# Lead-lag relationship between spot and futures stock indexes: intraday data and regime-switching models

Alemany, Nuria*; Aragó, Vicent** and Salvador, Enrique***

* Correspondence author, Associate Lecturer, Department of Accounting and Finance, Universitat Jaume I, Avda. Sos Baynat s/n, Castellón de La Plana, Spain. Tel: +34-964-387151, Fax: +34-964-728565, e-mail: nalemany@uji.es

** Associate Professor, Department of Accounting and Finance, Universitat Jaume I, Avda. Sos Baynat s/n, Castellón de La Plana, Spain. Tel: +34-964-387134, Fax: +34-964-728565, e-mail: arago@cofin.uji.es

*** Lecturer in Finance, Department of Accounting and Finance, Universitat Jaume I, Avda. Sos Baynat s/n, Castellón de La Plana, Spain. Tel: +34-964-387146, Fax: +34-964-728565, e-mail: esalvado@uji.es

**Acknowledgements**

The authors would like to thank Andreas Kaeck, Maria-Dolores Furió, Davide Avino and Thomas Conlon for their useful comments. We also thank the members of the Finance Department at University of Valencia and Cátedra Santander, and the participants at the Young Finance Scholars Conference 2016, World Finance Conference 2017, IINFINITI conference 2017 and the International Conference of the French Finance Association 2017 for helpful advice and suggestions. The authors acknowledge the financial support received from Universitat Jaume I of Castellón under the Research Personal Program PREDOC/2014/14 and the project UJI-B2017-14 and the Spanish Ministry of Economy and Enterprise under project ECO2014/55221-P and ECO2017-85746-P.

June, 2019

# Lead-lag relationship between spot and futures stock indexes: intraday data and regime-switching models

**Abstract**

This paper analyzes the impact of arbitrage opportunity changes on the price discovery process between the DAX30 index and the DAX30 index future within a short time scale. To this end, we use five-minute data, regime-switching models and the regime-dependent impulse response function. The results unveil the presence of nonlinearities in the cointegrating vector and the shortcomings of relying on linear assumptions. We also find that the presence of arbitrage opportunities alters the nature of the lead-lag dynamics: the more arbitrage opportunities, the greater the leading role of the futures market and the more pronounced the impact of unexpected shocks on prices.

**Keywords:** regime-switching models, arbitrage opportunities, lead-lag relationship, intraday data

**JEL Classification:** G12, G13, G14, G15

# 1 Introduction

According to the efficient market theory, asset prices reflect all relevant information available about their intrinsic value. This means that innovations are synchronously reflected both in spot and futures prices, and there is no lead-lag relationship between prices in these markets. However, due to market imperfections, such as asymmetric information, transaction costs, liquidity and other market restrictions, one market may reflect information faster than the other one, and as a result of that, a lead-lag relationship exists (Chan, 1992). Understanding this information flow across markets improves investment strategies, economic policy, asset valuation and hedging performance. It also provides investors with more efficient trading strategies (Kawaller *et al.*, 1987).

A large body of research, discussed in more detail in the next section, has investigated the relationship between spot and futures prices in the context of vector autoregressions using equilibrium correction models on a linear framework. Another strand of literature has reported that the dynamic relationship between spot and futures prices is characterized by nonlinear specifications, and neglecting these nonlinearities can lead to biased results (see among others, Hsieh, 1991; Perez-Quiros and Timmermann, 2001; Sarno and Valente, 2005; Balcilar *et al.*, 2015)[1]. Further, a number of researchers suggest that the lead-lag relationship is an important stylized fact at high-frequency data (HFD henceforth) and it vanishes when the frequency of observations decreases[2] (Harris *et al.*, 1995). Indeed, the use of intraday data can disclose new facts that cannot be detected at lower frequencies (Huth and Abergel, 2014). Although most of this literature states the leading effect of futures prices, results are ambiguous. Different conclusions are drawn from studies in different markets and time periods, by using different methodologies (Lien and Yang, 2003; Lien *et al.*, 2003; Gong *et al.*, 2016) or by considering asymmetries in arbitrage bounds (Dwyer *et al.*, 1996). To date, empirical evidence regarding the lead-lag relationship between spot and futures prices that considers both HFD and the presence of different regimes is scarce. To our knowledge, the only study that combines intraday data and regimes is Theissen (2012), who employs a threshold model to account for asymmetries in arbitrage.

Our paper departs from previous studies and makes several novel contributions to further understand the price discovery process[3] between the spot and futures markets. In particular, we focus on the effects of arbitrage opportunities on the lead-lag dynamics and extend the data and the methodology used in this field. First, we investigate the lead-lag phenomenon within a small time lag[4], using transaction prices on a five-minute interval basis from the $2^{nd}$ January 2014 to the $30^{th}$ September 2015 for the DAX30 index and the DAX30 index futures[5]. In contrast, the study of Theissen (2012) is based on observations of 15 seconds

[1] The focus of this study is on stock indexes, however, the importance of accounting for nonlinearities is not exclusive of stock indexes. Studies in other markets such as commodities also reveal that causal links between spot and futures prices are regime-varying (Balcilar *et al.*, 2015).

[2] Harris *et al.* (1995) show that frequency is crucial to test price dynamics between markets that are cointegrated.

[3] The price discovery process can be defined as the lead-lag relationship and information flows between two markets.

[4] Data at lower frequencies, such as daily, weekly or monthly, cannot detect an error correction process from high frequency trading strategies (Harris *et al.*, 1995).

[5] As Andersen (2000) highlights, *"the five-minute frequency is about the highest at which properties of the return series are not seriously distorted by irregular quoting, the discreteness of prices, and the tendency of*

and 1 minute, using DAX contracts, and covering 61 trading days in total. Since we are interested in studying short run and long run interactions, and testing for cointegration requires long datasets, our sample comprises 438 trading days in total.

Second, we employ regime-switching models to adequately characterize and capture unusual movements that appear in the relationship between spot and futures markets[6]. Unlike threshold models, regime-switching models have the advantage of identifying different regimes in the dynamic relationship without postulating exogenous structural changes. In particular, we estimate a state-dependent error correction model to determine whether asymmetric adjustments arise between spot and futures prices depending on the magnitude of the deviation from the long run equilibrium, and the market (spot or future) with more predictive capability. Since arbitrage has been described as an attempt to benefit from the long run trading opportunities involved in the cointegration relationship[7], we link the magnitude of the long-run disequilibrium, given by the error correction term (ECT henceforth), to arbitrage opportunities changes (Bondarenko, 2003; Hogan *et al.*, 2004)[8]. Further, to see to what extent results differ compared with a linear context, we also estimate the traditional vector error correction model (VECM henceforth). Results show the presence of two regimes in the cointegrating vector, identified as states of lower and higher arbitrage signals, in which the contribution to price discovery differs considerably. Concretely, in line with previous literature (Dwyer *et al.*, 1996; Theissen, 2012), results corroborate faster adjustments in states with higher arbitrage opportunities, and the leading role of the futures market in the price discovery process, especially when higher arbitrage opportunities emerge. Further, significant differences emanate from linear and nonlinear models: linear models find unilateral causality from the futures to the spot market and fail to detect cointegration, whereas nonlinear models reveal bilateral interaction with the leading role of the futures market and show that spot and futures prices may diverge temporarily but then readjust to the cointegrated pattern.

Third, to achieve a deeper comprehension of differences between states, we implement the regime-dependent impulse response function to investigate how fundamental disturbances affect the variables in the model depending on the regime (Gallant*et al.*, 1993; Koop *et al.*, 1996; Ehrmann *et al.*, 2003). To our knowledge, this feature has not been studied in this context before, but it yields very compelling results. We perform the analysis in the regimes previously identified (lower and higher ECT), and examine how markets react to shocks in each regime. This distinction enables to detect noticeable asymmetries across regimes after shocks in the system, so that the more arbitrage opportunities arise, the more pronounced the impact of unexpected shocks on prices. Particularly, results reveal that in states with higher arbitrage opportunities the impact of a shock is more than twice the impact of a

*foreign-exchange dealers to position their quotes with a view toward inventory control".*

[6]Previous literature finds inconsistent results regarding lead-lag effects among markets when nonlinear features are ignored (Sarno and Valente,2000; Li, 2009).

[7]The error correction term can be interpreted as an opportunity of arbitrage. When the future price is higher than the spot price, arbitragers will buy in the spot market and sell in the futures market. If the future price is lower, they will do the reverse. This trading forces prices back towards equilibrium so that at the time of futures contract expiration, the error correction term becomes zero.

[8]States with lower and higher ECT are identity with states with less and more arbitrage opportunities respectively. When spot and futures prices deviate from equilibrium levels, arbitrageurs have incentives to trade in order to yield profits. Our objective is to detect deviations in arbitrage bounds and see their impact in the price discovery process. What are the causes of this disequilibrium warrants future research (Aragó *et al.*, 2003).

shock in states with lower arbitrage opportunities.

Overall, our study reexamines the issue of price discovery process in spot and futures markets in a short time scale and unveils the shortcomings of neglecting that the presence of different arbitrage levels affects the dynamic of the adjustment between spot and futures prices. These findings provide new perspectives to understand how this lead-lag effect evolves when considering nonlinearities, which is valuable for its wide application in hedging and portfolio investments.

The remainder of this paper is organized as follows: section 2 contains a review of the literature; section 3 describes the data used while section 4 explains the methodology employed; section 5 presents the empirical results from the traditional VECM and the Markov-switching error correction model (MS-VECM henceforth) complemented by section 6 that addresses the regime-dependent impulse response function; finally, section 7 summarizes the results and concludes.

## 2 Review of literature

Literature related to lead-lag effects can be organized in several ways. In this paper, we first comment earlier literature on a linear context that fails to account for cointegration and uses either data at lower frequencies (Ng, 1987) or high-frequency data (Kawaller 1987; Stoll and Whaley, 1990; Frino *et al.*, 2000; Chan *et al.*, 1991; Chan, 1992). Then, we review literature that introduces cointegration in the analysis and uses daily data (Wahab and Lashgari, 1993) and data at higher frequencies (Tse, 1999; Pardo and Climent, 2000 and Blanco, 2003; Fassas and Siriopoulos, 2019). Finally, we offer an outline of relevant literature introducing nonlinearities, also for observations at lower (Sarno and Valente, 2000; Lien *et al.*, 2003; Li, 2009) and higher frequencies (Theissen, 2012).

Among studies on a linear framework that neglect cointegration and use daily data we find Ng (1987) who analyzes the price discovery process using the S&P 500 Index and the S&P 500 Futures Index on a daily basis, for approximately 5 years. This research concludes that futures prices lead rather than lag spot prices by one day, although the lead coefficients were weak in magnitude. Kawaller *et al.* (1987) examine the intraday price relationship between the S&P 500 Index and S&P 500 Futures Index using minute-to-minute data for all trading days during 1984 and 1985 and also conclude that the equity futures market leads the stock market, but effects in the opposite direction also exist. Stoll and Whaley (1990) examine the time series properties of five-minute returns for approximately 5 years of stock index and index futures (S&P500 and Major Market Index) and conclude that lagged stock index returns have a moderate predictive impact on futures returns. They argue that the effect is bidirectional but the futures market has more predictive capability. Frino *et al.* (2000) analyze the lead-lag relationship between the Share Price Index (SPI) futures contract and the All Ordinaries Index in presence of stock-specific information releases from August 1995 to December 1996. The study concludes that the lead-lag relationship is influenced by the release of macroeconomic and stock-specific information, and it is necessary to control for the effects of information releases in order to provide valid comparisons regarding the price discovery process. Chan *et al.* (1991) study the interdependence in price change and price-change volatility between returns in the S&P 500 stock index and its associated index futures from 1984 to 1989 using a five-minute frequency. They find much stronger bidirectional dependence when the volatility is also considered. Their evidence is consistent with the hypothesis that both markets contribute to price discovery. Chan (1992) studies the lead-lag relationship between returns in the Major Market cash index and returns of the Major Market Index futures and S&P 500 futures using a frequency of five minutes for two sample periods: August 1984–June 1985 and January 1987–September 1987. The author finds strong evidence of an asymmetric lead-lag relationship between the two markets with strong evidence that the futures index leads the cash index.

The literature previously presented fails to account for adjustments to the long run equilibrium relationship. Since spot and futures prices are generally cointegrated, an error correction term should be included in the model to correct in one period the disequilibrium detected in the previous one (Engle and Granger, 1987)[9]. The rationale behind the concept of cointegration is that two variables may deviate in the short run from each other, but market forces will bring them back together, so that a long run equilibrium relationship between these two variables exists. Neglecting this variable could lead to misspecified models. Wahab and Lashgari (1993) extend the study of lead-lag effects by applying the cointegration approach to investigate the robustness of previous studies including an alternative model parameterization: the error correction model. Their research uses daily spot and futures prices for both the S&P 500 Index and the Financial Times Index from January, 1988 to May, 1992, and shows that a two-way relationship exists between the cash and futures markets. The findings, consistent with Chan *et al.* (1991), show the important price discovery role served by both the stock and index futures markets and corroborate much stronger interdependence between stock index cash and futures market than found in earlier studies. Tse (1999) examines the intraday price discovery process and volatility spillovers between the DJIA index and index futures using minute-by-minute data for the six-month period of November 1997 to April 1998. The author uses the VECM to analyze the price discovery process and concludes that it mostly takes place at the futures market. Pardo and Climent (2000) and Blanco (2003) study the temporal relationship between the IBEX 35 index and IBEX 35 futures contracts applying a cointegration parameterization and using minute-by-minute data for the entire year 1996 and five-minute data from January, 1995 to October, 1995, respectively. These studies conclude that both markets contribute to price discovery, with greater predictive capability of the futures market. Fassas and Siriopoulus, 2019 analyse the lead-lag relationship between the cash and futures index prices in Athens Exchange using the error correction model and 30-minute observations from January, 2013 to December, 2014. The results provide evidence in support of the leading role of the futures market.

Motivated by an interesting strand of literature that suggests nonlinear dynamics in the lead-lag relationship between spot and futures prices[10], Sarno and Valente (2000) examine lead-lag effects between spot and futures prices in stock index futures markets using weekly data for the S&P500 and the FTSE100 indices from January 1988 to December 1997. To

[9]One interpretation of the error correction term is that it reflects the effect of arbitrage. If the futures price is too low compared to the index value, arbitragers will sell the stocks underlying the index and buy the futures contract. On the contrary, if the futures price is too high, they will sell the futures contract and buy the stocks underlying the index.

[10]See among others Dwyer *et al.*(1996); Gao and Wang (1999)

this end, they use nonlinear Markov-switching vector equilibrium correction models that allow for three regimes in the mean of the equilibrium correction model as well as in the variance-covariance matrix. Their research shows strong evidence against the hypothesis of linear dynamics and in favor of the better performance of regime-switching vector equilibrium models to capture the stylized behaviour of the financial series. Lien *et al.* (2003) implement a nonparametric genetic programming (GP) method to identify the structural changes in the Nikkei spot index and futures price using daily data from September 1995 to December 1999. The GP approach identifies "normal" regimes versus "extreme" regimes and finds that market behaviour depends on different regimes, with major market changes taking place first in the spot market followed by the futures market. For comparative purposes, these authors also consider a linear and nonlinear test for Granger causality. Contrariwise, the linear causality test suggests the leading role of the futures market, whereas the Granger nonlinear causality test finds no evidence of nonlinear lead-lead relationships. The study therefore concludes that conventional methods fail to capture relationships that arise from structural changes. Li (2009) studies the dynamics of the relationship between spot and futures markets of three mature markets (S&P500, FTSE100, DAX30) and two emerging markets (BOVESPA, BSI) from the period between April 1995 to December 2005 using daily data. The author uses a traditional VECM and a MS-VECM, in which the parameter of the deviation of spot-futures prices changes according to the state of the volatility regime. Results reveal that ignoring nonlinear features leads to inconsistent results among markets. Theissen (2012) examines the intraday price discovery process using DAX contracts with a frequency of observations of 15 seconds and 1 minute, covering 61 trading days in total. The author estimates a threshold error correction model that allow arbitrage opportunities to have an impact on the return dynamics. The findings reveal that arbitrage opportunities have a strong effect on the dynamics of the price discovery process, so that the more arbitrage signals arise, the greater the leadership of the futures market in the price discovery process.

# 3 Data and preliminary analysis

This study uses transaction prices on a five-minute interval basis from DAX30 for both the stock index and its corresponding futures contract. The sample period extends from the $2^{nd}$ January 2014 to the $30^{th}$ September 2015, and only data for the period of simultaneous operation of both markets are used.

We obtain the continuously compounded returns at every five-minute interval by taking the logarithm and subtracting the previous value, so that, returns at the $n$ interval at day $t$, for $n$=*1,2,...,N* and $t$=*1,2,...,T* are calculated as follows:

$$R_{n,t} = 100 \times log\Big(\frac{P_{n,t}}{P_{n-1,t}}\Big)$$

where $P_{n,t}$ represents the spot ($S_{n,t}$) and futures ($F_{n,t}$) price level on interval $n$, at day $t$.

Additionally, it is frequent in the literature to exclude some observations at the beginning of each trading day. When the negotiation in the spot markets begins, volatility generally reflects the adjustment to information accumulated overnight and displays the highest level.

Thus, to avoid deleterious effects on the econometric analysis we remove the first return of the trading day, that is, the 09:05 hour (see Andersen *et al.*, 2000 and Lee and Mathur, 1999).[11]

Table 1 presents some statistical tests for prices and returns series. As can be appreciated in panel A, returns are clearly not normally distributed due to asymmetric and leptokurtic patterns, and mean returns are close to zero. Panel B shows the Ljung-Box test statistic, which detects autocorrelation for both prices and returns. Panels C and D exhibit stationary and cointegration tests. Spot and futures prices are nonstationary, whereas returns series are stationary. This reveals that spot and future prices contain a unit root and are integrated of order one. To check the presence of unit root in the residuals $Z_t$ from the regression $S_t = \beta + \gamma F_t + Z_t$ (ECT regression), we use the Augmented Dickey Fuller (ADF) and the Kwiatkowski-Phillips-Schmidt-Shin (KPPS) test. According to the ADF test, the residuals from the ECT regression are stationary, that is, the ECT is $I(0)$ and spot and futures markets are cointegrated. In contrast, the KPPS test reveals that the ECT is not stationary, so that these markets are not cointegrated. Thus, the cointegration analysis shows contradictory results.

The presence of regimes in the ECT might explain this inconsistency. Indeed, as can be appreciated in Figure 1, two regimes are clearly visible in the ECT (lower and higher ECT). Since the power of unit root tests is sensitive to structural breaks in time series, neglecting them may lead to distorted values in the stationarity and cointegration tests (Bartley *et al.*, 2001). Perron (1989, 1990) shows that ignoring these structural changes in the data series leads to biased results in the ADF test. Further, Lee *et al.* (1997) highlight that the KPSS stationarity test proposed by Kwiatkowski *et al.* (1992) is biased towards rejecting the null hypothesis of stationarity repeatedly when the data generating process is stationary with a structural break. Given that linear models fail to account for this state-dependent feature, they might lead to inconsistent results regarding the lead-lag relationship. In contrast, nonlinear models enable to better capture the dynamic relationship between spot and futures prices. Thus, to account for the ECT dynamics, we use a state-dependent error correction model, while comparing it with the traditional VECM.

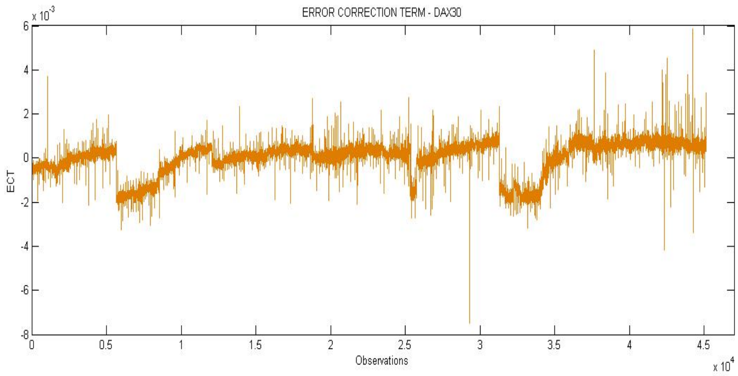


**Figure 1: Error Correction Term.** Figure 1 shows the error correction term for the DAX30 from the $2^{nd}$ January 2014 to the $30^{th}$ September 2015. The horizontal axis represents the observations on a five-minute interval basis, and the vertical axis gives the magnitude of the ECT

[11]This leaves us with a sample of 438 days with 103 intraday five-minute returns per day.

Table 1: **Summary statistics for the DAX30 index prices and returns**
Table 1 shows some statistical tests for DAX30 index prices and returns on a five-minute interval basis for both the spot and futures market from the $2^{nd}$ January 2014 to the $30^{th}$ September 2015. Returns at the $n$ interval at day $t$ have been calculated as follows: $R_{n,t} = 100 \times log(P_{n,t}/P_{n-1,t})$, where $P_{n,t}$ represents the price level on interval $n$, at day $t$. Panel A presents the main summary statistics and the Jarque-Bera test for normality. Panel B displays the results of the serial autocorrelation test Ljung-Box using 20 lags. Panel C performs the Augmented-Dickey Fuller and Kwiatkowski-Phillips-Schmidt-Shin stationary tests, and panel D shows the results of the cointegration tests (checking whether the error correction term is stationary).The * denotes significance at 5%.

| | Levels | | Returns | |
|---|---|---|---|---|
| | Spot | Futures | Spot | Futures |
| Panel A: Summary Statistics | | | | |
| Mean | 10,195.49 | 10,201.05 | -0.00034 | -0.00031 |
| Std. deviation | 927.25 | 930.61 | 0.099 | 0.100 |
| Skewness | 0.66 | 0.67 | -0.389 | -0.367 |
| Kurtosis | 2.12 | 2.15 | 16.78 | -17.02 |
| Minimum | 8,363.08 | 8,367.00 | -2.100 | -2.205 |
| Maximum | 12,385 | 12,422 | 1.245 | 1.292 |
| Jarque-Bera | 4,741* | 4,746* | 358,054* | 370,532* |
| Observations | 45,115 | 45,115 | 45,114 | 45,114 |
| Panel B: Autocorrelation test | | | | |
| $LB-Q(20)$ | 900,772.62* | 900,775.64* | 98.41 | 114.15 |
| $LB-Q^2(20)$ | 896,284.85* | 896,409.88* | 15,126.42* | 14,631.81* |
| Panel C: Stationarity test | | | | |
| ADF test | -0.0365 | -0.0494 | -215.42* | -216.99* |
| KPSS test | 347.51* | 348.13* | 0.1434 | 0.1336 |
| Panel D: Cointegration test | | | | |
| ADF test | | -38.537* | | |
| KPSS test | | 103.654* | | |

# 4 Methodology

This section explains the empirical models used. First, we describe the linear model, and then, we consider a nonlinear dynamic in the ECT[12].

## 4.1 Traditional VECM

The first approach to analyze the price discovery process is the traditional VECM, which assumes a permanent causal relationship between the spot and futures prices over the sample

[12]If two variables are cointegrated, they can be represented by a VECM that incorporates the last period error term as well as lagged returns of each variable. Thus, temporal causality can be evaluated by analysing the statistical significance and relative magnitude of lagged variables coefficients and ECT coefficients.

period.

$$\Delta S_t = a_s + \sum_{i=1}^{k} b_{si}\Delta S_{t-i} + \sum_{i=1}^{k} c_{si}\Delta F_{t-i} + \alpha_s Z_{t-1} + \varepsilon_{s,t} \tag{1}$$

$$\Delta F_t = a_f + \sum_{i=1}^{k} b_{fi}\Delta S_{t-i} + \sum_{i=1}^{k} c_{fi}\Delta F_{t-i} + \alpha_f Z_{t-1} + \varepsilon_{f,t} \tag{2}$$

In the study of the price discovery process, both the short run and the long run causality are analyzed. The short run relationship is captured by the coefficients $b_{si}$, $c_{si}$, $b_{fi}$ and $c_{fi}$. When these coefficients are significant, it implies that a lead-lag relationship exists, and therefore, lagged returns in one market can be used to predict futures returns in the other market. Additionally, the long run relationship is represented by the error correction term $Z_{t-1} = S_{t-1} - \beta - \gamma F_{t-1}$, where $S_{t-i}$ and $F_{t-i}$ are log lagged spot and futures prices, respectively.

The error correction coefficients $\alpha_s$, $\alpha_f$ collect information regarding the direction of the causal relationship between two series and show the speed at which the departures from equilibrium are corrected in the short run. Cointegrated variables may deviate from their relationship in the short run, but in the long run, their association will return. If a departure from equilibrium arises, prices in one or both markets should adjust to correct the deviation; otherwise, the series would wander apart without bound.

Additionally, $a_s$ and $a_f$ represent the unconditional return, and $\varepsilon_{s,t}$ and $\varepsilon_{f,t}$ are the residuals in the spot and futures equations, respectively. The optimal lag length will be determined using the Akaike information criteria (AIC) and the Bayesian information criteria (BIC).

## 4.2 Markov-Switching VECM

Recent years have witnessed a remarkable increase in the popularity of nonlinear modelling (see Sarno and Valente, 2000; Lien *et al.*, 2003; Ang and Timmermann 2012, among others). Most of the previous studies use linear specifications and neglect nonlinearities in their empirical models. To account for the dynamics of the ECT, we use a MS-VECM in which the parameter of the long run deviation of spot-futures prices is dependent on two regimes[13]. The nonlinear dynamic approach presented here implies that the degree and speed of adjustment towards the long run equilibrium depends on the size of the deviation. Moreover, the Markov-switching methodology offers the advantage of endogenously determining the changes in the dynamic relationship without postulating exogenous structural changes. Thus, instead of conjecturing a known regime in a certain period, its probability in each point of time is estimated based on the information extracted from the sample[14].

[13]To keep the number of parameters tractable, this investigation considers two regimes in the ECT.

[14]Our approach for using the MS-VECM has partly been due to the recent success of this methodology in describing the time series properties of the stock market data. But it is also worth mentioning that the MS-VECM is a parametric approach that can be complemented with a nonparametric methodology, that is, the genetic programming (GP) approach. The latter is also an effective methodology to capture structural

The methodology is basically the result of extending the Markovian regime shifts to time series analysis, considering changes in causality as random events governed by an exogenous Markov process (Hamilton, 1988,1989). Our MS-VECM, specified as follows, parameterizes that the ECT comes from a particular causality regime with a certain probability

$$\Delta S_t = a_s + \sum_{i=1}^{k} b_{si}\Delta S_{t-i} + \sum_{i=1}^{k} c_{si}\Delta F_{t-i} + \alpha_{s,s_t} Z_{t-1} + \varepsilon_{s,t,s_t} \tag{3}$$

$$\Delta F_t = a_f + \sum_{i=1}^{k} b_{fi}\Delta S_{t-i} + \sum_{i=1}^{k} c_{fi}\Delta F_{t-i} + \alpha_{f,s_t} Z_{t-1} + \varepsilon_{f,t,t,s_t} \tag{4}$$

$$\begin{pmatrix} \varepsilon_{s,t,s_t} \\ \varepsilon_{f,t,s_t} \end{pmatrix} |\psi_{t-1}\ BN(0, H_{t,s_t}) \tag{5}$$

$$H_{t,s_t} = \begin{bmatrix} h_{s,s,s_t} & h_{s,f,s_t} \\ h_{f,s,s_t} & h_{f,f,s_t} \end{bmatrix} \tag{6}$$

where $\psi_{t-1}$ refers to the information available at time $t-1$, $H_{t,st}$ is the regime-dependent variance-covariance matrix, and $s_t$ is an unobservable state variable that governs the regime shifts. The parameters $\alpha_{s,s_t}$, $\alpha_{f,s_t}$ accompanying the ECT also depend on the state[15] of the process $s_t = 1, 2$.

The MS-VECM is estimated using a two-step maximum likelihood procedure[16]. The error correction term $Z_{t-1}$ is determined in the first step, and is the same variable computed in the traditional VECM. The second step consists of the implementation of an expectation-maximization algorithm using maximum likelihood to estimate equations (3) to (6)[17].

## 5 Empirical findings

We analyze in this section how the presence of different regimes in the ECT affects the dynamics of the adjustment towards equilibrium. Section 5.1 examines the results for the traditional VECM, whereas section 5.2 discusses the main findings for the MS-VECM.

changes of prices and has been introduced in this area by Lien *et al.*, 2003. Implementing the GP remains in our agenda for future research.

[15]These states will be identified as State 1=Low ECT and State 2=High ECT.

[16]See Perlin, M. (2010) MS Regress - The MATLAB Package for Markov regime-switching Models. Available at SSRN: http://ssrn.com/abstract=1714016

[17]For more details about the algorithm used for drawing probabilistic inference about whether and when shifts might occur, see Hamilton (1989)

### 5.1 Estimates for linear models (VECM)

First, we use the AIC/BIC criteria to determine the appropriate lag length of the VECM.[18] Then, we estimate the model, whose parameters are exhibited in Table 2.

Table 2: **Parameter estimates of the linear VECM**

Table 2 presents the estimates of the linear VECM:

$\Delta S_t = a_s + \sum_{i=1}^{k} b_{si}\Delta S_{t-i} + \sum_{i=1}^{k} c_{si}\Delta F_{t-i} + \alpha_s Z_{t-1} + \varepsilon_{s,t}$ Spot return equation

$\Delta F_t = a_f + \sum_{i=1}^{k} b_{fi}\Delta S_{t-i} + \sum_{i=1}^{k} c_{fi}\Delta F_{t-i} + \alpha_f Z_{t-1} + \varepsilon_{f,t}$ Futures return equation

where $S_{t-i}$ and $F_{t-i}$ are log lagged spot and futures prices, respectively; $\Delta$ is the first-difference lag operator, $a_s$ and $a_f$ represent the unconditional return; coefficients $b_{si}$, $c_{si}$, $b_{fi}$, $c_{fi}$ capture the short-run relationship; $Z_{t-1}$ , computed as $Z_{t-1} = S_{t-1} - \beta - \gamma F_{t-1}$, is the error correction term (ECT); $\alpha_s$, $\alpha_f$ are the error correction coefficients that collect information regarding the long-run relationship; and $\varepsilon_{s,t}$ and $\varepsilon_{f,t}$ are the residuals in the spot and futures returns equations, respectively. The optimal lag length has been determined using the Akaike information criteria (AIC) and the Bayesian information criteria (BIC) and has been set in five lags. The * denotes significance at 0.05

| | | SPOT RETURNS EQUATIONS | FUTURES RETURNS EQUATION |
|---|---|---|---|
| Intercept $\alpha_s$,$\alpha_f$ | | -0.00037 | -0.00033 |
| ECT $\alpha_s$,$\alpha_f$ | | -0.011786 | 0.00222 |
| Lagged Spot Return Coeff. | 1 | -0.551384* | 0.00559 |
| $b_{si}$,$b_{fi}$ | 2 | -0.300798* | 0.02479 |
| | 3 | -0.182988* | 0.00462 |
| | 4 | -0.13084* | -0.03498 |
| | 5 | -0.06758* | -0.02765 |
| Lagged Futures Return Coeff. | 1 | 0.54140* | -0.02712 |
| $c_{si}$,$c_{fi}$ | 2 | 0.28443* | -0.04650 |
| | 3 | 0.20317* | 0.01171 |
| | 4 | 0.14388* | 0.04631 |
| | 5 | 0.07046* | 0.02782 |

Table 2 shows that short run causality is unidirectional from the futures to the spot market. The lagged spot return coefficients ($b_{fi}$) are not significant in the futures equation, whereas lagged future return coefficients ($c_{si}$) are significant in the spot equation. Additionally, contemporaneous spot returns are affected negatively by lagged spot returns ($b_{si}$) and positively by lagged futures returns[19] ($c_{si}$). This pattern is not observed in the futures market. Therefore, short run adjustments underscore the importance of futures prices in the price discovery process.

Regarding the long run relationship, the parameters accompanying the ECT ($\alpha_s$, $\alpha_f$) are not significant. This means that both series are not cointegrated and deviate without a bound in the long run. Contradictory results in the unit root test ADF vs. KPPS are

[18] According to these criteria the optimal lag length is 5. The same lag length is used in the VECM-MS.

[19] The same evidence is found by Theissen (2012).

hereby confirmed. Since spot and futures prices concerning the same index react to the same information, short run deviations might be possible, but in the long run, spot and futures prices are expected to strike a balance. In this regard, neglecting structural changes clearly visible in the ECT (see Figure 1), may lead us to inefficient estimates. This outcome reinforces the idea that a regime-switching ECT should be considered.

## 5.2 Estimates for nonlinear models (MS-VECM)

We postulate the existence of two regimes in the ECT (high ECT and low ECT regimes) and estimate the MS-VECM according to the procedure defined in section 4. Table 3 presents the parameter estimates of the MS-VECM[20]. The spot equation presents similar results to the linear estimation: lagged spot returns ($b_{si}$) have a negative impact while the impact of lagged futures returns is positive ($c_{si}$). In the futures equation, results differ substantially: both lagged spot returns ($b_{fi}$) and lagged futures returns ($c_{fi}$) are significant. Thus, after considering regimes in the ECT, a two-way causality is detected, which means that price innovations in either the cash or futures markets predict the arrival of new information in the other market and both markets play important price discovery roles. However, there is an asymmetric lead-lag relationship between the two markets with strong evidence that the futures index leads the cash index and weak evidence in the opposite direction (see in Table 3 that the magnitude of $c_{si}$ coefficients accompanying lagged futures returns in the spot equation is much greater, in absolute terms, than the lagged spot returns parameters $b_{fi}$ in the futures equation; for instance, $c_{s1} = 0.70512$, whereas $b_{f1} = 0.00213$). Therefore, the effect is bidirectional in the short run, but the futures market has more predictive capability. As one might expect, these findings are consistent with previous empirical studies that reinforce the idea of the leading role of the futures market (Kawaller *et al.*, 1987; Ng, 1987; Stoll and Whaley, 1990; Chan *et al.*, 1991; and Chan, 1992, among others).

The main difference between the estimates in the traditional VECM and the MS-VECM is encountered in the analysis of the long run relationship. Contrary to results found in the linear estimation, parameters accompanying the ECT ($\alpha_{s,s_t}$ and $\alpha_{f,s_t}$) have become statistically significant, suggesting that spot and futures prices may diverge temporarily but then readjust to the cointegrated pattern, and any mispricing is driven back to the equilibrium by arbitrage forces. In other words, a long run equilibrium relationship between spot and futures markets exists. The leading role of the futures market is also corroborated in the long run in both ECT states. Moreover, note that the $\alpha$ parameters measure the speed of the disequilibrium correction, so that the greater the value of this coefficient (in absolute terms), the more informationally efficient the market. Hence, according to the estimates in Table 3, the spot price makes the greater adjustment to re-establish the equilibrium, and this speed of convergence to equilibrium is faster in states with higher ECT than in states with lower ECT (in the spot return equation the parameter $\alpha_{s,st=2}$ accompanying the ECT in the high state is equal to 0.020933, more than four times the magnitude of the parameter $\alpha_{s,st=1}$ accompanying the ECT in the low state with a value of 0.004472). That is, the futures market leads the cash market in price discovery, but the dynamic of the adjustment

[20]For the estimate of the VECM-MS, we use the specification-robust estimator of the variance-covariance matrix suggested by Bollerslev and Wooldridge (1992) to prevent the effect of heteroskedasticity and autocorrelation in the residuals.

Table 3: **Parameter estimates of the MS-VECM**

Table 3 presents the estimates of the MS-VECM:

$\Delta S_t = a_s + \sum_{i=1}^{k} b_{si}\Delta S_{t-i} + \sum_{i=1}^{k} c_{si}\Delta F_{t-i} + \alpha_{s,s_t} Z_{t-1} + \varepsilon_{s,s_t,t}$ Spot return equation

$\Delta F_t = a_f + \sum_{i=1}^{k} b_{fi}\Delta S_{t-i} + \sum_{i=1}^{k} c_{fi}\Delta F_{t-i} + \alpha_{f,s_t} Z_{t-1} + \varepsilon_{f,s_t,t}$ Futures return equation

$$\begin{pmatrix}\varepsilon_{s,t,s_t}\\ \varepsilon_{f,t,s_t}\end{pmatrix} |\psi_{t-1}\ BN(0, H_{t,s_t})$$

$$H_{t,s_t} = \begin{bmatrix} h_{s,s,s_t} & h_{s,f,s_t} \\ h_{f,s,s_t} & h_{f,f,s_t} \end{bmatrix}$$

where $S_{t-i}$ and $F_{t-i}$ are log lagged spot and futures prices, respectively; $\Delta$ is the first-difference lag operator, $a_s$ and $a_f$ represent the unconditional return; coefficients $b_{si}$, $c_{si}$, $b_{fi}$, $c_{fi}$ capture the short-run relationship; $Z_{t-1}$ , computed as $Z_{t-1} = S_{t-1} - \beta - \gamma F_{t-1}$, is the error correction term (ECT); $\alpha_{s,s_t}$, $\alpha_{f,s_t}$ are the parameters accompanying the ECT that depend on the regime $s_t$*=1 (Low State), 2 (High State)* and are defined as the error correction coefficients that collect information regarding the long-run relationship; $H_{t,s_t}$ is the variance-covariance matrix (s=spot market, f=futures market); and $\varepsilon_{s,s_t,t}$ and $\varepsilon_{f,s_t,t}$ are the residuals in the spot and futures returns equations, respectively. The same lag length set in the VECM is used in the MS-VECM (five lags). The * denotes significance at 0.05

| | | | |
|---|---|---|---|
| **Non-switching parameters** | | | |
| | | SPOT RETURNS EQUATIONS | FUTURES RETURNS EQUATION |
| Intercept $\alpha_s$,$\alpha_f$ | | -0.00004* | -0.00026* |
| Lagged Spot Return Coeff. | 1 | -0.73230* | 0.00213 |
| $b_{si}$,$b_{fi}$ | 2 | -0.53412* | 0.02000* |
| | 3 | -0.39150* | 0.00601* |
| | 4 | -0.26318* | -0.00686* |
| | 5 | -0.12683* | 0.00244* |
| Lagged Futures Return Coeff. | 1 | 0.70512* | -0.03853* |
| $c_{si}$,$c_{fi}$ | 2 | 0.48213* | -0.07588* |
| | 3 | 0.00500* | 0.35822* |
| | 4 | 0.29444* | 0.03645* |
| | 5 | 0.14066* | 0.00860* |
| **Switching parameters** | | | |
| State with higher ECT $\alpha_{s,s_t}$,$\alpha_{f,s_t}$ | | 0.020933* | -0.000317* |
| State with lower ECT $\alpha_{s,s_t}$,$\alpha_{f,s_t}$ | | 0.004472* | 0.000403* |
| **Variance - covariance matrix $H_{t,s_t}$** | | | |
| State with higher ECT | | 0.025739 | 0.024996 |
| | | 0.024996 | 0.026798 |
| State with lower ECT | | 0.005084 | 0.005085 |
| | | 0.005085 | 0.005219 |
| **State duration** | | | |
| State with higher ECT | | 4.90 | |
| State with lower ECT | | 25.52 | |

differs in presence of arbitrage opportunities. These results are, therefore, consistent with those found by Theissen (2012), who documents the leading role of the futures market, specially when arbitrage opportunities arise.

Regarding to the expected duration of each regime, it is approximately two hours during states with lower ECT, whereas it decreases to 25 minutes in states with higher ECT (see in Table 3 that the state duration is 25.52 and 4.9 intervals of 5 minutes for the high and low states respectively). This implies that regimes with less arbitrage opportunities are more persistent and it takes longer to reach the new equilibrium during these periods.

To sum up, our findings reveal that linear models might be misspecified if structural changes present in the sample are neglected. Thus, considering the magnitude of the deviation from the long run equilibrium is critical to analyze in depth the contribution of the spot and futures prices to the price discovery process.

## 6 Regime-dependent impulse response function

Impulse response functions[21] (IRF henceforth) are useful tools to explain the dynamic interaction throughout time between the variables and the disturbances in the vector autoregressive model (VAR henceforth). The analysis of the IRF has been extensively implemented in linear models. However, the study of nonlinear cases has been less covered. To capture price dynamics properly is crucial to expand the analysis of the IRF to address changes in regimes (Gallant, 1993). To this end, the seminal article by Koop *et al.* (1996) generalizes the IRF to the nonlinear framework. Later, Ehrmann *et al.* (2003) propose the regime-dependent impulse response function (RD-IRF henceforth) to investigate how fundamental disturbances affect the variables in the model dependent on the regime.

To gain more insight into the idiosyncrasy of each regime in the MS-VECM, we implement here the RD-IRF and determine the response of the system conditioning on the Markov-switching regime in which the shock occurs. We follow a two-stage procedure to obtain the RD-IRF. First, we convert the MS-VECM back to a VAR [22] model, and then, the resulting VAR model is used to perform the RD-IRF[23].

The RD-IRF is defined as follows:

$$\theta_{k,s_t,h} = \frac{\partial E_t P_{t+h}}{\partial u_{k,t}} | s_t \quad \text{for h} \geq 0 \tag{7}$$

where $u_{k,t}$ is the structural shock to the $k-th$ variable, $P_{t+h}$ is the spot $S_{t+h}$ or future $F_{t+h}$ prices at time $t+h$[24] and $\theta_{k,s_t,h}$ is a $k$-dimensional response vector dependent on the regime $s_t$.

Given that two regimes are present in the ECT, our general model considers four regime-dependent impulse response functions, which includes the IRF for spot and futures markets in low ECT and high ECT regimes.

Figure 2 illustrates the RD-IRF for an unexpected shock in states with lower ECT (top plot) and higher ECT (bottom plot). Asymmetries in terms of magnitude and persistence of the

[21] Also known as *"error shock"* methodology

[22] The VECM(5) is converted into a VAR(6); then, we can express the VAR(6) as an infinite moving average model.

[23] The Generalized impulse response function by Pesaran and Shin (1998) is used.

[24] The dynamic response for each variable is traced out over a period of 30 intervals of five minutes (two and a half hours).

responses can be seen by comparing them. A noteworthy result is that the shock implies a temporary and permanent effect on prices, but the impact of a shock of one standard deviation magnitude has a greater reaction on regimes with higher ECT. After the shock hits the system, prices increase approximately 0.07% in states with lower ECT (see the top plot), whereas in regimes with more arbitrage opportunities prices rise between 0.15% and 0.16% (see the bottom plot). Moreover, the effect of the shock stabilizes after 40-45 minutes (8-9 periods), inducing a permanent increase in prices on both regimes, which rise to reach their new long run equilibrium level. Further, when a shock is applied to the futures prices, the market responses are more pronounced than when a shock is introduced into the spot market. These findings are coherent with the results of the MS-VECM. As Figure 2 illustrates, both futures and spot prices react to unexpected shocks in the spot and futures markets, which suggests bidirectional interaction between the DAX30 index and the DAX30 index futures. However, the response to a shock in futures prices is relatively larger than the response to a shock in spot prices (see that solid lines are above the dashed lines), indicating that the futures market leads the spot market and reinforcing the idea of the leading role of the futures markets in the price discovery process.

Summing up, the implications of the results on the RD-IRF analysis are threefold. First, the response to a shock in futures prices is relatively larger than the response to a shock in spot prices. Second, unexpected shocks induce a permanent change[25] in prices by moving the new long run equilibrium level. Third, the dynamic causal effect is remarkably different in low/high ECT regimes, so as the arbitrage opportunities increase, the impact of unexpected shocks on prices increases. These findings support the importance of considering regimes to capture asymmetric responses in states with different arbitrage opportunities.

# 7 Conclusions

This paper reconsiders the lead-lag relationships between spot and futures markets in a short time scale. To this end, we use the DAX30 stock index and DAX30 index futures on a five-minute interval basis, and the non-linear model MS-VECM to overcome the weakness of linear assumptions in the dynamic relationship between spot and futures prices. This model allows us to endogenously identify the presence of different regimes in the deviation from the long run equilibrium and analyze how the dynamics of the price discovery process change depending on the regime. We also implement the RD-IRF to foster deep understanding of the lead-lag effects.

The empirical results provide evidence in favor of the existence of different regimes in the cointegration vector, confirming the presence of asymmetries regarding arbitrage opportunities. After accounting for these asymmetries, we find that the dynamic causal effect differs among linear and non-linear models, and among regimes. The VECM finds unilateral interaction between these markets, so that the futures market leads the spot market. Further, it shows inconsistent results regarding the long run relationship, suggesting that spot and futures markets are not cointegrated. However, in the short run, the MS-VECM detects two-way causality in the price discovery process, with the leading role of the futures market,

[25]Prices are not stationary. If prices contain a unit root, the shocks have permanent effects, while if they are stationary, then the effects of shocks eventually die out.

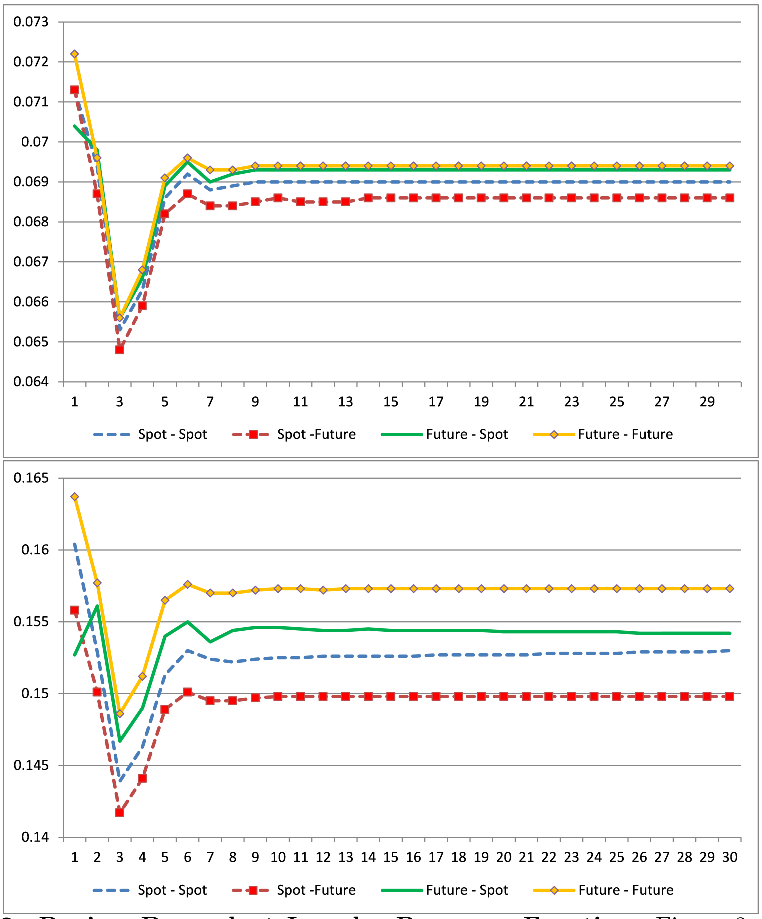


**Figure 2: Regime-Dependent Impulse Response Function.** Figure 2 gives the impulse responses to a one standard deviation shock in states with lower ECT (top plot) and higher ECT (bottom plot). The horizontal axis represents the period in intervals of five minutes. The vertical axis represents the magnitude of the shock expressed as a percentage of the price increase. The solid line and the solid line with markers denote the impulse response of the spot and futures markets to a shock in the futures market, respectively. In addition, the dashed line and the dashed line with markers represent the impulse response of the spot and futures markets to a shock in the spot market, respectively. All impulses are based on the Generalized Impulse Response Function by Pesaran and Shin (1998).

whereas in the long run, it reveals that spot and futures prices may diverge temporarily but both readjust to the cointegrated pattern, and it is the spot price that makes the greater adjustment to reestablish the equilibrium. Moreover, spot price adjustment accelerates in states with higher ECT, revealing faster convergence of the basis when more arbitrage opportunities arise. Finally, results from the impulse response analysis reinforce the idea that the dynamic causal effect changes in regimes with different arbitrage opportunities in such a way that as the arbitrage opportunities increase, the impact of unexpected shocks on prices increases.

This research reveals that traditional models fail to capture lead-lag relationships that emerge when structural changes occur and corroborates that the nature of price discovery process evolves in a different way under different regimes. Results found by traditional methodologies may change whether the sample contains “normal” or “extreme” regimes. Models allowing for time-varying causality (e.g. MS-VECM or GP) enable a better understanding of the lead-lag dynamic and a reduction in bias resulting from neglecting these nonlinearities. These findings are valuable to understand the dynamic evolution of the lead–lag relationship between stock index and stock index futures and can be applied to strategy trading in hedging and portfolio investments.

## References


[1] Andersen, T. G. (2000). Some reflections on analysis of high-frequency data. *Journal of Business and Economic Statistics*, 18(2), 146-153

[2] Andersen, T. G., Bollerslev, T. and Cai, J. (2000). Intraday and interday volatility in the Japanese stock market. *Journal of International Financial Markets, Institutions and Money*, 10(2), 107-130

[3] Ang, A. and Timmermann, A. (2012). Regime changes and financial markets . *Annual Review of Financial Econometrics*, 4(1), 313-37.

[4] Aragó, V., Corredor, P. and Santamaría, R. (2003). Transaction costs, arbitrage, and volatility spillover: a note. *International Review of Economics and Finance*, 12, 399–415

[5] Balcilar, M., Gungor, H., and Hammoudeh, S. (2015). The time-varying causality between spot and futures crude oil prices: A regime switching approach. *International Review of Economics and Finance*, 40, 51-71.

[6] Bartley, W. A., Lee, J. and Strazicich, M. C. (2001). Testing the null of cointegration in the presence of a structural break. *Economics Letters*, 73(3), 315-323

[7] Blanco, R. (2003). Transmisión de información entre el mercado de futuros sobre el Ibex 35 y el contado. *Revista de Economía Aplicada*, 11(31), 81-101.

[8] Bollerslev, T. and Wooldridge, J. M. (1992). Quasi-maximum likelihood estimation and inference in dynamic models with time-varying covariances. *Econometric Reviews*, 11(2), 143-172..

[9] Bondarenko, O. (2003). Statistical arbitrage and securities prices. *The Review of Financial Studies*, 16(3), 875-919.

[10] Chan, K. (1992). A further analysis of the lead-lag relationship between the cash market and stock index futures market. *The Review of Financial Studies*, 5(1), 123-152.

[11] Chan, K., Chan, K. C. and Karolyi, G. A. (1991). Intraday volatility in the stock index and stock index futures markets. *The Review of Financial Studies* 4(4), 657-684.

[12] Dwyer, G. P., Locke, P. and Yu, W. (1996). Index arbitrage and nonlinear dynamics between the S&P 500 futures and cash. *The Review of Financial Studies*, 9(1), 301-332.

[13] Ehrmann, M., Ellison, M., Valla, N. (2003). Regime-dependent impulse response functions in a Markov-switching vector autoregression model.*Economics Letters*, 78(3), 295-299.

[14] Engle, R. F. and Granger, C. W. (1987). Co-integration and error correction: representation, estimation, and testing. *Econometrica*, 55(2), 251-276.

[15] Fassas, A. P., and Siriopoulos, C. (2019). Intraday price discovery and volatility spillovers in an emerging market. International Review of Economics and Finance, 59, 333-346.

[16] Frino, A., Walter, T. and West, A. (2000). The lead–lag relationship between equities and stock index futures markets around information releases. *The Journal of Futures Markets*, 20(5), 467-487.

[17] Gallant, A. R., Rossi, P. E. and Tauchen, G. (1993). Nonlinear dynamic structures. *Econometrica*, 61(4), 871-907.

[18] Gao, A. H., and Wang, G. H. (1999). Modeling nonlinear dynamics of daily futures price changes. *The Journal of Futures Markets*, 19(3), 325-351.

[19] Gong, C. C., Ji, S. D., Su, L. L., Li, S. P., and Ren, F. (2016). The lead–lag relationship between stock index and stock index futures: A thermal optimal path method.*Physica A: Statistical Mechanics and its Applications*, 444, 63-72.

[20] Hamilton, J. D. (1988). Rational-expectations econometric analysis of changes in regime: An investigation of the term structure of interest rates. *Journal of Economic Dynamics and Control*, 12(2-3), 385-423.

[21] Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica*, 57(2), 357-384.

[22] Harris, F.H. de B, McInish, T. H., Shoesmith, G. L. and Wood, R. A. (1995). Cointegration, error correction, and price discovery on informationally linked security markets. *Journal of Financial and Quantitative Analysis*, 30(4), 563-579.

[23] Hogan, S., Jarrow, R., Teo, M. and Warachka, M. (2004). Testing market efficiency using statistical arbitrage with applications to momentum and value strategies. *Journal of Financial Economics*, 73(3), 525-565.

[24] Hsieh, D. A. (1991). Chaos and nonlinear dynamics: application to financial markets. *The Journal of Finance*, 46(5), 1839-1877.

[25] Huth, N. and Abergel, F. (2014). High-frequency lead/lag relationships-Empirical facts. *Journal of Empirical Finance*, 26, 41-58

[26] Kawaller, I. G., Kock, P. D. andKoch, T. W. (1987). The temporal relationship between S&P 500 futures prices and the S&P 500 index. *The Journal of Finance*, 42(5), 1309-1329.

[27] Koop, G., Pesaran, M. H. and Potter, S. M. (1996). Impulse response analysis in nonlinear multivariate models. *Journal of Econometrics*, 74(1), 119-147.

[28] Kwiatkowski, D., Phillips, P. C., Schmidt, P. and Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root?. *Journal of Econometrics*, 54(1-3), 159-178.

[29] Lee, J., Huang, C. J. and Shin, Y. (1997). On stationary tests in the presence of structural breaks. *Economics Letters*, 55(2), 165-172.

[30] Lee, C.I. and Mathur, I., (1999). The influence of information arrival on market microstructure: evidence from three related markets. *Financial Review*, 34(1), 1–26..

[31] Li, M. Y. L. (2009). The dynamics of the relationship between spot and futures markets under high and low variance regimes. *Applied Stochastic Models in Business and Industry*, 25(6), 696-718.

[32] Lien, D., Tse, Y. K., and Zhang, X. (2003). Structural change and lead-lag relationship between the Nikkei spot index and futures price: A genetic programming approach. *Quantitative Finance*, 3(2), 136-144.

[33] Lien D., and Yang, L. (2003). Contract settlement specification and price discovery: Empirical evidence in Australia individual share futures market.*International Review of Economics Finance*, 12(4), 495-512.

[34] Ng, N. (1987). Detecting spot prices forecasts in futures prices using causality tests. *Review of Futures Markets*, 6(2), 250-267.

[35] Pardo, Á. and Climent, F. (2000). Relaciones Temporales entre el Contrato de Futuro sobre Ibex35 y su Activo Subyacente. *Investigaciones Económicas*, 24(1), 219-236.

[36] Perez-Quiros, G., and Timmermann, A. (2000). Firm size and cyclical variations in stock returns. *The Journal of Finance*, 55(3), 1229-1262.

[37] Perlin, M. (2010). MS Regress – The MATLAB Package for Markov Regime Switching Models. Accessed 04.12.2015.

[38] Perron, P. (1989). The Great Crash, the Oil Price Shock and the unit root hypothesis. *Econometrica*, 57(6), 1361–1401.

[39] Perron, P. (1990). Testing for a unit root in time series with a changing mean. *Journal of Business and Economic Statistics*, 8(2), 153–162.

[40] Pesaran, H. H. and Shin, Y. (1998). Generalized impulse response analysis in linear multivariate models. *Economics Letters*, 58(1), 17-29.

[41] Sarno, L. and Valente G. (2000). The cost of carry model and regime shifts in stock index futures markets: An empirical investigation. *The Journal of Futures Markets*, 20(7), 603-624.

[42] Sarno, L., and Valente, G. (2005). Modelling and forecasting stock returns: exploiting the futures market, regime shifts and international spillovers. *Journal of Applied Econometrics*, 20(3), 345-376.

[43] Stoll, H. R. and Whaley, R.E. (1990). The dynamics of stock index and stock index futures returns. *The Journal of Financial and Quantitative Analysis*, 25(4), 441-468.

[44] Theissen, E. (2012). Price discovery in spot and futures markets: a reconsideration. *European Journal of Finance* 18(10), 969-987.

[45] Tse, Y. (1999). Price discovery and volatility spillovers in the DJIA index and futures markets. *The Journal of Futures Markets*, 19(8), 911-930.

[46] Wahab, M. and Lashgari, M. (1993). Price dynamics and error correction in stock index and stock index futures markets: A cointegration approach. *The Journal of Futures Markets*, 13(7), 711-742.