The Intraday Multivariate Structure of the Eurofutures Markets

Giuseppe Ballocchi, Michel M. Dacorogna,
Carl M. Hopman, Ulrich A. Müller,
Richard B. Olsen*

GBA.1997-11-25

September 11, 1998

A preprint by the O&A Research Group

This paper was presented at the Second International Conference on High Frequency Data in Finance, **HFDF - II**, 1-3 April 1998, in Zurich.

Abstract

We investigate the multivariate intraday structure in interest rates, focusing on implied forward rates from Eurofutures contracts. Since futures markets are the most liquid for interest rate instruments and they yield high-quality intraday data, it is somehow surprising that their intraday behavior has not been thoroughly studied in the literature.

We find interesting similarities with the foreign exchange market: scaling law, intraday patterns, all of which point to the heterogeneity of market participants. Other properties like asymmetric causal information flow between fine and coarse volatilities for the same time series are present in our data. There are also lead/lag correlation across maturities and currencies, but they that tend to disappear as markets mature.

A principal component analysis of the short end of the yield curve allows us to determine the most important components and to reduce the number of time series needed to describe the term structure. We find the decomposition rather stable over time. The first component which describes the curve level presents a HARCH effect while the remaining ones do not, having instead significant negative autocorrelations for the time series themselves. A HARCH model is applied to the first component and the impact of different market agents is discussed.

1 Introduction

The subject of interest rates can be approached from several different perspectives. Concepts such as money demand and supply, and the relation of interest rates to other macroeconomic factors have been prominent in the study of economics for decades. From a practitioners' viewpoint, the explosive growth of interest rate trading, together with the deregulation and globalization of the financial industry, have fueled the market for interest rate derivatives since the late seventies and early eighties. At around the same time interest rate models were introduced in finance which afforded the possibility to price interest rate contingent claims. It is somehow surprising that there has been little contact between the financial and the econometric literature on the subject, as pointed out by (Pagan et al., 1997). It is perhaps even more surprising that the intraday behavior of the interest rate markets has not been so far thoroughly investigated. Eurofutures contracts attracted less coverage in the academic literature than the US Treasury Bill futures, which seems odd, given than in 1997 the daily volume for just the Chicago Mercantile Exchange (CME) Eurodollar was about 450 times larger than for CME US Treasury Bill futures. This takes into account neither the Eurodollar futures traded in other markets, nor Eurofutures for many other currencies, while CME accounts for the entire US Treasury Bill volume ! The futures contract with the largest daily trading volume is the CME Eurodollar Time Deposit, with a daily volume of 461,098 contracts [1], with each contract corresponding to a notional one million dollar three-month deposit.

The bid/ask spread on the CME Eurodollar can be as small as half a basis point [2], compared with quoted spreads on Eurodollar deposits that are at least 10 times larger. Because of these small bid/ask spreads even tiny moves are tradable in the futures market, whereas they would

* **Olsen & Associates**, Research Institute for Applied Economics, Seefeldstrasse 233, CH–8008 Zürich, Switzerland; telephone +41-1-386-48-48, e-mail: bear@olsen.ch

[1] This is the averaged daily volume (all expiries combined) from January to September 1997, as reported in the January 1, 1998 issue of *Futures* magazine.

[2] A basis point corresponds to 1/100 of one percent, and its monetary value (in the case of three-month Eurodollar futures) is 25 dollars. The minimum price movement for the CME Eurodollar is half a basis point.

not be in the Eurodeposit market. This means that the futures market acts as a very sensitive measuring device for interest rates expectations. It is also transparent, because transaction prices are public knowledge, which is almost never the case for over-the-counter (OTC) markets. The interest rate futures markets are typically the most liquid markets for interest rate instruments, playing a crucial role in the price discovery mechanism [3]. They yield high-quality intraday data consisting of transaction prices, firm bid and ask quotes, and sometimes also volume information. Because of their liquidity and data availability, they are one of the best laboratories in which to investigate market microstructure models.

The goal of this paper is to provide elements for a model of the intraday behavior of the short-term futures markets, building upon the basic facts that emerge from an empirical study of those markets. The collection of basic facts summarized in the first part of this paper and presented in more detail in a companion paper (Piccinato et al., 1997) is a prerequisite for any detailed intraday model.

In order to extend the time series analysis beyond the lifetime of each contract, we construct time series of implied forward rates that are not affected by contract expiries. An empirical study of the implied forward rates is presented, focusing on the lead/lag correlation of returns and volatilities for each rate as well as across rates.

HARCH (Müller et al., 1997) models, using the convenient model formulation of (Dacorogna et al., 1997) (which we call EMAHARCH) are estimated for each rate. We then study the multivariate structure by examining the covariance matrix, and we find significant correlations that are stable in time and that are also relatively stable when changing the interval of observation from one day to intradaily.

Encouraged by the stability of the covariance structure, we perform a principal component analysis and we investigate the time series properties of the derived eigenvectors. We provide elements for the formulation of a parsimonious multivariate model, where the time series corresponding to the first eigenvector is modeled by EMAHARCH, whereas a simpler model is sufficient for the other eigenvectors.

This study shows that the Eurofutures markets present remarkable similarities with the other markets studied so far with high frequency data, such as the foreign exchange market (Guillaume et al., 1997; Andersen and Bollerslev, 1997).

2 Data sample and summary of basic facts

The data sample considered in this paper consists of intraday transaction prices for Eurofutures traded at CME [4], the London International Financial Futures Exchange (LIFFE) and the Singapore International Monetary Exchange (SIMEX). The results presented here come from the following contracts [5]:

1. Eurodollar Time Deposit from CME.
2. Three-Month Euromark and Three-Month Sterling from LIFFE. The other Eurofutures contracts traded at LIFFE (Three-Month Euroswiss, Three-Month Eurolira and Three-

[3] Futures markets dominate cash markets as far as price discovery is concerned, as shown by (Garbade and Silber, 1985) for commodity futures and by (Fung and Leung, 1993) for Eurofutures (the study is based on daily data).

[4] Financial futures are traded at the International Monetary Market Division (IMM) of the CME.

[5] The contract names used here are the ones specified by the exchanges.

Month ECU) were not studied at the same level of detail because of their lower liquidity and also shorter history (for the Eurolira).

3. Three-Month Eurodollar and Three-Month Euroyen from SIMEX.

The reason for a wide choice of markets and contracts is to uncover properties and formulate models that can lay claim to some degree of generality. The considered contracts are three-month Eurofutures with quarterly expiries (March, June, September and December). Serial expiry contracts [6] were not included, since typically they typically exhibit lower liquidity and have been introduced more recently.

Data prior to the end of 1996 was purchased from a variety of suppliers, and later data comes from the real-time collection at Olsen and Associates (O&A). Because the historical data always includes tick-by-tick transaction prices, no objections can be leveled at this study concerning the validity of the price information, which may be an issue for work relying on indicative (rather than firm) quotes, such as the ones typically coming from over-the-counter (OTC) markets (such as foreign exchange).

The acquisition and preparation of the data sample was a rather daunting task, whose complexity is easy to underestimate. Historical data from a number of data suppliers, each with their own data coding [7], was converted to a common format, with consistent time stamps expressed in Greenwich Mean Time. The same common format was used for the real-time data acquisition we developed at O&A. A number of data holes and other miscellaneous data problems was found by cross-checking different data sources. A filtering algorithm, along the lines of the one described in (Dacorogna et al., 1993), was developed, allowing an automatic rejection of outliers. In any filtering algorithm, a compromise must be achieved between the goal to eliminate outliers and the need not to reject valid prices. This papers errs on the conservative side by using a weak filter that is extremely unlikely to reject any valid price.

Strong intra-day and (intra-week) seasonalities are seen, similar to the famous U-shape for stock indices. The number of ticks can be substantial (a few thousand a day for CME Eurodollar). There is a strong day-of-the-week effect in the tick activity, with the activity on Fridays being almost double than on Mondays. The main difference with respect to FX (or cash interest rates), is that futures trading activity is localized in one market, or a few markets at most for some contracts.

2.1 Deterministic Volatility Analysis

An interesting question is whether there is a deterministic pattern in the volatility of the Eurofutures price, linked to the futures expiry. The answer can be found in (Piccinato et al., 1997), where the average daily volatility, estimated from intraday hourly returns, is studied as a function of the number of business days left before expiry. The results demonstrate a clear deterministic volatility pattern. First of all, the volatility tends to decrease when approaching expiry. This behavior is the opposite of what textbooks describe, which is the predicted behavior (Samuelson, 1965) observed in commodity futures. Another interesting pattern is represented by peaks of volatility that occur approximately every 90 days, before each quarterly expiry, although the contract in question is not expiring. A likely interpretation for these peaks is that they correspond to roll-over trading activity, i.e. traders switching from a given contract to the next

[6] Serial expiry contracts expire during months not included in the quarterly sequence.

[7] The rather ambiguous conventions that are sometimes used for the expiry month codes was one of the many sources of grief.

contract in the future. The goal of roll-over is to hold a position in the futures market with an (almost) constant time interval before expiry.

3 The construction of implied forward rates

Given that the time series of prices coming from a single contract is not sufficiently long to allow most statistical studies, it is necessary to construct longer samples joining several contracts together. There are several empirical prescriptions used by analysts and traders to join price histories corresponding to several futures contracts with successive expiries. Such prescriptions are typically based on roll-over schemes, i.e. they attempt to replicate the behavior of a trader holding a contract and switching ("rolling over") to the next one sometime before the expiry of the current contract. This approach is not well suited to Eurofutures, because several contracts with different expiries trade at the same time with comparable liquidity, unlike what happens in the commodity or bond futures markets, where basically only one contract (or at most two) are actively traded at any given time.

In the case of Eurofutures, the problem is inherently multivariate, because the interplay of several contracts with simultaneous high open interests cannot be neglected. The method followed in this paper is to infer the implied interest rates from the prices of futures and construct the corresponding term structure (yield curve). Futures prices alone are not sufficient to construct implied spot rates, because the Eurofutures market does not convey any information about the applicable spot rate for the period from the current time to the next futures expiry. To avoid the necessity to use data from other instruments, this study focuses on forward interest rates, with a minimum starting point of three months in the future.

There is an interesting application for the yield curve of implied interest rates: comparing at any point in time the curve derived from Eurofutures with the curve derived from other instruments (such as deposits or forward rate agreements), which allows an investigation of arbitrage opportunities.

3.1 The yield curve algorithms

To use the information from Eurofutures prices to construct a yield curve of forward (or spot) rates we need to overcome a timing problem: futures are defined in terms of the contract expiry date (a fixed date, every three months for the most important Eurofutures contracts (quarterly expiries)) and the maturity period of the underlying reference rate (a fixed time interval, which is three months for the instruments considered here), whereas the implied forward rates we intend to construct need to be defined in terms of fixed time intervals, not in terms of fixed calendar dates. Those time intervals (forward periods) can be written as $[t + \Delta T, t + \Delta T + \Delta M]$; they start at time $t + \Delta T$ (where t is the current time); we call ΔT the time-to-start for the implied forward rate and ΔM the rate maturity or simply the maturity [8]. The forward rates in this paper are labeled according to the market conventions for forward rate agreements: the IXJ forward rate (e.g. the 3X6 forward rate) is the forward rate quoted at time t and applicable

[8] Since there is no standard definition, the terminology can be quite confusing. For the sake of clarity, we use the term "expiry" to indicate the date and time when the futures contract expires (i.e. it is not traded any longer); we do not use the term "maturity" as a synonym of "expiry," but we reserve it to denote the length of the deposit period corresponding to spot or forward rates (in this paper the maturity is always three month). When we speak about "rates," we mean several different time series (e.g. with different time-to-starts or currencies), not several values within the same time series.

for the interval starting at time $(t + I)$ and ending at time $(t + J)$ (I and J are expressed in months); the corresponding time-to-start is I months and the maturity is $(J - I)$ months. We also sometimes use the notation (Im,Jm), which is equivalent to IXJ; e.g. (3m,6m) is equivalent to 3X6.

Using forward rates for fixed time intervals removes the systematic decrease in the deterministic volatility (see sect. 2.1) as we move closer to the expiry of a given futures contract (which is fixed in calendar time). On the other hand it can be expected that more subtle effects, such as the volatility peaks before each futures expiry, will still remain in the forward rates themselves.

Because there is no generally accepted method to construct yield curves, we used two different algorithms, and checked the stability of our results with respect to the choice of algorithm. These two different algorithms are representative of two alternative construction principles for yield curves:

1. The interpolation method: based on interpolations of rates between points on the expiry time axis. Here, we are using interpolation with polynomials of degree 2.
2. The segment method: based on an uninterrupted series of forward rates over short periods (segments).

An interesting by-product of this paper is the specification of a set of tests to study the possible biases introduced by yield curve algorithms.

The polynomial interpolation method is defined as follows. The continuously compounded forward interest rate ϱ is modeled as a function of the time T since "now"; T is actually the size of the time *interval* from the time when the quote was issued to the time point of interest. The annualized implied forward rate r corresponding to a futures quote can be expressed with the help of an integral over ϱ:

$$r \;=\; (1 - \frac{f}{100})\ 100\% \;=\; \frac{1\ \text{year}}{T_{\text{end}} - T_{\text{start}}}\ \{\exp[\int_{T_{\text{start}}}^{T_{\text{end}}} \varrho(T)\ \mathrm{d}T] - 1\}\ 100\% \tag{3.1}$$

where the forward period is from T_{start} to T_{end}. For a futures contract, we use $T_{\text{start}} = T_{\text{exp}}$, the expiry time, and $T_{\text{end}} = T_{\text{mat}}$, the "maturity time" terminating the maturity period: $T_{\text{mat}} = T_{\text{exp}} + \Delta M$, where the forward period ΔM is of 3 months here. Eq. 3.1 also shows the relation of the annualized rate r to the futures price f. Example: a value $r = 3\%$ implies a futures price of 97.00. The inverse formula is

$$\bar{\varrho} \;=\; \frac{\ln(1 + \frac{T_{\text{mat}} - T_{\text{exp}}}{1\ \text{year}}\ \frac{r}{100\%})}{T_{\text{mat}} - T_{\text{exp}}} \tag{3.2}$$

This gives the mean value of $\varrho(T)$ within the forward period. The function $\varrho(T)$ is constructed to consist of piecewise, continuously connected, quadratic polynomials:

$$\varrho(T) \;=\; a\,t^2 + b\,t + c\ , \quad \text{with } t \;=\; \frac{2\,T - T_{\text{exp}} - T_{\text{mat}}}{T_{\text{mat}} - T_{\text{exp}}}\ , \quad \text{for } T_{\text{exp}} \;\leq\; T \;\leq\; T_{\text{mat}} \tag{3.3}$$

The polynomial coefficients are determined by the requirement that eq. 3.1 must be fulfilled, so all quoted forward rates are reproduced by integration of eq. 3.3. The other requirement is the continuity of $\varrho(T)$ at the edge points T_{exp} and T_{mat}, where the forward period meets the forward periods of the neighbor contracts. The value of ϱ at the meeting point T_{exp} is determined from the four implied forward rates nearest to T_{exp} (in a regular sequence of three-month futures contracts) as follows:

$$\varrho(T_{\mathrm{exp}}) \quad = \quad \frac{9\ (\bar{\varrho}_2 + \bar{\varrho}_3) + \bar{\varrho}_1 + \bar{\varrho}_4}{16} \tag{3.4}$$

where $\bar{\varrho}$ is computed from eq. 3.2 and its index $(1 \ldots 4)$ indicates the position in the series of the four nearest forwards; $\bar{\varrho}_3$ refers to the forward from T_{exp} to T_{mat}, for example. Eq. 3.4 can be regarded as interpolated from a cubic polynomial going through the values $\bar{\varrho}_1 \ldots \bar{\varrho}_4$, located at the midpoints of the corresponding forward periods. The same equation, shifted by one forward period, leads to $\varrho(T_{\mathrm{mat}})$. If a contract at the edge ($\bar{\varrho}_1$ or $\bar{\varrho}_4$ in eq. 3.4) is not available from the data source, we extrapolate:

$$\bar{\varrho}_1 \quad = \quad \frac{3\ \bar{\varrho}_2 - \bar{\varrho}_3}{2} \tag{3.5}$$

and analogously for $\bar{\varrho}_4$. The numerical effect of an extrapolation error is small as $\bar{\varrho}_1$ and $\bar{\varrho}_4$ have little impact in eq. 3.4. By fulfilling all the discussed requirements, the coefficients of the polynomial $\varrho(T)$ of eq. 3.3 can be formulated:

$$a \quad = \quad \frac{3}{4}\ [\varrho(T_{\mathrm{exp}}) + \varrho(T_{\mathrm{mat}}) - 2\ \bar{\varrho}]\ , \tag{3.6}$$

$$b \quad = \quad \frac{1}{2}\ [\varrho(T_{\mathrm{mat}}) - \varrho(T_{\mathrm{exp}})]\ ,$$

$$c \quad = \quad \frac{1}{4}\ [6\ \bar{\varrho} - \varrho(T_{\mathrm{exp}}) - \varrho(T_{\mathrm{mat}})]$$

where $\bar{\varrho}$ refers to the forward period of the polynomial ($\bar{\varrho} = \bar{\varrho}_3$, using the indexing of eq. 3.4). Now we can compute the forward rate r or r_{ann} for any given forward period by inserting eqs. 3.2 - 3.6 in eq. 3.1 (but the integration may now extend over any interval from a T_{start} to a T_{end}, possibly covering several piecewise polynomials).

A consequence of the polynomial interpolation is the potential "overshooting" of the yield curve: if the forward rates implied by a series of Eurofutures have a maximum somewhere around medium-term expiries, an interpolated forward rate (for a period close to that maximum) may exceed all the underlying implied forward rates by the futures quotes. In the presented method, overshooting is however a local effect – distant contracts cannot influence the behavior of $\varrho(T)$. This is better than the case of the sometimes-proposed methods based on spline interpolation, where even very distant contracts have an influence on the local behavior.

The second yield curve computation method is a proprietary method used in real-time applications in the context of the interest rate service provided by the O&A Information System (OIS). The aim is to keep preconceived model assumptions on the yield curve at a minimum. The algorithm relies on an uninterrupted, internally stored sequence of continuously compounded forward rates for periods of ten days. The sequence of rates is updated with every new quote. There is hardly any limitation on the type of underlying interest rate instrument: three-month Eurofutures (also for "serial months" outside the quarterly expiry dates), one-month Eurofutures, short-term cash interest rates (to bridge the gap from "now" to the first Eurofutures expiry), etc. On the other hand, this method does not use polynomial interpolation and hence has no "overshooting" of the yield curve. The following computations are performed for each incoming quote:

1. The quote is expressed as a continuously compounded forward rate for its whole maturity period.
2. Because the maturity period of the new quote covers only a part of the segments on the maturity time axis, the old value of the quote is computed by extracting and cumulating the continuously compounded forward rates of all the covered segments, where the old rate is the mean of the forward rates of all affected segments [9].
3. The difference between the new continuously compounded rate (from the quote) and the old rate obtained from the old segment values is taken.
4. The continuously compounded forward rates of all affected segments are corrected by a constant shift proportional to the difference computed above (partially overlapping segments only with reduced weights). After this correction, the step (2) will return the new rate instead of the old one.

In both algorithms, the sequence of incoming quotes can be irregular in time and in the ordering of instruments.

Convexity corrections, which represent the difference between futures and forward contracts, and are due to the presence of margining arrangements for futures (Burghardt and Hoskins, 1995), are negligibly small in our case, because the futures contracts under consideration are never more than 18 months from expiry. Both of our algorithms neglect such corrections.

We have performed the analysis described in this paper with both algorithms and found the results to be in agreement, therefore adding robustness. The only noticeable difference is that the segment algorithm produces seasonalities in the return autocorrelation for the time series corresponding to the second and higher-order eigenvectors (see sect. 6.2). Because of this, we prefer the interpolation method and use it for all results presented in this paper.

3.2 Econometric characterization of the forward rate time series

It is known (and certainly not surprising) that Eurofutures prices are non-stationary (Fung and Leung, 1993), as are spot rates and forward rates (Pagan et al., 1997).

Since forward rates are (nearly) integrated (Pagan et al., 1997; Brenner et al., 1996), we focus on studying the return on continuously compounded forward rates:

$$R[\Delta t](t) = \rho(t) - \rho(t - \Delta t) \tag{3.7}$$

where $\rho = \frac{log(1+r\Delta M)}{\Delta M}$, r is the non-continuously compounded annualized forward rate applicable at time t for the forward period of length ΔM (time is measured in years; i.e. in our case $\Delta M = 0.25$). Because we study high-frequency data, we have the freedom to choose the interval Δt for the definition of the return from very short (intraday domain) to days or even weeks; we refer to Δt as the return interval.

[9] The partially overlapping segments at the beginning and the end enter the mean with only partial weights.

It should be noted that the volatility derived from the return defined in eq. 3.7 differs from the definition used by market practitioners, e.g. when pricing options on Eurofutures, or caps and floors. The reason is that to price such securities, a log-normal assumption is usually made for the corresponding forward rates, meaning that the input volatility for the Black-Scholes equation needs to be estimated from logarithmic returns. We did not adopt logarithmic returns, despite this market convention, because they are numerically less robust when rates are close to zero.

3.3 An appropriate time scale

In order to conduct the study, we need to introduce an appropriate business time scale, following the approach of (Dacorogna et al., 1993), with the very important qualification that the system under study here is multivariate. The ϑ-time scales were separately determined for every rate. [10] The application of a separate time scale transformation to every rate makes it impossible to perform synchronous multivariate observations in physical time for all the rates, which would prevent any multivariate analysis. Because the ϑ-time scales were found to be very similar for all the rates corresponding to the same Eurofutures, we use a common ϑ-time scale across all rates for each instrument, and determine it from the average activity for all those rates. The fits were performed using the standard O&A activity models (but for instance only the "European component" of the model is important for the LIFFE data).

This is the first time that a ϑ-time scale is used for multivariate analysis, and also for a localized market. The model suffers from several limitations:

1. The volatility changes at open/close are not sufficiently reproduced.
2. Day-of-the-week effects are not included.
3. Some exchange-specific holidays may differ from the country-wide holidays used in the ϑ-time model.

Those limitations however are immaterial for the conclusions of this study, as indicated by sensitivity analyses consisting of choosing different estimates for the ϑ-time scale parameters.

It has been reported in the literature (Ederington and Lee, 1993) that scheduled news releases have a substantial impact on the intraday volatility. The ϑ-time model could be easily extended in a natural way to incorporate both day-of-the-week effects and the impact of scheduled news releases. Both effects could be modeled as an additional seasonality, i.e. a scheduled news release would make the ϑ-time flow faster around the time of the announcement, according to the average impact of such an announcement [11].

It is easy to provide the order of magnitude of ϑ-time intervals for the futures markets, as compared to physical time. Considering that a ϑ-time interval of four years is by definition equal to a four-year physical time interval and that ϑ-time basically stops incrementing when the market is closed, the average ratio of a physical time interval to a ϑ-time interval is the ratio of the number of business days to the total number of days in a year, multiplied by the fraction of time the market is open in a business day, i.e. $\frac{365}{250}\frac{24}{8} = 4.38$, meaning that 4.38 ϑ-time hours elapse on average for every physical hour when the market is open.

[10] The mathematical details of the ϑ-time estimate are available from the authors upon request.

[11] In fact we could even take into account the different impact of each type of announcement.

4 The volatility of each forward rate

4.1 Autocorrelation

We study the autocorrelation function for the returns, the absolute values of the returns, and the square of the returns for implied rates corresponding to all the contracts listed in sect. 2, for different return intervals, using the ϑ-time scale (see sect. 3.3). Figure 1 presents the results for forward rates derived from the Three-Month LIFFE Euromark. The results are quite similar for all the rates, and are stable with respect to the interpolation method (performing a linear interpolation or taking the last available value of the rate (previous tick)):

- As to be expected from simple market efficiency arguments, the autocorrelation function for the return does not exceed the 95% confidence interval of a Gaussian random walk.
- The autocorrelation of the absolute value of the returns is larger than that of the square of the returns, in line with the results in the foreign exchange market (Müller et al., 1990; Dacorogna et al., 1993).
- The autocorrelation of the volatility seems to decay faster with increasing time-to-start for the forward rate. This result has implications for the EMAHARCH model and is discussed again in sect. 7.1.
- The rate of decay in the autocorrelation of the volatility varies for rates calculated from Eurofutures corresponding to different currencies.

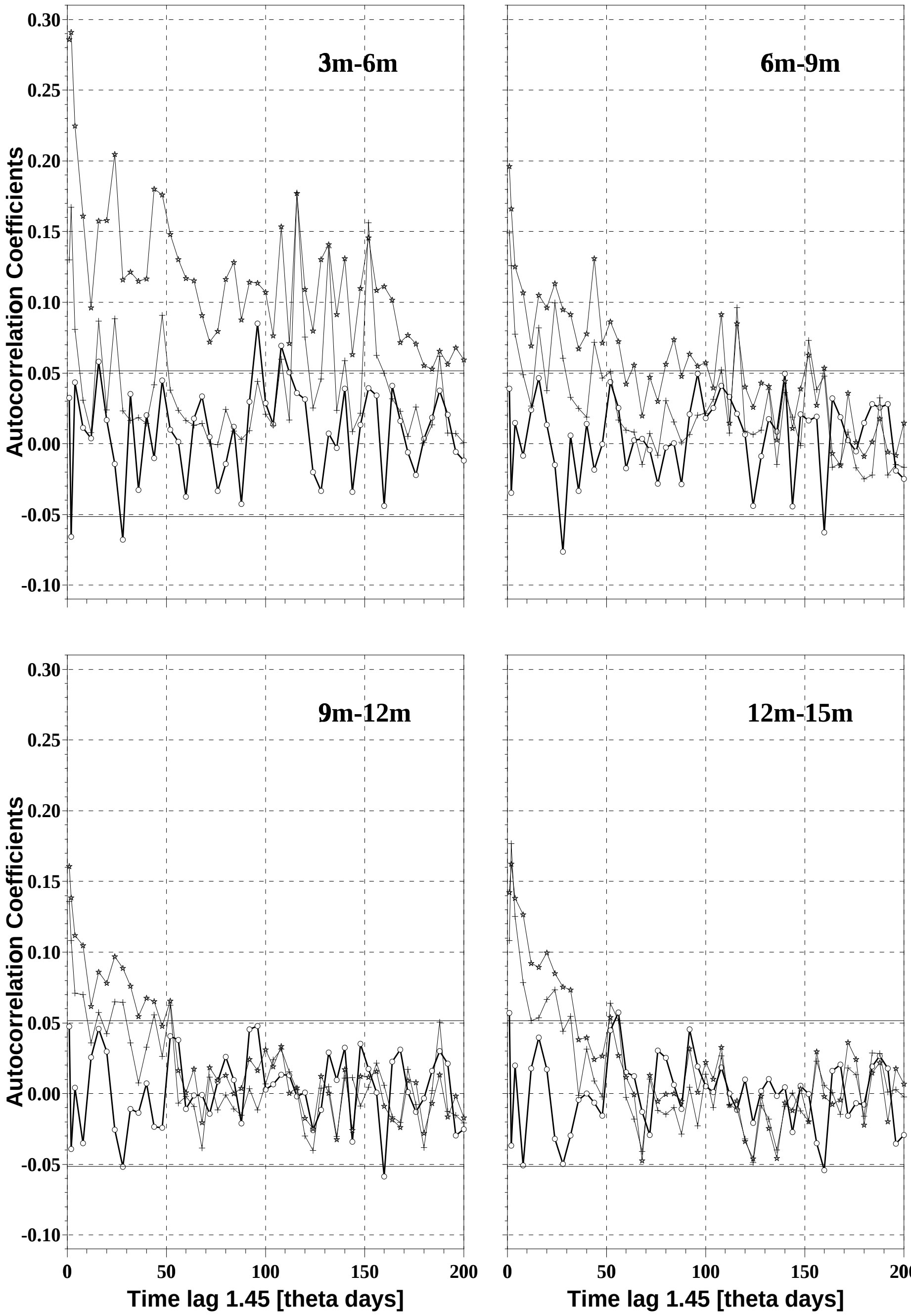
4.2 HARCH effect

Lagged correlation reveals causal relations and information flow structures in the sense of Granger causality. We perform such an analysis to look for the HARCH effect (Müller et al., 1997), i.e. the observation that in the foreign exchange markets the coarse-grained volatility predicts the fine-grained volatility better than the other way around. We find this effect for the implied forward rates derived from all the contracts listed in sect. 2. The effect is rather robust with respect to changes in the definition of the fine and coarse volatility. Figure. 2 shows the HARCH effect; the size of the effect seems to increase when increasing the time-to-start of the forward rate.

5 Lead-lag correlation study

The Eurofutures market is characterized by many instruments: several currencies, several contracts with different expiry dates for each currency. When studying these instruments, the question arises whether these instruments depend on each other.

A basic test for dependence is a linear correlation analysis of returns. In this study, we also investigate lead-lag correlations between different returns. This technique allows for identifying information flows from leading markets (instruments) to other instruments whose behaviors lag behind. The correlation studies are made on continuous time series over long samples, leading to significant results. In order to generate continuous time series, the raw Eurofutures data is converted to forward rates through the polynomial interpolation method as explained in



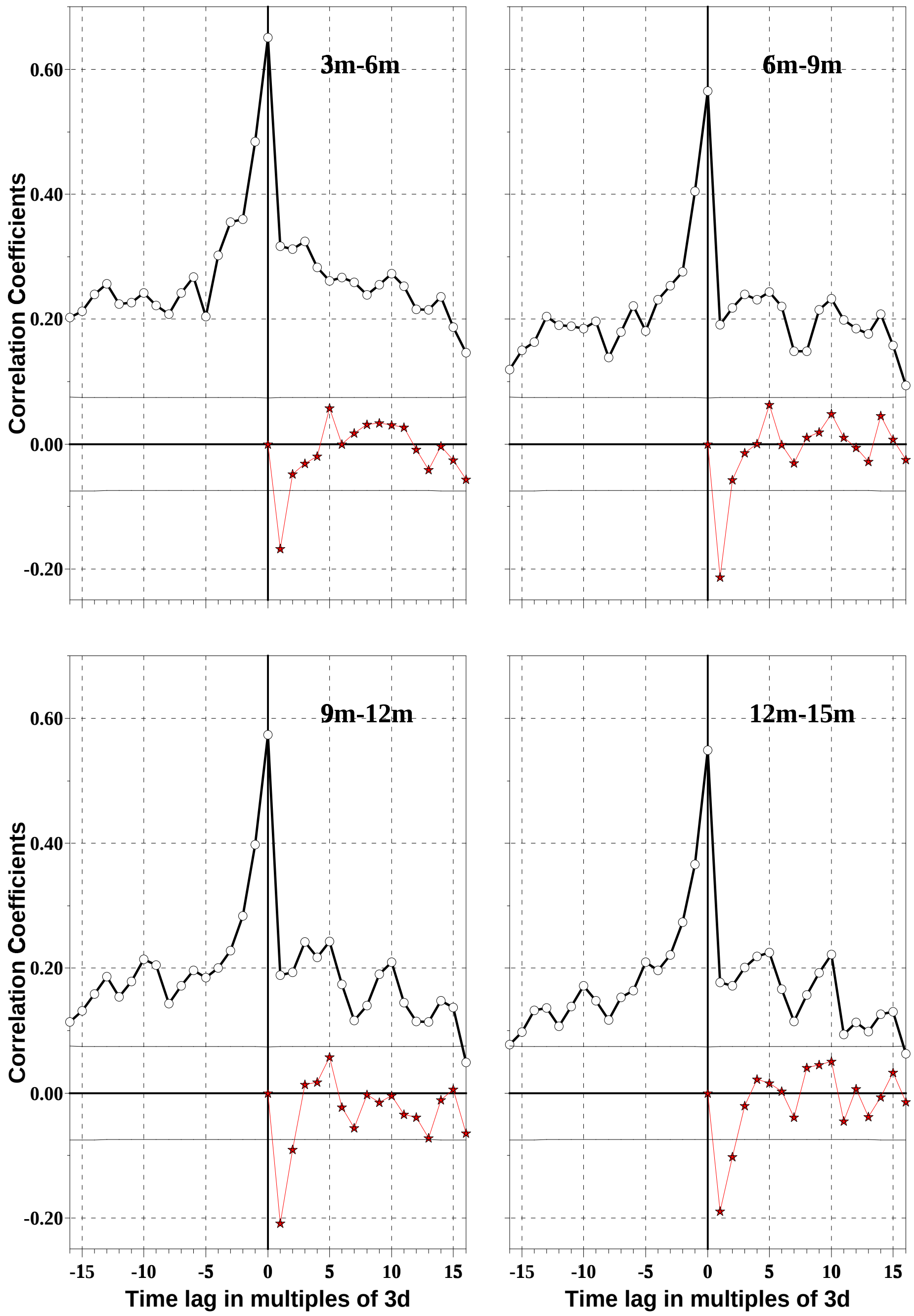
The autocorrelation functions of four different implied forward rates derived from the Three-Month LIFFE Euromark are shown on the same scale for the returns r_i (+), the absolute value of returns $|r_i|$ ($\star$) and the squared returns r_i^2 ($\circ$). The time scale is ϑ-time and the lag is 1.45 days which represents one business day. The horizontal lines represent the 95% confidence interval of a Gaussian random walk. The forward rate with the shortest time-to-start (3m-6m) has the longest volatility memory. The autocorrelation of the squares is significantly lower than for the absolute values. There is almost no significant autocorrelation in the returns themselves. (Sampling period: from April 6, 1992 to December 30, 1997, which represents 1445 independent observations.)

Figure 1: Autocorrelation

section 3.1. In addition to returns of the forward rates, absolute returns (which reflect volatility) are also investigated.

Lead-lag studies are done for forward rates corresponding to the same currency, but with different times-to-start (intra-currency analysis), for forward rates of different currencies (inter-currency analysis).

The maturity of the forward rates is always three months.



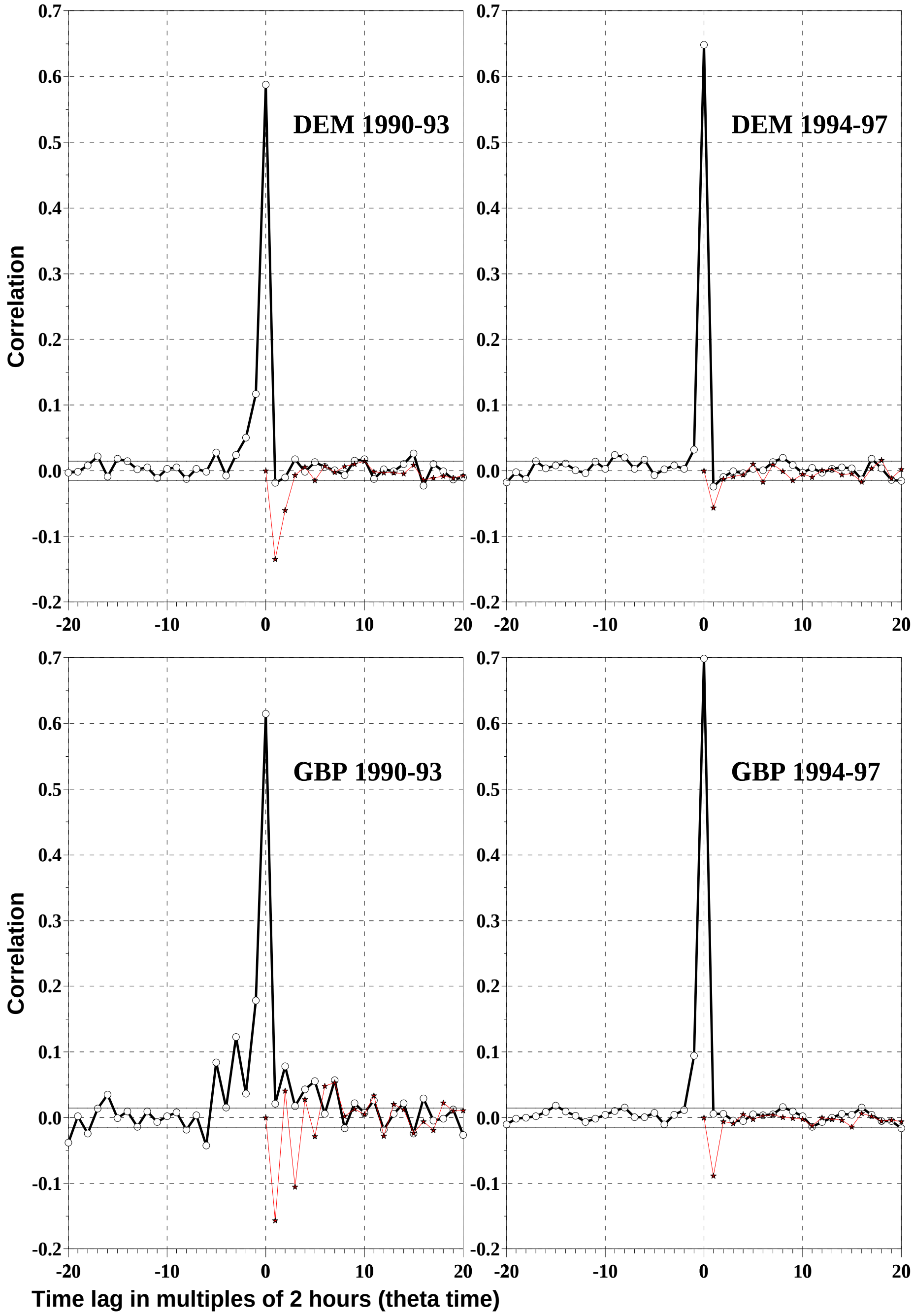
Lead/lag correlation of fine and coarse volatility for four different implied forward rates derived from the Three-Month LIFFE Euromark, with a three-hour grid in ϑ-time. *Fine volatility:* mean absolute return measured every three hours over three days.
Coarse volatility: mean absolute return over the whole three-day interval.
Thin curve: difference between correlations at positive and corresponding negative lags.
The band around the lag axis means 95% confidence for a Gaussian random walk. (Sampling period: from April 1, 1992 to December 30, 1997, which represents 700 independent observations.)

Figure 2: HARCH effect

5.1 Intra-currency lead-lag correlation

The forward rates as computed according to section 3 are analyzed in ϑ-time as explained in section 3.3. The ϑ-time scale has the effect of eliminating or at least strongly reducing intra-daily seasonalities of the volatility and the weekend effects. This is relevant because the lead-lag correlation analysis should not be overshadowed by seasonality.

The lead-lag study is done for two series: the 3X6 forward rate (with a forward period from 3 months to 6 months) and the 9X12 forward rate (with a forward period from 9 months to 1



Lead-lag correlation for the returns of 9X12 and 3X6 forward rates for different currencies and four-year periods. The thin curve indicates the asymmetry: the difference between correlations at positive and corresponding negative lags; negative values of this curve indicate that the 3X6 forward rate leads 9X12 forward rate. The confidence limits around zero correlation represent the 95% confidence interval of a Gaussian random walk.

The underlying Eurofutures data originates from LIFFE.

Figure 3: Asymmetric lead-lag correlation of 9X12 and 3X6 forward rates

year). This choice implies omitting the 6X9 forward rates (and also intermediary forward rates such as 4x7) and has two advantages: (1) the two forward rates are essentially based on different contracts (first and second position for the 3X6 forward rate, third and fourth position for the 9X12 forward rate), (2) the wide difference between the times-to-start makes the results more distinct. However, we also investigated the 6X9 forward rates and they always relate to the 3X6 and the 9X12 forward rates in the expected way.

The two time series, the 3X6 and the 9X12 forward rates, do not have the same quote frequency. The 9X12 forward rate is based on Eurofutures with more distant expiries that were rarely quoted in earlier years and that are still sparsely quoted today in some cases. In a lead-lag study, one could argue that a comparison between two time series with different quote frequencies is biased in favor of the more frequently quoted series. Therefore, we introduce the concept of time series *dilution*: we eliminate quotes from the more frequently quoted series so both series attain the same quote frequency. Example: we only take every second 3X6 forward rate quote if the 3X6 forward rates are quoted twice as frequently as the 9X12 forward rates. In reality, the dilution method is slightly more complex as the quote frequencies of the underlying Eurofutures have to be considered rather than those of the forward rates. The underlying Eurofutures quotes are exclusively those prices that have been used for real transactions. Fortunately, the lead-lag results are rather insensitive against changes in the dilution because the typical sampling frequency (one point per 2 ϑ-hours) is often much lower than typical raw data frequencies.

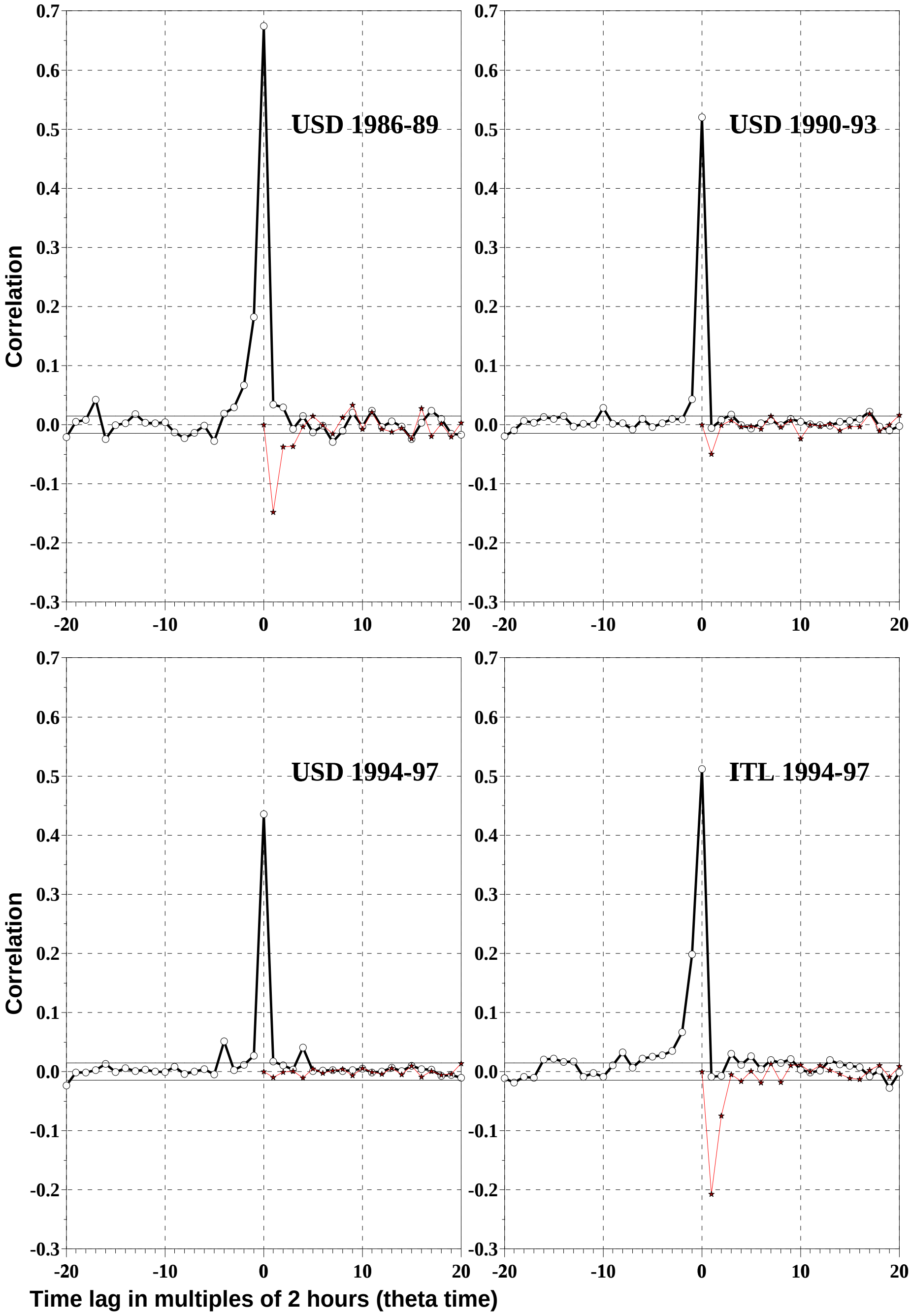
From the unequally spaced time series of forward rates, we generate equally spaced time series with a regular sampling frequency. Linear interpolation between original quotes is avoided because this could bias a lead-lag study. Instead, we assume that quotes are valid in the whole time interval from the quote time until the arrival of the next quote.

Figures 3 and 4 show a lead-lag analysis for returns observed over ϑ-time intervals of 2 hours, corresponding to about half an hour in physical time during the opening hours of the futures exchange. This choice is convenient; results obtained with other choices of the sampling interval do not provide essentially different information. The following phenomena are observed:

- A high simultaneous correlation between the 3X6 and the 9X12 forward rates. This is observed also for forward rates with other times-to-start and shows that there is a strong co-movement of all forward rates.
- A lead-lag effect at lags of 2 to 4 ϑ-hours ($\approx \frac{1}{2}$ to 1 hour of physical time when the futures exchange is open): the 3X6 forward rate leads the 9X12 forward rate. Wherever a thin line with stars in figures 3 and 4 is clearly below the 95% significance band, a lead-lag effect with asymmetric information flow from 3X6 forward rates to 9X12 forward rates is indicated. New movements in the rate tend to appear first for the rate with the shortest time-to-start and only afterwards for the rates with longer times-to-start. This effect is significant but weaker than the simultaneous correlation. It tends to weaken for the more recent samples as compared to the older ones.
- No significant correlation at larger positive or negative lags. In particular, there is no effect at lags of one full business day ($\approx$ 35 ϑ-hours).

The high simultaneous correlations among forward rates encourages us to identify principal components in the large set of future contracts, as it is done in sect. 6.

The significant lead-lag effect found in Figure 3 is in line with the degrees of liquidity found in a study of volume and open interest figures. The result is that the markets with the highest volume and open interest sizes lead the less-liquid ones. In earlier years, e. g. 1990-93, the volumes and open interest sizes of the contracts with distant expiries used to be small; in later years, e. g. 1994-97, they grew and came closer to those of of the contracts close to expiry. At the same time, the lead-lag effect became smaller. Figure 4 shows two extreme cases of the same phenomenon. The Eurodollar shows a very small lead-lag effect which has become entirely insignificant in the sample of 1994-97. We conclude that the CME Eurodollar market is more mature than the LIFFE Eurofutures markets because its trading activity is carried out evenly over a wider range of expiries. This is reflected also in the behavior of volume and open interest figures. By contrast, the LIFFE Eurolira shows an unusually strong lead-lag effect up to 1997. This market is less mature as most of its activity is concentrated on short expiries.



Lead-lag correlation (bold curve) for the returns of 9X12 and 3X6 forward rates for different currencies and four-year periods. The thin curve with stars indicates the asymmetry: the difference between correlations at positive and corresponding negative lags; negative values of this curve indicate that the 3X6 forward rate leads the 9X12 forward rate. The confidence limits around zero correlation represent the 95% confidence interval of a Gaussian random walk.

The underlying data consists of Eurolira futures from LIFFE for the plot on the lower right and of CME Eurodollar futures for the other three plots.

Figure 4: Asymmetric lead-lag correlation of 9X12 and 3X6 forward rates

Although significant lead-lag effects are in line with volume and open interest figures, they are not an artifact due to different quoting habits or frequencies. Significant lead-lag effects are real because a potential bias due to different quoting frequencies has been excluded through the dilution method and because all the underlying raw quotes are genuine transaction prices. Can the knowledge of such lead-lag effects be exploited in the form of arbitrage trades? If markets were frictionless, yes, but for Eurofutures with expiries further in the future, there might be some price slippage due to low liquidity which could erode potential arbitrage profits. Thus the arbitrage question cannot be answered with certainty.

In addition to returns of the forward rates, absolute values of returns (which reflect volatility) were also investigated. In many studies of empirical finance, absolute returns behave very differently from simple returns. However, this is not the case when forward rates with different times-to-start are compared. Here, we find the analogous lead-lag effect for absolute returns as for the simple returns – except that absolute returns have a small positive level of correlation at all lags, reflecting the clustering of volatility (heteroskedasticity). The investigation of absolute returns does not add much to our lead-lag analysis.

Significant correlations found at certain non-zero time lags reveal some interesting facts about information flows between contracts. However, they are usually much smaller than simultaneous correlations; they occur only at short lags (up to an hour or so during the opening time of the futures exchange) and they have decreased closer to insignificance in the liquid markets over the last years. Therefore, lead-lag effects are neglected in the modeling presented in sect. 6.

5.2 Inter-currency lead-lag correlation

The same lead-lag methodology can be use to determine information flows between the Eurofutures markets of different currencies. As in sect. 5.1, we artificially dilute the denser one of the two time series in order to arrive at an equitable comparison of time series with the same data frequency.

The only difference is related to the choice of the ϑ-time. The ϑ-time scales of the rates under comparison differ. Therefore, we perform the analysis twice, once choosing the ϑ-time for rates corresponding to the first currency, then again choosing that of rates corresponding to the second currency. The results show that the essential findings are the same in both cases and do not substantially depend on the choice of the ϑ-scale.

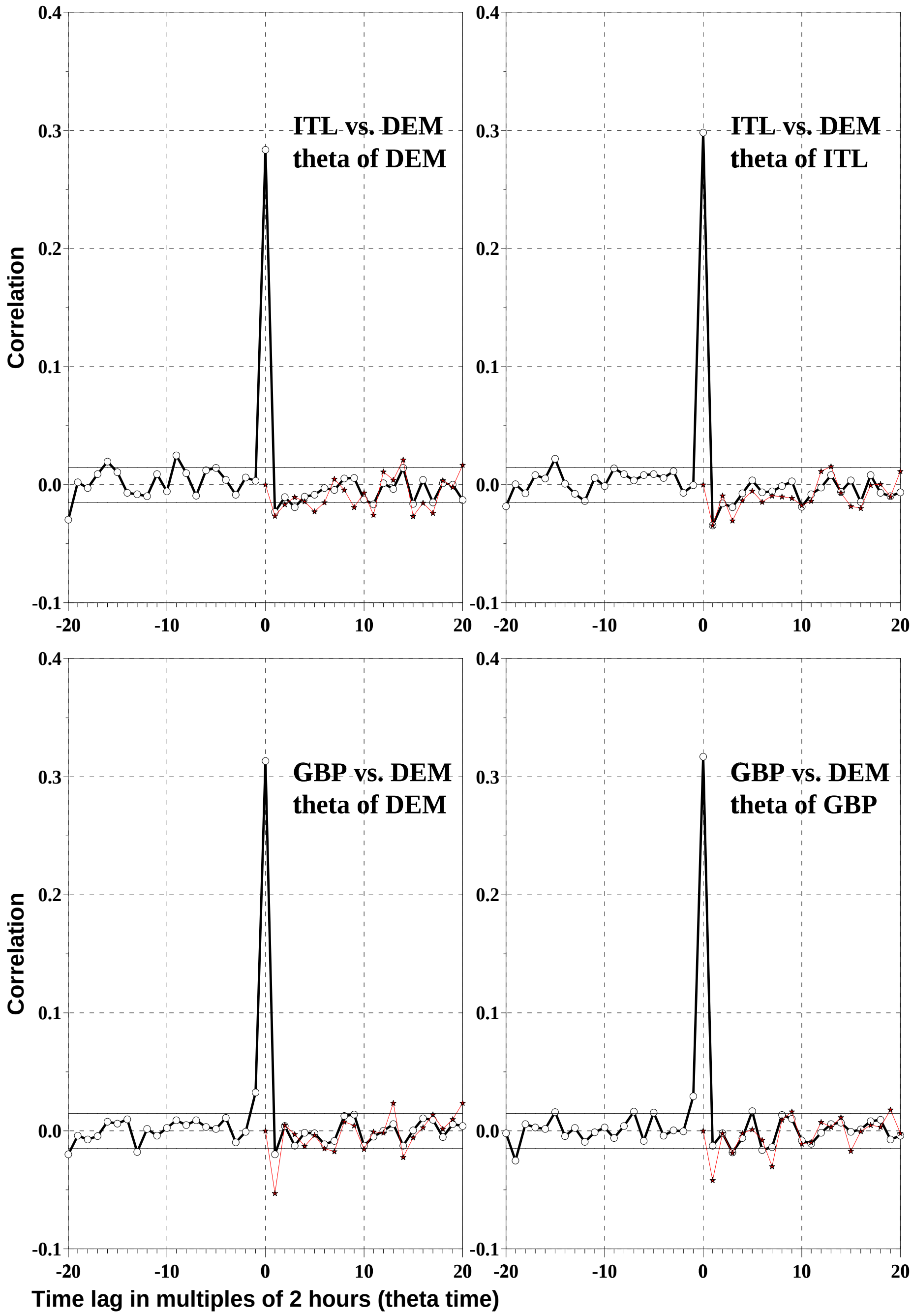
Figure 5 shows a lead-lag analysis of returns of the 3X6 forward rates: Eurosterling vs. Euromark and Eurolira vs. Euromark. The four-year period of 1994-97 is considered; the following phenomena are observed:

- A very significant simultaneous correlation between Eurosterling and Euromark and between Eurolira and Euromark. This correlation is less strong than in the intra-currency case but indicates a certain co-movement of major European Eurofutures traded at the same exchange (LIFFE).
- A significant but not overwhelming lead of Euromark futures over Eurosterling and Eurolira is observed. The effect is weak (especially for the Eurolira) and fades rapidly (especially for Eurosterling where it vanishes already at a lag of 4 ϑ-hours, i.e. at approximately 1 hour of physical time (when the futures exchange is open)).
- No significant correlation at larger positive or negative lags.

The obtained results are credible because they are not sensitive against changing the ϑ-time scale used, as can be seen when comparing the plots on the left and the right of Figure 5.

The positive simultaneous correlations indicate a worldwide co-movement of interest rate levels for different major currencies; this is supported also by positive correlations of CME Eurodollar futures to European Eurodollar futures[12]. However, no attempt is made in this paper to model a common factor across rates corresponding to different currencies.

[12] Because of the different opening hours of CME and LIFFE (different time zones), this correlation analysis was made only on a (working-)daily basis.



Lead-lag correlations of 3X6 forward rates for different currency pairs. Sample of four years: 1994-97. The thin curve indicates the asymmetry: the difference between correlations at positive and corresponding negative lags; negative values of this curve indicate that the first currency (always DEM, i. e. Euromark here) leads the second one. The confidence limits around zero correlation represent the 95% confidence interval of a Gaussian random walk.

Each analysis is done twice, using the ϑ-time scales of both involved instruments.

Figure 5: Asymmetric lead-lag correlation of 3X6 forward rates for different currencies

6 Principal component analysis

We concentrate now on the synchronous correlation between the considered forward rates. The small size and short duration of significant lead-lag correlation effects between forward rates return and volatility in mature markets, shown in Section 5.1, leads us to focus on the synchronous correlation structure. The results reported in this section and later are based on the following data samples of forward rates, constructed with the polynomial interpolation algorithm described in sect. 3.1: 3X6, 6X9, 9X12, 12X15 and 15X18 rates derived from the LIFFE Three-Month

(for 35 ϑ-hours)	3X6	6X9	9X12	12X15	15X18
3X6	1	0.983	0.963	0.944	0.935
6X9	0.983	1	0.987	0.971	0.958
9X12	0.963	0.987	1	0.988	0.974
12X15	0.944	0.971	0.988	1	0.989
15X18	0.935	0.958	0.974	0.989	1

(for 3 ϑ-hours)	3X6	6X9	9X12	12X15	15X18
3X6	1	0.956	0.910	0.889	0.868
6X9	0.956	1	0.946	0.908	0.881
9X12	0.910	0.946	1	0.934	0.889
12X15	0.889	0.908	0.934	1	0.933
15X18	0.868	0.881	0.889	0.933	1

Table 1: Correlation matrices between implied forward rates from LIFFE Euromark futures, using a measurement interval of 35 hours in ϑ-time corresponding to about one business day (matrix above) and an interval of 3 ϑ-hours corresponding to about 40 minutes of market opening time (matrix below). Computation sample: from April 6, 1992 until December 30, 1997

Euromark (starting from April 6, 1992), from the SIMEX Three-Month Eurodollar (starting from January 2, 1990), and from the SIMEX Three-Month Euroyen (starting from January 8, 1990). The end date for all the samples is December 30, 1997. We could not investigate with the same depth and the same degree of confidence forward rates further into the future (i. e. the 18X21 rate and above), because of the limited market liquidity for the underlying futures contracts at LIFFE and SIMEX. Results from 3X6, 6X9, and 9X12 forward rates derived from the CME Eurodollar, with a data sample starting on March 15, 1982, are in line with the findings from other rates, but are not reported in detail here; there we could not extend the study to 12X15 and 15X18 rates, because of holes in the data sample affecting expiries more than one year into the future.

By dividing the implied forward rate sample corresponding to a given instrument into several successive subsamples, we found that the intra-currency correlation is large and rather stable in time, whereas the inter-currency correlation is typically smaller and varies more in time.

Table 1 shows the correlation matrix between implied forward rates from LIFFE Euromark futures, using respectively a 35-hour interval in ϑ time (about one business day) and a three-hour ϑ-time interval (about 40 minutes in physical time) for the definition of the returns entering the calculation of the correlation. The correlation values are very high. Those of the 35-hour case are higher than those of the three-hour case (which are again higher than the corresponding simultaneous correlations in the two-hour case of Figure 3). This fading of correlation with small sampling interval sizes decreasing towards zero is in line with the results of (Lundin et al., 1998).

Principal component analysis (PCA) (Rao, 1997) provides a simple framework to study the correlation structure of the forward rates, with the advantage that no specific assumptions need to be introduced. Needless to say, this technique does not exploit non-linear dependences among the different forward rates, but any possible effects of this kind would be much smaller than the large observed linear correlation.

A typical aim of PCA is a dimensionality reduction, for instance in the context of derivative pricing models (Rebonato, 1996) and their empirical calibration. An easy way to have an intuitive grasp of the meaning of dimensionality reduction is to think of the N forward rates following an N-dimensional stochastic process, e.g. an N-dimensional Brownian motion; it is not relevant to the essence of this qualitative argument whether the real stochastic process is of another type. Let us take a system of two forward rates ($N = 2$): a large correlation between them means that the average angle between the two stochastic increments of the Brownian motion is small, in fact indicating that the knowledge of a single increment is almost sufficient to describe both, although strictly speaking, the two different increments are still needed in order to fully characterize the system.

Although the dimensionality reduction may be quite important when modeling the full yield curve (e.g. up to 30 years in the future for the US dollar), it is not the central aim of this work. Our goal is to to explore the existence of a set of latent variables, expressed as linear combinations of the original forward rates, that are better suited to capture the dynamics of the system. Instead of studying the original set of forward rate time series, we change the representation basis by diagonalizing the covariance matrix, thus arriving at a definition of eigenvectors, whose time series properties are investigated in sect. 6.2. We define linearly uncorrelated components by diagonalizing the covariance matrix between the different forward rates for the same currency (Rao, 1997). Because the inter-currency correlation is smaller and not as stable as the intra-currency correlation, the principal component analysis is applied on each set of forward rates derived from a single instrument, and not on all instruments and markets at once. The fact that trading is not conducted simultaneously in all the markets (in fact there is no time overlap between the business hours of CME and SIMEX) represents a formidable obstacle in attempting a global intraday PCA for all the time series under study.

The diagonalization is performed using different time intervals for the definition of the returns (we have used so far 3 ϑ-hours and 35 ϑ-hours (≈ 1 one business day); results for an interval of one week are also obtained).

The found eigenvectors are, as expected (Rebonato, 1996), in order of decreasing magnitude for the corresponding eigenvalues:

1. The first eigenvector is a sum of the original forward rates, with approximately the same positive weights, and can therefore be intuitively interpreted as the "average level" of the yield curve.

2. The second has opposite signs for maturities at the opposite end of the spectrum, with comparable magnitudes, and can therefore be interpreted as a measure of the slope of the yield curve. The corresponding eigenvalue is typically one twentieth of the first.

3. The third eigenvector is characterized by weights of the same sign at the opposite ends of the yield curve, with a weight of the opposite sign, and approximately twice as large at the center of the spectrum. This eigenvector can be deemed to represent a measure of the curvature of the yield curve, and its corresponding eigenvalue is typically one half of the second eigenvalue.

The qualitative picture that emerges is remarkably consistent across instruments and markets as demonstrated by table 2, which show the eigenvectors for the rates implied by the LIFFE Three-Month Euromark, by the SIMEX Three-Month Eurodollar and by the SIMEX Three-Month Euroyen. The diagonalization procedure also appears to be robust against changes of time periods (see table 3) and different return intervals (see e.g. table 4).

(LIFFE Euromark) Eigenvector number:	1	2	3	4	5
3X6	0.407	-0.464	-0.498	-0.357	0.494
6X9	0.436	-0.441	-0.114	0.189	-0.753
9X12	0.452	-0.194	0.609	0.475	0.403
12X15	0.466	0.334	0.407	-0.693	-0.161
15X18	0.472	0.664	-0.450	0.363	0.043
(SIMEX Eurodollar) Eigenvector number:	1	2	3	4	5
3X6	0.416	-0.402	-0.575	-0.458	0.354
6X9	0.448	-0.382	-0.071	0.238	-0.769
9X12	0.463	-0.271	0.447	0.496	0.516
12X15	0.463	0.306	0.539	-0.621	-0.124
15X18	0.444	0.725	-0.417	0.320	0.036

(SIMEX Euroyen) Eigenvector number:	1	2	3	4	5
3X6	0.334	-0.380	-0.489	-0.379	0.602
6X9	0.424	-0.465	-0.303	0.105	-0.708
9X12	0.501	-0.259	0.522	0.549	0.328
12X15	0.512	0.278	0.444	-0.661	-0.164
15X18	0.443	0.703	-0.447	0.328	0.042

Table 2: Eigenvectors for the implied forward rates for the LIFFE Three-Month Euromark (top), the SIMEX Eurodollar (middle) and the SIMEX Euroyen (bottom). The coefficients of the eigenvectors are listed in form of columns. The returns are always measured over intervals of three ϑ-hours.

This qualitative picture holds true for return intervals during the intraday period, although intraday correlations are typically smaller.

It should be emphasized that the interpretation for the PCA eigenvectors given above is heuristic in nature, but emerges without imposing any particular model assumption. If however the main intent was to explain the yield curve in terms of latent variables that are defined a priori (e.g. as the level, slope and curvature), the factor analysis (Rao, 1997) approach would be more suitable.

The stability of the diagonalization, both with respect to the choice of the sample and with respect to the time interval for the definition of the returns leads us to the conclusion that the diagonalization unveils fundamental characteristics of this multivariate system. Section 6.2 develops this point, through a detailed investigation of the time series properties of the eigenvectors.

6.1 A qualitative test of the dimensionality reduction with PCA

Here we intend to show qualitatively that by neglecting the eigenvectors corresponding to the smallest eigenvalues, not much information is lost. Table 7 shows the covariance matrix (the calculated matrix) for the implied forward rates for the LIFFE Three-Month Euromark, calculated from the first three eigenvectors, neglecting the last two, together with the original covariance

(First period) Eigenvector number:	1	2	3	4	5
3X6	0.362	-0.358	-0.296	-0.226	0.076
6X9	0.409	-0.448	-0.407	-0.265	0.630
9X12	0.457	-0.357	0.420	0.698	-0.014
12X15	0.499	0.276	0.608	-0.551	-0.033
15X18	0.495	0.683	-0.448	0.295	0.000

(Second period) Eigenvector number:	1	2	3	4	5
3X6	0.416	-0.640	-0.524	-0.291	0.239
6X9	0.443	-0.325	0.220	0.360	-0.721
9X12	0.457	-0.054	0.540	0.317	0.629
12X15	0.463	0.354	0.273	-0.747	-0.164
15X18	0.455	0.596	-0.557	0.357	0.019

Table 3: Eigenvectors for the implied forward rates for the the SIMEX Eurodollar for the period from January 2, 1990, until January 4, 1993 (top) and for the period from January 2, 1990, until December 30, 1997 (bottom). The coefficients of the eigenvectors are listed in columns. The returns are always measured over intervals of three ϑ-hours.

(LIFFE Euromark) Eigenvector number:	1	2	3	4	5
3X6	0.399	-0.653	-0.502	0.354	-0.191
6X9	0.430	-0.380	0.240	-0.536	0.571
9X12	0.450	-0.020	0.586	-0.075	-0.669
12X15	0.469	0.360	0.211	0.658	0.415
15X18	0.483	0.547	-0.550	-0.385	-0.130

Table 4: Eigenvectors for the implied forward rates for the the LIFFE Three-Month Euromark. The coefficients of the eigenvectors are listed in columns. The returns are measured over intervals of 35 ϑ-hours.

matrix. The values in the two matrices appear to be very close, although of course the calculated matrix is affected by multicollinearity because of the way it is constructed. Needless to say, the variances in the calculated matrix are smaller and the correlations larger than in the original matrix.

6.2 Time series properties of the eigenvectors

An original contribution of this paper is the combination of PCA with a time series study of the PCA eigenvectors, all being applied in the intraday domain with a ϑ-time transformation.

An intuitive understanding of the information content corresponding to each eigenvector can be gained by comparing its typical value (e.g. average of the absolute value) to the average

	DEM	JPY	USD
largest eigenvalue	5.899×10^{-7}	0.909×10^{-7}	8.847×10^{-7}
Ratio of largest to second largest eigenvalue	25.9	3.6	23.7
Ratio of second largest to third largest eigenvalue	1.9	2.1	2.0
Ratio of third largest to fourth largest eigenvalue	1.7	1.6	1.7
Ratio of fourth largest to fifth largest eigenvalue	1.7	1.6	2.0

Table 5: Eigenvalues for the implied forward rates, for a return interval of three ϑ-hours. DEM and GBP are the LIFFE Three-Month Euromark and Sterling; JPY and USD are the SIMEX Three-Month Euroyen and Three-Month Eurodollar.

	DEM	JPY	USD
largest eigenvalue	7.020×10^{-6}	1.478×10^{-6}	10.823×10^{-6}
Ratio of largest to second largest eigenvalue	58.1	9.6	52.9
Ratio of second largest to third largest eigenvalue	3.7	3.4	2.4
Ratio of third largest to fourth largest eigenvalue	2.9	2.2	2.4
Ratio of fourth largest to fifth largest eigenvalue	1.4	1.5	3.0

Table 6: Eigenvalues for the implied forward rates, for a return interval of 35 ϑ-hours. DEM and GBP are the LIFFE Three-Month Euromark and Sterling; JPY and USD are the SIMEX Three-Month Euroyen and Three-Month Eurodollar.

absolute return (e.g over a day): when going from the first to the last eigenvector the typical value decreases more than the corresponding average daily absolute change, suggesting a decrease in the signal/noise ratio.

We study the time series of returns for the eigenvectors, i.e. $R_i[\Delta t](t) = z_i(t) - z_i(t - \Delta t)$ (where R_i is the return time series for the ith eigenvector, and z_i is the ith eigenvector time series).

The intraday behavior of the returns corresponding to the different eigenvectors follows broadly the same patterns as for the forward rate returns, but the typical U-shaped distribution appears to flatten out in some cases when going to higher order eigenvectors.

The autocorrelation analysis for the returns of the eigenvectors shows that the autocorrelation function is comprised within the 95% significance level of a Gaussian random walk (i.e. $1.96/\sqrt{N}$, where N is the number of observations), as it is to be expected from naive market-efficiency arguments, except for the first lag autocorrelation, which appears to be significant. The first-lag autocorrelation is positive for the first eigenvector returns, but is negative for the remaining others, suggesting that there is a persistence in the co-movements, but opposite movements are dampened. This result holds for autocorrelations measured on three-hour ϑ-time intervals and on 35-hour ϑ-time intervals. This finding deserves further investigation.

The autocorrelation functions for the absolute value of the time-change for each eigenvector and the square of this change reveal the following characteristics:

- As noted in (Müller et al., 1990; Dacorogna et al., 1993) for FX rates, the autocorrelation

(Calculated matrix) Rates:	3X6	6X9	9X12	12X15	15X18
3X6	1.0568	1.1010	1.0705	1.0614	1.0892
6X9	1.1010	1.1675	1.1736	1.1605	1.1522
9X12	1.0706	1.1736	1.2566	1.2570	1.1953
12X15	1.0615	1.1605	1.2570	1.3273	1.3260
15X18	1.0892	1.1522	1.1953	1.3260	1.4357

(Original matrix) Rates:	3X6	6X9	9X12	12X15	15X18
3X6	1.0751	1.0819	1.0668	1.0752	1.0812
6X9	1.0819	1.1922	1.1678	1.1563	1.1556
9X12	1.0668	1.1678	1.2779	1.2320	1.2078
12X15	1.0752	1.1562	1.2320	1.3611	1.3085
15X18	1.0812	1.1556	1.2078	1.3085	1.4449

Table 7: The top matrix (calculated matrix) is the covariance matrix for the implied forward rates for the LIFFE Three-Month Euromark, calculated from the first three eigenvectors, neglecting the last two. The bottom matrix (original matrix) is the covariance matrix determined from the original time series of forward rates. All matrix elements (for both matrices) are given in units of 10^{-7}. The returns are always measured over intervals of three ϑ-hours.

of the absolute value is larger than the autocorrelation of the square and there is significant and persistent autocorrelation, with a decay that is slower than exponential.

- The autocorrelation of the absolute value of returns for the third and higher-order eigenvectors shows a seasonality, in the form of an approximately sinusoidal wave of decreasing amplitude with a three-month period (corresponding to the interval between futures expiry). Although the pattern is quite obvious, its order of magnitude is quite small (of the order of one basis point, as shown by a plot of the absolute value of the corresponding return over intervals of three ϑ-time hours), this small value being related to the small size of the higher-order eigenvectors. We have not yet established whether those seasonalities are an artifact of the yield curve algorithm, or whether perhaps they reflect a real dynamic effect coming from the deterministic volatility structure in the Eurofutures market described in sect. 2.1.
- There is also an intra-daily seasonality in the returns for the third and higher-order eigenvectors, showing perhaps that the ϑ-time scale derived from the average activity is not fully adequate, but again the order of magnitude of this effect is small.

An analysis of the scaling law for the mean absolute value of the eigenvector change shows that the results for the first eigenvector are in line with the results for the single forward rates (see table 8), whereas the scaling law for the higher-order eigenvector is characterized by a drift exponent (see (Müller et al., 1990) for the precise definition) that tends to the value of 0.5 expected from a Gaussian random walk.

The HARCH effect (see fig. 6) (namely that coarse-grained volatility predicts fine-grained volatility better than the other way around) is present in the time series corresponding to the

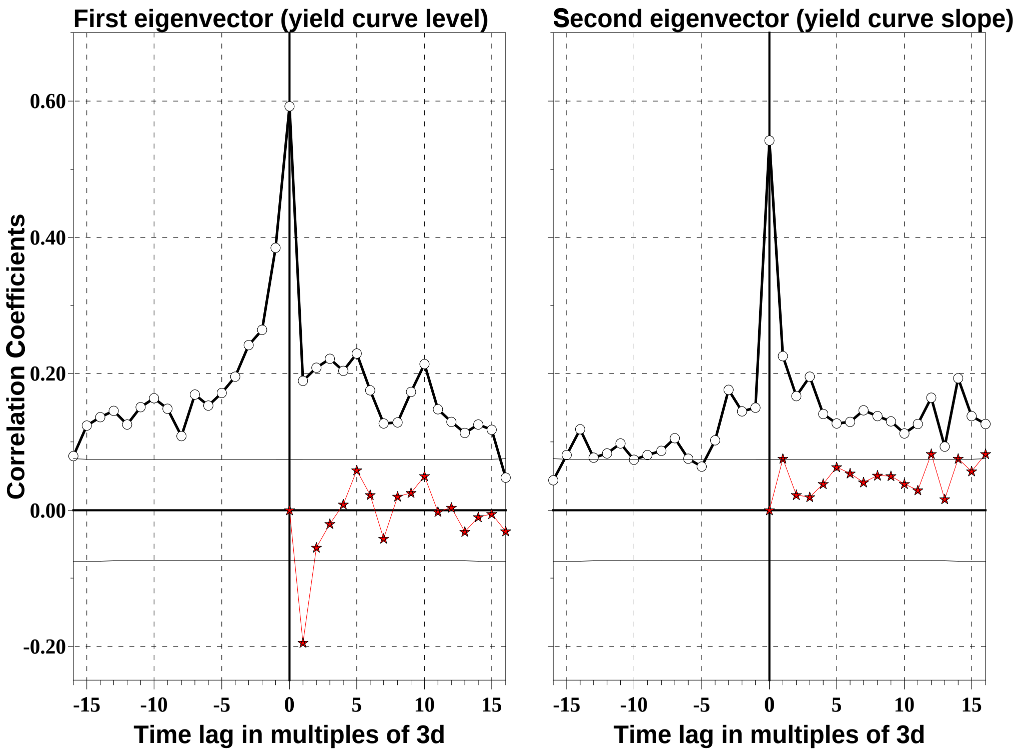


Figure 6: HARCH effect for eigenvectors

Lead/lag correlation of fine and coarse volatility for the time series corresponding to the first eigenvector (i.e. the eigenvector with the largest eigenvalue, corresponding to the yield curve level) and to the second eigenvector (i.e. the eigenvector with the second-largest eigenvalue, corresponding to the yield curve slope) for rates derived from the Three-Month LIFFE Euromark; a three-hour grid in ϑ-time is used. The time scale is ϑ-time. *Fine volatility:* mean absolute return measured every three hours over three days. *Coarse volatility:* mean absolute return over the whole three-day interval. Thin curve: difference between correlations at positive and corresponding negative lags. The band around the lag axis denotes the 95% confidence band for a Gaussian random walk. (Sampling period: 01.04.92 00:00:00 to 31.12.97 00:00:00 which represents 700 independent observations.)

3X6	6X9	9X12	12X15	15X18	v1	v2	v3	v4	v5
0.63	0.63	0.62	0.63	0.62	0.63	0.56	0.53	0.48	0.50

Table 8: Mean drift exponent for the scaling law of absolute price changes for different forward rates from LIFFE Three-Month Euromark and SIMEX Three-month Eurodollar and Three-Month Euroyen. v1, v2, v3, v4, and v5 label the eigenvector time series (from the largest eigenvalue to the smallest). The standard error on each value is of the order of 0.01.

first eigenvector, as shown in fig. 6, as it was for the single forward rates (see sect. 4.2). On the contrary there is no significant HARCH effect for higher-order eigenvectors.

The evidence presented so far in this section, when taken together, suggests that the PCA approach succeeds in concentrating a lot of the information from the N forward rates, all of which have non-trivial volatility dynamics (e.g. in terms of scaling law and HARCH effect) into the first eigenvector, which is the only one to show a clear HARCH effect, and has a drift exponent for the scaling law of absolute returns in line with the original forward rates, whereas the drift exponents for the other eigenvectors tend to the value expected from a Gaussian random walk.

7 Modeling the forward rates with a process equation

In the modeling of forward rates we would like to reproduce both the long memory of volatility and the HARCH effect, for which empirical evidence was presented in Section 4. HARCH (Müller et al., 1997) models, using the convenient model formulation of (Dacorogna et al., 1997) are estimated for each rate.

The EMAHARCH formulation consists in keeping in the process equation only a handful of representative interval sizes in the process equation and expressing the influence of the neighboring ones by an exponential moving average (EMA) of the returns measured on each interval, which has also the advantage of including a memory of the past intervals. The process is written in terms of partial volatilities σ_j^2, each of which can be regarded as the contribution of the jth component to the total market volatility σ^2. Here the volatility σ_j^2 is defined as the volatility observed over an interval of size $k_j \Delta t$. The EMAHARCH process equation is;

$$r(t) = \sigma(t)\, \varepsilon(t) \ ,$$

$$\sigma^2(t) = c_0 + \sum_{j=1}^{n} C_j\, \sigma_j^2(t) \tag{7.1}$$

where n is now the number of time components in the model. $\varepsilon(t)$ is a white noise with unit variance.

The partial volatility σ_j^2 has a memory of the volatility of *past* intervals of size $k_j \Delta t$ and can be written as:

$$\sigma_j^2(t) = \mu_j\, \sigma_j^2(t - \Delta t) + (1 - \mu_j) \left(\sum_{i=1}^{k_j} r(t - i\Delta t) \right)^2 \tag{7.2}$$

where k_j is the aggregation factor of the returns and takes n possible values following the relation

$$k_j = p^{j-2} + 1 \quad \text{for} \quad j > 1 \quad \text{with} \quad k_1 \equiv 1 \ . \tag{7.3}$$

The depth of the volatility memory is determined by the constant μ_j:

$$\mu_j = \mathrm{e}^{-\frac{\Delta t}{M(k_j \Delta t)}} \tag{7.4}$$

where the memory decay time constant of the component is given as the function M of the component's volatility interval $k_j \Delta t$. Instead of introducing new parameters for the characterization of $M(k_j \Delta t)$, it is simply chosen as

$$M(k_j \Delta t) = \frac{(k_{j+1} - k_j)\ \Delta t}{2} \ . \tag{7.5}$$

	3X6	6X9	9X12	12X15	15X18	v1
K_0	2.90 ± 0.15	4.54 ± 0.24	4.29 ± 0.24	6.45 ± 0.39	4.22 ± 0.37	18.86 ± 1.07
I_1	0.20 ± 0.01	0.17 ± 0.01	0.19 ± 0.01	0.18 ± 0.01	0.12 ± 0.01	0.17 ± 0.01
I_2	0.01 ± 0.01	0.00 ± 0.02	0.00 ± 0.02	0.01 ± 0.01	0.00 ± 0.02	0.02 ± 0.01
I_3	0.17 ± 0.02	0.16 ± 0.02	0.16 ± 0.02	0.15 ± 0.02	0.15 ± 0.02	0.13 ± 0.01
I_4	0.08 ± 0.02	0.11 ± 0.02	0.13 ± 0.02	0.14 ± 0.02	0.12 ± 0.02	0.16 ± 0.01
I_5	0.11 ± 0.02	0.15 ± 0.02	0.16 ± 0.02	0.15 ± 0.02	0.19 ± 0.02	0.14 ± 0.01
I_6	0.23 ± 0.02	0.22 ± 0.02	0.19 ± 0.02	0.08 ± 0.02	0.11 ± 0.02	0.11 ± 0.01
I_7	0.00 ± 0.01	0.03 ± 0.01	0.04 ± 0.01	0.02 ± 0.01	0.03 ± 0.01	0.00 ± 0.01
L	7.753	7.478	7.345	7.320	7.307	6.696

Table 9: Results of the EMAHARCH process estimate for three-hour ϑ-time intervals for the different forward rates and the first eigenvector (v1) for the LIFFE Three-Month Euromark [14]. L denotes the log-likelihood. The error given in the table is the standard error. The underlying data are from the LIFFE Three-Month Euromark. k_0 is equal to $c_0 \times 10^9$. Instead of the coefficients C_i (for $i > 0$), the corresponding impacts I_i are given. These provide a direct measure of the impacts of the market components on the HARCH variance. The market components are those defined in (Dacorogna et al., 1997) for EMAHARCH, with $p = 4$. The sum of the impacts is smaller than one in all cases, meaning that the estimated processes are stationary. The distribution of the random variable $\varepsilon(t)$ is normal with zero mean and unit variance. Data sample: from April 6, 1992 to December 30, 1997, representing 16774 observations.

The memory is defined by the start and the end point of the component interval k_j. The impact I_j of each component is given by:

$$I_j \;=\; k_j C_j \;. \tag{7.6}$$

7.1 Modeling each forward rate

The EMAHARCH process is estimated for each forward rate by maximizing the log-likelihood function using a two-step procedure as described in (Dacorogna et al., 1997). The first step is to use a genetic algorithm (GA) search (Pictet et al., 1995); then as a second step the Berndt, Hall, Hall and Hausman (BHHH) algorithm (Berndt et al., 1974) is applied. The results of table 9 show a decreasing impact of the longer-term components (corresponding to the market actors with the longest time-horizon) going from the first forward rate (i.e. that whose time-to-start is closest in the future) to the last one (i.e. that whose time-to-start is furthest in the future), reflecting the decrease in the volatility autocorrelation noted in Section 4.1. A plausible economic interpretation of these findings can be suggested. The first forward rate is determined from futures contracts that are close to expiry, and therefore are close to settlement based on the LIBOR rates, which are strongly influenced by the reference rates set by central banks (which can of course be deemed to be market agents with a long time horizon) as well as actual activity in the Eurodeposit market. The other forward rates come from futures contracts whose trading is based more on expectations, and with higher volatility allowing more profit potential for short-term players, attracting them to preferentially trade those rates.

7.2 Elements for a multivariate model

Section 6.2 shows that principal component analysis affords the possibility to concentrate the information content of the multivariate forward rate system in the first eigenvectors, although of course the other eigenvectors also play a role. The consequence is that the modeling effort can be focused on the first component, which we model by means of a EMAHARCH process, whereas a simpler model can be used for the other components. The resulting model for the N forward rate is much more parsimonious than the separate EMAHARCH modeling of each rate, as done in Section 7.1. Here we outline the steps to construct a complete model for the yield curve under study:

1. Find the principal components by diagonalizing the covariance matrix, and checking the stability of the transformation, with respect to time and to the chosen time interval of observation.
2. Model the time series corresponding to the first eigenvector as a EMAHARCH process, and use a simpler process equation for the other eigenvectors. There is an open problem in the fact that it may be difficult to reproduce the variation of the volatility memory with the time-to-start of the forward rate (see 4.1) using just a single EMAHARCH process.
3. Derive each forward rate from the models for the eigenvectors in point (2), by inverting the diagonalizing transformation.

The estimate of the EMAHARCH process for the first eigenvector time series for forward rates derived from the LIFFE Three-Month Euromark is given in table 9.

8 Conclusions

This paper is the first to address the multivariate volatility structure in the Eurofutures markets and also the first one to do so with intraday data and to propose elements for a model of this structure.

Our purpose is twofold:

1. To attract the attention of researchers to intraday studies of the Eurofutures markets, given their economic importance and better data availability than that of OTC markets, following in the steps of our companion paper (Piccinato et al., 1997).
2. To propose elements for a model for this market, reproducing the multivariate behavior of the yield curve.

In order to be able to study the futures contracts by means of a time series uninterrupted by contract expiries, we have constructed implied forward rates using two different algorithms, and examined them using a business time transformation. Throughout the paper, we emphasize studying different Eurofutures markets and contracts to reach conclusions that are as general as possible, and not due to specific contract/market details.

We have found evidence of interesting lead-lag correlation effects, with the returns of shorter rates leading those of longer rates, extending to lags of several ϑ-time hours. Those effects are found in data samples extending up to the early 90's (and even now for Eurolira), but disappear in more mature markets, where there is considerable open interest and liquidity even for contracts with expiry a year or more in the future.

We have found, as in the FX case, asymmetric causal information flow between fine and coarse volatilities for the same rate (the HARCH effect).

We have estimated EMAHARCH models for each rate and found that the impact of the longer-term components (the market actors with the longest time horizon) decreases with increasing time-to-start for the forward rate. The EMAHARCH formulation, motivated by the presence of the HARCH effect and the long memory in the volatility autocorrelation, shows interesting similarities with the foreign exchange market concerning the heterogeneity of market participants.

The principal component analysis provides results that are stable across time, markets and instruments and manages to concentrate many of the interesting dynamic effects, such as the HARCH effect, in the first eigenvector. PCA gives us the possibility to outline a parsimonious model for the entire multivariate forward rate system, consisting of an EMAHARCH model for the first component and simpler models for the others. An interesting question arising here is how many principal components we need to model, i.e. which dimensionality reduction we can achieve without suffering a significant loss of information. Our suggestion for the determination of the model dimensionality is through out-of-sample tests, using a specific performance measure (or loss function) chosen on the basis of the foreseen usage of the model. Such a model, whose details still need to be elaborated, can find applications for volatility forecasting, with possible uses in risk management, hedging, and option trading.

This paper does not pretend to have the final word on the subject; quite on the contrary, it is the authors' hope that it may be the first in the series of many extensive intraday studies of the Eurofutures markets. We intend to pursue our research in the subject by finalizing the model and applying it to volatility forecasting. The other direction we will follow is to extend this study to longer maturities in the yield curve, using the information provided by other financial instruments, such as bond futures, which are characterized by larger intraday volatility than Eurofutures.

Acknowledgements

This paper would not have been possible without the team work of many associates at O&A; a few of the many contributions are listed here. Bernard Héchinger helped in interpreting the data, providing information about the futures market and sharing his trading experience. Several software engineers were instrumental in making the data accessible: Beat Christen, Dr. Silvano Maffeis, Dan Smith and Leena Tirkkonen. Barbara Piccinato and Dr. Rakhal Davé contributed to the data preparation and data access tools. Dr. Olivier Pictet provided the process estimation and simulation software. Greg Babb reviewed this manuscript.

We thank Dr. Mico Loretan from the Federal Reserve Board, who discussed this paper at HFDF-II.

We also thank Paul MacGregor from the Market Data Services Department of the London Financial Futures Exchange for his kind help in providing and interpreting data from LIFFE.

References

Andersen T. G. and Bollerslev T., 1997, *Intraday periodicity and volatility persistence in* *financial markets*, Journal of Empirical Finance, **4**(2-3), 115–158.

Berndt E., Hall B., Hall R., and Hausman J., 1974, *Estimation and inference in nonlinear structural models*, Annals of Economic and Social Measurement, **3**, 653–665.

Brenner R. J., Harjes R. H., and Kroner K. F., 1996, *Another look at models of the short-term interest rate*, Journal of Financial and Quantitative Analysis, **31**(1), 85–107.

Burghardt G. and Hoskins B., 1995, *The convexity bias in eurodollar futures: Part 1"*, Derivatives Quarterly, 47.

Dacorogna M. M., Müller U. A., Nagler R. J., Olsen R. B., and Pictet O. V., 1993, *A geographical model for the daily and weekly seasonal volatility in the FX market*, Journal of International Money and Finance, **12**(4), 413–438.

Dacorogna M. M., Müller U. A., Olsen R. B., and Pictet O. V., 1997, *Modelling short-term volatility with GARCH and HARCH models*, Internal document MMD.1997-01-08, Olsen & Associates, Seefeldstrasse 233, 8008 Zürich, Switzerland.

Ederington L. L. and Lee J. H., 1993, *How markets process information: News releases and volatility*, The Journal of Finance, (4), 1161–1191.

Fung H.-G. and Leung W. K., 1993, *The pricing relationship of eurodollar futures and eurodollar deposit rates*, The Journal of Futures Markets, **13**(2), 115–126.

Garbade K. D. and Silber W. L., 1985, *Price movements and price discovery in futures and cash markets*, Review of Economics and Statistics, **65**, 289–297.

Guillaume D. M., Dacorogna M. M., Davé R. D., Müller U. A., Olsen R. B., and Pictet O. V., 1997, *From the bird's eye to the microscope: A survey of new stylized facts of the intra-daily foreign exchange markets*, Finance and Stochastics, **1**, 95–129.

Lundin M., Dacorogna M. M., and Müller U. A., 1998, *Correlation of High Frequency Financial Time Series*, Financial Markets Tick by Tick, edited by Pierre Lequeux. John Wiley & Sons, Ltd., London, England.

Müller U. A., Dacorogna M. M., Davé R. D., Olsen R. B., Pictet O. V., and von Weizsäcker J. E., 1997, *Volatilities of different time resolutions – analyzing the dynamics of market components*, Journal of Empirical Finance, **4**(2-3), 213–239.

Müller U. A., Dacorogna M. M., Olsen R. B., Pictet O. V., Schwarz M., and Morgenegg C., 1990, *Statistical study of foreign exchange rates, empirical evidence of a price change scaling law, and intraday analysis*, Journal of Banking and Finance, **14**, 1189–1208.

Pagan A. R., Hall A. D., and Martin V., 1997, *Modeling the Term Structure*, volume 14 of *HandBook of Statistics*, Elsevier Science B. V.

Piccinato B., Ballocchi G., and Dacorogna M., 1997, *A closer look at the eurofutures market: intraday statistical analysis*, Internal document BPB.1997-08-25, Olsen & Associates, Seefeldstrasse 233, 8008 Zürich, Switzerland.

Pictet O. V., Dacorogna M. M., Chopard B., Oussaidène M., Schirru R., and Tomassini M., 1995, *Using genetic algorithms for robust optimization in financial applications*, Neural Network World, **5**(4), 573–587.

Rao C. R., 1997, *Principal Component and Factor Analyses*, volume 14 of *HandBook of Statistics*, Elsevier Science B. V.

Rebonato R., 1996, *Interest-rate options models*, Wiley.

Samuelson P. A., 1965, *Proof that properly anticipated prices fluctuate randomly*, Industrial management review, **6:2**, 41–49.