No. 2008/12

A Partially Linear Approach to Modelling the Dynamics of Spot and Futures Prices

Jürgen Gaul and Erik Theissen

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. 2008/12

A Partially Linear Approach to Modelling the Dynamics of Spot and Futures Prices*

Jürgen Gaul[1] and Erik Theissen[2]

March 5, 2008

Abstract:
In this paper we consider the dynamics of spot and futures prices in the presence of arbitrage. We propose a partially linear error correction model where the adjustment coefficient is allowed to depend non-linearly on the lagged price difference. We estimate our model using data on the DAX index and the DAX futures contract. We find that the adjustment is indeed nonlinear. The linear alternative is rejected. The speed of price adjustment is increasing almost monotonically with the magnitude of the price difference.

JEL Classification: C32; C14; G13; G14

Keywords: Futures Markets, Cointegrated Systems, Partially Linear Models, Nonparametric Methods.

* We thank Don Andrews, J¨org Breitung and Michal Paluch, conference participants of the North American Summer Meetings of the Econometric Society in Durham and the 2007 Annual Meeting of Verein f¨ur Socialpolitik in Munich as well as participants of the Econometrics Research Seminar at Yale University and the Bonn-Frankfurt Econometrics Workshop for helpful comments. Moreover, we would like to thank Bloomberg, L. P. for providing the data.

1 Jürgen Gaul, University of Bonn, Bonn Graduate School of Economics, Adenauerallee 24- 42, 53113 Bonn - Germany, email: juergen.gaul@uni-bonn.de

2 Erik Theissen, (corresponding author) University of Bonn, and CFS, BWL I, Adenauerallee 24-42, 53113 Bonn - Germany, email: theissen@uni-bonn.de

Introduction

Prices in spot and futures markets are linked through the cost-of-carry relation. In a frictionless world arbitrage would eliminate any deviations from this relation. In practice, however, such deviations may and do occur for several reasons. First, the existence of transactions costs makes it unprofitable to exploit small deviations. Second, traders with access to private information may prefer to trade in a specific market. Consequently, prices in this market may reflect information earlier than prices in the other market. As transaction costs tend to be lower in the futures market (e.g. Berkmann et al. 2005) informed traders may prefer to trade in this market and it thus might reflect the information earlier than the spot market. The opposite may also occur, however. Consider a trader with information on the value of an individual stock. The trader can trade on that information in the spot market. In the futures market, on the other hand, he is restricted to trading a basket of securities (i.e., an index futures contract). Therefore, firm-specific information may be reflected in the spot market first.

The question of which market impounds new information faster is thus an empirical one, and it has been subject to academic research for about two decades.[1] The empirical methods have been considerably refined since the early work of Kawaller et al. (1987) and others. VAR models were introduced (e.g. Stoll and Whaley 1990) and soon thereafter replaced by error correction (ECM) models. A standard ECM implicitly assumes that deviations of prices from their long-run equilibrium (the pricing errors) are reduced at a speed that is independent of the magnitude of the price deviation. This is unlikely to be the case, however. Whenever the deviations are sufficiently large to allow for profitable arbitrage, the speed of adjustment should increase.[2] Some authors (e.g. Yadav et al. 1994, Dwyer et al. 1996 and Martens et al. 1998) have employed threshold error correction (TECM) models to address this issue. A TECM assumes a non-continuous transition function and allows for a discrete number of different speed of adjustment coefficients. If all traders would face identical transaction costs, a TECM with two different adjustment coefficients (i.e., a no-arbitrage regime and an arbitrage regime) would be a reasonable choice. If, on the other hand, traders are heterogeneous with respect to the transaction costs they face, a less restrictive model is warranted. An obvious candidate is a smooth transition error correction (STECM) model as applied by Taylor et al. (2000), Anderson and Vahid (2001) and Tse (2001).

A shortcoming of the STECM models is that the transition function must be exogenously specified, and there is no theory to guide the specification of the model. The researcher also has to decide for a symmetric transition function or one that allows for asymmetry. Such asymmetries may arise because short sales in the spot market are more expensive than short sales in the futures market.

The contribution of our paper is to propose a more flexible modelling framework. We estimate a partially linear ECM where the adjustment process is modelled non-parametrically. The short-run dynamics are estimated by density-weighted OLS based on the approach proposed by Fan and Li (1999a). The non-parametric function modelling the adjustment process is estimated by a Nadaraya-Watson estimator. The modelling approach that we use was proposed by Gaul (2005) but has as yet not been applied.

We implement our model using data from the German stock market. Specifically, we analyze the dynamics of the DAX index and the DAX futures contract. The results suggest that the speed of adjustment is indeed monotonically increasing in the magnitude of the price deviation. We test our specification against a standard ECM and clearly reject the latter. Estimates of the parameters governing the short-run dynamics are similar in the standard ECM and in our model.

These results have several implications. First, they confirm the intuition that the speed of adjustments of prices to deviations from equilibrium is increasing in the magnitude of the deviation. Second, they imply that a standard ECM as well as a TECM is unable to fully capture the dynamics of the adjustment process. Third, the form of the non-parametric adjustment function may guide the choice for a functional form in STECM models.

The remainder of the paper is organized as follows. Section 1 provides a description of the data set. In section 2 we describe the estimation procedure. In section 3 we describe a test for linearity. Section 4 is devoted to the presentation of the results, section 5 concludes.

1 Market Structure and Data

Our analysis uses DAX index level data and bid and ask quotes from the DAX index futures contract traded on Eurex. The DAX is a value-weighted index calculated from the prices of the 30 largest German stocks. The prices are taken from Xetra, the most liquid market for German stocks.[3] Index values are published in intervals of 15 seconds. The DAX is a performance index, i.e., the calculation of the index is based on the presumption that dividends are reinvested. As a consequence, the expected dividend yield does not enter the cost of carry relation. Besides an index calculated from the most recent transaction prices the exchange also calculates an index from the current best ask prices (ADAX) and an index calculated from the current best bid prices (BDAX). These indices are value-weighted averages of the inside quotes, and their mean is equivalent to a value-weighted average of the quote midpoints of the component stocks.

Futures contracts on the DAX are traded on the EUREX. The contracts are cash-settled and trade on a quarterly cycle. They mature on the third Friday of the months March, June, September, and December. The DAX futures contract is a highly liquid instrument. In the first quarter of 1999 (our sample period), more than 1,150,000 transactions were recorded. The open interest at the end of the quarter amounted to more than 290,000 contracts.

Both Xetra and EUREX are electronic open limit order books. Therefore, the results of our empirical analysis are unlikely to be affected by differences in market structure. The trading hours in the two markets are different, though. Trading in Xetra starts with a call auction held between 8.25 am and 8:30 am. After the opening auction, continuous trading starts and extends until 5 pm, interrupted by an intraday auction which takes place between 1:00 pm and 1:02 pm. Trading of the DAX futures contract starts at 9 am and extends until 5 pm.

We obtained all data from Bloomberg. Our sample period is the first quarter of 1999 and extends over 61 trading days. For this period we obtained the values of the DAX index and the two quote-based indices ADAX and BDAX at a frequency of 15 seconds. From the quote-based indices we calculate the midquote index $MQDAX_t = \frac{ADAX_t + BDAX_t}{2}$. We further obtained a time series of all bid and ask quotes and all transaction prices of the nearby DAX futures contract. We only use data for the period of simultaneous operation of both markets. We further discard all observations before 9 am and from 4:55 pm onwards. We also discard all observations within 5 minutes from the time of the intraday call auction (held between 1:00 pm and 1:02 pm). After these adjustments the sample consists of

100188 observations.

All estimations are based on quote midpoints. They are preferred to transaction prices because the use of midpoints alleviates the infrequent trading problem.[4] We match each index level observation whith the bid and ask quotes in the futures market that were in effect at the time the index level information was published.

The cost-of-carry relation implies that the cash index and the futures contract are cointegrated. In order to eliminate the time-variation of the cointegrating relation we discount the futures prices using daily observations on the one-month interbank rate as published by Deutsche Bundesbank.[5]

As a prerequisite for our empirical analysis we have to establish that the time series are I(1) and are cointegrated. Table 1 presents the results of augmented Dickey-Fuller tests and Phillips-Perron tests applied to p_t and Δp_t. p_t denotes a log price series observed at date t and the indices X and F identify observations relating to the cash market (X, Xetra) and the futures market (F), respectively. Δ is the difference operator. The results of the stationarity tests clearly suggest that all series are I(1).

	Level		First Difference	
	Augmented DF	Phillips / Perron	Augmented DF	Phillips / Perron
p^X	0.5773	0.6395	0.0001	0.0001
p^F	0.3964	0.4113	0.0001	0.0001

Table 1: Results of the Unit-Root tests for both time series

In equilibrium spot and futures prices are linked through the cost-of-carry relation. Consequently, the DAX index level and the discounted futures price should be equal in equilibrium, and their difference should be stationary. We test the latter hypothesis using both an augmented Dickey-Fuller test and a Phillips-Perron test and clearly reject the null of a unit root (p-value 0.0000 and 0.0001, respectively). This result confirms the theoretical prediciton that spot and futures prices are cointegrated with the cointegrating vector being $(1,-1)^\top$. We use this pre-specified cointegrating vector in our estimation.

2 Estimation procedure

For the reasons exposed in the Introduction, our model is characterized by a nonparametric function for the pricing error. In particular, we propose to use the model

$$\Delta y_t = \sum_{i=1}^{k} \Gamma_i \Delta y_{t-i} + F(\beta^\top y_{t-1}) + \epsilon_t, \qquad \text{t=1,..,T}, \tag{1}$$

where y_t denotes a vector process containing the variables p_t^X and p_t^F. The cointegrating vector is denoted by β and is pre-specified to $(1,-1)^\top$. The adjustment process is described by the unknown nonparametric function $F\colon \mathbb{R} \to \mathbb{R}^2$ and ϵ_t is a two-dimensional error process. By introducing the $2\times 2k$-matrix $\Gamma := (\Gamma_1 \ldots \Gamma_k)$ and the $2k$-dimensional vector $\xi_{t-1} := \left(\Delta y_{t-1}^\top \ldots \Delta y_{t-k}^\top\right)^\top$, model (1) can be written as

$$\Delta y_t = \Gamma \xi_{t-1} + F(\beta^\top y_{t-1}) + \epsilon_t. \tag{2}$$

Note that model (2) contains the linear VECM (Engle and Granger, 1987; Johansen, 1988), the threshold VECM (Hansen and Seo, 2002) and the smooth transition VECM (van Dijk and Franses, 2000) as special cases.

The estimation procedure described in the following involves two stages. First, we estimate the matrix Γ, then the function F.

2.1 Estimation of Γ

Taking expectations in (2) conditional on $\beta^\top y_{t-1}$, we have

$$E(\Delta y_t | \beta^\top y_{t-1}) = \Gamma E(\xi_{t-1} | \beta^\top y_{t-1}) + F(\beta^\top y_{t-1}), \tag{3}$$

using $E(\epsilon_t | \beta^\top y_{t-1}) = 0$. Subtracting (3) from (2) leads to

$$\Delta y_t - E(\Delta y_t | \beta^\top y_{t-1}) = \Gamma(\xi_{t-1} - E(\xi_{t-1} | \beta^\top y_{t-1})) + \epsilon_t, \tag{4}$$

which has the following form

$$\Delta y_t^* = \Gamma \xi_{t-1}^* + \epsilon_t, \tag{5}$$

where $\Delta y_t^* := \Delta y_t - E(\Delta y_t|\beta^\top y_{t-1})$ and $\xi_{t-1}^* := \xi_{t-1} - E(\xi_{t-1}|\beta^\top y_{t-1})$. If $E(\Delta y_t|\beta^\top y_{t-1})$ and $E(\xi_{t-1}|\beta^\top y_{t-1})$ were known, Γ could be estimated by OLS. Since $E(\Delta y_t|\beta^\top y_{t-1})$ and $E(\xi_{t-1}|\beta^\top y_{t-1})$ are usually unknown, an estimator based on Δy_t^* and ξ_{t-1}^* is not feasible. To obtain a feasible estimator, we will use the nonparametric kernel method, similar to Robinson (1988) and Fan and Li (1999a). In particular, the conditional means $E(\Delta y_t|\beta^\top y_{t-1})$ and $E(\xi_{t-1}|\beta^\top y_{t-1})$ are estimated by the Nadaraya-Watson estimator

$$\begin{aligned}
\hat{E}(\Delta y_t|\beta^\top y_{t-1}) &= \frac{1}{Th}\sum_{j=1}^{T} \Delta y_j K\left(\frac{\beta^\top y_{t-1} - \beta^\top y_{j-1}}{h}\right) / \hat{f}(\beta^\top y_{t-1}), \\
\hat{E}(\xi_{t-1}|\beta^\top y_{t-1}) &= \frac{1}{Th}\sum_{j=1}^{T} \xi_{j-1} K\left(\frac{\beta^\top y_{t-1} - \beta^\top y_{j-1}}{h}\right) / \hat{f}(\beta^\top y_{t-1}),
\end{aligned}$$

where

$$\hat{f}(\beta^\top y_{t-1}) = \frac{1}{Th}\sum_{j=1}^{T} K\left(\frac{\beta^\top y_{t-1} - \beta^\top y_{j-1}}{h}\right) \tag{6}$$

is the kernel density estimator for $f(\beta^\top y_{t-1})$, $K(\cdot)$ is a kernel function and h is a bandwidth parameter.

To avoid the random denominator problem in kernel estimation (i.e. the occurrence of small values of the estimated density function), we use density weighted estimates, similar to Fan and Li (1999a). Thus, we multiply (5) by $f(\beta^\top y_{t-1})$, the density function of $\beta^\top y_{t-1}$, and obtain

$$f(\beta^\top y_{t-1})\Delta y_t^* = \Gamma f(\beta^\top y_{t-1})\xi_{t-1}^* + f(\beta^\top y_{t-1})\epsilon_t. \tag{7}$$

We replace $E(\Delta y_t|\beta^\top y_{t-1})$, $E(\xi_{t-1}|\beta^\top y_{t-1})$ and $f(\beta^\top y_{t-1})$ in (7) by their estimates. This leads to the feasible estimator

$$\hat{\Gamma}^{\text{OLS}} = \left[\sum_{t=1}^{T} \Delta\hat{y}_t^* \hat{\xi}_{t-1}^{*\top} \hat{f}(\beta^\top y_{t-1})^2\right] \left[\sum_{t=1}^{T} \hat{\xi}_{t-1}^* \hat{\xi}_{t-1}^{*\top} f(\beta^\top y_{t-1})^2\right]^{-1}, \tag{8}$$

with $\Delta\hat{y}_t^* := \Delta y_t - \hat{E}(\Delta y_t|\beta^\top y_{t-1})$ and $\hat{\xi}_{t-1}^* := \xi_{t-1} - \hat{E}(\xi_{t-1}|\beta^\top y_{t-1})$. Besides some technical assumptions, we assume that $(\Delta y_t, \beta^\top y_{t-1})$ is β-mixing, $Th^2 \to \infty$ and $Th^8 \to 0$ for $T \to \infty$. Similar to Fan and Li (1999a), it can be shown that vec $(\hat{\Gamma}^{\text{OLS}} - \Gamma)$ is $\sqrt{T}$ consistent and asymptotically normally distributed. For a precise formulation of this statement and its assumptions we refer to Theorem 2 in Gaul (2005).

2.2 Estimation of F

Substituting $\hat{\Gamma}^{\text{OLS}}$ for Γ in model (2), one obtains the nonlinear, nonparametric model

$$\Delta\tilde{y}_t = F(\beta^\top y_{t-1}) + u_t, \tag{9}$$

where $\Delta\tilde{y}_t := \Delta y_t - \hat{\Gamma}^{\text{OLS}}\xi_{t-1}$.

Applying the Nadaraya-Watson estimator to (9), i.e.

$$\hat{F}(z) = \frac{\sum_{t=1}^{T} \Delta\tilde{y}_t K\left(\frac{z-\beta^\top y_{t-1}}{h}\right)}{\sum_{t=1}^{T} K\left(\frac{z-\beta^\top y_{t-1}}{h}\right)} \tag{10}$$

we get an estimator for the function F. It is well known that $\hat{F}(\cdot)$ has the same asymptotic distribution as if Γ were known. Later, we will use this statement for constructing pointwise confidence intervals.

2.3 Bandwidth Selection

In empirical applications we have to choose both the kernel function and the bandwidth parameter h. Whereas the influence of the kernel function is negligible, the choice of the bandwidth parameter plays a crucial role. Due to the enormous sample size, standard bandwidth selection procedures like cross-validation, are no longer applicable as the computational time increases at quadratic rate with the number of observations. In order to determine the bandwidth parameter h we use the method of *Weighted Averaging of Rounded Points (WARPing)* developed

by Härdle and Scott (1992). This technique is based on discretizing the data first into a finite grid of bins, then smoothing the binned data and finally selecting the optimal bandwidth using the binned data. The main advantage of WARPing is the substantial gain of computational efficiency. In particular, Härdle (1991) and Härdle and Scott (1992) show that the number of iterations increases at linear rate with the number of observations rather than quadratic.
In our application we determine the optimal bandwidth by using four different criteria, namely cross-validation, the Shibata's Model Selector, Akaike's Information Criterion and the Final Prediction Error Criterion. For a detailed discussion of them, we refer to Härdle, Müller, Sperlich and Werwatz (2004). The lower limit for h for the grid search is set to 0.000332, the upper to 0.005307 and the bandwidth d to $6.634 \cdot 10^{-5}$. The number of equidistant grid points is chosen to be 100. The analysis is carried out by using the software package XploRe. The results are given in the table below.

Bandwidth selection procedure	XDAX	FDAX
Cross Validation	0.000371	0.000492
Shibata's Model Selector	0.000351	0.000492
Akaike's Information Criterion	0.000361	0.000492
Final Prediction Error	0.000361	0.000492

Table 2: Results of bandwidth selection

The table shows that all methods lead to very similar results for the XDAX series. According to Akaike's Information Criterion and Final Prediction Error we choose $h^X = 0.000361$. For the FDAX series, all methods yield the same result. Hence, we choose $h^F = 0.000492$.

3 Test for linearity

The linear vector error correction model

$$\Delta y_t = \Gamma \xi_{t-1} + \alpha \beta^\top y_{t-1} + \epsilon_t \tag{11}$$

may be considered the baseline model in cointegration analysis. We now provide a statistical single-equation test to examine the hypothesis whether model (11) is as accurate a description of the data as model (1). Formally, we are interested in testing the hypotheses

$H_0:$ $E(\Delta y_{it}|\xi_{t-1}, \beta^\top y_{t-1}) = \Gamma_i \xi_{t-1} + \alpha_i \beta^\top y_{t-1}$ for some Γ_i and α_i against

$H_1:$ $E(\Delta y_{it}|\xi_{t-1}, \beta^\top y_{t-1}) = \Gamma_i \xi_{t-1} + F_i(\beta^\top y_{t-1})$ with $P(F_i(\beta^\top y_{t-1}) = \alpha_i \beta^\top y_{t-1}) < 1)$ for any $\alpha_i \in \mathbb{R}$.

To motivate an appropriate test statistic, we consider (2) with $\Gamma = 0$. Denote $u_{it} := \Delta y_{it} - \alpha_i \beta^\top y_{t-1}$ the residuals under H_0. Following Zheng (1996) and Li and Wang (1998), our test is based on $E\left[u_{it} E[u_{it}|\beta^\top y_{t-1}] f(\beta^\top y_{t-1})\right]$. Then under H_0, it follows

$$E\left[u_{it} E[u_{it}|\beta^\top y_{t-1}] f(\beta^\top y_{t-1})\right] = 0, \tag{12}$$

since $E[u_{it}|\beta^\top y_{t-1}] = 0$. Under H_1, we have $E[u_{it}|\beta^\top y_{t-1}] = F_i(\beta^\top y_{t-1}) - \alpha_i \beta^\top y_{t-1}$. Using the law of iterated expectations, we get under H_1

$$\begin{aligned} & E\left[u_{it} E[u_{it}|\beta^\top y_{t-1}] f(\beta^\top y_{t-1})\right] \\ = \; & E[E(u_{it} E(u_{it}|\beta^\top y_{t-1}) f(\beta^\top y_{t-1})|\beta^\top y_{t-1})] \\ = \; & E[E(u_{it}|\beta^\top y_{t-1}) E(u_{it}|\beta^\top y_{t-1}) f(\beta^\top y_{t-1})] \\ = \; & E[(F_i(\beta^\top y_{t-1}) - \alpha_i \beta^\top y_{t-1})^2 f(\beta^\top y_{t-1})] \\ > \; & 0. \end{aligned} \tag{13}$$

Due to (12) and (13) it is obvious to use the sample analogue of $E\left[u_{it} E[u_{it}|\beta^\top y_{t-1}] f(\beta^\top y_{t-1})\right]$ as the test statistic. The outer expected value is replaced by its mean, the inner expected value by the Nadaraya-Watson estimator

$$\hat{E}(u_{it}|\beta^\top y_{t-1}) = \frac{1}{(T-1)h} \sum_{j=1 j\neq t}^{T} K\left(\frac{\beta^\top y_{t-1} - \beta^\top y_{j-1}}{h}\right) u_{ij}/\hat{f}(\beta^\top y_{t-1}),$$

the density function $f(\cdot)$ by the kernel density estimator (6) and the residuals u_{it} by the empirical residuals under the null hypothesis, i.e. $\tilde{u}_{it} = \Delta y_{it} - \hat{\alpha}_i \beta^\top y_{t-1}$. Taking the lagged dependent values into account we substitute for $\tilde{u}_{it}$ the residuals $\hat{u}_{it} = \Delta y_{it} - \hat{\Gamma}_i^{\text{OLS}} \xi_{t-1} - \hat{\alpha}_i \beta^\top y_{t-1}$, where $\hat{\Gamma}_i^{\text{OLS}}$ denotes the estimator of the i-th row of Γ given by (8) and $\hat{\alpha}_i$ is the estimator of the i-th row of α under the null hypothesis. Thus, the test statistic is of the form

$$I_i := \frac{1}{T(T-1)h} \sum_{t=1}^{T} \sum_{j=1 j\neq t}^{T} K\left(\frac{\beta^\top y_{t-1} - \beta^\top y_{j-1}}{h}\right) \hat{u}_{it}\hat{u}_{ij}, \quad i = 1, \ldots, p$$

To derive the asymptotic distribution, it is important to note that I_i is a degenerate, second-order U-statistic. Combining the ideas of Fan and Li (1999b) and Li and Wang (1998), it can be shown that I_i is asymptotically normally distributed by applying a central limit theorem for U-statistics of β-mixing processes. Furthermore,

$$\hat{\sigma}_i^2 := \frac{2}{T(T-1)h} \sum_{t=1}^{T} \sum_{j=1 j\neq t}^{T} K^2\left(\frac{\beta^\top y_{t-1} - \beta^\top y_{j-1}}{h}\right) \hat{u}_{it}^2\hat{u}_{ij}^2, \quad i = 1, \ldots, p$$

is a consistent estimator for σ_i^2, the asymptotic variance of $Th^{1/2}I_i$. It is well known that the convergence speed to the normal distribution is quite low. Therefore, bootstrap methods are suggested to approximate the finite sample distribution, see e.g. Li and Wang (1998). Due to the enormous sample size in our application, however, it seems reasonable to rely on the asymptotic approximation given through the asymptotic distribution.

4 Results

We present the results in two steps. The starting point is the linear benchmark case. We then proceed to the partially linear model and also present the results for the test of linearity described in the previous section.

4.1 Linear error correction model

The following table shows the estimation results of the linear error correction model

$$
\begin{aligned}
r_t^F &= \mu^F + \sum_{i=1}^{20} \gamma_{1i}^F r_{t-i}^F + \sum_{i=1}^{20} \gamma_{1i}^X r_{t-i}^X + \alpha^F (p_{t-1}^X - p_{t-1}^F) + \epsilon_t^F \\
r_t^X &= \mu^X + \sum_{i=1}^{20} \gamma_{2i}^X r_{t-i}^X + \sum_{i=1}^{20} \gamma_{2i}^F r_{t-i}^F + \alpha^X (p_{t-1}^X - p_{t-1}^F) + \epsilon_t^X,
\end{aligned}
$$

where p denotes the log prices and r denotes a log return. The index X identifies variables and coefficients relating to the spot market (X, Xetra), the index F identifies variables (adjusted by a discount factor according to the cost-of-carry relation) and coefficients relating to the futures market. The cointegrating vector is pre-specified to $(1, -1)^\top$. The model is estimated by OLS with 20 lags, but to save space we present only the coefficients for lags 1-4. Standard errors are based on the heteroskedasticity-robust covariance estimator. The model is estimated based on quote midpoints and 100188 observations.

	XDAX		FDAX	
	Estimates	t-statistic	Estimates	t-statistic
Constant	3.385E-6	4.95	-4.427E-6	-3.80
EC	-0.0087	-14.85	0.0047	5.42
XDAX(-1)	-0.0876	-16.36	0.0542	7.36
XDAX(-2)	-0.0773	-16.22	0.0534	7.83
XDAX(-3)	-0.0632	-14.80	0.0573	7.69
XDAX(-4)	-0.0522	-12.14	0.0489	6.76
FDAX(-1)	0.2107	68.32	0.0358	7.97
FDAX(-2)	0.1572	58.18	-0.0166	-3.81
FDAX(-3)	0.1215	46.31	-0.0173	-3.97
FDAX(-4)	0.0989	37.38	-0.0079	-1.78
R^2	0.2244		0.0070	

Table 3: Estimation results of the linear ECM

Considering the short-run dynamics first, we find that the DAX returns depend negatively on their own lagged values but depend positively on lagged futures returns. Returns in the futures markets exhibit a similar pattern. There is one exception, however, as the coefficient on the first lag of the futures returns is positive and significant. The results of F-tests (not shown in the table) indicate that there is bivariate Granger causality.

The coefficients on the error correction term have the expected signs (negative for the spot market and positive for the futures market) and are both highly significant. The estimates can be used to construct the common factor weights

$$\theta^X = \frac{\alpha^F}{\alpha^F - \alpha^X}; \quad \theta^F = \left(1 - \theta^X\right) = \frac{-\alpha^X}{\alpha^F - \alpha^X}$$

The common factor weights measure the contributions of the two markets to the process of price discovery. The measure builds on Gonzalo and Granger (1995) and is discussed in more detail in Booth et a. (2002), deB Harris et al.

(2002) and Theissen (2002). In our linear error correction model the common factor weights are 0.3507 for the spot market and 0.6493 for the futures market. The futures market thus dominates in the process of price discovery. This result is consistent with previous findings.

4.2 Partially linear error correction model

The following table shows the estimation results of the partially linear error correction model

$$
\begin{aligned}
r_t^F &= \sum_{i=1}^{20} \gamma_{1i}^F r_{t-i}^F + \sum_{i=1}^{20} \gamma_{1i}^X r_{t-i}^X + F(p_{t-1}^X - p_{t-1}^F) + \epsilon_t^F \\
r_t^X &= \sum_{i=1}^{20} \gamma_{2i}^X r_{t-i}^X + \sum_{i=1}^{20} \gamma_{2i}^F r_{t-i}^F + F(p_{t-1}^X - p_{t-1}^F) + \epsilon_t^X,
\end{aligned}
$$

where the notation is as in the linear model. We estimate the model by the procedure described in section 3. Again, we use 20 lags, but only the coefficients for lags 1-4 are shown. Again, standard errors are based on the heteroskedasticity-robust covariance estimator. The cointegrating vector is pre-specified to $(1, -1)^\top$.

	XDAX		FDAX	
	Estimates	t-statistic	Estimates	t-statistic
XDAX(-1)	-0.0873	-15.25	0.0389	4.79
XDAX(-2)	-0.0693	-14.90	0.0475	6.15
XDAX(-3)	-0.0564	-13.57	0.0491	5.78
XDAX(-4)	-0.0435	-10.76	0.0449	5.54
FDAX(-1)	0.1571	70.98	0.0558	11.39
FDAX(-2)	0.1351	58.79	0.0020	0.39
FDAX(-3)	0.1063	47.14	-0.0053	-1.05
FDAX(-4)	0.0882	39.27	-0.0028	-0.54

Table 4: Estimation results of the partially linear ECM ($h = 2\hat{\sigma}T^{-0.2}$)

Applying the test for linearity developed in section 3, we obtain $I^F = 3.265$ and $I^X = 2.937$. We thus clearly reject the linear benchmark model in favor of our non-parametric specification. For the test we choose the bandwidth parameter to be $h = 2\hat{\sigma}T^{-0.2}$.

The results for the short-run dynamics are similar to those in the linear model. The spot market returns depend positively on their own lagged values and negatively on the lagged futures returns. Futures returns, on the other hand, depend positively on the lagged spot market returns. They also depend positively on their first lag. Coefficients for higher lags are insignificant.

Figure 1 presents the results for the adjustment process. The figure plots the value of the adjustment function F against the pricing error $\beta^\top y_{t-1}$. It also depicts the 95% confidence intervals. The upper panel shows the results for the futures market, the lower panel those for the spot market.The adjustment process is estimated very precisely, as evidenced by the narrow confidence intervals. In the outer regions (i.e., when pricing errors are large) estimation is less precise. This is a natural consequence of the low number of observations in these regions.

The speed of adjustment is almost monotonically related to the magnitude of the pricing error. This shape of the adjustment function is clearly at odds with a threshold error correction model. Adjustment is slow for small pricing errors, as is evidenced by the small slope of the adjustment function. When the pricing error becomes larger, the speed of adjustment increases sharply. This is consistent with arbitrage activities.

There is an asymmetry with respect to the level of the pricing error that triggers arbitrage. When the pricing error is negative (i.e., when the adjusted futures price is larger than the spot price) the trigger level is about -0.001. When the pricing error is positive, on the other hand, the trigger level is approximately 0.003. This pattern is explained by slight, but systematic deviations of prices from the cost-of-carry relation. On average, the difference between the discounted futures price and the DAX index is -2.8 index points. This pattern has been documented in previous research (e.g. Bühler and Kempf, 1995), and the most likely explanation is differential tax treatment of dividends in the spot and the futures market (see McDonald, 2001 for a detailed discussion).

In order to compare the predictive ability of the partially linear VECM with that of the linear VECM, the root mean squared error (RMSE) and the mean absolute error (MAE) are calculated for both models.[6] The RMSE and the MAE are defined for one-step ahead forecast errors by

$$\begin{aligned} RMSE &= \sqrt{\sum_{t=k}^{T} \left(\hat{E}_{t-1} p_t^X - p_t^X\right)^2}, \\ MAE &= \sum_{t=k}^{T} \left|\hat{E}_{t-1} p_t^X - p_t^X\right|. \end{aligned}$$

We set $k = 80000$ to ensure that the parameter estimates are based on a sufficiently large numbers of observations. The results are shown in Table 5.

	Linear VECM (A)	Partially Linear VECM (B, B/A)
RMSE	0.025	0.023 (0.919)
MAE	2.276	2.067 (0.908)

Table 5: Prediction ability of the linear VECM and the partially linear VECM

Table 5 shows that the root mean squared error (RMSE) of the partially linear VECM is about 10% lower than that of the linear VECM. A similar result is obtained for the mean absolute error (MAE). Hence, the partially linear VECM clearly improves the forecasting ability.

5 Conclusion

The present paper extends the literature on the joint dynamics of prices in spot and futures markets by modelling the price-adjustment process non-parametrically using the methodology developed in Gaul (2005).

We apply our partially linear error correction model to data for the German blue chip index DAX and the DAX futures contract traded on the EUREX. We find that the adjustment process is indeed nonlinear. The linear benchmark case is rejected at all reasonable levels of significance. Consistent with economic intuition, the speed of adjustment is almost monotonically increasing in the magnitude of the pricing error (the deviation between discounted futures price and spot price). This pattern is inconsistent with a simple threshold error correction model. It is consistent with a smooth transition model, and in fact the shape of the adjustment process in our non-parametric model may guide the choice of the transition function in future empirical research.

Notes

[1]Given the nature of our empirical analysis we restrict the brief survey of the literature to papers analyzing the relation between stock price indices and stock index futures contracts.

[2]The width of the arbitrage bounds is likely to depend on the liquidity of the market. In a recent paper Roll et al. (2007) have documented a relation between liquidity and the futures-cash basis for the NYSE composite index futures contract over the period 1988-2002.

[3]The DAX stocks are traded on Xetra, on the floor of the Frankfurt Stock Exchange and on several regional exchanges. The market share of Xetra amounted to 90% during our sample period.

[4]Spot market index levels are calculated using the last available transaction price for each of the component stocks. As stocks do not trade simultaneously, some of the prices used to calculate the index are stale. This may induce positive serial correlation in the index returns. Quote midpoints, on the other hand, are based on tradable bid and ask prices and should be less affected by the infrequent trading problem. See Shyy et al. (1996) or Theissen (2005).

[5]Given the margin requirements in the futures market, the rate for overnight deposits is an alternative choice. However, the time series of overnight deposit rates exhibits peaks which may be due to bank reserve requirements. Besides, the term structure at the short end was essentially flat during the sample period, making the choice of the interest rate less important.

[6]We restrict the analysis of the forecasting errors to the XDAX equation. This equation lends itself to forecasting because of the high R^2 and the large and significant coefficients on

the lagged futures returns documented in table 3.

References

[1] Anderson, H. M. and F. Vahid (2001), Market Architecture and Nonlinear Dynamics of Australian Stock and Futures Indices, *Australian Economic Papers* **40**, 541-566.

[2] Berkmann, H., Brailsford, T. and A. Frino (2005), A Note on Execution Costs for Stock Index Futures: Information Versus Liquidity Effects, *Journal of Banking and Finance* **29**, 565-577.

[3] Booth, G., Lin, J., Martikainen, T. and Y. Tse (2002) Trading and Pricing in Upstairs and Downstairs Markets, *Review of Financial Studies* **15**, 1111-1135.

[4] Buehler, W. and A. Kempf (1995), DAX Index Futures: Mispricing and Arbitrage in German Markets, *Journal of Futures Markets* **15**, 833-859.

[5] deB Harris, F., McInish, T. and R. Wood (2002), Common factor components versus information shares: A reply, *Journal of Financial Markets* **5**, 341-348.

[6] Dwyer Jr., G. P., Locke, P. and W. Yu (1996), Index Arbitrage and Nonlinear Dynamics between the S&P 500 Futures and Cash, *Review of Financial Studies* **9**, 301-332.

[7] Engle, R. F. and C. W. J. Granger (1987), Co-integration and Error Correction: Representation, Estimation and Testing, *Econometrica* **55**, 251-276.

[8] Fan, Y. and Q. Li (1999a), Root-N-Consistent Estimation of Partially Linear Time Series Models, *Journal of Nonparametric Statistics* **11**, 251-269.

[9] Fan, Y. and Q. Li (1999b), Central Limit Theorem for Degenerate U-Statistics of Absolutely Regular Processes with Applications to Model Specification Testing, *Journal of Nonparametric Statistics* **10**, 245-271.

[10] Gaul, J. (2005), A Partially Linear Vector Error Correction Model, Unpublished Manuscript, University of Bonn.

[11] Gonzalo, J. and C. W. J. Granger (1995), Estimation of Common Long-Memory Components in Cointegrated Systems, *Journal of Business and Economic Statistics* **13**, 27-35.

[12] Härdle, W. (1991), *Smoothing Techniques - With Implementation in S*, Springer: New York.

[13] Härdle, W., Müller, M., Sperlich, S. and A. Werwatz (2004), *Nonparametric and Semiparametric Models*, Springer: New York.

[14] Härdle, W. and D. W. Scott (1992), Smoothing in by Weighted Averaging Using Rounded Points, *Computational Statistics* **7**, 97-128.

[15] Hansen, B. E. and B. Seo (2002), Testing for Two-Regime Threshold Cointegration in Vector Error Correction Models, *Journal of Econometrics* **110**, 293-318.

[16] Johansen, S. (1988), Statistical Analysis of Cointegration Vectors, *Journal of Economics Dynamic and Control* **12**, 231-254.

[17] Kawaller, I. G., Koch, P. D. and T. W. Koch (1987), The Temporal Price Relationship between S&P 500 Futures and the S&P 500 Index, *Journal of Finance* **42**, 1309-1329.

[18] Li, Q. and S. Wang (1998), A Simple Consistent Bootstrap Test for a Parametric Regression Function, *Journal of Econometrics* **87**, 145-165.

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

[20] McDonald, R. L. (2001), Cross-Border Investing with Tax Arbitrage: The Case of German Dividend Tax Credits, *Review of Financial Studies* **14**, 617-657.

[21] Robinson, P. (1988), Root-N Consistent Semiparametric Regression, *Econometrica* **56**, 931-954.

[22] Roll, R., Schwartz, E. and A. Subrahmanyam (2007), Liquidity and the Law of One Price: The Case of the Futures-Cash Basis, *Journal of Finance* **62**, 2201-2234.

[23] Shyy, G., Vijayraghavan, V. 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.

[24] Stoll, H. and R. Whaley (1990), The Dynamics of Stock Index and Stock Index Futures Returns, *Journal of Financial and Quantitative Analysis* **25**, 441-468.

[25] Taylor, N., van Dyck, D., Franses, P. H. and A. Lucas (2000), SETS, Arbitrage Activity and Stock Price Dynamics, *Journal of Banking and Finance* **24**, 1289-1306.

[26] Theissen, E. (2002), Price Discovery in Floor and Screen Trading Systems, *Journal of Empirical Finance* **9**, 455-474.

[27] Theissen, E. (2005), Price Discovery in Spot and Futures Markets: A Reconsideration. Unpublished Manuscript. University of Bonn.

[28] Tse, Y. (2001) Index Arbitrage with Heterogenous Investors: A Smooth Transition Error Correction Analysis, *Journal of Banking and Finance* **25**, 1829-1855.

[29] van Dijk, D. and P. H. Franses (2000), Nonlinear Error-Correction Models for Interest Rates in the Netherlands, in *Nonlinear Econometric Modelling in Time Series: Proceedings of the Eleventh International Symposium in Economic Theory*, Barnett, W. A., Hendry, D., Hylleberg, S., Teräsvirta, T., Tjostheim, D. and A. Wurtz (eds), Cambridge University Press: Cambridge, 203-227.

[30] Yadav, P. K., Pope, P. F. and K. Paudyal (1994), Threshold Autoregressive Modelling in Finance: The Price Differences of Equivalent Assets, *Mathematical Finance* **4**, 205-221.

[31] Zheng, J. X. (1996), A Consistent Test of Functional Form via Nonparametric Estimation Techniques, *Journal of Econometrics* **75**, 263-289.

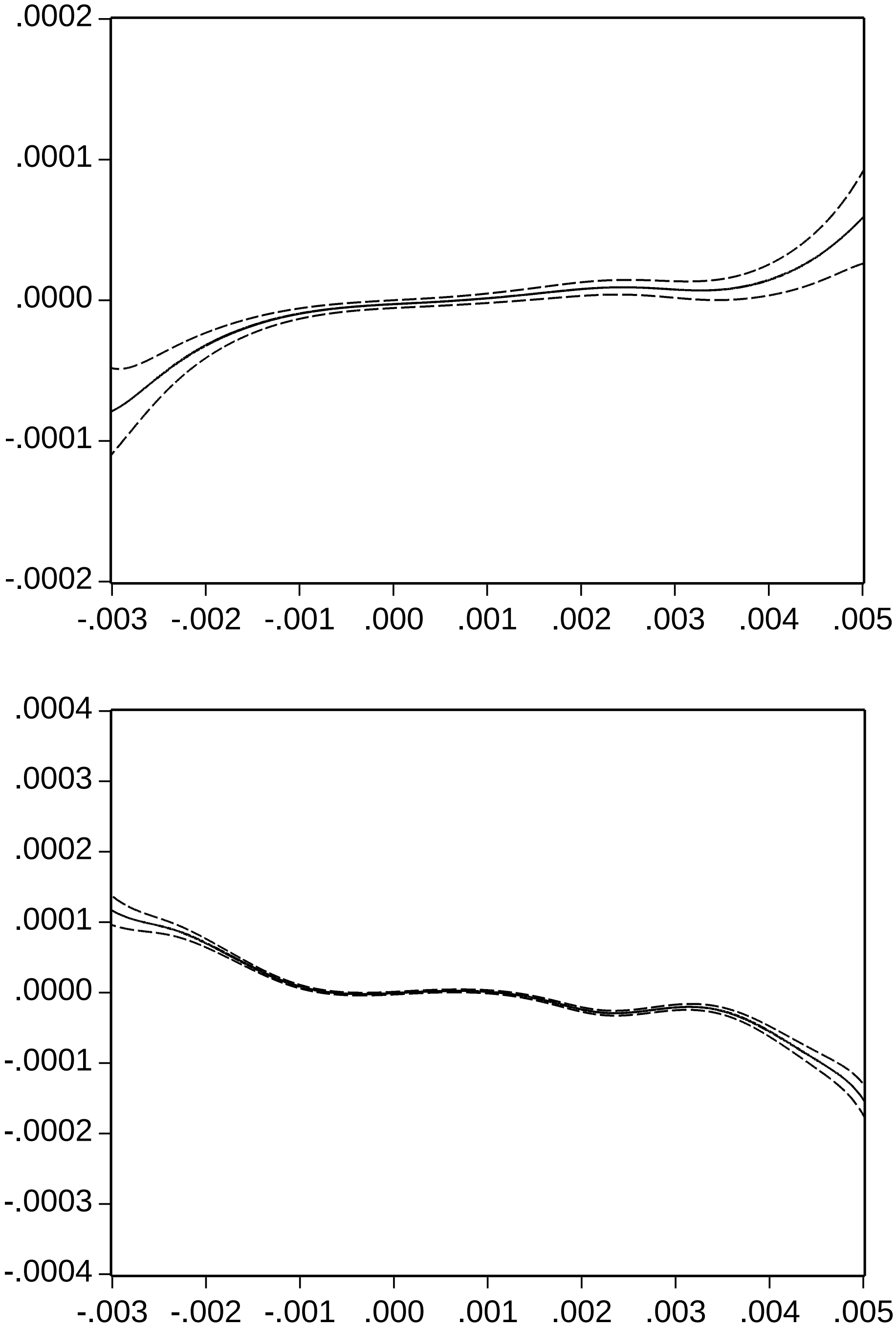

Figure 1: Estimated adjustment process (solid line) and pointwise 95% confidence interval (dashed line) for FDAX (upper panel) and XDAX (lower panel) as a function of the error correction term $p^X - p^F$. A Gaussian kernel and the bandwidths $h^F = 0.000492$ and $h^X = 0.000361$ have been used.

CFS Working Paper Series:

No.	Author(s)	Title
2008/11	**Roman Kräussl** **Niels van Elsland**	Constructing the True Art Market Index - A Novel 2-Step Hedonic Approach and its Application to the German Art Market
2008/10	**Alan Muller** **Roman Kräussl**	Do Markets Love Misery? Stock Prices and Corporate Philanthropic Disaster Response
2008/09	**Christopher D.Carroll** **Jirka Slacalek** **Martin Sommer**	International Evidence on Sticky Consumption Growth
2008/08	**Markus Haas** **Stefan Mittnik**	Multivariate Regime–Switching GARCH with an Application to International Stock Markets
2008/07	**Markus Haas** **Stefan Mittnik** **Mark S. Paolella**	Asymmetric Multivariate Normal Mixture GARCH
2008/06	**Charles Grant** **Christos Koulovatianos** **Alexander Michaelides** **Mario Padula**	Evidence on the Insurance Effect of Marginal Income Taxes
2008/05	**Dimitris Christelis** **Dimitris Georgarakos** **Michael Haliassos**	Economic Integration and Mature Portfolios
2008/04	**Elena Carletti** **Philipp Hartmann** **Steven Onega**	The Economic Impact of Merger Control Legislation
2008/03	**Annamaria Lusardi** **Olivia S. Mitchell**	Planning and Financial Literacy: How Do Women Fare?
2008/02	**Bannier** **Hirsch**	The Economics of Rating Watchlists: Evidence from Rating Changes

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