

# Statistical evidence for short-run dynamics between spot and futures equity markets and economic value of basis trade strategies

The case of DAX

Konstantinos I. Sakellariou \*

This draft: October 3, 2009

## Abstract

The aim of this paper is to examine the existence of short-run lead-lag effects between the spot and the futures market in the German stock exchange and particularly the DAX index, over the period 1/2000 - 12/2003, using intraday data. As long as such relationships are established we investigate the economic value of spread trades motivated by the abovementioned relationships. To model the lead-lag effects we employ the multivariate Threshold Regression Model (TRM) of Tsay (1998) and for the economic value we use a reduced version of the methodology introduced by Fleming *et al.* (2001). Our main finding is that there exist quite robust short-run effects between the two markets across time. Also, there is some evidence for economic value in the basis-trade strategies which is totally diluted when realistic transaction costs apply.

---

\*Warwick Business School, Sakellariou.K@Gmail.com, +44-(0)-7527918382, 48 Gale Close Street, Mitcham, CR4 3QJ, UK.

I am thankful to Dr. Ingmar Nolte for providing me the data for this paper.

## 1. Introduction – Literature Review

In frictionless markets and without transaction costs, the intraday spot index and index futures prices should move in lockstep, because of the no-arbitrage condition. However, in the presence of frictions or even stochastic interest rates, one market could consistently lead the other so there would be a lead-lag relationship. That issue has attracted the attention of both academics and practitioners and the most popular explanations suggested are: non-frequent trading of some of the index constituents and the presence of lower transaction costs in the futures market (see e.g. Stoll and Whaley (1990a)).

Especially the US market has been widely investigated. In their early study Kawaller *et al.* (1987) examined the S&P500 and they found that futures market leads the cash by between 20 to 45 minutes. Moreover, Chan (1992) investigated the same index for the period 1985 - 1987 and provided evidence that futures lead the spot index, while there is no evidence for the inverse relation. In addition, he found that the degree of leadership is stronger when there is market-wide information, i.e. news that affect the entire market. In the same group of authors who have investigated the S&P500 and found that futures market responds significantly faster to new information, we could include Stoll and Whaley (1990a), Ghosh (1992) and Dwyer *et al.* (1996) amongst others. Examining the case of Dow Jones for the period 1997 - 1998, Tse (1999) found consistent results, meaning that futures market leads the spot for 88.3% of new information.

While the majority of empirical research has focused on the US capital market, there are very few studies focusing on the Asian or European Market, especially the German exchange. With respect to Asian exchanges, Min and Najand (1999) investigated the Korean market and they found consistent results with previous studies as the futures market leads the cash by as long as 30 minutes. Jiang *et al.* (2001) investigated the Hong Kong market and they showed that the futures market is faster than the spot, in the accumulation of new information especially in falling markets. Focusing on the European exchanges, Abhyankar (1995) and Shyy *et al.* (1996) investigated the FTSE-100 and CAC Indices respectively and they all confirmed the general pattern of futures market leadership. The German index DAX, which is the case we will study, has been investigated amongst others by Grublichler *et al.* (1995) and Booth *et al.* (1999) and they all concluded that futures market has more information share than the spot one.

In this study, we will also look for lead-lag relationship in the DAX index, using intradaily data for the period 2000 - 2003. Apart from employing a more recent data set, we contribute to the literature by using the relatively new Threshold Regression Model (TRM hereafter) (see e.g. Tse and Chan (2009)). The advantage of this approach is that it enables us to capture the short-run dynamics under different market conditions, which are proxied by the threshold variable and stratify the data

set into several piecewise linear fractions. Also, contrary to some researchers who have examined the abovementioned relationship only when arbitrarily set extreme market conditions occur, we adopt a data-based method to divide our sample into different regimes.

Another important contribution of our analysis is the use of economic criteria to assess the significance of lead-lag effects. With respect to that issue, the majority of the studies has focused only on statistical predictability, while very few studies have investigated its economic significance. The evaluation based on economic criteria is important because evidence from statistical measures, like the Mean Squared Error (MSE) or the Sharpe ratio (1968) and the Jensen's Alpha (1968), do not guarantee that an investor could earn profits from a strategy which is based on them. Besides, it has been shown that in many cases they suffer from biases (see e.g. Sharpe (1994), Goetzmann *et al.* (2002) and Goetzmann *et al.* (2007) for critique on several ratios). This could be due to the actual distribution of the returns, which is usually not Gaussian, or because of the transaction costs, among many possible explanations. Thus, in an attempt to overcome those problems, especially for known market anomalies (like Holidays, Mondays or end of year effects) many authors have tried to assess the identified predictabilities in an economic framework.

The traditional way is by setting up a strategy and then trying to see how it performs over time when transaction costs have been taken into consideration (see e.g. Pesaran and Timmermann (1995) for stock returns, or Bollen and Busse (2004) for mutual funds performance). In the same context of evaluating strategies based on economic criteria, West *et al.* (WEC hereafter) introduced the idea of comparing strategies based on the utility they "offer" to the investor and since then, many authors have also tried to incorporate the mean-variance portfolio framework of Markowitz (1952) in the utility based valuation. For instance, Fleming *et al.* (2001) using quadratic utilities, examined if there is economic value in volatility timing. Following the same approach, Della Corte *et al.* (2008) and (2009) investigated the importance of volatility and correlation timing respectively. Tsiakas (2009) and Marquering and Verbeek (2004) are two more examples which evaluated stock market patterns in a utility based framework. For the purposes of our analysis, on a weekly basis, we are going to set up three spread-trade strategies based on the previous week's lead-lag relationship analysis and we will use the approach proposed by WEC, modified à la Fleming *et al.* to examine if the potential benefits from such interrelations have been actually exploited.

To preview our results, we find evidence in favour of the usage of the Threshold Regression to model the short-run dynamics between the spot and the futures market. In our data set the futures market consistently leads the spot for up to 6 minutes, while there are cases where the lead-lag effects last up to 10 minutes. On the other hand the spot market doesn't lead the spot for more than 2 minutes in the great majority of the cases and when this happens it is economically insignificant. Every

spread-trade strategy outperforms the DAX index (without including transaction costs), however only one beats the risk-free rate in the long run. With respect to the economic valuation, the evidence is mixed depending on the used benchmark and the period under examination. In general, the implemented spread-trade strategies are economically profitable in the short run and not in the long run (but for one case), though any profitability diminishes when transaction costs apply.

The rest of the paper is organized as follows. In the next section we describe the microstructural characteristics of the DAX index and we focus on issues which will affect our analysis. Section 3. describes our data set and the adopted filtering methods. In Section 4. we discuss the way we model the lead-lag relations, build the strategies and evaluate their performance. Our empirical results are reported in Section 5. , whilst Section 6. concludes.

## 2. Microstructure Issues

The German stock market is one of the biggest equity markets in the world in terms of liquidity and capitalisation. According to the European Central Bank<sup>1</sup> its capitalisation and trading volume for the year 2007 were 1,429,955 and 3,144,150 million Euros respectively. Within this market, the DAX (Deutscher Aktien Index) is the most popular German equity market index and consists of the 30 biggest German stocks in terms of order book volume and market capitalisation. Unlike other value weighted major indices (e.g. S&P500), the DAX reflects stock distribution in its calculation, so it is considered to be a “Total return index”. Table 1 presents the index constituents and weighting as of Close of Business 30/12/2003, when our sample ends, while Table 2 presents the evolution of the market value and trade volume of both the spot and the futures market throughout the entire sample.

The DAX initiated on the 31<sup>st</sup> of December 1987 with a starting value of 1000 units. In its level calculation price from electronic trading and floor trading through all the eight exchanges located in Germany, were taken into account. However, more than 95% of the total trading volume of German stocks takes place in the Frankfurt Stock Exchange and thus, since the 18<sup>th</sup> of June 1999 only XETRA equity prices are used in the calculation of the index. In our sample, XETRA trading day lasts from 9.00 a.m.<sup>2</sup> till 5.30 p.m., while Deutsche Boerse also offers an indication for the after-hours trading period between 5.45 p.m. till 10.00 p.m., the so-called X-DAX. In our analysis, we employ data only from the former trading session as it is by far more active.

---

<sup>1</sup>Annual Report (2007): Securities trading: market capitalisation of listed companies and executed trades.

<sup>2</sup>In this paper time is always quoted in German local time only.

Table 1: DAX Index component stocks

#	Company Name	No of shares		#	Company Name	Index Weight		No of shares in the Index
		Index Weight	in the Index			Index Weight	in the Index	
1	Adidas	0.972%	3	16	Deutsche Telekom	8.255%	136	
2	Allianz	8.005%	19	17	E.ON	7.993%	37	
3	Altana	0.791%	4	18	Fresenius Medical Care & Co	0.460%	2	
4	BASF	6.021%	32	19	Henkel & Co	0.872%	3	
5	Bayer	3.768%	39	20	Infineon Technologies	1.134%	25	
6	Bayer Schering Pharma	1.671%	10	21	Linde	0.816%	5	
7	Bayerische Hypo & Vereinsbank	1.570%	20	22	MAN	0.576%	6	
8	Bayerische Motoren Werke	2.892%	19	23	Metro	1.188%	8	
9	Commerzbank	1.792%	28	24	Muenchener Rueckversicherungs	3.883%	10	
10	Continental	0.961%	8	25	RWE	2.996%	23	
11	Daimler	7.179%	46	26	SAP	6.491%	12	
12	Deutsche Bank	9.053%	33	27	Siemens	12.525%	47	
13	Deutsche Boerse	1.148%	6	28	ThyssenKrupp	1.527%	23	
14	Deutsche Lufthansa	1.077%	19	29	TUI	0.478%	7	
15	Deutsche Post	1.611%	24	30	Volkswagen	2.294%	12	

This table presents the Constituents of DAX Index as of COB 30/12/2003.

Source: Bloomberg

Table 2: DAX and FDAX price and trade volume evolution from 1/2000 – 12/2003

Date	Price	% Growth	DAX			Price	% Growth	FDAX		
			Volume	% Growth	% Growth			Contracts	% Growth	% Growth
30/12/1999	6958.14	-	11,356,744	-	-	7027.00	-	11,055	-	-
31/12/2000	6433.61	-7.538	15,941,576	40.371	-	6500.00	-7.500	11,986	8.422	8.422
28/12/2001	5160.10	-19.795	22,827,002	43.192	-	5188.80	-20.177	8,578	-28.433	-28.433
30/12/2002	2892.63	-43.942	29,283,860	28.286	-	2904.50	-44.020	11,911	120.459	120.459
30/12/2003	3965.16	37.078	34,660,664	18.361	-	3995.00	37.545	20,765	9.804	9.804

This table presents the evolution of price and trade volume of the DAX Index and the FDAX.

As trade volume for FDAX the number of daily traded contracts is used.

Source: Bloomberg

The first futures contract in DAX index (FDAX) initiated on the 23<sup>rd</sup> of November 1990, the multiplier was set to DM100 per index point and the tick size to 0.5 points. However, since the passage to the Euro-era, the contract size was changed to €25 per index point. Contrary to the components stocks, FDAX contracts were initially traded electronically through the Deutsche Terminborse, located in Frankfurt, but since September 1998 trading is performed through the newly formed cross border exchange EUREX. As usual, contracts are offered for the next two quarterly expiration months (March, June, September and December) and they expire at the third Friday of the respective settlement month if that is an exchange trading day, otherwise the exchange trading day immediately preceding this Friday. At the end of the examined period, the trading day is from 8.30 a.m. till 10.00 p.m., but we mention that there are several changes in the auction plan during our sample.

### 3. Data

In our analysis we use intradaily data of DAX spot index and DAX futures index for the period 1/1/2000 till 31/12/2003. The spot index<sup>3</sup> time series consists of the last transaction prices given every 1 minute. For the futures index we always take the closest to maturity contract, which is generally more actively traded and our time series consists of transactions data which are irregularly spaced. Because the spot time series is homogenous, the futures one should be the same, so we transform our data into one-minute spaced observations using the previous-tick interpolation proposed by Wasserfallen and Zimmermann (1985), according to which:

$$P_j = P_m^*, \text{ where } t_m^* = \max\{t_i^* | t_i^* < t_j\} \quad (1)$$

where:  $P_j$  is the regularly spaced observation,  $P_m^*$  is the irregularly spaced price,  $t_m^*$  is the execution time of the  $m^{th}$  trade and  $t_j$  is the time stamp for the  $j^{th}$  regularly spaced price. As mentioned by Dacorogna *et al.* (2001), this method of interpolation is preferable, compared to other methods such as the next-point, cubic or linear interpolation because it is not using information which is not available at time  $j$ .

After homogenising our time series, the next step is to filter them for erroneous inputs. It is well documented in the literature that false records (e.g. misplaced decimal points) can be found even in fully automated systems (e.g. Falkenberry (2002)) and several methods have been suggested to overcome that issue, like the Olsen and Associates algorithm (Muller (2001)) or Oomen (2006). We employ the

---

<sup>3</sup>Contrary to prior studies like Chan (1992) which calculate the index level from its component stocks, we use the price quoted by XETRA since DAX index is value weighted and it is practically impossible for us to collect its constituents weighting in 1-minute intervals for 4 years and dynamically replicate it.

fairly simple, yet parsimonious method suggested by Brownlees and Gallo (2006), according to which the observation  $P_t$  is kept only if:

$$|P_t - \bar{P}_t(k)| < 3\sigma_t(k) + \zeta \quad (2)$$

where:  $\bar{P}_t(k)$  and  $\sigma_t(k)$  denote respectively the  $\delta$ -trimmed sample mean and sample standard deviation (see Appendix A.1 for a brief description of those measures) of a neighbourhood of  $k$  observations around time  $t$  and  $\zeta$  is the granularity parameter. For our analysis  $k$  has been set equal to  $100^4$  and we repeat that process day by day. So, for the beginning and the ending of each trading session we use a window of the first and the last observations of the day respectively, while for every other observation a symmetric window around that point is applied.

Furthermore, in our analysis we use observations only when both the spot and the futures market are jointly open. However, it is common in the literature to exclude some observations at the beginning and the ending of each session, because there is not sufficient trading volume or because their informational content is affected by closing and opening overnight positions (see e.g. Jiang *et al.* (2001) or Tse and Chan (2009)). For our study, we exclude 30 minutes at the beginning and at the end of each day, as in that interval very few trades occur in the futures market, hence in our sample we include trades taking place every day between 9.30 a.m. and 5.00 p.m., but for the New Year's eve when the session ends at 1.30 p.m..

Next, we transform the price time series into 1-minute returns using the following formula:

$$r_{M,t} = \ln(P_{M,t}) - \ln(P_{M,t-1}) \quad (3)$$

where:  $r_{M,t}$  and  $P_{M,t}$  are respectively the return and the price of asset  $M$  at time  $t$  and  $M$  is either the spot index or the futures contract. We mention that our data set doesn't seem to suffer by the infrequent trading problem mentioned by Stoll and Whaley (1990a) as, within the 4 years sample, when both markets are active, less than 0.02% of the durations between the trades are higher than one minute. Table 3 reports the frequency of observing more than one minutes of inactivity in both markets. With respect to the futures market, such findings are more frequent during the X-DAX operating hours.

---

<sup>4</sup>Different values, like 50 or 200, have been tested but the results were not considerably different.

Table 3: **Frequency of more than one minutes of inactivity**

Year	Spot Market				Futures Market			
	Total Observations	Inactive Minutes	Inactivity at Joint Hours	Percentage ( <i>in bps</i> )	Total Observations	Inactive Minutes	Inactivity at Joint Hours	Percentage ( <i>in bps</i> )
2000	154,164	20	18	1.168	149,447	986	16	1.071
2001	170,603	79	36	2.110	164,925	914	12	0.728
2002	170,415	63	30	1.760	165,056	188	8	0.485
2003	164,850	86	33	2.002	165,762	124	20	1.207

This table reports the frequency of observing more than one minutes without update in the prices for each year and market in our sample.

Columns Percentage present the ratio of Inactive minutes during joint trading hours over the Total number of Observations, expressed in basis points.

Finally, we use the Euro Overnight Index Average (EONIA), as a proxy of the risk-free interest rate, which is needed to evaluate the daily returns of our strategies. It is an effective overnight rate computed as the weighted average of all overnight lending transactions in the interbank market, initiated within the euro area by a contributing panel of banks. Also, we use the Total Return Index of the EONIA rate, provided by Deutsche Bank, as an indicator of the performance of a Money Market portfolio with a very short investment horizon.

## 4. Methodology

### 4.1 Lead-Lag Effects

We investigate the existence of lead-lag effects using a Threshold Regression Model as was introduced by Tsay (1989) and extended to the multivariate case by Tsay (1998). According to this family of models, the relation between the two assets might be different depending on the value of an exogenous threshold variable.

To be consistent with the existing literature and concerning our data set, as the threshold variable we use the basis spread between the spot and the futures market, the volatility of the spot market and the trading volume of the closest to expiry futures contract, calculated as the number of contracts transacted. Obviously, increased volatility renders market more risky and certain intraday patterns have been documented, so participants might behave differently when they face extreme volatility values (e.g. Harris (1986a), Muller *et al.* (1990) and Andersen and Bollerslev (1997)).

Similar patterns in intraday trading volume have been found as well (see Admati and Pfleiderer (1988) and Easley and O'Hara (1992) amongst others) and Kyle (1985) pointed out the different informational content of different trading volumes. Also, as mentioned by Engle and Russell (1998) the different time intervals between two consecutive trades and their autocorrelation affect the dynamics of the price dis-



covery process. However, we don't have irregularly spaced data for both markets, so we cannot perform such analysis. Nevertheless, in very small intervals we expect Conditional Duration to be highly correlated (in absolute terms) with the number of contracts transacted in that period and this is one more reason supporting the use of trading volume as a regime variable.

Last, it is known that arbitrage opportunities occur under very high or very low values of the basis between the spot and the futures markets (e.g. Brennan and Schwartz (1990)), so we expect different mean-relation of the two markets in such conditions. Particularly, we define the basis as  $F_t - S_t$ , where:  $F_t$  and  $S_t$  are the price of the futures contract and the spot index at time  $t$  respectively, because of the contango situation, which is usually observed in financial markets.

In general, an  $i$ -regime TRM has the form:

$$Y_t = a_0^i + \sum_{h=1}^p a_h^i Y_{t-h} + \sum_{j=1}^p b_j^i X_{t-j} + e_t^i \quad (4)$$

where:  $i$  is the regime indicator,  $p$  is the maximum number of lagged observations<sup>5</sup>,  $Y_t$  and  $X_t$  are the returns of either the spot or the futures market depending on which market we are checking, while statistically significant  $b$  coefficients indicate lead-lag relations. For our analysis we run two sort of regressions, for futures market leading the spot and vice versa and we focus on estimates which are not only statistically, but economically significant as well (in our case with absolute value higher than 0.1). Also, we do not include in our analysis any significant coefficients which come after insignificant ones (i.e. if we get  $b_1$ ,  $b_2$ ,  $b_5$  and  $b_6$  to be different from zero, then we will say that there exist lead-lag effects for two minutes only).

Before estimating the aforementioned model we formally check for piecewise linear dynamics between the variables, in order to ensure that the TR model is the appropriate one. Several methods have been suggested to perform that test, like that of Chan and Tong (1990) who used a likelihood ratio approach and Monte Carlo simulations to obtain the critical values. On the contrary, Tsay (1998) proposed a test using re-arranged regressions and predictive residuals to construct the test statistics. The benefit of the latter approach is that it has a known asymptotic chi-squared distribution, it is not affected by the specification of the alternative model and it has been proven to have satisfying detecting power of threshold non-linearity. For those reasons, we use the Tsay (1998) test in our analysis and Appendix A.2 illustrates the steps to perform the test.

If the null hypothesis that  $Y_t$  is linear in time, versus the alternative that it follows a TRM, is not rejected, we proceed directly to the estimation of a single regime

<sup>5</sup>Some authors used different number of lags for the spot and futures contract observations. However, we use the same number of lags in order to be consistent with Tse and Chan (2009).

model. Otherwise, we have to determine the number of different regimes and the exact regime-switching values. Several approaches have been suggested in order to identify the optimal number of segments. We adopt the Bai and Perron (2003) algorithm for modeling partial structural changes models, which is a generalisation of the Bellman and Roth (1969) dynamic programming algorithm to optimise the fitting of a continuous curve by a sequence of straight line segments. Briefly, that approach applies to the threshold regression case as follows: after arranging the dependent variable and the regressors based on the threshold variable, we calculate all the possible combinations of the starting and ending points of different number of regimes within the sample. Then using the OLS methodology we should estimate Equation (4) for all the previously computed combinations and we keep the sum of squared residuals (SSR). The optimal regime structure will be the feasible linear combination of break points which minimises the SSR. Practically, all the possible combinations fall within a triangular matrix with dimension  $T$ , which is the number of observations in our full sample. In that case, the maximum number of combinations is reduced to  $T(T + 1)/2^6$ , which is considerably lower than the counterpart number needed in other approaches like the “Grid methodology” of Tsay (1989) and (1998).

The majority of empirical findings (see e.g. Forbes *et al.* (1997), Tsay (1998) and Tse and Chan (2009), amongst others) support that at most two break points provide sufficient fit of the data, dividing the data set in 3 sub-samples corresponding to good, normal and bad market conditions. So, although there is no particular theory behind the proper number of different regimes, for computational reasons we limit our analysis to the case of maximum three distinct segments. Also, we slightly modify the Bai and Perron approach in the following way: instead of looking directly for the global minimum of SSR by trying small changes in the dimensions of the segments, we allow for bigger steps in the changes of potential break-points, we locate the fragment which minimizes the SSR and finally look for the exact break points by trying very small changes around those fragments<sup>7</sup>.

After the optimal number of segments has been specified, we estimate the parameters of the TRM for each different regime. Tsay (1989) and (1998) have examined in detail that issue and the properties of the estimates. They showed that, conditional least squares estimators are the appropriate method, as they become strongly consistent as the sample size increases. Tsay (1998) also provided explicit parametric formulation for the multivariate model coefficients, as follows:

---

<sup>6</sup>A further reduction of the minimum number of combinations is doable when restrictions like the minimum number of observations within each segment are imposed.

<sup>7</sup>For different small step values our results are the rather similar, however the algorithm runs dramatically faster.

$$\hat{\Phi}_i(BPi) = \left( \sum_t^{(i)} \mathbf{X}_t \mathbf{X}_t' \right)^{-1} \left( \sum_t^{(i)} \mathbf{X}_t \mathbf{y}_t' \right) \quad (5)$$

where:  $i$  is the regime indicator (1, 2 or 3 in our study),  $BP_i$  is the breakpoint for the operative regime,  $\sum_t^{(i)}$  is the summation across observation in regime  $i$ ,  $\mathbf{X}_t$  is the matrix of the regressors (i.e.  $p$  lagged observations of spot and futures market returns in our case) and  $\mathbf{y}_t$  is the vector of the dependent variable.

With respect to the correct model specification, we determine the proper number of lagged values to be included in the regression based on the properties of the residuals and particularly their serial correlation<sup>8</sup>. For each regime we investigate the ACF and perform the Breusch – Godfrey (1978) test and once we have managed to get serially uncorrelated residuals for the entire week, we use the Akaike information criterion (AIC) to select the length of our model. The exact formula for the multivariate case is:

$$AIC(p, s) = \sum_{i=1}^s \left[ n_i \ln \left( \left| \frac{1}{n_i} \sum_t^{n_i} \hat{\varepsilon}_t^{(i)} \hat{\varepsilon}_t^{(i)'} \right| \right) + 2(2p + 1) \right] \quad (6)$$

where:  $i$  is the regime indicator,  $s$  is the total number of regimes,  $n_i$  is the number of observations in regime  $i$ ,  $p$  is the lag length and  $\hat{\varepsilon}_t^{(i)}$  are the residuals. As mentioned by Tsay (1998) “...when  $p$  and  $s$  are fixed, AIC is asymptotically equivalent to selecting the model that has the smallest generalised residual variance using the conditional least squares method...” which applies in our case as well, as every sample we use has approximately 2100 data points.

## 4.2 Dynamic Strategies

After investigating the existence of lead–lag relation between the spot and the futures market, the next step is to set up a strategy which aims to exploit those effects, as well as to assess its statistical and economic significance. We suggest the following, relatively simple, spread–based strategy: at the beginning of each week we examine the existence of short–run dynamics between the two markets based on information of the previous week, using each one of the three different threshold variables, in order to find which market systematically precedes the other and for how long. After that, during the week, every time the basis exceeds a particular value (to be defined later) we initiate a basis–trade (i.e. go long on one market and short the other) and we

---

<sup>8</sup>Other properties of the residuals, like excess skewness or kurtosis, are not examined as they can be affected by the conditional volatility model, while volatility spillovers are out of the scope of this project.

stick to this strategy for a period equal to the estimated lead-lag effects. Also, when there is no spread-trade active, we maintain a long position on the spot index.

As an example, suppose that our Monday-morning analysis has shown that there are significant lead-lag effects and that the futures market leads the spot for up to 6 minutes. During the day, whenever we see a rapid increase in the spread, for instance because the futures market has fallen, we would expect the spot market to fall in the next few minutes as well. So, if this increase exceeds that week's critical value, then there is a signal to begin a spread-trade, which would be to go short the spot market and long the futures one. In this case, the return of the dynamic strategy at time  $t$ ,  $r_{DS,t}$ , would be calculated as follows:

$$r_{DS,t} = \begin{cases} r_{S,t}, & \text{if } t \leq \tilde{t} \\ r_{S,t}, & \text{if } t = \tilde{t} + 1 \\ r_{F,t+} - r_{S,t}, & t \in \{\tilde{t} + 2, \tilde{t} + 6\} \end{cases}$$

where:  $r_{S,t}$  is the return of the spot index at time  $t$ ,  $r_{F,t}$  is the return of the futures contract at the same time and  $\tilde{t}$  is the time when the signal to initiate a basis-trade is observed. We mention that we assume investors who need only 1 minute to implement the strategy, so the first minute immediately after the signal is observed, the return of the strategy is still the return of the spot index.

We determine the critical value to initiate the spread-trades as the 7<sup>th</sup> percentile of the previous week's distribution of the absolute basis percentage changes. Also, in an attempt to keep our distribution up-to-date we add to it all the available information of the operative week. We decided to use the 7% point, as (to the extent of our knowledge) the average intraday trader wants to do approximately 20–30 full trades per day. We mention that the usage of the previous week's distribution allows the actual number of trades to be potentially different from the previous range, which is, of course, a desirable property. In that point we would like to mention that during the period into consideration no short-selling restrictions were imposed like today, thus it was theoretically possible for an investor to have only a purely short positions portfolio. In addition, to circumvent the problem of replicating and shorting the entire spot index, one could have traded Contracts for Difference (CFDs) in DAX (which initiated in early 2000), in order to get the same pay-off, instead of trading individual stocks.

In the rest of this paper, with respect to the notation of Appendix A.1, we name each dynamic strategy based on the used threshold variable,  $z_t$ . For instance, when the market volatility is used to rank the observations in the re-arranged regression, then the strategy which is based on those estimates is called the “volatility strategy”.

### 4.3 Economic Valuation

In this section we describe how to assess the value of the previously mentioned “spread” strategies.

The traditional method employed in ranking different strategies or assets in general, is the use statistical criteria like the Sharpe ratio (1966), Jensen’s Alpha (1968) or the Sortino ratio (1999), which are all risk-adjusted performance measures. However, these ratios are linear in return and several studies have shown that they suffer either from biases in the case of time-varying parameters (like volatility or portfolio weights) and dynamic strategies<sup>9</sup>, or they provide no insight for economic value. On the other hand, West *et al.* (1993) showed that when shorting is done based on statistical criteria the result might be different than when economic criteria are used. By evaluating conditional variance estimates and using exponential utility function, they showed that an estimate is superior to another when it derives higher utility for the same amount of initial wealth. Given the latter, our choice is to apply the WEC methodology as it was generalized Fleming *et al.* (2001). In their alternative, they used a quadratic utility function of the form:

$$U_{(W_{t+1})} = W_t R_{t+1} - \frac{\varphi W_t^2}{2} R_{t+1}^2 \quad (7)$$

where:  $W_t$  is the wealth at time  $t$ ,  $R_{t+1}$  is the wealth relative from time  $t$  to  $t+1$  (i.e.  $\frac{W_{t+1}}{W_t}$ ) and  $\varphi$  is the degree of absolute risk aversion. In this family of utility functions the degree of relative risk aversion,  $\gamma$ , takes the form:

$$\gamma_t = \frac{\varphi W_t}{1 - \varphi W_t}. \quad (8)$$

In order to allow for comparison across different portfolios we keep the product  $\varphi W_t$  constant, so the entire Equation (7) can be described by the  $\gamma$  coefficient of Equation (8). We consider the quadratic utility function to be superior to the exponential one used by WEC, as it can be viewed as the second order approximation of the representative investor’s utility function, therefor yielding a satisfactory proxy of a wide range of more sophisticated cases (see e.g. Hlawitschka (1994)). In this particular framework, it is shown that the average realised utility over a time period, consistently estimates the expected utility generated by a strategy. Assuming for simplicity initial investment of €1, we have:

$$\bar{U}(\text{strategy}) = \sum_{t=0}^{T-1} \left\{ R_{p,t+1} - \frac{\gamma}{2(1+\gamma)} R_{p,t+1}^2 \right\} \quad (9)$$

---

<sup>9</sup>see Marquering and Verbeek (2004) and Han (2006) among others.

where  $\bar{U}(\cdot)$  is the average realised utility. One step further, if proportional to trade size performance fees are applicable, Fleming *et al.* (2001) suggest that a strategy adds value to the investor if and only if a positive performance fee can apply to it. That fee can be interpreted as the amount which the representative investor would be willing to pay to switch from the benchmark strategy to an alternative one. Consequently, the best strategy is the one to which the highest expenses will be charged. In the utility ranking context, the appropriate performance fee for each strategy is the one which equates investors average utility with the one of the benchmark portfolio. Algebraically this is expressed as:

$$\sum_{t=0}^{T-1} \left[ (R_{t+1}^* - \Phi) - \frac{\gamma}{2(1+\gamma)} (R_{t+1}^* - \Phi)^2 \right] = \sum_{t=0}^{T-1} \left[ R_{t+1} - \frac{\gamma}{2(1+\gamma)} R_{t+1}^2 \right] \quad (10)$$

where  $R_t^*$  is the wealth relative (as described before) of the dynamic strategy at time  $t$  and  $\Phi$  is the above mentioned performance fee. We mention that the left and the right-hand-sides of Equation (10) represent the average realised utility derived from the dynamic strategy and the benchmark portfolio respectively.

#### 4.4 Transaction Costs

Finally, we examine how the presence of transaction costs affects the performance of strategies which aim to exploit the lead-lag relationship between the spot and the futures market.

As it is done in several recent studies (see e.g. Neely and Weller (2003), Marquering and Verbeek (2004), Della Corte *et al.* (2008) and Rime *et al.* (2009) amongst others) we calculate the Break-Even transaction costs ( $\tau^{BE}$ ), i.e. the level of transaction costs which, in utility terms, renders the investor indifferent between the benchmark and each alternative strategy. In the majority of the previous studies only fixed or proportional transaction costs were taken into account. However, in our study both types of expenses are applicable: usually there is a fixed cost for each futures contract traded and a variable cost proportional to the value traded in the spot market. Thus, for each strategy, we estimate the frontier of bundles of fixed and proportional execution costs, and this strategy should be preferable than the benchmark if and only if individual investor's actual costs are located in the area below the aforementioned frontier. Proportional transaction fees are calculated as percentage points of each trade size in the spot market, while the fixed costs of the futures market are measured in index points (or equivalently €5) charged per contract traded.

Table 4: **Non-linearity tests results**

Critical values	Statistically significant Tsay-statistics					
	Spot on Futures			Futures on Spot		
	Volatility	Volume	Basis	Volatility	Volume	Basis
10%: 34.3815	208	208	208	208	208	208
5%: 37.6524	207	208	208	207	208	206
3%: 39.8803	200	198	208	201	200	205
1%: 44.3141	196	194	204	193	194	205

This table reports the number of weeks, out of 208 in our sample, for which the Null Hypothesis of  $y_t$  being linear in time has been rejected at different levels of statistical significance.

Columns Volatility, Volume and Basis present the test results when Volatility, Volume and Basis have been used as the threshold variable  $z_t$ , respectively.

Under the Null, the Tsay-statistic follows a  $\chi^2$  distribution with  $2p+1$  degrees of freedom, which is 25 in our case.

## 5. Empirical Findings

As mentioned before our sample spans from January 2000 to December 2003 and we perform our analysis in weekly sub-samples. Thus, we have 208 weeks and each one consists of approximately 2100 observations. The first step is to check whether a linear or a TR model is more appropriate. The test is performed on a weekly basis and on the entire sample as well, using the three different threshold variables. A preliminary analysis shows that there exist lead-lag effects up to 12 minutes in our data set, hence with respect to the notation of Equation (4) we use  $p = 12$  to perform the test. Table 4 summarises the results. With respect to the “Spot on Futures” regression, from which we examine if the futures market leads the spot, we observe that our data support the usage of the TRM as all the Tsay-statistics are higher than their critical values at 10% statistical significance level. Most of them are significant even at 5% level. As expected, the results when using the intradaily volatility and the trading volume as threshold variables are quite close. We believe this is the effect of the sample correlation being quite high ranging from 68% to 79% in our sample. With respect to the 3<sup>rd</sup> threshold variable, although its correlation with the other two is negative (this finding is consistent with Chen *et al.* (1995)), ranging from -23% to -31% the results are still the same: the TRM is more appropriate than the linear model. The same pattern is observed in the case of “Future on Spot” group of regressions.

As long as we have evidence for non-linear in time dynamics, the next step is to define the number of different regimes and the exact value of breakpoints, for each possible threshold variable. For the Bai-Perron algorithm we have imposed

Table 5: **Regime break points**

Threshold Variable	Break points		Statistics		No of Observations		
	1 <sup>st</sup>	2 <sup>nd</sup>	Minimum	Maximum	1 <sup>st</sup>	2 <sup>nd</sup>	3 <sup>rd</sup>
Volatility	5.484	7.082	2.852	11.189	899	580	587
Trade Volume	95.468	149.572	35.031	284.822	845	612	608
Basis	19.541	23.931	1.184	104.482	615	921	530

This table reports the average values of the 3 threshold variables across the 208 weeks of our sample. The maximum and the minimum value of each threshold variable at the break-points and the average number of observations in each regime are presented as well.

Volatility is calculated in one minute intervals and quoted in basis points, Trade Volume is quoted in number of futures contract and Basis is presented in index points.

only two restrictions: each segment should have at least 200 observations, in order to enable us to use asymptotic properties of the estimators and the initial step value for changing regimes is 20 observations at a time. Once the global minimum of the process is located, we run the algorithm one more time around that area with step value equal to 1. After analysing our data set on a weekly basis, we find no sub-samples supporting the two regimes only scenario. Table 5 reports the average values of the 1<sup>st</sup> and 2<sup>nd</sup> break points in our data set for the 3 different regime variables. The maximum and the minimum value of each threshold variable and the average number of observations in each regime are reported as well<sup>10</sup>. First of all, we believe it is worth to mention that as we find no sub-samples with one breakpoint only, combined with one regime including too few observations, we have more evidence supporting our previous result of non-linearity of  $y$  in time.

Again, as expected, the results for the Volatility and the Trade Volume are quite close to each other. Volatility is calculated by minute observations and it is quoted in basis points. The minimum and maximum values correspond approximately to 8.55% and 33.56% annual volatility respectively, depending on the annualisation factor. Both the volatility and the trade volume patterns are usual for a financial index. We observe that the two breakpoints are located approximately in the same percentiles of the range between the minimum and the maximum value of each variable. On the contrary the majority of the observations are located in the first segment. We believe this evidence supports our intuition that market participants change their

<sup>10</sup>The exact values for each week are available upon request, but are not reported for the sake of presentation simplicity.



behaviour according to the market conditions, especially in shallow markets. On the other hand, the pattern with respect to the basis as threshold variable is slightly different: here the middle regime has consistently more observations than the other two regimes, however its values ranged between 19.51 and 23.93 which consists a very small portion of the total range of [1.18, 104.48]. Having transactions costs and costs of carry in mind, when the difference in the values between the two assets falls outside of a particular range, arbitrage opportunities occur, thus the dynamics change. Our findings support the existence of that mechanism and suggest that there is a rather tight bound (i.e. the middle regime) within which the majority of transactions takes place, while there are some observations with different dynamics which fall outside of those bounds yet with much higher dispersion. Also, we have to mention that the purpose of our initial definition of Basis (as *Futures* – *Spot* instead of the opposite) has been fulfilled as even the smallest break–point value is positive and equal to 1.184.

The next step is to estimate our TRM and Equation (4) using  $Y$  as the spot index return and  $X$  as the futures contract return, for the case that futures market lead the spot. The same procedure is repeated with  $Y$  being the futures contract return and  $X$  being the spot index returns. Tables 6 and 7 present the results for the former and the latter case respectively. We mention that, although the preliminary analysis indicated lead–lag effects for up to 12 minutes, when we correct for the optimal lag length based on the properties of the residuals, there are no cases including more than 10 lags, consequently in our tables we report results up to that lag length only.

For each threshold variable and regime we report the average values of the model estimates across the 208 weeks of our sample. With respect to the p–value we use the median observation instead of the average, because the latter is affected by extreme values. We present the percentage of the estimated coefficients which are statistically significant as well as the  $R^2$ .

With respect to the “Futures leading the Spot” case, our first finding is that all the constant coefficients of the model are statistically insignificant. The only exception is the case of the Basis being the threshold variable, where most of the coefficients are statistically significant, but their value range from  $10^{-4}$  to  $10^{-6}$ , so they are economically zero and the general picture is still the same. This result is consistent with all the previous studies and the theory, as returns of efficient markets should not have constant components. The  $R^2$  is found in a reasonable range for

Table 6: **Threshold Regression Model Estimates: Future market leads Spot**

Statistic	Const.	Spot observations lag length										Futures observations lag length									
		1	2	3	4	5	6	7	8	9	10	1	2	3	4	5	6	7	8	9	10
Coefficient	0.00	-0.07	-0.14	-0.11	-0.09	-0.08	-0.05	-0.03	-0.02	-0.02	-0.01	0.17	0.14	0.12	0.10	0.08	0.06	0.04	0.02	0.03	0.02
p-value	0.55	0.02	0.03	0.05	0.11	0.16	0.24	0.35	0.41	0.50	0.53	0.00	0.00	0.01	0.02	0.06	0.14	0.26	0.37	0.32	0.40
Significant Coef.	0.00	-0.10	-0.19	-0.16	-0.15	-0.14	-0.11	-0.07	-0.06	-0.03	-0.00	0.18	0.17	0.15	0.14	0.12	0.10	0.08	0.06	0.05	0.03
% Significant	0.13	0.65	0.68	0.62	0.51	0.46	0.35	0.29	0.20	0.18	0.19	0.90	0.79	0.73	0.65	0.59	0.50	0.35	0.25	0.24	0.13
R <sup>2</sup>	0.08																				
Threshold Variable: Volatility - Regime: 2																					
Coefficient	0.00	-0.20	-0.22	-0.17	-0.15	-0.12	-0.08	-0.06	-0.03	-0.03	-0.08	0.32	0.23	0.18	0.14	0.12	0.09	0.06	0.03	0.01	0.02
p-value	0.49	0.00	0.00	0.03	0.08	0.11	0.15	0.27	0.35	0.48	0.53	0.00	0.00	0.02	0.05	0.08	0.15	0.26	0.39	0.20	0.42
Significant Coef.	0.00	-0.28	-0.31	-0.24	-0.24	-0.19	-0.15	-0.14	-0.07	-0.05	-0.01	0.37	0.28	0.23	0.21	0.18	0.14	0.12	0.06	0.11	0.02
% Significant	0.17	0.73	0.69	0.67	0.55	0.52	0.45	0.34	0.26	0.16	0.10	0.86	0.77	0.73	0.62	0.57	0.45	0.39	0.30	0.25	0.20
R <sup>2</sup>	0.13																				
Threshold Variable: Volatility - Regime: 3																					
Coefficient	0.00	-0.16	-0.21	-0.17	-0.14	-0.11	-0.08	-0.06	-0.03	-0.05	-0.01	0.26	0.20	0.17	0.13	0.10	0.08	0.05	0.03	0.03	0.00
p-value	0.54	0.01	0.01	0.03	0.06	0.17	0.23	0.35	0.42	0.49	0.56	0.00	0.00	0.01	0.06	0.12	0.21	0.30	0.45	0.50	0.60
Significant Coef.	0.00	-0.25	-0.28	-0.25	-0.21	-0.19	-0.16	-0.13	-0.10	-0.12	-0.15	0.28	0.24	0.23	0.20	0.17	0.16	0.11	0.07	0.05	0.04
% Significant	0.08	0.64	0.71	0.61	0.61	0.46	0.39	0.28	0.19	0.09	0.14	0.89	0.79	0.69	0.57	0.52	0.39	0.32	0.22	0.18	0.15
R <sup>2</sup>	0.10																				
Threshold Variable: Trade Volume - Regime: 1																					
Coefficient	0.00	-0.09	-0.14	-0.10	-0.10	-0.08	-0.06	-0.04	-0.02	0.00	-0.01	0.17	0.14	0.12	0.10	0.08	0.06	0.04	0.02	0.00	0.01
p-value	0.47	0.02	0.01	0.07	0.07	0.14	0.28	0.29	0.39	0.47	0.51	0.00	0.00	0.01	0.05	0.07	0.18	0.24	0.33	0.39	0.41
Significant Coef.	0.00	-0.13	-0.20	-0.16	-0.14	-0.13	-0.11	-0.09	-0.06	0.00	0.00	0.19	0.16	0.15	0.13	0.11	0.10	0.08	0.05	0.02	0.00
% Significant	0.09	0.67	0.67	0.59	0.60	0.49	0.43	0.29	0.19	0.05	0.03	0.90	0.81	0.74	0.67	0.62	0.47	0.39	0.26	0.20	0.05
R <sup>2</sup>	0.08																				
Threshold Variable: Trade Volume - Regime: 2																					
Coefficient	0.00	-0.19	-0.22	-0.17	-0.14	-0.10	-0.07	-0.05	-0.02	-0.01	-0.01	0.31	0.22	0.18	0.14	0.11	0.08	0.05	0.02	0.01	0.01
p-value	0.55	0.00	0.00	0.02	0.06	0.12	0.20	0.26	0.31	0.38	0.47	0.00	0.00	0.01	0.05	0.12	0.16	0.30	0.42	0.55	0.61
Significant Coef.	0.00	-0.24	-0.30	-0.23	-0.20	-0.17	-0.13	-0.09	-0.05	-0.04	-0.06	0.36	0.28	0.24	0.19	0.17	0.13	0.09	0.06	0.04	0.21
% Significant	0.17	0.76	0.72	0.70	0.62	0.51	0.42	0.41	0.26	0.23	0.19	0.85	0.75	0.71	0.61	0.52	0.48	0.32	0.20	0.18	0.03
R <sup>2</sup>	0.13																				
Threshold Variable: Trade Volume - Regime: 3																					
Coefficient	0.00	-0.16	-0.20	-0.17	-0.14	-0.11	-0.08	-0.05	-0.03	-0.02	-0.04	0.26	0.19	0.17	0.13	0.11	0.08	0.05	0.02	0.02	0.01
p-value	0.55	0.01	0.01	0.05	0.09	0.18	0.24	0.32	0.37	0.35	0.50	0.00	0.00	0.02	0.05	0.13	0.25	0.33	0.38	0.43	0.49
Significant Coef.	0.00	-0.23	-0.27	-0.24	-0.23	-0.20	-0.17	-0.12	-0.07	-0.04	-0.01	0.29	0.25	0.23	0.20	0.19	0.16	0.12	0.06	0.06	0.04
% Significant	0.09	0.66	0.70	0.62	0.56	0.42	0.38	0.29	0.22	0.19	0.17	0.86	0.75	0.68	0.60	0.49	0.39	0.30	0.21	0.10	0.05
R <sup>2</sup>	0.10																				
Threshold Variable: Basis - Regime: 1																					
Coefficient	0.00	-0.16	-0.20	-0.16	-0.13	-0.10	-0.08	-0.05	-0.03	-0.03	-0.02	0.28	0.22	0.17	0.14	0.11	0.09	0.05	0.03	0.01	0.02
p-value	0.00	0.01	0.00	0.04	0.07	0.12	0.12	0.27	0.37	0.51	0.48	0.00	0.00	0.01	0.02	0.04	0.11	0.19	0.35	0.53	0.68
Significant Coef.	0.00	-0.23	-0.26	-0.23	-0.20	-0.17	-0.13	-0.11	-0.07	-0.11	-0.07	0.32	0.27	0.23	0.18	0.16	0.13	0.09	0.07	0.04	0.01
% Significant	0.85	0.67	0.73	0.61	0.57	0.51	0.51	0.36	0.28	0.14	0.11	0.83	0.77	0.71	0.67	0.63	0.55	0.46	0.36	0.29	0.18
R <sup>2</sup>	0.10																				
Threshold Variable: Basis - Regime: 2																					
Coefficient	0.00	-0.19	-0.22	-0.17	-0.15	-0.11	-0.08	-0.05	-0.03	-0.01	-0.01	0.32	0.23	0.18	0.15	0.12	0.09	0.06	0.03	0.03	0.01
p-value	0.29	0.00	0.00	0.00	0.01	0.03	0.09	0.21	0.28	0.33	0.39	0.00	0.00	0.00	0.01	0.02	0.02	0.10	0.22	0.35	0.46
Significant Coef.	0.00	-0.23	-0.24	-0.21	-0.18	-0.14	-0.13	-0.09	-0.06	-0.04	-0.02	0.33	0.24	0.20	0.17	0.14	0.12	0.09	0.06	0.05	0.09
% Significant	0.35	0.77	0.84	0.78	0.71	0.67	0.54	0.41	0.31	0.23	0.09	0.93	0.92	0.87	0.86	0.79	0.71	0.54	0.40	0.25	0.19
R <sup>2</sup>	0.09																				
Threshold Variable: Basis - Regime: 3																					
Coefficient	0.00	0.04	-0.03	-0.02	-0.02	-0.01	0.00	-0.01	-0.01	-0.01	-0.02	0.03	0.02	0.02	0.02	0.01	0.00	0.01	0.00	-0.01	0.01
p-value	0.00	0.09	0.31	0.41	0.44	0.47	0.44	0.41	0.39	0.36	0.40	0.31	0.37	0.45	0.50	0.50	0.49	0.51	0.51	0.48	0.52
Significant Coef.	0.00	0.07	-0.09	-0.05	-0.09	-0.05	0.01	-0.05	-0.01	-0.01	-0.02	0.07	0.08	0.08	0.07	0.08	0.02	0.03	-0.02	-0.01	0.01
% Significant	0.96	0.55	0.29	0.27	0.23	0.22	0.18	0.23	0.18	0.12	0.09	0.38	0.25	0.24	0.18	0.20	0.16	0.16	0.16	0.16	0.16
R <sup>2</sup>	0.06																				

This table reports the estimates and several statistics of the TRM model for the case of Futures market leading the Spot. Fields: Coefficient, Significant Coefficients and R<sup>2</sup> are the average values across the 208 weeks of our sample, while the field p-value is the median observation.

Table 7: Threshold Regression Model Estimates: Spot market leads Future

Statistic	Const.	Spot observations lag length										Futures observations lag length									
		1	2	3	4	5	6	7	8	9	10	1	2	3	4	5	6	7	8	9	10
		Threshold Variable: Volatility - Regime: 1										Threshold Variable: Volatility - Regime: 2									
Coefficient	0.00	0.62	0.44	0.31	0.21	0.17	0.09	0.04	0.05	0.02	0.01	-0.64	-0.53	-0.44	-0.36	-0.29	-0.22	-0.16	-0.08	-0.01	0.07
p-value	0.61	0.00	0.00	0.08	0.29	0.32	0.58	0.60	0.68	0.70	0.68	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.10	0.11	0.14
Significant Coef.	0.00	0.64	0.40	0.17	0.06	0.07	0.08	0.05	0.11	0.11	0.11	-0.67	-0.57	-0.48	-0.41	-0.33	-0.26	-0.20	-0.12	-0.03	-0.05
% Significant	0.04	0.91	0.60	0.41	0.29	0.15	0.04	0.04	0.04	0.02	0.02	0.93	0.90	0.88	0.86	0.83	0.78	0.70	0.48	0.26	0.04
R <sup>2</sup>	0.18																				
Coefficient	0.00	0.59	0.45	0.25	0.18	0.15	0.09	0.13	0.08	0.08	0.06	-0.58	-0.53	-0.43	-0.35	-0.27	-0.22	-0.15	-0.09	-0.03	0.03
p-value	0.58	0.00	0.00	0.06	0.28	0.22	0.46	0.45	0.23	0.37	0.52	0.67	0.00	0.00	0.00	0.00	0.01	0.04	0.16	0.27	0.38
Significant Coef.	0.00	0.70	0.56	0.27	0.22	0.19	0.05	0.03	0.02	0.02	0.02	0.02	-0.68	-0.63	-0.53	-0.45	-0.36	-0.32	-0.22	-0.16	-0.03
% Significant	0.07	0.79	0.44	0.27	0.15	0.05	0.04	0.12	0.08	0.08	0.05	0.82	0.79	0.77	0.72	0.69	0.61	0.54	0.32	0.11	0.11
R <sup>2</sup>	0.28																				
Coefficient	0.00	0.57	0.43	0.26	0.19	0.15	0.09	0.12	0.05	0.12	0.09	-0.57	-0.50	-0.41	-0.35	-0.27	-0.20	-0.14	-0.07	-0.01	0.06
p-value	0.64	0.00	0.06	0.13	0.13	0.27	0.28	0.34	0.40	0.46	0.52	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.12	0.23	0.33
Significant Coef.	0.00	0.64	0.53	0.27	0.30	0.16	0.10	0.02	0.02	0.02	0.01	-0.65	-0.57	-0.49	-0.42	-0.34	-0.27	-0.20	-0.13	-0.06	-0.01
% Significant	0.05	0.86	0.55	0.28	0.16	0.16	0.14	0.10	0.10	0.09	0.09	0.86	0.84	0.80	0.76	0.69	0.64	0.56	0.38	0.27	0.14
R <sup>2</sup>	0.14																				
Coefficient	0.00	0.61	0.45	0.27	0.30	0.15	0.07	0.04	0.07	0.04	0.04	-0.62	-0.52	-0.43	-0.35	-0.29	-0.22	-0.15	-0.08	-0.01	0.06
p-value	0.47	0.00	0.00	0.12	0.22	0.29	0.44	0.63	0.69	0.76	0.62	0.00	0.00	0.00	0.00	0.00	0.00	0.04	0.11	0.18	0.25
Significant Coef.	-0.00	0.63	0.50	0.15	0.07	0.05	0.06	0.12	0.12	0.12	0.11	-0.65	-0.56	-0.47	-0.39	-0.33	-0.26	-0.19	-0.12	-0.04	-0.03
% Significant	0.02	0.94	0.59	0.38	0.27	0.16	0.05	0.03	0.05	0.04	0.06	0.94	0.91	0.89	0.86	0.81	0.77	0.66	0.43	0.21	0.19

This table reports the estimates and several statistics of the TRM model for the case of Spot market leading the Future.

Fields: Coefficient, Significant Coefficients and  $R^2$  are the average values across the 208 weeks of our sample, while the field p-value is the median observation.

time series data, around 8%. For the volatility and volume cases it takes slightly higher values which seems reasonable, because as we mentioned before we expect participants' behaviour to change in very unstable or shallow markets. Interestingly, the result for the basis case is different: the  $R^2$  dumps out as the Basis increases. After examining our results week by week, we found that usually the smallest spread between the two assets appears near the expiry of the futures contract, so it seems logical to have higher  $R^2$  during that period as the two prices should theoretically converge to the same value. With respect to the coefficient estimates, the lagged spot observations are usually significant up to 4–5 lags and as it was expected they are all negative, otherwise market bubbles would have occurred. The majority of the futures observations are significant up to lag 6, so there exists a 6-minute lead-lag relationship. This finding is very close to the recent study of Tse and Chan (2009), who find such effects to exist up to 9 minutes. Also, we observe that when we use the basis as the regime switching variable, in Regime 3, most of the  $b$ 's are insignificant, which is consistent with our previous result of very low  $R^2$  in that case.

In the “Spot leading the Futures” case, the findings with respect to the constant coefficients are the same as in the previous case: almost all of them are statistically zero and those which are significant, are economically zero. As now the futures return is the dependent variable, then  $a$ 's (with respect to notation of Equation (4)) for the lagged futures returns are again negative. However, in this case the lead-lag effects are quite powerful usually up to two minutes only. After the 2<sup>nd</sup> lag, the p-values indicate insignificant coefficients and their loadings become very small. The only exception is found in the 3<sup>rd</sup> regime when the basis is used as threshold variable. In that case the lead-lag effects last more and the  $R^2$  is considerably higher. This is in accordance with our previous finding of almost no short-run dynamics in that regime, when we were examining the futures market leading the spot. Finally, there are cases where the lead-lag relationship lasts more time, but these are isolated cases corresponding to less than 5% of our sample (i.e. less than 10 weeks).

After estimating the model described by Equation (4) with 3 different regime switching variables the next step is to build the spread-trade strategies. As we found that the spot market leads the futures only by 2 minutes and each strategy needs (at least) one minute to be implemented, we base our strategies only on the futures market leadership. We are setting up 3 different strategies based on those estimates and we

Table 8: Descriptive statistics of portfolios' returns

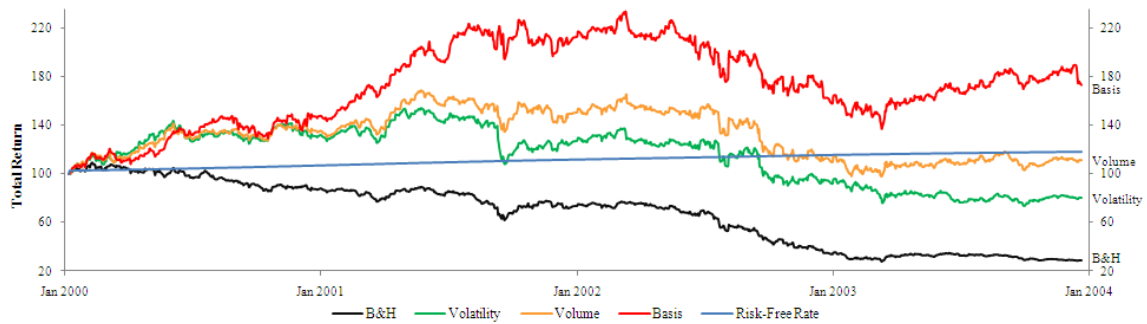
Statistic	Portfolio Strategy				Portfolio Strategy			
	B&H	Volatility	Volume	Basis	B&H	Volatility	Volume	Basis
Year 1					Year 2			
Daily Mean	-5.105	11.804	12.644	16.714	-5.766	-0.366	5.224	15.393
Annual Volatility	16.444	20.510	18.075	21.902	21.503	21.349	20.267	21.804
Skewness	0.329	0.141	0.163	0.222	-1.300	-1.350	-0.850	-0.810
Sharpe ratio	-1.036	1.306	1.667	1.971	-0.974	-0.397	0.313	1.788
Year 3					Year 4			
Daily Mean	-28.518	-10.964	-10.800	-9.806	-6.294	-5.560	0.427	3.568
Annual Volatility	27.226	27.865	27.170	28.628	21.884	20.945	21.080	21.628
Skewness	0.231	-0.133	0.679	0.641	0.506	0.344	0.461	0.292
Sharpe ratio	-2.115	-1.142	-1.154	-1.041	-0.968	-0.929	-0.267	0.110
Full Sample								
Daily Mean	-11.271	-1.151	2.014	6.622				
Annual Volatility	22.149	22.872	21.925	23.700				
Skewness	-0.104	-0.280	0.219	0.085				
Sharpe ratio	-1.008	-0.420	-0.085	0.579				

This table reports some descriptive statistics of the Buy and Hold portfolio and our 3 alternative strategies for each year separately and the entire sample.

Daily Mean is expressed in basis point, Volatility is presented in annual percentage points, on the basis of 250 Business days and Sharpe ratio is expressed in annual basis as well.

compare their results with a passive Buy & Hold index strategy (B&H hereafter). Figure 1 presents the Total Returns of the 3 + 1 portfolios and the suggested Risk-Free rate during the 4 years of our sample, while Table 8 presents some descriptive statistics of the portfolios.

Figure 1: Total Return Indices of the strategies



This graph illustrates the total performance of the Buy and Hold strategy and the 3 alternative ones across time.

As Risk-Free Rate the EONIA Total Return Index is used.

Interestingly, we observe the examined during that period the B&H portfolio loses 70.72% of its value over the last 4 years, while all of the alternative strategies exhibit overwhelming performance and two of them have gains, ranging from 2.32% to 18.37% per annum. Also, only one of the alternative strategies (the Basis) beats the risk-free target return for the entire period, however, in a year-by-year basis the evidence is mixed. All the dynamic strategies outperform the overnight rate during the first two years and only the Basis is profitable during the last two. With respect to the risk of the alternative strategies, their volatility is very close to the B&H strategy (less than  $\pm 2\%$  difference) all the time. Also, the skewness of the distribution of the returns is moving in the same direction across the years for all the strategies, which was expected, nonetheless it fluctuates less for the alternative strategies. Furthermore, excluding the year 2001, dynamic portfolios have almost all the times positively skewed returns, which is a point of superiority of those strategies. Finally, the ranking order suggested by the Sharpe ratio of the full sample is: Basis, Volume, Volatility based strategies and last the passive one. Exactly the same result is reached with year-over-year investigation, with the only exception of the raking reversal of the Volume and Volatility strategies in the 3<sup>rd</sup> year.

To provide additional insight about the profitability of the dynamic strategies and the appropriateness of the TRM we record how many trades occur in each regime for every threshold variable, how many of them are profitable and the cumulative return generated by those trades in the entire sample. Table 9 reports these findings and we observe that the majority of the trades is done in the extreme regimes (i.e. the 1<sup>st</sup> and the 3<sup>rd</sup>). Approximately, one out of two trades were profitable and this result is consistent across all regimes and threshold variables. Nevertheless for the Volatility and the Volume cases, transactions in the extreme regimes are slightly more frequently successful (3% – 6% of the times) rather than in the middle regime, while the opposite holds for the Basis case. Also, regarding the returns, in every case the 1<sup>st</sup> or the 3<sup>rd</sup> regime are by far more profitable than the middle regime and we believe this is another fact supporting the use of the TRM to capture the lead-lag effects between the two markets. Considerably fewer trades and smaller returns occur in the 3<sup>rd</sup> regime when the Basis is used as the Threshold variable, which is consistent with the more loose dynamics between the two markets which were reported earlier.

We proceed with evaluating the strategies' performance based on economic cri-

Table 9: Spread–trades decomposition

Regimes	Volatility			Volume			Basis		
	Trades	Profitable	Return	Trades	Profitable	Return	Trades	Profitable	Return
1 <sup>st</sup>	10.908	54.94%	36.54%	7.903	52.52%	19.20%	19.589	51.89%	72.12%
2 <sup>nd</sup>	8.531	48.06%	−1.15%	8.012	46.71%	23.02%	4.627	53.04%	19.28%
3 <sup>rd</sup>	12.256	51.27%	16.88%	16.062	52.52%	32.26%	3.810	51.21%	0.29%
Total	31.705	51.29%	57.78%	31.977	49.60%	94.45%	28.026	51.99%	105.89%

This table reports the break–down of the executed trades in each regime for every threshold variable used.

Column Profitable is the percentage of the trades executed in each regime which have gained profits and column Returns shows the total return of the trades executed in that regime.

teria. Table 10 presents the daily performance fees that a risk averse investor with degree of relative risk aversion,  $\gamma$ , equal to 2 or 6 is willing to pay to switch from the passive B&H strategy to each of the dynamic alternatives. Again, we calculate the same figures for both the entire sample and yearly sub–samples, in an attempt to identify how the strategies perform not only on average but also across the years. With respect to the entire sample’s results, we are expecting positive  $\Phi$ ’s when the benchmark strategy is the B&H because of the significant overperformance of the dynamic strategies and  $\Phi$ ’s negative or at best close to zero when the benchmark portfolio is the EONIA Total Return Index mainly because of the excess volatility of the strategies. This intuition is confirmed by the data and the performance fees vary from 9.68 to 17.48 bps on a daily basis with the B&H as benchmark, depending on the degree of relative risk aversion and the followed dynamic strategy. We mention that these results appear very supportive of the alternative strategies merely because the benchmark portfolio has suffered important losses and not because each of the alternative strategies has actually created impressive utility benefits for the investor. Year–by–year results support that view. During the first half of the first year the dynamic portfolios outperform the benchmark by about 25% - 30% and the 1<sup>st</sup> year’s fees are around 15 bps as the underlying index was in good levels as well. However, the pattern is the same at the 3<sup>rd</sup> year when the index loses more than 50% of its value and although each alternative strategy loses approximately 20%, that year the fees jump to 18 bps per day, merely because the dynamic strategies losses are smaller than those of the passive portfolio’s. Respectively, the performance fees when using the EONIA as target return are almost all the time negative, varying from -1.323 bps to -21.863, which means that the dynamic strategies do not add value to the investor,

Table 10: Performance fees for the dynamic strategies

DRA: $\gamma$	Volatility	Volume	Basis	Volatility	Volume	Basis
<b>Benchmark: B&amp;H Portfolio</b>						
Year 1			Year 2			
2	16.206	17.437	20.331	5.477	11.246	21.283
6	14.998	16.972	18.706	5.418	11.617	20.940
Year 3			Year 4			
2	17.405	17.732	18.404	0.914	6.913	9.954
6	17.093	17.759	17.801	1.300	7.298	10.072
Full Sample						
2	9.971	13.315	17.486			
6	9.685	13.403	16.877			
<b>Benchmark: EONIA Total Return Portfolio</b>						
Year 1			Year 2			
2	8.816	10.048	12.940	-4.133	1.637	11.674
6	5.475	7.455	9.180	-7.711	-1.507	7.808
Year 3			Year 4			
2	-15.709	-15.382	-14.712	-7.306	-1.307	1.734
6	-21.863	-21.184	-21.166	-10.799	-4.801	-2.030
Full Sample						
2	-4.667	-1.323	2.847			
6	-8.827	-5.102	-1.640			

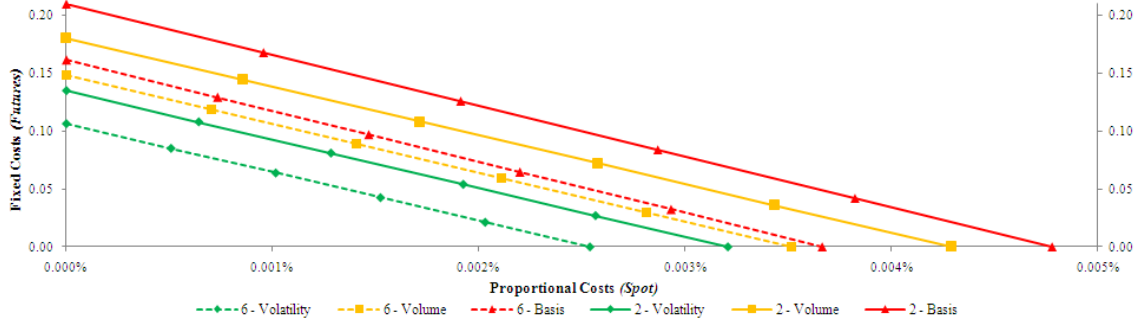
This table presents the daily performance fees ( $\Phi$ ), in basis points, which a risk averse investor would be willing to pay in order to switch from the benchmark portfolio to the dynamic one. We have used 2 different degrees of relative risk aversion ( $\gamma = 2$  or 6) and two different benchmark portfolios (the passive B&H and the EONIA Total Return Index).

in terms of utility, as he should be paid to undertake that strategy. However, the results are mixed when we examine them on a yearly basis. The 1<sup>st</sup> year and in half of the cases the 2<sup>nd</sup> year, are economically valuable for the investor with the positive fees ranging from 5.47 to 12.94 bps. On the other hand, the losses of the 3<sup>rd</sup> year and the excess volatility of the 4<sup>th</sup> are significant enough to dilute the benefits of the entire sample, as the negative fees are consistently close to -15 bps and only one case generated positive fees of 1.73 bps. Finally, we mention that the case of the 3<sup>rd</sup> year supports our previous comment about the very high performance fees when the benchmark is the B&H portfolio, which had very big losses.

Finally, we calculate the possible combinations of execution costs which render the investor indifferent, in terms of utilities, between each dynamic strategy and the benchmark. Figure 2 presents our results. Again, we assume two different risk averse investors, with degrees of relative risk aversion 2 and 6 respectively. We mention that our purpose is to investigate how the profitability of each strategy is affected by



Figure 2: Break-even transaction costs frontiers



This graph illustrates the bundles of fixed and proportional transaction costs which equate the utilities of each dynamic strategy with that of the simple B&H portfolio, for two different degrees of risk aversion  $\gamma = 2$  (solid line) and 6 (dotted line).

transaction costs and therefore we employ only the B&H portfolio as a benchmark and not the EONIA index as before. As anticipated, higher transaction costs can be applied to more profitable strategies (graphically, the frontier of the Basis strategy is “higher” than the frontier of the Volume strategy) and vice versa. Also, break-even transaction costs fall as the degree of risk aversion increases. With respect to the magnitude of our results, we observe that even the highest possible transaction costs are 0.23 index points (or €1.15) per futures contract trade and 4.7 bps per Euro of trading volume in the spot market. These fees are significantly smaller than what an institutional investor faces (about €3 and 10bps for the futures and spot trades respectively) so even the most profitable of our alternative strategies (i.e. the Basis) becomes loss making when realistic transaction costs are applied to it. Although our fees look very small to have such a big impact on the strategy’s profitability, they become rational when we recall that we have approximately 30 full trades per day, so we are rebalancing the portfolios 60 times. This is also the reason why, our findings are not directly comparable with the majority of previous studies (e.g. Tsiakas (2009)) which calculate break-even transaction costs and rebalance their portfolios once per day.

## 6. Summary – Conclusions

Using intradaily data of the DAX spot and futures indices, for the period 1/1/2000 - 31/12/2003, we investigated the existence of lead-lag effects between those markets. Firstly, we checked whether the dynamics are piecewise linear, for 3 different threshold variables, in order to examine if a TRM is more appropriate than a linear model. The Tsay (1998) test suggested that there are non-linear dynamics in all the 208 weeks of our sample in 10% statistical significance level and this result was quite robust in higher confidence intervals as well. Then, using the Bai and Perron (2003) algorithm we determined the exact break points for each week. The main finding was that in every sub-sample the 3-regime case is the most appropriate. We, subsequently, estimated the TRMs with lag length for each market varying from 3 to 12 (we started from 3 because we need at least that time for our strategies and we stopped at 12 because our preliminary analysis has shown that lead-lag effects occur in our sample up to that time). With respect to the case of the Futures market leading the Spot, we found significant lead-lag relationship up to 6 minutes on average and the result were similar for all the regimes and all the threshold variables. Regarding the Spot leading the Futures case, we found that there exist statistically significant short-run dynamics for up to 2 minutes. We proceeded with the implementation of our strategies, which were based only on the Futures market leadership, because it was more evident and we evaluated their performance. It was surpassing that all the alternative strategies outperformed the passive B&H portfolio and one achieved to beat the EONIA Total Return Index in the long run. Additionally, we evaluated those strategies with the traditional Sharpe ratio and with the Economic criteria methodology of Fleming *et al.* (2001) and the general results remained the same. There were only minor differences in the ranking of the strategies on a yearly basis. Finally, we calculated the frontiers of fixed and proportional break even transaction costs and we found that when realistic fees are applied all the profits are diluted. To sum up, our main finding is that the futures market of DAX index was consistently leading the spot and there where proper dynamic spread-trade strategies which could outperform the index (gross of transaction costs) both statistically and economically, while there was no profitability with realistic fees.

There are many directions in which this empirical analysis can be improved. We

just mention some of them: transactions data can be used for the spot time series and they would probably provide better insight about the dynamics of the two markets in the very short-run. Additional threshold variables can be used, probably to identify market wide information. In another version, we can model not only the conditional mean, but the conditional variance relation as well. Also, we believe it would be interesting to try to model the way investors decide to initiate a trade when they consider the power of the lead-lag effects (i.e. not only for how many minutes they last, but the levels of the coefficients as well) combined with the “neutral return band” which is caused by transaction costs. And finally, a more wide sample with some bullish periods in it can be used, in an attempt to find if such strategies are actually generating profits for the investor and not simply avoiding losses.

## APPENDICES

### A.1 *Estimation of trimmed sample mean and standard deviation*

Let  $X_1, X_2, \dots, X_N$  be a random sample and let  $X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(N)}$  be the observations written in ascending order. Suppose the desired percentage of trimming has been chosen to be  $\delta$ , where  $0 \leq \delta < 0.5$  and let  $d$  be the value of  $\delta \times N$  rounded down to the nearest integer.

Then, the  $\delta$ -trimmed mean is calculated excluding the smallest and the biggest  $d$  values, as:

$$\bar{X}_\delta = \frac{1}{N - 2d} \sum_{i=d+1}^{N-d} X_i \quad (11)$$

Respectively, to calculate the  $\delta$ -trimmed standard deviation, we have to remove from the sample the smallest and biggest  $d$  observations and to replace them with the numbers next in line for trimming.

As an example, suppose we need to calculate the 5%-trimmed standard deviation of the ordered sample:  $X_{(1)}, X_{(2)}, \dots, X_{(10)}$ .

We proceed as follows: we replace the 1<sup>st</sup> and the 10<sup>th</sup> observations with the 2<sup>nd</sup> and 9<sup>th</sup> respectively (as  $2 \times 5\% \times 10 = 1$ ).

Now we have the new set  $X_{(1)}, X_{(1)}, X_{(3)}, \dots, X_{(9)}, X_{(9)}$  for which we calculate the standard deviation using the traditional formula:

$$\sigma_X = \sqrt{\frac{1}{N - 2} \sum_{i=1}^N (X_i - \bar{X})^2} \quad (12)$$

where  $\bar{X}$  is the arithmetic mean of the new set.

### A.2 *Non-linearity test of Tsay*

To test for a linear model between the variables, versus the alternative of non-linear in time dynamics, Tsay (1998) suggested a statistic based on the predictive residuals of a re-arranged regression using the recursive least squares methodology.

Consider the model:

$$\mathbf{y}_t = \mathbf{X}_t \boldsymbol{\xi} + \boldsymbol{\varepsilon}_t, \quad i = 1, \dots, T \quad (13)$$

where:  $\mathbf{y}$  is the  $(T \times 1)$  vector of the dependent variable,  $\mathbf{X}$  is the  $(T \times 2p + 1)$  matrix of the regressors (because we use  $p$  lagged values of  $y$ 's and  $x$ 's respectively and a constant coefficient),  $\boldsymbol{\xi}$  is the  $((2p + 1) \times 1)$  vector of estimated parameters, and  $\boldsymbol{\varepsilon}$  is the vector of the residuals. If there is a model change in the above system then the recursive least squares estimators are biased, so the predictive residuals are not white noise and Tsay (1998) has shown that testing for model changes can be done by checking whether those residuals are correlated with the regressors or not.

To perform the test, the first step is to sort our observations based on the value of the exogenously given threshold variable,  $z_t$ . The key issue here is that even in the re-arranged form the dynamics between our variables are still the same. Now the initial regression can be represented as:

$$\mathbf{y}_{t(i)+d} = \mathbf{X}_{t(i)+d} \boldsymbol{\xi} + \boldsymbol{\varepsilon}_{t(i)+d}, \quad i = h + 1, \dots, T \quad (14)$$

where:  $t_{(i)}$  is the time index of the  $i^{th}$  element of the re-ordered observations based on  $z_t$  and  $h = \max(p, d)$ . In our case  $d$  is equal to one, because we assume that each investor needs 1 minute to rebalance her portfolio.

The next step is to calculate the predictive residuals from the aforementioned equation. Let  $\hat{\boldsymbol{\xi}}$  be the estimate of  $\boldsymbol{\xi}$ , then they are:

$$\hat{\boldsymbol{\varepsilon}}_{t(m+1)+d} = \mathbf{y}_{t(m+1)+d} - \mathbf{X}_{t(m+1)+d} \hat{\boldsymbol{\xi}}_m \quad (15)$$

where the  $m$  smallest values of  $z$  are used. The initial value  $m_0$  for the recursive least squares is set to  $3\sqrt{T}$ , rounded to the closest integer, where  $T$  is the full sample size. After that, we standardise the residuals and we get:

$$\hat{\boldsymbol{\eta}}_{t(m+1)+d} = \frac{\hat{\boldsymbol{\varepsilon}}_{t(m+1)+d}}{\sqrt{[1 + \mathbf{X}'_{t(m+1)+d} (\sum_{i=1}^m \mathbf{X}_{t(i)+d} \mathbf{X}'_{t(i)+d}) \mathbf{X}_{t(m+1)+d}]}} \quad (16)$$

Finally, we estimate the auxiliary regression:

$$\hat{\boldsymbol{\eta}}_{t(l)+d} = \mathbf{X}_{t(l)+d}\boldsymbol{\psi} + \mathbf{w}_{t(l)+d}, \quad l = m_0 + 1, \dots, T - h \quad (17)$$

where:  $\boldsymbol{\psi}$  is the vector of the estimated coefficients and  $\mathbf{w}$  are the residuals.

As mentioned before, under the Null hypothesis of liner relationship between the variables the standardised residuals should be a white noise and hence uncorrelated with the explanatory variables. So, with respect to notation in equation (17) the hypotheses can be expressed as:

$$H_0 : \boldsymbol{\psi} = 0 \quad vs \quad H_1 : \boldsymbol{\psi} \neq 0$$

To perform the previous test, Tsay (1998) proposed a statistic which, in our case of a single regressor with  $p$  lags, asymptotically follows a chi-squared distribution with  $2p + 1$  degrees of freedom. The exact formula for the statistic is:

$$C = [T - h - m_0 - (2p + 1)] \{ \ln(|S_0|) - \ln(|S_1|) \} \quad (18)$$

where:

$$S_0 = \frac{1}{T - h - m_0} \sum_{l=m_0+1}^{T-h} \hat{\boldsymbol{\eta}}'_{t(l)+d} \hat{\boldsymbol{\eta}}_{t(l)+d} \quad (19)$$

and

$$S_1 = \frac{1}{T - h - m_0} \sum_{l=m_0+1}^{T-h} \hat{\mathbf{w}}'_{t(l)+d} \hat{\mathbf{w}}_{t(l)+d}. \quad (20)$$

## References

- ABHYANKAR, A. H. (1995): "Return and Volatility Dynamics in the FTSE-100 Stock Index and Stock Index Futures Markets," *Journal of Futures Markets*, 15, 457–488.
- ADMATI, A., AND P. PFLEIDERER (1988): "A Theory of Intraday Patterns: Volume and Price Variability," *Review of Financial Studies*, 1, 3–40.
- AKAIKE, H. (1973): "Information Theory and an Extension of Maximum Likelihood Principle," *Second International Symposium on Information Theory*, Akademiai Kiado, 267–281.
- ANDERSEN, T., AND T. BOLLERSLEV (1997): "Intraday Periodicity and Volatility Persistence in Financial Markets," *Journal of Empirical Finance*, 4, 114–158.
- BAI, J., AND P. PERRON (2003): "Computation and Analysis of Multiple Structural Change Models," *Journal of Applied Econometrics*, 18, 1–22.
- BELLMAN, R., AND R. ROTH (1969): "Curve Fitting by Segmented Straight Lines," *Journal of the American Statistical Association*, 64, 1079–1084.
- BOLLEN, N. P., AND J. A. BUSSE (2004): "Short-Term Persistence in Mutual Fund Performance," *Review of Financial Studies*, 18, 569–597.
- BOOTH, G., R. SO, AND Y. TSE (1999): "Price Discovery in the German Equity Index Derivatives Markets," *Journal of Futures Markets*, 19, 619–643.
- BRENNAN, M., AND E. SCHWARTZ (1990): "Arbitrage in Stock Index Futures," *The Journal of Business*, 63, S7–S31.
- BREUSCH, T. (1978): "Testing for Autocorrelation in Dynamic Linear Models," *Australian Economic Papers*, 17, 334–355.
- BROWNLEES, C., AND G. GALLO (2006): "Financial Econometric Analysis at Ultra-High Frequency: Data Handling Concerns," *Computational Statistics & Data Analysis*, 51, 2232–2245.
- CHAN, K. (1992): "A Further Analysis of the Lead-Lag Relationship Between the Cash Market and Stock Index Futures Market," *Review of Financial Studies*, 5, 123–152.
- CHAN, K., AND H. TONG (1990): "On Likelihood Ratio Tests for Threshold Autoregression," *Journal of the Royal Statistical Society*, 52, 469–476.

- CHEN, N.-F., C. CUNY, AND R. HAUGEN (1995): "Stock Volatility and the Levels of the Basis and Open Interest in Futures Contracts," *Journal of Finance*, 50, 281–300.
- DACOROGNA, M., R. GENCAI, A. MULLER, R. OLSEN, AND V. PICTET (2001): "An Introduction to High Frequency Finance," *Academic Press*.
- DELLACORTE, P., L. SARNO, AND I. TSIAKAS (2008): "An Economic Evaluation of Empirical Exchange Rate Models," *Review of Financial Studies*, Forthcoming.
- (2009): "Correlation Timing in Asset Allocation with Bayesian Learning," *Working Paper*.
- DWYER, G. J., P. LOCKE, AND W. YU (1996): "Index Arbitrage and Nonlinear Dynamics Between the S&P 500 Futures and Cash," *Review of Financial Studies*, 9, 301–332.
- EASLEY, D., AND M. O'HARA (1992): "Time and the Process of Security Price Adjustment," *Journal of Finance*, 47, 577–605.
- ENGLE, R. F., AND J. R. RUSSELL (1998): "Autoregressive Conditional Duration: A New Model for Irregularly Spaced Transaction Data," *Econometrica*, 66, 1127–1162.
- FALKENBERRY, N. (2002): "High Frequency Data Filtering," *Tick Data*, Technical Report.
- FLEMING, J., C. KIRBY, AND B. OSTDIEK (2001): "The Economic Value of Volatility Timing," *Journal of Finance*, 56, 329–352.
- FORBES, C. S., G. R. J. KALB, AND P. KOFMAN (1997): "Bayesian Arbitrage Threshold Analysis," *Journal of Business & Economic Statistics*, 17, 364–372.
- GHOSH, A. (1992): "Cointegration and Error Correction Models: Intertemporal Causality Between Index and Futures Prices," *Journal of Futures Markets*, 13, 193–198.
- GODFREY, L. (1978): "Testing Against General Autoregressive and Moving Average Error Models when the Regressors Include Lagged Dependent Variables," *Econometrica*, 46, 1293–1302.
- GOETZMANN, W., J. INGERSOLL, AND M. SPIEGEL (2007): "Portfolio Performance Manipulation and Manipulation-Proof Performance Measures," *Review of Financial Studies*, 20, 1503–1546.



- GOETZMANN, W. N., J. E. INGERSOLL JR., M. I. SPIEGEL, AND I. WELCH (2002): “Sharpening Sharpe Ratios,” *NBER Working Paper*.
- GRUNBICHLER, A., F. A. LONGSTAFF, AND E. S. SCHWARTZ (1995): “Electronic Screen Trading and the Transmission of Information: An Empirical Examination,” *Journal of Financial Intermediation*, 3, 166–187.
- HAN, Y. (2006): “Asset Allocation with a High Dimensional Latent Factor Stochastic Volatility Model,” *Review of Financial Studies*, 19, 237–271.
- HARRIS, L. (1986a): “A Transaction Data Study of Weekly and Intradaily Patterns in Stock Returns,” *Journal of Financial Economics*, 16, 99–117.
- HLAWITSCHKA, W. (1994): “The Empirical Nature of Taylor–Series Approximations to Expected Utility,” *The American Economic Review*, 84, 713–719.
- JENSEN, M. C. (1968): “The Performance of Mutual Funds in the Period 1945–1964,” *Journal of Finance*, 23, 389–416.
- JIANG, L., J. K. FUNG, AND L. T. CHENG (2001): “The Lead–Lag Relation Between Spot and Futures Markets Under Different Short–Selling Regimes,” *Financial Review*, 36, 63–88.
- KAWALLER, I. G., P. D. KOCH, AND T. W. KOCH (1987): “The Temporal Price Relationship Between S&P 500 Futures and S&P 500 Index,” *Journal of Finance*, 42, 1309–1329.
- KYLE, A. S. (1985): “Continuous Auctions and Insider Trading,” *Econometrica*, 53, 1315–1335.
- MARKOWITZ, H. (1952): “Portfolio Selection,” *Journal of Finance*, 7, 77–91.
- MARQUERING, W., AND M. VERBEEK (2004): “The Economic Value of Predicting Stock Index Returns and Volatility,” *Journal of Financial and Quantitative Analysis*, 39, 407–429.
- MIN, J. H., AND M. NAJAND (1999): “A Further Investigation of the Lead–Lag Relationship Between the Spot Market and Stock Index Futures: Early Evidence from Kor,” *Journal of Futures Markets*, 19, 217–232.
- MULLER, U. (2001): “The Olsen Filter for Data in Finance,” *Olsen & Associates Technical Notes*.

- MULLER, U. A., M. M. DACOROGNA, R. B. OLSEN, O. V. PICTET, M. SCHWARZ, AND C. MORGENEGG (1990): "Statistical Study of Foreign Exchange Rates, Empirical Evidence of a Price Change Scaling Law and Intraday Analysis," *Journal of Banking & Finance*, 14, 1189–1208.
- NEELY, C. J., AND P. A. WELLER (2003): "Intraday Technical Trading in the Foreign Exchange Market," *Journal of International Money and Finance*, 22, 223–237.
- OOMEN, R. C. A. (2006): "Properties of Realized Variance Under Alternative Sampling Schemes," *Journal of Business & Economic Statistics*.
- PESARAN, M. H., AND A. TIMMERMANN (1995): "Predictability of Stock Returns: Robustness and Economic Significance," *Journal of Finance*, 50, 1201–1228.
- RIME, D., L. SARNO, AND E. SOJLI (2009): "Exchange Rate Forecasting, Order Flow and Macroeconomic Information," *Journal of International Economics*, Forthcoming.
- SHARPE, W. F. (1966): "Mutual Fund Performance," *Journal of Business*, 39, 119–138.
- (1994): "The Sharpe Ratio," *Journal of Portfolio Management*, 21, 49–58.
- SHYY, G., V. VIJAYRAGHAVAN, AND B. SCOTT-QUINN (1996): "A Further Investigation of the Lead–Lag Relationship Between The Cash Market and Stock Index Futures Market with the Use of Bid/Ask Quotes: The Case of France," *Journal of Futures Markets*, 16, 405–420.
- SORTINO, F. A., R. VAN DER MEER, AND A. PLANTINGA (1999): "The Dutch Triangle," *Journal of Portfolio Management*, 26, 50–58.
- STOLL, H. R., AND R. E. WHALEY (1990a): "The Dynamics of Stock Index and Stock Index Futures Returns," *Journal of Financial and Quantitative Analysis*, 25, 441–468.
- TSAY, R. S. (1989): "Testing and Modeling Threshold Autoregressive Processes," *Journal of the American Statistical Association*, 84, 231–240.
- (1998): "Testing and Modeling Multivariate Threshold Models," *Journal of the American Statistical Association*, 84, 1188–1202.
- TSE, Y. (1999): "Price Discovery and Volatility Spillovers in the DJIA Index and Futures Markets," *Journal of Futures Markets*, 19, 911–9300.

- TSE, Y., AND W. CHAN (2009): “The Lead–Lag Relation Between the S&P 500 Spot and Futures Markets: An Intraday - Data Analysis Using a Threshold Regression Model,” *Japanese Economic Review*.
- TSIAKAS, I. (2009): *The Economic Gains of Trading Stocks around Holidays* Forthcoming.
- WASSERFALLEN, W., AND H. ZIMMERMANN (1985): “The Behavior of Intra–daily Exchange Rates,” *Journal of Banking & Finance*, 9, 55–72.
- WEST, K., H. EDISON, AND D. CHO (1992): “A Utility Based Comparison of Some Models of Exchange Rate Volatility,” *NBER*, Technical Paper No 128.