Regarding the post financial crisis sample, the chosen error correction models used to explain the logarithmic return on the Nikkei 225 Stock Average Index and the return on the futures contracts imply that the contemporaneous traits of the relationship between the prices remain. However, they also suggest that past prices have a price discovery role. The test for Granger causality uncovers a bi-directional relationship between the prices, underlining the simultaneous feedback effect between the variables.

The implications for market efficiency in regards to the Japanese market and more specifically the Nikkei 225 Stock Average Index spot and futures prices are many. In both sample periods it is possible to create models based on a co-integrating relationship that use past and present prices to predict future prices, as confirmed by a significant coefficient for the error correction term in the ECM and the ECM-COC models. Thus, implying that the market is not perfectly semi-strong efficient. Furthermore, in the post financial crisis sample where short-term relations also are significant, the indication of semi-strong market inefficiency is even stronger. The lower speed of adjustment coefficients for the post financial crisis period also indicate that the market has become less efficient than in the pre financial crisis period. As the post financial crisis sample period also includes the actual financial crisis, the results are interesting as they might describe some of the reactions of the markets to the global financial crisis. The implication of a less efficient market in the post financial crisis period suggests that the global recession and economic despair reached the Japanese market and changed the market dynamics. Speculative investments in the futures market became more risky as the Nikkei 225 index fell and these investors most likely reduced their exposure in the market, following the argumentation of Tang et al. (1992). Institutional investors who used the market for hedging remained as the hedging protected their exposure in the Nikkei 225 spot market. This might be the reason that the overall trading volume of the Nikkei 225 futures contract decreased after the financial crisis, as can be observed in figure 2.

Figure 2*: Daily trading volume for OSE traded Nikkei 225 futures contracts, 01.07.2004 through 30.06.2014. Pre and post financial crisis sample. Extracted from DATASTREAM.*

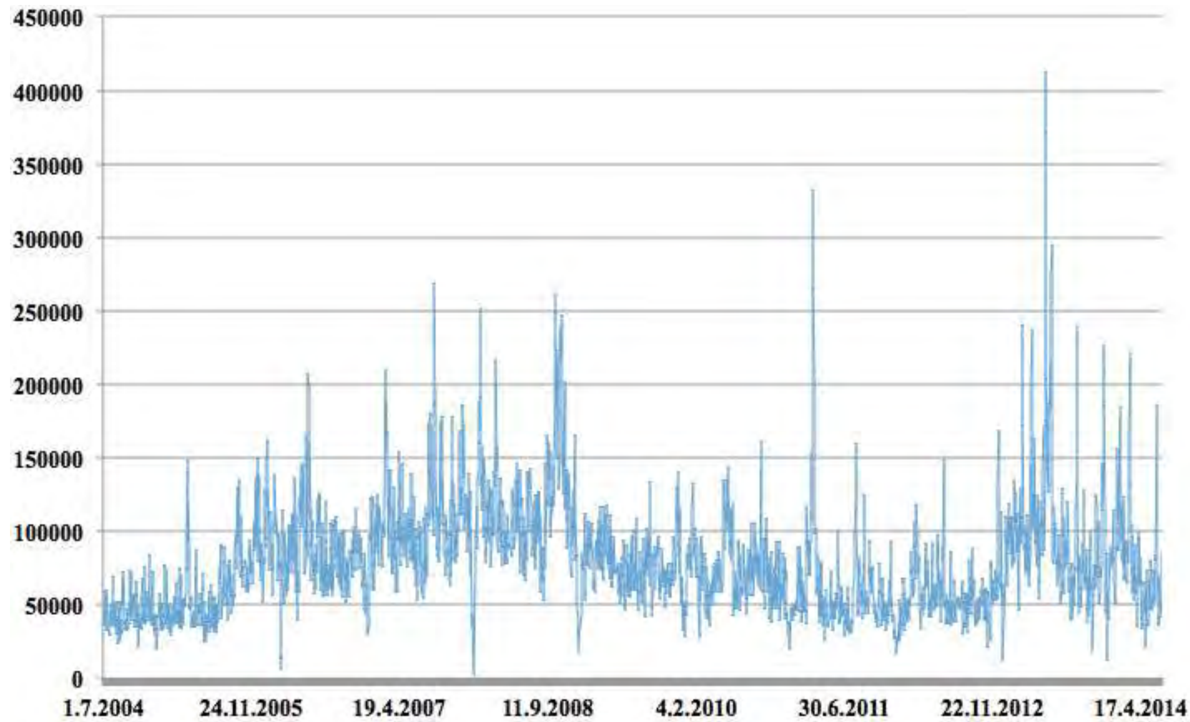


Table 16: *Mean number of daily OSE traded Nikkei 225 futures contracts*

	Pre financial crisis sample	Post financial crisis sample
Mean n^{o} daily traded futures contracts	84,775.76	76,680.07

Table 16 further describes the decrease in trading volume. The average daily trading volume of the Nikkei 225 futures contracts decreased by 9.55% from the pre financial crisis period to the post financial crisis period. Fewer trades reduce the ability of the market to reflect available information, and therefore reduce the efficiency of the market. This might help explain the differences that are observed in this study between the pre and the post financial crisis period.

However, in order to provide any decisive conclusion or implication on market efficiency, the models must be tested further and transaction costs must be taken into account. In the next section, the predictive power of the models is explored, different trading strategies based on the predictions of selected models are tested and transaction costs are considered.

4.9 Model comparison

The comparisons are done using different measures on how well the different models perform in an out-of-sample period. The best-fitted model from each approach is further investigated. The root mean squared error (RMSE) and mean average error (MAE) indicate how accurate each model estimates the current spot or futures price based on historical information. The lower these measures are, the better fitted the model is said to be. The percentage of correct prediction of direction of the change in the prices indicates how well the models utilize historical price information to predict the direction of a one step ahead price change.

4.9.1 The pre financial crisis sample

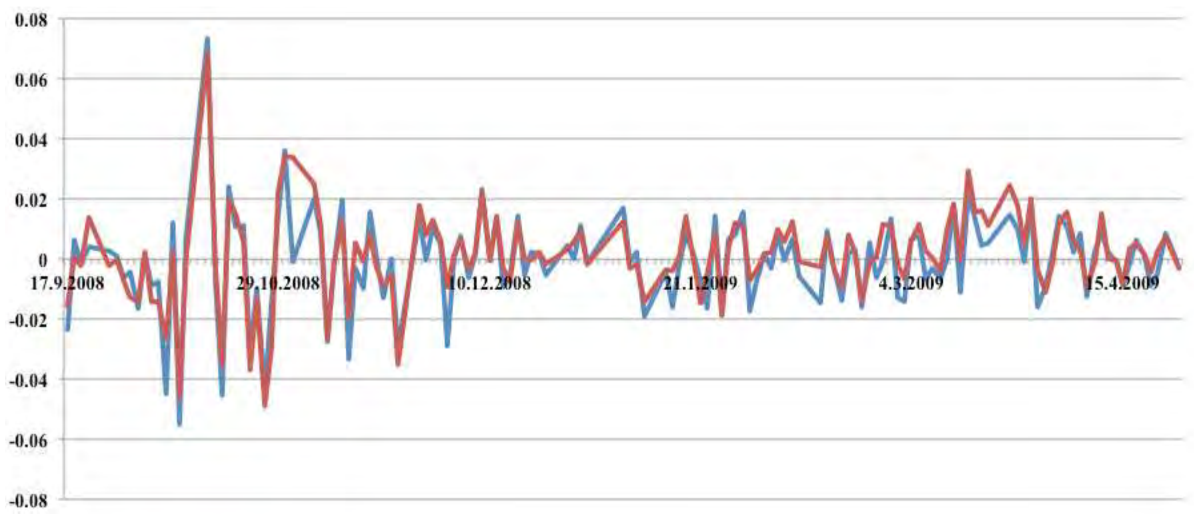
The models from the pre financial crisis period where the spot price is the dependent variable were tested on the first out-of-sample period, which is extracted from the post financial crisis sample. The out-of-sample period reaches from 16.09.2008 to 27.04.2009. The results of the out-of-sample predictions gave values for RMSE, MAE and percentage of correct prediction of direction as described in table 17:

***Table 17**: Results comparison pre financial crisis sample, predicting spot price*

Model comparison	RMSE	MAE	Correct prediction of direction (%)
ECM	0.01109	0.00972	62.8
ECM - COC	0.00613	0.00459	82
ARIMA (1, 0, 0)	0.01569	0.01183	56.6
Best model			**ECM - COC**

The error correction model based on the cost-of-carry equation (ECM-COC) has the lowest score for both RMSE and MAE. It also has the highest percentage of correct prediction of direction of movement with 82% correct direction predictions. Figure 3 illustrates how well the best ECM-COC forecasts the direction of successive spot price movements compared to the actual price changes during the pre financial crisis sample period. From the graph it becomes evident that the shifts forecasted by the model match those of the actual spot price relatively well, and that the changes in the price forecasts also are close to the level of the actual spot price changes.

***Figure 3**: Comparison of actual percentage spot price changes (blue line) and forecasts of spot price changes (red line) produced by the ECM-COC*



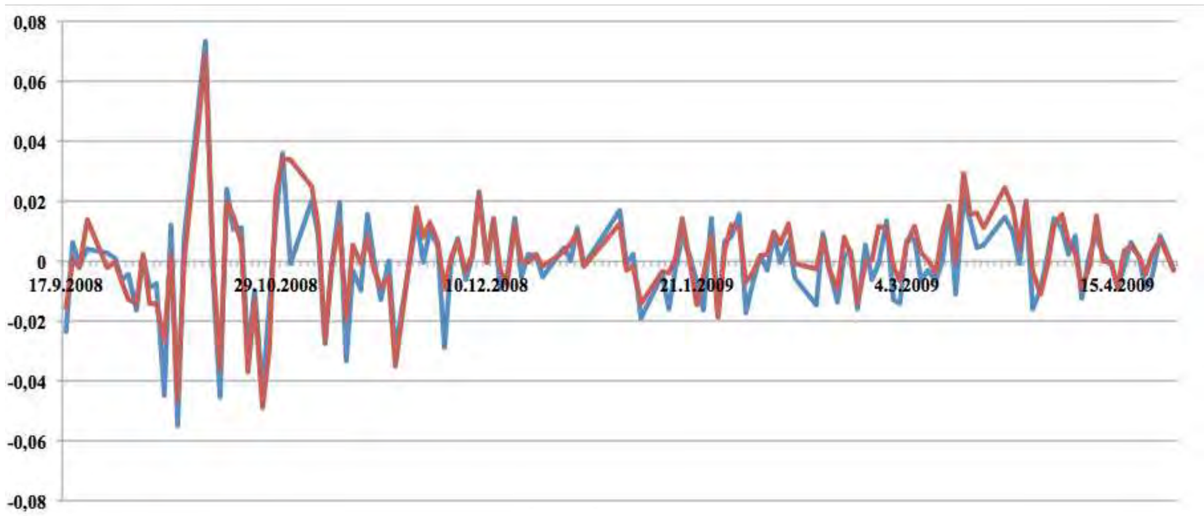
The regular ECM is the next best model with 62.8% prediction of correct direction, which is somewhat expected as the ECM and the ECM-COC models are quite similar except for the calculation of the error correction term. The ARIMA model with one lag performs the worst with 56.6% correct predictions of direction.

The results from the models where the futures price is the dependent variable have similar results as those concerned with predicting the spot price, as seen in table 18 below. The only notable remark on this comparison is that the ARIMA model performs even worse by only predicting the correct direction of movement of the price in 43.9% of the cases. When using the ECM-COC to forecast successive futures price changes, the forecasts produced in the pre financial crisis sample are again relatively close to the actual futures price movements both in magnitude and level. This is illustrated in figure 4.

Table 18*: Results comparison pre financial crisis sample, predicting futures price*

Model comparison	RMSE	MAE	Correct prediction of direction (%)
ECM	0.00899	0.00752	72.7
ECM – COC	0.00665	0.00509	81.3
ARIMA (1, 0, 0)	0.02437	0.01693	43.9
Best model			**ECM – COC**

Figure 4*: Comparison of actual percentage futures price changes (blue line) and forecasts of futures price changes (red line) produced by the ECM-COC*



4.9.2 The post financial crisis sample

The models from the post financial crisis period where the spot price is the dependent variable were tested on the second out-of-sample period, reaching from 01.07.2014 to 10.02.2015. Table 19 gives the results from the comparison.

***Table 19**: Results comparison post financial crisis sample, predicting spot price*

Model comparison	RMSE	MAE	Correct prediction of direction (%)
ECM	0.01816	0.01807	57.9
ECM – COC	0.00893	0.00699	22.3
ARIMA (1, 1, 0)	0.00450	0.00363	49.2
Best model			**ECM**

In table 19, it is clear that the results are quite different from the results from the pre financial crisis sample. None of the models are able to get as high percentages of correct prediction of direction as in the first sample. The ECM-COC model is now the worst in terms of prediction, with only 22.3% correct predictions of direction, and the second worst in terms of RMSE and MAE. The ARIMA (1,1,0) model has the lowest RMSE and MAE score, but the ECM has the highest percentage of correct prediction of direction with 57.9%. In spite of not having the lowest RMSE and MAE, the ECM is chosen as the best model because of the predictive ability of the model compared to the alternative models. According to Leitch and Tanner (1991) models that can accurately predict the sign of future returns usually are more profitable models. Because the best model will be tested further by applying different trading strategies the profitability of the model is highly relevant.

Figure 5 illustrates how well, in terms of correct prediction of the direction of successive price changes, the ECM produces spot price forecasts. Clearly, the ECM gives a relatively poor prediction of the level of each successive spot price change. However, as the purpose is only to produce forecasts of the direction of price change, the ECM proves to be superior. The model forecasts price movements that follow a pattern remarkably similar to the path of the actual development in the spot price during the post financial crisis sample.

Figure 5: Comparison of actual percentage spot price changes (blue line) and forecasts of spot price changes (red line) produced by the ECM

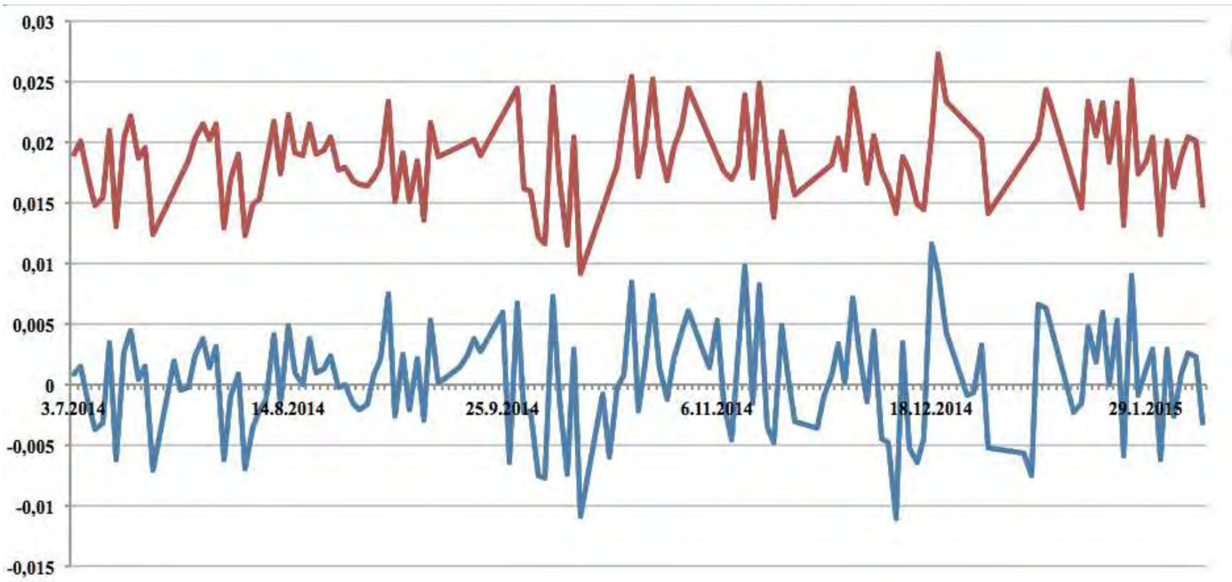
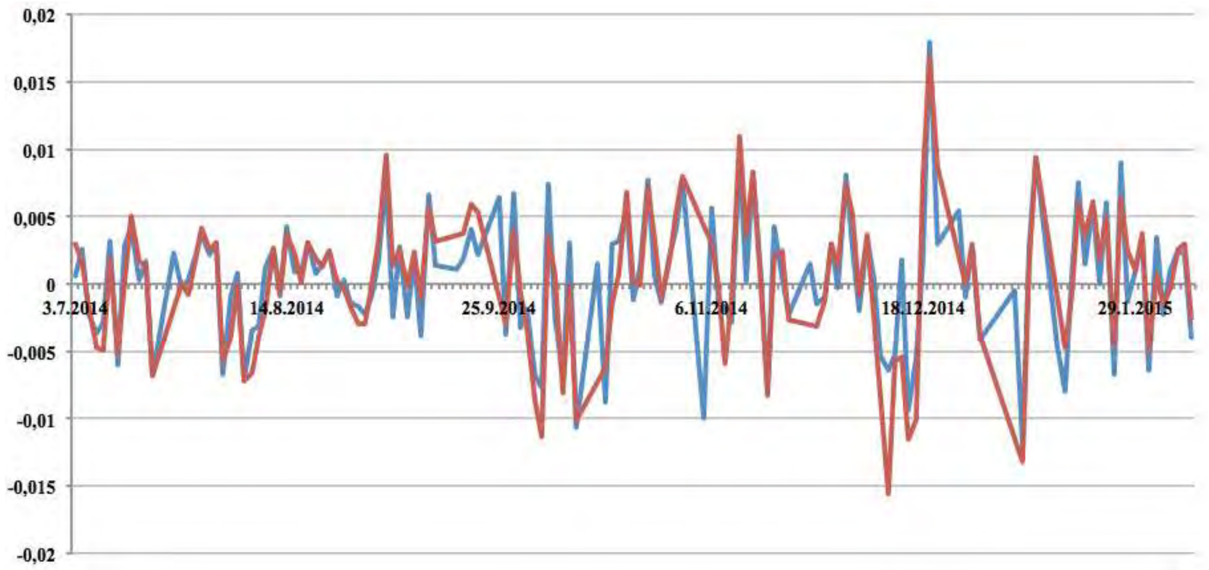


Table 20: Results comparison post financial crisis sample, predicting futures price

Model comparison	RMSE	MAE	Correct prediction of direction (%)
ECM	0.00219	0.00159	86.2
ECM – COC	0.00219	0.00160	85.4
ARIMA (1, 0, 0)	0.00673	0.00538	50
Best model			**ECM**

The models predicting the futures price in the post financial crisis sample provide much better predictions than the models predicting the spot price. The best model, the ECM, has a percentage of correct prediction of direction of 86.2%, with the ECM-COC model as the runner-up with 85.4% correct predictions of direction. All of the models have fairly low and similar values for RMSE and MAE. The ECM model has the lowest scores, which makes it the best overall model. The performance of the ECM is illustrated by a graph in figure 6. Here, it also becomes evident that the model produces forecasts of successive futures price changes that, in addition to being in the correct direction, are approximately of the same magnitude and at the same level as the actual price change.

Figure 6: *Comparison of actual percentage futures price changes (blue line) and forecasts of futures price changes (red line) produced by the ECM*



As the previous results from the analysis show, stronger signals in favour of the market efficiency hypothesis are found in the pre financial crisis sample than in the post financial crisis sample. Therefore, it may be reasonable to expect that the ECM-COC model based on the theoretical cost-of-carry relationship between spot and futures prices performs the best among the models in the pre financial crisis sample. In the post financial crisis sample the support for market efficiency is less evident, and the traditional ECM proves to be the best model. As the cost-of-carry theory is based on strong assumptions about the market, it is natural that the ECM model is the best model in the post financial crisis period.

The absence of the variables included in the cost-of-carry component (RDT) used in estimating the ECM-COC models, may be a reason why the traditional ECM performs better than the ECM-COC in the post financial crisis sample. These variables include both a risk free interest rate and the dividend yield of the stock index. Both parameters are expected to change in financial markets during a financial crisis. During national and global financial distress, it is expected that the risk free interest rate will increase to account for the increased investment risk (Vernimmen et al., 2005). This, alongside an assumption of decreasing dividend yield during uncertain economic times, might have an increasing effect on the cost-of-carry component that may bias the ECM-COC model, and therefore make the traditional ECM model a better fitted model.

4.10 Trading

4.10.1 Trading

The models from the post financial crisis sample with the highest percentage of correct prediction of direction are further investigated by exploring different trading strategies. The models from the pre financial crisis sample are not investigated as profitable trading strategies only have practical implications for future periods.

The traditional error correction model, as described in the previous section, proved to be the best model for forecasting both the spot price and the futures price in the post financial crisis period. The profitability of trading according to strategic trading rules is explored in the following section. Assuming that the investor starts off with ¥100 000 on the first trading day, the question is whether strategic trading will be able to outperform a passive benchmark trading strategy.

The passive benchmark strategy involves investing the whole amount of ¥100 000 in the Nikkei 225 Stock Average Index at the first out-of-sample trading day at a price equal to the spot price on that day. Alternatively, the ¥100 000 is invested in OSE traded Nikkei 225 futures contracts on the first trading day. Either way, the position is then held throughout the out-of-sample period before the position is closed and profits are calculated. The strategic trading rules that are applied are a filter strategy that says to buy when the predicted return is higher than a given filter threshold, and a liquid trading strategy that instructs the investor to buy whenever the predicted return for the next period is positive. For both strategies, the rule is to sell out and close the position whenever the trading rule does not give instructions to buy. If the investor closes his position, he is able to earn the risk free rate of return. The profitability of an alternative approach by placing the investment amount of ¥100 000 in a risk free investment throughout the trading period is also included for comparison.

The impact on profitability of a range of transaction costs is also tested. Following Sutcliffe (2006), a 1.5 % range of transaction costs is tested when trading on the spot price of Nikkei 225 Stock Average Index and when trading futures contracts based on the index. Brooks et al. (2001) used a transaction cost for the spot price of 1.70% and a transaction cost of 0.116% for the futures price in their study of the FTSE 100. These transaction costs have been used as a starting point to determine the range of transaction costs. Furthermore, the transaction costs for trading futures contracts are assumed to be lower than those for trading the spot index (Fleming et al., 1996). The range is set slightly higher in both cases to perform a strict test of abnormal returns. The transaction costs are computed from the total of each transaction amount, i.e. ¥100 000 adding any cumulated profits or losses. The risk free interest rate is assumed to be the daily averaged Tokyo Interbank three month offered rate during the out-of-sample period.

4.10.2 Trading with the Nikkei 225 Stock Average Index (spot)

When trading according to the filter strategy, the computed spot threshold filter becomes 0.00108. Therefore, each time the error correction model forecasts an increase in the logarithmic spot price that is greater than the filter threshold, which is established by the average predicted positive return, the trading rule says to buy. Otherwise, the position is closed. Regarding the liquid trading rule, it takes a particularly short-term horizon into account when deciding what position an investor should take. Therefore, this strategy becomes the trading rule with the highest number of transactions during the trading period.

As can be seen from table 21, the passive buy-and-hold strategy yields higher returns both when excluding and including a range of transaction costs. The large number of transactions dictated by the liquid trading strategy causes this trading rule to create the biggest loss for the spot-trading investor in this trading period.

Table 21*: Results trading strategies, predicting spot, post financial crisis sample*

Trading strategies	Nº of transactions	Gross return (%)	Net return (%) after transaction cost equal to:			
			1.5%	2.0%	2.5%	3.0%
Passive buy-and-hold	1	14	11.9	11.4	10.8	10.2
Liquid trading	37	- 23.6	- 55.7	- 63.1	- 69.3	- 74.5
Filter	17	- 13.9	- 32.4	- 37.6	- 42.5	- 47.1
Risk free investment	0	0.3	0.3	0.3	0.3	0.3
Best strategy		Passive buy-and-hold	Passive buy-and-hold	Passive buy-and-hold	Passive buy-and-hold	Passive buy-and-hold

4.10.3 Trading with futures contracts based on the Nikkei 225 Stock Average Index

When using the best-fitted traditional error correction model to predict the futures price, the threshold of the filter strategy is computed to be 0.00043. As implied by the many transactions decided by the liquid trading strategy, which suggests buying whenever the price change is positive, the out-of-sample trading period reflects a bullish market. The number of transactions can be seen in table 22. Thus, when using such a low filter threshold, having a large number of transactions is as expected.

From table 22, it seems that both the liquid and the filter trading strategy yield substantial profits before transaction costs are taken into account. However, since both strategies result in a large number of transactions during the trading period, the net return becomes negative as the transaction costs increases. The liquid trading strategy and the filter strategy may be profitable only at very low transaction costs. Thus, when model forecasts are used to trade futures contracts, the passive buy-and-hold strategy outperforms strategic trading when transaction costs above 0.82% - 0.83% are considered.

Table 22: *Results trading strategies, predicting futures, post financial crisis sample*

Trading strategies	Nº of transactions	Gross return (%)	Net return (%) after transaction cost equal to:			
			0.5%	1.0%	1.5%	2.0%
Passive buy-and-hold	1	10.8	10.4	9.8	9.3	8.7
Liquid trading	54	72.3	31.5	0.2	- 23.8	- 42.1
Filter	58	62.4	23.2	- 8.0	- 31.4	- 49.0
Risk free investment	0	0.3	0.3	0.3	0.3	0.3
Best strategy		Liquid trading	Liquid trading	Passive buy-and-hold	Passive buy-and-hold	Passive buy-and-hold

4.10.4 Conclusions trading

By using a traditional error correction model to forecast spot and futures prices, it is shown that strategic trading can generate substantial profits in the case of trading based on futures prices and no transaction costs. A liquid trading strategy that trades on futures prices yields the highest gross return. On the other hand, when using the model forecasts to trade the index at the spot price, no strategic trading rule is able to outperform the passive benchmark even in the absence of transaction costs.

When including transaction costs and calculating the net return from each of the trading strategies and the passive benchmark, none of the active trading strategies are able to outperform the passive buy-and-hold strategy when the trading is based on the spot price. When trading futures contracts, strategic trading may be profitable, but only at very low transaction costs, lower than 0.82% - 0.83%. These results imply that strategic trading is quite ineffective for an investor since all strategies seem to make substantial losses over the trading period when more realistic levels of transaction costs are taken into account.

5 Conclusion

This thesis was centred on the following thesis questions:

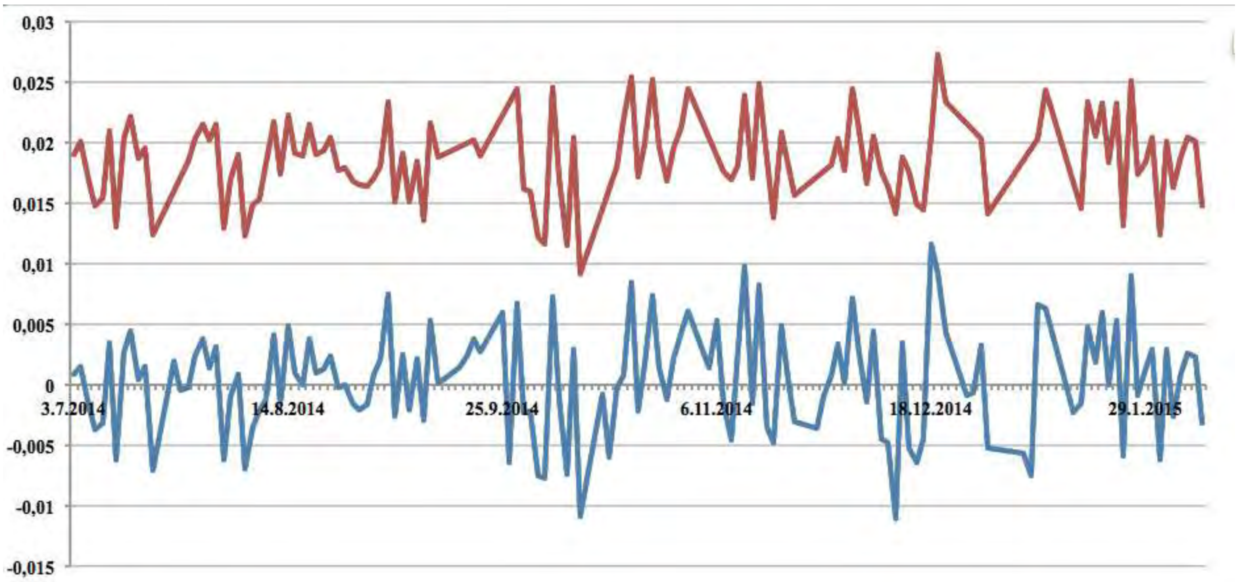
- *Is there a lead-lag relationship between the spot and the futures price based on the Nikkei 225 Stock Average Index?*
- *How has the relationship between the spot and the futures price changed after the financial crisis of 2008?*
- *If there is in fact a lead-lag relationship, does it allow for abnormal profits when taking transaction costs into account?*
- *Which implications do the findings have for market efficiency in the Japanese market as reflected by the Nikkei 225 Stock Average Index?*

Both prices were found to be stationary at first difference for both periods. Results from the Engle and Granger two-step test for co-integration both based on the regular co-integration equation and the cost-of-carry equation suggest that the spot and the futures price were co-integrated both before and after the financial crisis of 2008. This co-integrating relationship is described further by various error correction models and seems to be strongly contemporaneous because of strong links between the current price changes. Furthermore, there is clearly a long-term relationship between the prices as the coefficient of the error correction term is significant and large for all models. However, in the post financial crisis sample more short-term relations have become significant as including more lagged variables of the dependent and the independent variable seems to improve the model. Therefore, in the post financial crisis period, past prices seem to play a more important role in determining future prices, implying a slight deterioration of market efficiency.

The speed of adjustment coefficient is also lower in the post financial crisis period, implying that corrections of deviations from the long-run equilibrium relationship have slowed down from the pre financial crisis period. Considering that the financial crisis was included in the post financial crisis sample, the signs of a less efficient market are interesting as they can be linked to the financial crisis. Granger causality tests confirm that there is a lead-lag relationship between the prices with a feedback effect, namely a bi-directional relationship. This causality is consistent throughout both periods, and might be explained by a large participation of institutional investors that use the futures market to hedge investments in the spot market, as argued by Tang et al. (1992). A lower trading volume in the post financial crisis period, as documented by numbers from DATASTREAM, might also explain the less efficient flow of information. The lower trading volume might be caused by less speculation in the markets as a consequence of the more risky post financial crisis environment.

An ARIMA model was added as a robustness check when the best-fitted ECM models were compared in terms of RMSE, MAE and percentage of correct prediction of direction of the movement of the price. The above-mentioned measures were based on an out-of-sample period. The ECM model based on the cost-of-carry relationship had the highest percentage of correct predictions for both the spot and the futures price in the pre financial crisis sample, with 82% correct predictions for the spot price and 81.3% correct predictions for the futures price. In the post financial crisis sample, the regular ECM had the highest percentage of correct predictions. For the spot price the model had 57.9% correct predictions, and for the futures price the model had 86.2% correct predictions. A comparison of predicted prices and actual prices from the post financial crisis period can be observed in the following figures 7 and 8:

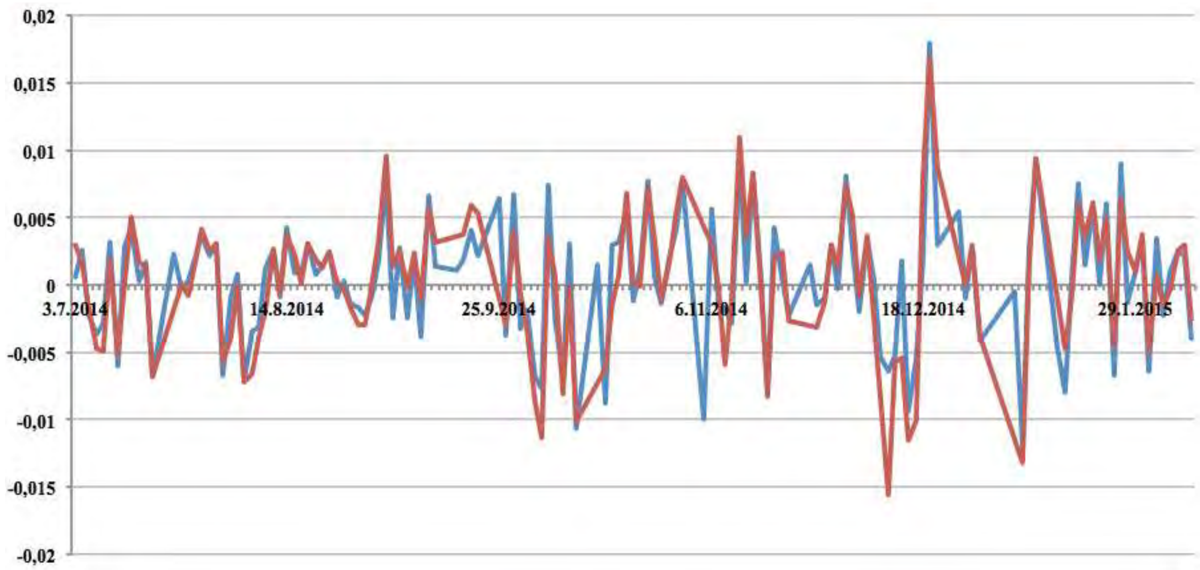
***Figure** 7: Comparison of actual percentage spot price changes (blue line) and forecasts of spot price changes (red line) produced by the ECM*



Although the forecasts seem to perform poorly in terms of predicting the level of the spot price, the predicted direction of change is quite similar to the actual price change as shown in

Figure 7. Figure 8 shows how the chosen ECM model is able to predict a pattern of changes in the futures price that is quite similar both in terms of direction and level to that of the actual change in futures prices.

***Figure 8**: Comparison of actual percentage futures price changes (blue line) and forecasts of futures price changes (red line) produced by the ECM*



Finally, different trading strategies were tested based on the models with the highest percentage of correct prediction of direction of change. The chosen strategies were a passive buy-and-hold strategy, a liquid trading strategy, a filter strategy and holding a risk free investment. This step was essential in order to answer the main research question:

Is it possible to earn an abnormal return in the Japanese market, as represented by the Nikkei 225 Stock Average Index, by trading in the spot and futures market based on an error correction model?

None of the strategies outperformed the passive buy-and-hold strategy when transaction costs were included in the case of predicting spot prices. A specter of 1.5% - 3.0% was tested. The opposite direction, where the aim was to predict futures prices, provided somewhat different results. Since transaction costs generally are assumed to be lower for trading in the futures market than for trading in the spot market (Fleming et al, 1996), the specter was set to 0.5% - 2.0%. Both the liquid trading strategy and the filter strategy proved more profitable than the passive buy-and-hold strategy for transaction costs of 0.5% with a net return of 31.5% and 23.2%. However, as the transaction costs increased to 1% the passive buy-and-hold strategy once again proved to be the most profitable with the changing point at transaction costs of 0.82% - 0.83%. Regarding the research question, the conclusion is for the most part that it is not possible to earn an abnormal return when transaction costs are deducted. However, if the transaction costs are low enough it is possible to gain an abnormal profit from active trading strategies used to trade futures contracts.

Finally, in regards to implications for market efficiency in the Japanese market, it seems to be efficient in the sense that arbitrage is not possible unless transaction costs are set very low or excluded altogether. However, there seems to have been a change in the way the market absorbs new information after the financial crisis, and past prices have more pronounced roles in determining future prices. This change might be the result of less frequent trade and less speculation in the futures market because of greater perceived risk (as suggested by Tang et al. 1992). There is a bi-directional relationship between the spot and the futures market for the Nikkei 225 Stock Average Index, implying that information reaches both markets at approximately the same time and that there is a feedback of information from one market to the other. High involvement from investors that use the futures market to hedge their investments in the spot market might explain this information flow. Thus, there seems to be no clear choice of go-to market for new information as both markets will adjust to the information in a similar way.

6 Criticism and suggestions for further research

In this thesis, different methodological approaches have been used to examine if there is any informational inefficiency in the Japanese financial markets, which would allow for the existence of a lead-lag relationship between spot and futures prices. Based on the findings in the analysis and based on macro economical circumstances, different suggestions for further research and improvements on the topic of this thesis are presented in the following section.

The in-sample time series that are used contain information on spot prices for the Nikkei 225 Stock Average Index and the OSE traded Nikkei 225 futures contracts from July 2004 to June 2014. A graphical inspection of the time series evolvement makes it evident that dividing the sample into a pre and post financial crisis sample is natural as the crisis turned the pre crisis bull market in Japan into a post crisis bear market. Towards the end of the post financial crisis sample, which also includes the time span of the actual crisis, the Japanese financial market is yet again showing signs of being more bullish. However, because an out-of-sample period was needed to perform trading based on the in-sample estimated models, creating a third sub sample period with the natural beginning in mid 2013 was not ideal. This was due to the necessity of a sufficient number of valid observations in each sub sample. Therefore, for further future research, the use of a third sub sample period that takes account for the natural change of regime in the post financial crisis period is recommended.

Furthermore, the choice to include the actual financial crisis in the in-sample period used to estimate models might have affected the applicability of the results. Since the change of regime in mid 2013 is not taken into account, the out-of-sample forecasting is made on a sample period with characteristics that are somewhat different from the sample period used in the model estimation. For example, this may have contributed to the weak predictive abilities of the ARIMA models. However, because of the practical reasons discussed above, a small number of well-predicting models were preferred rather than having very few observations in a potential third in-sample period used for estimation. The comments on the choice of sample periods may also apply in an argument suggesting that the forecasting ability of the models estimated in the pre financial crisis sample may be biased. They might be biased because they were tested in an out-of-sample period that took place after an evident structural break imposed by the outburst of the financial crisis in mid 2008.

7 Bibliography

<u>Articles:</u>

Abhyankar, A. H. 1995. Return and volatility dynamics in the FT-SE 100 stock index and stock index futures markets. *Journal of Futures Markets,* 15**,** 457-488.

Antoniou, A. & Holmes, P. 1996. Futures market efficiency, the unbiasedness hypothesis and variance-bounds tests: The case if the FTSE-100 futures contract*. *Bulletin of Economic Research,* 48**,** 115-128.

Bachelier, L. 1900. Théorie de la spéculation, *Annales scientifiques de l'École Normale Supérieure*, 3, 21-86.

Brooks, C., Garrett, I. & Hinich, M. J. 1999. An alternative approach to investigating lead-lag relationships between stock and stock index futures markets. *Applied Financial Economics,* 9**,** 605-613.

Brooks, C., Rew, A. G. & Ritson, S. 2001. A trading strategy based on the lead–lag relationship between the spot index and futures contract for the FTSE 100. *International Journal of Forecasting,* 17**,** 31-44.

Campbell, J. Y. & Perron, P. 1991. Pitfalls and opportunities: what macroeconomists should know about unit roots. *NBER Macroeconomics Annual 1991, Volume 6.* MIT press.

Chan, K. 1992. A Further Analysis of the Lead-Lag Relationship between the Cash Market and Stock Index Futures Market. *Review of Financial Studies,* 5**,** 123-152.

Chiang, M.-H. & Wang, J.-Y. 2008. Regime switching cointegration tests for the Asian stock index futures: evidence for MSCI Taiwan, Nikkei 225, Hong Kong Hang-Seng, and SGX Straits Times indices. *Applied Economics,* 40**,** 285-293.

Chung, Y. P., Kang, J.-K. & Rhee, S. G. 1994. Index futures arbitrage in Japan. *University of California, Riverside, working paper*.

Clare, A. & Miffre, J. 1995. A note on forecasting the CAC 40 and DAX stock index futures. *Applied Economics Letters,* 2**,** 327-330.

Cornell, B. & French, K. R. 1983. Taxes and the pricing of stock index futures. *The Journal of Finance,* 38**,** 675-694.

Cowles, A. 1933. Stock market forecasting. *Econometrica, Journal of the Econometric Society***,** 1, 206-214.

Dickey, D. A. & Fuller, W. A. 1979. Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American statistical association,* 74**,** 427-431.

Engle, R. F. & Granger, C. W. J. 1987. Co-Integration and Error Correction: Representation, Estimation, and Testing. *Econometrica,* 55**,** 251-276.

Fama, E. F. 1965. The behavior of stock-market prices. *Journal of business***,** 38, 34-105.

Fama, E. F. 1970. Efficient Capital Markets: A Review of Theory and Empirical Work. *Journal of Finance,* 25**,** 383-417.

Fama, E. F. 1991. Efficient capital markets: II. *The journal of finance,* 46**,** 1575-1617.

Fama, E. F., Fisher, L., Jensen, M. C. & Roll, R. 1969. The adjustment of stock prices to new information. *International economic review,* 10**,** 1-21.

Fleming, J., Ostdiek, B. & Whaley, R. E. 1996. Trading costs and the relative rates of price discovery in stock, futures, and option markets. *Journal of Futures Markets,* 16**,** 353-387.

Floros, C. & Vougas, D. 2008. Lead-lag relationship between futures and spot markets in Greece: 1999-2001. *International Research Journal of Finance and Economics***,** 34, 168-174.

Ghosh, A. 1993. Cointegration and Error Correction Models: Intertemporal Causality between Index and Futures Prices. *Journal of Futures Markets,* 13**,** 193-198.

Granger, C. W. 1969. Investigating causal relations by econometric models and cross-spectral methods. *Econometrica: Journal of the Econometric Society***,** 37, 424-438.

Grossman, S. J. & Stiglitz, J. E. 1980. On the impossibility of informationally efficient markets. *The American economic review***,** 70, 393-408.

Herbst, A. F., Mccormack, J. P. & West, E. N. 1987. Investigation of a lead-lag relationship between spot stock indices and their futures contracts. *Journal of Futures Markets,* 7**,** 373-381.

Hiraki, T., Maberly, E. D. & Takezawa, N. 1995. The information content of end-of-the-day index futures returns: International evidence from the Osaka Nikkei 225 futures contract. *Journal of Banking & Finance,* 19**,** 921-936.

Iihara, Y. & Kato, K. 1996. Intraday return dynamics between the cash and the futures markets in Japan. *Journal of Futures Markets,* 16**,** 147-162.

Jensen, M. C. 1968. The performance of mutual funds in the period 1945–1964. *The Journal of finance,* 23**,** 389-416.

Jensen, M. C. 1978. Some anomalous evidence regarding market efficiency. *Journal of financial economics,* 6**,** 95-101.

Johansen, S. 1988. Statistical analysis of cointegration vectors. *Journal of economic dynamics and control,* 12**,** 231-254.

Judge, A. & Reancharoen, T. 2014. An empirical examination of the lead–lag relationship between spot and futures markets: Evidence from Thailand. *Pacific-Basin Finance Journal,* 29**,** 335-358.

Kawaller, I. G., Koch, P. D. & Koch, T. W. 1987. The temporal price relationship between S&P 500 futures and the S&P 500 index. *The Journal of Finance,* 42**,** 1309-1329.

Kendall, M. G. & Hill, A. B. 1953. The analysis of economic time-series-part i: Prices. *Journal of the Royal Statistical Society. Series A (General),* 116**,** 11-34.

Leitch, G. & Tanner, J. E. 1991. Economic forecast evaluation: profits versus the conventional error measures. *The American Economic Review***,** 81, 580-590.

Lim, K. G. 1992. Arbitrage and price behavior of the Nikkei stock index futures. *Journal of Futures Markets,* 12**,** 151-161.

Mackinnon, J. G. 2010. Critical values for cointegration tests. Queen's Economics Department Working Paper.

Niederhoffer, V. & Osborne, M. F. M. 1966. Market making and reversal on the stock exchange. *Journal of the American Statistical Association,* 61**,** 897-916.

Osborne, M. F. 1959. Brownian motion in the stock market. *Operations research,* 7**,** 145-173.

Patra, G. C. & Mohapatra, S. R. 2014. Role of Futures in Price Discovery Process in Indian Stock Market. *Vilakshan: The XIMB Journal of Management,* 11**,** 1-18.

Phillips, P. C. & Perron, P. 1988. Testing for a unit root in time series regression. *Biometrika,* 75**,** 335-346.

Pizzi, M. A., Economopoulos, A. J. & O'neill, H. M. 1998. An examination of the relationship between stock index cash and futures markets: A cointegration approach. *Journal of Futures Markets,* 18**,** 297-305.

Roberts, H. V. 1959. Stock-Market “Patterns” And Financial Analysis: Methodological Suggestions. *The Journal of Finance,* 14**,** 1-10.

Schermelleh-Engel, K., Moosbrugger, H. & Müller, H. 2003. Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of psychological research online,* 8**,** 23-74.

Sinha, T. 1991. Relation Between Spot and Futures An Analysis of Nikkei Index and Nikkei Futures During the October 1987 Crash. *Actuarial Approach for Financial Risks,* 3**,** 441-51.

Stoll, H. R. & Whaley, R. E. 1990. The dynamics of stock index and stock index futures returns. *Journal of Financial and Quantitative Analysis,* 25**,** 441-468.

Tang, G. Y., Mak, S. & Choi, D. F. 1992. The causal relationship between stock index futures and cash index prices in Hong Kong. *Applied Financial Economics,* 2**,** 187-190.

Tse, Y. K. 1995. Lead-lag relationship between spot index and futures price of the nikkei stock average. *Journal of Forecasting,* 14**,** 553-563.

Turkington, J. & Walsh, D. 1999. Price discovery and causality in the Australian share price index futures market. *Australian Journal of Management,* 24**,** 97-113.

Wahab, M. & Lashgari, M. 1993. Price dynamics and error correction in stock index and stock index futures markets: A cointegration approach. *Journal of Futures Markets,* 13**,** 711-742.

Walter, C. 2003. The Efficient Market Hypothesis, the Gaussian Assumption, and the Investment Management Industry.

Yang, J., Yang, Z. & Zhou, Y. 2012. Intraday price discovery and volatility transmission in stock index and stock index futures markets: Evidence from China. *Journal of Futures Markets,* 32**,** 99-121.

Yen, G. & Lee, C.-F. 2008. Efficient market hypothesis (EMH): past, present and future. *Review of Pacific Basin Financial Markets and Policies,* 11**,** 305-329.

Yule, G. U. 1926. Why do we sometimes get nonsense-correlations between Time-Series?--a study in sampling and the nature of time-series. *Journal of the royal statistical society***,** 89, 1-63.

Zakaria, Z. & Shamsuddin, S. 2012. Relationship between Stock Futures Index and Cash Prices Index: Empirical Evidence Based on Malaysia Data. *Journal of Business Studies Quarterly,* 4**,** 103-112.

Zeckhauser, R. & Niederhoffer, V. 1983. The Performance of Market Index Futures Contracts. *Financial Analysts Journal,* 39**,** 59-65.

Books:

Asteriou, D. & Hall, S. G. 2007. *Applied Econometrics: a modern approach using eviews and microfit*, New York, Palgrave Macmillan.

Box, G. E. & Jenkins, G. M. 1976. *Time series analysis: forecasting and control, revised ed*, San Francisco, Holden-Day.

Brooks, C. 2008. *Introductory Econometrics for Finance,* New York, Cambridge University Press.

Fuller, W. A. 1996. *Introduction to statistical time series*, Canada, John Wiley & Sons.

Mcdonald, R. L., Cassano, M. & Fahlenbrach, R. 2006. *Derivatives markets*, Boston, Pearson Education Inc.

Montgomery, D. C. & Johnson, L. A. 1976. *Forecasting and Time Series Analysis*, New York, McGraw Hill Higher Education.

Ohno, K. 2006. *The Economic Development of Japan,* Tokyo, GRIPS Development Forum.

Vernimmen, P., Quiry, P., Dallocchio, M., Fur Y.L. & Salvi, A. 2005. *Corporate finance Theory and Practice,* West Sussex, John Wiley&Sons Ltd.

Sutcliffe, C.M.S. 2006. *Stock index futures*, Aldershot UK, Ashgate.

Wooldridge, J. 2012. *Introductory econometrics: A modern approach,* Mason, Cengage Learning.

Webpages:

Bergmann, A. 2015. *World's largest economies* [Online]. CNN Money. Available: http://money.cnn.com/news/economy/world_economies_gdp/ [Accessed 05/08 2015].

Bloomberg Business. 2015. *NKY Quote - Nikkei 225 Index - Bloomberg* [Online]. Available: http://www.bloomberg.com/quote/NKY:IND [Accessed 03/16 2015].

Central Intelligence Agency. 2014. *The World Factbook* [Online]. Central Intelligence Agency. Available: https://www.cia.gov/library/publications/the-world-factbook/geos/ja.html [Accessed 05/08 2015].

Fukao, K. & Yuan, T. 2009. *Why is Japan so heavily affected by the global economic crisis? An analysis based on the Asian international input-output tables* [Online]. VOX CEPR's Policy Portal. Available: http://www.voxeu.org/article/why-has-japan-been-so-hard-hit-global-crisis [Accessed 04/21 2015].

Irwin, N. 2013. Why Japan is the most interesting story in global economics right now. Available: http://www.washingtonpost.com/blogs/wonkblog/wp/2013/04/08/why-japan-is-the-most-interesting-story-in-global-economics-right-now/ [Accessed 03/24/2015].

Nikkei Inc. 2014. *Nikkei Stock Average FactSheet* [Online]. Available: http://www.ose.or.jp/e/derivative/225futures [Accessed 03/16 2015].

Nikkei Inc. 2015. *Index Information - Nikkei Indexes* [Online]. Nikkei Inc. Available: http://indexes.nikkei.co.jp/en/nkave/index/profile?idx=nk225 [Accessed 03/16 2015].

Osaka Securities Exchange. 2015. *Nikkei 225 Futures* [Online]. Available: http://www.ose.or.jp/e/derivative/225futures [Accessed 05/08 2015].

Pham, L. 2013. Time Series Analysis with ARIMA - ARCH/GARCH model in R. Available: https://talksonmarkets.files.wordpress.com/2012/09/time-series-analysis-with-arima-e28093-arch013.pdf [Accessed 03/20/2015].

Riley, C. 2015. Japanese stocks are on fire. But why? *How high can the Nikkei go? 19,000? 20,000? Maybe even 25,000?* [Online]. Available: http://money.cnn.com/2015/03/12/investing/japan-nikkei-stocks/ [Accessed 03/25/2015].

Trading Economics. 2015. *Japan Unemployment Rate* [Online]. Trading Economics. Available: http://www.tradingeconomics.com/japan/unemployment-rate [Accessed 05/08 2015].

8 Appendices

A.1: Formulas

A.1.1 Root mean squared error (RMSE)

$$RMSE = \left(m^{-1} \sum_{h=0}^{m-1} \hat{e}_{n+h+1}^{2} \right)^{1/2}$$

where *n* observations are used to estimate the model parameters and *m* observations are used for forecasting. If $\hat{f}_{n+h}$ is the one-step ahead forecast of x_{n+h+1} for $h = 0,1,\dots,m-1$, *x* being the time series variable, the *m* forecasts errors are $\hat{e}_{n+h+1} = x_{n+h+1} - \hat{f}_{n+h}$. When RMSE is computed for two or more forecasting models, the preferred model is the one with the lowest out-of-sample RMSE.

A.1.2 Mean absolute error (MAE)

$$MAE = m^{-1} \sum_{h=0}^{m-1} |\hat{e}_{n+h+1}|$$

The formula parameters are defined as for the RMSE formula, and again, a lower out-of-sample MAE is preferred to a larger one.

A.1.3 Akaike Information Criterion (AIC)

$$AIC = \ln(\hat{\sigma}^2) + \frac{2k}{N}$$

where *N* is the sample size, $\hat{\sigma}^2$ is the residual variance and k is the number of parameters estimated.

A.1.4 Schwartz Bayesian Information Criterion (SBIC)

$$SBIC = \ln(\hat{\sigma}^2) + \frac{k}{N} lnN$$

The formula parameters are defined as for the AIC formula. According to the AIC and SBIC formulas, both information criteria include a term k that is a function of the residual sum of squares (RSS). Further, the SBIC criterion includes a penalty for the loss of degrees of freedom when adding extra parameters in the model. When using information criteria in the decision-making, the objective is to choose the model with the number of parameters that minimizes the selected information criterion.

A.2: Methodology

A.2.1 Stationarity and unit roots

Valid statistical inference drawn from estimating regression models on time series variables assumes that the series are stationary. If they are not, the regressions may be spurious, suggesting significant correlations even if the time series variables in reality share nothing but a common time trend factor (Yule, 1926). Thus, to ensure BLUE[1] OLS regression estimates, only stationary variables should be included in the model.

A time series variable x_t that is covariance stationary, exhibits long-run mean reversion (A), has a finite variance (B) and a constant auto covariance structure (C). It is also necessary that the time series exhibits weak persistence, meaning the covariance $\gamma_s \to 0$ when s increases:

$$E(x_t) = \mu \tag{A}$$
$$var(x_t) = \sigma^2 \tag{B}$$
$$cov(x_t, x_{t-s}) = cov(x_t, x_{t+s}) = \gamma_s \tag{C}$$

Hence, a shock to a stationary time series is necessarily temporary as the effects of the shock will dissipate over time and the series will revert to its long-run mean (Asteriou and Hall, 2007).

An autoregressive AR(p) model is a model where the independent variables are lagged values of the dependent variable. The parameter p denotes the order of the autoregressive process, i.e. the number of lags of the dependent variable included in the regression. An autoregressive model of order p can be expressed as:

$$x_t = \sum_{i=1}^{p} \phi_i \, x_{t-i} + e_t \tag{D}$$

where e_t is a sequence of normal independent random variables ($e_t \sim NI(0, \sigma^2)$). A constant is left out for simplicity, but it should be included if the data suggest that there should be one.

[1] OLS being the Best Linear Unbiased Estimator

The simplest statistical time series model, an AR(1) model, is used to explain the concept of stationarity. One lag of the dependent variable is included as a regressor in this model. The model then becomes:

$$x_t = \phi x_{t-1} + e_t \tag{E}$$

The assumption that $|\phi| < 1$ rules out the possibility of an explosive series where x_t tends to increase over time. Using the lag operator L with the property $L^n x_t = x_{t-n}$, an AR(1) process equation (E) can be rewritten as:

$$x_t = \phi L x_t + e_t$$

$$(1 - \phi L) x_t = e_t \tag{F}$$

In equation (F), $(1 - \phi L)$ is the characteristic equation used to find the root, i.e. the value of L that will set the characteristic equation equal to zero:

$$L = \frac{1}{\phi} \tag{G}$$

When $|L| > 1$, equivalent to $|\phi| < 1$, the AR(1) process will be mean reverting. In equation (G), if $L = 1$ (i.e. $\phi = 1$) the process contains one unit root and is therefore not stationary, but a pure random walk (RW). Hence, a time series variable x_t is stationary when the characteristic equation has no unit roots. Then, $x_t \sim I(0)$.

A.2.2 ARIMA models

The general AR(p) model is described by equation (D). An MA process of order q, MA(q), describes the present value of a times series by a linear function of its past and current error terms. It can be described as:

$$x_t = \beta - \sum_{i=1}^{q} \theta_i e_{t-i} + e_t \tag{H}$$

where e_t is a white noise process with a zero mean and $var(e_t) = \sigma^2$. An MA(q) process with finite weights ($\theta's$) is always stationary regardless of the value of the weights (Montgomery and Johnson, 1976). More specifically, an MA(1) process can be described as:

$$x_t = \beta - \theta e_{t-1} + e_t \tag{I}$$

Then, by combining an AR(p) and an MA(q) process, the ARMA(p,q) model becomes:

$$x_t = \beta + \sum_{i=1}^{p} \phi_i x_{t-i} - \sum_{i=1}^{q} \theta_i e_{t-i} + e_t \tag{J}$$

where the $\phi's$ are the autoregressive parameters to be estimated, the $\theta's$ are the moving average parameters to be estimated, the $x's$ are the original series and the $e's$ are a series of unknown random residuals, which are assumed to follow the normal probability distribution, and $E(e_t, e_s) = 0$ when $t \neq s$. In equation (I) and (J), the MA parameters (the $\theta's$) have negative signs. Even though some software programs define these with a positive sign, the convention employed here is according to Box and Jenkins (1976).

Using the lag operator L, the general ARMA(p,q) model can be rewritten as:

$$\left(1 - \phi_1 L - \cdots - \phi_p L^p\right) x_t = \beta + (1 - \theta_1 L - \cdots - \theta_q L^q) e_t \tag{K}$$

equivalent to:

$$\phi_p(L) x_t = \beta + \theta_q(L) e_t \tag{L}$$

where $\phi_p(L)$ and $\theta_q(L)$ are polynomials in L of orders p and q respectively. The reason for writing an ARMA model in terms of the lag operator is to more easily see how several models may be equivalent. This is to obtain the model with the least number of parameters.

In practice, many time series are non-stationary. Then, since the Box-Jenkins method is for stationary models only, such time series need to be transformed into stationary time series before the methodology is applied. An ARMA(p,q) model is equivalent to an ARIMA(p,d,q)

model fitted to a time series variable that is I(d), i.e. needs to be differenced d times to become stationary. Using the lag operator, an ARIMA(p,d,q) model has the following structure:

$$\phi_p(L)(1-L)^d x_t = \beta + \theta_q(L)e_t \quad \text{(M)}$$

where d is the order of differencing. If the time series exhibits an occasional change of mean, first differences will most often result in a stationary series. For seasonal series, the Box-Jenkins methodology provides a modification that is not discussed here.

The Box-Jenkins methodology (Box and Jenkins, 1976) for estimating ARIMA models involves a three-step process: identification, estimation and diagnostic checking.

Step 1: Identification

The aim in the first step is to determine the order of the ARIMA model. Obtaining an appropriate order is necessary to capture the dynamic features of the data. In the identification step of the procedure the sample autocorrelation function (ACF) and the sample partial autocorrelation function (PACF) are useful. The ACF gives the correlation between different lags of a series, while the PACF is constructed to measure the correlation between an observation s periods ago and the current observation while controlling for intermediate lags.

The ACF becomes particularly useful in indicating whether a series exhibit a trend that causes it to be non-stationary. Large autocorrelations that persist even after several lags would be such an indication, suggesting that the series should be differenced in order to become stationary. An autocorrelation coefficient larger than $2/\sqrt{N}$ in absolute value, N being the number of observations, is statistically significant (Montgomery and Johnson, 1976). Differencing a series once usually reduces the number of large autocorrelations. However, if this is not the case and the time series is still non-stationary in the first difference, the series would have to be differenced again. Thus, when the autocorrelation die out quickly, the appropriate value of d has been found. An alternative approach to find the necessary number of times the series must be differenced in order for it to become stationary, is conducting the ADF test presented in the methodology chapter.

The PACF of the appropriately differenced series is convenient when determining the autoregressive order p of an ARIMA model. If the PACF cuts off after a few lags, the last lag with a large value would be the estimated value of p. If the PACFs do not cut off, the reason could either be that the model is a moving average model ($p = 0$) or an ARIMA model with positive p and q. The order of the MA component q is also found from an inspection of the PACFs. The appropriate q is the last lag with a large PACF value. Now, if the PACFs do not cut off, the reason is either that the model is autoregressive ($q = 0$) or that it is an ARIMA model with a positive p and q.

To determine whether to include a constant or not in the ARIMA model, the order of differencing is considered. A constant is usually included if the order of differencing is one or less. When the time series is I(1), the constant indicates the average trend in the forecast provided by the ARIMA model.

In practice, the ACF and the PACF might not exhibit simple patterns to be interpreted in order to determine the order of the model. In these cases information criteria such as the Akaike information criterion (AIC) or the Schwartz's Bayesian information criterion (SBIC) is often used to evaluate the ARIMA models (Brooks, 2008).

<u>Step 2: Estimation</u>

Model estimation involves estimating the parameters in several alternative ARIMA models using maximum likelihood (ML) techniques. Using an ML technique to estimate regression parameters involves maximizing the log-likelihood function (Wooldridge, 2012). Then, the parameters are chosen to maximize the likelihood that the empirical covariance matrix is drawn from a population for which the implied covariance matrix is valid. The ML estimates and their standard errors will be asymptotically unbiased, consistent and efficient, given the fulfilment of certain conditions (Schermelleh-Engel et. al., 2003). One of these conditions is that the time series sample is drawn from a multivariate normal distribution. The other conditions are that the model is correctly specified and that the sample size is sufficiently large.

For technical estimation reasons one must notice that including lagged error terms as regressors in an ARIMA model causes the prediction of the model to not be a linear function of the coefficients. This in spite of the fact that they are linear functions of past time series data. This implies that the coefficients in ARIMA models where $q > 0$ must be estimated by nonlinear optimization methods, and not just by solving a system of equations.

Step 3: Diagnostic checking

To determine whether the fitted model is adequately specified and estimated, the third step involves checking the model using residual diagnostics. The aim is to check for linear dependence in the residuals. If present, this would suggest that the estimated ARIMA model is insufficient in capturing the features of the time series employed. Thus, the key element in the third step of the Box-Jensen method is to make sure that the residuals of the selected model are normally distributed. The ACF and PACF could be used in testing for autocorrelation in the residuals. An information criterion (i.e. AIC or SBIC) is used to decide between parsimonious models.

A.3: Variable definitions

A.3.1 Notation

The following notation is used:

Variable name in Eviews	Variable name in Excel	Symbol in DATASTREAM	Name in text	Description
Spot	Spot	JAPDOWA	Nikkei 225 or Spot price	Nikkei 225 Stock Average Index
Futures	Futures	ONACS00	Nikkei 225 futures or Futures price	Nikkei 225 futures contract traded on the OSE
	Rf interest	Y79259	Risk free interest rate	Tokyo Interbank Euroyen 3 month offered interest rate
	DY		Dividend yield	Nikkei 225 Stock Average Index dividend yield
RDT			RDT	Cost-of-carry component in the co-integration model for the ECM-COCs.

The average annual Nikkei 225 Stock Average Index dividend yield derived from the years 2010-2014 is assumed to be an approximation of the annual dividend yield of the Nikkei 225 Stock Average Index for the total of the sample period and out-of sample period. From the average annual dividend yield, the daily dividend yield is found by using the formula:

$$DY_d = (1 + DY_y)^{\frac{1}{365}} \qquad \text{(N)}$$

where DY_d is the daily dividend yield and DY_y is the annual dividend yield.

When using the best fitted models for trading in the out-of-sample period, the daily three month risk free interest rate for the period is averaged and re-computed to make an approximation of the average daily risk free interest rate. This is done in a similar way as in equation (N).

A.3.2 The Nikkei 225 Stock Average Index

The Nikkei 225 Stock Average Index was first computed on May 16th 1949, at the opening of Tokyo Stock Exchange (TSE) after the Second World War (Bloomberg, 2015). The index has since become the globally most popular benchmark on the Japanese stock market (Nikkei Inc., 2015). The Nikkei 225 is a price-weighted averaged equity index for domestic common stocks listed in the First Section of the TSE and it consists of the 225 most actively traded stocks listed.

Once a year, the Nikkei 225 index is subject to a Periodic Review. By adding stocks with high liquidity to the index and removing those that are less liquid, the goal is to maintain the representativeness of the market. In addition to liquidity, sector balance is also an aspect that is considered at each Periodic Review. An Extraordinary Review is conducted when specific events require the delisting of some stocks (Nikkei Inc., 2015).

When re-computing the Nikkei 225, each stock is weighted based on its presumed par value. Since index constituents may be changed, the "Dow method" is used to modify the divisor each time changes are made in the constituents or in the case of events such as stock splits. The Nikkei 225 Index is recalculated every 15 seconds as long as the TSE is open. At the close of the market, the index value is usually computed by the last traded price of each constituent stock.

A.3.3 OSE futures contracts on the Nikkei 225 Stock Average Index

Nikkei 225 Stock Average Index futures contracts began trading at the Osaka Securities Exchange (OSE) in 1988. The derivative was the first index futures contract listed on a Japanese exchange. The futures contracts have a maturity of three months and are settled quarterly, in March, June, September and December. Each contract unit is for Nikkei 225 x ¥1 000, and the contracts are cash settled (OSE, 2015).

A.4: Tables error correction models

A.4.1 Pre financial crisis sample

ECM*	α_0 Constant (t-statistic)	δ ECT (t-statistic)	β_0 DlogSpot(t) (t-statistic)	β_1 DlogSpot(t-1) (t-statistic)	γ_0 DlogFutures(t) (t-statistic)	γ_1 DlogFutures(t-1) (t-statistic)	Adj.R^2	Durbin Watson statistic	SIC
Nº 1	-1.84E-05 (-0.131245)	-0.810695 (-18.72634)			0.846836 (39.03045)		0.863547	1.966748	-8.103685
Nº 2	-7.28E-05 (-0.516592)	-0.807976 (-18.75196)		0.022817 (1.506645)	0.845838 (38.54271)		0.861235	2.022816	-8.081212
Nº 3	2.52E-06 (0.016851)	-0.892738 (-16.54425)	1.015743 (35.82419)				0.865087	2.017536	-7.922119
Nº 4	6.20E-05 (0.409031)	-0.875868 (-19.83738)	1.014520 (34.59173)			-0.022026 (-1.085643)	0.863929	2.005112	-7.899169

* Model 1 and 2: Futures leads spot. Model 3 and 4: Spot leads futures

A.4.2 Post financial crisis sample – futures leads spot

ECM	α_0 Constant (t-statistic)	δ ECT (t-statistic)	β_1 DlogSpot(t-1) (t-statistic)	β_2 DlogSpot(t-2) (t-statistic)	γ_0 DlogFutures(t) (t-statistic)	γ_1 DlogFutures(t-1) (t-statistic)	γ_2 DlogFutures(t-2) (t-statistic)	Adj.R^2	Durbin Watson statistic	SIC
Nº 1	-5.72E-05 (-0.287299)	-0.847255 (-15.40424)			0.862525 (36.84628)			0.8645	1.7628	-7.4548
Nº 2	-0.000133 (-0.699024)	-0.839356 (-13.85129)	0.058588 (2.233350)		0.858012 (33.47885)			0.8634	1.9213	-7.4789
Nº 3	-0.000118 (-0.722581)	-0.670174 (-8.153388)	-0.119456 (-2.346311)		0.864242 (33.41682)	0.181840 (3.222805)		0.8681	1.9934	-7.5093
Nº 4	-6.49E-05 (-0.155683)	-0.302453 (-1.492946)	0.165211 (1.111420)			-0.047824 (-0.370821)		0.0233	2.0002	-5.5116
Nº 5	-5.12E-05 (-0.121641)	-0.165501 (-1.152786)				0.099207 (2.460353)		0.0205	1.9669	-5.5135
Nº 6	-0.000124 (-0.734249)	-0.673243 (-7.794486)	-0.146471 (-2.512987)		0.867922 (32.00070)	0.205955 (3.060938)	0.033661 (2.670650)	0.8704	1.9515	-7.4914
Nº 7	-0.000128 (-0.794475)	-0.561222 (-6.342320)	-0.284488 (-3.033371	-0.171577 (-1.509994)	0.869417 (32.96472)	0.345600 (4.128932)	0.188706 (1.852354)	0.8740	1.9038	-7.5145

A.4.3 Post financial crisis sample – spot leads futures

ECM	α_0 Constant (t-statistic)	δ ECT (t-statistic)	β_0 DlogSpot(t) (t-statistic)	β_1 DlogSpot(t-1) (t-statistic)	β_2 DlogSpot(t-2) (t-statistic)	γ_1 DlogFutures(t-1) (t-statistic)	γ_2 DlogFutures(t-2) (t-statistic)	Adj.R^2	Durbin Watson statistic	SIC
Nº 1	8.71E-05 (0.409526)	-0.953409 (-13.69279)	1.001636 (29.23605)					0.8745	1.8806	-7.3068
Nº 2	0.000141 (0.710931)	-0.862530 (-11.31874)	1.004585 (31.80603)			-0.071584 (-2.940412)		0.8709	1.9996	-7.3460
Nº 3	0.000127 (0.716605)	-0.726984 (-8.761696)	1.000914 (30.82724)	0.163430 (3.305851)		-0.216751 (-3.996456)		0.8738	2.0394	-7.3635
Nº 4	6.23E-05 (0.138037)	-0.427495 (-2.463222)		0.327744 (2.105129)		-0.263727 (-2.000474)		0.0642	2.0471	-5.3649
Nº 5	8.35E-05 (0.180402)	-0.678049 (-4.605668)		0.070326 (1.608497)				0.0565	2.0458	-5.3615
Nº 6	0.000123 (0.659673)	-0.753436 (-8.679218)	1.000328 (29.56721)	0.153901 (2.820531)	-0.021201 (-0.858018)	-0.205677 (-3.249746)		0.8747	2.0161	-7.3394
Nº 7	0.000121 (0.650934)	-0.720457 (-8.423843)	0.999004 (29.74058)	0.203199 (3.667699)		-0.252309 (-3.906071)	-0.046789 (-2.988549)	0.8762	1.9910	-7.3515
Nº 8	0.000126 (0.715458)	-0.602648 (-7.058287)	1.001473 (28.62939)	0.348084 (3.525229)	0.181711 (1.489953)	-0.399434 (-4.508134)	-0.210754 (-1.943759)	0.8796	1.9332	-7.3739

A.5: Tables error correction models based on the cost-of-carry relationship

A.5.1 Pre financial crisis sample

ECM*	α_0 Constant (t-statistic)	δ ECT (t-statistic)	β_0 DlogSpot(t) (t-statistic)	β_1 DlogSpot(t-1) (t-statistic)	γ_0 DlogFutures(t) (t-statistic)	γ_1 DlogFutures(t-1) (t-statistic)	Adj.R^2	Durbin Watson statistic	SIC
Nº 1	-1.56E-05 (-0.113113)	-0.811928 (-18.81607)			0.846855 (38.46664)		0.864199	1.973178	-8.108477
Nº 2	-6.99E-05 (-0.501975)	-0.809417 (-18.74719)		0.020329 (1.368925)	0.845784 (37.95007)		0.861810	2.023789	-8.085363
Nº 3	1.15E-06 (0.007912)	-0.895487 (-27.84173)	1.016814 (78.09894)				0.865556	2.019117	-7.925603
Nº 4	6.04E-05 (0.399496)	-0.881076 (-19.74187)	1.015600 (34.39114)			-0.019548 (-0.969806)	0.864319	2.006726	-7.902045

*Model 1 and 2; Futures leads spot. Model 3 and 4; Spot leads futures

A.5.2 Post financial crisis sample – futures leads spot

ECM	α_0 Constant (t-statistic)	δ ECT (t-statistic)	β_1 DlogSpot(t-1) (t-statistic)	β_2 DlogSpot(t-2) (t-statistic)	γ_0 DlogFutures(t) (t-statistic)	γ_1 DlogFutures(t-1) (t-statistic)	γ_2 DlogFutures(t-2) (t-statistic)	Adj.R^2	Durbin Watson statistic	SIC
Nº 1	-4.41E-05 (-0.219321)	-0.839681 (-15.50409)			0.865404 (36.94230)			0.8635	1.7673	-7.4472
Nº 2	-0.000112 (-0.588326)	-0.828650 (-14.01645)	0.058168 (2.186326)		0.861699 (33.73491)			0.8622	1.9305	-7.4699
Nº 3	-0.000102 (-0.620156)	-0.655735 (-8.245705)	-0.125218 (-2.486218)		0.867287 (33.83751)	0.187407 (3.351027)		0.8672	2.0014	-7.5023
Nº 4	-5.05E-05 (-0.121223)	-0.227134 (-1.054458)	0.130568 (0.881049)			-0.011590 (-0.089635)		0.0204	1.9993	-5.5086
Nº 5	-4.22E-05 (-0.100564)	-0.120043 (-0.775777)				0.104620 (2.609334)		0.0189	1.9729	-5.5119
Nº 6	-9.95E-05 (-0.589698)	-0.655499 (-7.820942)	-0.154682 (-2.686029)		0.871249 (32.43552)	0.214781 (3.237643)	0.033708 (2.696553)	0.8695	1.9619	-7.4841
Nº 7	-0.000101 (-0.590037)	-0.677474 (-8.036602)	-0.117197 (-2.071887)	0.009342 (0.433742)	0.868942 (31.93178)	0.179882 (2.753409)		0.8684	1.9902	-7.4760
Nº 8	-0.000107 (-0.668305)	-0.542381 (-5.860949)	-0.295106 (-3.017507)	-0.175295 (-1.507656)	0.872162 (33.21310)	0.356811 (4.056579)	0.192100 (1.841035)	0.8732	1.9104	-7.5083

A.5.3 Post financial crisis sample – spot leads futures

ECM	α_0 Constant (t-statistic)	δ ECT (t-statistic)	β_0 DlogSpot(t) (t-statistic)	β_1 DlogSpot(t-1) (t-statistic)	β_2 DlogSpot(t-2) (t-statistic)	γ_1 DlogFutures(t-1) (t-statistic)	γ_2 DlogFutures(t-2) (t-statistic)	Adj.R²	Durbin Watson statistic	SIC
Nº 1	7.66E-05 (-0.357887)	-0.949972 (-13.78897)	0.997652 (29.59847)					0.8743	1.8826	-7.3051
Nº 2	0.000124 (0.627145)	-0.856170 (-11.66495)	0.999850 (32.52845)			-0.071525 (-2.929976)		0.8708	2.0090	-7.3451
Nº 3	0.000113 (0.637519)	-0.720690 (-9.057079)	-0.996901 (31.41168)	0.164746 (3.315083)		-0.217851 (-4.044720)		0.8737	2.0467	-7.3631
Nº 4	6.14E-05 (0.136682)	-0.493638 (-2.613309)		0.295200 (1.894804)		-0.229718 (1.738384)		0.0683	2.0447	-5.3693
Nº 5	7.57E-05 (0.164973)	-0.710465 (-4.564799)		0.070904 (1.648036)				0.0626	2.0441	-5.3680
Nº 6	0.000101 (0.542468)	-0.744137 (-8.909447)	0.995944 (30.13347)	0.157217 (2.873839)	-0.021203 (-0.859567)	-0.209732 (-3.352629)		0.8748	2.0252	-7.3396
Nº 7	0.000100 (0.538446)	-0.711539 (-8.730939)	0.994818 (30.29649)	0.206240 (3.762633)		-0.256080 (-4.058899)	-0.046620 (-3.091581)	0.8762	1.9993	-7.3516
Nº 8	0.000108 (0.616561)	-0.595197 (-6.713719)	0.997947 (28.99840)	0.350231 (3.401594)	0.180855 (1.455405)	-0.402194 (-4.352358)	-0.209866 (-1.901357)	0.8796	1.9394	-7.3736

A.6: Spot and futures autocorrelations and partial autocorrelations

A.6.1 Pre financial crisis sample

Table: *Autocorrelations for the logarithmic Nikkei 225 spot price*

Lag	Autocorrelation	Std. Error[a]	Box-Ljung Statistic Value	df	Sig.[b]
1	.950	.030	982.798	1	.000
2	.940	.030	1954.118	2	.000
3	.937	.030	2920.884	3	.000
4	.933	.030	3878.031	4	.000
5	.928	.030	4825.975	5	.000
6	.924	.030	5768.908	6	.000
7	.919	.030	6703.636	7	.000
8	.914	.030	7630.771	8	.000
9	.911	.030	8553.301	9	.000
10	.910	.030	9473.979	10	.000
11	.908	.030	10391.089	11	.000
12	.908	.030	11309.779	12	.000
13	.902	.030	12219.271	13	.000
14	.902	.030	13125.334	14	.000
15	.899	.030	14024.777	15	.000
16	.893	.030	14917.492	16	.000

a. The underlying process assumed is independence (white noise).
b. Based on the asymptotic chi-square approximation.

Figure: *Autocorrelation function (ACF) for the logarithmic Nikkei 225 spot price*

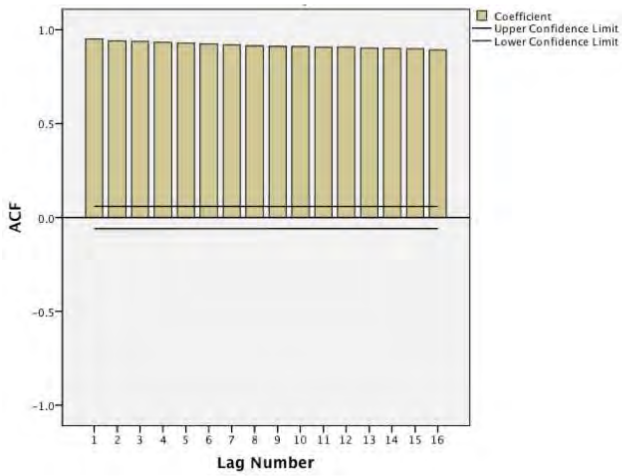


Table: *Partial autocorrelations for the logarithmic Nikkei 225 spot price*

Lag	Partial Autocorrelation	Std. Error
1	.950	.031
2	.391	.031
3	.264	.031
4	.160	.031
5	.100	.031
6	.076	.031
7	.047	.031
8	.022	.031
9	.035	.031
10	.063	.031
11	.049	.031
12	.076	.031
13	-.006	.031
14	.037	.031
15	.010	.031
16	-.039	.031

Figure: *Partial autocorrelation function (PACF) for the logarithmic Nikkei 225 spot price*

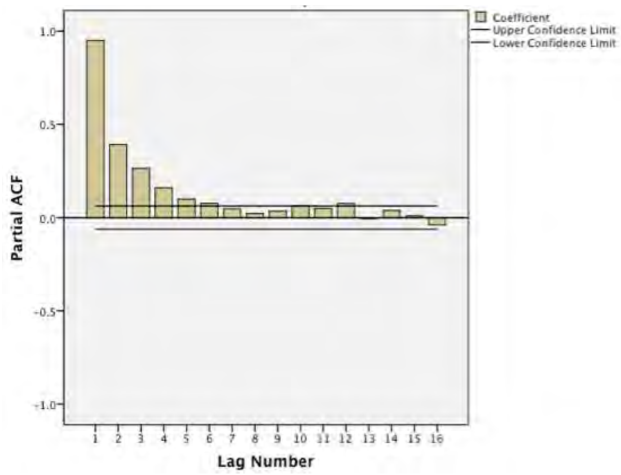


Table: *Autocorrelations for the logarithmic Nikkei 225 futures price*

Lag	Autocorrelation	Std. Error[a]	Box-Ljung Statistic		
			Value	df	Sig.[b]
1	.949	.030	981.914	1	.000
2	.940	.030	1952.424	2	.000
3	.937	.030	2917.783	3	.000
4	.932	.030	3873.976	4	.000
5	.927	.030	4820.722	5	.000
6	.923	.030	5762.375	6	.000
7	.919	.030	6695.826	7	.000
8	.914	.030	7622.086	8	.000
9	.910	.030	8543.617	9	.000
10	.909	.030	9462.904	10	.000
11	.907	.030	10378.971	11	.000
12	.907	.030	11296.230	12	.000
13	.902	.030	12204.858	13	.000
14	.901	.030	13109.417	14	.000
15	.898	.030	14007.388	15	.000
16	.892	.030	14898.688	16	.000

a. The underlying process assumed is independence (white noise).
b. Based on the asymptotic chi-square approximation.

Figure: *Autocorrelation function (ACF) for the logarithmic Nikkei 225 futures price*

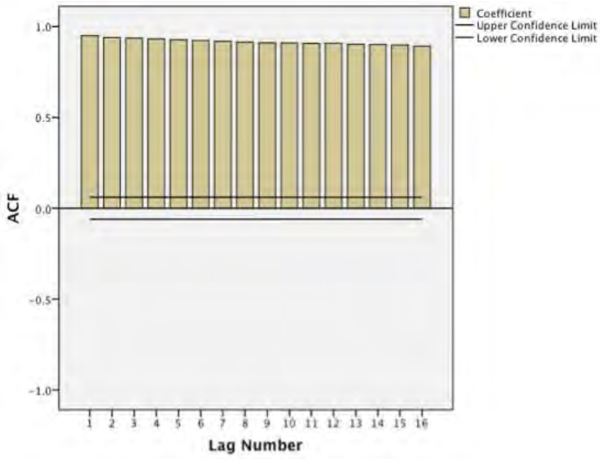
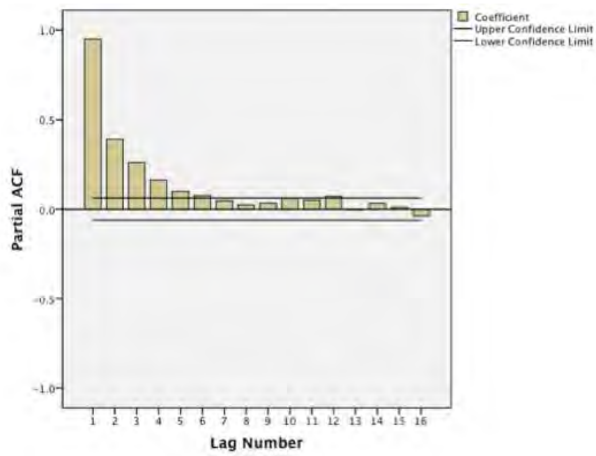


Table: *Partial autocorrelations for the logarithmic Nikkei 225 futures price*

Lag	Partial Autocorrelation	Std. Error
1	.949	.031
2	.392	.031
3	.261	.031
4	.163	.031
5	.099	.031
6	.076	.031
7	.047	.031
8	.025	.031
9	.035	.031
10	.061	.031
11	.050	.031
12	.073	.031
13	-.002	.031
14	.033	.031
15	.010	.031
16	-.039	.031

Figure: *Partial autocorrelation function (PACF) for the logarithmic Nikkei 225 futures price*



A.6.2 Post financial crisis sample

Table: *Autocorrelations for the logarithmic Nikkei 225 spot price*

Lag	Autocorrelation	Std. Error[a]	Box-Ljung Statistic		
			Value	df	Sig.[b]
1	.945	.026	1334.684	1	.000
2	.932	.026	2645.996	2	.000
3	.924	.026	3938.314	3	.000
4	.920	.026	5220.108	4	.000
5	.916	.026	6492.447	5	.000
6	.913	.026	7754.739	6	.000
7	.908	.026	9003.122	7	.000
8	.903	.026	10241.263	8	.000
9	.898	.026	11470.341	9	.000
10	.887	.026	12672.228	10	.000
11	.885	.026	13868.942	11	.000
12	.880	.026	15052.352	12	.000
13	.875	.026	16224.907	13	.000
14	.871	.026	17382.079	14	.000
15	.875	.026	18549.215	15	.000
16	.862	.026	19688.549	16	.000

a. The underlying process assumed is independence (white noise).

b. Based on the asymptotic chi-square approximation.

Figure: *Autocorrelation function (ACF) for the logarithmic Nikkei 225 spot price*

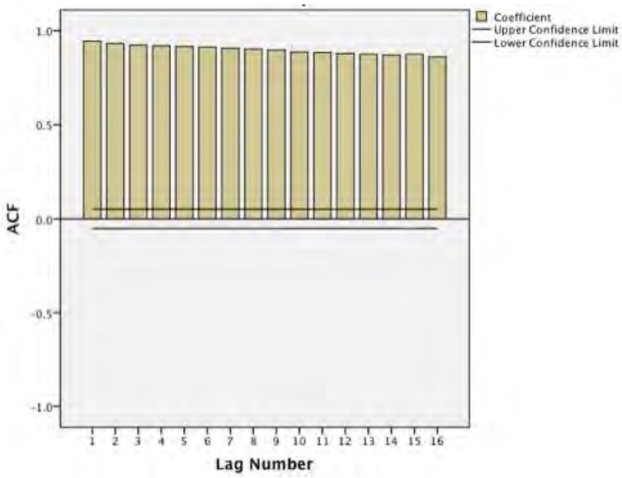


Table: *Partial autocorrelations for the logarithmic Nikkei 225 spot price*

Lag	Partial Autocorrelation	Std. Error
1	.945	.027
2	.369	.027
3	.208	.027
4	.172	.027
5	.121	.027
6	.105	.027
7	.051	.027
8	.030	.027
9	.026	.027
10	-.059	.027
11	.036	.027
12	.006	.027
13	.009	.027
14	.009	.027
15	.120	.027
16	-.088	.027

Figure: *Partial autocorrelation function (PACF) for the logarithmic Nikkei 225 spot price*

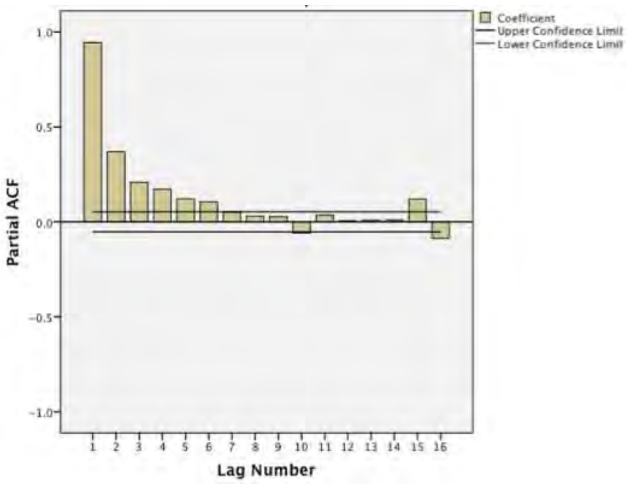


Table: *Autocorrelations for the logarithmic Nikkei 225 futures price*

Lag	Autocorrelation	Std. Error[a]	Box-Ljung Statistic Value	df	Sig.[b]
1	.944	.026	1331.373	1	.000
2	.932	.026	2641.173	2	.000
3	.924	.026	3932.845	3	.000
4	.920	.026	5213.249	4	.000
5	.915	.026	6483.507	5	.000
6	.913	.026	7743.301	6	.000
7	.907	.026	8989.894	7	.000
8	.902	.026	10226.322	8	.000
9	.897	.026	11453.461	9	.000
10	.886	.026	12653.896	10	.000
11	.884	.026	13848.734	11	.000
12	.879	.026	15030.705	12	.000
13	.875	.026	16201.503	13	.000
14	.869	.026	17355.168	14	.000
15	.875	.026	18520.165	15	.000
16	.861	.026	19658.064	16	.000

a. The underlying process assumed is independence (white noise).
b. Based on the asymptotic chi-square approximation.

Figure: *Autocorrelation function (ACF) for the logarithmic Nikkei 225 futures price*

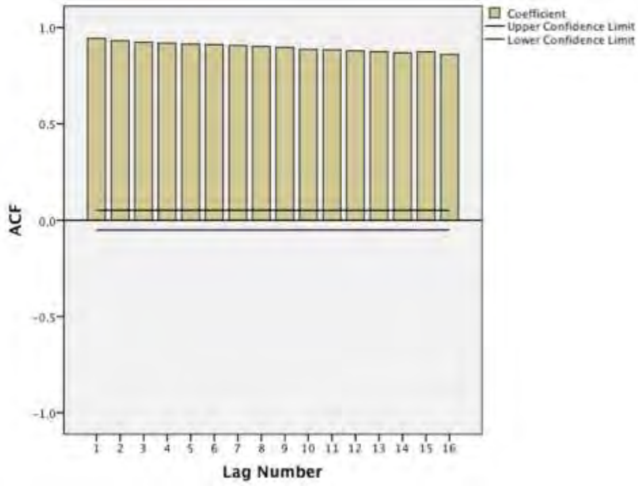
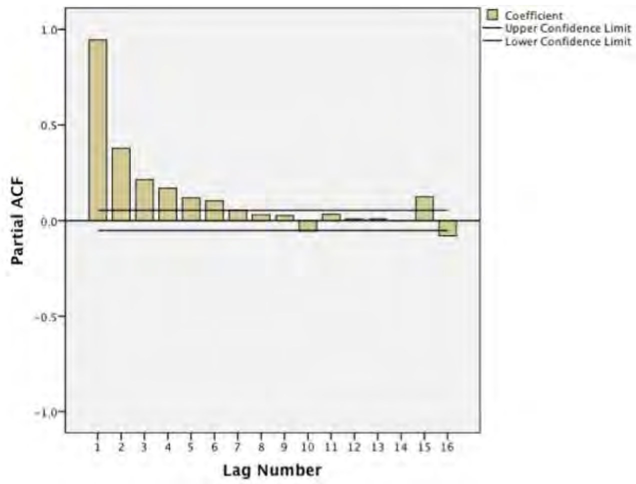


Table: *Partial autocorrelations for the logarithmic Nikkei 225 futures price*

Lag	Partial Autocorrelation	Std. Error
1	.944	.027
2	.377	.027
3	.213	.027
4	.169	.027
5	.118	.027
6	.103	.027
7	.055	.027
8	.030	.027
9	.025	.027
10	-.055	.027
11	.032	.027
12	.008	.027
13	.008	.027
14	.001	.027
15	.124	.027
16	-.081	.027

Figure: *Partial autocorrelation function (PACF) for the logarithmic Nikkei 225 futures price*



A.7: Alternative ARIMA models

A.7.1 Pre financial crisis sample

Table: Univariate spot price ARIMA models (Dependent variable: log($Spot_t$))

ARIMA	$\bar{R}^2$	Constant (t-statistic)	AR(1) (t-statistic)	AR(2) (t-statistic)	SIC	LM (p value)
(1,0,0)	0.995	9.517** (73,104)	0.997** (454.337)		-6.118521	1.459* (0.018)
(1,1,0)	-1.19E-04	-1.55E-04 (-0.419)	0.031 (0.940)		-6.118560	0.941 (0.967)
(2,1,0)	-1.31E-03	-1,72E-04 (-0.441)	0.030 (0,895)	-0.007 (-0.199)	-6.095994	2.161 (0,826)
(1,2,0)	0.229	1.44E-05 (0.086)	-0.484** (-17.702)		-5.706879	191.415** (0.000)

***1 % significant, *5 % significant*

Table: Univariate futures price ARIMA models (Dependent variable: log($Futures_t$))

ARIMA	$\bar{R}^2$	Constant (t-statistic)	AR(1) (t-statistic)	AR(2) (t-statistic)	SIC	LM (p value)
(1,1,0)	5.96E-03	-9.11E-05 (-0.245)	-0.084** (-2.106)		-5.923014	8.856 (0.115)
(2,1,0)	8.00E-03	-1.86E-04 (-0.451)	-0.077** (-2.286)	0.061* (1.801)	-5.906286	10.482* (0.063)
(1,2,0)	0.321	-6.76E-06 (-0.021)	-0.570*** (-20.553)		-5.529213	193.827*** (0.000)
(1,0,0)	0.994	9.525*** (83.467)	0.997*** (530.767)		-5.925810	12.504** (0.013)

****1 % significant, **5 % significant, *10 % significant*

A.7.2 Post financial crisis sample

Table: Univariate spot price ARIMA models (Dependent variable: log(Spot$_t$))

ARIMA	$\bar{R}^2$	Constant (t-statistic)	AR(1) (t-statistic)	AR(2) (t-statistic)	SIC	LM (p value)
(1,1,0)	0.015	-2.33E-05 (-0.049)	0.124* (2.784)		-5.512769	4.106 (0.534)
(2,1,0)	0.018	-7.56E-05 (-0.159)	0.132* (4.508)	-0.073** (-2.508)	-5.481667	21.396* (0.001)
(1,2,0)	0.155	3.33E-06 (0.008)	-0.392* (-14.923)		-5.094894	272.885* (0.000)
(1,0,0)	0.993	9.315* (45.292)	0.998* (452.838)		-5.460483	26.645* (0.000)

**1 % significant, **5 % significant*

Table: Univariate futures price ARIMA models (Dependent variable: log(Futures$_t$))

ARIMA	$\bar{R}^2$	Constant (t-statistic)	AR(1) (t-statistic)	AR(2) (t-statistic)	SIC	LM (p value)
(1,1,0)	0.003	-6.93E-06 (-0.015)	-0.058* (1.623)		-5.311076	3.250 (0.661)
(1,2,0)	0.256	-9.54E-05 (-0.441)	-0.509*** (-8.861)		-4.841793	316.054* (0.000)
(2,1,0)	0.002	-1.8E-04 (-0.253)	-0.059 (-1.453)	-0.028 (-0.393)	-5.279647	28.629* (0.000)
(1,0,0)	0.992	9.299*** (69.927)	0.997*** (416.018)		-5.236529	6.562 (0.255)

****1 % significant, **5 % significant, *10 % significant*

A.8: Tables of critical values

A.8.1 Dickey-Fuller critical values

Table: Dickey-Fuller critical values for different significance levels, α

Sample size T	0.01	0.025	0.05	0.10
		τ		
25	−2.66	−2.26	−1.95	−1.60
50	−2.62	−2.25	−1.95	−1.61
100	−2.60	−2.24	−1.95	−1.61
250	−2.58	−2.23	−1.95	−1.62
500	−2.58	−2.23	−1.95	−1.62
∞	−2.58	−2.23	−1.95	−1.62
		τ_μ		
25	−3.75	−3.33	−3.00	−2.63
50	−3.58	−3.22	−2.93	−2.60
100	−3.51	−3.17	−2.89	−2.58
250	−3.46	−3.14	−2.88	−2.57
500	−3.44	−3.13	−2.87	−2.57
∞	−3.43	−3.12	−2.86	−2.57
		τ_τ		
25	−4.38	−3.95	−3.60	−3.24
50	−4.15	−3.80	−3.50	−3.18
100	−4.04	−3.73	−3.45	−3.15
250	−3.99	−3.69	−3.43	−3.13
500	−3.98	−3.68	−3.42	−3.13
∞	−3.96	−3.66	−3.41	−3.12

Source: Originally Fuller (1996), reprinted by Brooks (2008).

A.8.2 Engle and Granger co-integration critical values

Table: Critical values for the Engle and Granger co-integration test on regression residuals with no constant in test regression

Number of variables in system	Sample size T	0.01	0.05	0.10
	50	−4.32	−3.67	−3.28
2	100	−4.07	−3.37	−3.03
	200	−4.00	−3.37	−3.02
	50	−4.84	−4.11	−3.73
3	100	−4.45	−3.93	−3.59
	200	−4.35	−3.78	−3.47
	50	−4.94	−4.35	−4.02
4	100	−4.75	−4.22	−3.89
	200	−4.70	−4.18	−3.89
	50	−5.41	−4.76	−4.42
5	100	−5.18	−4.58	−4.26
	200	−5.02	−4.48	−4.18

Source: Originally Engle and Yoo (1987), reprinted in Brooks (2008).

A.8.3 MacKinnon critical values co-integration

Table: Critical values for co-integration test

N	Variant	Size (%)	Obs.	β_∞	(SE)	β_1	β_2
1	No constant	1	600	−2.5658	(.0023)	−1.960	−10.04
		5	600	−1.9393	(.0008)	−0.398	0.0
		10	560	−1.6156	(.0007)	−0.181	0.0
1	No trend	1	600	−3.4335	(.0024)	−5.999	−29.25
		5	600	−2.8621	(.0011)	−2.738	−8.36
		10	600	−2.5671	(.0009)	−1.438	−4.48
1	With trend	1	600	−3.9638	(.0019)	−8.353	−47.44
		5	600	−3.4126	(.0012)	−4.039	−17.83
		10	600	−3.1279	(.0009)	−2.418	−7.58
2	No trend	1	600	−3.9001	(.0022)	−10.534	−30.03
		5	600	−3.3377	(.0012)	−5.967	−8.98
		10	600	−3.0462	(.0009)	−4.069	−5.73
2	With trend	1	600	−4.3266	(.0022)	−15.531	−34.03
		5	560	−3.7809	(.0013)	−9.421	−15.06
		10	600	−3.4959	(.0009)	−7.203	−4.01
3	No trend	1	560	−4.2981	(.0023)	−13.790	−46.37
		5	560	−3.7429	(.0012)	−8.352	−13.41
		10	600	−3.4518	(.0010)	−6.241	−2.79
3	With trend	1	600	−4.6676	(.0022)	−18.492	−49.35
		5	600	−4.1193	(.0011)	−12.024	−13.13
		10	600	−3.8344	(.0009)	−9.188	−4.85
4	No trend	1	560	−4.6493	(.0023)	−17.188	−59.20
		5	560	−4.1000	(.0012)	−10.745	−21.57
		10	600	−3.8110	(.0009)	−8.317	−5.19
4	With trend	1	600	−4.9695	(.0021)	−22.504	−50.22
		5	560	−4.4294	(.0012)	−14.501	−19.54
		10	560	−4.1474	(.0010)	−11.165	−9.88
5	No trend	1	520	−4.9587	(.0026)	−22.140	−37.29
		5	560	−4.4185	(.0013)	−13.641	−21.16
		10	600	−4.1327	(.0009)	−10.638	−5.48
5	With trend	1	600	−5.2497	(.0024)	−26.606	−49.56
		5	600	−4.7154	(.0013)	−17.432	−16.50
		10	600	−4.4345	(.0010)	−13.654	−5.77
6	No trend	1	480	−5.2400	(.0029)	−26.278	−41.65
		5	480	−4.7048	(.0018)	−17.120	−11.17
		10	480	−4.4242	(.0010)	−13.347	0.0
6	With trend	1	480	−5.5127	(.0033)	−30.735	−52.50
		5	480	−4.9767	(.0017)	−20.883	−9.05
		10	480	−4.6999	(.0011)	−16.445	0.0

Source: MacKinnon (1991)

A.8.4 Critical values Student's t distribution

Table: Critical values of Student's t distribution for different probability levels, α, and degrees of freedom, ν

α	0.4	0.25	0.15	0.1	0.05	0.025	0.01	0.005	0.001	0.0005
ν										
1	0.3249	1.0000	1.9626	3.0777	6.3138	12.7062	31.8205	63.6567	318.3087	636.6189
2	0.2887	0.8165	1.3862	1.8856	2.9200	4.3027	6.9646	9.9248	22.3271	31.5991
3	0.2767	0.7649	1.2498	1.6377	2.3534	3.1824	4.5407	5.8409	10.2145	12.9240
4	0.2707	0.7407	1.1896	1.5332	2.1318	2.7764	3.7469	4.6041	7.1732	8.6103
5	0.2672	0.7267	1.1558	1.4759	2.0150	2.5706	3.3649	4.0321	5.8934	6.8688
6	0.2648	0.7176	1.1342	1.4398	1.9432	2.4469	3.1427	3.7074	5.2076	5.9588
7	0.2632	0.7111	1.1192	1.4149	1.8946	2.3646	2.9980	3.4995	4.7853	5.4079
8	0.2619	0.7064	1.1081	1.3968	1.8595	2.3060	2.8965	3.3554	4.5008	5.0413
9	0.2610	0.7027	1.0997	1.3830	1.8331	2.2622	2.8214	3.2498	4.2968	4.7809
10	0.2602	0.6998	1.0931	1.3722	1.8125	2.2281	2.7638	3.1693	4.1437	4.5869
11	0.2596	0.6974	1.0877	1.3634	1.7959	2.2010	2.7181	3.1058	4.0247	4.4370
12	0.2590	0.6955	1.0832	1.3562	1.7823	2.1788	2.6810	3.0545	3.9296	4.3178
13	0.2586	0.6938	1.0795	1.3502	1.7709	2.1604	2.6503	3.0123	3.8520	4.2208
14	0.2582	0.6924	1.0763	1.3450	1.7613	2.1448	2.6245	2.9768	3.7874	4.1405
15	0.2579	0.6912	1.0735	1.3406	1.7531	2.1314	2.6025	2.9467	3.7328	4.0728
16	0.2576	0.6901	1.0711	1.3368	1.7459	2.1199	2.5835	2.9208	3.6862	4.0150
17	0.2573	0.6892	1.0690	1.3334	1.7396	2.1098	2.5669	2.8982	3.6458	3.9651
18	0.2571	0.6884	1.0672	1.3304	1.7341	2.1009	2.5524	2.8784	3.6105	3.9216
19	0.2569	0.6876	1.0655	1.3277	1.7291	2.0930	2.5395	2.8609	3.5794	3.8834
20	0.2567	0.6870	1.0640	1.3253	1.7247	2.0860	2.5280	2.8453	3.5518	3.8495
21	0.2566	0.6864	1.0627	1.3232	1.7207	2.0796	2.5176	2.8314	3.5272	3.8193
22	0.2564	0.6858	1.0614	1.3212	1.7171	2.0739	2.5083	2.8188	3.5050	3.7921
23	0.2563	0.6853	1.0603	1.3195	1.7139	2.0687	2.4999	2.8073	3.4850	3.7676
24	0.2562	0.6848	1.0593	1.3178	1.7109	2.0639	2.4922	2.7969	3.4668	3.7454
25	0.2561	0.6844	1.0584	1.3163	1.7081	2.0595	2.4851	2.7874	3.4502	3.7251
26	0.2560	0.6840	1.0575	1.3150	1.7056	2.0555	2.4786	2.7787	3.4350	3.7066
27	0.2559	0.6837	1.0567	1.3137	1.7033	2.0518	2.4727	2.7707	3.4210	3.6896
28	0.2558	0.6834	1.0560	1.3125	1.7011	2.0484	2.4671	2.7633	3.4082	3.6739
29	0.2557	0.6830	1.0553	1.3114	1.6991	2.0452	2.4620	2.7564	3.3962	3.6594
30	0.2556	0.6828	1.0547	1.3104	1.6973	2.0423	2.4573	2.7500	3.3852	3.6460
35	0.2553	0.6816	1.0520	1.3062	1.6896	2.0301	2.4377	2.7238	3.3400	3.5911
40	0.2550	0.6807	1.0500	1.3031	1.6839	2.0211	2.4233	2.7045	3.3069	3.5510
45	0.2549	0.6800	1.0485	1.3006	1.6794	2.0141	2.4121	2.6896	3.2815	3.5203
50	0.2547	0.6794	1.0473	1.2987	1.6759	2.0086	2.4033	2.6778	3.2614	3.4960
60	0.2545	0.6786	1.0455	1.2958	1.6706	2.0003	2.3901	2.6603	3.2317	3.4602
70	0.2543	0.6780	1.0442	1.2938	1.6669	1.9944	2.3808	2.6479	3.2108	3.4350
80	0.2542	0.6776	1.0432	1.2922	1.6641	1.9901	2.3739	2.6387	3.1953	3.4163
90	0.2541	0.6772	1.0424	1.2910	1.6620	1.9867	2.3685	2.6316	3.1833	3.4019
100	0.2540	0.6770	1.0418	1.2901	1.6602	1.9840	2.3642	2.6259	3.1737	3.3905
120	0.2539	0.6765	1.0409	1.2886	1.6577	1.9799	2.3578	2.6174	3.1595	3.3735
150	0.2538	0.6761	1.0400	1.2872	1.6551	1.9759	2.3515	2.6090	3.1455	3.3566
200	0.2537	0.6757	1.0391	1.2858	1.6525	1.9719	2.3451	2.6006	3.1315	3.3398
300	0.2536	0.6753	1.0382	1.2844	1.6499	1.9679	2.3388	2.5923	3.1176	3.3233
∞	0.2533	0.6745	1.0364	1.2816	1.6449	1.9600	2.3263	2.5758	3.0902	3.2905

Source: Originally Biometrika Tables for Statisticians (1966), volume 1, 3rd ed. Reprinted in Brooks (2008).