THE INTRADAY LEAD-LAG RELATIONSHIP OF SPOT AND FUTURES MARKETS IN TURKEY: CO-INTEGRATION AND CAUSALITY ANALYSES

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF APPLIED MATHEMATICS
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

NEŞE ABUK

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
FINANCIAL MATHEMATICS

MAY 2011

Approval of the thesis:

THE INTRADAY LEAD-LAG RELATIONSHIP OF SPOT AND FUTURES MARKETS IN TURKEY: CO-INTEGRATION AND CAUSALITY ANALYSES

submitted by **NEŞE ABUK** in partial fulfillment of the requirements for the degree of **Master of Science in Department of Financial Mathematics**, **Middle East Technical University** by,

Prof. Dr. Ersan Akyıldız ____________
Director, Graduate School of **Applied Mathematics**

Assoc. Prof. Dr. Ömür Uğur ____________
Head of Department, **Financial Mathematics**

Assist. Prof. Dr. Seza Danışoğlu ____________
Supervisor, **Department of Business Administration, METU**

Examining Committee Members:

Assoc. Prof. Dr. Nuray Güner ____________
Department of Business Administration, METU

Assist. Prof. Dr. Seza Danışoğlu ____________
Department of Business Administration, METU

Dr. Hande Ayaydın Hacıömeroğlu ____________
Department of Business Administration, METU

Date: 26.05.2011

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last name: NEŞE ABUK

Signature :

ABSTRACT

THE INTRADAY LEAD-LAG RELATIONSHIP OF SPOT AND FUTURES MARKETS IN TURKEY: CO-INTEGRATION AND CAUSALITY ANALYSES

Abuk, Neşe
M.Sc., Department of Financial Mathematics
Supervisor : Assist. Prof. Dr. Seza Danışoğlu

May 2011, 72 pages

This study is concerned with the lead-lag relationship between Turkish spot equity and derivatives markets. In the study, the spot equity market is represented by the ISE-30 Index. In order to compare the structure of the two markets, the futures contract written on the ISE-30 Index, namely TURKDEX-ISE 30, is chosen to represent the derivatives market. The analysis is performed over the sample period beginning February 4, 2005 and ending on December 10, 2010 which actually covers the entire time span from the establishment of the TURKDEX market until the end of last year. This sample period is examined on the basis of 5-minute intervals during the trading day, enabling a more detailed and accurate evaluation of the lead-lag power of the markets. The main methods applied to examine the structure of information flow between the markets are co-integration and causality analyses. Different approaches of these basic methods are employed as well in order to provide robust results. An additional robustness check is provided through examining the relationship between the markets by using both raw and filtered prices. ARMA filtering is performed on the prices and these findings are compared to those obtained by raw prices in order to avoid the problem of infrequent trading. Outcomes of both raw and filtered price analyses reveal that in 2006, 2007 and 2009 the relationship between the markets is bi-directional, whereas in 2008 and 2010, futures market strictly leads the spot market. Filtered and raw analyses do not have a definitive conclusion regarding the lead-lag relationship in 2005. For this year, while the raw data support a bi-directional relationship, ARMA filtering indicates that the spot market leads the derivatives market.

Keywords: Lead-lag Relationship, Turkish Derivatives Market, Infrequent Trading, Co-integration, Granger Causality

ÖZ

TÜRKİYE NAKİT VE TÜREV ÜRÜN PİYASALARININ GÜN İÇİ ‘LEAD-LAG’ İLİŞKİSİ: EŞBÜTÜNLEŞME VE NEDENSELLİK ANALİZLERİ

Abuk, Neşe

Yüksek Lisans, Finansal Matematik Bölümü

Tez Yöneticisi : Yrd. Doç. Dr. Seza Danışoğlu

Mayıs 2011, 72 sayfa

Bu çalışma Türkiye’deki nakit ve türev ürün piyasaları arasındaki lead-lag ilişkisi ile ilgilidir. Çalışmamızda, nakit piyasa IMKB 30 Endeks fiyatları ile temsil edilmektedir. Söz konusu iki marketi karşılaştırabilmek için IMKB 30 Endeksi ürezine yazılan vadeli işlem sözleşmesi (VOB IMKB 30 vadeli işlem sözleşmesi) türev piyasayı temsil etmek üzere seçilmiştir. Araştırma, 4 Şubat 2005’ten başlayıp 10 Aralık 2010’da biten, yani VOB’un kuruluşundan geçtiğimiz senenin sonuna kadar süren geniş kapsamlı bir veri seti üzerinden yapılmaktadır. Dahası, tüm bu zaman dilimi lead-lag yapı hakkında daha detaylı ve kesin sonuçlara imkan tanıyan 5 dakikalık aralıklarla incelenmiştir. Piyasalar arası bilgi akışını anlamak için başvurulan temel yöntemler eşbütünleşme ve nedensellik analizleridir. Ancak, sağlam sonuçlar elde etmek için bu yöntemlerin farklı yaklaşımları kullanılmıştır. Bir diğer denetim hem ham hem de filtrelenmiş verileri analiz ederek sağlanmıştır. Seyrek iş hacmi sorunundan kaçınmak amacıyla ARMA filtrelemesi kullanılmış ve sonuçları ham veriden elde edilen sonuçlarla kıyaslanmıştır. Hem ham hem filtrelenmiş analiz neticesinde 2006, 2007 ve 2009 yıllarında iki piyasa arasındaki ilişki çift yönlüdür, halbuki 2008 ve 2010 yıllarına ait bulgular türev piyasanın tek yönlü olarak nakit piyasaya öncülük ettiğini göstermektedir. Ne var ki, veri analizleri 2005 yılı sonuçları konusunda ortak bir kanıya varamamıştır: Ham veri çift yönlü ilişki öngörürken ARMA filtreli veri nakit piyasa liderliğini işaret etmektedir.

Anahtar kelimeler: Lead-lag İlişkisi, Türk Türev Piyasası, Seyrek İş Hacmi, Eşbütünleşme, Granger Nedensellik Analizi

To my family and my love for their unconditional love…

ACKNOWLEDGMENTS

Firstly, I would like to thank my supervisor Assist. Prof. Dr. Seza Danışoğlu for patiently guiding and encouraging me throughout this study. I am also very thankful for her compassionate and friendly behavior towards me.

I would also like to thank Prof Dr. Nuray Güner and Dr. Hande Ayaydın Hacıömeroğlu for allocating their valuable time and effort for reviewing my thesis.

I would like to thank The Scientific and Technological Research Council of Turkey (TÜBİTAK) for the financial support during my graduate education.

In addition, I would like to thank a number of special people:

I render thanks to my colleagues at SGK for their close friendship, understanding and support.

Special thanks need to be given to Ela Uysal, Nazlı Sarıaslan, Nesrin Mevsim, Özlem Yalçın, Şirin Ocaklı and Tuğba Erdem for their sincere and cheerful friendship.

I would like to express my heartfelt thanks to Göksel Tüker for being with me all the time, guiding me wisdomly and sharing hard times. Without his love and assistance, I could barely complete this work.

Finally, I want to express my sincere gratitude to my family for their unconditional love, endless patience and complete reliance. My beautiful mother Sevgi and my admirable father Kaşif make me believe in me and my potential. Every goal I achieved is by courtesy of them.

TABLE OF CONTENTS

LIST OF TABLES

TABLES

LIST OF FIGURES

FIGURES

CHAPTER 1

INTRODUCTION

1.1 Motivation

Brooks (1999) states that 'in a perfect capital market with non-stochastic interest rates and dividend yields' prices of the derivative and the corresponding underlying ought to be completely correlated which brings the result that there should not exist a leading or a lagging market. That is to say in an ideal efficient market, incoming news should be reflected to both markets simultaneously, not letting any delays in either one of the markets. However, many studies show that practice does not concur with the theory. Due to some market imperfections, derivatives market seems to lead the cash market. This situation is mostly explained by the effects of asynchronous trading or lower transaction costs. Asynchronous trading may mask the efficiency of the markets and pretends as if futures prices lead spot prices. Lower transaction cost is the most popular justification of the lead/lag relation which first introduced by Fleming (1996) going under the name of 'transaction cost hypothesis'. Researchers claim that cheaper trading costs in derivatives market make it charming to investors. Consequently, when new information arrive, investors first trade in derivatives market which bring about the price leadership of derivatives prices or returns. On the other hand, against the possibility that the leadership is not real but seem to occur favoring the derivatives, the analysis should be handled taking the effects of asynchronous trading into account. If derivatives insist on its role as a price discovery tool even when we consider the effects of asynchronous trading, it should be accepted that one market significantly leads/lags the other which makes it possible for the investors to alter their hedging strategies and takes positions to maximize profits. Hence, when possible gainings are thought the topic is really worth the efforts.

1.2 Aim of the Study

The aim of this study is mainly to identify the price discovery structure of Turkish Spot and Futures Markets. In other words, this study aims to provide precious information about how to hedge their saving or even help them to make use of the speed of the information flow between markets. Another important aim is comparing the obtained results with literature and find countries with a similar market structure. Moreover, this study answers whether one of these markets strictly lead the other, or not revealing the trading habits of Turkish traders.

1.3 Contributions of the Thesis

In literature there exist many studies which examine the lead-lag pattern of spot and derivative markets. From the time that derivatives market transactions become popular, traders fell the need for understanding the price discovery process. The reason for the efforts on price leadership: profit. It is for sure that once a trader can get a clue about the time delays of information between markets, it enables him to profit from this foreseen price interaction. Even trader cannot have the chance to profit, it at least provides traders with protection on their savings.

The inferences from this study may guide the investors since there are a few papers investigating the lead/lag relation of Turkish markets and also this study have some distinguishing properties of the previous studies conducted in Turkey. Firstly, the analysis is done through a wide time period starting from the establishment of the derivatives market (VOB) until last year. This wide period helps us understand the development of the Turkish market through years. Moreover, the price discovery process is examined year by year in order to check the robustness of the lead/lag decision. It also enables us to see whether the structure in the early years of market differs from the recent years or not. Another innovation is the time span chosen. The early papers work on the daily values of VOB and IMKB prices, however, the results of this study will be reported based on intervals of 5-minute providing more explanatory results. The last extra property is about one of the methods implemented. Co-integration and Granger linear causality are put into practice by the authors of the previous papers but non-linear Granger causality will be applied to Turkish data set in the aim of price discovery for the first time. In conclusion, considering the improvements, this thesis study may contribute to the finance literature and may guide the Turkish investors to take better positions and also allows foreign investors understand the market structure of Turkey.

1.4 Scope of the Thesis

This thesis on lead-lag relationship between spot and derivatives markets in Turkey is organized in seven chapters. The contents of these chapters are summarized below:

Chapter 1 This chapter basicly introduces the study. Motivation, aims and possible contributions of this thesis study are described in Introduction Chapter.

Chapter 2 The general market structure in Turkey is explained in this chapter. The two markets Istanbul Stock Exchange (ISE) and Turkish Derivatives Exchange (TURKDEX) are introduced.

Chapter 3 Previous works and efforts on the lead-lag relationship are introduced. Foreign and Turkish literature examined in a detailed manner.

Chapter 4 The dataset is described in this chapter. Neccessary adjustments are performed to be ready for the analysis. Moreover, preliminary statistics are presented.

Chapter 5 The methods and different applications of these methods are represented. Important linkages between the methods are noted.

Chapter 6 This chapter is composed of the empirical results of the study.

Chapter 7 In this last chapter, the main points are summarized briefly and the study is concluded.

CHAPTER 2

MARKET STRUCTURE IN TURKEY

The first exchange market activity in the world is accepted as the commerce of the precious metals. In time, trading precious metals has grown by means of involving agents and become a market in which commercial papers are traded. The first market in Europe was the one established in Anvers in 1487. After that, markets in Amsterdam, Lyon and London were established. However, the first real stock exchange market was founded in London in 1801 and it took until 1875 that London Stock Exchange becomes a legal institution.

New York Stock Exchange, the largest market today, was founded on March 8, 1817. New York Stock Exchange was the prototype for all of its subsequent American and Canadian markets.

Word's markets are mainly categorized as developed and emerging markets. Developed markets are the ones functioning in industrialized countries which constitute international financial markets. Investors mainly think that developed markets are less risky because of the political and economic stability. On the other hand, emerging markets are not steady as developed markets but they show a significant progress or they just have the admirable potential compared to small markets. Some examples of developed markets are America, England, Japan, German, France, Hong Kong, Holland and Denmark. Countries like China, Brazil, India and Egypt are fall under the emerging markets category. Turkey is also an emerging market with its rapidly growing trading volume.

In a general form, markets are divided into two main categories, namely spot and derivatives markets, according to differences in delivery times and price determination dates. In spot markets, trading occurs immediately. In other words, the purchaser makes the payment and the seller delivers the instrument at the same time. Nonetheless, in derivatives markets, parties agree on trading at a today-determined price in a specified future date. In Turkey, both spot and derivatives markets are functioning. Spot market is Istanbul Stock Exchange (ISE) which is founded in 1985 and the derivatives transactions are done through TURKDEX.

2.1 Istanbul Stock Exchange (ISE)

Stock market activities in Turkey dates back to second half of the 19th century. The first market formation was an unofficial one named as the Galata Stock Market. Then in 1871, semi-official 'Dersaadet Tahvilat Borsası' was founded. In 1929, official market is established with the name of 'İstanbul Menkul Kıymetler Kambiyo Borsası' and after displaying activity for over 120 years, at January 3, 1986 'İstanbul Menkul Kıymetler Borsası – İstanbul Stock Exchange (ISE)'is established. At the beginning, ISE had only 40 companies contributed. While trading volume was $50000 initially, today trading volume has reached $2 billion with stocks of 337 companies trading. In 2010, Turkey ranked third among the members of World Federation of Exchanges. Since its inception, ISE designated as being the cheapest, the most profitable and the most easily accessible market. Moreover, ISE is acknowledged to be 'investible foreign market' by US Securities & Exchange Commission (SEC), Japan Securities Dealers Association (JSDA) and Austria Ministry of Finance. In addition, ISE is accepted to be a member of some international federations:

- The World Federation of Exchanges (WFE),
- Federation of Euro-Asian Stock Exchange (FEAS),
- Federation of European Securities Exchanges (FESE),
- International Securities Services Association (ISSA),
- International Capital Market Association (ICMA),
- European Capital Markets Institute (ECMI),
- International Organizations of Securities Commissions (IOSCO).

The basic products trading in ISE markets are stocks, bonds, exchange traded funds and warrants. Besides, stock indices are formed in order that the investors can easily keep track of the joint movements of market products. 13 stock indices are traded in ISE market:

Table 2.1 Indices traded in ISE

ISE Indices	**Explanation**
ISE 100 Index	100 selected stocks
ISE 50 Index	50 selected stocks
ISE 30 Index	30 selected stocks
ISE 10 Bank Index	10 selected stocks of banks
ISE 100-30 Index	70 stocks of ISE 100 excluding ISE 30 stocks
ISE Corporate Governance Index	stocks with min. required corp. govern. rate
ISE All Index	all ISE stocks except Investment Trusts
ISE All-100 Index	stocks of ISE All excluding ISE 100 stocks
Sector & Subsector Indices	selected ISE stocks except Investment Trusts
ISE National Index	all ISE stocks traded in National Market
ISE Second National Index	all ISE stocks traded in Second National Market
ISE Investment Trusts	all Investment Trusts
ISE City Index	indexes categorized by location of companies

The most popular indices traded are ISE 100 index and ISE 30 index. For a stock to be included in ISE 100 and ISE 30 indexes, it must be traded at least 60 days as of the end of the basic ratio periods. Moreover, the selected stock should have a high trading volume as well as the sector-specific representation ability. In other words, these indexes are composed of the most actively traded, so they contain more information compared to other indexes. The lead/lag relationship analysis is conducted on the data of ISE 30 index. Choice of ISE 30 index over ISE 100 stems from the fact that the derivative corresponding to ISE 30 is more actively traded than that of ISE 100.

2.2 Turkish Derivatives Exchange (TURKDEX)

Until 1970s, only the derivatives written on agricultural products are traded in America. First foreign exchange derivatives and interest rate derivatives are traded in 1973. The first derivative market 'London International Financial Futures and Options Exchange' is established in 1982. Since 1982, many derivatives markets are began functioning and in 2007 trading volume of word-wide derivatives markets reached \$2.2 quadrillion.

The first private derivatives exchange in Turkey, TURKDEX, is founded in İzmir on July 4, 2002 and the first transactions of TURKDEX are started on February 4, 2005. TURKDEX has 11 shareholders as illustrated in Table 2.2:

Table 2.2 Shareholders of TURKDEX

Shareholders	Percentage
The Union of Chambers & Commodity Exchanges of Turkey	25%
Istanbul Stock Exchange	18%
Izmir Mercantile Exchange	17%
Yapı Kredi Bankası A. Ş.	6%
Akbank T. A. Ş.	6%
Vakıfbank Investment Securities	6%
Türkiye Garanti A. Ş.	6%
Is Investment Securities	6%
The Association of Capital Market Intermediary Instutitions of Turkey	6%
ISE Settlement & Custody Bank	3%
Industrial Development Bank of Turkey	1%

Derivatives transactions are getting more popular is day due to the opportunities offered. The derivatives markets provide investors with the possibility of:

- Hedging,
- Speculation,
- Arbitrage,
- Gaining the profit of big investment with a small payment,
- Profiting not just in bull markets but also bear markets,
- Taking a risk that each investor can tolerate.

TURKDEX provides the Turkish investors to buy/sell contracts on currency, index, interest rate, commodity and gold. Since we investigate the causality relation between spot and derivatives markets, the underlying of the derivative and the spot market instrument should be same in order that we can compare them. In this study, spot market is represented by ISE 30 index hence futures contract written on ISE 30 is selected to be the representative of the derivatives market.

TURKDEX-ISE 30 futures contract is the derivatives contract written on the ISE 30 national index. Properties of TURKDEX-ISE 30 futures contract are as shown in Table 2.3:

Table 2.3 Properties of TURKDEX-ISE 30

Specification	**Explanation**
Underlying Asset	ISE 30 national index
Contract Size	Index Value / 1000 * 100 TRY
Price Quotation	Index Value / 1000 quoted with 3 decimals
Daily Price Limit	+/- 15% of the base price
Contract Months	February, April, June, August, October, December
Settlement Method	Cash settlement

CHAPTER 3

REVIEW OF LITERATURE

Various studies are devoted to analyze the lead/lag relation between derivatives and the spot market. It is known that both markets react to same information, but the question of interest is which one reacts first. The general consensus reveals that futures market is the leader of both options and and spot markets with little or no feedback. However, no such consensus can be reached for the relation between options and spot markets. Below are the review of the selected literature on the lead/lag relation by different techniques, data and markets.

In 1982, Manaster and Rendleman (M&R) investigate the lead/lag relation between individual stock and stock options markets in USA covering the time period from April 1973 to June 1976, using daily closing prices. M&R attempt to discover the lead/lag pattern via the difference in observed stock prices and implied stock prices by options. They test the hypothesis that implied prices provide no information regarding the future movements of the observed stock prices. Their study results in rejecting this hypothesis and they declare that options prices lead spot prices up to one day. However they are regardless of the fact that options and spot markets do not close simultaneously, and that may cause spurious results about the leader of the market.

Correspondingly Bhattacharya (1987) copes with the same issue with some different properties from the work of MR (1982). He works on the intermarket relations for the USA stock and stock options market through June 2, 1977 to Aug. 15, 1978. During this period, he uses intraday transaction data which generates more sensible search given that most probably options and stock market changes occur in a shorter time than a day. In order to capture the lead/lag properties, he compares the actual and implied bid/ask call prices. At the end of the study, he confirms the results of MR (1982) stating that options prices are tend to lead the individual stock prices. A critical aspect of MR (1982) and Bhattacharya (1987) is that they fail to notice the effect of the stock prices on options market. In other words, by the hypotheses they test, it is questioned whether options market lead the spot market but not the other way around. Therefore, both studies are weak in detecting the real information flow process between options and spot markets.

Following the works of MR (1982) and Bhattacharya (1987), Anthony (1988) bring a new approach by examining the common stock and options markets in terms of trading volumes rather than prices. He works on the daily data from the beginning of 1982 till end of June in 1983. He choses Granger – Sims causality as a method to detect whether trading in options market causes trading in spot market or vice versa. His study eventuates in that call option trading leads stock market trading up to 1 day.

Another study focused on the lead/lag relation is suggested by Stephan and Whaley (1990). They examined the relation via Chicago Board of Options Exchange (CBOE) data of first quarter of 1986. By breaking price data into 5-minute-intervals, they aim to determine the lead/lag effect more precisely than the previous works. Before starting the analysis possible bid/ask spread price effect is purged from the price observations using a moving average (MA) process. In empirical study, implied stock price changes are patterned on a Roll-style American option formula and then price changes are computed using multivariate time series regression. They claim that, unlike previous studies, stock prices lead option prices about 20 minutes. This study causes some suspect about the studies done before since Stephan and Whaley (SW) reveal their results by intraday data and more direct methodology. Clearly, this study of SW keeps away from the two major drawbacks of the previous studies. First biases due to non simultaneity of closing prices in two markets are avoided by the use of transaction-by-transaction data. Second, the analysis directly concentrates on the lead/lag relation rather than simulated trading strategies as MR (1982) and Bhattacharya (1987).

Fleming et al. (1996) is the first researcher to examine the options, futures and spot markets interrelations together. They study price discovery process among the S&P 500 stock index, S&P 500 index futures and S&P 100 stock index and S&P 100 index options with minute-by-minute data from June 1988 to March 1991. Via multipe regression techniques they search for linear causality among markets. Results provide that both index futures and index options prices lead spot index prices on average by 5 minutes. A further important result is that index futures lead index options prices. They put a wide interpretation on these results and figure out that these lead/lag pattern may be due to transaction cost effect. Among those three narkets futures market is the less costly one due to high liquidity, low trading costs and low margin. Thus information flow from futures to other markets is expected. This hypothesis introduced by Fleming et al. (1996) called as ‘trading cost hypothesis’ and they stress that the lead relation should alter when trading cost conditions change in one of those markets.

Second research on interrelation of three markets is conducted by De Jong & Donders (1998). In order to reach more robust conclusions, they investigate two samples of data one from Jan. 20 to July 17 of 1992 and the other includes the first quarter of 1993. The data are obtained from European Options Exchange and consist of Amsterdam Stock Index (AEX), AEX index futures and AEX index options. Whole price observations are arranged to compose 5 and 10-minute-intervals. Having high frequency data with short intervals bring on missing or zero-valued observations and that causes correlation and covariance structures of the markets to be biased. De Jong & Donders (1998) avoid from nontrading problem using a new estimator developed by De Jong & Nijman (1997). With that adjusted estimator, regression and cross correlation analysis are conducted. Evidence show that index futures lead index market and index options market by 10 minutes, while there exists a bidirectional symmetric relation between options and spot markets.

In the following years, some other researchers examine the interrelation among these three markets including Booth et al. (1999); Gwilym and Buckle (2001) and Kang, Lee and Lee (2006). Those three papers generate quite different results about the structure of the markets. This contradiction may occur by reason of the fact that those papers investigate the data coming from different countries' markets impling completely different trading habits. Booth et al. (1999) deal with German DAX stock index, index futures and index options price observations for the time Dec 5, 1994 through July 11, 1997 with 15-min-intervals. To overcome nonsynchroncity problem REPLACE ALL and MINSPAN approaches developed by Harris et al.(1995) are employed. Price discovery process is analyzed by cointegration and Error Correction Models (ECM). It is found that futures market leads both options and spot markets and also spot market seems to lead the options market. Gwilym and Buckle (2001) write the first paper on intermarket relations including options market in UK. They work on hourly data of FTSE stock market index and its corresponding futures and options contracts between 1993 and 1996. Selection of hourly intervals grows out of the fact that options market does not support any higher frequency. Due to hourly data usage nontrading problem is minimized and since data is collected on quoted prices rather than transaction prices, bid/ask spread is no more problem. But nevertheless, ARMA filtering is performed to purge any forgotten effects. To detect price relation, they implement multiple regression models with error correction term acoounted for cointegration. Although their study lead up to bidirectional relations, unlike Booth et al. (1999), call options market leads futures market and futures market lead put options market strongly than the reverse. FTSE 100 derivatives altogether lead spot market. These results indicate some depatures from the Fleming et al.'s 'trading cost hypothesis' since both call and put options should have similar trading costs. Thus, they comment that there must be some other factors driving the intermarket price relation rather than trading costs.

Korean KOSPI 200 spot, futures and options markets investigation for the last quarter of 2001 through 2002 is provided by Kang, Lee and Lee (2006). The study contends with lead/lag relations of not only the prices/returns but also the volatilities. They implement multivariate time series models presented by Stephan and Whaley (1990) with a modification in derivation of implied prices. Stephan and Whaley compute the implied prices according to American formula however Kang, Lee and Lee prefer to use put-call parity which bring the advantage of model-free approach. In conclusion, Korean futures and options markets lead spot market by 10 minutes in returns and 5 minutes in volatilities. Although analysis on returns reveals that options lag futures, in terms of volatilities no lead/lag relation is detected.

Numerous studies are performed to analyze the link between futures and the corresponding spot market. Looking at the literature, those studies can be distinguished upon the methodology selected. Starting from the 1990's, cointegration techniques become a trend investigating the financial time series data. Lead/lag detection and price discovery studies follow the trend rapidly since this new method allows the relation to be evaluated based on short-run and long-run deviations separately. Before cointegration, studies are conducted using multiple regression and correlation analysis with different adjustment techniques. Although the methodologies are various as well as the research countries, the wast majority of the results point out that futures market is tend to lead the spot market with no or little feedback.

The first efforts on the temporal relation between index futures and the cash index markets are by Finnerty & Park (1987) and Kawaller et al. (1987). Both studies explore intraday USA data. Finnerty and Park (1987) examine MMI & MMMI cash and futures prices over two-year period beginning from Aug. 1984. Minute-by-minute S&P500 index and index futures prices between 1984 and 1985 are delved by Kawaller et al.(1987). Both Finnerty & Park and Kawaller et al. reach to the conclusion that futures market leads the spot market. In Kawaller et al. (1987), it is found that index futures prices lead cash index prices up to 40 minutes althoug cash index of only one minute is observed at times. To reach this conclusion they perform 3-stage LS regression along with Granger-Sims causality. But the miss the fact that by minute-by-minute data lead of futures market may be caused by infrequent trading problem.

In 1987, Herbst et al. conduct a study on daily closing prices of Value Line spot index and its 4 futures contracts. Spectral analysis reveals that futures are tend to lead cash index less than a day. In order to validate the study and to put boundaries on the lagging time, they also investigate intraday tick-by-tick data on VL index futures contract and S&P 500 index. With the new data previous results are validated and it is specified that lead time is actually up to 16 minutes.

Harris (1989) examines 5-minute changes of S&P 500 index and futures contract prices during October 1987 stock market crash. In the times of crash, 5-minute intervals contain large number of mising values. To solve this infrequent trading problem, he derives new estimators. Outcomes show that even after the removal of the infrequent trading effects, strong lead occurs from futures to cash.

Stoll and Whaley (1990) suggest that during the five years period from 1982 till 1987 S&P 500 and MM index futures have the leadership by 5 minutes on average, but there is weak evidence that spot market leads the futures market. Strong leadership of the index futures market exists after they take out the nontrading and bid/ask spread effects. These microstructural problems are handled via ARMA filtering. Following the adjustment procedure, unidirectional relation is found using multiple regression method similar to Sims (1972) with additional regressor of contemporaneous variable.

S&P 500 index and index futures from Aug. 1, 1984 through the end of 1989 and MM index and futures for one year period beginning in July 1984 are used in the empirical study of Chan, Chan and Karolyi (1991). They not only work on the returns but also the return volatilities by auto-cross correlation analysis and bivariate GARCH models respectively. They suggest that futures returns lead spot returns by 5 minutes, with strong intermarket dependence in volatility. Similar study is performed by Min and Najand (1999) on Korean market. They construct 10-minute intervals of KOSPI index and its nearby futures contracts over the third quarter of 1996. Temporal relation is detected by Granger causality test using Simultaneous Equations method and VAR analysis. Parallel to foundings of Chan (1991), this study reveals that futures market leads spot market up to 30 minutes in returns, but regarding volatility, two markets show strong bidirectional causality relation.

A further inquiry on both temporal return and volatility relations of futures and spot markets is accomplished by Iihara, Takunaga and Kato (1996) in Japan with NSA index and its futures contracts. The data set covers the time between March 1, 1989 and Feb. 26, 1991 with 5-minute intervals of transaction prices. Choice of time period is decent because it composes of both bull (1989) and bear (1990, 1991) market times. Before modeling the temporal relation AR model is used to purge the drawbacks caused by infrequent trading. After adjustment, multivariate regression is performed bringing about that futures market leads the spot market by about 20 minutes.

Silvapulle and Moosa (1999) contribute to literature by a work on Crude Oil Market. Daily price information examined to see the pattern between Crude Oil spot and futures prices. Linear and nonlinear causality tests for returns and EGARCH model for volatilities conclude that Crude Oil spot and futures market show bidirectional association but with stronger lead of futures market on spot market.

An elaborate work on on the lead/lag relation is introduced by Chan (1992). He uses S&P 500 and MM index and index futures data from Aug. 1984 to June 1985 and Jan. to Sept. 1987 by means of 5-minute intervals of trading days. The second time period covering 9 months of 1987 is used to verify the robustness of the relation. He also stratified the observations to detect the changes of lead/lag structure under good news vs bad news; changing market information and different intensity levels of trading activity.To examine the temporal link, a linear regression model is proposed close to one implemented by Stoll and Whaley (1990). But for this model, he proposes that error terms of the model are probably time-varying heteroskedastic due to previous evidence by Chan et al. (1991) supporting that 'volatilities in the two markets are not only time-varying, but also related.' In the study of Chan (1992) the dynamics regarding volatilities are not modeled expilicitly but heteroskedasticity problem owing to related error terms are handled by adjusting t-ratios and estimates according to Hansen's Variance-Covariance Matrix. This new approach is employed to MM and S&P 500 cash index and index futures returns as well as all component stocks of MM index. Results reveal that there exists an asymmetric lead/lag relation. Although futures returns lead cash returns strongly, the relation is not completely unidirectional. Moreover, findings suggest that under bad news, futures market loses the lead effect on spot market. It is clearly stated that when the number of stocks moving together increases providing wide market information, spot market lags futures market. In terms of intensity of relative trading, no such evidence found referring lead or lag pattern.

De Jong and Nijman (1997) propose a new method to calculate the correlations and covariances from irregular observations of price data. This method removes the bias in correlations induced by imputation techniques in case of nontrading problem. In the empirical study, they work with S&P 500 index and futures prices with 1 and 5-minute intervals. Since their new estimator is applied to intervals without price observations, any higher frequency than 5 minutes can easily be investigated without adding any bias. Results of the study are consistent with the previous work. The lead/lag relation is found to be bidirectional with stronger evidence in behalf of futures. Index futures prices lead cash index prices by 10 minutes; whereas spot prices lead futures prices by at most 2 minutes.

High frequency data of FTSE market are examined for the first time by Abhyankar (1998). He studies FTSE 100 index and index futures prices during 1992. In his paper, temporal structure is investigated through two diverse methods: linear and nonlinear causality. Before performing those methods, the data are filtered by ARMA model to get rid of the nontrading bias as Stoll and Whaley (1990) suggested. Along with ARMA, EGARCH filtering is also performed to be able to catch the neglected nonstationarities taking root from heteroskedasticity. Then multivariate time series regression is carried out. Linear tests result in futures market leadership by 5 to 15 minutes. Once he performs linear causality tests, peresence of nonlinear causality is examined using adjusted Baek & Brock (1992) test. Opposite to linear test indications, nonlinear causality implies that no lead/lag connection between the returns of cash index and index futures in UK is remained.

Hasan (2005) examines daily FTSE 100 data for UK and S&P 500 data for USA. Cross correlation and Cross bidirectional analysis findings indicate bidirectional relation in returns.

In a very recent paper by Tse and Chan (2009) temporal relation between cash index and index futures in USA is reexamined using Treshold regression model (TRM). The choice of TRM is supported as ‘The TRM enables us to capture the lead/lag relation under different scenarios or market conditions, which determine different linear regression regimes.’ Factors composing the different market conditions are proxied by treshold variables. In the study of Tse and Chan, three treshold variables are used to represent effect of short selling, market wide information and good vs bad news condition in the market. This method is employed to S&P 500 index and index futures for March through July 2004 with 3-minute intervals data. They realize that short selling restrictions reduce the lead effect of spot market on futures. In terms of increasing market wide information, futures market leading spot market becomes stronger. Lastly under good or bad news futures have a tendency to lead spot market.

Finance literature is rich in the methods performed to investigate the lead/lag structure between derivatives and their underlying stocks. Up to middle of 1990s, numerous methods are applied such as multivariate time series regression, spectral analysis, correlation analysis, causality tests. However, from 1993, cointegration together with ECM, become a new trend in intermarket relation investigation. It is pointed out that cointegration is superior to other methods in that this method is able to differentiate short-run and long-run deviations easily. With the two papers written in 1993 by Wahab&Lashgari and Antoniou& Garrett cointegration is introduced to the literature of lead/lag relation history.

Wahab & Lashgari (1993) investigated daily closing prices of USA and UK index and index futures markets between 1988 and 1992. The new method, cointegration and ECM resulted in that spot market lead on futures market is strong than the vice versa, as oppose to the most of the literature suggesting that leadership of the futures is more pronounced than that of spot market. Similarly Pradhan and Bhat (2009) discover that spot prices are likely to lead futures prices and price discovery occurs in spot market more rapidly than the futures market. This conclusion is reached by analysing daily observations of India's Nifty index futures and spot index between 2000 and 2007 through Johansen cointegration model.

In 1993, Antoniou & Garrett investigated the two days (October 19, 20) of October 1987 stock market crash period with minute-by-minute data for FTSE100 index and futures prices. Engle-Granger cointegration method with error correction representation of VAR is employed yielding that on October 19, although futures lead index, the relation is not completely unidirectional. They find weak evidence that spot market leads futures market on that day. However on October 20, the lead/lag relation is turned out to be completely unidirectional in favor of the futures market.

After 1995, many futures-spot lead/lag interaction studies are emerged from the markets of different countries such as Spain, Greece, Korea, Taiwan, and France. Nieto et al. (1998) work on the daily observations from Spanish Stock index IBEX 35 and its futures contract. The sample data covers the period March 1, 1994 – Sept. 30, 1996. Johansen cointegration with VAR representation is employed as a method. Empirical findings of the study reveal that futures prices lead spot prices in short run. However, in long run no lead/lag pattern is observed indicating that market is efficient as theory desires.

Another study by Mattos and Garcia (2004) explored the relation in Brazil with daily data of 1997 till 2001. They used the exact same methods with Nieto et al. (1998), but they conclude that in short run no lead/lag structure is present. Whereas, long run analysis signalizes that futures market leads spot market in a tough manner.

Pizzi et al. (1998) examined USA market with cointegration techniques between Jan. and March 1987. S&P 500 stock index and the three months and six months futures contracts with minute-by-minute price informations are documented. Empirical study denotes that while the futures market tends to have a stout lead effect, unidirectional lead/lag pattern is refuted. Their paper was the first in literature as it implements cointegration with a much finer grid in USA data.

Similar to Pizzi et al. (1998), Alphonse (2000); Chung & Chuang (2003); Raju & karande (2003); Kenourgious (2004) and Kavussanos (2008); Ryoo & Smith (2004) and Florous & Vougas (2007) carried out cointegration with ECM bringing about bidirectional lead/lag relationship supporting futures market lead. In his survey, Alphonse (2000) explores French index and futures market interrelation for the first four months of 1995. Transaction prices are organized to create intervals of 5- minute length. The experimental facts derived from CAC 40 index and index futures are in harmony with Pizzi et al. (1998) telling futures- dominant bidirectional relation.

An alternative evidence of two-sided pattern evolves in Taiwan market by Chung and Chuang (2003). They attempt to clarify the price discovery and volatility spillover processes among MSCI and TAISEX index and index futures contrats with daily prices data. Intermarket conditions of returns are assessed via cointegration and volatility effects are identified by EGARCH error correction model which ends up with the decision of bidirectional relation.

In a working paper of Raju and Karande (2003), like Chung and Chuang (2003), both price discovery and volatility relations are discovered by cointegration and GARCH analysis respectively. The data period covers the daily prices of Indian Nifty cah index and index futures from June, 2000 till October, 2002. The major foundings of the study are that there is bidirectional feedback between the markets and high frequency data should be analyzed to detect any lead/lag pattern in a more robust manner.

Kenourgious (2004) and Kavussanos et al. (2008) deal with information linkage between Greek derivatives and spot markets by means of daily observations. Kenourgious (2004) examines FTSE/ASE-20 spot and futures prices for the duration of 1999 – 2002, while Kavussanos et al. (2008) deal with FTSE/ATHEX-20 and FTSE/ATHEX-Mid40 index futures and their underlying indices for 2000 through 2003. Both researchers execute cointegration and ECM analysis producing bidirectional findings.

Futures market leadership is propped up in Korean markets by Ryoo and Smith (2004). Korean KOSPI 200 index and index futures prices from Sept. 1, 1993 till Dec. 28, 1998 are collected so as to discover the lead/lag connection of the two markets. The price information is reorganized to have 5-minute intervals and then modeled by cointegration with error correction structure. They declare bidirectional causality between markets. It is also stressed that robust evidence of futures-to-spot is observed as well as the weak edvidence of the reverse.

Above studies of cointegration support bidirectional relation between futures and spot markets. However, in literature, there are several papers indicating completely unidirectional link from futures market to cash market. Some examples supporting the one-sided structure emerge from various country markets such as UK, Spain, Taiwan, India, Greece, etc. Brooks et al. (2001) and Bhatia (2007) delve the lead/lag relation using intraday data. Brooks et al. (2001) explore one-year data between June 1996 and June 1997 by 10-minute periods of FTSE 100 index and mid-point quoted index futures prices. Three different methods, namely cointegration with ECM, ARMA analysis and VAR analysis are considered in practice. Empirical investigation eventuates in leadership of futures prices by about half an hour with no feedback from spot market. Bhatia (2007) deals with Indian cash index and index futures markets with intraday data. In the same way as Brooks et al. (2001), his study shows that index futures market leads the spot market strongly.

Lafuente (2002) works on the hourly data on returns and volatilites of Spain IBEX 35 index ad index futures during one-year period. In his paper, returns and volatilities are examined jointly, not separately. After he confirms that there exists a long run relation between markets, Bivariate Error Correction GARCH Model is performed. This specific model is preferred since it captures inter-market depence of returns as well as volatility cross-interactions. Results confirm that unidirectional lead/lag pattern is observed from futures to spot in returns. On the other hand, unlike returns, volatility analysis reveals a bidirectional relation between markets.

In the same year, Asche and Guttormsen put cointegration into practice to study futures on Gas Oil market. Prices from April 1981 till September 2001 are taken from International Petrolleum Exchange on a monthly basis. As a method, they employ two distinct approaches of cointegration. First, analysis is done by Engle-Granger approach specifying the possible shortcomings therein. Then Johansen Cointegration is carried and it is cited that by this method shortcomings of Engle-Granger cointegration can easily be avoided. Empirical study affirms the long-run relation between gas oil spot and futures prices. In addition, the results declare that gas oil futures market leads the gas oil spot market.

In Turkey, the relationship between spot and derivatives markets is first studied by Özen et al. (2009). Futures transactions from Izmir Derivatives Exchange (VOB) and Istanbul Stock Exchange (ISE) national 30 spot index prices are examined through co-integration and Granger causality. The research data are closing prices of 1024 days regarding the period February 4, 2005 – February 27, 2009. Their paper differentiates between the short and long run causal relationship on the basis of co-integration and then determines the direction of the relation by Granger causality applied over ECMs. Long run results indicate bidirectional causality whereas, short run deviations signal an effect from spot towards futures market. Following Özen et al. (2009), Kapusuzoğlu and Tasdemir (2010) try to explain the impact of VOB futures market on ISE national 100 index prices through market efficiency. Similar to the previous study of Özen et al. (2009), co-integration and Granger causality are performed on the daily closing prices beginning from November 1, 2005 until June 30, 2009. Empirical study reveals that both VOB derivatives and ISE spot markets are efficient in a weak form. What they find is on the contrary to the expected result of dominant futures market. Parallel to Özen et al. (2009), spot market is found to lead futures market significantly.

CHAPTER 4

DATA DESCRIPTION AND ADJUSTMENT

The aim of this thesis study is to understand the interactive behavior of derivatives and spot market based on the price information corresponding to each of these markets. Derivatives market is represented by the futures market prices. Data regarding spot market are the price series of IMKB 30 index. The data used in this study are supplied from two different sources. The requested market information is sent on the form of CD's. Futures prices are supplied from Turkish Derivatives Exchange (TURKDEX) and spot prices are provided by Istanbul Stock Exchange (ISE).

4.1 Data Properties and Preliminary Statistics

Futures transactions data files coming from TURKDEX are contained of prices between February 4, 2005 and December 10, 2010. The earlier data are not available since there are no VOB transactions until February 4. TURKDEX data set provide us with trade date, trade time, security type, security name, price and quantity. Trade date corresponds to the date that transaction is occurred and trade time is the exact time of that transaction. Security type shows which index futures is traded. More clearly, security type states whether the futures contracts are written on ISE 30 or ISE 100 index. Security name reveals the maturity date of the futures contract. Price simply represents the money level that the transaction is occurred and quantity is the number of futures contracts in transaction. The price is quoted at every second that a new transaction is occurred.

ISE data are available beginning from January 5, 1998 up to December 10, 2010. ISE 30 index data is in a simpler form compared to futures prices. ISE 30 index files contain the information about trade date, trade time, session and price of the index. Trade date, trade time and price have the same meaning with that of futures information. Session indicates the part of the day (morning or afternoon) that the transaction takes place.

To be able to investigate the lead/lag relation of futures and spot prices a number of necessary adjustments should be made. The basic modification is about the irregularly spaced observations. Price observations are recorded at the time that a new transaction occurs together with the exact time of that transaction. Hence, the data is in a very irregular form. However, unless prices are defined over the same period, the econometric analysis cannot be performed. Therefore, to cope with this problem both futures and spot prices must be re-defined in an interval basis. The selection of interval length is an important issue since if interval is too short; some intervals may cover no information. On the other hand, long intervals make it hard to identify precision of the lead/lag pattern. Some researchers use daily data in their studies like Antony (1998), Wahab and Lashgari (1993) and Silvapulle and Moosa (1999). Their results indicate a one day lead or lag between markets but they are criticized in that this result is not very informative since for the market a day is a very long time period. During a day time, most probably there are some periods that one market leads (lags) the other. Thus to reach a deeper information, shorter time periods should be defined. Antoniou and Garret (1993), Pizzi, Economopoulos and O'niell (1998) and Savor (2009) work with 1-minute data. Results of the analysis on 1-minute data are informative but data frequency should be carefully examined to decide on 1-minute intervals. For our study, derivatives market prices are infrequent in years 2005 and 2006 since those years where the first years of derivatives market transactions. Stephan and Whaley (1990), Abhyankar (1998) and Ryoo and Smith (2004) are just some of the authors that divide the data into 5-minute intervals. Joining to the majority, in this study, 5-minute intervals will be formed. 5-minute is a fine grid to identify the leading/lagging time. Also, infrequent trading problem is avoided since interval is not too short. Moreover, 5-minute intervals provide us with the opportunity of comparison. Put another way, because of the large numbers of papers using 5-minute intervals, it is possible to compare the results of this study with the literature.

Another point to decide is the interval price. When constructing intervals, many price observations will be reduced to one observation. For example, assume 100 transactions of futures contracts occurred between 14:55 and 14:59. The 5-minute interval (14:55-14:59), should be represented by just one value. The decision of which value to use is significant. There are many candidates of the representative values. It may be mean, mode, median or the first or the last observation in the interval. In this point literature guides us. Nearly all of the papers select the last observation of the interval as the price that characterizes the corresponding 5-minute interval since it contains more information compared to the preceding. ISE 30 observations can easily be converted into intervals. However, this is not such simple for futures data. As mentioned above, one of the variables regarding futures market is the 'security name'. Security name shows the maturity of the contract, or more clearly different names mean different contracts. Prices of different contracts cannot be used in the analysis because they contain different information. To avoid possible problems resulting from usage of different contracts, nearest contract is to be practiced because the nearest contract is the one highly transacted. Thus the nearest contract has more information due to its high trading volume. In TURKDEX, nearest three contracts of months February, April, June, August, October and December can be dealt with at the same time. To illustrate, in May, contracts that mature in June, August and October can be transacted. For this example the nearest contract is the June contract. Only prices of the June contract are necessary to conduct the analysis.

There is one more need to be met for the data to be ready for the econometric analysis. The causality relation can be examined as long as both spot and futures prices are synchronized. Spot and futures prices should have synchronous time intervals so that we can compare them through the investigation. But unfortunately, opening and closing hours of ISE and TURKDEX do not match with each other. Moreover, during the past years, trading hours of ISE and TURKDEX are extended by changing opening and closing times. To be able to start the econometric process it is inevitable to come up with a solution on these timing problems.

From the beginning of the spot market, ISE has changed normal trading hours for two times. In one of these changes, the trading hours of the second session is extended and in the other change both morning and afternoon session become longer as illustrated below Table 4.1:

Table 4.1 Session Hours of ISE

	Morning Session	**Afternoon Session**
until 07.09.2007	09:30-12:00	14:00-16:30
07.09.2007-1810.2009	09:30-12:00	14:00-17:00
after 19.10.2009	09:30-12:30	14:00-17:30

From the time that TURKDEX is established its trading hours are altered for six times. The biggest renewal is the ‘no session break’ application that put into practice in 2008. Development of trading hours in TURKDEX is summarized in Table 4.2:

Table 4.2 Session Hours of TURKDEX

	Morning Session	**Afternoon Session**
until 19.12.2005	10:00-12:00	13:00-15:00
19.12.2005-19.03.2006	09:15-12:00	13:00-16:00
20.30.2006-06.09.2007	09:15-12:00	13:00-16:40
07.09.2007-23.03.2008	09:30-12:00	13:00-17:10
24.03.2008-12.10.2008	09:30	17:10
13.10.2008-15.10.2009	09:15	17:15
after 16.10.2009	09:15	17:35

While constructing synchronous intervals for ISE and TURKDEX prices, some intervals will be removed from the beginning and end. The reason is that at the beginning and at the end, prices become more volatile showing a different pattern from the rest of the series. In addition to this, early morning prices undertake overnight effects. If investors overreact to news released evening, this reaction will be reflected to early morning prices. Similarly, the last prices of the day are better to be deleted because of possible reaction sales. Due to mentioned complications, inclusion of early morning and late afternoon prices may yield wrong results about causality. To tackle these problems, data will be fixed by disposing some intervals so that ISE 30 index and index futures have synchronized intervals.

After the required adjustments, number of working days and number of data corresponding to each year is computed as shown in Table 4.3:

Table 4.3 Yearly Working Days and Observations

	Number of Working Days	**Number of Observations**
2005	232	5132
2006	250	13039
2007	238	13301
2008	251	14869
2009	238	14186
2010	235	13259
Total	1444	73786

Naturally, 2005 is the year with least observations since it is the establishment year of TURKDEX. In the first year of TURKDEX, the derivatives transactions can only be performed between 10:00 and 14:00 thus the number of observations are quite low when compared to other years.

In order to understand the general structure of the price series below descriptive statistics tables are prepared as in Table 4.4.

For both spot and futures prices means increase in 2006 and 2007. Then in 2008 and 2009 means fall down drastically. In 2010, means of futures and index prices pass beyond the means of the previous years. Considering the extreme values of the data, in 2005 and in 2008, the range is quite large. The tables give the idea that in 2005 prices begin from low values and move upward significantly. Conversely, starting prices of 2008 are quite high when compared to the rest of the year. Kurtosis and skewness give idea about the shape of the price series. For 2005, 2006, 2007, 2009 and 2010 the prices are leptokurtic, in other words prices have a sharp peak and fat tails. However, 2008 spot and futures prices are platykurtic meaning that prices have a wide peak around the mean and thin tails. All price observations show a right skewed pattern. Apart from that J-B test result will reveal whether the observations follow a normal distribution or not. According to corresponding p-values, none of the prices are normally distributed.

Table 4.4 Descriptive Statistics of Spot Index and Index Futures Prices

Index Futures						
	2005	2006	2007	2008	2009	2010
Mean	38832,51	49575,29	60627,34	47551,97	46573,68	74239,59
Median	37340	48575	58575	49575	44900	72850
Maximum	50650	61250	75675	71300	65625	92300
Minumum	29700	39700	54175	26850	28450	59850
Std. Deviation	5975,16	4763,11	7564,06	9220,92	12035,41	774,08
Kurtosis	0,3243	0,3040	0,2183	-0,2381	0,0635	0,6560
Skewness	1,8941	2,1287	1,7934	2,7137	1,4984	2,5303
J-B Statistic	351,46	613,30	912,50	191,30	1342,26	1072,87
p-value	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
Spot Index						
	2005	2006	2007	2008	2009	2010
Mean	39246,84	49915,12	60557,55	47171,44	46517,67	74072,14
Median	38289,32	48503,09	58450,31	49211,62	44977,53	72741,52
Maximum	50748,99	61271,95	75371,67	69984,39	65553,55	91350,99
Minumum	29412,91	39589,48	45118,43	26538,30	28649,72	59803,29
Std. Deviation	6040,38	4766,32	7284,79	9034,80	1181,22	753,70
Kurtosis	0,2386	0,3508	0,1640	-0,2900	0,0584	0,6758
Skewness	1,7948	2,1453	1,8811	2,5991	1,5104	2,5168
J-B Statistic	359,30	664,33	753,43	308,02	1313,55	1138,47
p-value	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000

After interpreting the yearly numeric statistics in Table 4.4, the overall behavior of the two market prices through 2005 until the end of 2010 are illustrated in Figure 4.1. Futures and spot prices are plotted together which helps to see whether they move in a harmony or not. Examining Figure 4.1, X-axis shows the years of the corresponding observations meaning that when the year label changes, observations belonging to that specific year are visualized until the next label change. To illustrate, observations between 2005 and 2006 correspond to values of 2005. Y-axis resembles the prices of the two market prices in terms of TL. Coloring helps to differentiate between the series as TURKDEX-ISE 30 prices are blue and ISE 30 Index prices are purple. What below figure tell is that futures and spot prices have a very similar pattern during years and there is no sign that price series deviate from each other in long-run.

After all adjustments, above summarized data series are obtained and they are named as ‘raw data’. Eventually, raw data are ready to be investigated. Nevermore, raw data will again be treated through filtering and results of filtered data analysis will be reported as well as the results of the raw data examination.

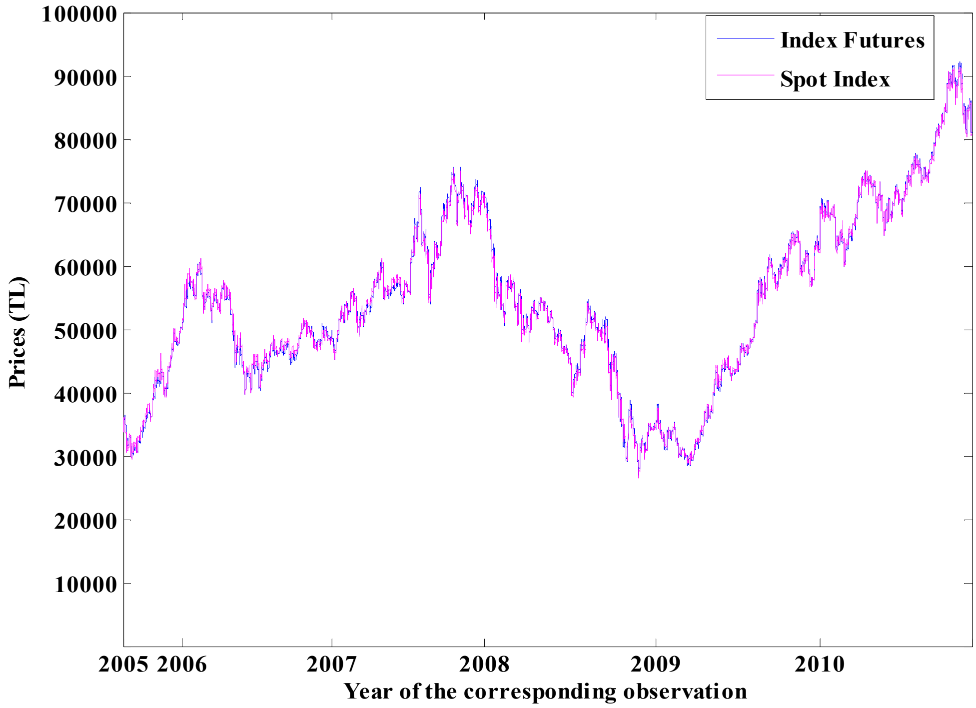

Figure 4.1 Index Futures and Spot Index Price Movements

4.2 Infrequent Trading and ARMA Filtering

In most of the studies, futures prices are found to lead the cash market but this tendency may not be caused by economic reasons. Using price transaction information makes the data prone to spurious lead/lag decision owing to possible infrequent trading effects. The individual stocks composing the index do not trade continuously, so the reported index level does not completely image the true index value. Some of the stocks contained in the index react to new information rapidly whereas; some may slowly adjust to the unexpected news. The more detailed explanation is that prices of some individuals stocks contained in ISE 30 index may change as information flows though others may stay at the same level at a specific time. Even if the price level of ISE 30 changes, it cannot truly reflect the expected value due to the unchanged individual stock prices, Stemming from this imperfection, prices behave as if spot market lags the futures.

In order to get rid of the drawbacks of infrequent trading, Stoll and Whaley (1990) suggest filtering the transaction price series with an ARMA process and using the innovation series gain by the fitted ARMA model through the rest of the statistical analysis. ARMA modeling purges the complications caused by infrequent trading. This suggestion is put into application by many researchers such as Pizzi (1998), Abhyankar (1998) and Kang, Lee and Lee (2006). Nevertheless, their empirical study does not seem to agree with the idea introduced by Stoll and Whaley

(1990). To illustrate, Pizzi (1998), Abhyankar (1998) and Kang, Lee and Lee (2006), find that raw prices of futures lead the spot prices but considering the suspect that 'effects of infrequent trading make futures appear to lead spot market', they implement ARMA filtering as suggested. Consequently, the decision about the direction of causality remains unchanged after filtering which verifies the counterview of Chan (1992). Chan (1992) states that infrequent trading cannot solely explain the lead/lag structure because even 'actively traded' stocks are found to lag the futures prices. Against the possibility that 'the effects of infrequent trading make futures appear to lead spot market' is right, the analysis will be conducted with using first raw and then ARMA filtered series. At the end of this study, inferences may show whose idea is supported by the Turkish market structure; Stoll and Whaley (1990) or Chan (1992).

While performing the ARMA filtering, determining the AR and MA orders is important. As suggested by Stoll and Whaley (1990), orders are to be provided from the correlation structure of the raw data series. Statistically significant lags of ACF and PACF help to condition the MA and the AR orders, respectively. Nonetheless, ARMA models will be constructed with the suggested lag structure determined by Schwarz Information Criterion (SIC).Thereby the best model will be created and the residual series of the ARMA model will be recorded in order to use in the causality test later.

Filtering is applied to futures and spot price series in a yearly basis. AR and MA orders of the accepted models are tabulated below:

Table 4.5 ARMA Model Parameters for Filtering

		MA orders	**AR orders**
2005	**futures**	33 63 73 90 164 168	7 90 105 153
	index	29 69 174	28 152 174
2006	**futures**	54 110 165	54 110 164 165
	index	48 57 96	23 110
2007	**futures**	24 31 55 60 134 165	10 24 165 180
	index	–	55 165 180
2008	**futures**	120 180	60 84 120 180
	index	120 180	60 120 180
2009	**futures**	36 60	36 42 120
	index	–	61
2010	**futures**	89 114	11 57 89
	index	138	138

After we complete the analysis on the raw price observations, the outcomes will be reported and interpreted. Yet, considering the infrequent trading issue, the analysis will be repeated through ARMA filtered data and outcomes will be compared to raw data results.

CHAPTER 5

METHODOLOGY

5.1 Stationarity and Unit Root Tests

A stationary series is a stochastic process whose mean, variance and auto-covariance structure stays unchanged as time shifts. The concept of stationarity is quite important to researchers for some reasons that Brooks states in his book. Before all else, stationary series strongly reflects the past behavior of the series to future. To clarify, if a shock hits a stationary series, effects of the shock gradually die away. In other words, the effect of the shock at t+1 is smaller than that of t and shock power will decrease from t+1 to t+2, and it will go on like that. But, on the other hand, a non-stationary series does not move like a stationary one towards shocks. Effects of the shock persist and never die away. Moreover, it can never be anticipated that how the non-stationary series will behave. Another shortcoming of a non-stationary process is that it may lead spurious regression problem. That is to say, in regression analysis if the two variables are trending through time, R^2 could be very high although series are completely unrelated. Thus usage of non-stationary data may lead to meaningless results. One other weakness arises in estimation process. In a model containing non-stationary components, standard assumptions of asymptotic analysis fail. In other words, t-statistics do not follow t-distribution, F-statistics not anymore come from F-distribution and hypothesis testing becomes impracticable. Hence, estimates cannot be relied on. To be able to avoid all those imperfections, before starting analysis, stationarity conditions should be investigated and if necessary some improvements are supposed to be performed.

Prior to working on lead-lag relationship between markets, stationarity conditions of each market component need to be tested. This testing is urgent since the methods to be applied require some specific stationarity conditions. In some stages of the study stationarity series will be required however some techniques need non-stationarity. In non-stationary series it is also very essential to specify the order of integration. Thus these requirements make it crucial to correctly define the series on the basis of stationarity. Owing to these facts univariate properties of both series will be analyzed by means of statistical tests.

Non-stationary series basically have two forms, namely, random walk with drift model and trend stationary model.

The general form to describe the non-stationarity can be written as:

$$Y_t = \mu + bt + \varphi Y_{t-1} + u_t \, , \tag{5.1}$$

where Y_t is a time series process, b is trend term and μ is the intercept coefficient. When $\varphi = 0$, equation (5.1) becomes trend stationary model. But if $b = 0$, the remaining equation is random walk with drift model as:

$$Y_t = \mu + \varphi Y_{t-1} + u_t \, , \tag{5.2}$$

There exist three possible cases for model (5.2) regarding stationarity:

1. $\varphi < 1$, meaning that shocks to the system gradually die away, so series is stationary.

2. $\varphi = 1$, meaning a unit root is present at the series. Shocks stick with the system thus the system is non-stationary.

3. $\varphi > 1$. It is the explosive case. Here effects of shocks become severe as time goes but this type of system does not any reasonable description in time series analysis.

Case 1 represents the series which we prefer to work with. However, if the data is in form of case 2, it must be converted to case 1 to be able to conduct an analysis. One way to achieve this is differencing the series. Differenced series is defined as,

$$\Delta Y_t = Y_t - Y_{t-1} \, , \tag{5.3}$$

If we think of a random walk with drift model and then difference it, the differenced series will be stationary.

$$Y_t = \mu + Y_{t-1} + u_t \, , \tag{5.4}$$

$$Y_t - Y_{t-1} = \mu + Y_{t-1} + u_t - Y_{t-1} \, , \tag{5.5}$$

then, substitute (5.3) in (5.5), we have

$$\Delta Y_t = \mu + u_t \, , \tag{5.6}$$

which is a stationary series. Since (5.4) become stationary after one differencing, it is denoted as I(1) (integrated of order 1).

It is revealed that some methods are proper only when the data is non-stationary in their levels but the rest of the study will be conducted on the stationary series. On account of this information, above procedure called differencing is quite important and it will be used throughout the empirical work frequently.

The first effort to test the existence of unit root is by Dickey and Fuller (1979). They state in the null hypothesis that the series has a unit root against the alternative hypothesis of stationarity. Simply, they work on the hypotheses:

$$\mathrm{H}_0 : \varphi = 1 (\Delta Y_t = u_t) \, , \tag{5.7}$$

$$\mathrm{H}_1 : \varphi < 1(\Delta Y_t = \psi Y_{t-1} + \mu + bt + u_t), \tag{5.8}$$

The test statistic is $\hat{\psi} / s\hat{e}(\hat{\psi})$, but due to non-stationarity, this test statistic does not follow t-distribution under the null hypothesis. To solve this problem MacKinnon (1991) recomputed the critical values of ADF unit root test after suitable simulation studies.

One drawback of Dickey-Fuller (1979) test is that the error series is considered not to be auto-correlated. However, when the autocorrelation in dependent variable, ΔY_t, is neglected, error terms undertake those auto-correlation effects. In this situation, test will end in oversized and misleading results. Not to allow such inefficiency, Dickey and Fuller enhance the procedure by adding lags of ΔY_t to the model. Eventually, the model (5.8) is upgraded to below model:

$$\Delta Y_t = \psi Y_{t\text{-}1} + \sum_{i=1}^{p} \alpha_i \Delta Y_{t\text{-}i} + \varepsilon_t \, , \tag{5.9}$$

The test with the new model is called Augmented Dickey-Fuller (ADF) test based on same test statistic and critical values. In ADF test, it is crucial to determine the optimal lag length denoted as p in (5.9). Unfortunately, there is no one specific criterion to decide on the length p although optimal selection is quite essential. If few lags are included, part of auto-correlation will remain in the model, on the other hand, too much lags increase the standard errors of coefficients. Therefore the robust results will be reached only after determining the correct lag length using an information criterion that suits the data properties. The three famous information criteria are Akaike (AIC), Schwarz (SIC) and Hannan-Quinn Information Criteria. Among those, SIC is selected to specify the true lag length due to its superior large sample properties and its strong consistency.

After a year that Dickey and Fuller (1979) find ADF test, a similar but improved unit root test is introduced by Philips and Perron (1988). The basic modification of PP test is controlling for the serial correlation. What Philips and Perron actually do is that they modify the t-ratio in DF test serial correlation cannot ruin the distribution of the test statistic. PP test is frequently practiced like ADF test, but it is accepted to be a more developed test owing to its serial correlation resistance and heterosekdasticity robustness.

Both ADF and PP tests have some weakness in deciding the non-stationarity especially when the coefficient φ is close to 1. These tests are also poor if sample size is small. Hence, to be sure about the stationarity decision, one confirmatory test may be used. As stated, ADF and PP tests are unit root tests, in other words null hypothesis states non-stationary process. Contrary to those, KPSS test by Kwaitkawski et. al in 1992 introduced having the stationarity in the null hypothesis. To be able to say that results are robust, while ADF/PP resulting in rejecting the null hypothesis, KPSS should fail to reject H_0 or if ADF/PP fails to reject, null hypothesis must be rejected in KPSS test.

At the beginning of all methods the series of spot and futures will be tested to see if they have a unit root. ADF and PP test hypothesis states that futures (spot) series has a unit root signaling non-stationary. Therefore, rejecting the null hypothesis means that the futures (spot) series is stationary. On the other hand KPSS test states that futures (spot) process is stationary and in this case rejection means non-stationarity. Hence correct decision requires opposite results from unit root and stationarity tests. To be able to get the exact result, throughout the empirical analysis, both unit root and stationarity tests will be performed.

5.2 Co-integration and Error Correction Models (ECM)

One of the methods to be used in application to determine the lead/lag relation between the derivatives and the spot market is co-integration. With the work of Wahab and Lashgari (1993), co-integration is introduced to the finance world as a tool to analyze the connection between spot and derivatives markets. In our empirical study, two different sub-methods of co-integration will be applied to the Turkish data set. By both sub-methods, basically, the long run behavior of both markets will be examined. If both markets support a very parallel behavior across time, then short run deviations from the long run behavior are studied and the results of this short run movements form the basics of the lead/lag relation.

Co-integration can be defined as a modern technique to describe the movements of multidimensional economic time series data. Numerous financial studies especially on prices and exchange rates frequently take advantage of co-integration in literature. The reason behind is that co-integration allows the researchers to differentiate between short run and long run deviations from equilibrium providing information on price discovery, lead/lag relation and market efficiency. With the same reason, in this study, co-integration is one of the applied methods to discover the temporal lead/lag relation between spot index and index futures markets in Turkey.

The theory implies that non-stationary time series processes integrated of same order are said to be co-integrated if their linear combination is stationary without differencing and the co-integrated series are expected to arrive to equilibrium level after some short run deviations. This procedure can be explained for the current study in the following way: If both spot and futures prices are non-stationary in their levels and if both have stationary structure after differencing of order 1, then a co-integration regression is formed. In this regression model point of interest is the residual series. In case that unit root tests reach a conclusion of stationary residuals, futures and spot prices are then said to be co-integrated. More clearly, two series of interest shape long run equilibrium relation; and by the short run pattern, temporal relation between the markets is revealed.

Co-integration is first introduced by Granger (1981) and further developed by Engle & Granger (1987) and Johansen (1988). Those improvements represent the two distinct types of the co-integration technique. In this study, both procedures will be performed to analyze the temporal interactions between spot index and index futures markets.

5.2.1 Engle-Granger Approach

Engle and Granger (1987) propose the long run equilibrium relation by the below equations:

$$F_t = \beta_0 + \beta_1 S_t + u_t, \tag{5.10}$$

$$F_t - \beta_0 - \beta_1 S_t = u_t, \tag{5.11}$$

where F_t and S_t represent the index futures and spot index prices at time t, respectively and u_t is the error term, i.e. deviation from equilibrium. Engle and Granger (1987) express that if both F_t and S_t are non-stationary in their levels, but stationary after first differencing the futures and spot prices are said to be co-integrated of order 1 and denoted I(1) with β_1 being the co-integration coefficient.

Steps of estimating Engle and Granger (1987) Co-integration:

Step 1: Before modeling the co-integration relation, univariate properties of the price series must be revealed. To be able to search for the co-integration, first the unit root structure of the series should be tested. EG approach requires both series to be integrated of order 1. To check this necessity, ADF test by Dickey and Fuller (1981) and PP test by Philips and Perron (1987) along with KPSS test will be performed. In case that both series are found to be I(1), then model (5.11) is constructed. Put another way, index futures prices (F_t) and stock index prices (S_t) need to be non-stationary in their levels. However, after differencing once, both series should become stationary. If this property is satisfied, co-integration regression by equation (5.11) is formed.

Once the price series (F_t and S_t) are stationary after first differencing and the model is built, then the error terms will be analyzed through ADF and PP tests. Provided that error terms (u_t) follow a mean-reverting & constant variance structure, i.e. I(0), futures and spot prices are co-integrated. Thus, to be able to say that futures and spot markets are co-integrated, error term of model (5.11) should be stationary without differencing. Model (5.11) represents the price relation, but classical statistical inferences based on OLS are not valid due to violations of regression assumptions by non-stationarity. Thus, although it is known that price series are co-integrated and an equilibrium relation exists, coefficient of (5.11) cannot be interpreted in that sense.

Once model (5.11) produces stationary residuals, it is stated that futures and spot prices move in company in the long run. In the next step, short term differences of these series will be checked to catch the lead/lag pattern.

Step 2: In step 1 long run relation is verified, then, in this step, short run deviation structure will be examined through Error Correction Models (ECM). The basic idea behind ECMs is that co-integration assures the long run equilibrium between two economic variables; in short run, however, synchronization may not be maintained. In other words, for some short time periods there may be drifts from long run accordance. The significance of short run disequilibrium can be formulated by the below models:

$$\Delta S_t = \alpha_{0s} + \theta_s \hat{u}_{t-1} + \sum_{i=1}^{p} \alpha_{1s,i} \Delta S_{t-i} + \sum_{j=1}^{p} \alpha_{2s,j} \Delta F_{t-j} + v_t^s, \qquad (5.12)$$

$$\Delta F_t = \alpha_{0f} + \theta_f \hat{u}_{t-1} + \sum_{i=1}^{p} \alpha_{1f,i} \Delta S_{t-i} + \sum_{j=1}^{p} \alpha_{2f,j} \Delta F_{t-j} + v_t^f, \qquad (5.13)$$

where ΔS_t and ΔF_t stands for the differenced forms of the corresponding price series at time t and $\hat{u}_t$ is called the "error correction term" representing the residual series of (5.10). According to Engle and Granger (1987), a drift occurred in one period is adjusted in the next period and this adjustment is represented via error correction term. Here, error correction term measures the speed of the short run drifts to return to equilibrium. In those equations, estimates of α_{2s} and α_{1f} indicate the lead/lag pattern along with θ_s and θ_f. It is stated by Granger (1988) that if there is a co-integration relationship between two variables, then at least uni-directional causality must exist between them. Thus in the empirical study, once co-integration is verified, lead/lag pattern can be investigated through (5.12) and (5.13). When futures market leads the spot market, some of coefficients α_{2s} and θ_s should found to be significant. Significant results of coefficients α_{1f} and θ_f indicate spot market leadership over futures market. On the other hand, bi-directional causality is implied if α_{2s}, α_{1f} are jointly significant with meaningful θ_s and θ_f.

To get a clearer picture of the analysis, Step 2 can be re-expressed in the following manner. As stated by Granger (1988), there should be at least one-way causality between spot and futures prices since they are found to be co-integrated. This causality relation is investigated in Step 2 by models (5.12) and (5.13). As literature suggests, causality must be studied via stationary series, which is differenced price series rather than price observations themselves. Since multi-dimensional causality (whether spot causes futures or futures causes spot or bi-directional causality) is questioned, two models, one for causality from futures to spot and one for causality from spot to futures, are formed. The existence of the lead/lag pattern is decided upon the coefficients of suggested models. In model (5.12), roughly futures market is regressed on spot market, therefore if coefficients corresponding futures are significant, "futures market leads spot market" is the result. In a similar way, significant coefficients corresponding spot prices mean that "spot market leads the futures market". Alternatively, both above conclusions may happen at the same time signaling bi-directional lead/lag relation.

In Engle and Granger (1987) approach, the hypothesis with respect to the parameters in the long run relationship cannot be tested. However, this problem is removed by Johansen Co-integration Approach.

5.2.2 Johansen Approach

Unlike Engle & Granger (1987), Johansen (1988) is a multivariate approach with n variables all integrated of same order. Since this method has multivariate components, the long run model is in form of Vector Autoregressive (VAR) Model as:

$$X_t = \beta_0 + \sum_{i=1}^{k} \beta_i X_{t-i} + v_t, \qquad (5.14)$$

where X_t is the vector of differenced forms of futures and spot prices $[F_t, S_t]$, β_0 is the intercept vector and v_t is the error term .

Differently from Engle and Granger (1987), in this approach long run behavior of at least two variables can be investigated. It means that Johansen's (1988) method enables researchers to study on more than two time series processes. However, throughout empirical analysis behavior of two components will be examined as in Engle and Granger (1987) since the temporal relation of one series of index futures prices and one series for stock index prices are the components of interest. Even there exist only two series, for the error correction representation VAR model will be constructed as method requires. X_t of model (5.14) will simply be composed of index futures and stock index prices.

Steps of estimating Johansen's (1988) Co-integration:

Step 1: As in Engle & Granger (1987) case, at the beginning of the study, stationarity conditions of the components must be investigated. Both spot and futures prices series should be integrated of order 1. This limitation is tested via ADF, PP and KPSS tests. If the necessary condition is satisfied, suitable ECM of (5.14) is constructed in a very same manner that stated in Engle and Granger procedure.

In the first steps of both techniques stationarity conditions are to be checked to see the two processes are integrated of the same order or not. Before all else, it is to be proved that futures and spot price series are integrated of order 1 to say that the differenced series are stationary. Thereafter the corresponding ECM form can be constructed by model (5.15).

Step 2: Re-parameterization of (5.14) results in the below ECM:

$$\Delta X_t = \beta_0 + \sum_{j=1}^{k-1} \tau_j \Delta X_{t-j} + \pi X_{t-k} + \varepsilon_t, \qquad (5.15)$$

where $\tau_j = \sum_{p=1}^{j} \beta_p - \mathrm{I}$ is short run adjustment and $\pi = -(\mathrm{I} - \sum_{m=1}^{k} \beta_m)$ is long run response matrix.

Mathematical formulation of Johansen ECM in (5.15) can actually be written as system of two equations, such as:

$$\Delta S_t = \beta_s + \sum_{i=1}^{k-1} \beta_{1s,i} \Delta S_{t-i} + \sum_{j=1}^{k-1} \beta_{1f,j} \Delta F_{t-j} + \theta_s \hat{u}_{t-1} + v_{s,t}, \qquad (5.16)$$

$$\Delta F_t = \beta_f + \sum_{i=1}^{k-1} \beta_{2s,i} \Delta S_{t-i} + \sum_{j=1}^{k-1} \beta_{2f,j} \Delta F_{t-j} + \theta_f \hat{u}_{t-1} + v_{f,t}, \qquad (5.17)$$

Above error correction models (5.16) and (5.17) are the open forms of model (5.15). It can be more clearly seen from these ECMs that differenced price series are regressed on its own lags, on the lags of the other differenced price series and on the error correction term.

After requirements stated in the first step are ensured then the relation is expressed in the form of models (5.16) and (5.17). These models are established in a way that they harbor the long relation correlation as well as the short run dynamics. That is to say analysis of models (5.16) and (5.17) will reveal the equilibrium relation as well as the temporal structure between spot and futures prices.

Step 3: Johansen (1988) procedure can be highly affected by the lag length. Thus the lag length k in equations (5.16) and (5.17) is specified by Schwarz's Bayesian Information Criterion (SBIC).

In the third step of the Johansen's strategy, the lag length is specified. Lag length selection is crucial since all estimates are affected by this value leading wrong results on the lead/lag structure. To avoid deficient estimates, lag length is determined by the help of famous statistical criteria.

Step 4: In this approach, two tests of co-integration relation are performed, namely, trace test and maximal eigenvalue test. Null hypothesis of both tests indicate that there are at most r co-integration vectors.

The test of co-integration conducted on the rank of π matrix through its eigenvalues. The rank r is important since it determines the number of co-integrated vectors. Let n denote the number of variables of X matrix. If r = n, then we say all the variables in X are stationary. If r = 0, there are no stationary linear combinations of components of X. When 0<r<n, there exist r co-integration vectors, i.e. r stationary linear combinations of components of X.

In empirical analysis, it is expected to identify 1 co-integration vector to be able to prove co-integration. Since X_t composes of two vectors one is futures prices F_t and the other is spot market prices S_t, n is equal to 2. If r = 2, spot and futures prices will be stationary thus it is not sensible to work on co-integration. If r = 0, co-integration relation cannot be constructed due to the fact that no such linear combination is found which is stationary. However, when r = 1, it is clear that there exists one co-integration vector by which the linear combination of futures and spot prices compose a stationary structure.

Given that the rank of a matrix is equal to the number of its eigenvalues, eigenvalues (λ_i) are computed first and build in ascending order as:

$$\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n, \tag{5.18}$$

If there is no co-integrated vectors then rank r must be 0 meaning that λ_i statistically will not be different than 0 for all i. When there exist r co-integrated vectors, eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_r$ are statistically non-zero, λ_1 being the largest, while others get smaller.

Two test statistics are calculated in the light of eigenvalues. First test is trace test and the other is maximal eigenvalue test. Both state maximum r eigenvectors in the null hypothesis against the alternative that there are r+1eigenvectors in for trace test and more than r eigenvectors for maximal eigenvalue test. The corresponding test statistics are given below:

$$\lambda_{trace}(r) = -T\sum_{i=r+1}^{n} \ln(1-\hat{\lambda}_i), \qquad (5.19)$$

$$\lambda_{\max}(r, r+1) = -T\ln(1-\hat{\lambda}_{r+1}), \qquad (5.20)$$

where $\hat{\lambda}_i$ is the estimation of the i^{th} eigenvalue from π matrix. Those test statistics are compared to the special critical values introduced by Johansen and Juselius (1990). If the test statistic is greater than the table value, then decline that there are at most r co-integration vectors. This process follows a sequence of test steps. At the first stage, begin testing the hypothesis that r = 0. In case that the hypothesis is rejected, then test for r = 1 and continue in this way until the value r that the hypothesis cannot be rejected.

The exact procedure of hypothesis testing in empirical work is summarized as follows. At the first stage H_0:r=0 versus H_1:r=1for trace test and H_0:r=0 versus H_1:0<r≤1 for maximal eigenvalue test will be stated. If null hypotheses cannot be rejected, then the conclusion is no co-integration and the testing procedure is finished. But if both null hypotheses are rejected then new hypotheses are set up for both trace and maximal eigenvalue tests as H_0:r=1 versus H_1:r=2. Rejection of the null indicates that spot and futures series are stationary. So that no need for co-integration technique. However if H_0 cannot be rejected, test reveals that there is 1 co-integration vector, i.e. series are co-integrated. For this study, expectation is to find 1 co-integration vector making this method available for investigating lead/lag pattern. After confirming 1 co-integration vectors, the exact same VEC models as in Engle and Granger approach can be constructed. It is proper to conduct the same analyses and tests to come up with the lead/lag pattern.

The basic difference between the approaches of Engle&Granger (1987) and Johansen (1988) is the fact that Johansen procedure allows to test the co-integration relationship directly between variables rather than working on the residuals. In Johansen's test we are free to test the long run and short run relation on the variable properties. Moreover Johansen's method is a multivariate process, while only bivariate tests can be performed via Engle&Granger approach. Owing to those extra properties, Johansen's method is accepted to be more efficient than the preceding procedure in literature. Therefore, after co-integration tests are performed according to both methods, the rest of the analysis will be grounded on the findings of Johansen approach.

5.3 Causality

While examining the lead/lag relation between derivatives market and spot market, Granger causality analysis allows us to understand whether futures or the spot prices rule the market. In other words, by performing causality analysis, price discovery process can be explained.

Causality is a method to work on multivariate data, investigating whether the changes in one variable have impact on the changes of the other variables. This idea is first suggested by Granger (1969). The study is focused on the predicting ability of the past observations of one time series process on the present and future values of the other ones.

5.3.1 Linear Granger Causality

Granger's linear method works on stationary time series data. Thus differenced forms of index futures and spot index prices will again be in use. Granger states that if lags of ΔF_t contain information to predict the future values of variable ΔS_t, then ΔF_t 'granger causes' ΔS_t. Put another way, if changes in spot prices (ΔS_t) are better forecasted by adding lags of change of futures prices (ΔF_{t-i}) to the past spot prices (ΔS_{t-i}) compared to past spot prices alone (ΔS_{t-i}), then futures market is said to "Granger cause" the spot market. Granger models this relationship by the following Vector Autoregressive Representation (VAR):

$$\Delta S_t = \mu + \sum_{i=1}^{p} \alpha_i \Delta S_{t-i} + \sum_{j=1}^{p} \beta_j \Delta F_{t-j} + v_{s,t}, \tag{5.21}$$

as well as;

$$\Delta F_t = \varphi + \sum_{i=1}^{p} \lambda_i \Delta F_{t-i} + \sum_{j=1}^{p} \gamma_j \Delta S_{t-j} + v_{f,t}, \tag{5.22}$$

Estimation of model (5.21) with the reverse model (5.22) reveals the granger causality relation between these two time series processes. The null hypothesis of both (5.21) and (5.22) states that there exists no granger causality by testing $\beta_1 = \beta_2 = \ldots = \beta_p = 0$ and $\gamma_1 = \gamma_2 = \ldots = \gamma_p = 0$ respectively. Testing these hypotheses result in either one of below four alternatives:

- Spot market Granger causes derivatives market if H_0: $\gamma_1 = \gamma_2 = \ldots = \gamma_p = 0$ is rejected;
- Derivatives market Granger causes spot market if H_0: $\beta_1 = \beta_2 = \ldots = \beta_p = 0$ is rejected;
- "A feedback relation" exists between the two markets if both hypotheses H_0: $\gamma_j = 0$ and H_0: $\beta_j = 0$ *for all j* are rejected;
- Spot and derivatives markets are independent if we fail to reject both hypotheses H_0: $\gamma_j = 0$ and H_0: $\beta_j = 0$ *for all j.*

As understood, in case that one of the hypotheses is rejected, then there exists linear Granger causality at least in one direction. While applying linear Granger causality analysis, one point that should be correctly specified is the lag number, p. Inadequate lags result in auto-correlated error terms or redundant lags reduce the power of the test unless the correct number of lags is chosen. However, there is not any specific criterion defined to select the necessary lag number in Granger causality. Thus, in order to identify the suitable number of lags the length will be set to 10 and then by the help of Schwarz Information Criteria the most parsimonious model will be selected which consists the optimum lag length.

As clarified, a variable A Granger causes B provided that inclusion of A's past values contributes to the predictability of B. For the purpose of testing this contribution, unrestricted models (5.21) and (5.22) are not sufficient by themselves since they should be compared to models containing its own lags only. Thus, restricted forms of these models are constructed as:

$$\Delta S_t = \mu + \sum_{i=1}^{p} \alpha_i \Delta S_{t-i} + h_{s,t}, \tag{5.23}$$

$$\Delta F_t = \varphi + \sum_{i=1}^{p} \lambda_i \Delta F_{t-i} + h_{f,t}, \tag{5.24}$$

After building restricted models, now we are ready to form the test statistic. While testing the null hypothesis that "spot prices does not cause futures prices", coefficient of determination estimates (R^2) of (5.22) and (5.24) are calculated and the following Wald F test statistics are obtained as suggested by Sims (1972):

$$F = ((RSS_{(5.24)} - RSS_{(5.22)})/q) / (RSS_{(5.22)}/(n-k)), \tag{5.25}$$

where n is the total sample size, q and k are number of parameters in (5.24) and (5.22) respectively. *RSS* stands for the residual sum of squares. If results imply the rejection of the null hypothesis, then the conclusion is that cash market Granger causes futures market.

Another way of testing the Granger Linear Causality can be performed through Wald Chi-Square Test. Similar to F-test explained above, Wald Chi-Square procedure also tests the joint significance of the coefficients corresponding to the market that suspected to lead the other market. In the empirical study of lead/lag structure, both Wald F and Wald Chi-Square test will be applied to the spot and futures market observations.

As stated before in lead/lag structure detection, Granger causality method will be applied to the differenced price series corresponding to futures and spot markets due to the fact that causality models can only be constructed on the stationary variables. So as to test the hypothesis that futures does not cause spot market, test-statistic based on (5.22) and (5.24) is calculated as in (5.25). If computed statistic is greater than the critical value, then the null hypothesis is rejected and it is concluded that futures returns cause (lead) spot returns. The very same procedure is implemented using (5.21) and (5.23) to test that spot returns do not cause futures returns. If both tests result in rejecting the corresponding hypotheses, bi-directional causality is the relation between the markets.

When analyzing the lead/lag pattern, co-integration is always the first method to perform since existence of co-integration affects the application procedure of the following methods. Put another way, inclusion of co-integration necessitates making adjustments in the succeeding steps. For this reason, we must take existence of co-integration into consideration. In a regular analysis if the series are not co-integrated, Granger causality is detected by VAR models shown by (5.21) and (5.22). However, in case that co-integration is established, lagged Error Correction Term (ECT) should be added to VAR models, meaning that VECMs designed in Johansen approach are supposed to be used rather than VAR models. If series are co-integrated but ECT is not included in VAR causality model, only the causality stemming from short run deviations will be caught. Unfortunately, existent causality of common trend cannot be detected using VAR models. On the other hand, using VECM having ECT can catch the causal relationship originating from the long run equilibrium even if short run information does not have any causality power. Therefore, if spot and futures prices are co-integrated, linear causality will be tested by error correction models suggested below:

$$\Delta S_t = \beta_s + \sum_{i=1}^{k-1} \beta_{1s,i} \Delta S_{t-i} + \sum_{j=1}^{k-1} \beta_{1f,j} \Delta F_{t-j} + \theta_s \hat{u}_{t-1} + v_{s,t}, \tag{5.26}$$

$$\Delta F_t = \beta_f + \sum_{i=1}^{k-1} \beta_{2s,i} \Delta S_{t-i} + \sum_{j=1}^{k-1} \beta_{2f,j} \Delta F_{t-j} + \theta_f \hat{u}_{t-1} + v_{f,t}, \tag{5.27}$$

Thus, in application if the series will be found to be co-integrated, then the causality test will be performed not on VAR models (5.21) and (5.22) but on VEC models numbered (5.26) and (5.27).

5.3.2 Nonlinear Granger Causality

It is stated in literature that Linear Granger Causality Test is unable to detect nonlinear interactions of variables. In this study, the link between spot and futures markets may be non-linear rather than linear. Due to this suspect, non-linear Granger causality test will also be applied to the data set. Non-linear causality is fed from the results of linear causality analysis. That is to say, non-linear causality should be worked on the basis of finding of the linear work. The classical VAR model or the error correction model in case of co-integrated series only measures the linear association. In case of the existent non-linear tie between markets, the error terms are to undertake this relation. As a result, the residuals collected from VAR models should be evaluated to see whether any remaining non-linear structure is present or not. Therefore testing the residuals will ensure to catch any nonlinear relation that cannot be identified by linear procedures.

The first test to identify the nonlinear structure which cannot be revealed by traditional linear tests is developed by Baek and Brock (1992). Their method is a nonparametric testing technique dealing with correlation integral to catch the nonlinearities within and across time series. Employing their test nonlinear structure is exposed between income level and money. Then their test is improved by Hiemstra and Jones (1994) such that small sample properties are modified and identically and independently distribution assumption is eliminated. That is to say the test developed by Baek and Brock (1992) is not practical due to the severe assumptions based on strict stationarity. With the latter study of Hiemstra and Jones (1994), the test is expanded by debugging the strict stationarity and mutual independency assumptions. After these adjustments, the new test is found to be robust to sample size problems and structural breaks. The new nonparametric test of Hiemstra and Jones (1994) is named as Modified Baek-Brock Test.

The procedure of Modified Baek-Brock Test can be explained in the following way. To be able to perform the test, there must be two stationary and weakly dependent time series processes. This need will be filled by the residuals of linear causality models representing the spot and futures price series. The residuals coming from VAR (VECM) model of spot series is denoted by s and that of futures series is symbolized by f. After that let the series s_t^m, s_{t-ls}^{ls} and f_{t-lf}^{lf} correspond to m-length lead vector of s_t, ls length lagged vector of s_t and lf length lagged vector of f_t respectively. Then these series are defined as:

$$s_t^m = (s_t, s_{t+1}, \ldots, s_{t+m-1})$$

$$s_{t-ls}^{ls} = (s_{t-ls}, s_{t-ls+1}, \ldots, s_{t-1})$$

$$f_{t-lf}^{lf} = (f_{t-lf}, f_{t-lf+1}, \ldots, f_{t-1})$$

For defined values of parameters m, ls and $lf \geq 1$and for some $e \geq 0$, the series f_t does not strictly Granger cause s_t, if the following equation holds:

$$\begin{aligned} &\Pr(\| s_t^m - s_g^m \| < e \mid \| s_{t-ls}^{ls} - s_{g-ls}^{ls} \| < e, \| f_{t-lf}^{lf} - f_{g-lf}^{lf} \| < e) \\ &= \Pr(\| s_t^m - s_g^m \| < e \mid \| s_{t-ls}^{ls} - s_{g-ls}^{ls} \| < e), \end{aligned} \tag{5.28}$$

where Pr stands for probability and ||.|| denotes maximum norm. Explanation of equation (1) is stated in Hiemstra and Jones (1994) as follows: "The left side of (5.28) is the conditional probability that two arbitrary m-length lead vectors of s_t are within a small distance e of each other, given that the corresponding s_t and f_t (lag vectors) are within e of each other. The probability on the right side of (5.28) is the conditional probability that two arbitrary m-length lead vectors of s_t are within a distance e of each other."

To obtain a test statistic, we first write conditional probability equation (5.28) in terms of ratios of joint probabilities:

$$\frac{H_1(m+ls, lf, e)}{H_2(ls, lf, e)} = \frac{H_3(m+ls, e)}{H_4(ls, e)}, \tag{5.29}$$

where,

$$H_1(m+ls, lf, e) = \Pr(\| s_{t-ls}^{m+ls} - s_{g-ls}^{m+ls} \| < e, \| f_{t-lf}^{lf} - f_{g-lf}^{lf} \| < e), \tag{5.30}$$

$$H_2(ls, lf, e) = \Pr(\| s_{t-ls}^{ls} - s_{g-ls}^{ls} \| < e, \| f_{t-lf}^{lf} - f_{g-lf}^{lf} \| < e), \tag{5.31}$$

$$H_3(m+ls, e) = \Pr(\| s_{t-ls}^{m+ls} - s_{g-ls}^{m+ls} \| < e), \tag{5.32}$$

$$H_4(ls, e) = \Pr(\| s_{t-ls}^{ls} - X_{g-ls}^{ls} \| < e), \tag{5.33}$$

Now, let $I(A, B, e)$ be an indicator function which is 1 if A and B are "within the maximum norm distance e of each other" and 0 in all other cases. Thereby correlation integral estimators of equations (5.30), (5.31), (5.32) and (5.33) are calculated as:

$$H_1(m+ls,lf,e,n)=\frac{2}{n(n-1)}\sum_{t<g}\sum I(s_{t-ls}^{m+ls},s_{g-ls}^{m+ls},e).I(f_{t-lf}^{lf},f_{g-lf}^{lf},e), \quad (5.34)$$

$$H_2(ls,lf,e,n)=\frac{2}{n(n-1)}\sum_{t<g}\sum I(s_{t-ls}^{ls},s_{g-ls}^{ls},e).I(f_{t-lf}^{lf},f_{g-lf}^{lf},e), \quad (5.35)$$

$$H_3(m+ls,e,n)=\frac{2}{n(n-1)}\sum_{t<g}\sum I(s_{t-ls}^{m+ls},s_{g-ls}^{m+ls},e), \quad (5.36)$$

$$H_4(ls,e,n)=\frac{2}{n(n-1)}\sum_{t<g}\sum I(s_{t-ls}^{ls},s_{g-ls}^{ls},e), \quad (5.37)$$

where $n=T+1-m-\max(ls,lf)$ and $t,g=\max(ls,lf)+1,...,T-m+1$

Above estimates of correlation integral now make it possible to test the Granger non-causality represented by equation (5.28). Hiemstra and Jones (1994) reports that "For given values of *m*, *ls* and *lf* and scale parameter *e*, if f_t does not strictly cause s_t then,"

$$\sqrt{n}[\frac{H_1(m+ls,lf,e,n)}{H_2(ls,lf,e,n)}-\frac{H_3(m+ls,e,n)}{H_4(ls,e,n)}]\rightarrow N(0,\sigma^2(m,ls,lf,e)), \quad (5.38)$$

To be able to apply the nonlinear test, one should decide on the parameters *m*, *ls*, *lf*, *e* and *σ*. Unfortunately, there is not any criterion to exactly select the optimal values for those parameters. Hiemstra and Jones (1994) perform the test assigning values to parameters in the light of their previous study. On account of Hiemstra and Jones (1993) Monte Carlo Simulations for the modified Baek and Brock test propose some specific values and conditions for parameters *m*, *ls*, *lf*, *e* and *σ*. According to the simulation results, lead-length m=1. Another important issue is determination of ls and lf. Hiemstra end Jones accept that ls=lf and they generally range between values 1 and 8. Since ls and lf are mainly represent the lag structure of the nonlinear testing, in our study these values will be equated to the calculated lag-lengths. Moreover they suggest standardizing each series to have a common σ=1 while e=1.5σ. Apart from their study conducted in 1994, many researchers implement these suggested parameter values. Fujihara and Mougoue (1997) make use of pre-determined parameter values. In addition Abhyankar (1998), Silvapulle and Moosa (1999) and Ciner (2001) who inspect the lead lag relationship between spot and futures market, directly exercise the recommended values. In light of the previous studies, simulation results will be practiced in application of the nonlinear causality analysis.

The Modified Baek-Brock test is uni-directional, in other words, it detects only one-way causality. Thus to be able to correctly specify the causality relation, the hypothesis that "futures prices do not strictly cause spot prices" will be tested as well as that of "spot prices do not strictly cause futures prices". For both hypotheses, the test statistics defined by (5.38) are calculated following the nonlinear causality algorithm and then, compared to critical t-values. If statistics exceed the critical values; the decision is futures (spot) prices cause spot (futures) prices significantly.

Nonlinear causality is the last step of price discovery process. It basically takes the residuals coming from linear causality models and studies them to see if nonlinearity shapes the causality. Similar to case in linear causality, in nonlinear analysis returning to equilibrium level in long run changes the procedure significantly. The standard behavior is valid unless the co-integration is established. But in case that the series will be found to co-integrate, nonlinear causality can no more be examined through VAR residuals but through the ECM residuals.

CHAPTER 6

APPLICATION AND EMPIRICAL RESULTS

Before performing the statistical tests, it is beneficial to visualize the raw price series in order to detect the price fluctuations and patterns. In addition, by plotting the spot and futures prices on the same graph we will be able to see whether they move accordingly, or not. The yearly graphs of raw price series are illustrated in Figure 6.1:

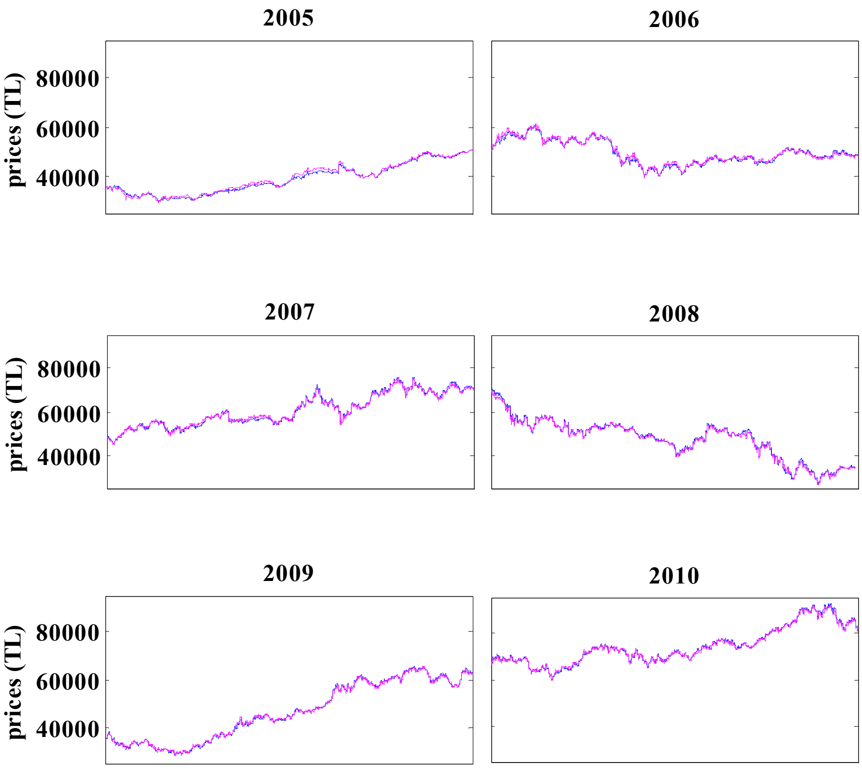

Figure 6.1 Yearly Movements of Index Futures and Spot Index Prices

From the Figure 6.1 it can be observed that spot and futures prices move accordingly in long run. In 2005, 2007, 2009 and 2010, prices move upward trend in long run. Conversely, in 2008, the possible trend seems to change its direction. In 2006, the price movement structure is not similar to other years because in this year, prices do not seem to follow noticible upward or downward trend. While performing the analysis, trend structure of the data is explained by statistical tests since Figure 6.1 cannot directly reveal the existence of the trend.

6.1 Stationarity Results

The long run structure of futures and spot prices will be investigated through co-integration and according to the decision of co-integration, causality analysis will be performed. But in the first step, stationarity should be tested since co-integration requires I(1) series. From the Figure 6.1 index futures and index prices belonging to each year seem to be non-stationary. Although non-stationarity is obvious by visual inspection, a statistical test is needed to say that the prices are non-stationary.

Raw spot index and index futures prices are tested through ADF, PP and KPSS tests. If the prices are non-stationary, the differenced forms will also be tested. The stationarity condition of the differenced prices will determine the order of integration.

The testing procedure is sensitive to the structure of the data. Calculation of the test statistic and the critical value differs if the series have trend or/and intercept. Hence, in the first step, we seek the significance of intercept and trend component.

Table 6.1 Trend and Intercept Tests

	Intercept		**Trend**	
	coefficient	p-value	coefficient	p-value
futures05	29178,4	0,000	3,763	0,000
index05	29415,68	0,000	3,832	0,000
future06	54206,6	0,000	-0,710	0,000
index06	54924,23	0,000	-0,768	0,000
futures07	48600,23	0,000	1,808	0,000
index07	48948,92	0,000	1,745	0,000
futures08	61572,58	0,000	-1,886	0,000
index08	61005,32	0,000	-1,860	0,000
futures09	26753,09	0,000	2,794	0,000
index09	27022,59	0,000	2,760	0,000
futures10	62650,79	0,000	1,748	0,000
index10	62765,93	0,000	1,705	0,000

Parallel to visualization, statistical models also indicate that both futures and spot index series have an intercept and trend component according to results in Table 6.1. In the light of the foregoing, ADF, PP and KPSS tests will be performed considering the presence of trend and intercept.

The test results of raw series are tabulated in Table 6.2 in a yearly basis:

Table 6.2 Unit Root and Stationarity Tests for TURKDEX-ISE 30 and ISE 30 Prices

	ADF		**PP**		**KPSS**	
	Test Statistic	CV	Test Statistic	CV	Test Statistic	CV
futures05	-2,98809	-3,41068	-3,02341	-3,41068	0,89057	0,14600
index05	-3,08069	-3,41068	-3,09282	-3,41068	0,58971	0,14600
futures06	-1,68387	-3,41015	-1,87051	-3,41015	2,52526	0,14600
index06	-2,02182	-3,41015	-2,01617	-3,41015	2,44991	0,14600
futures07	-2,92505	-3,41014	-3,01336	-3,41014	0,37476	0,14600
index07	-3,28880	-3,41014	-3,17515	-3,41014	0,28694	0,14600
futures08	-2,75989	-3,41011	-2,82809	-3,41011	0,93882	0,14600
index08	-2,86914	-3,41011	-2,82656	-3,41011	0,94683	0,14600
futures09	-2,45125	-3,41010	-2,47111	-3,41010	1,23526	0,14600
index09	-2,61685	-3,41012	-2,42921	-3,41012	1,17980	0,14600
futures10	-2,05190	-3,41014	-2,14944	-3,41014	1,59792	0,14600
index10	-2,19029	-3,41014	-2,20227	-3,41014	1,61697	0,14600

The lag lengths of all three tests are determined by Schwarz Information Criteria. ADF and PP procedures state that if |CV| > |Test Statistic|, the series are non-stationary. According to the statement of unit root tests, all price series are found to be non-stationary as expected. Co-integration needs pairs of I(1) futures and index series, so the differenced forms of the series are tested expecting stationary series. When Figure 6.2 examined, the differenced series seem to be stationary.

All series seem to move around an imaginary mean line and variation does not seem to change drastically. Although graphs sign that the series are stationarity, this intuition should be supported by statistical tests. ADF and PP results are summarized in Table 6.3.

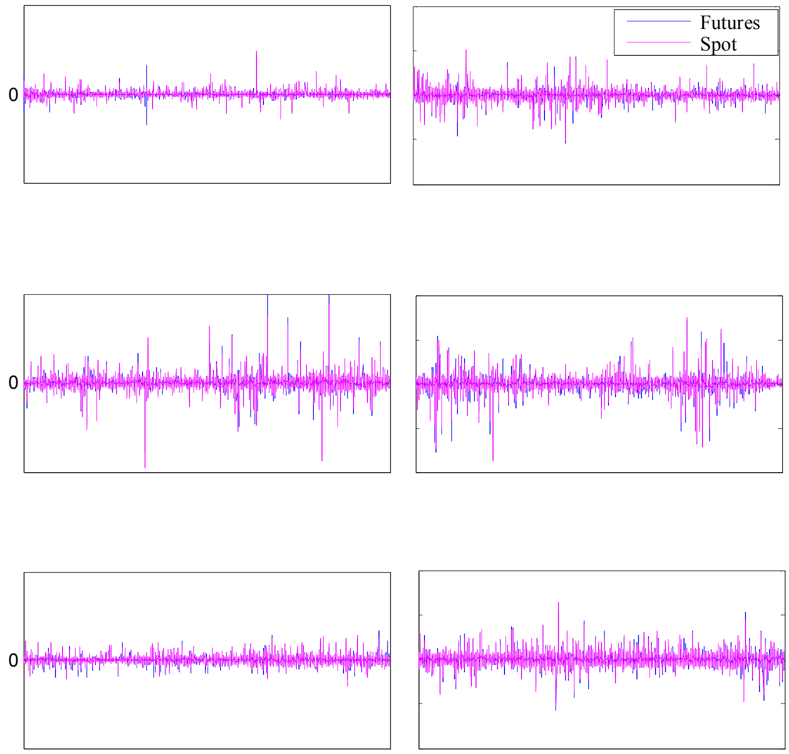

Figure 6.2 Yearly Differenced Index Futures and Spot Index Prices

Table 6.3 Unit Root and Stationarity Tests for Differenced TURKDEX-ISE 30 and ISE 30 Prices

	ADF		**PP**	
	Test Statistic	CV	Test Statistic	CV
dfutures05	-79,70470	-1,94089	-79,28570	-1,94089
dindex05	-73,33540	-1,94089	-73,31750	-1,94089
dfutures06	-112,77270	-1,94085	-113,43950	-1,94085
dindex06	-117,72260	-1,94085	-117,77650	-1,94085
dfutures07	-114,21130	-1,94085	-114,23620	-1,94085
dindex07	-85,27770	-1,94085	-119,96240	-1,94085
dfutures08	-121,63840	-1,94085	-121,72520	-1,94085
dindex08	-125,06060	-1,94085	-125,05180	-1,94085
dfutures09	-114,92250	-1,94084	-115,10440	-1,94084
dindex09	-89,84194	-1,94084	-125,67840	-1,94084
dfutures10	-114,60590	-1,94084	-114,71420	-1,94084
dindex10	-89,23477	-1,94084	-131,90820	-1,94084

ADF and PP test statistics corresponding to differenced forms of both spot and futures prices indicate stationarity. However, KPSS test is not performed, because this test is valid only if the data includes intercept. In other words, since |CV| > |Test Statistic| for all of the cases, series are proved to be integrated of order 1. Thereby the basic requirement of co-integration is satisfied.

6.2 Co-integration Results

Long-run equilibrium of prices tested according to two different algorithms.

6.2.1 Results of Engle-Granger Approach

Engle-Granger (1987) claimed that if two series are co-integrated then the residuals coming from the simple linear regression on the mentioned series should be stationary. Thus, futures and spot index series will be regressed on each other and obtained residuals will be tested with unit root tests to understand if they are stationary.

First futures data is regressed on index data and the residuals belonging to this model is named as 'futres'. Then the reverse model is constructed and coming residuals are called 'indres'. Unit root results of these innovations will say whether futures and spot prices are co-integrated.

Table 6.4 Unit Root and Stationarity Tests for Engle-Granger Approach Residuals

	ADF Test Stat.	**PP Test Stat.**	**CV**
futres05	-3,44763	-4,93196	-1,94088
indres05	-3,47048	-4,94789	-3,41068
futres06	-6,64855	-7,93585	-3,41015
indres06	-6,92079	-8,24887	-3,41015
futres07	-7,18158	-10,39548	-1,94085
indres07	-7,20257	-10,13800	-1,94085
futres08	-6,91083	-11,39177	-3,41011
indres08	-6,94133	-11,58186	-3,41011
futres09	-8,26293	-17,98170	-3,41012
indres09	-8,28415	-18,06117	-3,41012
futres10	-7,65949	-32,76044	-3,41014
indres10	-7,67381	-32,87890	-3,41014

Table 6.4 reveals that futures and index prices are co-integrated following Engle-Granger approach.
Howsoever Engle-Granger procedure reveals co-integration, more detailed and trustable results are offered by Johansen's approach.

6.2.2 Results of Johansen Approach

Johansen procedure firstly requires series integrated of the same order just as in the case with Engle-Granger approach. Unit root tests state that index futures and spot index price series are integrated of order 1 as duly. Since differenced forms are found to be stationary, trace and maximum eigenvalue tests are ready to be performed. These tests will show if the series have a long-run equilibrium level or not. In case that the series are co-integrated, Johansen suggests that there should be at least one way causality relation between these two variables. In order to capture the causality structure, EC models are supposed to be built. Significant lags of ECMs reveal the direction and the magnitude of the causality relation. The key point in error correction models is the designation of the lag lengths. As stated before, lag lengths are to be determined through SIC which has the most superior large sample properties. SIC suggests the lag lengths to be used in ECMs as 4 in 2005; 3 in 2006 and 5 in 2007 and 2008, 6in 2009 and 7 in 2010. In accordance with the specified SIC lag lengths, EC models are built. Then, the series are tested through trace and maximum eigenvalue statistics to decide if series have a long run relationship. Results of these tests are summarized below yearly:

Table 6.5 Johansen Test Results of 2005

Test	H_0	H_1	Eigenvalue	Test Stat.	p-value
Trace	r=0	r≤1	0,005597	40,23718	0,0004
	r≤1	r=2	0,002233	11,46112	0,0746
Max. Eigenvalue	r=0	r≤1	0,005597	28,77606	0,0016
	r≤1	r=2	0,002233	11,46112	0,0746

Trace and maximum eigenvalue tests conclude that there exist exactly 1 co-integration vector meaning that the two price series are co-integrated. In other words prices are expected to reach an equilibrium level. However, in the short run, some deviations may occur as Johansen procedure. These possible deviations are investigated by the results of ECMs. The proposed error correction model of 2005 is:

$$\Delta F_t = 3,14 - 0,38\Delta F_{t-1} - 0,15\Delta F_{t-2} - 0,11\Delta F_{t-3} - 0,04\Delta F_{t-4}$$
$$+0,31\Delta S_{t-1} + 0,14\Delta S_{t-2} + 0,09\Delta S_{t-3} + 0,06\Delta S_{t-4} - 0,01 ect_{t-1}$$

$$\Delta S_t = -2,83 + 0,09\Delta F_{t-1} + 0,07\Delta F_{t-2} + 0,02\Delta F_{t-3} + 0,02\Delta F_{t-4}$$
$$-0,08\Delta S_{t-1} - 0,06\Delta S_{t-2} - 0,01\Delta S_{t-3} + 0,006\Delta S_{t-4} - 0,0006 ect_{t-1} \tag{6.1}$$

Table 6.6 ECM Significance Statistics of 2005

Dependent:	**dfutures**			**dindex**	
	T	p-val		t	p-val
dfutures(-1)	-19,682	0,0000		4,2614	0,0000
dfutures(-2)	-7,4662	0,0000		3,0517	0,0023
dfutures(-3)	-5,4874	0,0000		0,8949	0,3709
dfutures(-4)	-2,5762	0,0100		1,0228	0,3064
dindex(-1)	19,5288	0,0000		-4,2817	0,0000
dindex(-2)	8,3584	0,0000		-2,8929	0,0038
dindex(-3)	5,4612	0,0000		-0,6995	0,4843
dindex(-4)	3,7856	0,0002		0,3238	0,7461
ect(-1)	-3,9584	0,0001		-1,1983	0,8428
Intercept	-2,7821	0,0054		0,4501	0,6527

Table 6.6 represents the outcomes of the two error correction models; one regressed on differenced futures and other on the differenced spot index prices as formulated in (6.1). When the model is regressed on futures, coefficients of index determine the power of the causality and when the regressand is index, futures coefficients explain the causality. For the year 2005, first four lags of the index are statistically significant. Every lag corresponds to a 5-minute interval as stated very early, thus spot index prices lead index futures about 20 minutes. In contrast, first two futures lags are significant meaning that futures market leads spot market by about 10 minutes. The relation is bidirectional but index prices are stronger in point of magnitude.

Table 6.7 Johansen Test Results of 2006

Test	**H_0**	**H_1**	**Eigenvalue**	**Test Stat.**	**p-value**
Trace	r=0	r≤1	0,004911	67,18453	0,0000
	r≤1	r=2	0,000231	3,01238	0,8751
Max. Eigenvalue	r=0	r≤1	0,004911	64,17216	0,0000
	r≤1	r=2	0,000231	3,01238	0,8751

According to Table 6.7, spot and derivative market prices move in a balanced way in the long run. Since the series are found to be co-integrated, the following ECM is constructed in 6.2:

$$\begin{aligned}\Delta F_t &= -0{,}11 - 0{,}09\Delta F_{t-1} - 0{,}01\Delta F_{t-2} + 0{,}01\Delta F_{t-3} \\ &+0{,}11\Delta S_{t-1} + 0{,}007\Delta S_{t-2} - 0{,}01\Delta S_{t-3} - 0{,}003ect_{t-1} \\ \Delta S_t &= -0{,}11 + 0{,}22\Delta F_{t-1} + 0{,}10\Delta F_{t-2} + 0{,}06\Delta F_{t-3} \\ &-0{,}18\Delta S_{t-1} - 0{,}10\Delta S_{t-2} - 0{,}06\Delta S_{t-3} - 0{,}006ect_{t-1}\end{aligned} \tag{6.2}$$

Table 6.8 ECM Significance Statistics of 2006

	dfutures			Dindex	
	T	p-val		t	p-val
dfutures(-1)	-6,5148	0,0000		14,1062	0,0000
dfutures(-2)	-0,7891	0,4300		6,5080	0,0000
dfutures(-3)	0,7732	0,4394		4,0163	0,0001
dindex(-1)	9,6733	0,0000		-13,3297	0,0000
dindex(-2)	0,5826	0,5601		-7,0394	0,0000
dindex(-3)	-1,0377	0,2994		-4,3682	0,0000
ect(-1)	-1,6033	0,1089		2,7567	0,0058
Intercept	-1,2750	0,2023		1,9868	0,0467

Bi-directional causality occurs between the two markets, with a stronger lead from futures to spot market. Table 6.8 uncovers that index leads futures by 5 minutes; while, futures prices lead index prices about 15 minutes. Outcomes expose that the direction of the inner causality is different in 2005 and 2006.

Table 6.9 Johansen Test Results of 2007

Test	H_0	H_1	Eigenvalue	Test Stat.	p-value
Trace	r=0	r≤1	0,004906	74,23369	0,0000
	r≤1	r=2	0,000665	8,84646	0,1900
Max. Eigenvalue	r=0	r≤1	0,004906	65,38723	0,0000
	r≤1	r=2	0,000665	8,84646	0,1900

Table 6.9 reveals that just as in the case of the years 2005 and 2006, in 2007 spot and derivatives markets are discovered to have a long run stable relationship. As suggested by Granger (1988), existence of 'at least one way causality' will be examined through error correction mechanism:

$$\Delta F_t = 1,54 - 0,02\Delta F_{t-1} - 0,01\Delta F_{t-2} + 0,03\Delta F_{t-3} + 0,01\Delta F_{t-4} + 0,01\Delta F_{t-5}$$
$$+0,04\Delta S_{t-1} - 0,07\Delta S_{t-2} - 0,02\Delta S_{t-3} - 0,01\Delta S_{t-4} - 0,04\Delta S_{t-5} - 0,005ect_{t-1} \tag{6.3}$$
$$\Delta S_t = 1,68 + 0,28\Delta F_{t-1} + 0,13\Delta F_{t-2} + 0,09\Delta F_{t-3} + 0,04\Delta F_{t-4} + 0,03\Delta F_{t-5}$$
$$-0,24\Delta S_{t-1} - 0,14\Delta S_{t-2} - 0,1\Delta S_{t-3} - 0,06\Delta S_{t-4} - 0,02\Delta S_{t-5} - 0,009ect_{t-1}$$

Table 6.10 ECM Significance Statistics of 2007

	Dfutures			**dindex**	
	T	p-val		t	p-val
dfutures(-1)	-1,9451	0,0518		18,7021	0,0000
dfutures(-2)	-0,7754	0,4381		8,2264	0,0000
dfutures(-3)	0,2674	0,7892		6,1146	0,0000
dfutures(-4)	0,9059	0,3650		3,0178	0,0026
dfutures(-5)	1,1169	0,2640		2,3526	0,0187
dindex(-1)	3,2408	0,0006		-17,6779	0,0000
dindex(-2)	-0,5890	0,5559		-9,7899	0,0000
dindex(-3)	-0,1576	0,8758		-7,1096	0,0000
dindex(-4)	-1,0251	0,3053		-4,8048	0,0000
dindex(-5)	-0,3213	0,7480		-1,6163	0,1061
ect(-1)	-2,0350	0,0419		3,2390	0,0012
Intercept	-1,7599	0,0785		3,4082	0,0007

In 2007, the pattern is two-sided but the leadership of futures prices becomes evident by Table 6.10. As index leads futures just by 5 minutes, derivative prices lead spot prices approximately 25 minutes. Price information of 2007 indicates growing price discovery power of futures market. Howsoever futures seem to be the leader of the market; the relation is not completely uni-directional.

Table 6.11 Johansen Test Results of 2008

Test	**H_0**	**H_1**	**Eigenvalue**	**Test Stat.**	**p-value**
Trace	r=0	r≤1	0,004577	76,65343	0,0000
	r≤1	r=2	0,000569	8,46190	0,2162
Max. Eigenvalue	r=0	r≤1	0,004577	68,19153	0,0000
	r≤1	r=2	0,000569	8,46190	0,2162

Similar to 2005, 2006 and 2007, existence of co-integration between markets is proved by Table 6.11. Test statistics points out that there is one co-integration vector present. Hence the linear combination of the prices is stationary. Thus, now we are able to construct the below model.

$$\begin{aligned}
\Delta F_t &= -2,42 - 0,02\Delta F_{t-1} - 0,04\Delta F_{t-2} + 0,03\Delta F_{t-3} - 0,01\Delta F_{t-4} + 0,04\Delta F_{t-5} \\
&+0,03\Delta S_{t-1} + 0,08\Delta S_{t-2} + 0,08\Delta S_{t-3} + 0,02\Delta S_{t-4} + 0,07\Delta S_{t-5} - 0,006ect_{t-1} \\
\Delta S_t &= -2,27 + 0,31\Delta F_{t-1} + 0,16\Delta F_{t-2} + 0,09\Delta F_{t-3} + 0,05\Delta F_{t-4} + 0,03\Delta F_{t-5} \\
&-0,29\Delta S_{t-1} - 0,16\Delta S_{t-2} - 0,09\Delta S_{t-3} - 0,06\Delta S_{t-4} - 0,03\Delta S_{t-5} + 0,007ect_{t-1}
\end{aligned} \tag{6.4}$$

Table 6.12 ECM Significance Statistics of 2008

	Dfutures			**Dindex**	
	T	p-val		t	p-val
dfutures(-1)	-1,7820	0,7480		22,0367	0,0000
dfutures(-2)	-0,3044	0,7608		10,6852	0,0000
dfutures(-3)	0,2157	0,8293		6,2275	0,0000
dfutures(-4)	-0,7555	0,4500		3,5147	0,0004
dfutures(-5)	0,2900	0,7718		2,6647	0,0077
dindex(-1)	2,4420	0,0146		-20,2843	0,0000
dindex(-2)	0,5471	0,5843		-10,7600	0,0000
dindex(-3)	0,5575	0,5772		-6,1132	0,0000
dindex(-4)	0,1521	0,8791		-4,5397	0,0000
dindex(-5)	0,5134	0,6077		-2,2602	0,0238
ect(-1)	-2,2920	0,0219		2,5333	0,0113
Intercept	-2,9129	0,0036		0,7717	0,4403

The significance analysis of model (6.4) is tabulated. Table 6.12 basically produces the same results with Table 6.10. Index prices lead derivatives about 5 minutes and lags about 25 minutes.

Table 6.13 Johansen Results of 2009

Test	**H_0**	**H_1**	**Eigenvalue**	**Test Stat.**	**p-value**
Trace	r=0	r≤1	0,006246	95,25183	0,0000
	r≤1	r=2	0,000547	7,64668	0,2819
Max. Eigenvalue	r=0	r≤1	0,006246	87,60515	0,0000
	r≤1	r=2	0,000547	7,64668	0,2819

In accordance with the results of the previous years, Johansen procedure concludes that there exists one co-integration vector which proves the stable long-run behavior of futures and spot markets. Revealing the equilibrium between markets, short-run behavior is the question of interest. This question is easily answered by the EC models below:

$$\Delta F_t = 1,86 + 0,009\Delta F_{t-1} + 0,001\Delta F_{t-2} - 0,005\Delta F_{t-3}$$
$$+0,01\Delta F_{t-4} + 0,02\Delta F_{t-5} + 0,01\Delta F_{t-6} + 0,03\Delta S_{t-1} - 0,01\Delta S_{t-2}$$
$$+0,008\Delta S_{t-3} + 0,01\Delta S_{t-4} - 0,02\Delta S_{t-5} - 0,01\Delta S_{t-6} - 0,009ect_{t-1}$$

(6.5)

$$\Delta S_t = 2,01 + 0,41\Delta F_{t-1} + 0,21\Delta F_{t-2} + 0,11\Delta F_{t-3}$$
$$+0,08\Delta F_{t-4} + 0,07\Delta F_{t-5} + 0,02\Delta F_{t-6} - 0,32\Delta S_{t-1} - 0,22\Delta S_{t-2}$$
$$-0,11\Delta S_{t-3} - 0,09\Delta S_{t-4} - 0,08\Delta S_{t-5} - 0,03\Delta S_{t-6} + 0,01ect_{t-1}$$

Table 6.14 ECM Significance Statistics of 2009

	Dfutures			**dindex**	
	T	p-val		t	p-val
dfutures(-1)	-0,0333	0,9734		29,1033	0,0000
dfutures(-2)	0,0973	0,9224		13,7093	0,0000
dfutures(-3)	-0,4965	0,6195		7,0983	0,0000
dfutures(-4)	0,6527	0,5140		5,6890	0,0000
dfutures(-5)	1,9438	0,0519		4,7703	0,0000
dfutures(-6)	1,2721	0,2034		1,6974	0,0896
dindex(-1)	4,1286	0,0000		-26,8778	0,0000
dindex(-2)	-0,9399	0,3473		-16,2936	0,0000
dindex(-3)	0,8060	0,4202		-8,6587	0,0000
dindex(-4)	0,2075	0,4356		-6,7816	0,0000
dindex(-5)	-1,7362	0,0825		-6,3964	0,0000
dindex(-6)	-0,1927	0,8471		-2,6855	0,0072
ect(-1)	-2,9473	0,0032		3,2870	0,0010
İntercept	2,0834	0,0372		1,9074	0,0565

Model (6.5) investigated through the significance table tell that index market leads the futures market about 5 minutes and lags it approximately 25 minutes exactly as in the case of 2007 and 2008 according to Table 6.14.

Table 6.15 Johansen Results of 2010

Test	**H_0**	**H_1**	**Eigenvalue**	**Test Stat.**	**p-value**
Trace	r=0	r≤1	0,006086	85,27211	0,0000
	r≤1	r=2	0,00033	4,37622	0,6872
Max. Eigenvalue	r=0	r≤1	0,006086	80,89590	0,0000
	r≤1	r=2	0,00033	4,37622	0,6872

Not surprisingly, existence of one co-integration vector is proved by the help of Trace and Maximum Eigenvalue Tests for 2010 market structure. Assuring that in long-run two markets will move accordingly, the short-run movements need to be discovered in order to differentiate between the markets:

$$\begin{aligned}
&\Delta F_t = 0,97 - 0,01\Delta F_{t-1} - 0,008\Delta F_{t-2} - 0,003\Delta F_{t-3} - 0,006\Delta F_{t-4} \\
&+0,002\Delta F_{t-5} - 0,01\Delta F_{t-6} - 0,001\Delta F_{t-7} + 0,02\Delta S_{t-1} - 0,003\Delta S_{t-2} \\
&-0,003\Delta S_{t-3} + 0,01\Delta S_{t-4} + 0,01\Delta S_{t-5} + 0,008\Delta S_{t-6} + 0,01\Delta S_{t-7} - 0,005 ect_{t-1} \\
&\Delta S_t = 0,96 + 0,47\Delta F_{t-1} + 0,31\Delta F_{t-2} + 0,18\Delta F_{t-3} + 0,12\Delta F_{t-4} \\
&+0,09\Delta F_{t-5} + 0,08\Delta F_{t-6} + 0,03\Delta F_{t-7} - 0,45\Delta S_{t-1} - 0,29\Delta S_{t-2} \\
&-0,18\Delta S_{t-3} - 0,12\Delta S_{t-4} - 0,08\Delta S_{t-5} - 0,07\Delta S_{t-6} - 0,02\Delta S_{t-7} + 0,02 ect_{t-1}
\end{aligned} \tag{6.6}$$

Table 6.16 ECM Significance Statistics of 2010

	Dfutures			**dindex**	
	T	p-val		t	p-val
dfutures(-1)	-1,8913	0,0586		34,0326	0,0000
dfutures(-2)	-0,6074	0,5436		19,3901	0,0000
dfutures(-3)	-0,0231	0,9815		10,9540	0,0000
dfutures(-4)	-0,4409	0,6593		7,7294	0,0000
dfutures(-5)	0,1798	0,8573		5,6538	0,0000
dfutures(-6)	-1,1684	0,2427		5,0416	0,0000
dfutues(-7)	-0,0903	0,9280		2,0290	0,0425
dindex(-1)	3,1334	0,0017		-37,6791	0,0000
dindex(-2)	-0,2449	0,8065		-20,4376	0,0000
dindex(-3)	-0,0741	0,9409		-12,2259	0,0000
dindex(-4)	1,3618	0,1733		-8,2510	0,0000
dindex(-5)	1,0638	0,2874		-5,5803	0,0000
dindex(-6)	0,7395	0,4596		-5,6468	0,0000
dindex(-7)	1,1611	0,2456		-1,7075	0,0000
ect(-1)	-1,2957	0,1951		5,1697	0,0877
intercept	0,8144	0,4154		0,6980	0,4852

According Table 6.16, 2010 is the year that leading role of the futures market becomes more notable. Table shows that futures market is about 35 minutes ahead of the spot market whereas, spot market can only lead futures market by 5 minutes.

6.3 Causality Results

In every single year of the analysis, cash and futures market prices are found to be co-integrated. Due to the existence of this long-run behavior, the causality models should be built on the VEC models rather than VAR models.

6.3.1 Results of Linear Granger Causality

Granger Linear Causality test allows us to test whether one market lag the other in a one step testing procedure. Findings of the ECMs will become more meaningful after gaining the consequence of the linear causality test. F or the Wald test may bring about; bidirectional relation, no relation or one way causality. The test is performed once for the searching the futures market leading ability and once for that of spot market. To clarify, the hypotheses to be tested are:

- H_{01} : Futures prices do not cause spot index prices,
- H_{02} : Spot index prices do not cause futures prices.

Both Wald and F statistics are computed and Table 6.17 is obtained:

Table 6.17 Linear Granger Causality Tests regarding 'H_0: F does not cause S'

H_0: F does not cause S	**Wald Test**			**F Test**	
	Test Statistic	**Df**	**p-value**	**Test Statistic**	**CV**
2005	20,8899	4	0,0003	3,8838	2,1003
2006	208,4316	3	0,0000	44,9721	2,2147
2007	376,2351	5	0,0000	58,2153	2,0102
2008	504,0758	5	0,0000	77,6679	2,0102
2009	821,7448	6	0,0000	115,5287	1,9390
2010	1096,389	7	0,0000	141,4281	1,8805

In respect of Wald and F tests, for all of the years, the stated null hypothesis is rejected at 95% confidence level as shown in Table 6.17. This denotes that from 2005 up to 2010, futures market prices lead the spot market prices. However, without testing the second hypothesis, it is not proper to express the price leadership of the derivatives market.

Table 6.18 Linear Granger Causality Tests regarding ‘S does not cause F’

H_0: S does not cause F	Wald Test			F Test	
	Test Statistic	Df	p-value	Test Statistic	CV
2005	390,9830	4	0,0000	70,2441	2,1003
2006	100,9712	3	0,0000	20,9479	2,2147
2007	15,9875	5	0,0069	2,7640	2,0102
2008	6,5417	5	0,2570	0,9596	2,0102
2009	22,2803	6	0,0011	3,6135	1,9390
2010	12,7032	7	0,0797	1,8245	1,8805

Table 6.18 reveals that apart from 2008 and 2010, spot market prices also lead the futures market prices significantly. Combining the consequences of Tables 6.17 and 6.18, the final decision of linear causality is easily given. For the years 2005, 2006, 2007 and 2009 the causality relation is bi-directional indicating that each market has the ability to lead/lag the other at some periods of time. On the other hand, in 2008 and 2010, the relationship turns out to be a uni-directional and market leadership belongs to futures prices in both of the years.

In the light of the findings of Granger linear causality, the lead/lag time measurements coming from ECM tables are to be re-interpreted. Although the relation is two-sided through 2005, 2006 and 2007, in 2005 cash index prices are more likely to lead the futures prices. Conversely, in 2006 and 2007 the lead of futures market is quite stronger compared to lag of futures prices. The year 2008 brings a change in the price discovery process and futures market significantly becomes the only price leader. Table 6.19 summarizes the price discovery structure.

Table 6.19 Lead-lag Structure of Raw Data

Year	Relation	F → C	C → F
2005	bi-directional	10 min.	20 min.
2006	bi-directional	15 min.	5 min.
2007	bi-directional	25 min.	5 min.
2008	uni-directional	25 min.	–
2009	bi-directional	25 min.	5 min.
2010	uni-directional	35 min.	–

As discussed earlier in Data Description and Adjustment Chapter, the existing short run linear causality may arise from the infrequent trading problem. If this hypothesis is correct, the found bi-directional relation may turn out to be uni-directional favoring spot market or at least futures market become less powerful in price discovery process.

6.3.1.1 Linear Granger Causality of ARMA Filtered Prices

Filtering does not affect the co-integration property of the series but the linear causality and ECM results are different from the previous ones. First the Granger Linear Causality test findings of the filtered price series are illustrated to see the overall pattern of causality and then the corresponding ECM tables are investigated to highlight the time based relationship.

Table 6.20 Linear Granger Causality Tests regarding ‘H_0: F does not cause S’ of Filtered Data

H_0: F does not cause S	Wald Test			F Test	
	Test Statistic	df	p-value	Test Statistic	CV
2005	6,0054	4	0,1987	1,6756	2,1004
2006	16,9008	3	0,0007	5,9357	2,2147
2007	25,4785	5	0,0001	7,6404	2,0102
2008	343,0392	5	0,0000	73,1661	2,0102
2009	105,5114	6	0,0000	20,4295	1,9390
2010	61,5335	7	0,0000	19,1420	1,8805

Table 6.21 Linear Granger Causality Tests regarding ‘H_0: S does not cause F’ of Filtered Data

H_0: S does not cause F	Wald Test			F Test	
	Test Statistic	df	p-value	Test Statistic	CV
2005	121,1758	4	0,0000	30,1569	2,1004
2006	12,6842	3	0,0054	10,2838	2,2147
2007	84,6154	5	0,0000	15,1792	2,0102
2008	4,6386	5	0,4615	2,0054	2,0102
2009	159,3072	6	0,0000	18,8425	1,9390
2010	3,3934	7	0,7581	1,4711	1,8805

Tables 6.20 and 6.21 indicate that after ARMA filtering, the direction of the causality does not change drastically. For the years 2006 and 2007 the bi-directional relation stays unchanged. On the other hand, in 2005, futures prices are found not to lead the cash market prices. Even though the direction of the relationship survives, the effect of the cash market on the derivatives becomes notable.

The detailed lead/lag pattern of ARMA filtered prices is revealed yearly by significance statistics of the error correction models as below:

Table 6.22 Significance Statistics of 2005 ARMA Filtered ECM

	darmafut			darmaind	
	t	p-val		t	p-val
darmafut(-1)	-0.7192	0.4720		-1.0198	0.3079
darmafut(-2)	7.4616	0.0000		-0.9175	0.3589
darmafut(-3)	6.0955	0.0000		-2.0816	0.0374
darmafut(-4)	7.7793	0.0000		0.4167	0.6769
darmaind(-1)	9.4722	0.0000		-1.6664	0.0957
darmaind(-2)	2.4194	0.0156		-1.9882	0.0468
darmaind(-3)	2.8901	0.0039		0.9964	0.3191
darmaind(-4)	4.7396	0.0000		0.6143	0.5390
ect(-1)	-24.5652	0.0000		4.6125	0.0000
intercept	-0.0492	0.9607		0.0490	0.9609

Linear Granger causality results illustrated by Table 6.22 points out that in the year 2005, the causality is one-sided, information moving from spot prices to futures prices. In the light of the foregoing, above ECM table can be interpreted. As stated in Table 6.20, futures market does not have any predictive power on spot market but spot market information flow happens about 20 minutes earlier in spot compared to futures market. The leadership of the cash prices is obvious in 2005.

Table 6.23 Significance Statistics of 2006 ARMA Filtered ECM

	darmafut			darmaind	
	t	p-val		t	p-val
darmafut(-1)	-4.0716	0.0000		3.5349	0.0004
darmafut(-2)	-3.3719	0.0007		3.8208	0.0084
darmafut(-3)	-1.7295	0.0838		2.1909	0.0285
darmaind(-1)	2.5149	0.0119		-1.0190	0.3082
darmaind(-2)	-2.0312	0.0423		0.9676	0.3333
darmaind(-3)	-1.1973	0.2312		0.3260	0.7444
ect(-1)	-11.6607	0.0000		14.9366	0.0000
intercept	-0.0614	0.9510		0.1119	0.9109

As understood from Table 6.23 the bidirectional relationship persists between spot and futures prices after filtering. ISE 30 index leads ISE 30 index futures about 10 minutes and lags approximately 15 minutes. As compared to raw data outcomes, lead of index gets stronger but still futures lead is longer.

Table 6.24 Significance Statistics of 2007 ARMA Filtered ECM

	darmafut			**darmaind**	
	t	p-val		t	p-val
darmafut(-1)	6,2589	0,0000		4,1999	0,0000
darmafut(-2)	5,1833	0,0000		-2,1900	0,0103
darmafut(-3)	6,5344	0,0000		-2,1143	0,0090
darmafut(-4)	7,2210	0,0000		-2,4517	0,0142
darmafut(-5)	8,6414	0,0000		-0,7339	0,4630
darmaind(-1)	6,1805	0,0000		-6,6763	0,0000
darmaind(-2)	4,2199	0,0000		-4,0138	0,0001
darmaind(-3)	1,8020	0,0623		-4,6503	0,0000
darmaind(-4)	1,6313	0,0594		-3,2622	0,0011
darmaind(-5)	1,8216	0,0685		2,1044	0,0354
ect(-1)	-32,0275	0,0000		5,5616	0,0000
intercept	0,0355	0,9716		-0,0109	0,9913

ECM of the filtered data, in 2007, concludes that futures market leads cash prices by 25 minutes while, spot prices leads the futures prices about 5 minutes. After filtering, futures lead by 20 minutes while lead of cash market rises to 10 minutes.

Table 6.25 Significance Statistics of 2008 ARMA Filtered ECM

	darmafut			**darmaind**	
	t	p-val		t	p-val
darmafut(-1)	-0.3052	0.7602		18.3999	0.0000
darmafut(-2)	0.1975	0.8434		7.5516	0.0000
darmafut(-3)	1.8928	0.0584		4.8442	0.0000
darmafut(-4)	-0.5612	0.5746		2.7010	0.0190
darmafut(-5)	2.3866	0.0170		2.3627	0.0182
darmaind(-1)	0.6040	0.5459		-17.4213	0.0000
darmaind(-2)	-0.6515	0.5147		-8.6688	0.0000
darmaind(-3)	-0.4318	0.6659		-4.2601	0.0000
darmaind(-4)	-0.1995	0.8419		-3.1573	0.0016
darmaind(-5)	-1.7822	0.0747		-2.3716	0.0177
ect(-1)	-4.9227	0.0000		6.1875	0.0000
intercept	0.0154	0.9877		0.0255	0.9796

The strict uni-directional structure is captured in 2008. Futures market is the price discovery tool for the spot market prices by 25 minutes exactly as stated earlier by the raw data analysis.

Table 6.26 Significance Statistics of 2009 ARMA Filtered ECM

	darmafut			**darmaind**	
	t	p-val		t	p-val
darmafut(-1)	1,2920	0,1963		11,5105	0,0000
darmafut(-2)	1,3229	0,1836		-0,9084	0,0363
darmafut(-3)	1,8681	0,0618		-3,3454	0,0008
darmafut(-4)	2,3808	0,0173		-1,8677	0,0186
darmafut(-5)	2,3321	0,0197		-2,0506	0,0403
darmafut(-6)	0,9652	0,3344		-1,8073	0,0707
darmaind(-1)	2,9481	0,0032		-11,4071	0,0000
darmaind(-2)	-0,5839	0,5592		-6,6609	0,0000
darmaind(-3)	-0,2328	0,8159		-1,9434	0,0520
darmaind(-4)	1,2664	0,2054		-1,7975	0,0723
darmaind(-5)	0,4078	0,6834		-3,1495	0,0016
darmaind(-6)	0,2004	0,8411		-0,9460	0,3441
ect(-1)	-11,0607	0,0000		15,9173	0,0000
intercept	0,0139	0,9889		-0,1329	0,8942

Table 6.27 Significance Statistics of 2010 ARMA Filtered ECM

	darmafut			**darmaind**	
	t	p-val		t	p-val
darmafut(-1)	12,9189	0,0000		-9,5238	0,0000
darmafut(-2)	2,6595	0,0078		5,1627	0,0000
darmafut(-3)	-1,5152	0,1297		5,7231	0,0000
darmafut(-4)	-1,5536	0,1203		6,0816	0,0000
darmafut(-5)	-1,7918	0,0732		8,3534	0,0000
darmafut(-6)	-1,7998	0,0719		5,0261	0,0000
darmafut(-7)	-1,8317	0,0670		6,8706	0,0000
darmaind(-1)	-0,2011	0,3321		7,8275	0,0000
darmaind(-2)	-0,7149	0,6987		5,0717	0,0000
darmaind(-3)	-0,3409	0,4017		4,6203	0,0000
darmaind(-4)	-0,3043	0,4009		6,7216	0,0000
darmaind(-5)	0,4418	0,6586		6,3272	0,0000
darmaind(-6)	-2,0742	0,0381		5,6197	0,0000
darmaind(-7)	-0,5138	0,6074		5,4379	0,0000
ect(-1)	7,9791	0,0000		-32,7975	0,0000
intercept	-0,2740	0,7840		0,0506	0,9596

Table 6.26 produces exactly same results as discovered in non-filtered case, futures market leads cash market about 25 minutes while cash market leads futures market only 5 minutes.

ARMA filtered results hold with the non-filtered analysis results. In both examinations, spot prices do not have any leading power on derivatives prices. On the other hand, futures prices of the year 2010 lead spot prices about 35 minutes.

ARMA filtering can be accepted successful because the significant autocorrelation in each series seem to die out after filtering. In other words, filtering removes the drawbacks of infrequent trading successfully as expected. The results of the analysis free of infrequent trading are more dependable.

The outcomes of the analysis through ARMA filtering are summarized in table 6.28 ARMA filtering changed the results of the first three years. For 2005, the relationship turns out to be uni-directional after adjustment. Leading effect of futures market becomes unimportant while futures lagging affect stays the same. For 2006 and 2007, futures power declines 5 minutes whereas, cash market strength increases as the same amount.

Table 6.28 Lead-lag Structure of ARMA Filtered Data

Year	Relation	F → C	C → F
2005	uni-directional	–	20 min.
2006	bi-directional	15 min.	10 min.
2007	bi-directional	20 min.	10 min.
2008	uni-directional	25 min.	–
2009	bi-directional	25 min.	5 min.
2010	uni-directional	35 min.	–

We know that the results of this study can be compared to many studies from literature because of the common 5-minute time-span used. The empirical findings will be discussed and similarities of Turkish Market and the foreign markets will be revealed. However, since, previous studies in Turkey differ from our study both in time-span and length of dataset, the direct comparison may not be fair. Thus, apart from yearly analysis, we also investigate the dataset as a whole and collect the results to compare studies from Turkey.

According to Table 6.29, if we perform the analyisis through the whole dataset, we reach the conclusion that there exists bidirectional relation between markets. The detailed investigation of ECM of the 'whole-period' indicates that lead of spot market to futures market is stronger. The general analysis results that futures prices lead spot prices approximately 20 minutes while they lag the spot prices only about 5 minutes. In the light of this information, previous Turkish studies and our study can be compared considering the dataset they cover. Nevertheless, without having the same time-intervals still one cannot directly make a comparison. But we can freely say that the most probable reason of having completely different outcomes is the time-span used.

Table 6.29 Linaer Causality Findings of 'whole-period' Investigation

Tested Hypothesis	Wald Test			F Test	
	Test Statistic	df	p-value	Test Statistic	CV
H_0: F does not cause S	3033,727	11	0,0000	127,112014	1,517441
H_0: S does not cause F	164,104	11	0,0000	10,392767	1,517441

6.3.2 Results of Nonlinear Granger Causality

Linear causality and ECM results revealed the initials of the inter-market structure but the linear analysis misses one important point. The underlying relation between markets may be nonlinear rather than being linear. This suspect will be examined through nonlinear causality analysis. The basic idea of the analysis is that if linear investigation is not sufficient to explain the relation, the corresponding residuals keep the true inner connection. For this reason, nonlinear structure will be derived from the VECM residuals.

Table 6.30 shows that in our results, no evidence of nonlinear causality from futures market to cash market or from cash prices to futures prices. In other words, there exists no nonlinear predictivity power of either market on the other.

Table 6.30 Non-linear Causality Results

CV = ±1,96		Lags (Lx=Ly)						
	Time Span	1	2	3	4	5	6	7
null hypothesis: F does not cause C	2005	0,5709	-1,8499	-1,2900	0,0000	0,0000	0,0000	0,0000
	2006	4,6578	-1,3546	-1,8324	0,0000	0,0000	0,0000	0,0000
	2007	1,5253	-2,7365	-1,4645	-1,0521	0,0000	0,0000	0,0000
	2008	1,6092	-1,6374	-1,4078	-2,9506	0,0000	0,0000	0,0000
	2009	2,1842	1,9450	1,1384	-1,8723	0,0000	0,0000	0,0000
	2010	1,9336	1,7824	-1,0506	-0,8621	0,0000	0,0000	0,0000
null hypothesis: C does not cause F	2005	-0,0753	-0,4433	0,0000	0,0000	0,0000	0,0000	0,0000
	2006	0,0793	-1,0958	-1,4271	0,0000	0,0000	0,0000	0,0000
	2007	0,1853	4,1658	-1,2704	0,0000	0,0000	0,0000	0,0000
	2008	0,1206	-1,0883	0,0000	0,0000	0,0000	0,0000	0,0000
	2009	0,9873	1,4253	0,0049	0,0000	0,0000	0,0000	0,0000
	2010	0,5318	-0,7246	-0,5302	0,0000	0,0000	0,0000	0,0000

CHAPTER 7

SUMMARY AND CONCLUSION

Nowadays, derivatives transactions become very popular in Turkey. Increase in the number of derivatives market investors makes it charming to study the initials of market structure. In this work, we study the interrelation of Turkish spot and futures markets. The relationship between these markets is quite important because the strength of this bound may lead the investors to profit or hedge themselves. Despite the richness of the foreign literature on this topic, in Turkey, there are only a few projects and efforts in this scope. For this reason, Turkish cash and futures market structures are analyzed in terms of the lead-lag patterns.

In accordance with this purpose, a detailed data set covering transaction date, time, prices and many more is collected from ISE and TURKDEX which are the institutions of cash and derivatives markets, respectively. The data set ranges from the day that the derivatives market is founded in Turkey (2005) to the end of 2010. In other words, the information belonging to every single time that both marketsare current can be reached over this data set. For the first time, intraday data is used in lead-lag pattern investigation through our study. The daytime is divided into 5-minutes interval and econometric analysis held on the prices which correspond to the related intervals. Moreover, not only the raw data but also the filtered prices examined in order to avoid misleading results due to the infrequent trading problem.

The lead-lag relationship is handled with co-integration and causality analysis. Co-integration aims to explain whether the prices of the two markets reach an equilibrium level. If this is true, we guarantee that in the long-run, these markets will behave very similarly but they may just act differently in short-term period. The strength of the short-run deviations is the key to investors to take quicker positions to hedge their investments. In other words, knowing that the prices will follow the same pattern in the long-run, the investors have the chance to take the advantage of the short-run differences. To be able to benefit from this chance, the short-run structure is needed to be analyzed through causality tests. Causality results mainly reveal that which market flow information quicker than the other. That is to say, empirical findings of causality study provide Turkish investors with the approximate time-delay between markets to take suitable positions.

In this work, the empirical analysis is divided into three parts. First, unit root tests are performed. Then, co-integration relation is investigated and in the last step, causality between markets is tested. Order of the steps of methodology is quite important because occurances in one step drastically change the application of the following steps. To illustrate, co-integration requires I(1) series, therefore unit root conditions are very crucial. Moreover, the way we test the existence of causality changes according to co-integration relation. In other words, testing co-integration is essential before analyzing the causality.

Findings of this study are quite attracting since the study seem to reveal the main differences between Turkish spot and derivatives markets. Firstly, existence of co-integration between prices is proven by eigenvalue and trace tests. Yearly results proclaim that from the establishment of TURKDEX until the end of 2010, equilibrium level is reached meaning that not in a single year long-run expectations of the markets diverge from each other. Moving to short-run deviations, except from 2008 and 2010, bi-directional relation exists between markets. In 2005, the relation is bi-directional but cash market prices lead futures prices by 20 minutes, wheras; futures lead spot markets by 10 minutes. In 2006, 2007 and 2009, the bi-directional relation is observed similar to 2005 but with one crucial difference that for these years futures market leadership is stronger. While futures prices lead spot prices by at least 15 minutes, they lag spot prices only by 5 minutes. This direction change occurs most probably due to the trading inequalities between markets. In the year of 2005, which is the establishment year of TURKDEX, the daily trading period covers only 4 hours. Thus it is normal to observe that a new information flows in spot market quicker than the futures market. When we look at 2008 and 2010 results, the picture is completely different owing to the uni-directional relationship. Futures market leads spot market by approximately 25 and 35 minutes in 2008 and 2010, respectively with no feedback from the spot market.

Empirical analysis on the raw data gives a clear picture of the intermarket linkage but in order to avoid infrequent trading effects, analyses on the filtered data are performed and results are collected. Raw and filtered data findings match for 2008, 2009 and 2010. For 2005, only price discoverer is found to be cash market. Filtered data say that cash markets leads futures market by 20 minutes without lagging the futures prices. According to ARMA filtering, futures prices laed spot prices about 20 and 25 minutes for 2006 and 2007, repectively while futures lag the spot market by 10 minutes in both years. When we compare the findings with the literature, we see that Turkey market structure have similar patterns with other country markets. Harris (1989), Chan (1992), Abhyankar (1998), Alphonse (2000), Ryoo and Smith (2004), Bhatia (2007) and many more have reached the same conclusion with us that examined through 5-minute intervals, futures market leadership is significantly stronger than the other way. However, two main studies conducted in Turkey do not reach a consensus with the results of our study. Özen (2008) finds that cash market leads the spot market. Similarly, Kapusuzoğlu (2010) claims that at the end of ISE 100 based analyses, cash market is found to be the price discoverer. But our study is significantly different from both two papers in the time-span worked. Unlike us, Özen (2008) and Kapusuzoğlu (2010) work on the daily closing prices. As mentioned earlier, we keep away from using daily prices since a day is not an informative time-span to specify the lead-lag relationship and most probably there are times in a day that futures market lead the spot market or the other way around.

To summarize, Turkish Derivatives and Spot Markets are efficient in long-term letting only short term deviations. Short-run deviations signal that futures transactions flow information quicker than spot transactions. The time delay resulting from the pace of information flow is the key point enabling investors to take positions. This thesis work gives clue to Turkish investors that futures prices lead spot prices about 15-35 minutes in Turkey.

REFERENCES

[1] Abhyankar A., *Linear and nonlinear Granger causality: Evidence from the U.K. stock index futures market*. The Journal of Futures Markets, vol. 18, no. 5, pp. 519-540, 1998.

[2] Alphonse P., *Efficient price discovery in stock index cash and futures markets*. Annales D'économie et de Statistique, vol. 60, 2000.

[3] Ansı A. and Ouda O. B., *How options markets affect price discovery on the spot markets: A survey of the empirical literature and synthesis*. International Journal of Business and Management, vol. 4, no. 8, 2009.

[4] Anthony J. H., *The interrelation of stock and options market trading-volume data.* The Journal of Finance, vol. 43, no. 4, pp. 949-964, 1988.

[5] Antoniou A. and Garrett I., *To what extent did stock index futures contribute to the October 1987 stock market crash?*. The Economic Journal, vol. 103, no. 421, pp.1444-1461, 1993.

[6] Asche F. and Guttormsen A. G., *Lead lag relationship between Futures and spot prices*. Working Paper Series, Institute for research in economics and business administration Bergen, 2002.

[7] Baek E. and Brock W., *A General Test for Nonlinear Granger Causality: Bivariate Model. Working Paper*, Iowa State University and University of Wisconsin, Madison.

[8] Bhattacharya M., *Price changes of related securities: The case of call options and stocks*. The Journal of Financial and Quantitative Analysis, vol. 22, no. 1, pp.1-15, 1987.

[9] Booth G. G., So R. W.and Tse Y., *Price discovery in the German equity index derivatives markets*. The Journal of Futures Markets, vol. 19, no. 6, pp.619-643, 1999.

[10] Chan K., Chan K. C., Karolyi G. A., *Intraday volatility in the stock index and stock index futures markets.* The Review of Financial Studies, vol. 4, no. 4, pp.657-684, 1991.

[11] Chan K., *A further analysis of the lead-lag relationship between the cash market and stock index futures market*. The Review of Financial Studies, vol. 5, no. 1, pp.123-152, 1992.

[12] Chung and Chuang C., *International information transmissions between stock index futures and spot markets: The case of futures contracts related to Taiean index.* Tamsui Oxford Journal of Management Sciences, vol. 19, no. 1, pp.51-78, 2003.

[13] Ciner C., *Energy Shocks and Financial Markets: Nonlinear Linkages.* Studies in Nonlinear Dynamics and Econometrics, vol. 5, no. 3, 2001.

[14] DeJong F. and Nijman T., *High frequency analysis of lead-lag relationships between financial markets*. Journal of Empirical Finance, vol. 4, pp. 259-277, 1997.

[15] DeJong F. and Donders M. W. M., *Intraday lead-lag relationships between the futures, options and stock market.* European Finance Review, vol. 1, pp. 337-359, 1998.

[16] Finnerty J. E. and Park H. Y., *Stock index futures: Does the tail wag the dog?.* Financial Analysts Journal, 1987.

[17] Fleming J., Ostdiek B. and Whaley R. E., *Trading costs and the relative rates of price discovery in stock futures and options markets.* The Journal of Futures Markets, vol. 16, no. 4, pp. 353-387, 1996.

[18] Floros C. and Vougas D. V., *Led-lag relationship between futures and spot markets in Greece: 1999-2001*. International Research Journal of Finance and Economics, vol. 7, 2007.

[19] Fujihara R. A. and Mougoue M., *An Examination of Linear and Nonlinear Causal Relationships between Variability and Volume in Petroleum Futures Markets.* The Journal of Futures Market, vol. 17, no. 4, pp. 385-416, 1997.

[20] Gwilym O. A. and Buckle M., *The lead-lag relationship between the FTSE100 stock index and its derivative contracts.* Applied Financial Economics, vol. 11, pp. 385-393, 2001.

[21] Harris L., *The October 1987 S&P 500 stock-futures basis*. The Journal of Finance, vol. 44, no. 1, pp. 77-99, 1989.

[22] Hasan M., *An alternative approach in investigating lead-lag relationships between stock and stock index futures markets – comment.* Applied Financial Economics Letters, vol. 1, pp. 125-130, 2005.

[23] Herbst F., McCormack J. P. and West E. N., *Investigation of a lead-lag relationship between spot stock indices and their futures contracts*. The Journal of Futures Markets, vol. 7, no. 4, pp. 373-381, 1987.

[24] Hiemstra C. and Jones J. D., *Monte Carlo Results for a Modified Version of the Baek and Brock Nonlinear Granger Causality Test.* Working Paper, University of Stratchclyde and Securities and Exchange Commision, 1993.

[25] Hiemstra C. and Jones J. D., *Testing for Linear and Nonlinear Granger Causality in the Stock Price – Volume Relation.* The Journal of Finance, vol. 49, no. 5, pp. 1639-1664, 1994.

[26] Hsieh D. A., *Implications of Nonlinear Dynamics for Financial Risk Management*. Journal of Financial and Quantitative Analysis, vol. 28, pp. 41-64, 1993.

[27] Iıhara Y., Kato K. and Tokunaga T., *Intraday return dynamics between the cash and the futures markets in Japan.* The Journal of Futures Markets, vol. 16, no. 2, pp. 147-162, 1996.

[28] Kang J., Lee C. J. and Lee S., *An empirical investigation of the lead-lag relations of returns and volatilities among the KOSPI 200 spot, futures and options markets and their explanations*. Journal of Emerging Market Finance, vol. 5, no. 3, pp. 235-261, 2006.

[29] Kapusuzoğlu A. and Demir A., *The Analysis of the Effects of Derivatives Exchange (DE) Transactions on the Market Efficiency of Istanbul Stock Exchange (ISE) National 100 Index and on Spot Market Transaction Prices*. African Journal of Business Management, vol. 4, no. 2, pp. 242-247, 2010.

[30] Kavussanos M. G., Visvikis I. D. and Alexakis P. D., *The lead-lag relationship between cash and stock index futures in a new market.* European Financial Management, vol. 14, no. 5, pp. 1007-1025, 2008.

[31] Kawaller I. G., Koch P. D. and Koch T. W., *The temporal price relationship between S&P 500 futures and the S&P 500 index*. The Journal of Finance, vol. 42, no. 5, pp. 1309-1329, 1987.

[32] Kenourgios D. F., *Price discovery in the Athens derivatives exchange: Evidence for the FTSE/ASE-20 futures market*. Economic and Business Review, vol. 6, no. 3, pp. 229-243, 2004.

[33] Khoury N., Perrakis S., Savor M. and Czerwonko M., *Price discovery in internationally cross-listed options and underlying equities: The Canadian case.*

[34] Lafuente J. A., *Intraday price and volatility relationships between the Ibex 35 spot and futures markets*. Spanish Economic Review, vol. 4, pp. 201-220, 2002.

[35] Manaster S. and Rendleman R. J., *Option prices as predictors of equilibrium stock prices.* The Journal of Finance, vol. 37, no. 4, pp. 1043-1057, 1982.

[36] Mattos F. and Garcia P., *Price discovery in thinly traded markets: Cash and futures relationships in Brazilian agricultural futures markets*. NCR-134 Conference on Applied Commodity Price Analysis, Forecasting and Market Risk Management, St. Louis, Missouri, 2004.

[37] Milunovich G. and Joyeux R., *Testing market efficiency and price discovery in European carbon markets.* Macquaire Economics Research Papers, 2007.

[38] Min J. H. and Najand M., *A further investigation of the lead-lag relationship between the spot market and stock index futures: Early evidence from Korea.* The Journal of Futures Markets, vol. 19, no. 2, pp. 217-232, 1999.

[39] Nelson D. B., *Conditional Heteroskedasticity in Asset Returns: A New Approach.* Econometrica, vol. 59, no. 2, pp. 347-370, 1991.

[40] Nieto M. L., Fernandez A. and Munoz M. J., *Market efficiency in the Spanish derivatives markets: An empirical analysis*. IAER, vol. 4, no. 4, pp. 349-355, 1998.

[41] Özen E., Bozdoğan T. and Zügül M., *the Relationship of the Causality between Price of Futures Transactions Underlying Stock Exchange and Price of Cash Market: The Case of Turkey*. Middle Eastern Finance and Economics, vol. 4, 2009.

[42] Pizzi M. A., Economopoulos A. J. and O'Neill H. M., *An examination of the relationship between stock index cash and futures markets: A co-integration approach.* The Journal of Futures Markets, vol. 18, no. 3, pp. 297-305, 1998.

[43] Pradhan K. C. and Bhat K. S., *An empirical analysis of price discovery, causality and forecasting in the Nifty futures markets*. International Research Journal of Finance and Economics, vol. 26, pp. 83-92, 2009.

[44] Raju M. T. and Karande K., *Price discovery and volatility on NSE futures market.* Working Paper Series, Securities and Exchange Board of India, 2003.

[45] Roope M. and Zurbruegg R., *The intraday price discovery process between the Singapore exchange and Taiwan futures exchange.* The Journal of Futures Markets, vol. 22, no. 3, pp. 219-240, 2002.

[46] Ryoo H. J. and Smith G., *The impact of stock index futures on the Korean stock market.* Applied Financial Economics, vol. 14, pp. 243-251, 2004.

[47] Silvapulle P. and Moosa I. A., *The relationship between spot and futures prices: Evidence from the crude oil market.* The Journal of Futures Markets, vol. 19, no. 2, pp. 175-193, 1999.

[48] Srinivasan P. and Bhat K. S., *Spot and futures markets of selected commercial banks in India: What causes what?*. International Research Journal of Finance and Economics, vol. 31, pp. 28-40, 2009.

[49] Stephan J. A. and Whaley R. E., *Intraday price change and trading volume relations in the stock and stock option markets.* The Journal of Finance, vol. 45, no. 1, pp. 191-220, 1990.

[50] Stoll H. R. and Whaley R. E., *The dynamics of stock index and stock index futures returns*. The Journal of Finance and Quantitative Analysis, vol. 25, no. 4, pp. 441-468, 1990.

[51] Tse Y. K. and Chan W. S., *The lead-lag relation between the S&P500 spot and futures markets: An intraday-data analysis using a threshold regression model*. The Japanese Economic Review, 2009.

[52] Turkington J. and Walsh D., *Price discovery and causality in the Australian share price index futures market*. Australian Journal of Management, vol. 24, no. 2, pp. 97-113, 1999.

[53] Wahab M. and Lashgari M., *Price dynamics and error correction in stock index and stock index futures markets: A co-integration approach*. The Journal of Futures Markets, vol. 13, no. 7, pp. 711-742, 1993.

APPENDIX A

MATLAB CODES OF NONLINEAR CAUSALITY

```
function [c1 c2 c3 c4]= lx8(x,y)
c1=0; c2=0; c3=0; c4=0;
for i=9:(size(x)+1)
   for j=9:(size(x)+1)
      if j>i
         q=[abs(x(j-1)-x(i-1));abs(x(j-2)-x(i-2));abs(x(j-3)-x(i-3));abs(x(j-4)-x(i-
4));abs(x(j-5)-x(i-5));abs(x(j-6)-x(i-6));abs(x(j-7)-x(i-7));abs(x(j-8)-x(i-8))];
         maxx1=max(q);
         w=[abs(y(j-1)-y(i-1));abs(y(j-2)-y(i-2));abs(y(j-3)-y(i-3));abs(y(j-4)-y(i-
4));abs(y(j-5)-y(i-5));abs(y(j-6)-y(i-6));abs(y(j-7)-y(i-7));abs(y(j-8)-y(i-8))];
         maxy1=max(w);
         if maxx1<1.5 && maxy1<1.5
            c2=c2+1;
         else
         end
            if maxx1<1.5
               c4=c4+1;
            else
            end
      else
      end
   end
end
for p=9:size(x)
   for r=9:size(x)
      if r>p
         e=[abs(x(r-8)-x(p-8));abs(x(r-7)-x(p-7));abs(x(r-6)-x(p-6));abs(x(r-5)-x(p-
5));abs(x(r-4)-x(p-4));abs(x(r-3)-x(p-3));abs(x(r-2)-x(p-2));abs(x(r-1)-x(p-
1));abs(x(r)-x(p))];
         maxx2=max(e);
         d=[abs(y(r-8)-y(p-8));abs(y(r-7)-y(p-7));abs(y(r-6)-y(p-6));abs(y(r-5)-y(p-
5));abs(y(r-4)-y(p-4));abs(y(r-3)-y(p-3));abs(y(r-2)-y(p-2));abs(y(r-1)-y(p-1))];
         maxy2=max(d);
         if maxx2<1.5 && maxy2<1.5
            c1=c1+1;
         else
         end
            if maxx2<1.5
               c3=c3+1;
```

```
                else
                end
            else
            end
        end
    end
```

APPENDIX B

LITERATURE REVIEW TABLE

AUTHOR,YEAR	DATA DESCRIPTION	METHODOLOGY	RESULT
MANASTER & RENDLEMAN, 1982	Daily closing prices April, 73 - June, 76 USA	Ex post and Ex ante tests	Opt. → C
FINNERTY & PARK, 1987	Intraday data Aug. 23, 1984 - Aug. 15, 1986 USA	Multiple regression method	F → C
KAWALLER, KOCH & KOCH, 1987	One min. intervals 1984 - 1985 USA	3 stage LS regression Granger - Sim's causality	F → C up to 45 min. C → F about 1 min.
BHATTACHARYA, 1987	15 min. intervals June 2, 1977 - Aug. 15, 1978 USA	Granger - Sim's causality	C → Opt.
HERBST, MCCORMACK & WEST, 1987	Daily data Feb. 24, 1982 - Sept. 18, 1982 USA	Spectral analysis	F→ C up to 16 min.
ANTHONY, 1988	Daily data Jan. 1, 1982 - June 30, 1983 USA	Granger - Sim's causality	Call opt. → C up to 1 day
HARRİS, 1989	5 min. intervals Oct. 12, 1987 - Oct. 23, 1987 USA	Weighted LS regression	F → C
STEPHAN & WHALEY, 1990	5 min. intervals Jan. 2, 1986 - March 31, 1986 USA	Multivariate T. S. regression	C → Opt. about 15 min.
STOLL & WHALEY, 1990	5 min. intervals 1982 - 1987 USA	Multiple Regression ARMA model	F → C by average of 5 min. C → F about 1 min.
CHAN, CHAN & KAROLYI, 1991	5 min. intervals Aug.1, 1984 - Dec. 31, 1989 Jul. 23, 1984 - June 30, 1985 USA	Bivariate GARCH model	F → C strong C → F weak

AUTHOR,YEAR	DATA DESCRIPTION	METHODOLOGY	RESULT
CHAN, 1992	5 min. intervals Aug. 1984 - June 1985 Jan. - Sept. 1987 USA	Regression analysis adjusted by Hansen's Covariance Matrix	F → C strong C → F weak
WAHAB & LASHGARI, 1993	Daily data 1988 - 1992 USA & UK	Cointegration & ECM	C → F strong F → C weak
ANTONIOU & GARRETT, 1993	One min. intervals Oct. 19 - 20, 1987 UK	Cointegration & ecm Granger causality	F → C strong C → F weak
FLEMİNG, OSTDIEK & WHALEY, 1996	One min. intervals Jan. 1988 - March 1991 USA	Multiple regression model	F → C about 5 min. Opt. → C about 5 min.
IIHARA, TAKUNAGA & KATO,1996	5 min. intervals March 1, 1989 - Feb. 26, 1991 Japan	Multivariate regression	F → C by 20 min.
DE JONG & NIJMAN, 1997	1 and 5 min. intervals Last quarter of 1993 USA	Regression and cross correlation	F → C by 10 min. C → F by at most 2 min.
DE JONG & DONDERS, 1998	5 and 10 min. intervals Jan. 20 - July 17, 1992 Jan. 4 - June 18, 1993 Netherlands	Regression and auto-cross correlations	F→ C by 10 min. C ↔ Opt.
ABHYANKAR, 1998	5 min. intervals During 1992 Stock and futures markets in UK	Linear and Nonlinear causality	Linear: F → C by 5 -15 min. Nonlinear: F ↔ C
NIETO & FERNANDEZ, 1998	Daily data March 1, 1994 - Sept. 30, 1996 Spain	Cointegration and Granger causality	F → C
PIZZI & ECONOMOPOULOS, 1998	One min. intervals Jan. - March, 1987 USA	Cointegration and ECM	F → C by at least 20 min. C → F by approximately 3-4 min.

AUTHOR,YEAR	DATA DESCRIPTION	METHODOLOGY	RESULT
BOOTH, SO & TSE, 1999	15 min. intervals Dec. 5, 1994 - July 11, 1997 Germany	Cointegration and ECM	F→ C, Opt. C → Opt.
MIN & NAJAND, 1999	10 min. intervals May 3 - Oct. 16, 1996 Korea	VAR analysis	F→ C about 30 min.
SILVAPULLE & MOOSA, 1999	Daily data Jan.2, 1985 - July 11, 1996 Australia	Linear and Nonlinear causality	No lead/lag relationship
ALPHONSE, 2000	5 min. intervals Jan. 3 - March 31, 1995 France	Cointegration and ECM	F → C strong C → F weak
GWILYM & BUCKLE, 2001	Hourly intervals Jan. 4, 1993 - Dec. 31, 1996 UK	Multiple regression	F, Opt. → C Call opt. → F, Put opt.
BROOKS, REW & RİTSON, 2001	10 min. intervals June, 1996 - June, 1997 UK	Cointegration and ECM, ARMA, VAR analysis	F → C by 30 min.
LAFUENTE, 2002	Hourly intervals Dec. 20, 1993 - Dec. 20, 1996 Spain	Cointegration and ECM	F → C
ASCHE & GUTTORMSEN, 2002	Monthly intervals April, 1981 - Sept. ,2001 International Petroleum Exchange	Cointegration and ECM	F → C
ROOPE & ZURBRUEGG, 2002	5 min. intervals Jan. 11 - June 31, 1999 Taiwan&Singapore	Cointegration and exogeniety tests	Singapore → Taiwan
RAJU & KARANDE, 2003	Daily data June, 2000 - Oct., 2002 India	Cointegration and ECM	F → C

AUTHOR,YEAR	DATA DESCRIPTION	METHODOLOGY	RESULT
CHUNG & CHUANG, 2003	Daily data July 21, 1998 - Aug. 20, 1999 Taiwan	Cointegration and ECM	$F \leftrightarrow C$
KENOURGIOUS, 2004	Daily data Aug., 1999 - June, 2002 Greece	Cointegration and ECM	$F \leftrightarrow C$
RYOO & SMITH, 2004	5 min. intervals Sept. 1, 1993 - Dec. 28, 1998 Korea	Cointegration and ECM	$F \rightarrow C$ strong $C \rightarrow F$ weak
MATTOS & GARCIA, 2004	Daily data 1997 - 2001 Brazil	Cointegration and ECM VAR analysis	No lead/lag relationship
HASAN, 2005	Daily data June 24, 1992 - Aug. 11, 1999 USA & UK	VAR analysis	$F \leftrightarrow C$ in both USA and UK
KANG, LEE & LEE, 2006	5 min. intervals Oct. 1, 2001 - Dec. 30, 2002 Korea	Multiple regression	F, Opt. $\rightarrow C$ up to 10 min.
FLOROS & VOUGAS, 2007	Daily data 1999 - 2001 Greece	Bivariate GARCH	$F \rightarrow C$
MILUNOVICH & JOYEUX, 2007	Daily data June 25, 2005 - Nov. 27, 2006 Australia	Cointegration and ECM, Granger causality	$F \rightarrow C$ strong $C \rightarrow F$ weak
BHATIA, 2007	Intraday data April, 2005 - March, 2006 India	Cointegration and ECM	$F \rightarrow C$ about 10 - 25 min.
KAVUSSANOS & VISVIKIS, 2008	Daily data Feb., 2000 - June, 2003 July, 2000 - June, 2003 Greece	Cointegration and ECM	$F \rightarrow C$ strong $C \rightarrow F$ weak

AUTHOR,YEAR	DATA DESCRIPTION	METHODOLOGY	RESULT
PRADHAN & BHAT, 2009	Daily data June 12, 2000 - Nov. 28, 2007 India	Cointegration and ECM	C $\rightarrow$ F
TSE & CHAN, 2009	3 min. intervals March 5 - July 1, 2004 USA	Treshold regression	F $\rightarrow$ C
SAVOR, 2009	One min. intervals Sept., 2007 - Feb., 2008 USA & Canada	Cointegration, Hasbrouck's approach	USA opt. $\rightarrow$ Canadian opt. Canadian C $\rightarrow$ Canadian opt.