No. 2007/20

Electronic Trading Systems and Intraday Non-Linear Dynamics: An Examination of the FTSE 100 Cash and Futures Returns

Bea Canto and Roman Kräussl

Center for Financial Studies
an der Universität Frankfurt ▪ Mertonstraße 17-21 ▪ D-60325 Frankfurt am Main
Telefon: +49 (0)69 798-30050
Fax: +49 (0)69 798-30077 ▪ http://www.ifk-cfs.de ▪ E-Mail: info@ifk-cfs.de

Center for Financial Studies

The *Center for Financial Studies* is a nonprofit research organization, supported by an association of more than 120 banks, insurance companies, industrial corporations and public institutions. Established in 1968 and closely affiliated with the University of Frankfurt, it provides a strong link between the financial community and academia.

The CFS Working Paper Series presents the result of scientific research on selected topics in the field of money, banking and finance. The authors were either participants in the Center´s Research Fellow Program or members of one of the Center´s Research Projects.

If you would like to know more about the *Center for Financial Studies*, please let us know of your interest.

Prof. Dr. Jan Pieter Krahnen

Prof. Volker Wieland, Ph.D.

Center for Financial Studies
an der Universität Frankfurt ▪ Mertonstraße 17-21 ▪ D-60325 Frankfurt am Main
Telefon: +49 (0)69 798-30050
Fax: +49 (0)69 798-30077 ▪ http://www.ifk-cfs.de ▪ E-Mail: info@ifk-cfs.de

CFS Working Paper No. 2007/20

Electronic Trading Systems and Intraday Non-Linear Dynamics: An Examination of the FTSE 100 Cash and Futures Returns*

Bea Canto[1] and Roman Kräussl[2]

March 15, 2007

Abstract:
This paper focuses on dynamic interactions of equity prices among theoretically related assets. We explore the existence of intraday non-linearities in the FTSE 100 cash and futures indices. We test whether the introduction of the electronic trading systems in the London Stock Exchange in 1997 and in the London International Financial Futures and Options Exchange (LIFFE) in 1999 has eliminated the non-linear dynamic relationship in the FTSE 100 markets. We show that the introduction of the electronic trading systems in the FTSE 100 markets has increased the efficiency of the markets by enhancing the price discovery process, namely by facilitating the increase of the speed of adjustment of the futures and cash prices to departures of the mispricing error from its non-arbitrage band. Nevertheless, we conclude that the automation of the markets has not completely eliminated the non-linear properties of the FTSE 100 cash and futures return series.

JEL Classification: G12, G14, G15

Keywords: Intraday non-linearities, Dynamic Spillovers, Electronic Trading Systems, Price Discovery Process, Cost of Carry Model, Regime Switching Model, Vector Error Correction Mechanism, SETAR Model

* We are grateful to Dick van Dijk, Charles Goodhart, Richard Payne, Mark Salmon, and seminar participants at London School of Economics, Financial Markets Group, Tinbergen Institute, and Bank of England.

1 Bea Canto, Watson Wyatt Brans & Co., Postbus 75201, 1070 AE Amsterdam, The Netherlands, Tel.: +31 20, 5433000, Fax: +31 20 5433995, email: Bea.Canto@watsonwyatt.nl

2 Corresponding author: Roman Kräussl, Free University of Amsterdam and Center for Financial Studies, Frankfurt/Main. Address at FU: Department of Finance and Financial Sector Management, FEWEB, de Boelelaan, 1105, 1081 HV Amsterdam, The Netherlands, Tel.: +31 20 5986102, Fax: +31 20 5986020, email: rkraeussl@feweb.vu.nl

I. Introduction

This paper focuses on dynamic interactions of equity prices among theoretically related assets. We analyze the dynamic spillovers between the FTSE 100 futures and cash indices and investigate the effects of arbitrage activity on shaping the observed dynamic interactions. In particular, we explore the existence of intraday non-linearities in the FTSE 100 cash and futures indices during the month of July 2001 using minute-by- minute data. We test if the introduction of the electronic trading systems in the London Stock Exchange in 1997 and in the London International Financial Futures and Options Exchange (LIFFE) in 1999 has eliminated the non-linear dynamic relationship in the FTSE 100 markets. To this aim, we use a regime switching model that allows the interactions to behave differently according to whether arbitrage opportunities are present in the market. Since the introduction of the screen trading in both exchanges, to our knowledge no study has analyzed the non-linear dynamics of the FTSE 100 index and futures returns.

The cost of carry model is often assumed to describe the non-arbitrage relation between the futures and index prices (see Brenner and Kronner (1995) as well as Dwyer, Locke and Yu (1996)). From a theoretical perspective, transaction costs and arbitrage activity in stock markets motivate the use of non-linear specifications to model the lead-lag relationship between an stock index and its futures markets. However, in the last years, the introduction of electronic trading systems to replace the traditional floor trading in many markets has significantly reduced the transaction costs and has accelerated the price discovery process in these markets.[4] As a consequence, we expect screen trading to have importantly reduced or even eliminated the non-linear dynamics between stock and futures returns. In the case of Australia, Anderson and Vahid (2001) find strong evidence of non-linearities in returns before the electronic trading

[4] See Grunbilchler, Longstaff and Schwartz (1994), Franses, Lucas, Taylor and van Dijk (2000) and Anderson and Vahid (2001) for studies on the ways in which the introduction of electronic trading affected the lead-lag relationship between futures and cash prices.

in the futures market was introduced and weaker evidence of non-linearities after the online trading. Their analysis suggests that the automation of the markets has removed the non-linear properties of the basis.

Grünbichler et al. (1994) extensively examine the effect of electronic screen trading on the lead-lag relation between futures and index levels. They highlight that the introduction of electronic trading lowers the trading costs for market participants. They also point out that price information is captured and disseminated more rapidly with screen trading, which accelerates the price discovery process. More recent studies also examine the effects of electronic trading in different markets. For instance, Hasbrouck (2003) analyses the effect of the introduction of the electronically-traded futures contracts in the U.S. equity indexes on price formation.

With respect to the dynamic interactions between the FTSE 100 stock index and its futures contracts, Abhyankar (1998) provides an extensive survey of the empirical evidence on the lead-lag relationship between cash and futures prices. Additionally, several studies document the lead-lag relationship in the British context, for instance Gwilym, McMillan and Speight (1999) and Gwilym and Buckle (2001), but little research has been conducted on examining non-linearities in the U.K. markets. As our findings will show that non-linearities are important in explaining the short-term dynamics between the FTSE 100 futures and the cash index, the former studies fail to capture the effects of the arbitrage activity in these markets.

To our knowledge two studies elaborates on the non-linear intraday dynamics in the FTSE 100 markets using regime-switching models: Garrett and Taylor (2001) and Franses et al. (2000). Both studies find strong evidence of non-linearities in the U.K. markets. Garrett and Taylor (2001) examine the intraday and interday dynamics of both the level of and changes in the FTSE 100 basis. In particular, they investigate if the first-order autocorrelation in basis changes is a result of arbitrage behavior or a manifestation of market microstructure effects such as non-trading in the underlying stock index. In their analysis, they apply a Self Exciting Threshold Autoregressive model (SETAR) to the mispricing. Our paper also analyses the dynamics of the basis using a SETAR specification. We extend the analysis of Garrett and Taylor (2001) analysis since we additionally focus on the effects of the arbitrage opportunities on the futures and stock index returns dynamics.

Franses et al. (2000) examine the impact of the introduction of the electronic trading system in the London Stock Exchange on stock price dynamics. They find strong evidence of non-linearities before the introduction of the electronic trading system and much weaker evidence of non-linearities with on-line trading. They suggest that the automation of markets may remove the non-linear properties of the basis.

Our paper builds upon this last point. We investigate the existence of non-linearities in electronically trading markets. In particular, we extend Franses et al. (2000) analysis to examine the non-linear dynamic behavior of the FTSE 100 index and its futures. They explore the non-linear dynamic relationship in the U.K. markets in 1997, at the time of the introduction of the electronic trading platform in the London Stock Exchange. After the introduction of the automated trading system in the LIFFE exchange in 1999, we expect that the transaction costs faced by investors in the British markets are even lower. An interesting unanswered question that we investigate in this paper is whether this further reduction in transaction costs has eliminated the non-linear dynamics between the FTSE 100 cash and futures returns.

This paper has two main contributions. First, as mentioned before, this is the first study that presents a discrete regime-switching model to analyze the index arbitrage in the FTSE 100 markets after the introduction of electronic trading platform in its futures market. Second, from an econometric perspective, this study generalizes previous models as we use an integrated approach suggested by Tsay (1998) in which the threshold values that define the different regimes are endogenously determined in the model.

Our results show that arbitrage activity is of some significance in shaping the short-term dynamic relationship between the FTSE 100 cash and futures prices. The empirical evidence confirms the presence of non-linearities in the behavior of the basis and the returns when using one-minute frequency data. We conclude that the introduction of the electronic trading systems in the FTSE 100 markets has increased the efficiency of the markets by enhancing the price discovery process, namely by facilitating the increase of the speed of adjustment of the futures and cash prices to departures of the mispricing error from its non-arbitrage band. Nevertheless, the automation of the markets has not completely eliminated the non-linear properties of the return series.

The remainder of this paper is organized as follows. Section II introduces the cost of carry model and describes the econometric model. Section III provides details on the dataset and descriptive statistics. Section IV contains the empirical results of the non-linearity tests for the basis and the returns. It also presents the estimation of the non-linear model for the basis, elaborates upon the results of the Threshold Error Correction Model and extends the analysis using different frequency sub-samples. Concluding remarks are given in Section V.

II. Methodology

II.1 Cost of Carry Model with Transaction Costs

According to the cost of carry model, the basis or the mispricing error is defined as

$$z_t = \ln F_{t,T} - \ln S_t - (r_{t,T} - q_{t,T})(T - t) \tag{1}$$

where $F_{t,T}$ is the price at time t of a future contract with maturity T. S_t is the index value in period t, $r_{t,T}$

stands for the risk free interest rate for the period $T\text{-}t$ and $q_{t,T}$ is the dividend yield on the index.

The introduction of transaction costs in the cost of carry model provides the motivation for the non-linear behavior of the basis. Transaction costs include the bid-offer spread, stamp duty, market commissions and any impact costs which reflect the size of the trade and the liquidity of the markets. For arbitrage to be profitable in equation (1), the basis z_t must be sufficiently large to offset the transaction costs. We therefore propose to use a Self Exciting Threshold Autoregressive framework (SETAR) to model the behavior of the basis with three different regimes. This specification reflects that arbitragers react to a large enough negative mispricing error that was observed d periods in advance, $z_{t-d} \leq c_1$, as well as to a large positive mispricing error, $z_{t-d} > c_2$. In these regimes the deviations of the basis from zero are big enough to offset the transaction costs, c_1 and c_2. When the deviations of the basis are smaller than the transaction costs, $c_2 < z_{t-d} \leq c_1$, there are no arbitrage opportunities. With the above considerations, the SETAR specification for the basis can be written in three different regimes as

$$
\begin{aligned}
z_t &= \delta^{(1)} + \sum_{i=1}^{I} \alpha_i^{(1)} z_{t-i} + \xi_t^{(1)} \text{ if } z_{t-d} \leq c_1 \\
z_t &= \delta^{(2)} + \sum_{i=1}^{I} \alpha_i^{(2)} z_{t-i} + \xi_t^{(2)} \text{ if } c_1 < z_{t-d} \leq c_2 \\
z_t &= \delta^{(3)} + \sum_{i=1}^{I} \alpha_i^{(3)} z_{t-i} + \xi_t^{(3)} \text{ if } z_{t-d} > c_2
\end{aligned}
\tag{2}
$$

where c_1 and c_2 are the threshold values for the variable z_{t-d} that define the regime switching. We examine the hypothesis that, because of arbitrage, any mean reversion in the basis is stronger in regimes one and three than in the middle regime, i.e., $\alpha_i^{(1)} << \alpha_i^{(2)}$ and $\alpha_i^{(3)} << \alpha_i^{(2)}$.

The arbitrage trade in regime 1 consists of simultaneously buying index futures and short-selling the security index while an arbitrage trade in regime 3 consists of simultaneously buying the security index and selling the index futures. We specify the threshold variable as z_{t-d} instead of z_t because it takes time (minutes) for arbitragers to take appropriate positions in the stock and stock index futures contracts. Consequently, we do not expect arbitrage to occur and affect the futures and the stock index in the same minute as when the arbitrage opportunity appears. This threshold lag, d, gives an indication of the speed at which the market responds to deviations from the no-arbitrage relation. As previously mentioned, c_1 and c_2 are endogenously determined in the model using Tsay (1998) technique.

III.2 Econometric Model

The cointegration relation between the futures and the cash indices documented in the empirical literature implies that an Error Correction Mechanism characterizes the relationship between them (Engle and Granger (1987)). In our case, equation (2) suggests three regimes to characterize the dynamic relationship between the FTSE 100 index and its futures contracts. If arbitrage activity affects the size of the responses of the futures and index levels to lagged variables and their adjustment process to the long-term equilibrium, the values of the parameters in the Error Correction Model will depend on the regimes. Together, the cointegration, the arbitrage opportunities and the transaction costs suggest a Threshold Vector Error Correction Mechanism (TVECM) to model the dynamics of the cash and futures. This means that current futures and index returns are explained by past futures, past index returns and by the deviation from the no-arbitrage relation d periods in advance. The effects of lagged variables, as well as the effect of the mispricing error are in our specification different for each regime. The VECM for each of the three regimes, j, is specified as

$$\begin{aligned}
\Delta \ln F_{t,T} &= \phi_{10}^{(j)} + \sum_{i=1}^{p} \phi_{11,i}^{(j)} \Delta \ln F_{t-i,T} + \sum_{i=1}^{p} \phi_{12,i}^{(j)} \Delta \ln S_{t-i} + \beta_1^{(j)} z_{t-d} + \varepsilon_{1,t}^{(j)} \\
\Delta \ln S_t &= \phi_{20}^{(j)} + \sum_{i=1}^{p} \phi_{21,i}^{(j)} \Delta \ln F_{t-i,T} + \sum_{i=1}^{p} \phi_{22,i}^{(j)} \Delta \ln S_{t-i} + \beta_2^{(j)} z_{t-d} + \varepsilon_{2,t}^{(j)}
\end{aligned} \qquad (3)$$

where Δ is the difference operator, i.e., $\Delta X_t = X_t - X_{t-1}$, $\beta_1^{(j)}$ and $\beta_2^{(j)}$ are the error correction coefficients and $\varepsilon_t = (\varepsilon_{1,t}, \varepsilon_{2,t})$ are zero mean and serially uncorrelated error terms that can be contemporaneously correlated. As in equation (2), the regimes are determined by

$$\begin{aligned} j=1 &\quad \text{if} \quad z_{t-d} \le c_1 \\ j=2 &\quad \text{if} \quad c_1 < z_{t-d} \le c_2 \\ j=3 &\quad \text{if} \quad z_{t-d} > c_2 \end{aligned} \qquad (4)$$

In this specification, the parameters of the ECM depend on the level of mispricing. The thresholds are signals for index arbitrage. To test if regime 2 reflects the non-arbitrage band, we can test if the effects of the correction term in this regime are smaller than in the outer regimes. Thus, in equations (3) and (4) we test $\left|\beta_1^{(1)}\right| > \left|\beta_1^{(2)}\right|$ and $\left|\beta_1^{(3)}\right| > \left|\beta_1^{(2)}\right|$ for the futures equation, and $\left|\beta_2^{(1)}\right| > \left|\beta_2^{(2)}\right|$ and $\left|\beta_2^{(3)}\right| > \left|\beta_2^{(2)}\right|$ for the cash equation. In addition, note that there can be differences in the impact of arbitragers in the lower and upper regimes as the arbitrage strategies are different in both regimes.

III. Data and Descriptive Statistics

The empirical analysis is based on the FTSE 100 stock index. The FTSE 100 index comprises the 100 largest U.K. companies listed on the London Stock Exchange (LSE). The LSE trades between 08:00 am and 16:30h (London time) from Monday to Friday (excluding the public holidays). Stock trading has been fully automated since 1997, when the LSE introduced its electronic trading system (SETS). SETS enables traders to place buy and sell orders for any of the FTSE 100 shares in an electronic order book. These orders are then automatically matched with other orders placed. The futures contracts on the FTSE indices are traded in the London International Financial Futures and Options Exchange (LIFFE). The LIFFE Connect is the automated trading system in the derivatives exchange and was introduced in May 1999. This electronic trading platform also matches orders, disseminates prices and reports trades. Trading in the stock index futures occurs between 08:00 and 17:30h.

The sample period used in this study covers the month of July 2001. The index data are intraday minute-by-minute snapshots of the FTSE 100 index values obtained from the LIFFE Exchange. The FTSE 100 index value is updated approximately four times a minute. The data is converted to one observation per minute by using the last observation for each minute. Our futures data correspond to the transaction prices of the FTSE 100 futures maturing on 21 September 2001.

The overlapping trading hours for both markets are between 08:00 and 16:30h. However, to avoid anomalies related to stale cash prices at the beginning of the trading day, the first thirty minutes of each day are discarded. Using the remaining observations, the one minute returns for each market are calculated as the difference of the natural log of the prices, i.e., the futures returns equal to $\Delta \ln F_{t,T} = \ln F_{t,T} - \ln F_{t-1,T}$ and the index level returns equal to $\Delta \ln S_t = \ln S_t - \ln S_{t-1}$. This results in 478 (or less when the trading starts after 08:00h or finishes earlier) returns per day. When stacking several days, overnight returns are removed. Each of our data series contains 10,470 observations.

We follow Dwyer et al. (1996) to calculate the cost of carry. First, we subtract daily means from the logarithms of the futures and cash indexes. Demeaning the futures removes any constant in the logarithms of the futures due to the constant part of dividends and interest rates for that day. The difference between the demeaned logarithms of the futures and cash indexes is the deviation of the basis from its daily mean. If dividends and interest rates are relatively constant during the day, this adjusted basis is an estimate of a mispricing series that does not require other explicit assumptions about expected dividend or interest rates.

This point is vital as the validity of the mispricing series relies heavily on the use of appropriate ex-ante dividends and interest rates.[5]

It is useful to examine the properties of the basis and the returns prior to modeling their dynamics. Some summary statistics are provided in Table I. We observe that the futures returns are more volatile and have a higher average than the cash returns. There is evidence of positive first-order autocorrelation in index returns. As demonstrated by Lo and McKinley (1990), this pattern occurs if stocks trade infrequently. The futures returns exhibit negative first order autocorrelation. A likely explanation is that transaction prices bounce between the bid and ask levels (see Glosten and Milgrom (1985)).

PLEASE INSERT TABLE I ABOUT HERE

The mispricing changes also exhibit negative first-order autocorrelation. Taking into account the 'infrequent trading' effect and the 'bid-ask bounce' effect, Miller, Mutshuswamy and Whaley (1994) analytically demonstrate that negative first-order autocorrelation in mispricing changes is likely to occur under quite general conditions. Table I also shows that the basis is more volatile than the futures and cash returns. Additionally, Figure I presents the time series plots of one-minute returns of the FTSE 100 futures, the index values and the associated basis. We observe that all the series fluctuate around a fixed mean and within a fixed range.

PLEASE INSERT FIGURE I ABOUT HERE

[5] An alternative would be to use the actual dividends yield on the FTSE 100 index reported by FT Interactive Data. However, they are realized dividends, not expected dividends. Therefore, we prefer to substract the daily means from the series. Henceforth, the mispricing error will be denoted by z_t and we will present the values of the basis as 100 * z_t for notational reasons.

To test for non-stationarity, Augmented Dickey Fuller (ADF) tests are performed on the one-minute frequency log price series and on the basis. The results of the tests are given in Table II. Panel A shows that both the futures and cash prices have a unit root, while the returns on these assets are stationary. However, the null hypothesis of a unit root can be rejected at the 1 percent level of significance for the basis equation. This means that the basis is a stationary process rather than a random walk. Miller et al. (1994) argue that in the absence of arbitrage activity, if index levels and futures prices follow a random walk, then the basis should follow a random walk as well. By contrast, if arbitragers exist in the market, then mispricing will be removed within a very short period of time. Consequently, the basis will follow a mean reverting process. The test results in Panel A show that the basis follows a stationary process which indicates that arbitrage activity is of some significance in the FTSE 100 markets.

PLEASE INSERT TABLE II ABOUT HERE

Possible cointegration between these prices is investigated by applying the Johansen Cointegration test to the futures and index price series. The results of the test are presented in Panel B. The first part of the table presents the results of the cointegration test between the futures price and the index value. The second part of the table reports the cointegration test between the futures price and the theoretical futures price, i.e., the futures price implied by the cost of carry model. The results in both parts indicate the existence of one cointegration equation at the 5 percent significance level. This means that the futures and the index price, adjusted for the cost of carry and without adjusting for it, are cointegrated.

Given that the Johansen Cointegration test does not reject the existence of one cointegration equation, the last row of each part of the table presents the stationary linear combination that exists between the futures and the index prices, namely, the cointegration relation or the Error Correction term. These results indicate that there is some evidence that the cointegrating vector is not strictly (1,-1). However, if we restrict the vector according to (1,-1) we still find evidence of cointegration. To facilitate the interpretation, we will use the (1,-1) vector as the cointegration vector and hence, the mispricing error as defined in equation (1).[6]

IV. Discussion of Empirical Results

IV.1 Non-linearity Test for the Basis and for the Returns

We start by testing the SETAR behavior for the basis z_t. We examine the hypothesis that the basis follows a linear AR(I) process against the alternative hypothesis that the basis follows a non-linear model. We start selecting the AR order I for the basis. Following Martens et al. (1998), we use the partial autocorrelation function of z_t and we choose the lag order for the basis $I=4$. Next, we choose the set D of possible threshold lags. We assume that $d \in D$ can be chosen by practical experience. The electronic trading system in the LSE allows the possibility of simultaneous trading in both index and futures markets. Therefore, we expect the arbitrage opportunities to be observed almost immediately and we use $d \in \{1,2,3,4,5\}$.[7]

Table III presents the results of the test statistic *C(d)*.[8] We test the null hypothesis that the basis follows a linear AR(4) process, so that the model in equation (2) reduces to a univariate model. The test statistic

---

[6] The qualitative results of our estimations are robust to the use of (1,-1) as the cointegration vector or to the use of (1,-α), where alpha is in this case the cointegration coefficient.

[7] Notice that minute-by-minute transaction prices are used. $d \in \{1,2,3,4,5\}$ indicates that any arbitrage trading order is executed within five minutes.

[8] *C(d)* is the test statistic to test for non-linearity. Tsay (1998) defines the test statistic and demonstrates that *C(d)* is asymptotically a chi-square random variable.

follows an asymptotic chi-square distribution with 5 degrees of freedom. The p-values of the test-statistic are also presented in the table. The recursive estimation starts with $m_0 = 250$, which is about $2.5\sqrt{10,470}$.[9] The results of the tests in Panel A indicate that p-values are close to zero for the threshold lags $d = 1, 2$ and 3 and thus, the tests reject linearity for these lags. Moreover, the maximum value of the test statistic corresponds to $d = 1$, indicating that 1 is the optimal delay for the threshold variable. These results point out that a SETAR model like the one suggested in equation (2) is a sensible representation of the behavior of the basis.[10]

PLEASE INSERT TABLE III ABOUT HERE

Non-linearities in the basis require a TVECM to model the behavior of the futures and index returns. As a consequence, when applying the linearity test to the system $y_t = \{\Delta \ln F_{t,T}, \Delta \ln S_t\}$, we expect that the test rejects linearity and that the threshold variable is the same as the one found in the previous subsection for the basis, i.e., $d = 1$.

For the linear Error Correction representation, we choose a lag-length $p = 9$ based on the significant coefficients at the 10 percent level. This long lag structure provides a broader picture of the lead-lag relationship between the futures and the index returns. As in the previous subsection, $d \in \{1,2,3,4,5\}$ is used as the possible set of values for d. Panel B in Table III presents the test results of the multivariate

---

[9] The choice of m_0 is explained in Tsay (1998): Small m_0 may introduce bias in the empirical distribution of $C(d)$. He suggests a starting value for the recursive autoregression around $2.5\sqrt{N}$, where N is the total number of observations.

[10] We want to mention that the results of an ARCH test performed on the residuals from the estimated models indicated that there is significant heteroskedasticity present. Therefore, White heteroskedasticity consistent standard errors are presented in the estimations and in the tests of this analysis.

linearity test applied to the futures and index returns. The null hypothesis is that the return series are linear, so that model in equations (3) and (4) reduces to a bivariate linear Vector Error Correction model. The alternative hypothesis is that the return series present non-linear patterns. The test statistic *C(d)* is carried out with $p = 9$ and $d \in \{1,2,3,4,5\}$ and follows a chi-square distribution with 40 degrees of freedom. The results of the test reject linearity more clearly for the returns' system than for the basis equation. Consequently, our results point to a non-linear specification for the behavior of the futures and index returns. Furthermore, the test statistic *C(d)* reaches its maximum value when $d = 1$, which also confirms that the optimal threshold variable is z_{t-1}.

IV.2 The Dynamics of the Basis

Next we estimate the implied SETAR model for the basis described in equation (2) with three different regimes. Given the complicated nature of the non-linearity, we use a two stage estimation process. The first stage involves a grid search to locate the threshold values c_1 and c_2; in a second stage, we estimate the implied SETAR model taken c_1 and c_2 as fixed parameters in the estimation.

Based on the empirical range of z_{t-1}, we assume that the candidates for the threshold values are $c_1 \in$ [-0.078, -0.041] and $c_2 \in$ [0.038, 0.082].[12] The minimum value for c_1 and the maximum value for c_2 are chosen such that there are at least 500 observations, approximately 5 percent of the total observations, included in the outer regimes. Using a grid search of 300 points on each of the intervals, the minimum Akaike Information Criterion (AIC) selects $\hat{c}_1 = -0.060$ and $\hat{c}_2 = -0.049$, which correspond to the

[12] Note that the selection of *I* and *d* beforehand dramatically reduces the state space of the grid search to choose c_1 and c_2.

values that trigger the arbitrage. Such values leave 1,003 observations in the lower regime, 7,600 observations in the middle regime and 1,867 observations in the upper regime. The minimum AIC is −166,087.

Our optimal threshold values indicate that the non-arbitrage range lies between −6.0 and 4.9 basis points. These estimated values of the transaction costs are very low compared to the results of previous studies.[13] Several points are worth noting. First, the small magnitude of the transaction costs is consistent with the fact that the electronic trading system has significantly reduced the magnitude of the transaction costs that investors face. Second, as the FTSE 100 markets are among the most liquid markets in Europe, we do not expect to find large bid-ask spreads in these markets.[14] Finally, the mispricing estimates of Deutsche Bank[15] for the month of July 2001 range between −1.3 and 12.7 basis points, which also points to very small deviations of the basis from its equilibrium value.

We turn next to present the estimates of the SETAR model for the basis as stated in system (2). Table IV displays the results of the AR(4) estimation for each regime. The results show strong support for the notion that the basis follows a different process depending on whether arbitrage opportunities are present. The estimates of the coefficients $\alpha_1^{(j)}$ corresponding to z_{t-1} are 0.490, 0.615 and 0.436 for regimes j = 1, 2 and 3, respectively: the further the mispricing is away from the equilibrium, the stronger is the reversion

[13] For instance, Garrett and Taylor (2001) analyze FTSE 100 data from the period January to April 1998 and find that the symmetric transaction costs for the markets during 12:00 to 16:00h is 26.23 basis points.

[14] The fact that institutional investors trade within the spread and they do not pay stamp duty justifies the small magnitude of the threshold values.

[15] Deutsche Bank Derivatives Research Group produces a daily Global Fair Value sheet for European Futures; see: "Deutsche Bank Portfolio, Index and Futures Research".

back to the equilibrium.[16] This fitted model confirms the expectations that z_t has stronger mean-reverting tendency in the outer regimes, where arbitrage is presumably possible. This result indicates that, as soon as arbitrage opportunities are observable, the arbitragers enter the market to take advantage of such opportunities. These empirical findings show that the U.K. markets respond to deviations from the non-arbitrage relation in just a few minutes.

PLEASE INSERT TABLE IV ABOUT HERE

Dwyer et al. (1996) argue that it is possible for the basis to be mean-reverting outside the arbitrage bounds but not within them. We analyze this proposition more carefully using a Dickey Fulley type regression test applied to the middle regime sub-sample. In particular, we run the following regression for the subsample of regime 2 and test if $\beta_1 = 0$.

$$\Delta z_t = \beta_0 + \beta_1 z_{t-1} + u_t$$

The results of this regression are reported in Table V. The test statistic of β_1 with 0.281 is different from zero, which implies that the basis is also mean-reverting in regime 2 and, thus, the mispricing does not persist indefinitely in this regime.

PLEASE INSERT TABLE V ABOUT HERE

IV.3 Non-linear Impulse Response Functions for the Basis

---

[16] The changes in the dynamic pattern of z{t} are robust to different threshold values in the neighborhood of $\hat{c}_1$ and $\hat{c}_2$.

To further evaluate the dynamic properties of the estimated regime switching model for the basis, we analyze its Impulse Response Functions. These functions examine the effects of shocks ξ_t on the evolution of the time series z_t as defined in system (2).[17] The Generalized Impulse Response Functions are illustrated in Figure II. A shock of size ±1 percent and ±2 percent is introduced in date $t = 0$. The graphs are just a representative example of many possible impulse response functions depending on the history. Panel A plots the impulse response function after a shock in regime 1. Panel B depicts the response after a shock in regime 2 and Panel C draws the adjustment path after a shock in regime 3. Even though the effects of all shocks almost disappear within ten minutes of the introduction of the shock, we observe that the degree of persistence of the shocks is higher in regime 2, within the non-arbitrage band, than in regimes 1 and 3. This result confirms the finding that the further the mispricing error is away from its equilibrium, the stronger is the reversion back to its equilibrium due to the activity of the arbitrageurs.

PLEASE INSERT FIGURE II ABOUT HERE

Panels A, B and C of Figure II visualize that the system remains in the same regime after a shock. This is not the case in Panel D, where an example of non-linear behavior is illustrated. The negative shock implies a switch in regime, in particular, it moves the system from regime 3 into regime 2. Thus, the Generalized Impulse Response Function is also affected by the difference between the parameter estimates in regimes 3 and 2 explaining the rapid increase to zero and negative values after the shock.

---

[17] As noted by Koop, Pesaran and Potter (1996), non-linear models produce impulse response functions that depend on the sign and size of the shock, as well as on the history of the time series. They introduce the Generalized Impulse Response Function (GIRF) which provides a solution to the problems involved in defining impulse responses in nonlinear models. The GIRF for an arbitrary impulse $\xi_t = \delta$ and a history w_{t-1} is defined as

$$GIRF_z(h, \delta, w_{t-1}) = E\left[z_{t+h} \mid \xi_t = \delta, w_{t-1}\right] - E\left[z_{t+h} \mid w_{t-1}\right] \qquad (5)$$

Overall, we can conclude that, even with a narrow arbitrage band, our SETAR estimates and the Impulse Response Functions support evidence of non-linearities in the dynamic behavior of the mispricing error.

IV.4 The Dynamics of the Futures and Cash Indices

In the following we estimate a Threshold Error Correction Mechanism (TVECM) to characterize the non-linear dynamic dependence between the FTSE 100 cash and futures returns described in equations (3) and (4). As in the previous section, we start with searching the threshold values. The threshold candidates are assumed to be in the intervals $c_1 \in$ [-0.078, -0.041] and $c_2 \in$ [0.038, 0.082]. Using a grid search of 300 points in these intervals, the minimum AIC provides $\hat{c}_1 = -0.057$ and $\hat{c}_2 = 0.059$, with the minimum AIC equal to −346,436. These values leave 1,134 observations in the lower regime, 7,844 observations in the middle regime and 1,491 observations in the upper regime. These selected optimal threshold values are consistent with those obtained for the basis.

Given z_{t-1} and the three regimes defined by $\hat{c}_1$ and $\hat{c}_2$, then we estimate the conditional Error Correction Model for each regime. The lag-length in each regime and for each equation is based on significant coefficients, at the 10 percent level, with a minimum of one lag. The results of the estimation are presented in Table VI. Panel A presents the coefficient estimates of the futures equation $\Delta \ln F_{t,T}$. Our empirical results, however, show different outcomes. First, the error correction coefficient is not significant in regimes 1 and 3. Furthermore, the futures returns do not depend on past futures returns in regimes 1 and 3 as the estimates of $\phi_{11}^{(1)}$ and $\phi_{11}^{(3)}$ are not statistically significant.

PLEASE INSERT TABLE VI ABOUT HERE

Panel B displays the coefficient estimates of the cash equation $\Delta \ln S_t$. The results show that the Error

Correction term is statistically significant in all the regimes, $\beta_2^{(j)} = 0.181$, 0.099 and 0.222 in regimes j = 1, 2 and 3, respectively. The magnitude of this coefficient is approximately twice as large in regimes 1 and 3 as in regime 2. This increase in the dependence on the Error Correction term on regimes 1 and 3 reflects that the index prices immediately react to departures of the mispricing error from its non-arbitrage band. In addition, we observe that the lag dependence of the cash returns to its own returns and to the futures returns tends to be lower in regime 2, $\phi_{21,i}^{(2)} < \phi_{21,i}^{(1)}, \phi_{21,i}^{(2)} < \phi_{21,i}^{(3)}$ and $\phi_{22,i}^{(2)} < \phi_{22,i}^{(1)}, \phi_{22,i}^{(2)} < \phi_{22,i}^{(3)}$. In particular, the coefficient $\phi_{21,i}^{(j)}$ corresponding to $\Delta \ln F_{t-1,T}$ increases from 0.191 in regime 2 to 0.273 and 0.269 in regimes 1 and 3, respectively. This evidence suggests that the cash index adjusts more quickly to the future market movements when arbitrage opportunities are available in the market.

Our empirical results point out that new information coming into the markets is first impounded in the futures prices. The futures market fixes the value of the mispricing error and the cash market adapts to the futures movements. In this sense, the lead-lag dependence between the FTSE 100 futures and cash markets is best described by the cash equation as described in Panel B.

In a final step, we want to describe the main common stylized facts across the regimes in Table VI to compare them with the empirical findings of previous linear studies of the lead-lag relationship between derivatives and cash markets in the U.K. First, not surprisingly the error correction term is negative in the futures equation and positive in the stocks equation, i.e., $\beta_1^{(j)} < 0$ and $\beta_2^{(j)} > 0$ for j = 1, 2 and 3. Only the estimates of the cash equation are statistically significant different from zero. This result indicates that the adjustment of the cash market to a mispricing disequilibrium is very rapid. Second, the index returns depend negatively on their own past returns and positively on the future returns, i.e., $\phi_{21,i}^{(j)} > 0$ and $\phi_{22,i}^{(j)} > 0$ for j = 1, 2 and 3. Third, it is apparent that the FTSE 100 futures market generally leads the cash market in all the regimes by 5 to 9 minutes, i.e., $\phi_{21,1}^{(j)}, ..., \phi_{21,5}^{(j)}, ..., \phi_{21,9}^{(2)}$ are statistically significant. Finally, the fitted equations perform better in the cash equation than in the futures equation as the larger adjusted R^2 indicates.

All these results are in line with previous linear studies on the relationship between the FTSE derivatives markets and the cash market; see for instance Gwilym and Buckle (2001) and Abhyankar (1995). All the studies on linear lead-lag relationship in the stock index futures markets state that the index futures returns generally lead the stock index returns with little or no feedback from the cash to the futures markets. A possible explanation for this finding is that informed traders are more likely to trade in stock index futures as a consequence of the leverage and transaction costs benefits offered by these markets and thus, price movements of stock index futures are likely to lead price index movements.[18] However, as the empirical linear studies do not take into account the transaction costs that define the different regimes, they fail to capture the different behavior of the dynamic relationship between the FTSE 100 futures and the cash market due to the arbitrage activity in the markets. Related to this last point, our results indicate that arbitrage activity is of some significance in the FTSE 100 markets.

IV.5 Robustness Analysis

The analysis presented so far has used one-minute frequency data. In the following we repeat the analysis using lower frequency data over the same sample period to assess if our results are robust to changes in the frequency of the data. In particular, we repeat the analysis with two- and five-minute frequency data over the same sample period.

To begin with, the Augmented Dickey Fuller (ADF) tests and the Johansen cointegration tests are

[18] Fleming, Ostdiek and Whaley (1996) demonstrate that the cost of taking a position in a stock index futures is considerably lower than the cost of taking an equivalent position in stocks.

performed on the new frequency series. The results of the tests and the cointegration equations are reported in Table VII. For both cases, the results of the tests are robust with those obtained using one-minute frequency data; namely, the futures and cash prices contain a unit root and both price series are cointegrated.

PLEASE INSERT TABLE VII ABOUT HERE

The second step is to calculate the non-linearity test *C(d)*. To make the analysis comparable with the one-minute frequency results, we set $d = \{1, 2\}$ when two-minute frequency data is used, which corresponds with actual delays of two and four minutes. In the same way, when we use five-minute frequency data, we set a delay parameter $d = \{1\}$, which is equivalent to a delay of five minutes. Table VIII reports the test statistic and the *p*-values. Several interesting features stand out from this table. First, with two-minute frequency data the test suggests threshold non-linearity in the basis series and the return series when $d = 1$ (*p*-values = 0.000). However, the test does not reject linearity in the basis series when $d = 2$ (*p*-value = 0.120). These results imply that the optimal delay for the threshold variable is $d = 1$. Second, with five-minute frequency data, the test statistics do not reject linearity (*p*-value = 0.107 and 0.378 for the basis and the returns, respectively). Third, these outcomes are robust with the test-statistics obtained in Table III using one-minute frequency data. In that case, the test did not reject linearity for the delays *d* equal to 4 and 5 minutes.

PLEASE INSERT TABLE VIII ABOUT HERE

We can conclude that the regime-switching models are not the appropriate specification to describe the dynamics of the FTSE 100 futures and cash returns when we work with five-minute frequency data, but they are appropriate when we analyze higher frequency returns dynamics. Accordingly, we suggest estimating a regime-switching model for the two-minute frequency data sample and a linear model when using the five-minute frequency subsample.

Next, we estimate the non-linear regime-switching models using the two-minute frequency data; namely a SETAR for the basis and a TVECM for the returns dynamics. The new dataset contains 5,124 observations. To make it consistent with the previous estimation, we choose the lag order of the SETAR model for the basis I = 2, which corresponds to four minutes and the lag order of the TVECM equal to ten minutes, p = 5. The candidates for the threshold values are also the same as the ones selected for the one-minute frequency analysis, i.e., $c_1 \in [-0.078, -0.041]$ and $c_2 \in [0.038, 0.082]$. The minimum AIC criterion for the SETAR selects $\hat{c}_1 = -0.051$ and $\hat{c}_2 = 0.045$, with the AIC value equal to −38,317. The AIC criterion for the TVECM selects $\hat{c}_1 = -0.058$ and $\hat{c}_2 = 0.049$, with the AIC value equal to −91,126. Table IX reports the estimated parameters of the SETAR model for the basis. The results are very similar to those obtained using one minute frequency data: the mean reversion of the basis is stronger in regimes 1 and 3 where arbitrage is presumably profitable, $\alpha_1^{(j)} = 0.377, 0.472$ and 0.439 in regimes j = 1, 2 and 3, respectively.

PLEASE INSERT TABLE IX ABOUT HERE

Table X displays the estimates of the TVECM for the returns system. We observe that the estimated coefficients of the Error Correction term in the futures equation are non-significant. On the opposite, the estimated coefficients $\beta_2^{(j)}$ in the cash equation point out that cash prices are the ones that react to any disequilibrium movements. This fact is especially remarkable in regime 1, where the error correction coefficient is more than four times larger than the one in regime 2, $\beta_2^{(j)} = 0.467, 0.111$ and 0.199 for j = 1, 2 and 3, respectively.

PLEASE INSERT TABLE X ABOUT HERE

Regarding the five-minute frequency data, we next present the linear AR(1) and the VECM(2) estimation. The new dataset contains 2,028 observations. As in the previous samples, we select the lag order to account for delays of up to ten minutes, in particular I = 1 and p = 2. The estimated AR(1) model for the basis is: $z_t = 2.6 \cdot 10^{-5,***} + 0.343^{***} z_{t-1} + \xi_t$ .The estimated results indicate that the mispricing error follows a stationary process. [19] Table XI presents the estimates of the linear VECM(2) for the FTSE 100 cash and futures returns. The empirical findings are in line with those of previous linear studies on the lead-lag relationship between futures and cash prices. The signs of the adjustment coefficients in the VECM are those expected and significantly different from zero, $\beta_1 = -0.189$ in the futures equation and $\beta_1 = 0.308$ in the cash equation. In addition, our results point out that the index futures returns lead the stock index returns, $\phi_{21}^1 = 0.300$ and $\phi_{21}^2 = 0.115$.

PLEASE INSERT TABLE XI ABOUT HERE

To summarize the findings in this section, the non-linear properties of the FTSE 100 cash and futures returns are not robust to changes in sample frequencies. Non-linearities are still present in the FTSE 100 markets when we work with frequencies higher than five minutes. This finding indicates that the introduction of screen trading has accelerated the price discovery process in the FTSE 100 markets, namely, the information is incorporated more rapidly into prices.

V. Conclusions

This paper has analyzed the dynamic interactions between stock prices that are theoretically related; in

[19] We also estimated the linear models with longer lag orders. The new coefficients turned out to be not significant.

particular, between futures and cash indices for the FTSE 100 using one minute frequency data. We have analyzed the role of transaction costs and arbitrage activity to explain the non-linear dynamics observed between these contracts. In addition, we have investigated whether the introduction of the electronic trading platforms has eliminated the non-linear dynamics in the FTSE 100 markets. To this aim, we suggested a discrete regime-switching framework to define the bands within which arbitrage may be profitable. First, we estimated a Self Exciting Vector Autoregressive for the basis. In a second step, we specified a Threshold Error Correction Model which explicitly modeled the behavior of arbitragers and allowed for non-linear adjustments of the returns towards the long-term equilibrium. Intuitively, index-futures arbitrage only occurs when the deviations from the non-arbitrage relationship are sufficiently large to compensate for the transaction costs. In this context, the TVECM provides the bands within which arbitrage is not profitable and the effects of arbitrage on the convergence of futures and cash values.

The main conclusion from our empirical investigation is that arbitrage activity is of some significance in shaping the short-term dynamic relationship between the FTSE 100 cash and futures prices. Our evidence confirms the presence of non-linearities in the behavior of the basis and the returns when using one minute frequency data. The main findings of this paper can be summarized as follows:

1. The basis or mispricing follows different processes depending on whether arbitrage opportunities are present. In particular, the mean reversion of the basis to the cost of carry in the regimes in which arbitrage is profitable is stronger than in the regime where there are no arbitrage opportunities.

2. As for the dynamic relationship between the futures and cash prices, our results show that the parameters of the Error Correction Mechanism depend on the level of mispricing. In particular, the adjustment process of the FTSE 100 cash index to deviations from the mispricing equilibrium exhibits clear non-linearities. New information coming into the market is first included in futures prices. The index market then responds to arbitrage opportunities pushing the mispricing error back to the non-arbitrage

band. This behavior is particularly strong in the arbitrage regime where the deviations of the basis are large and positive. In such situations, the arbitrage strategy consists of selling futures contracts on the FTSE 100 and simultaneously buying the stocks underlying the index.

3. We extended the analysis to assess whether our results are robust to changes in the frequency of the data. In particular, we repeated the analysis with two and five minute frequency data series over the same sample period. We find that the non-linear dynamic behavior is not robust to changes in data frequencies. When using five minute frequency data, the non-linearities are not present and thus, the regime-switching models are not an appropriate specification to model the lead-lag relationship between the FTSE 100 cash and futures indices, indicating that index arbitrage opportunities in the FTSE markets vanish within five minutes.

Overall, the introduction of the electronic trading systems in the FTSE 100 markets has increased the efficiency of the markets by enhancing the price discovery process, namely by facilitating the increase of the speed of adjustment of the futures and cash prices to departures of the mispricing error from its non-arbitrage band. Nevertheless, the automation of the markets has not completely eliminated the non-linear properties of the return series.

References

Abhyankar, A. (1995): “Returns and Volatility Dynamics in the FTSE Stock Index and Stock Index Futures Markets”, *Journal of Futures Markets*, 15(4), pp. 457-488.

Abhyankar, A. (1998): “Linear and Nonlinear Granger Causality: Evidence from the UK Stock Index Futures Market”, *Journal of Futures Markets*, 18(5), pp. 519-540.

Anderson, H., and F. Vahid (2001): “Market Architecture and Nonlinear Dynamics of Australian Stock and Futures Indices”, *Australian Economic Papers*, 40(4), pp. 541-566.

Balke, N., and T. Fomby (1997): “Threshold Cointegration”, *International Economic Review*, 38, pp. 627-646.

Balke, N., and M. Wohar (1998): “Nonlinear Dynamics and Covered Interest Rate Parity”, *Empirical Economics*, 23, pp. 535-559.

Baum, C., M. Karkuolas and M. Caglayan (2001): “Nonlinear Adjustment to Purchasing Power Parity in the Post-Bretton Woods Era”, *Journal of International Money and Finance*, 20, pp. 379-399.

Brenner, R., and K. Kroner (1995): “Arbitrage, Cointegration and Testing the Unbiasedness Hypothesis in Financial Markets”, *Journal of Quantitative and Financial Analysis*, 30, pp. 23-42.

Dwyer, G., P. Locke and W. Yu (1996): “Index Arbitrage and Nonlinear Dynamics between the S&P 500 Futures and Cash”, *Review of Financial Studies*, 9(1), pp. 301-332.

Engle, R., and C. Granger (1987): "Cointegration and Error-Correction Representation, Estimation and Testing", *Econometrica*, 55, pp. 251-276.

Flemming, J., B. Ostdiek and R. Whaley (1996): "Trading Costs and the Relative Rates of Price Discovery in Stocks, Futures and Options Markets", *Journal of Futures Markets*, 16, pp. 353-387.

Franses, P., A. Lucas, N. Taylor and D. van Dijk (2000): "SETS, Arbitrage Activity and Stock Price Dynamics", *Journal of Banking and Finance*, 24, pp. 1289-1306.

Garrett, I., and N. Taylor (2001): "Intraday and Interday Basis Dynamics: Evidence from the FTSE 100 Index Futures Market", *Studies in Nonlinear Dynamics and Econometrics*, 5(2), pp. 133-152.

Glosten, L., and P. Milgrom (1985): "Bid, Ask and Transaction Prices in a Specialist Market with Heterogeneously Informed Traders", *Journal of Financial Economics*, 13, pp. 71-100.

Grünbichler, A., F.A. Longstaff, and E.S. Schwartz (1994): "Electronic Screen Trading and the Transmission of Information: An Empirical Investigation", *Journal of Financial Intermediation*, 3, pp. 166-187.

Gwilym, O., D.G. McMillan, and A.E.H. Speight (1999): "The Intraday Relationship between Volume and Volatility in LIFFE Futures Markets", *Applied Financial Economics*, 9, pp. 593-604.

Hansen, B., and B. Seo (2001): "Testing for Two-Regime Threshold Cointegration in Vector Error Correction Models", *Journal of Econometrics*, 110(2), pp. 293-318.

Hasbrouck, J. (2003): "Intraday Price Formation in US Equity Index Markets", *Journal of Finance*, 58(6), pp. 2375-2400.

Kofman, P., M. Martens and T. Vorst (1998): "A Threshold Error Correction Model for Intraday Futures and Index Returns", *Journal of Applied Econometrics*, 13, pp. 245-263.

Koop, G., H. Pessaran and S. Potter (1996): "Impulse Response Analysis in Non-linear Multivariate Models", *Journal of Econometrics*, 74, pp. 119-147.

Lo, A.W., and A.C. MacKinley (1990): "An Econometric Analysis of Nonsynchronous Trading", *Journal of Econometrics*, 45, pp. 181-211.

Martens, M., P. Kofman, and T.C.F. Vorst (1998): "A Threshold Error-Correction Model for Intraday Futures and Index Returns", *Journal of Applied Econometrics*, 13(3), pp. 245-263.

Michael, P., R. Nobay and D. Peel (1997): "Transaction Costs and Nonlinear Adjustment in Real Exchange Rates: An Empirical Investigation", *Journal of Political Economy*, 105(4), pp. 862-879.

Miller, M., J. Muthuswamy and R. Whaley (1994): "Mean-Reversion of Standard & Poor's 500 Index Basis Changes: Arbitrage Induced or Statistical Illusion", *Journal of Finance*, 49, pp. 479-513.

Tsay, R. (1998): "Testing and Modeling Multivariate Threshold Models", *Journal of the American Statistical Association*, 93, pp. 1188-1202.

Yadav, P., P. Pope and K. Paudyal (1994): "Threshold Autoregressive Modeling in Finance: the rice Differences of Equivalent Assets", *Mathematical Finance*, 4(2), pp. 205-221.

Table I
Summary Statistics

Notes: The number of observations for each time series is 10,470. The basis is calculated according to equation (1) $z_t = \ln F_{t,T} - \ln S_t - (r_{t,T} - q_{t,T})(T - t)$. ρ_1 is the first order autocorrelation coefficient which is calculated for the first differences in the basis.

	$\ln F_{t,T}$	$\ln S_t$	$100*z_t$
Maximum	0.0025	0.0023	0.220
Minimum	-0.0027	-0.0016	-0.300
Mean	$1.1\cdot10^{-6}$	$3.7\cdot10^{-7}$	$2.7\cdot10^{-3}$
Median	0.000	0.000	$3.0\cdot10^{-3}$
Std. Dev.	$3.9\cdot10^{-4}$	$4.0\cdot10^{-4}$	$4.9\cdot10^{-4}$
ρ_1	-0.013	0.189	-0.499

Table II

Panel A: Augmented Dickey Fuller Unit Root Test

Notes: The unit root regressions for the futures and index prices contain a constant and 10 lags, while the unit root regression for the basis contains a constant and 4 lags.

	$\ln F_{t,T}$	$\ln S_t$	z_t	Critical Value 1 percent
ADF Levels	-1.17	-1.13	-18.4	-3.41
ADF Differences	-45.5	-30.1	-21.1	-3.41

Panel B: Johansen Cointegration Test

Notes: The test is carried assuming that the series have linear trends. λ_i refers to the Eigen values, the second column displays the Likelihood Ratio test statistic. For each part of the table, the first row tests the hypothesis of no cointegration, the second row tests the hypothesis of one cointegration relation, the third row presents the cointegration vector. Standard errors are reported in parentheses.

Cointegration between $\ln F_{t,T}$ and $\ln S_t$			
λ_i	Likelihood Ratio	Critical Value 5 percent	H_0
0.0159	172.7	15.41	$r = 1$
0.0002	3.147	3.76	$r \leq 1$
EC term:	$\ln F_{t,T} - 1.053 \ln S_t$ (0.003)	$\ln S_t - 0.949 \ln F_{t,T}$ (0.004)	
Cointegration between $\ln F_{t,T}$ and $\ln S_t^{(c)}$ adjusted for cost of carry			
λ_i	Likelihood Ratio	Critical Value 5 percent	H_0
0.0388	421.3	15.41	$r = 0$
0.0002	3.144	3.76	$r \leq 1$
EC term:	$\ln F_{t,T} - 1.017 \ln S_t^{(c)}$ (0.001)	$\ln S_t^{(c)} - 0.983 \ln F_{t,T}$ (0.001)	

Table III
Non-linearity Tests, *C(d)*

Notes: The sample size is 10,470 and the starting point of the recursive least squares is 250. The non-linearity tests present heterokedasticity consistent results.

Panel A: C(d) tests H_0: "z_t follows a linear AR(4)" against H_1: "z_t is non-linear"

$d =$	1	2	3	4	5
$C(d) \sim \chi^2(5)$	23.07	15.37	13.77	9.243	8.810
p – value	0.000	0.008	0.017	0.099	0.117

Panel B: C(d) tests H_0: "$y_t = \{\Delta \ln F_{t,T}, \Delta \ln S_t\}$ follows a linear VECM(9)" against H_1: "y_t is non-linear"

$d =$	1	2	3	4	5
$C(d) \sim \chi^2(40)$	94.21	74.40	64.90	44.13	46.96
p – value	0.000	0.000	0.008	0.301	0.209

Table IV
Self Exciting Threshold Autoregressive Model for the Basis

Notes: The model estimated is given in equation (2):

$$z_t = \delta^{(j)} + \sum_{i=1}^{I} \alpha_i^{(j)} z_{t-i} + \xi_t^{(j)} \qquad c_{j-1} < z_{t-1} \le c_j$$

where j = 1, 2, 3 and the threshold lag equals to 1. The optimal threshold values are $\hat{c}_1$ = -0.060 and $\hat{c}_2$ = 0.049, which define the three regimes. The number of observations is 1,003, 7,600 and 1,867 in regimes 1, 2 and 3, respectively. White heteroskedasticity consistent standard errors are given. $\hat{\sigma}^{2,(j)}$ is the sum of squared residuals in the regression. *; ** and *** indicate significance at 10, 5 and 1 percent levels, respectively.

	Regime 1		**Regime 2**		**Regime 3**	
	$z_{t-1} \le$ -0.060		-0.060 < $z_{t-1} \le$ 0.049		z_{t-1} > 0.049	
	Coef.	Std. Error	Coef.	Std. Error	Coef.	Std. Error
z_{t-1}	0.490***	0.066	0.615***	0.017	0.436***	0.046
z_{t-2}	0.046	0.044	0.090***	0.015	0.095***	0.032
z_{t-3}	0.104**	0.043	0.023	0.014	-0.013	0.031
z_{t-4}	0.113***	0.039	0.055***	0.012	0.107***	0.026
$\delta^{(j)}$	$-2.8\cdot10^{-5}$	$5.1\cdot10^{-5}$	$1.1\cdot10^{-5,***}$	$3.8\cdot10^{-6}$	$1.2\cdot10^{-4,***}$	$3.3\cdot10^{-5}$
$\hat{\sigma}^{2,(j)}$	$4.1\cdot10^{-4}$		$3.5\cdot10^{-4}$		$3.7\cdot10^{-4}$	
Adj. R^2 (%)	18.1		27.8		11.1	

Table V
Mean Reversion of the Basis in Regime 2

Notes: Dickey Fuller type regression applied to regime 2 with 7,845 observations. *** indicates significance at 1 percent level.

Independent variable: Δz_t		
	Coefficient	Std. Error
β_0	$1.2 \cdot 10^{-5}$ ***	$4.0 \cdot 10^{-6}$
β_1	-0.281***	-0.014

Table VI
Threshold Error Correction Model for the returns

Notes: The estimated TVECM is given in equations (3) and (4). The lag-length in each regime and for each equation is based on significant coefficients. Number of observations is 1,134, 7,845 and 1,491 in regimes 1, 2 and 3 respectively. White heteroskedasticity consistent standard errors are reported. *, ** and *** stand for significance at 10, 5 and 1 percent levels, respectively.

	Regime 1		**Regime 2**		**Regime 3**	
	$z_{t-} \leq -0.057$		$-0.057 < z_{t-1} \leq 0.059$		$z_{t-1} > 0.059$	
	Coef.	Std. Error	Coef.	Std. Error	Coef.	Std. Error
Panel A: Futures Equation, $\Delta\ln F_{t,T}$						
z_{t-1}	-0.042	0.070	-0.054***	0.015	-0.069	0.065
$\Delta\ln F_{t-1,T}$	0.013	0.050	-0.045***	0.015	-0.019	0.038
$\Delta\ln S_{t-1}$	0.153***	0.059	0.163***	0.018	0.115***	0.043
Constant	$3.6\cdot10^{-5}$	$5.3\cdot10^{-5}$	$6.1\cdot10^{-6}$	$4.2\cdot10^{-6}$	$-8.5\cdot10^{-5}$	$4.9\cdot10^{-5}$
Adj. R^2(%)	1.51		1.50		0.65	
Panel B: Cash Equation, $\Delta\ln S_t$						
z_{t-1}	0.181***	0.051	0.099***	0.012	0.222***	0.047
$\Delta\ln F_{t-1,T}$	0.273***	0.034	0.191***	0.012	0.269***	0.032
$\Delta\ln F_{t-2,T}$	0.187***	0.034	0.167***	0.012	0.188***	0.029
$\Delta\ln F_{t-3,T}$	0.081**	0.032	0.129***	0.011	0.116***	0.031
$\Delta\ln F_{t-4,T}$	0.050*	0.027	0.104***	0.011	0.118***	0.030
$\Delta\ln F_{t-5,T}$	0.094***	0.024	0.084***	0.011	0.092***	0.029
$\Delta\ln F_{t-6,T}$			0.068***	0.011	0.087***	0.027
$\Delta\ln F_{t-7,T}$			0.051***	0.010	0.053**	0.023
$\Delta\ln F_{t-8,T}$			0.032***	0.010	-0.137***	0.040
$\Delta\ln F_{t-9,T}$			0.025***	0.009	-0.048	0.041
$\Delta\ln S_{t-1}$	-0.099**	0.043	-0.094***	0.015	-0.089**	0.037
$\Delta\ln S_{t-2}$	-0.055	0.043	-0.105***	0.015	-0.113***	0.034
$\Delta\ln S_{t-3}$	-0.092**	0.045	-0.104***	0.014	-0.083**	0.033
$\Delta\ln S_{t-4}$			-0.086***	0.014	-0.052*	0.031
$\Delta\ln S_{t-5}$			-0.052***	0.014		
$\Delta\ln S_{t-6}$			-0.075***	0.013		
$\Delta\ln S_{t-7}$			-0.029**	0.012		
$\Delta\ln S_{t-8}$			-0.039***	0.012		
Constant	$5.0\cdot10^{-5}$	$3.9\cdot10^{-5}$	$-2.5\cdot10^{-6}$	$2.8\cdot10^{-6}$	$-1.0\cdot10^{-4,***}$	$3.2\cdot10^{-5}$
Adj. R^2 (%)	15.7		10.8		14.4	

Table VII
Unit Root Tests and Cointegration Tests – Two- and Five-Minute Frequency Series

Notes: Tests are applied to two- and five-minute frequency datasets. The 1 percent critical value of the ADF test is -3.43. The 5 percent critical values of the Likelihood Ratio test are 15.41 and 3.76 respectively. The unit root regressions for the futures and index prices contain a constant and 10 lags, while the unit root regression for the basis contains a constant and 4 lags. The test is carried assuming that the series have linear trends. λ_i refers to the Eigen values, the second column displays the Likelihood Ratio test statistic. For each part of the table, the first row tests the hypothesis of no cointegration, the second row tests the hypothesis of one cointegration relation, the third row presents the cointegration vector. Standard errors are reported in parentheses.

Panel A: **ADF unit root test on the prices**			
	$\ln F_{T,t}$	$\ln S_t$	$100^* z_t$
2-minute sample			
Levels	-1.80	-1.81	-20.9
Differences	-31.4	-30.6	-47.7
5-minute sample			
Levels	-1.74	-1.72	-15.1
Differences	-21.4	-21.9	-32.5
Panel B: **Johansen cointegration test**			
	λ_i	Likelihood Ratio	H_0
2-minute	0.0147	79.3	$r = 0$
	0.0006	3.27	$r<=1$
	EC term	$\ln F_{t,T} - \underset{(0.004)}{1.053} \ln S_t$	$\ln S_t - \underset{(0.004)}{0.949} \ln F_{t,T}$
5-minute	0.0146	32.8	$r = 0$
	0.0014	3.76	$r<=1$
	EC term	$\ln F_{t,T} - \underset{(0.007)}{1.052} \ln S_t$	$\ln S_t - \underset{(0.006)}{0.949} \ln F_{t,T}$

$d =$	1	2	3	4	5
$C(d) \sim \chi^2(5)$	23.07	15.37	13.77	9.243	8.810
p – value	0.000	0.008	0.017	0.099	0.117

Table VIII
Non-Linearity Tests – Two- and Five-Minute Frequency Series

Notes: Two minute frequency series: sample size is 5,124 observations. The starting point of the recursive OLS is 175. Five minute frequency series: sample size is 2,028 observations. The starting point of the recursive OLS is 110. Tests present heterokedasticity consistent results. All the delays are chosen to include up to 10 minutes in the estimations.

Panel A: H_0: "z_t follows a linear AR(I)"			
	$d =$	1	2
2-minute	$C(d) \sim \chi^2(3)$	17.43	5.803
	p-value	0.000	0.120
5-minute	$C(d) \sim \chi^2(2)$	4.46	
	p-value	0.107	
Panel B: H_0: "y_t follows a linear VECM(p)"			
	$d =$	1	2
2-minute	$C(d) \sim \chi^2(24)$	66.20	34.01
	p-value	0.000	0.084
5-minute	$C(d) \sim \chi^2(12)$	12.87	
	p-value	0.378	

Table IX
Self Exciting Threshold Autoregressive (SETAR) Model for Basis – Two-Minute Frequency Series

Notes: The model estimated is given in equation (2):

$$z_t = \delta^{(j)} + \sum_{i=1}^{I} \alpha_i^{(j)} z_{t-i} + \xi_t^{(j)} \qquad c_{j-1} < z_{t-1} \leq c_j$$

where j = 1, 2, 3 and the threshold lag equals to 1. The optimal threshold values are $\hat{c}_1$ = -0.051 and $\hat{c}_2$ = 0.045, which define the three regimes. The number of observations is 675, 3,456 and 993 in regimes 1, 2 and 3, respectively. White heteroskedasticity consistent standard errors are given. $\hat{\sigma}^{2,(j)}$ is the sum of squared residuals in the regression. *; ** and *** indicate significance at 10, 5 and 1 percent levels, respectively.

	Regime 1		**Regime 2**		**Regime 3**	
	$z_{t-1} \leq$-0.051		-0.051 < $z_{t-1} \leq$ 0.045		z_{t-1} > 0.045	
	Coef.	Std. Error	Coef.	Std. Error	Coef.	Std. Error
z_{t-1}	0.377***	0.089	0.472***	0.028	0.439***	0.071
z_{t-2}	0.153	0.041	0.152***	0.019	0.157***	0.034
$\delta^{(j)}$	$-2.8 \cdot 10^{-5}$	$6.4 \cdot 10^{-5}$	$1.4 \cdot 10^{-5,**}$	$8.8 \cdot 10^{-6}$	$7.6 \cdot 10^{-6}$	$5.0 \cdot 10^{-5}$
$\hat{\sigma}^{2,(j)}$	$4.4 \cdot 10^{-4}$		$4.0 \cdot 10^{-4}$		$4.1 \cdot 10^{-4}$	
Adj. R^2 (%)	7.13		13.9		9.31	

Table X
Threshold Error Correction Model for the Returns – Two-Minute Frequency Series

Notes: The estimated TVECM is given in equations (3) and (4). The lag-length in each regime and for each equation is based on significant coefficients. Number of observations is 681, 3,584 and 922 in regimes 1, 2 and 3 respectively. White heteroskedasticity consistent standard errors are reported. *, ** and *** stand for significance at 10, 5 and 1 percent levels, respectively.

	Regime 1		**Regime 2**		**Regime 3**	
	$z_{t-1} \leq 0.058$		$-0.058 < z_{t-1} \leq 0.049$		$z_{t-1} > 0.049$	
	Coef.	Std. Error	Coef.	Std. Error	Coef.	Std. Error
Futures equation, $\Delta\ln F_{t,T}$						
z_{t-1}	-0.032	0.141	-0.134***	0.032	-0.062	0.100
$\Delta\ln F_{t-1,T}$	0.133**	0.067	-0.022	0.026	0.014	0.049
$\Delta\ln S_{t-1}$	0.040	0.078	0.164***	0.028	0.088*	0.054
Constant	$1.3\cdot10^{-4}$	$1.0\cdot10^{-4}$	$8.6\cdot10^{-6}$	$8.7\cdot10^{-6}$	$-3.0\cdot10^{-5}$	$7.3\cdot10^{-5}$
Adj. R^2 (%)	2.31		1.92		0.72	
Cash equation, $\Delta\ln S_t$						
z_{t-1}	0.467***	0.111	0.111***	0.027	0.199***	0.084
$\Delta\ln F_{t-1,T}$	0.450***	0.057	0.288***	0.023	0.318***	0.044
$\Delta\ln F_{t-2,T}$	0.095	0.056	0.209***	0.022	0.205***	0.041
$\Delta\ln F_{t-3,T}$	0.158**	0.052	0.143***	0.022	0.138**	0.036
$\Delta\ln F_{t-4,T}$			0.066***	0.021	0.039	0.040
$\Delta\ln F_{t-5,T}$			0.045***	0.018	0.058*	0.033
$\Delta\ln S_{t-1}$	-0.177**	0.073	-0.139***	0.028	-0.183**	0.051
$\Delta\ln S_{t-2}$	-0.054	0.074	-0.135***	0.026	-0.214***	0.045
$\Delta\ln S_{t-3}$	-0.170**	0.052	-0.106***	0.025	-0.106**	0.040
$\Delta\ln S_{t-4}$	-0.165***	0.038	-0.080***	0.024	-0.075***	0.031
$\Delta\ln S_{t-5}$			-0.042**	0.018		
Constant	$2.1\cdot10^{-4,**}$	$7.4\cdot10^{-5}$	$-3.0\cdot10^{-6}$	$6.1\cdot10^{-6}$	$-1.3\cdot10^{-5}$	$5.7\cdot10^{-5}$
Adj. R^2 (%)	18.0		10.7		9.98	

Table XI
Linear VECM(2) for the Returns – Five-Minute Frequency Series

Notes: The system estimated is given as

$$\Delta \ln F_{t,T} = \phi_{10} + \sum_{i=1}^{2} \phi_{11,i} \Delta \ln F_{t-i,T} + \sum_{i=1}^{2} \phi_{12,i} \Delta \ln S_{t-i} + \beta_1 z_{t-1} + \varepsilon_{1,t}$$

$$\Delta \ln S_t = \phi_{20} + \sum_{i=1}^{2} \phi_{21,i} \Delta \ln F_{t-i,T} + \sum_{i=1}^{2} \phi_{22,i} \Delta \ln S_{t-i} + \beta_2 z_{t-1} + \varepsilon_{2,t}$$

White heteroskedasticity consistent standard errors are given in parenthesis. *; ** and *** stand for significance at 10, 5 and 1 percent levels, respectively.

	Coef.	Std. Error
Panel A: Futures equation, $\Delta \ln F_{t,T}$		
z_{t-1}	-0.189***	0.059
$\Delta \ln F_{t-1,T}$	0.1085	0.057
$\Delta \ln S_{t-1}$	0.017	0.061
Constant	$6.7 \cdot 10^{-6}$	$1.9 \cdot 10^{-5}$
Adj. R^2(%)	1.07	
Panel B: Cash equation, $\Delta \ln S_t$		
z_{t-1}	0.308***	0.051
$\Delta \ln F_{t-1,T}$	0.300***	0.048
$\Delta \ln F_{t-2,T}$	0.115**	0.041
$\Delta \ln S_{t-1}$	-0.176***	0.052
$\Delta \ln S_{t-2}$	-0.138***	0.041
Constant	$-1.4 \cdot 10^{-5}$	$1.6 \cdot 10^{-5}$
Adj. R^2(%)	14.0	

Figure I

Time Plots of One-Minute FTSE 100 Index and Futures Returns and Associated Threshold Variable

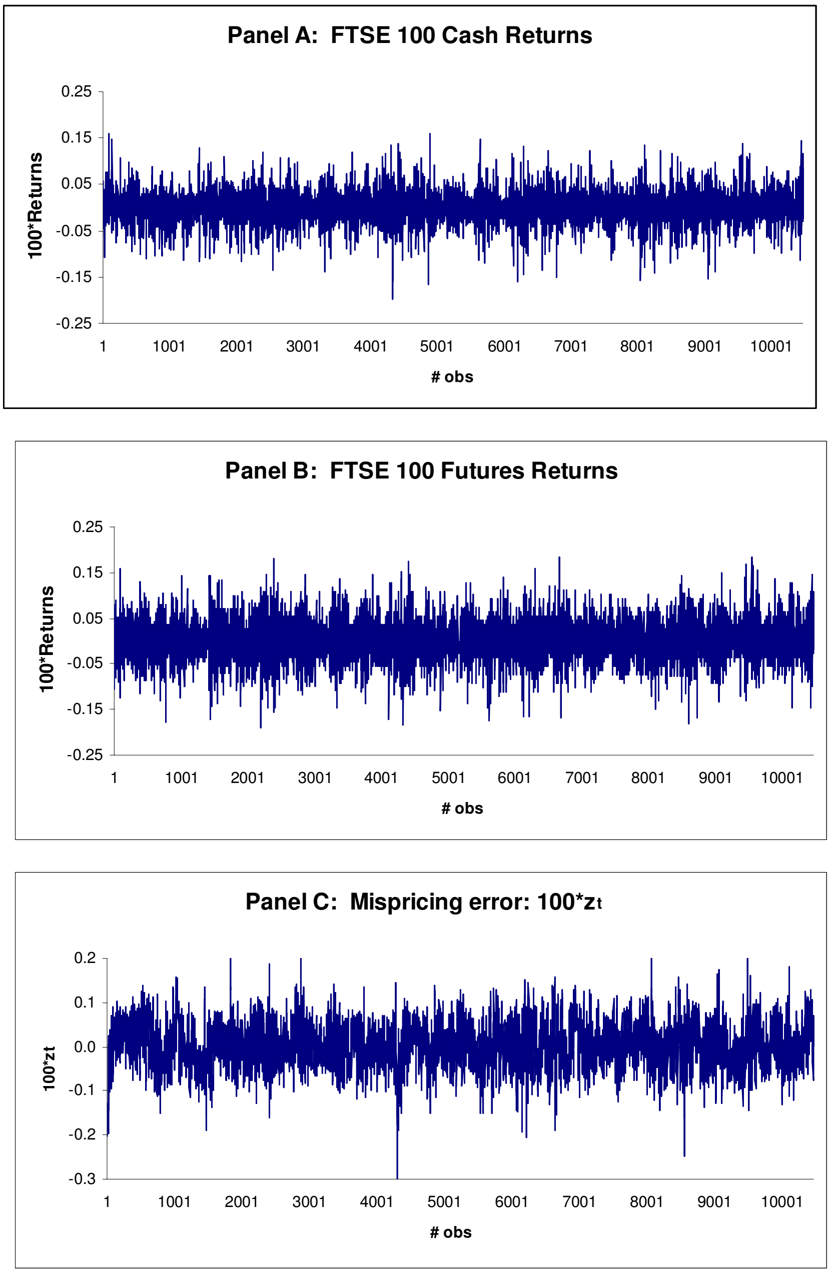
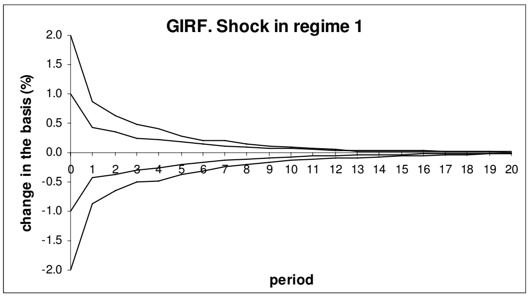
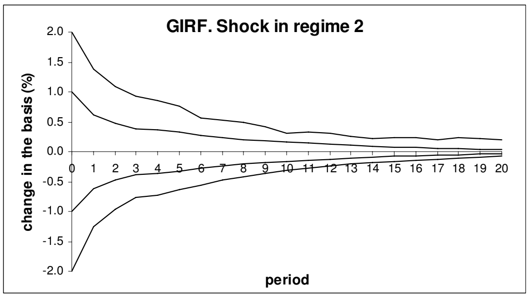


Figure II
Generalized Impulse Response Functions

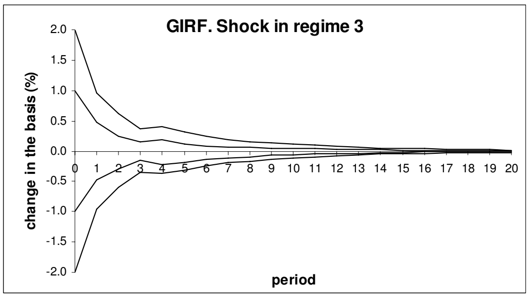
Panel A: Observation t = 5,545. History (t-4, …, t) = –0.007, 0.022, –0.003, –0.012, –0.082



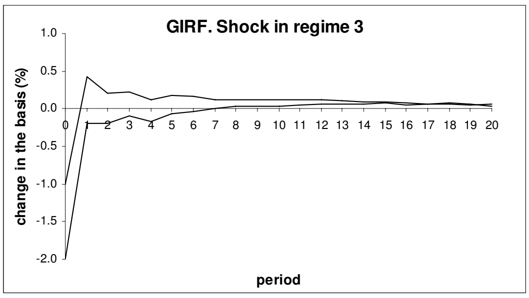
Panel B: Observation t = 6,029. History (t-4, …, t) = –0.094, –0.016, –0.029, 0.014, 0.016



Panel C: Observation t = 1,020. History (t-4, …, t) = 0.074, 0.101, 0.101, 0.051, 0.158



Panel D: Observation t = 7,687. History (t-4, …, t) = –0.007, 0.016, 0.038, 0.033, 0.055



CFS Working Paper Series:

No.	Author(s)	Title
2007/19	**Maria Kasch-Haroutounian** **Erik Theissen**	Competition between Exchanges: Euronext versus Xetra
2007/18	**Günter W. Beck** **Volker Wieland**	Money in Monetary Policy Design under Uncertainty: A Formal Characterization of ECB-Style Cross-Checking
2007/17	**Günter W. Beck** **Volker Wieland**	Money in Monetary Policy Design under Uncertainty: The Two-Pillar Phillips Curve versus ECB-Style Cross-Checking
2007/16	**Silvio Colarossi** **Andrea Zaghini**	Gradualism, Transparency and Improved Operational Framework: A Look at the Overnight Volatility Transmission
2007/15	**Annamaria Lusardi** **Olivia S. Mitchell**	Financial Literacy and Retirement Preparedness: Evidence and Implications for Financial Education Programs
2007/14	**Jean Boivin** **Marc P. Giannoni** **Ilian Mihov**	Sticky Prices and Monetary Policy: Evidence from Disaggregated U.S. Data
2007/13	**Virgiliu Midrigan**	Menu Costs, Multi-Product Firms, and Aggregate Fluctuations
2007/12	**Michael Woodford**	Robustly Optimal Monetary Policy with Near-Rational Expectations
2007/11	**Lars E. O. Svensson** **Noah Williams**	Bayesian and Adaptive Optimal Policy under Model Uncertainty
2007/10	**Alessandro Calza** **Tommaso Monacelli** **Livio Stracca**	Mortgage Markets, Collateral Constraints, and Monetary Policy: Do Institutional Factors Matter?

Copies of working papers can be downloaded at http://www.ifk-cfs.de