Scuola Normale Superiore

Perfezionamento in Matematica
Indirizzo in Scienze Finanziarie e Assicurative

Volatility estimate via Fourier analysis

Roberto Renò

Contents

Preface **7**

1 Theory of quadratic variation **11**

1.1 Quadratic variation 11

1.1.1 Preliminaries 11

1.1.2 Stopping times and subdivisions 12

1.1.3 Martingales 13

1.1.4 Quadratic variation for locally square-integrable martingales 15

1.1.5 Semimartingales 16

1.1.6 Stochastic integral 17

1.1.7 Quadratic Variation 19

1.1.8 Stochastic differential equations 24

1.2 Convergence to a martingale 26

1.3 Estimating quadratic variation 28

1.4 Quadratic variation in financial economics 29

2 Volatility estimate via Fourier Analysis **33**

2.1 Univariate case 33

2.2 Multivariate case 40

3 Univariate applications **41**

3.1 Introduction 41

3.2 Monte Carlo experiments 43

3.2.1 GARCH(1,1) process 44

3.2.2 SR-SARV(1) process 49

3.2.3 Volatility forecasting evaluation 50

3.3 Foreign exchange rate analysis 51

3.3.1 Forecasting daily exchange rate volatility using intraday returns . . . 56

3.4 A linear model for volatility 60

3.5 The volatility of the Italian overnight market 63

3.5.1 The data set 67

3.5.2 Testing the martingale hypothesis 68

3.5.3 Interest rate volatility and market activity 72

3.6 Conclusions 78

4 Multivariate applications **81**

4.1 Introduction 81

4.2 Performance on simulated data 82

4.3 Investigating the Epps effect 85

4.3.1 Monte Carlo experiments 86

4.3.2 Data analysis 90

4.4 Dynamic principal component analysis 95

4.4.1 Empirical Results 97

4.5 Conclusions 101

5 Nonparametric estimation **105**

5.1 Introduction 105

5.2 Nonparametric estimation of the diffusion coefficient 109

5.3 Small sample properties 116

5.4 Data Analysis 119

5.5 Conclusions 120

Bibliography **127**

Preface

The aim of this Thesis is to study some selected topics on volatility estimation and modeling. Recently, these topics received great attention in the financial literature, since volatility modeling is crucial in practically all financial applications, including derivatives pricing, portfolio selection and risk management. Specifically, we focus on the concept of *realized volatility*, which became important in the last decade mainly thanks to the increased availability of high-frequency data on practically every financial asset traded in the main marketplaces. The concept of realized volatility traces back to an early idea of Merton (1980), and basically consists in the estimation of the daily variance via the sum of squared intraday returns, see Andersen et al. (2003). The work presented here is linked to this strand of literature but an alternative estimator is adopted. This is based on Fourier analysis of the time series, hence the term *Fourier estimator*, which has been recently proposed by Malliavin and Mancino (2002). Moreover, we start from this result to introduce a nonparametric estimator of the diffusion coefficient.

The Thesis has two main objectives. After introducing the concept of quadratic variation and the Fourier estimator, we compare the properties of this estimator with realized volatility in a univariate and multivariate setting. This leads us to some applications in which we exploit the fact that we can regard volatility as an observable instead of a latent variable. We pursue this objective in Chapters 3 and 4. The second objective is to prove two Theorems on the estimation of the diffusion coefficient of a stochastic diffusion in a univariate setting, and this is pursued in Chapter 5.

The detailed structure of the work is the following.

In Chapter 1 we review basic facts about quadratic variation, thus we need to introduce semimartingales, stochastic integrals and stochastic differential equations. Moreover, some results on convergence of processes to semimartingales are sketched. The Chapter ends with a brief review of the literature concerning the issues that will be discussed in the subsequent chapters.

In Chapter 2 the Fourier estimator (Malliavin and Mancino, 2002) is introduced and proved to be consistent, both for the univariate and multivariate case. Details on the asymptotic distribution are given in the case of constant variance.

In Chapter 3 we deal with univariate applications. We first compare the Fourier estimator to alternative estimators, mainly realized volatility, on simulated data. Comparison is done according to precision in estimating the integrated variance and reliability in measuring the forecasting performance of a GARCH(1,1) model. Indeed Andersen and Bollerslev (1998a) pointed out that in order to thoroughly assess the forecasting performance of a volatility model, reliable volatility estimates are necessary. We repeat the analysis on exchange rate data. On real data, the high-frequency analysis is distorted by microstructure effects, so that we have to select a proper cut-off frequency, paralleling the choice of the grid in realized volatility measurement. We show that, when evaluating the forecasting performance of the GARCH(1,1), we get better results than those obtained with realized volatility. This part is mainly taken from Barucci and Renò (2002a,b). We then show that, treating volatility as an observable, we can simply model it via an auto-regressive process, and this simple model performs better than GARCH(1,1) and Riskmetrics (Barucci and Renò, 2002c). We then use these ideas to the analyze the Italian money market, following Barucci et al. (2004). The Italian money market, which can be viewed as a proxy of the Euro market for liquidity, displays several interesting features. We show that, as for the U.S. markets, the martingale hypothesis for the overnight rate has to be rejected. This is due to the fact that banks do not trade for speculative reasons, but only for hedging reasons. Efficient managing of reserves leads to the violation of the martingale hypothesis. We then estimate a model for volatility, which is estimated with a unique data set consisting of all transactions for the four years following EMU. We show that volatility, as for stock markets, has a strong autoregressive pattern, we find calendar effects and, most important, we show that volatility is linked to the number of contracts instead of trading volume or average volume.

In Chapter 4, we follow the same route on multivariate applications. We first show, on simulated data, that the Fourier estimator performs significantly better than the classical estimator in measuring correlations. The reason for this fact is inherent in the nature of correlation measurement. When two time series are not observed exactly at the same time, and some interpolation techniques have to be used, a downward bias is introduced in the absolute value of the measured correlation. With the Fourier estimator, we integrate the time series, and the interpolation rule is used only to compute integrals. With classical estimators, the interpolated time series is directly used for estimation. For the same reason, due to the uneven nature of high-frequency transactions when increasing frequency a downward bias in correlations is observed. This phenomenon is called *Epps effect*. Here, following Renò

(2003) we provide evidence for this effect and relate this to lead-lag relationships and non-synchronicity, both on simulated and real data. Finally, we depict a geometric interpretation of the time-dependent variance-covariance matrix (Mancino and Renò, 2002), and illustrate these ideas on a set of 98 U.S. stocks.

In Chapter 5, we turn to the problem of measuring the diffusion coefficient of an univariate diffusion where the diffusion depends on the state variable only. This problem can be approached in a parametric fashion, as well as with nonparametric techniques (Florens-Zmirou, 1993; Stanton, 1997; Ait-Sahalia, 1996a; Bandi and Phillips, 2003). Following the last strand, we introduce a non-parametric estimator of the diffusion coefficient which borrows from the theory of Chapter 2 (Renò, 2004). We derive the asymptotic properties of this estimator, which turns out to be consistent and asymptotically normally distributed (Theorems 5.6 and 5.9). We then compare our non-parametric estimator with those introduced in the literature. After studying, via Monte Carlo simulation, the small sample properties of the estimator. we use it to model the short rate, and we estimate the diffusion coefficient on several interest rate time series. The results obtained are in line with those in the literature, with some differences for large and small values of the short rate.

Many people contributed to his project with their work and ideas. I would like first to thank Mavira Mancino and Paul Malliavin for their help, suggestions, and trust.

Then I would like to acknowledge Emilio Barucci and Claudio Impenna, who co-authored most of the material presented in Chapter 3.

I would like to acknowledge the Department of Economics of University of Siena, and in particular Carlo Mari, Claudio Pacati and Antonio Roma, since they allowed me to fully pursue this project during my first years as an Assistant Professor at University of Siena.

I am also indebted to Scuola Normale Superiore, to the *Associazione degli Amici della Scuola Normale Superiore* and to Carlo Gulminelli for scientific and financial support. In particular I would like to thank Maurizio Pratelli, Anna Battauz and especially Marzia De Donno.

Finally, I would like to thank Rosario Rizza and Fulvio Corsi, as well as several anonymous referees.

Chapter 1

Theory of quadratic variation

In this introductory Chapter, we first define quadratic variation and its properties as a tool in stochastic process theory. We subsequently discuss, through a brief review of literature, the importance of the theory of quadratic variation in the financial literature.

1.1 Quadratic variation

1.1.1 Preliminaries

In what follows, we work in a filtered probability space $(\Omega, (\mathcal{F}_t)_{t\in\mathbb{R}^+}, P)$ satisfying the usual conditions, see Jacod and Shiryaev (1987); Protter (1990). We start by the definition of a process, and some other useful definitions.

Definition 1.1 *A process is a family $X = (X_t)_{t\in\mathbb{R}_+}$ of random variables from Ω to some set E.*

In applications, the set E will be usually $\mathbb{R}^d$. A process can be thought as a mapping from $\Omega \times \mathbb{R}_+$ into E.

Definition 1.2 *A trajectory of the process X is the mapping $t \to X(\omega)$ for a fixed $\omega \in \Omega$.*

Definition 1.3 *A process X is called càdlàg if all its trajectories are right-continuous and admit left-hand limits. It is called* **càg** *if all its trajectories are left-continuous.*

If a process is càdlàg, we can naturally define two other processes, X_- and ΔX as follows:

$$X_{0-} = X_0, \quad X_{t-} = \lim_{s \to t} X_s \tag{1.1}$$

$$\Delta X_t = X_t - X_{t-} \tag{1.2}$$

If the trajectory is continuous in t, then $X_{t-} = X_t$ and $\Delta X_t = 0$.

Definition 1.4 *A process is said to be adapted to the filtration $\mathcal{F}$ if X_t is $\mathcal{F}_t$-measurable for every $t \in \mathbb{R}_+$.*

Since our aim is to study quadratic variation, it is important to identify processes of finite quadratic variation for every trajectory. We denote by $\mathcal{V}$ the set of all real-valued processes X that are càdlàg, adapted, with $X_0 = 0$ and whose each trajectory $X_t(\omega)$ has a finite variation over each finite interval $[0, t]$, which implies

$$Var(X) = \lim_{n \to \infty} \sum_{1 \leq k \leq n} \left| X_{tk/n}(\omega) - X_{t(k-1)/n}(\omega) \right| < \infty. \tag{1.3}$$

We then abbreviate by $X \in \mathcal{V}$ the fact that X is an adapted process with finite variation.

We end this subsection with the definitions of increasing and predictable processes.

Definition 1.5 *A process X is said to be increasing if it is càdlàg, adapted, with $X_0 = 0$ and such that each trajectory is non-decreasing.*

Definition 1.6 *The predictable σ-field is the σ-field on $\Omega \times \mathbb{R}_+$ that is generated by all càg adapted processes. A process is said to be predictable if it is measurable with respect to the predictable σ-field.*

1.1.2 Stopping times and subdivisions

The concept of stopping time is very useful in econometric analysis, since economic data are recorded at discrete points in time.

Definition 1.7 *A stopping time is a mapping $T : \Omega \to \mathbb{R}_+$ such that $\{\omega | T(\omega) \leq t\} \in \mathcal{F}_t$ for all $t \in \mathbb{R}_+$.*

Given a process X and a stopping time T, we define the stopped process as $X_t^T = X_{T\wedge t}$. Among other things, stopping times are necessary to introduce the localization procedure.

Definition 1.8 *If $\mathcal{C}$ is a class of processes, we define the localized class $\mathcal{C}_{loc}$ as follows: a process X belongs to $\mathcal{C}_{loc}$ if and only if there exists an increasing sequence T_n of stopping times such that $\lim_{n\to\infty} T_n = \infty$ a.s. and that each stopped process $X^{T_n} \in \mathcal{C}$.*

The sequence T_n is called a *localizing sequence*. It is clear that $\mathcal{C} \subset \mathcal{C}_{loc}$.

We call an *adapted subdivision* a sequence τ_n of stopping times with $\tau_0 = 0$, $\sup_{n\in\mathbb{N}} \tau_n < \infty$ and $\tau_n < \tau_{n+1}$ on the set $\{\tau_n < \infty\}$. Among subdivisions, we will consider the Riemann sequence.

Definition 1.9 *A sequence $\tau_{n,m}, m \in \mathbb{N}$ of adapted subdivisions is called a Riemann sequence if $\lim_{n\to\infty} \sup_{m\in\mathbb{N}} [\tau_{n,m+1} - \tau_{n,m}] = 0$ for all $t \in \mathbb{R}_+$.*

1.1.3 Martingales

Among processes, a very important role is played by martingales.

Definition 1.10 *A martingale is an adapted process X whose P-a.s trajectories are càdlàg such that every X_t is integrable and such that, for every $s \leq t$:*

$$X_s = \mathbb{E}\left[X_t \,|\mathcal{F}_s\right] \tag{1.4}$$

Definition 1.11 *A martingale X is square-integrable if $\sup_{t\in\mathbb{R}_+} \mathbb{E}[X_t^2] < \infty$.*

In the forthcoming analysis, an important role is played by two special classes of martingales: local martingales and locally squared-integrable martingales.

Definition 1.12 *A locally square-integrable martingale is a process that belongs to the localized class constructed from the space of square integrable martingales.*

Definition 1.13 *A local martingale is a process that belongs to the localized class of uniformly integrable martingales, that is of martingales X such that the family of random variables X_t is uniformly integrable.*

We obviously have that if a martingale X is locally squared-integrable, than it is a local martingale. The class of local martingale can be obtained by localization of the class of martingales also. Indeed we have the following:

Proposition 1.14 *Each martingale is a local martingale*

Proof. Let X be a martingale, and consider the sequence of stopping times $T_n = n$. Then, for every $t \in \mathbb{R}^+$, we have $X_t^{T_n} = \mathbb{E}[X_n|\mathcal{F}_t]$. Since the class of uniformly integrable martingales is stable under stopping, we have that X^{T_n} is uniformly integrable as well. □

Local martingales, that is martingales, can be decomposed in a continuous and discontinuous part. This concept will be very useful when defining quadratic variation.

Definition 1.15 *Two local martingales M, N are called orthogonal if their product MN is a local martingale. A local martingale X is called a purely discontinuous local martingale if $X_0 = 0$ and if it is orthogonal to all continuous local martingales.*

The following properties help the intuition:

Proposition 1.16
1. *A local martingale X is orthogonal to itself if and only if X_0 is square integrable and $X = X_0$ up to null sets*
2. *A purely discontinuous local martingale which is continuous is a.s. equal to 0.*
3. *A local martingale X with $X_0 = 0$ is purely discontinuous if and only if it is orthogonal to all continuous bounded martingales Y with $Y_0 = 0$.*
4. *A local martingale in $\mathcal{V}$ is purely discontinuous.*

Proof. 1. Let X be a local martingale such that X^2 is a local martingale. By localization, we can assume that X, X^2 are uniformly integrable, so that X is square integrable. Thus $\mathbb{E}[X_t] = \mathbb{E}[X_0]$ and $\mathbb{E}[X_t^2] = \mathbb{E}[X_0^2]$, and these fact imply $X_t = X_0$ a.s. 2. Is a consequence of point 1. 3. X is orthogonal to Y if and only if it is orthogonal to $Y - Y_0$. Since Y is continuous, $Y - Y_0$ is locally bounded, then the claim follows from localization. 4. See Jacod and Shiryaev (1987), Lemma I.4.14 (b). □

The concept of orthogonality, which can be proved to be equivalent to orthogonality in a suitable Hilbert space, allows the following decomposition:

Theorem 1.17 *Any local martingale X admits a unique (up to null sets) decomposition:*

$$X = X_0 + X^c + X^d \tag{1.5}$$

where $X_0^c = X_0^d = 0$, X^c is a continuous local martingale and X^d is a purely discontinuous local martingale.

Proof. See Jacod and Shiryaev (1987), Theorem I.4.18. □

We call X^c the *continuous part* of X and X^d its *purely discontinuous part.* We have also the following:

Proposition 1.18 *Let X, Y be two purely discontinuous local martingales such that $\Delta M = \Delta N$ (up to null sets). Then $M = N$ (up to null sets).*

Proof. Apply Theorem 1.17 to $M - N$. □

1.1.4 Quadratic variation for locally square-integrable martingales

We start defining the quadratic variation of two locally square-integrable martingales, see Definition 1.8. We first need the following:

Lemma 1.19 *Any predictable local martingale which belongs to $\mathcal{V}$ is equal to 0 a.s.*

Proof. See Jacod and Shiryaev (1987), Corollary I.3.16. □

Theorem 1.20 *For each pair M, N of locally square-integrable martingales there exists a unique, up to null measure sets, predictable process $< M, N > \in \mathcal{V}$ such that $MN - < M, N >$ is a local martingale.*

Proof. The uniqueness comes from Lemma 1.19. For the existence, see Jacod and Shiryaev (1987), Theorem I.4.2.

The process $< M, N >$ is called the predictable quadratic variation of the pair (M, N).

Proposition 1.21 *The following property holds: $< M, N > = < M - M_0, N - N_0 >$*

A fundamental example is the Wiener process.

Definition 1.22 *A Wiener process is a continuous adapted process W such that $W_0 = 0$ and:*

1. $\mathbb{E}[W_t^2] < \infty, \mathbb{E}[W_t] = 0, \forall t \in \mathbb{R}_+$
2. *$W_t - W_s$ is independent of the σ-field $\mathcal{F}_s$, $\forall\ 0 \leq s \leq t$.*

It can be proved that the Wiener process is Gaussian. The function $\sigma^2(t) = \mathbb{E}[W_t^2]$ is called the variance function of W_t. If $\sigma^2(t) = t$ then W is called a *standard Wiener process*. In the literature, the Wiener process is also called a Brownian motion. For a proof of the existence of the Wiener process, see Da Prato (1998). The most important properties of the Wiener process can be found in Karatzas and Shreve (1988).

We can now prove the following proposition about the quadratic variation of the Wiener process, which is a locally square-integrable martingale. The result is very intuitive:

Proposition 1.23 *If W is a Wiener process, then $< W, W >_t = \sigma^2(t)$.*

Proof. By Theorem 1.20, we have to prove that $X_t = W_t^2 - \sigma^2(t)$ is a local martingale. We have:

$$X_t - X_s = W_t^2 - W_s^2 - \sigma_t^2 + \sigma_s^2 = (W_t - W_s)^2 - 2W_s^2 + 2W_tW_s - \sigma_t^2 + \sigma_s^2.$$

Then $\mathbb{E}[X_t - X_s|\mathcal{F}_s] = 0$, hence the result. □

Note that $\sigma^2(t)$ is continuous, null at 0 and increasing.

1.1.5 Semimartingales

Let us denote by $\mathcal{L}$ the set of all local martingales M such that $M_0 = 0$.

Definition 1.24 *A semimartingale is a process X of the form $X = X_0 + M + A$ where X_0 is finite-valued and $\mathcal{F}_0$-measurable, $M \in \mathcal{L}$ and $A \in \mathcal{V}$ (see the discussion of equation 1.3). If there exists a decomposition such that A is predictable, X is called a special semimartingale.*

From the definition is clear that if $X \in \mathcal{V}$ then it is a semimartingale. Obviously the decomposition $X = X_0 + M + A$ is not unique, but if X is a special semimartingale then there is a unique decomposition with A predictable (Jacod and Shiryaev, 1987). Given that a semimartingale is the sum of a local martingale and a process of finite variation, we can naturally decompose it in a continuous and discontinuous part in the same fashion of Theorem 1.17:

Proposition 1.25 *Let X be a semimartingale. Then there is a unique (up to null sets) continuous local martingale X^c such that $X^{c,0} = 0$ and any decomposition $X = X_0 + M + A$ of type 1.24 meets $M^c = X^c$ up to null sets.*

Proof. It is enough to use Theorem 1.17 and Proposition 1.16(4).

We then follow the above terminology and call X^c the *continuous martingale part* of the semimartingale X. The following Proposition shows that all deterministic processes with finite variation are semimartingales:

Proposition 1.26 *Let $F(t)$ be a real-valued function on $\mathbb{R}_+$, and define the process $X_t(\omega) = F(t)$. Then X is a semimartingale if and only if F is càdlàg, with finite-variation over each finite interval.*

Proof. For the sufficiency, it is enough to use the definition of semimartingales. For the converse, see Jacod and Shiryaev (1987), Proposition I.4.28. □

1.1.6 Stochastic integral

If a process $X \in \mathcal{V}$, it is easy to define the integral of another process H with respect to X. We define the integral process $\int HdX$ by:

$$\int_0^T h_s dX_s(\omega) = \begin{cases} \int_0^t H_s(\omega) dX_s(\omega) & if \int_0^t |H_s(\omega)| d[Var(X)]_s(\omega) < \infty \\ +\infty & otherwise \end{cases} \tag{1.6}$$

This definition stems from the fact that, if $X \in \mathcal{V}$, then its trajectories are the distribution functions of a signed measure. We want now to define the stochastic integral when X is a semimartingale. In this case, the trajectories do not define a measure; for example, the Wiener process has infinite variation over each finite interval. Now consider a generic process X. The stochastic integral can be naturally defined for processes H such that $H = Y1_{[0]}$

where Y is bounded and $\mathcal{F}_0$-measurable, or $H = Y1_{]r,s]}$, where $r < s$ and Y is bounded and $\mathcal{F}_r$-measurable. In this case we can define:

$$\int_0^t H_s dX_s = \begin{cases} 0 & if\ H = Y1_{[0]} \\ Y(X_{s\wedge t} - X_{r\wedge t}) & if\ H = Y1_{]r,s]} \end{cases} \tag{1.7}$$

The distinctive property of semimartingales is that this definition can be extended to any locally bounded predictable process H if and only if X is a semimartingale. The feasibility of the extension for semimartingales is stated in the following theorem.

Theorem 1.27 *Let X be a semimartingale. Then the mapping 1.7 has an extension to the space of all locally bounded predictable processes H, with the following properties:*

1. *$G_t = \int_0^t H_s dX_s$ is a càdlàg adapted process*
2. *The mapping $H \to \int H dX$ is linear*
3. *If a sequence H^n of predictable processes converges pointwise to a limit H, and if $|H^n| \leq K$, where K is a locally bounded predictable process, then $\int_0^t H_s^n dX_s$ converges to $\int_0^t H_s dX_s$ in measure for all $t \in \mathbb{R}_+$.*

Moreover this extension is unique, up to null measure sets, and in iii) above the convergence is in measure, uniformly on finite intervals: $\sup_{s\leq t} |\int_0^s H_u^n dX_u - \int_0^s H_u dX_u| \to 0$.

A complete proof of the above Theorem can be found in Dellacherie and Meyer (1976). It is important to state the following properties, which we state without proof.

Proposition 1.28 *Let X be a semimartingale and H, K be locally bounded predictable process. Then the following properties hold up to null sets:*

1. *The mapping $X \to \int H dX$ is linear.*
2. *$\int H dX$ is a semimartingale; if X is a local martingale, then $\int H dX$ is a local martingale.*
3. *If $X \in \mathcal{V}$ then $\int H dX \in \mathcal{V}$ and it is given by (1.6).*
4. *$(\int H dX)_0 = 0$ and $\int H dX = \int H d(X - X_0)$.*

5. $\Delta(\int HdX) = H\Delta X$.

6. $\int Kd(\int HdX) = \int HKdX$.

The stochastic integral of a predictable process that is left-continuous can be approximated by Riemann sums. Consider a subdivision τ_n. Then the τ-Riemann approximant of the stochastic integral $\int HdX$ is defined as the process $\tau(\int HdX)$ defined by

$$\tau\left(\int HdX\right)_t = \sum_{n\in\mathbb{N}} H_{\tau_n}(X_{\tau_{n+1}\wedge t} - X_{\tau_n\wedge t}) \tag{1.8}$$

We than have the following:

Proposition 1.29 *Let X be a semimartingale, H be a càg adapted process and τ_n a Riemann sequence of adapted subdivisions. Then the τ_n-Riemann approximants converge to $\int HdX$, in measure uniformly on each compact interval.*

Proof. Consider $\tau_{n,m}$ and define H_n by

$$H_n = \sum_{m\in\mathbb{N}} H_{\tau_{n,m}} 1_{]\tau_{n,m},\tau_{n,m+1}]} \tag{1.9}$$

Then H_n is predictable, converges pointwise to H, since H is càg. Now consider $K_t = \sup_{s\leq t}|H_s|$. The process K is adapted, càg, locally bounded and $|H_n| \leq K$. Then the result follow from Theorem 1.27 and from the property $\tau_n(\int HdX) = \int H_n dX$. □

1.1.7 Quadratic Variation

We now define the quadratic variation of two semimartingales, and state its most important properties.

Definition 1.30 *The quadratic variation of two semimartingales X and Y is defined by the following process:*

$$[X,Y]_t := X_tY_t - X_0Y_0 - \int_0^t X_{s-}dY_s - \int_0^t Y_{s-}dX_s \tag{1.10}$$

From the definition itself, it is straightforward to verify the following properties:

Proposition 1.31 *The quadratic variation of two semimartingales X, Y has the following properties:*

1. $[X, Y]_0 = 0$
2. $[X, Y] = [X - X_0, Y - Y_0]$
3. $[X, Y] = \frac{1}{4}\big([X + Y, X + Y] - [X - Y, X - Y]\big)$ *(polarization)*

The following analysis is crucial for at least two reason. First, the name quadratic variation comes after Theorem 1.32. Second, it is the basis for realized volatility, a concept which will be illustrated in the following chapters. Indeed, it allows an estimation of quadratic variation.

Theorem 1.32 *Let X and Y be two semimartingales. Then for every Riemann sequence $\tau_{n,m}$ of adapted subdivisions, the process $S_{\tau_n}(X, Y)$ defined by:*

$$S_{\tau_n}(X, Y)_t = \sum_{m \geq 1} \left(X_{\tau_{n,m+1} \wedge t} - X_{\tau_{n,m} \wedge t}\right) \left(Y_{\tau_{n,m+1} \wedge t} - Y_{\tau_{n,m} \wedge t}\right) \tag{1.11}$$

converges, for $m \to \infty$, to the process $[X, Y]_t$, in measure and uniformly on every compact interval.

Proof. By polarization, it suffices to prove the claim for $X = Y$. From equation (1.8) we get:

$$S_{\tau_n}(X, X) = X^2 - X_0^2 - 2\tau_n \left(\int X_- dX \right).$$

The last term converges to $\int X_- dX$ by Proposition 1.29, then $S_{\tau_n}(X, X)$ converges to $[X, X]$. □

An immediate consequence of Theorem 1.32 is that the quadratic variation of the Wiener process is $[W, W]_t = \sigma^2(t)$.

Let us provide now useful properties of the quadratic variation:

Proposition 1.33 *Let X and Y be two semimartingales.*

1. $[X, Y] \in \mathcal{V}$.
2. $[X, X]$ *is increasing.*

3. $\Delta[X,Y] = \Delta X \Delta Y$.

4. *If T is a stopping time, then* $[X^T, Y] = [X, Y^T] = [X^T, Y^T] = [X,Y]^T$.

The property 3 *implies that if X or Y is continuous, then $[X,Y]$ is continuous as well.*

Proof. We prove the properties for $X = Y$, then we can generalize by polarization. 1. $[X,X]$ is càdlàg, adapted and with $[X,X]_0 = 0$, thus $[X,X] \in \mathcal{V}$. 2. Comes directly from Theorem 1.32, since $S_{\tau_n}(X,X)$ is increasing. 3. Using Proposition 1.28 (5) we get $\Delta[X,X] = \Delta(X^2) - 2X_- \Delta X$. Then, since $\Delta(X^2) = (\Delta X)^2 + 2X_- \Delta X$, we have $\Delta[X,X] = (\Delta X)^2$. 4. It is a simple consequence of Theorem 1.32. □

Proposition 1.34 *If X is a special semimartingale and $Y \in \mathcal{V}$ then:*

1. $[X,Y]_t = \int_0^t \Delta X_s dY_s$

2. $X_t Y_t = \int_0^t Y_{-s} dX_s + \int_0^t X_s dY_s$

3. *if Y is predictable, then* $[X,Y]_t = \int_0^t \Delta Y_s dX_s$

4. *If X or Y is continuous, then* $[X,Y] = 0$.

Proof. See Jacod and Shiryaev (1987), Proposition I.4.49. □

We now provide a very useful result.

Lemma 1.35 *Let X be a purely discontinuous square-integrable martingale. Then* $[X,X]_t = \sum_{s \le t} (\Delta X_s)^2$

Proof. This is Lemma I.4.51 in Jacod and Shiryaev (1987). □

Theorem 1.36 *If X and Y are semimartingales, and if X^c, Y^c denote their continuous martingale parts, then:*

$$[X,Y]_t = < X^c, Y^c >_t + \sum_{s \le t} \Delta X_s \Delta Y_s \tag{1.12}$$

Proof. We prove the theorem in the case $X = Y$, then polarization yields the result. We can use the decomposition of Proposition 1.25, $X = X_0 + X^c + M + A$, where $A \in \mathcal{V}$ and M is locally square-integrable and purely discontinuous. By localization we can assume that M is square-integrable. Then:

$$[X, X] = [X^c, X^c] + 2[X^c, M] + 2[X^c, A] + [M, M] + 2[M, A] + [A, A]. \tag{1.13}$$

We have $[X^c, X^c] =< X^c, X^c >$. Moreover we have $[M, M] = \sum(\Delta M_s)^2$ from 1.35, while $[A, A] = \sum(\Delta A_s)^2$ and $[M, A] = \sum \Delta M_s \Delta A_s$ from Proposition 1.34(1). Then the sum of the last three terms is $\sum(\Delta X_s)^2$. From 1.34, 4 we have $[X^c, A] = 0$. Finally, since X^c and M are orthogonal, then $< X^c, M >= 0$. But $[X^c, M]$ is continuous, by Proposition 1.33(3), thus it is equal to $< X^c, M >= 0$. This ends the proof. □

Corollary 1.37 *Let X, Y be local martingales. Then*

1. *$[X, Y] = 0$ if X is continuous and Y purely discontinuous.*
2. *$[X, Y] =< X, Y >= 0$ if X and Y are continuous and orthogonal.*
3. *Let H be a locally bounded predictable process. If X is continuous, then $\int H_s dX_s$ is a continuous local martingale. If X is purely discontinuous, then $\int H_s dX_s$ is a purely discontinuous local martingale.*

Proof. See Jacod and Shiryaev (1987), Corollary I.4.55. □

Maybe the most important application of quadratic variation in the field of stochastic processes is Ito's lemma. We state the univariate result, multivariate extension is straightforward. Both proofs can be found in Protter (1990), Chapter II.

Theorem 1.38 *Let X be a semimartingale and $f \in \mathcal{C}^2$. Then $f(X)$ is a semimartingale and:*

$$f(X_t) - f(X_0) = \int_0^t f'(X_{s-}) dX_s + \frac{1}{2} \int_0^t f''(X_{s-}) d[X, X]_s^c + \sum_{0<s\leq t} [f(X_s) - f(X_{s-}) - f'(X_s-)\Delta X_s] \tag{1.14}$$

The following Theorems provide a characterization of Wiener processes and provide a change of time result which will be useful in Chapter 5.

Theorem 1.39 *A stochastic process X is a standard Wiener process if and only if it is a continuous local martingale with $[X, X] = t$.*

Proof. The fact that, if W is a standard Wiener process then $[W, W] = t$ is already known. To show sufficiency, define $Z_t = exp(iuX_t + \frac{u^2}{2}t)$ for some $u \in \mathbb{R}$. Using Ito's lemma we get:

$$Z_t = 1 + iu\int_0^t Z_s dX_s + \frac{u^2}{2}\int_0^t Z_s ds - \frac{u^2}{2}\int_0^t Z_s d[X, X]_s = 1 + iu\int_0^t Z_s dX_s.$$

Then Z is a continuous complex local martingale, as well as any stopping of Z is a martingale. Then we have, $\forall u \in \mathbb{R}$:

$$\mathbb{E}\left[exp(iu(X_t - X_s))|\mathcal{F}_s\right] = exp\left(-\frac{u^2}{2}(s - t)\right)$$

hence X is a standard Wiener process. □

Theorem 1.40 *Let M be a continuous local martingale with $M_0 = 0$ and such that $\lim_{t\to\infty}[M, M]_t = \infty$ a.s. and $T_s = \inf_{t>0}[M, M]_t > s$. Define $\mathcal{G}_s = \mathcal{F}_{T_s}$ and $B_s = M_{T_s}$. Then B_s is a standard Wiener process with respect to the filtration $\mathcal{G}_s$. Moreover $[M, M]_t$ are stopping times for $\mathcal{G}_s$ and $M_t = B_{[M,M]_t}$.*

Proof. See Protter (1990), Chapter II, Theorem 41. □

We finally state the following Theorem which is due to Knight (1971). It allows to transform a vector of orthogonal square-integrable continuous martingales into a vector of independent Brownian motions via a suitable time change.

Lemma 1.41 *(Knight's Theorem) Let $M_1, \ldots, M_n$ be orthogonal square-integrable martingales, and consider the time changes:*

$$T_i(t) = \begin{cases} \inf_s[B_i, B_i]_s > t & if\ this\ is\ finite \\ +\infty & otherwise \end{cases} \tag{1.15}$$

Then the transformed variables:

$$X_i(t) = \begin{cases} B_i(T_i(t)) & if\ T_i(t) < \infty \\ B_i(\infty) + W_i(t - [B_i, B_i]_\infty) & otherwise \end{cases} \tag{1.16}$$

where $W_1, \ldots W_n$ is an n-dimensional Brownian motion independent of X_i, are an n-dimensional Brownian motion relative to their generated filtration.

1.1.8 Stochastic differential equations

In this subsection, we are concerned with the following equation:

$$X(t) = \eta + \int_0^t \beta(s, X(s))ds + \int_0^t \sigma(s, X(s))dW(s), \tag{1.17}$$

where $W(s)$ is the standard Wiener process, as defined in 1.22, and we look for an adapted process $X(t) \in L^2(\Omega)$. The functions β, σ are applications from $[0,T] \times L^2(\Omega) \to L^2(\Omega)$, while η is an $\mathcal{F}_0$ measurable process in $L^2(\Omega)$, that is the boundary condition. It is common to write equation (1.17) in the shorthand notation:

$$\begin{cases} dX(t) = \beta(t, X(t))dt + \sigma(t, X(t))dW(t) \\ X(0) = \eta \end{cases} \tag{1.18}$$

For a review of theory of stochastic differential equation of the kind (1.17), see Da Prato (1998); Karatzas and Shreve (1988). For our purposes, it is sufficient to state the following existence and uniqueness result.

Theorem 1.42 *Assume the following assumptions hold:*

1. *β and σ are continuous.*

2. *There exists $M > 0$ such that:*

$$\begin{array}{ll} ||\beta(t,\zeta)||^2 + ||\sigma(t,\zeta)||^2 \leq M^2(1 + ||\zeta||^2) & \forall t \in [0,T], \zeta \in L^2(\Omega) \\ ||\beta(t,\zeta_1) - \beta(t,\zeta_2)|| + ||G(t,\zeta_1) - G(t,\zeta_2)|| \leq M||\zeta_1 - \zeta_2|| & \forall t \in [0,T], \zeta_1, \zeta_2 \in L^2(\Omega) \end{array} \tag{1.19}$$

3. *$\forall t \in [0,T], \zeta \in L^2(\Omega)$ such that ζ is $\mathcal{F}_t$-measurable, we have that $\beta(t,\zeta), \sigma(t,\zeta) \in L^2(\Omega)$ and are $\mathcal{F}_t$-measurable.*

Let $\eta \in L^2(\Omega)$ and $\mathcal{F}_0$-measurable. Then there exist a unique (up to null sets) adapted process $X(t) \in L^2(\Omega)$ fulfilling equation (1.17).

Corollary 1.43 *Assume the hypothesis of Theorem 1.42 hold. Then the (unique) solution process X is a continuous semimartingale, and*

$$[X, X]_t = \int_0^t \sigma^2(s, X(s))ds \tag{1.20}$$

Proof. The result come from the fact that η is $\mathcal{F}_0$-measurable and finite-valued, $\int_0^t \beta(s, X(s))ds$ is of finite variation and $\int_0^t \sigma(s, X(s))dW(s)$ is a local martingale, since Wiener process is a local martingale and using 1.28. For the continuity, see Da Prato (1998). □

We want now to investigate the link between quadratic variation and the covariance function of the difference process, following Andersen et al. (2003). Consider an $\mathbb{R}^d$ valued semimartingale $p(t)$ in $[0, T]$, and its unique decomposition according to Theorem 1.25, $p(t) = p(0) + M(t) + A(t)$. Let $t \in [0, T]$, h such that $t + h < T$, and denote the difference process in the interval $[t, t + h]$ by $r(t, h) = p(t + h) - p(t)$. In financial economics, if $p(t)$ is the process of logarithmic prices, $r(t, h)$ are called logarithmic returns. We can also define the cumulative difference process $r(t) = p(t) - p(0)$. It is clear that $[r, r]_t = [p, p]_t$. We then have the following:

Proposition 1.44 *Consider a semimartingale $p(t)$. The conditional difference process covariance matrix at time t over $[t, t + h]$ is given by*

$$Cov(r(t, t+h)|\mathcal{F}_t) = \mathbb{E}\left[[r, r]_{t+h} - [r, r]_t|\mathcal{F}_t\right] + \Gamma_A(t, t+h) + \Gamma_{AM}(t, t+h) + \Gamma'_{AM}(t, t+h) \quad (1.21)$$

where $\Gamma_A(t, t + h) = Cov(A(t + h) - A(t)|\mathcal{F}_t)$ and $\Gamma_{AM}(t, t + h) = \mathbb{E}[A(t + h)(M(t + h) - M(t))'|\mathcal{F}_t$.

Proof. See Proposition 2 of Andersen et al. (2003).

Proposition 1.44 decomposes the covariance matrix of the difference process in three parts. The first is the contribution of quadratic variation, and it is simply given by its conditional expectation. The second is the contribution of the drift term. The third is the contribution of the covariance between drift and diffusion term. The last two terms disappear if, for instance, the drift term is not stochastic. Even if the drift term is stochastic, so that the last two terms are not null, they are still less relevant when compared to the quadratic variation contribution. For example, we have:

$$\Gamma^{ij}_{AM}(t, t + h) \leq \left(Var[A^i(t + h) - A^i(t)|\mathcal{F}_t]\right)^{\frac{1}{2}} \cdot \left(Var[M^j(t + h) - M^j(t)|\mathcal{F}_t]\right)^{\frac{1}{2}}$$

and the latter terms are of order h and $h^{\frac{1}{2}}$ respectively, thus Γ_{AM} is at most of order $h^{\frac{3}{2}}$.

We finally state a proposition on the distribution of returns:

Proposition 1.45 *Let X be the process satisfying (1.17), and consider the difference process $r(t, t + h)$ of X, and assumptions of Theorem 1.42 hold, and that β, σ are independent of $W(s)$ in the interval $[t, t + h]$. Then the law of $r(t, t + h)$ conditional to $\mathcal{F}_t$ is $\mathcal{N}\left(\int_t^{t+h} \beta(s)ds, \int_t^{t+h} \sigma^2(s)ds\right)$.*

Proof. See Andersen et al. (2003), Theorem 2.

1.2 Convergence to a martingale with given quadratic variation

Definition 1.46 *A process with independent increments (PII) in a filtered probability space is a cadlag adapted $\mathbb{R}$-valued process X such that $X_0 = 0$ and that $\forall\ 0 \leq s \leq t$ the variable $X_t - X_s$ is independent of $\mathcal{F}_s$.*

Definition 1.47 *A truncation function $h(x)$ is a bounded Borel real function with compact support which behaves like x near the origin.*

For every semimartingale X, we define its characteristics (B, C, ν) as follows. Let h be a truncation function. We define $X(h) = X - \sum_{s \leq t}[\Delta X_s - h(\Delta X_s)]$. $X(h)$ is a special semimartingale (since it has bounded jumps) and we can write its canonical decomposition:

$$X(h) = X_0 + M(h) + B(h) \tag{1.22}$$

where $M(h)$ is a local martingale and $B(h)$ a predictable process of finite variation.

Definition 1.48 *The characteristics of X is the triplet (B, C, ν) defined by:*

1. *$B = B(h)$ in (1.22)*
2. *$C = [X^c, X^c]$ i.e. the quadratic variation of the continuous martingale part of X*
3. *ν is the compensator of the random measure associated with the jumps of X.*

We then have that B is a predictable process of finite variation, C is a continuous process of finite variation and ν is a predictable measure on $\mathbb{R}^+ \times \mathbb{R}$. Extension to the multivariate case is straightforward. If X is a PII with $X_0 = 0$ and without fixed times of discontinuity, then Levy-Kinthchine formula holds:

$$\mathbb{E}[e^{iuX_t}] = \exp\left(iuB_t - \frac{u^2}{2}C_t + \int_{\mathbb{R}^+}\left(e^{iux} - 1 - iuh(x)\right)\nu_t(dx)\right). \tag{1.23}$$

Then next theorem provides the characteristics of the processes of the following kind:

$$Y_t = \sum_{i=1}^{[nt]} U_i \tag{1.24}$$

where $[x]$ is the integer part of x and $(U_i)_{i \in \mathbb{N}}$ is an adapted process.

Theorem 1.49 *Let h be any truncation function and $g \geq 0$ Borel. Then*

$$\begin{cases} B_t = \sum_{i=1}^{[nt]} \mathbb{E}[h(U_i)|\mathcal{F}_{i-1}] \\ C_t = 0 \\ g * \nu = \int\int_{[0,T]\times\Omega} g d\nu = \sum_{i=1}^{[nt]} \mathbb{E}[g(U_i) I_{\{U_i \neq 0\}} | \mathcal{F}_{i-1}] \end{cases} \tag{1.25}$$

If $h^2 * \nu < \infty, \ \forall t \in [0,T]$, we can define the following:

$$\tilde{C}_t = C_t + h^2 * \nu_t - \sum_{s \leq t} (\Delta B_s)^2 \tag{1.26}$$

We then have the following convergence theorem:

Theorem 1.50 *Fix a truncation function h. Assume that X^n is a sequence of semimartingales, and X is a PII semimartingale without fixed time of discontinuity. Denote by (B^n, C^n, ν^n) the characteristics of X^n and by (B, C, ν) the characteristics of X. Define $\tilde{C}$ by equation (1.26).Moreover assume the following:*

1. $\sup_{s \leq t} |B^n_s - B_s| \to 0$ *in probability*, $\quad \forall t \in [0,T]$
2. $\tilde{C}^n_t \to \tilde{C}_t$ *in probability* , $\quad \forall t \in [0,T]$
3. $g * \nu^n_t \to g * \nu_t$ *in probability*, $\quad \forall t \in [0,T], g \in \mathcal{C}_1(\mathbb{R})$

Then $X_n \to X$ in distribution.

Proof. This is Theorem VIII.2.17 in Jacod and Shiryaev (1987). □

We then show the following Theorem, which will be useful in our analysis:

Theorem 1.51 *Consider the process Y^n_t defined in (1.24), with U_i bounded, and assume the following:*

1. $\sum_{i=1}^{[nt]} \mathbb{E}[U_i|\mathcal{F}_{i-1}] \to 0$ *in probability*

2. $\sum_{i=1}^{[nt]} \mathbb{E}[U_i^2|\mathcal{F}_{i-1}] \to V_t$ *in probability*

3. $\forall\ \varepsilon > 0$, $\sum_{i=1}^{[nt]} \mathbb{E}[U_i^2 I_{\{|U_i|>\varepsilon\}}|\mathcal{F}_{i-1}] \to 0$ *in probability (conditional Lindeberg condition)*

Then Y_t converges in distribution to the continuous martingale M_t with quadratic variation $[M, M]_t = V_t$.

Proof. We have to prove conditions $1 - 3$ of Theorem 1.50. We compute the characteristics (B^n, C^n, ν^n) of Y^n by theorem 1.49, with $h(x) = x \wedge \sup U_i$ and $\tilde{C}$ by (1.26). The characteristics of M_t is $(0, V_t, 0)$.

1. We have $B'^n = B^n$. Since U_i is bounded, this follows directly.
2. Comes directly from the definition of C'^n and the fact that $\sum_{s\leq t}(\Delta B_s)^2 = \sum_{i=1}^{[nt]} \mathbb{E}[U_i^2|\mathcal{F}_{i-1}] \to 0$ from 1.
3. If $g \in \mathcal{C}_1$ there exist real numbers k, K such that $|g(x)| \leq Kx^2 I_{\{|x|>k\}}$, thus the conditional Lindeberg condition implies $g * \nu^n \to 0$.

□

1.3 Estimating quadratic variation

Different estimators for the integrated volatility have been proposed. Nowadays, the most popular is *realized volatility*, which will be discussed thoroughly throughout. The idea behind realized volatility hinges on Theorem 1.32. Let $p(t) \in \mathbb{R}^d$ be driven by the SDE (1.17) in the interval $[0, T]$, and consider equally spaced observations $p_0^i, p_1^i, \ldots, p_n^i$, for $i = 1, \ldots d$. Then define $r_k^i = p_k^i - p_{k-1}^i$. Then realized volatility is given by:

$$RV^{ij} = \sum_{k=1}^{m} r_{t+k/m}^i \cdot r_{t+k/m}^j. \tag{1.27}$$

The drift component is ignored since it can be set to zero for typical application. On the statistical properties of realized volatility, see Andersen et al. (2003); Barndorff-Nielsen and

Shephard (2002a). On the effectiveness of realized volatility as a measure of integrated volatility, see Meddahi (2002).

The *range* is based on the following observation of Parkinson (1980): if $p(t)$ is a one-dimensional solution of of $dp(t) = \sigma dW(t)$, with $\sigma \in \mathbb{R}$, and it is observed in $[0, T]$, then

$$\sigma^2 = 0.361 \cdot \mathbb{E}\left[\left(\max_{t\in[0,T]} p(t) - \min_{t\in[0,T]} p(t)\right)^2\right] \tag{1.28}$$

This immediately provides an estimate of the variance, which is very popular among finance practitioners, since the maximum and the minimum of the price (so-called high and low) are always recorded.

It is simple to show that this idea can be extended to the full variance-covariance matrix, see e.g. Brandt and Diebold (2004). This method has been refined using also open and close price, see Garman and Klass (1980); Rogers and Satchell (1991); Yang and Zhang (2000).

Other methods have been proposed in the literature. Ball and Torous (1984, 2000) regard volatility as a latent variable and estimate it via maximum likelihood; Genon-Catalot et al. (1992) develop a wavelet estimator. Spectral methods have been devised by Thomakos et al. (2002) and Curci and Corsi (2003). Finally, a spectral method has been worked out by Malliavin and Mancino (2002), see Chapter 2. Andersen et al. (2003) is a nearly complete review of this topic.

1.4 Quadratic variation in financial economics

The importance of quadratic variation in financial economics is widely recognized. The main reason stems from the seminal contribution of Black and Scholes (1973) and Merton (1973), who showed that option prices are a function of asset price volatility. In this Thesis, we will circumvent the issue of derivative pricing, since we are more interested in the estimation of quadratic variation from the observation of asset prices, which in the derivative field is called *historic volatility*. It is well known that the *implicit volatility*, that is the volatility which "prices" options, is very different from the historic one, and one very well known example is the smile effect. In particular, we will concentrate on the use of the so-called high-frequency data, whose use became customary in the last decade, see Goodhart and O'Hara (1997).

Historic volatility was paid a great attention in the financial economics literature. Here we give just few examples of the main problems raised. Christie (1982) analyzes the relation

between variance and leverage and variance and interest rates. The leverage effect has been longly studied, since the contributions of Black (1976) and Cox and Ross (1976). The asymmetric link between realized volatility and returns is studied in a recent paper by Bekaert and Wu (2000), where a model of volatility feedback is introduced, see also Duffee (1995); Wu (2001). French and Roll (1986) pose the problem that asset prices variance during trading periods is higher than variance during non-trading periods, and link this finding to the role of private information. The same approach has been followed by Amihud and Mendelson (1987). French et al. (1987), assess the relation between volatility and expected risk premium of stock returns. In the same line, Schwert (1989) studies volatility over more than a century, shows that it is stochastic and tries to explain its movements with regard to macroeconomic variables. In the same spirit, Campbell et al. (2001) study the phenomenon of increasing volatility of stocks, explaining this via macroeconomic variables. Intraday volatility has been studied by Lockwood and Linn (1990) and in Andersen and Bollerslev (1997), where a link is posed between intraday periodicity and persistence.

Maybe the most important stylized fact on volatility is its persistence, or clustering. This idea can be found already in Mandelbrot (1963) or Fama (1965). Poterba and Summers (1986) highlight the importance of persistence on the data used by French et al. (1987). Schwert and Seguin (1990) relate the degree of heteroskedasticity to size. Heteroskedasticity leads to modeling persistence in order to get a good picture of asset prices evolution. The result are the ARCH model of Engle (1982) and the GARCH model of Bollerslev (1986), which are very popular nowadays, see Bollerslev et al. (1992) and Bollerslev et al. (1994) for a review. Nelson (1992) assesses the relation between the variance estimated by an ARCH model and the true quadratic variation, showing that the difference between the two converges to zero when the time interval shrinks. The interest in volatility persistence stems from its consequent predictability. Forecasting volatility is probably the main application of the use of the concept of quadratic variation. A quite extensive review of this topic is Poon and Granger (2003). On the importance of volatility forecasting for risk management, see also Christoffersen and Diebold (2000); we will deal with this topic in Chapter 3. Quadratic variation has been used in assessing the informational efficiency of implied volatility, see e.g. Christensen and Prabhala (1998); Blair et al. (2001).

The financial literature on quadratic variation renewed after the contribution of Andersen and Bollerslev (1998a). They show that the low forecasting performance of GARCH(1,1) models, as found e.g. in Jorion (1995), is not due to the poor forecasting ability of these models, but to the poor estimation of integrated volatility. Dating back to an idea of Merton (1980), they show, using simulations and FX data, that it is possible to estimate daily volatility using intraday transactions (high-frequency data), and that these estimates

are by far more precise than just the daily squared return, and that GARCH forecasting performance is good. They called the measure of volatility via cumulative squared returns *realized volatility*. This parallels the work of Poterba and Summers (1986); French et al. (1987); Schwert (1989); Schwert and Seguin (1990) who compute monthly volatility using daily returns. In some sense, it introduces a new econometric variable, and this leaded to a very large literature.

Within the same strand, Barndorff-Nielsen and Shephard (2002b) study the statistical properties of realized volatility. Hansen and Lunde (2004) compare a large class of autoregressive models using realized volatility measures, concluding that GARCH(1,1) is very difficult to be outperformed. Andersen et al. (2001) and Andersen et al. (2001) study the statistical properties of realized volatility of stock prices and exchange rates respectively. Andersen et al. (2000a) study the distribution of standardized returns. The purpose of this studies is to assess the unconditional and conditional properties of volatility, for instance long memory. Similar studies have been conducted for different markets: see Taylor and Xu (1997), Zhou (1996), Areal and Taylor (2002) for the FTSE, Andersen et al. (2000) for the Nikkei, Bollerslev et al. (2000) for an application to interest rates, Bollen and Inder (2002), Martens (2001), Martens (2002), Thomakos and Wang (2003) for futures markets and Renò and Rizza (2003); Pasquale and Renò (2005); Bianco and Renò (2005) for the Italian futures market.

Using integrated volatility as an observable leads to modeling it directly. The simplest idea to capture persistence is an autoregressive model. Andersen et al. (2003) propose an autoregressive model with long memory, end estimate it on foreign exchange rates and stock returns. A similar model is proposed by Deo et al. (2003). The HAR-RV model of Corsi (2003) is similar, but economically significant restrictions are imposed to the autoregressive structure; long memory is attained using the intuition of Granger (1980), that is as the sum of short memory components of different frequencies. Maheu and McCurdy (2002) study the importance of non-linear components in the autoregressive structure of volatility dynamics, while Maheu and McCurdy (2004) study the impact of jumps on volatility. Fleming et al. (2001) show that using a dynamic volatility model instead of a static one, portfolio management can improve substantially. Then, they refine their research using realized volatility as an observable (Fleming et al., 2003) and find even better results.

Finally, integrated volatility has been used as an observable for estimating stochastic models. One example is Bollerslev and Zhou (2002), which estimates a model for exchange rates using realized volatilities and GMM. In the same spirit, Pan (2002) uses GMM to estimate a model for stock prices, using as observables the stock prices, option prices and realized volatility. Barndorff-Nielsen and Shephard (2002a) suggest a maximum likelihood estimator which uses realized volatilities; Galbraith and Zinde-Walsh (2000) use realize volatility to es-

timate GARCH-like models. Alizadeh et al. (2002); Gallant et al. (1999) estimate integrated volatility using the range, that is the squared difference between the high and low of an asset price during a day, and show how to estimate stochastic volatility models including the range as an observable. The range has been used to get more efficient estimates of EGARCH models (Brandt and Jones, 2002). Similar studies on the range have been conducted by Brunetti and Lildholdt (2002a,b).

Chapter 2

Volatility estimate via Fourier Analysis

2.1 Univariate case

We work in the filtered probability space $(\Omega, \mathcal{F}_t, \mathcal{P})$ satisfying the usual conditions (Protter, 1990), and define X_t as the solution of the following process:

$$\begin{cases} dX_t = \mu(t)dt + \sigma(t)dW_t \\ X_0 = x_0 \end{cases} \tag{2.1}$$

where $\sigma(t), \mu(t)$ are bounded deterministic functions of time, and W_t is a standard Brownian motion.[1] We will write $X_t(\omega)$ to explicit the dependence of X from $t \in [0, T]$ and $\omega \in \Omega$.

In this case X_t is a semimartingale, and its quadratic variation (Jacod and Shiryaev, 1987) is given by:

$$[X, X]_t = \int_0^t \sigma^2(s)ds \tag{2.2}$$

We can assume, without loss of generality, that the time interval is $[0, 2\pi]$, and define the

[1]The restrictions that the drift and the diffusion coefficients be deterministic functions of time can be relaxed, see Malliavin and Mancino (2005).

Fourier coefficients of dX and σ^2 as follows:

$$\begin{aligned} a_0(dX) &= \frac{1}{2\pi}\int_0^{2\pi} dX_t & a_0(\sigma^2) &= \frac{1}{2\pi}\int_0^{2\pi} \sigma^2(t)dt \\ a_k(dX) &= \frac{1}{\pi}\int_0^{2\pi} \cos(kt)dX_t & a_k(\sigma^2) &= \frac{1}{\pi}\int_0^{2\pi} \cos(kt)\sigma^2(t)dt \\ b_k(dX) &= \frac{1}{\pi}\int_0^{2\pi} \sin(kt)dX_t & b_k(\sigma^2) &= \frac{1}{\pi}\int_0^{2\pi} \sin(kt)\sigma^2(t)dt \end{aligned} \tag{2.3}$$

There are many ways to reconstruct $\sigma^2(t)$ given its Fourier coefficients. One way is the Fourier-Fejer formula:

$$\sigma^2(t) = \lim_{M\to\infty} \sum_{k=0}^{M} \left(1 - \frac{k}{M}\right) \left[a_k(\sigma^2)\cos(kt) + b_k(\sigma^2)\sin(kt)\right] \tag{2.4}$$

Convergence of Fourier sums is in $\mathcal{L}^2([0, 2\pi])$ norm and it is pointwise where $\sigma^2(t)$ is analytic[2]. We now state the main result:

Theorem 2.1 *Consider a process X_t satisfying (2.1), and define the Fourier coefficients of dX and σ^2 as in (2.3). Given an integer $n_0 > 0$, we have in $\mathcal{L}^2$:*

$$a_0(\sigma^2) = \lim_{N\to\infty} \frac{\pi}{N+1-n_0} \sum_{k=n_0}^{N} a_k^2(dX) = \lim_{N\to\infty} \frac{\pi}{N+1-n_0} \sum_{k=n_0}^{N} b_k^2(dX) \tag{2.5}$$

$$a_q(\sigma^2) = \lim_{N\to\infty} \frac{2\pi}{N+1-n_0} \sum_{k=n_0}^{N} a_k(dX)a_{k+q}(dX) = \lim_{N\to\infty} \frac{2\pi}{N+1-n_0} \sum_{k=n_0}^{N} b_k(dX)b_{k+q}(dX) \tag{2.6}$$

$$b_q(\sigma^2) = \lim_{N\to\infty} \frac{2\pi}{N+1-n_0} \sum_{k=n_0}^{N} a_k(dX)b_{k+q}(dX) = -\lim_{N\to\infty} \frac{2\pi}{N+1-n_0} \sum_{k=n_0}^{N} b_k(dX)a_{k+q}(dX) \tag{2.7}$$

Proof. We follow the proof of Malliavin and Mancino (2002).

Consider first the case $\mu(t) = 0$.

[2]Actually there are looser request for punctual convergence, but it is important to stress that continuity is not sufficient.

We choose $k, h \in \mathbb{N}$ such that $k > h \geq 1$. We have:

$$\begin{aligned}\mathbb{E}\left[a_k(dX)a_h(dX)\right] &= \mathbb{E}\left[\frac{1}{\pi}\int_0^{2\pi}\cos(kt)dX_t \cdot \frac{1}{\pi}\int_0^{2\pi}\cos(ht)dX_t\right] = \\ &= \mathbb{E}\left[\frac{1}{\pi^2}\int_0^{2\pi}\cos(kt)\sigma(t)dW(t) \cdot \int_0^{2\pi}\cos(ht)\sigma(t)dW(t)\right] = \\ &= \frac{1}{\pi^2}\int_0^{2\pi}\sigma^2(t)\cos(kt)\cos(ht)dt.\end{aligned} \tag{2.8}$$

by the contraction formula.

Using the following identity:

$$2\cos(kt)\cos(ht) = \cos[(k-h)t] + \cos[(k+h)t] \tag{2.9}$$

we get:

$$\mathbb{E}\left[a_k(dX)a_h(dX)\right] = \frac{1}{2\pi}\left[a_{k-h}(\sigma^2) + a_{k+h}(\sigma^2)\right] \tag{2.10}$$

Moreover we have:

$$\left\|\sigma^2\right\|_{\mathcal{L}^2}^2 = \sum_{k=0}^{+\infty}\left(a_k^2(\sigma^2) + b_k^2(\sigma^2)\right) \tag{2.11}$$

Now fix an integer $n_0 > 0$ and define, for $q \in \mathbb{N}$:

$$U_N^q = \frac{1}{N+1-n_0}\sum_{k=n_0}^{N} a_k(dX)a_{k+q}(dX) \tag{2.12}$$

Using (2.10) after taking expectations we get:

$$\mathbb{E}[U_N^q] = \frac{1}{N+1-n_0}\frac{1}{2\pi}\sum_{k=n_0}^{N}\left(a_q(\sigma^2) + a_{2k+q}(\sigma^2)\right) = \frac{1}{2\pi}a_q(\sigma^2) + R_N. \tag{2.13}$$

Where

$$|R_N| = \frac{1}{N+1-n_0}\frac{1}{2\pi}\left|\sum_{k=n_0}^{N} a_{2k+q}(\sigma^2)\right| \leq \frac{1}{\sqrt{N+1-n_0}}\left\|\sigma^2\right\|_{\mathcal{L}^2} \tag{2.14}$$

by Schwartz inequality, thus

$$a_q(\sigma^2) = 2\pi\lim_{N\to\infty}\mathbb{E}[U_N^q]. \tag{2.15}$$

We want now to prove that $a_q(\sigma^2) = 2\pi\lim_{N\to\infty} U_N^q$. To do so we compute:

$$\mathbb{E}^2[U_N^q] = \frac{1}{(N+1-n_0)^2}\sum_{n_0\leq k_1,k_2\leq N}\mathbb{E}\left[a_{k_1}(dX)a_{k_1+q}(dX)\right]\mathbb{E}\left[a_{k_2}(dX)a_{k_2+q}(dX)\right] \tag{2.16}$$

Using the fact that $a_k(dX)$ is a Gaussian random variable with mean 0, we use a well known formula for the product of four Gaussian random variables to compute:

$$\begin{aligned}\mathbb{E}[(U_N^q)^2] &= \frac{1}{(N+1-n_0)^2}\sum_{n_0\le k_1,k_2\le N} \mathbb{E}\left[a_{k_1}(dX)a_{k_1+q}(dX)a_{k_2}(dX)a_{k_2+q}(dX)\right] = \\ &= \frac{1}{(N+1-n_0)^2}\sum_{n_0\le k_1,k_2\le N} \left(\mathbb{E}\left[a_{k_1}(dX)a_{k_1+q}(dX)\right]\mathbb{E}\left[a_{k_2}(dX)a_{k_2+q}(dX)\right] + \right. \\ &\quad +\mathbb{E}\left[a_{k_1}(dX)a_{k_2}(dX)\right]\mathbb{E}\left[a_{k_1+q}(dX)a_{k_2+q}(dX)\right] \\ &\quad \left. +\,\mathbb{E}\left[a_{k_1}(dX)a_{k_2+q}(dX)\right]\mathbb{E}\left[a_{k_1+q}(dX)a_{k_2}(dX)\right]\right)\end{aligned} \tag{2.17}$$

We now use equation (2.10) to get:

$$\begin{aligned}\mathbb{E}[(U_N^q-\mathbb{E}[U_N^q])^2] &= \frac{1}{4\pi^2(N+1-n_0)^2}\cdot \\ &\quad \cdot\sum_{n_0\le k_1,k_2\le N}\left[\left(a_{k_1+k_2}(\sigma^2)+a_{|k_1-k_2|}(\sigma^2)\right)\left(a_{k_1+k_2+2q}(\sigma^2)+a_{|k_1-k_2|}(\sigma^2)\right)+\right. \\ &\quad \left.+\left(a_{k_1+k_2+q}(\sigma^2)+a_{|k_1-k_2-q|}(\sigma^2)\right)\left(a_{k_1+k_2+q}(\sigma^2)+a_{|k_1-k_2+q|}(\sigma^2)\right)\right]\end{aligned} \tag{2.18}$$

Finally we use Cauchy-Schwartz:

$$\begin{aligned}&\mathbb{E}[(U_N^q-\mathbb{E}[U_N^q])^2]\le \\ &\le \frac{1}{4\pi^2(N+1-n_0)^2}\cdot \\ &\cdot\left[\left(\sum_{n_0\le k_1,k_2\le N}\left(a_{k_1+k_2}(\sigma^2)+a_{|k_1-k_2|}(\sigma^2)\right)^2\cdot\sum_{n_0\le k_1,k_2\le N}\left(a_{k_1+k_2+2q}(\sigma^2)+a_{|k_1-k_2|}(\sigma^2)\right)^2\right)^{\frac{1}{2}}+\right. \\ &\left.+\left(\sum_{n_0\le k_1,k_2\le N}\left(a_{k_1+k_2+q}(\sigma^2)+a_{|k_1-k_2-q|}(\sigma^2)\right)^2\cdot\sum_{n_0\le k_1,k_2\le N}\left(a_{k_1+k_2+q}(\sigma^2)+a_{|k_1-k_2+q|}(\sigma^2)\right)^2\right)^{\frac{1}{2}}\right]\le \\ &\le \frac{2}{\pi^2(N+1-n_0)}\left\|\sigma^2\right\|_{\mathcal{L}^2}^2\end{aligned} \tag{2.19}$$

The above inequality proves convergence in $\mathcal{L}^2$, then in probability.

If we now repeat the calculation (2.8) replacing a_k, a_h with a_k, b_h, we have:

$$\mathbb{E}\left[a_k(dX)b_h(dX)\right] = \int_0^{2\pi}\sigma^2(t)\cos(kt)\sin(ht)dt. \tag{2.20}$$

We now use the identity:

$$2\cos(kt)\sin(ht) = \sin[|k-h|t] + \sin[(k+h)t] \tag{2.21}$$

and we get:

$$\mathbb{E}\left[a_k(dX)b_h(dX)\right] = \frac{1}{2\pi}\left[b_{k-h}(\sigma^2)+b_{k+h}(\sigma^2)\right] \tag{2.22}$$

We then get formula (2.7) by computing the expected value of:

$$V_N^q = \frac{1}{N+1-n_0}\sum_{k=n_0}^{N} a_k(dX)b_{k+q}(dX),\ W_N^q = \frac{1}{N+1-n_0}\sum_{k=n_0}^{N} b_k(dX)a_{k+q}(dX) \quad (2.23)$$

The second part of formula (2.6) comes in the same way from the identity:

$$2\sin(kt)\sin(ht) = \cos[|k-h|t] - \cos[(k+h)t] \quad (2.24)$$

Formula (2.5) comes in the same way from:

$$\mathbb{E}[a_k^2(dX)] = \mathbb{E}[b_k^2(dX)] = \frac{1}{2\pi}\left[2a_0(\sigma^2) - a_{2k}(\sigma^2)\right] \quad (2.25)$$

If $\mu(t) \neq 0$, then in all previous computation we replace dX with dv defined by $dv = dX - \mu(t)dt$. Now, all the extra terms depending on μ vanish asymptotically since:

$$\int_0^{2\pi} \mu^2(t)dt = \sum_{k=0}^{+\infty}\left(a_k^2(\mu) + b_k^2(\mu)\right). \quad (2.26)$$

□

Corollary 2.2 *The Fourier coefficients of $\sigma^2(t)$ can be computed in the $\mathcal{L}^2$ sense as:*

$$a_0(\sigma^2) = \lim_{N\to\infty} \frac{\pi}{N+1-n_0}\sum_{k=n_0}^{N}\frac{1}{2}\left(a_k^2(dX) + b_k^2(dX)\right) \quad (2.27)$$

$$a_q(\sigma^2) = \lim_{N\to\infty} \frac{\pi}{N+1-n_0}\sum_{k=n_0}^{N}\left(a_k(dX)a_{k+q}(dX) + b_k(dX)b_{k+q}(dX)\right) \quad (2.28)$$

$$b_q(\sigma^2) = \lim_{N\to\infty} \frac{\pi}{N+1-n_0}\sum_{k=n_0}^{N}\left(a_k(dX)b_{k+q}(dX) - b_k(dX)a_{k+q}(dX)\right) \quad (2.29)$$

From Theorem 2.1 we get immediately an estimator of the integrated volatility. Indeed:

$$\int_0^{2\pi} \sigma^2(s)ds = 2\pi a_0(\sigma^2) \quad (2.30)$$

where $a_0(\sigma^2)$ is given by formula (2.5). The following Theorem provides asymptotic confidence intervals for the Fourier coefficients of volatility, in the case of constant σ:

Theorem 2.3 *Assume volatility is a constant, $\sigma(\cdot) = \sigma \in \mathbb{R}$. As $N \to \infty$, we have:*

$$\sqrt{N+1-n_0}\left(a_0(\sigma^2) - \sigma^2\right) \to \mathcal{N}\left(0, 2\sigma^4\right) \tag{2.31}$$

$$\sqrt{N+1-n_0}\; a_q(\sigma^2) \to \mathcal{N}\left(0, \sigma^4\right) \tag{2.32}$$

$$\sqrt{N+1-n_0}\; b_q(\sigma^2) \to \mathcal{N}\left(0, \sigma^4\right) \tag{2.33}$$

where the above limit is in distribution.

Proof. We have already shown in the proof of Theorem 2.1 that $a_k(dX), b_k(dX)$ can be replaced by $a_k(dv), b_k(dv)$ where $dv = dX - \mu(t)dt$, since all the contribution of the coefficients of μ vanish a.s. as $N \to \infty$. We start from the fact that $a_k(dv), b_k(dv)$ are Gaussian random variables with zero mean. Let $\sigma(\cdot) = \sigma$. We then have:

$$\mathbb{E}[a_k^2(dv)] = \mathbb{E}[b_k^2(dv)] = \frac{1}{\pi}\sigma^2 \tag{2.34}$$

and

$$\mathbb{E}[a_k^4(dv)] = \mathbb{E}[b_k^4(dv)] = \frac{3}{\pi^2}\sigma^4 \tag{2.35}$$

Moreover, from the orthogonality of the trigonometric base, if $k \neq h, \mathbb{E}[a_k(dv)a_h(dv)] = \mathbb{E}[b_h(dv)b_k(dv)] = 0$ and, for every $k, h, \mathbb{E}[a_k(dv)b_k(dv)] = 0$. Thus $a_k(dv), b_h(dv)$ are all independent, thus if $k \neq h$, $a_k^2(dv) + b_k^2(dv)$ is independent of $a_h^2(dv) + b_h^2(dv)$. Then standard central limit theorem yields the result. We get the result for $a_q(\sigma^2), b_q(\sigma^2)$ with the same reasoning, since $\mathbb{E}[a_k(dv)a_{k+q}(dv)] = \mathbb{E}[a_k(dv)b_{k+q}(dv)] = 0$ and $\mathbb{E}[(a_k(dv)a_{k+q}(dv))^2] = \sigma^4/\pi^2$. □

It is sometimes convenient to rewrite equation (2.4) as:

$$\sigma^2(t) = \lim_{M\to\infty} \sum_{k=-M}^{M} \left(1 - \frac{k}{M}\right) A_k(\sigma^2) e^{ikt}, \tag{2.36}$$

where

$$A_k(\sigma^2) = \begin{cases} \frac{1}{2}(a_k(\sigma^2) - ib_k(\sigma^2)), & k \geq 1 \\ \frac{1}{2}a_0(\sigma^2), & k = 0 \\ \frac{1}{2}(a_{|k|}(\sigma^2) + ib_{|k|}(\sigma^2)) & k \leq -1 \end{cases} \tag{2.37}$$

For the implementation of the estimator, we adopt the following procedure. Since we observe the process X_t only at discrete times $t_1, \ldots, t_n$, we set $X_t = X_{t_i}$ in the interval $t_i \leq t < t_{i+1}$. Using interpolation techniques different from this we get a bias in the volatility measurement (Barucci and Renò, 2002b). Then the Fourier coefficients of the price can be computed as:

$$a_k(dX) = \frac{1}{\pi}\int_0^{2\pi} \cos(kt)dX_t = \frac{X_{2\pi} - X_0}{\pi} - \frac{k}{\pi}\int_0^{2\pi} \sin(kt)X_t dt, \tag{2.38}$$

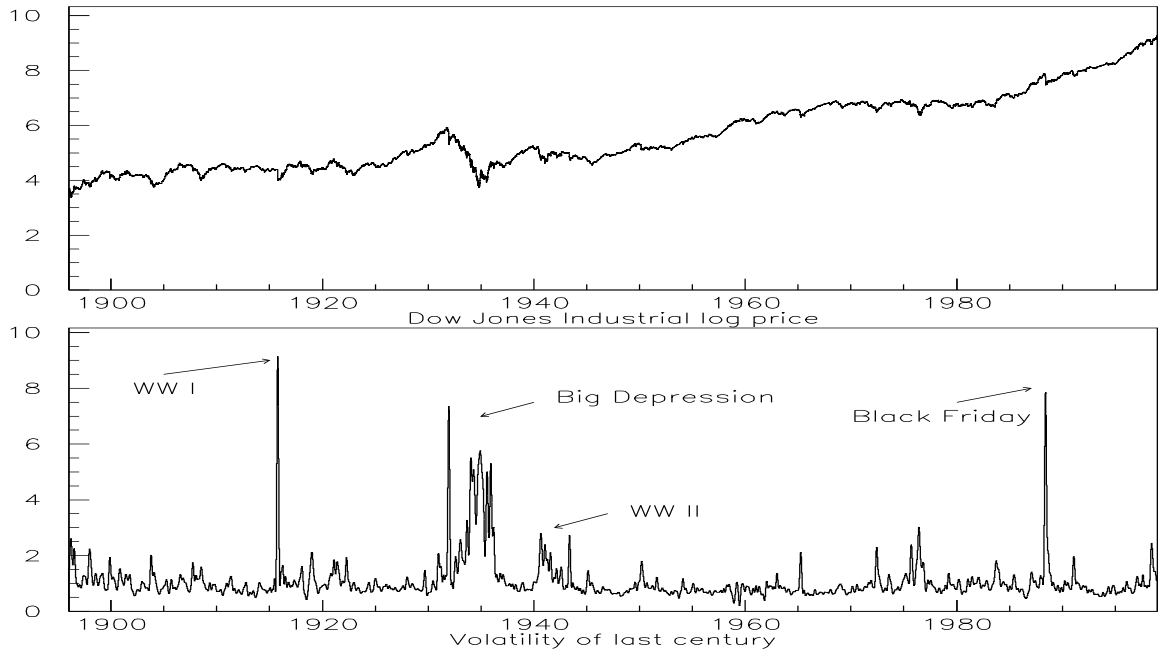

Figure 2.1: Top: Daily log-price of the Dow Jones Industrial average from 1896 to 1999. Bottom: Daily volatility of the same index, computed with the Fourier method.

then using:

$$\frac{k}{\pi}\int_{t_i}^{t_{i+1}} \sin(kt)X_t\ dt = X_{t_i}\frac{k}{\pi}\int_{t_i}^{t_{i+1}} \sin(kt)dt = X_{t_i}\frac{1}{\pi}\left[\ \cos(kt_i) - \cos(kt_{i+1})\ \right]. \qquad (2.39)$$

Before computing (2.38), we add a linear trend such that we get $X_{2\pi} = X_0$, which does not affect the volatility estimate. Then we stop the expansions (2.27-2.29) at a properly selected frequency N. For equally spaced data, the maximum N which prevents aliasing effects is $N = \frac{n}{2}$, see Priestley (1979). Finally, we have to select the maximum M in (2.36). M should be a function of N such that $M(N) \to \infty$ when $N \to \infty$.

In order to illustrate the potential of the method, we compute the volatility $\sigma^2(t)$ on a one century long time series, that is the daily close price of the Dow Jones Industrial index, from 26 May 1896 to 29 April 1999. We implemented the method with $5,000$ coefficients for the price and 500 for volatility. Figure 2.1 shows the result, and how it is possible to link volatility bursts and clustering to well defined periods.

2.2 Multivariate case

The multivariate case is a straightforward extension of the previous analysis. In the usual filtered probability space $(\Omega, \mathcal{F}_t, \mathcal{P})$, define $X_t \in \mathbb{R}^d$ as the solution of the following process:

$$\begin{cases} dX_t = \mu(t)dt + \sigma(t)dW_t \\ X_0 = x_0 \end{cases} \tag{2.40}$$

where $\sigma(t) \in \mathbb{R}^d \times \mathbb{R}^d, \mu(t) \in \mathbb{R}^d$ are bounded deterministic functions, and W_t is an $\mathbb{R}^d$ valued Brownian motion. In this case, with continuous trajectories, quadratic variation is simply:

$$[X, X]_t = \int_0^t \sigma^T(s)\sigma(s)ds \tag{2.41}$$

We will then set:

$$\sigma_{ij}^2(t) = \sum_{k=1}^{d} \sigma_{ik}(t)\sigma_{kj}(t), \tag{2.42}$$

and we write $\sigma_{ij}^2(t)$ in the following way:

$$\sigma_{ij}^2(t) = \lim_{n\to\infty} \sum_{k=0}^{n} \left(1 - \frac{k}{n}\right) \cdot \left[a_k(\sigma_{ij}^2)\cos(kt) + b_k(\sigma_{ij}^2)\sin(kt)\right]. \tag{2.43}$$

We then have the following:

Corollary 2.4 *Consider a process X_t satisfying (2.40), and define the component-wise Fourier coefficients of dX and σ^2 as in (2.3). Given an integer $n_0 > 0$, we have in $\mathcal{L}^2$:*

$$a_0(\sigma_{ij}^2) = \lim_{N\to\infty} \frac{\pi}{N + 1 - n_0} \sum_{k=n_0}^{N} \frac{1}{2} \left(a_k^i(dX)a_k^j(dX) + b_k^i(dX)b_k^j(dX)\right) \tag{2.44}$$

$$a_q(\sigma_{ij}^2) = \lim_{N\to\infty} \frac{\pi}{N + 1 - n_0} \sum_{k=n_0}^{N} \left(a_k^i(dX)a_{k+q}^j(dX) + b_k^i(dX)b_{k+q}^j(dX)\right) \tag{2.45}$$

$$b_q(\sigma_{ij}^2) = \lim_{N\to\infty} \frac{\pi}{N + 1 - n_0} \sum_{k=n_0}^{N} \left(a_k^i(dX)b_{k+q}^j(dX) + b_k^i(dX)a_{k+q}^j(dX)\right) \tag{2.46}$$

Proof. We can use the polarization property (Proposition 1.31, 3), that is:

$$[X^i, X^j] = \frac{1}{4}([X^i + X^j, X^i + X^j] - [X^i - X^j, X^i - X^j]) \tag{2.47}$$

Using the fact that $a_k(dX^i \pm dX^j) = a_k(dX^i) \pm a_k(dX^j)$, we get $a_0(\sigma_{ij}^2)$ from (2.27) after substituting $a_k^2(dX)$ with

$$\frac{1}{4}(a_k(dX^i) + a_k(dX^j))^2 - (a_k(dX^i) - a_k(dX^j))^2 = a_k(dX^i)a_k(dX^j). \tag{2.48}$$

which yields the result. □

Chapter 3

Univariate applications

3.1 Introduction

Volatility estimation and forecasting is a critical topic in the financial literature. Indeed, it plays a crucial role in many different fields, e.g., risk management, time series forecasting and contingent claim pricing.

In the last twenty years, starting out from empirical investigations showing that volatility in financial time series is highly persistent with clustering phenomena, many models have been proposed to describe volatility evolution. The literature is now quite large with several specifications of auto-regressive models. Since empirical analysis have shown a high degree of intertemporal volatility persistence, then forecasting with an autoregressive specification should provide satisfactory results, but in many papers it has also been observed that forecasting with GARCH models can be extremely unsatisfactory when the daily volatility is measured ex post by the squared (daily) return, see Andersen and Bollerslev (1998a); Andersen et al. (1999); Figlewski (1997); Pagan and Schwert (1990). Andersen and Bollerslev (1998a) suggest that the main motivation of this failure is that the squared daily return is a very noisy estimator of volatility. Monte Carlo experiments of diffusion processes, whose parameters have been estimated on exchange rate time series (DM-$ and Yen-$), show that the noise of the high frequency volatility estimator is much smaller than that of the daily squared return. Then, they show as the forecasting performance of a GARCH(1,1) model is improved when the daily volatility (*integrated volatility*) is measured by means of the cumulative squared intraday returns. On this topic see also Barndorff-Nielsen and Shephard (2002a,b).

In this Chapter we address volatility estimation and forecasting in a GARCH setting with high frequency data by applying the algorithm described in Chapter 2 to compute the volatility of a diffusion process. This method is based on Fourier analysis techniques (hereafter *Fourier method*). The volatility of a diffusion process is defined as the limit of its quadratic variation. This definition motivates standard volatility estimation methods based on a *differentiation procedure*: equation (1.11) of a process with a given frequency (day, week, month) is taken as a volatility estimate, see e.g. French et al. (1987). Extending this technique to intraday data presents some drawbacks due to the peculiar structure of high frequency data. For example, tick-by-tick data are not equally spaced. In the above cited papers an equally spaced time series for intraday returns is constructed by linearly interpolating logarithmic midpoints of bid-ask adjacent quotes or by taking the last quote before a given reference time (henceforth called imputation method). This procedure induces some distortions in the analysis, e.g. it may generate spurious returns autocorrelation, and it reduces the number of observations. The method adopted in this Chapter avoids these problems; it is based on *integration* of the time series, and it employs all the (irregularly spaced) observations. To compute integrals, we assume the price to be piecewise constant, i.e. the price is constant between two subsequent observations.

Volatility computation by using all the data with the Fourier method should then be more precise. We illustrate this fact through Monte Carlo simulations of a continuous-time GARCH(1,1) model with the parameters estimated in Andersen and Bollerslev (1998a). We also extend the simulation framework to representative models belonging to the SR-SARV(1) class (Andersen, 1994; Meddahi and Renault, 2004; Fleming and Kirby, 2003), which includes GARCH(1,1) as a particular case. We show that, in some settings, the variance of our integrated volatility estimator is smaller than that of the cumulative squared intraday returns. Moreover, the precision of the cumulative squared intraday returns in measuring volatility depends on the procedure employed to build an equally spaced time series. Linear interpolation causes a downward bias which increases with sampling frequency. The imputation method is immune from these drawbacks. When implementing the Fourier method with linearly interpolated observations instead of assuming the price to be piecewise constant, a strong downward bias arises as well.

Through Monte Carlo simulations, we show that, by measuring integrated volatility according to the Fourier method, the forecasting performance of the GARCH(1,1) model, and other models belonging to the SR-SARV(1) class, is better than that obtained by computing volatility according to the cumulative squared intraday returns.

These results are confirmed when the method is applied to compute volatility of exchange rate high frequency time series. We apply the Fourier method to the evaluation of the

forecasting performance of the daily GARCH model and of the intraday GARCH model, as in Andersen et al. (1999). For both the time series considered, the GARCH model forecasts are evaluated to be better if the Fourier method is employed as a volatility estimate instead of the cumulative squared intraday returns.

We then turn to directly modeling volatility measures. A good forecasting model for daily integrated volatility is crucial for VaR estimates. Traditional models regard volatility as a latent factor; here we model it as an observable quantity through an $AR(n)$ model estimated by ordinary least squares. In spite of its simplicity, this model performs better than traditional models (GARCH(1,1) and Riskmetrics).

Finally we investigate the relationship between market activity and volatility in the Italian interbank overnight market. The interesting point is that this market is almost free of information asymmetry-heterogeneity. As a matter of fact, the overnight interest rate market is affected by liquidity conditions and by European Central Bank decisions, hence it is almost impossible for a bank to detain private information on them. Banks trade for pure hedging reasons. We show that definitely the number of contracts, and not trading volume, is associated with interest rate volatility. Our results confirm that liquidity management is the driving force of interest rate movements.

This Chapter is organized as follows. Section 3.2 describes the Monte Carlo experiments and assesses the performance of the Fourier method in measuring and forecasting volatility of simulated data. In Section 3.3 we turn to the analysis of foreign exchange rate data. Section 3.4 describes an application to Value at Risk measurement. In Section 3.5 we analyze the interest rate and volatility dynamics of the Italian money market. Section 3.6 concludes.

3.2 Monte Carlo experiments

In Barucci et al. (2000) we tested the Fourier method on equally spaced data. Monte Carlo experiments simulating a diffusion process with constant volatility showed that the method allows to consistently estimate volatility in a univariate setting and cross-volatilities in a multivariate setting. The precision of the estimate is similar to that of classical methods. When applied to the daily time series of the Dow Jones Industrial and Dow Jones Transportation Index, the Fourier method replicates the volatility estimates obtained by the classical method. Then there is no difference between the Fourier method and classical methods on equally spaced data.

3.2.1 GARCH(1,1) process

Let $p(t) = \log S(t)$, where $S(t)$ is a generic asset price. Following a large literature, we model the asset price through the continuous-time GARCH model (Drost and Werker, 1996):

$$\begin{aligned} dp(t) &= \sigma(t)dW_1(t) \\ d\sigma^2(t) &= \theta\left[\omega - \sigma^2(t)\right]dt + \sqrt{2\lambda\theta}\sigma^2(t)dW_2(t), \end{aligned} \tag{3.1}$$

where θ, ω, λ are constants and W_1, W_2 are two independent Brownian motions. Provided $\int \sigma(t)dW(t)$ is a continuous martingale, our method allows for jumps in the process $\sigma(t)$. Jumps inserted directly in the differential equation driving the price evolution are not allowed, since it would be impossible to disentangle the contribution to quadratic variation of the diffusion coefficient from that of jumps, see Theorem 1.36.[1].

Given a time window $[0, 1]$ (a day, week, month), we wish to compute the integrated volatility of the process, i.e. $\int_0^1 \sigma^2(t)dt$. An unbiased estimator of this quantity is provided by $[p(1) - p(0)]^2$. However, this estimator is very noisy. As early examples, in Andersen and Bollerslev (1998a) for exchange rates and in Martens (2002) for a stock index, it is shown that an estimator with smaller noise is provided by the sum of squared intraday returns:

$$\hat{\sigma}^2(m) = \sum_{i=2}^{m}\left[p\left(\frac{i}{m}\right) - p\left(\frac{i-1}{m}\right)\right]^2. \tag{3.2}$$

The daily squared return corresponds to the case $m = 2$. As described in Muller et al. (1990), if p is not observed at time i/m, $p(i/m)$ is computed as the linear interpolation of two adjacent observations (one before and one after the time i/m). Theoretically, by increasing the frequency of observations, an arbitrary precision in the estimate of the integrated volatility can be reached. This could not be our case because of the interpolation procedure. In most of the papers estimating volatility with high frequency data, (3.2) is computed with $m = 288$, corresponding to 5 minute returns.

We adopt an alternative methodology. The method described in Chapter 2 gives us an estimator of integrated volatility, that is equation (2.30). The computation of $a_0(\sigma^2)$ provides an estimate of the integrated volatility using all the observations. To illustrate the validity of the Fourier approach, we simulate the diffusion process (3.1) by a first-order Euler discretization scheme, see Kloeden and Platen (1992). We use the parameters $\theta = 0.035, \omega = 0.636, \lambda = 0.296$ estimated in Andersen and Bollerslev (1998a) on the daily return time series of the Deutsch Mark-U.S. Dollar exchange rate. Existence of the exact discretization of the

[1]Recently, econometric techniques based on bipower variation have been developed to disentangle the jump component from the volatility component: see Andersen et al. (2004); Barndorff-Nielsen and Shephard (2004c,b).

process (3.1) is guaranteed by Drost and Werker (1996) in a weak sense. On this topic, see also the comprehensive results on temporal aggregation in Meddahi and Renault (2004).

Taking a day as a reference measure, we simulate 24 hours of trading[2] with $dt = 1/86400$ when discretizing (3.1), which corresponds to an observation every second. In order to simulate high frequency unevenly sampled observations, we select a subset of $[1, 86400]$ by extracting the time differences from an exponential distribution with the mean equal to 14 seconds. This choice is motivated by the fact that the empirical distribution of $t_i - t_{i-1}$ can be approximated with an exponential shape and 14 seconds is the mean duration between quotes in the DM-$ exchange rate time series. As a result, we get a data set $(t_k, p(t_k), k = 1, \ldots, N)$ with t_k unevenly sampled, and $\sigma(t)$ recorded every second, so that the generated value of the integrated volatility can be computed. Then we compute the integrated volatility according to three estimators: the squared daily return, the cumulative five minute squared intraday returns with linear interpolation of adjacent observations, and the Fourier estimator (2.27). The implementation of the Fourier estimator is accomplished with the imputation method. Since we observe the process $p(t)$ only at discrete times $t_1, \ldots, t_n$, we set $p(t) = p(t_i)$ in the interval $t_i \leq t < t_{i+1}$, see equation (2.39).

Results are illustrated in Figure 3.1, where the distribution of the normalized difference between the integrated volatility and its estimate is shown. As expected, the squared daily return is a very noisy estimator. As argued in Andersen and Bollerslev (1998a), when estimating volatility with the cumulative squared intraday returns, we get a smaller variance. However, by measuring volatility according to the Fourier method we get a further reduction of the variance, as well as of the measurement bias.

With real data, as those analyzed later in this Chapter, when increasing the sampling frequency of a financial time series we encounter problems related to microstructure effects. In our setting (simulated time series) such effects are not present; as a consequence, by increasing the return sampling frequency from five minutes to, say, one minute the cumulative squared intraday return estimator should perform better. This is definitely not the case. Figure 3.1 shows that increasing the frequency from five minutes to one minute the variance of the cumulative squared intraday return estimator is reduced but a downward bias comes in.

This downward bias does not depend on the estimation method. If we implement the Fourier method by computing the integrals in (2.27) interpolating linearly the prices in the interval $[t_i, t_{i+1}]$, instead of assuming the price to be constant, we get again a downward biased

[2]Foreign exchange rates are traded 24 hours per day, weekends excluded. Roughly, each of the three major markets (London, Tokyo and New York) is open when the other two are closed.

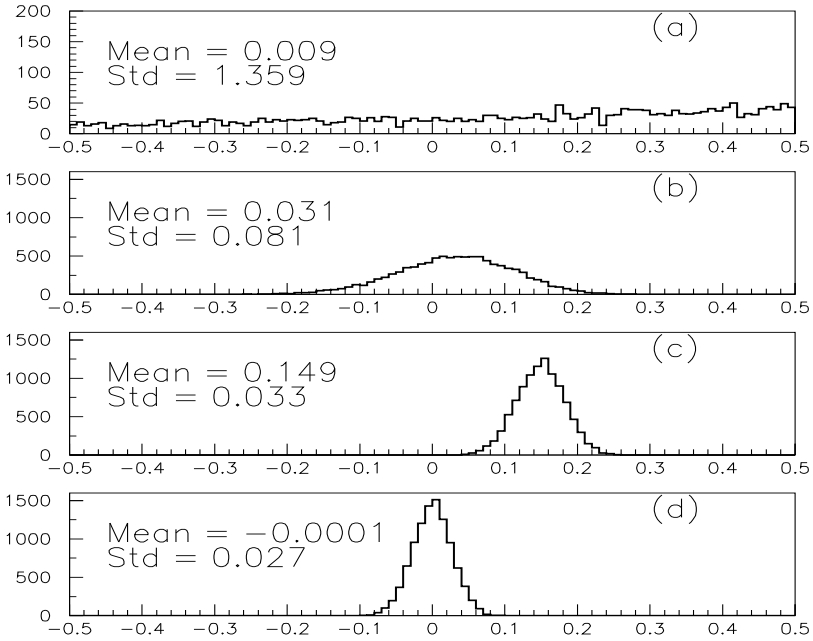

Figure 3.1: Distribution of $\frac{\int_0^1 \sigma^2(t)dt - \hat{\sigma}^2}{\int_0^1 \sigma^2(t)dt}$ where $\hat{\sigma}^2$ is obtained with four different estimators of the integrated volatility: (a) $\hat{\sigma}^2 = [p(1) - p(0)]^2$; (b) $\hat{\sigma}^2 = \sum_{i=2}^{288}[p(\frac{i}{288}) - p(\frac{i-1}{288})]^2$ (five minute estimator); (c) $\hat{\sigma}^2 = \sum_{i=2}^{1440}[p(\frac{i}{1440}) - p(\frac{i-1}{1440})]^2$ (one minute estimator); (d) $\hat{\sigma}^2 = 2\pi a_0(\sigma^2)$ (Fourier estimator). In (b) and (c) returns are linearly interpolated. For each distribution, we indicate its mean and its standard deviation (Std). The distributions are computed with 10,000 "daily" replications of model (3.1).

estimator; on the simulation sample of Figure 3.1, we get a downward bias in the mean of 0.426. Linear interpolation induces spurious autocorrelation; in some sense, a straight line is the "minimum variance" path between two observations.

The downward bias effect of linear interpolation was also conjectured in Corsi et al. (2001). In that paper the authors suggest to use an imputation algorithm which coincides with our piecewise constant assumption. As a matter of fact, taking the last observation before time t as $p(t)$ is equivalent to assume $p(t) = p(t_i)$ in the interval $[t_i, t_{i+1}]$. With this imputation scheme, the cumulative squared intraday return estimator is unbiased, as shown in Figure 3.2 where the cumulative squared intraday returns with the adoption of the imputation scheme is computed for a sampling frequency of two minutes, one minute and 14 seconds. As suggested by the theory, increasing the sampling frequency the variance of this estimator

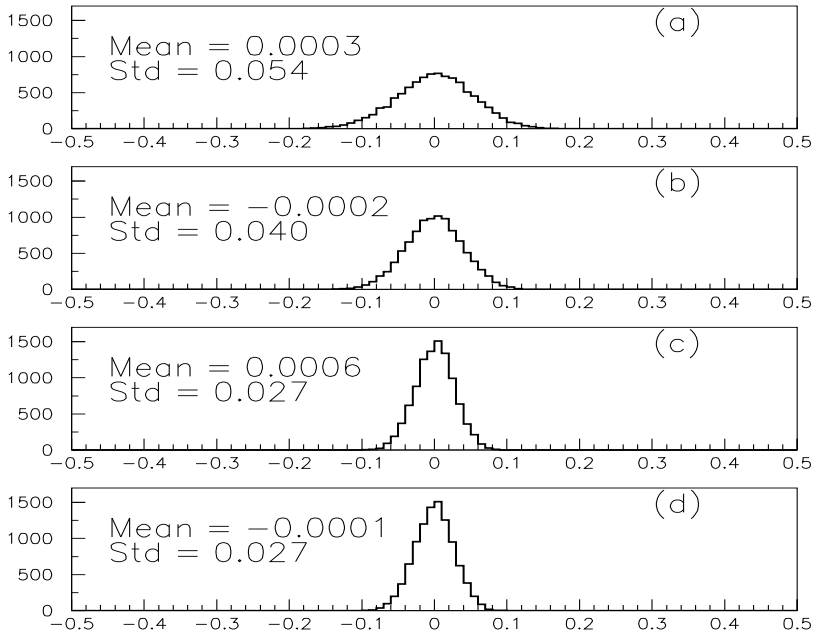

Figure 3.2: Distribution of $\frac{\int_0^1 \sigma^2(t)dt - \hat{\sigma}^2}{\int_0^1 \sigma^2(t)dt}$ where $\hat{\sigma}^2$ is obtained with four different estimators of the integrated volatility: (a) 2-minute cumulative squared intraday returns; (b) 1-minute cumulative squared intraday returns; (c) 14-second cumulative squared intraday returns; (d) Fourier estimator. In (a),(b),(c) returns are obtained with an imputation scheme. For each distribution, we indicate its mean and its standard deviation (*Std*). The distributions are computed with 10,000 "daily" replications of model (3.1).

decreases. In the limit, it converges from above to the variance of the Fourier estimator.

The Fourier estimator is characterized by the smallest variance for any sampling frequency, regardless of the adoption of the imputation or of the linear interpolation scheme for the cumulative squared intraday return estimator. In what follows, when volatility is computed according to the cumulative squared intraday returns, we will aggregate the data through linear interpolation, as it is done in large part of literature.

We now resort to a new set of simulations. First note that the model (3.1) is closed under temporal aggregation in a weak sense, see Drost and Werker (1996); Corradi (2000), and its

discrete time analogous is given by:

$$\begin{aligned} r_t &= \zeta_t \epsilon_t \\ \zeta_{t+1}^2 &= \psi + \alpha \cdot r_t^2 + \beta \cdot \zeta_t^2 \end{aligned} \tag{3.3}$$

where ϵ_t are i.i.d. Normal random variables. The exact relation between (ψ, α, β) and $(\theta, \omega, \lambda)$ is derived in Drost and Werker (1996). We want to test the reliability of our estimates when we change the values of α, β in (3.3). For these new experiments, we extract the observation times in such a way that the time differences are drawn from an exponential distribution with mean equal to $\tau = 45$ seconds.

In this new simulation setting we will also compute the estimator (3.2) with $m = 144$, corresponding to ten-minutes returns, and $m = 720$ corresponding to two-minutes returns; we don't increase further m since the mean time between transactions is 45 seconds. We also include the range, as defined in equation (1.28). We will evaluate the performance of the estimators (2.27) and (3.2) with $m = 144, 288, 720$ by the statistics:

$$\mu = \mathbb{E}\frac{\int_0^1 \sigma^2(s)ds - \hat{\sigma}^2}{\int_0^1 \sigma^2(s)ds}, \quad std = \left[\mathbb{E}\left(\frac{\int_0^1 \sigma^2(s)ds - \hat{\sigma}^2}{\int_0^1 \sigma^2(s)ds}\right)^2\right]^{\frac{1}{2}},$$

where $\hat{\sigma}^2$ is the estimate and $\int_0^1 \sigma^2(s)ds$ is the integrated volatility generated in a simulation, whose value is known in our simulation setting.

We recall that without manipulating the data, we should observe smaller μ and std when increasing the frequency. Figure 3.3 shows the results on simulated time series with $\alpha = 0.25, \beta = 0.7, \psi(1-\alpha-\beta) = 1$. Results are in line with those in Figure 3.1. The range is very noisy, since it uses only two observations (the maximum and the minimum), and it is also very skewed. The ten-minutes estimator (not shown) provides a downward biased estimate of the integrated volatility, with a standard deviation larger than the bias. The five-minutes is also downward biased, with a standard deviation of the same order of the bias in mean. Increasing further the frequency, the estimator is characterized by less variance but a larger bias is observed. The bias is removed if we use previous-tick interpolation instead of linear interpolation. The Fourier estimator is characterized by the smallest bias in mean and by a variance smaller than that of the 5-10 minutes estimate and slightly larger than that of the 2 minutes estimate with linear interpolation. To check the robustness of these results, we repeated the Monte Carlo experiments on a grid of values $(\alpha, \beta, \psi = (1-\alpha-\beta)^{-1})$ with 2 and 5 minutes returns. The results, reported in Table 3.1, can be summarized as follows: the estimator (3.2) turns out to be downward biased ($\mu > 0$), with a bias increasing with m, while the bias of the Fourier estimator is statistically null. If m is chosen in such a way that the bias of (3.2) is less than its standard deviation, then the Fourier estimate provides a smaller standard deviation.

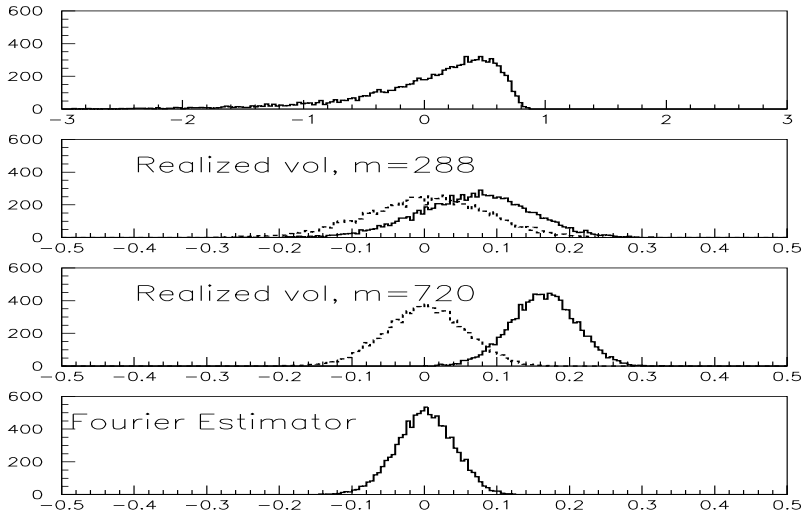

Figure 3.3: Distribution of $\frac{\int_0^1 \sigma^2(t)dt - \hat{\sigma}^2}{\int_0^1 \sigma^2(t)dt}$, where $\hat{\sigma}^2$ are three different estimators of the integrated volatility: (a) the range, equation (1.28); (b) estimator (3.2) with $m = 288$; (the solid line refers to linear interpolation, the dashed line to previous-tick interpolation) (c) estimator (3.2) with $m = 720$; (d) Fourier estimator (2.27). The distribution is computed with 10,000 "daily" replications of model (3.1).

3.2.2 SR-SARV(1) process

For completeness, we checked these results on the following auto-regressive diffusion models:

$$\begin{aligned} &\text{NGARCH}(1,1) \text{ model, Engle and Ng (1993) :} \\ &\sigma_{t+1}^2 = \psi + \alpha(r_t - \gamma\sigma_t)^2 + \beta \cdot \sigma_t^2 \end{aligned} \tag{3.4}$$

$$\begin{aligned} &\text{GJR} - \text{GARCH model, Glosten, Jagannathan and Runkle (1989) :} \\ &\sigma_{t+1}^2 = \psi + \alpha \cdot r_t^2 + \beta \cdot \sigma_t^2 + \delta\theta(-r_t)r_t^2 \end{aligned} \tag{3.5}$$

$$\begin{aligned} &\text{EGARCH model, Nelson (1991) :} \\ &\log(\sigma_{t+1}^2) = \psi + \alpha\sigma_t \cdot (|r_t| + \gamma \cdot r_t) + \beta \cdot \log(\sigma_t^2) \end{aligned} \tag{3.6}$$

where the θ-function is given by $\theta(x) = 1$ if $x > 0$ and $\theta(x) = 0$ if $x \leq 0$. All these models fall in the general class of the $SR - SARV(1)$ models of Andersen (1994), which have the nice property to be closed under temporal aggregation, see Meddahi and Renault (2004); Duan (1997), as well as Fleming and Kirby (2003) for its relation with the GARCH(1,1), so that a continuous diffusion process exists, with the property that the corresponding discrete process

Table 3.1: μ, std, R^2 (multiplied by 100) for the three estimators: (3.2) with $m = 720$ denoted by $2'$, (3.2) with $m = 188$ denoted by $5'$ and (2.27) denoted by F, on a grid of values for (α, β) in (3.3), and $\psi \cdot (1 - \alpha - \beta) = 1$. All the values are computed with 10000 "daily" replications.

$\beta \downarrow$	$\alpha \rightarrow$	0.05			0.1			0.15			0.2			0.25		
		$2'$	$5'$	F	$2'$	$5'$	F	$2'$	$5'$	F	$2'$	$5'$	F	$2'$	$5'$	F
0.5	μ	23.9	10.0	0.1	23.9	10.0	0.1	23.9	10.0	0.1	23.9	10.0	0.1	23.9	10.0	0.1
	std	4.4	7.7	4.9	4.5	7.8	5.0	4.5	7.9	5.0	4.6	7.9	5.1	4.6	8.0	5.1
	R^2	4.92	4.79	4.98	9.58	9.44	9.70	14.0	13.9	14.2	18.6	18.5	18.8	23.8	23.8	24.0
0.6	μ	23.9	10.0	0.1	23.9	10.0	0.1	23.9	10.0	0.1	23.9	10.0	0.1	23.9	10.0	0.1
	std	4.4	7.7	4.9	4.5	7.8	4.9	4.5	7.8	5.0	4.5	7.8	5.0	4.5	7.8	5.0
	R^2	6.79	6.62	6.88	13.8	13.7	14.0	21.9	21.8	22.1	32.5	32.5	33.0	44.6	44.5	45.6
0.7	μ	23.9	10.0	0.1	23.9	10.0	0.1	23.9	10.0	0.1	23.9	10.0	0.1	23.9	10.0	0.1
	std	4.4	7.7	4.9	4.4	7.7	4.9	4.4	7.7	4.9	4.5	7.8	4.9	4.5	7.8	4.9
	R^2	9.67	9.43	9.79	21.5	21.3	21.7	38.2	37.9	38.7	51.3	51.1	52.2	48.2	48.1	48.8
0.8	μ	23.9	10.0	0.1	23.9	10.0	0.1	23.9	10.0	0.1						
	std	4.4	7.6	4.9	4.4	7.7	4.9	4.4	7.7	4.9						
	R^2	15.2	14.7	15.3	36.9	36.5	37.2	54.3	53.8	54.5						
0.9	μ	23.9	10.0	0.1												
	std	4.4	7.6	4.8												
	R^2	30.8	29.9	30.9												

Table 3.2: μ, std, R^2 (multiplied by 100) for the three estimators, (3.2) with $m = 720$ denoted by $2'$, (3.2) with $m = 288$ denoted by $5'$ and (2.27) denoted by F, for the diffusion processes (3.4-3.5-3.6) with the reported parameter values. All the values are computed with 10000 "daily" replications.

	GJR			NGARCH			EGARCH		
	$\psi = 0.0587,\ \alpha = 0.0312$			$\psi = 0.0554,\ \alpha = 0.0952$			$\psi = -0.1491,\ \alpha = 0.1786$		
	$\beta = 0.8275,\ \delta = 0.1271$			$\beta = 0.8001,\ \gamma = 0.6048$			$\beta = 0.9512,\ \gamma = -0.4815$		
	$2'$	$5'$	F	$2'$	$5'$	F	$2'$	$5'$	F
μ	23.9	10.0	-0.3	23.9	10.0	-0.5	23.9	10.0	-0.2
std	4.4	7.6	4.9	4.4	7.6	5.0	4.4	7.6	4.9
R^2	37.8	37.0	37.9	47.3	46.4	47.4	46.6	45.5	46.8

is its exact discretization. We simulated these processes with the parameters estimated in Duan (1997) and reported in Table 3.2. Table 3.2 reports also μ, std, R^2 for each model. The results in Table 3.2 confirm those obtained with the continuous GARCH model, i.e. the Fourier estimator has a smaller bias and an higher precision than the estimator (3.2).

3.2.3 Volatility forecasting evaluation

Following Drost and Nijman (1993); Drost and Werker (1996), the GARCH continuous time diffusion (3.1) can be discretized, obtaining the weak GARCH(1,1) process (3.3). In this setting, ζ_{t+1}^2 provides us with an unbiased forecast of $\int_0^1 \sigma^2(t+\tau)d\tau$. While there is a strong support in favor of a high persistence in the volatility dynamics, the one day ahead forecasting

performance of the above model has been evaluated to be very poor in the literature (Jorion, 1995).

Using simulated time series, we employ the Fourier method to measure the realized integrated volatility, and we evaluate the GARCH(1,1) model forecasting performance according to it. The GARCH model performance with respect to this estimator is compared to that associated with the cumulative squared intraday returns with linearly interpolated observations. Following Andersen and Bollerslev (1998a), the forecasting performance of the GARCH model can be evaluated according to the R^2 of the linear regression:

$$\hat{\sigma}_t^2 = a + b \cdot \zeta_t^2 + \epsilon_t, \tag{3.7}$$

which is given by

$$R^2 = \left[corr\left(\hat{\sigma}_t^2, \zeta_t^2\right)\right]^2, \tag{3.8}$$

where $\hat{\sigma}^2$ is the ex post integrated volatility estimator.

We use the simulation setting described in Section 3.2.1. For the Y-$ time series we set the mean duration equal to 52 seconds. The forecasting model is given by (3.3) with the parameters as in Table 3.5 for $m = 1$, corresponding to those employed in the above Section according to Drost and Werker (1996).

We point out that the R^2 obtained in (3.8) must be compared to the R^2 obtained when the ex post integrated volatility measure is perfectly known; its value is given by

$$R_\infty^2 = \left[corr\left(\int_{t-1}^{t} \sigma^2(s)ds, \zeta_t^2\right)\right]^2, \tag{3.9}$$

a value that can be computed in our simulation setting.

Results are shown in Table 3.3: the Fourier estimator gives an R^2 which is very close to R_∞^2. Employing the Fourier estimator, the GARCH forecasting performance is better than that obtained by measuring integrated volatility through the sum of squared intraday returns.

The same analysis has been performed for different parameter values of the GARCH(1,1) and the models (3.4),(3.5),(3.6). Results, shown in Tables 3.1 and 3.2 respectively confirm the previous conclusions.

3.3 Foreign exchange rate analysis

The data set at hand consists of the one year (from October, 1^{st} 1992 to September 30^{th} 1993) collection of tick-by-tick quotes (bid and ask) of the Deutsch Mark-U.S. Dollar exchange rate

Table 3.3: R^2 obtained on simulated data with different estimators of integrated volatility. The values are computed through 50,000 "daily" replications. The table reports also the estimates $\hat{a}, \hat{b}$ of the regression (3.7), together with their standard errors (s.e.) and the Mean Square Error (MSE). Ljung-Box test on the residuals (not reported) strongly reject the null of zero auto-correlation.

DM-$						
Estimator	R^2	$\hat{a}$	s.e.	$\hat{b}$	s.e.	MSE
$(p(1)-p(0))^2$	0.062	-0.620	0.011	1.954	0.015	1.33
$\sum_{i=2}^{288}\left[p\left(\frac{i}{288}\right)-p\left(\frac{i-1}{288}\right)\right]^2$	0.476	-0.003	0.003	0.971	0.004	0.088
Fourier	0.489	-0.002	0.003	1.004	0.004	0.090
R^2_∞	0.491	-0.002	0.003	1.005	0.004	0.089
Y-$						
Estimator	R^2	$\hat{a}$	s.e.	$\hat{b}$	s.e.	MSE
$(p(1)-p(0))^2$	0.092	-0.438	0.007	1.869	0.011	1.31
$\sum_{i=2}^{288}\left[p\left(\frac{i}{288}\right)-p\left(\frac{i-1}{288}\right)\right]^2$	0.488	-0.050	0.003	1.095	0.004	0.147
Fourier	0.501	-0.049	0.003	1.096	0.004	0.147
R^2_∞	0.505	-0.052	0.003	1.104	0.004	0.148

and of Japanese Yen-U.S. Dollar exchange rate, with time stamps rounded to the nearest even second. The data set was collected and delivered by Olsen & Associates. This data set has been extensively studied, e.g. see Andersen and Bollerslev (1997, 1998b,a); Andersen et al. (2001, 1999); Guillaume et al. (1997); Muller et al. (1990, 1997); Zumbach (2000).

We define the price to be the mid-price between bid and ask quotes. The trading day is chosen to begin and to end at 21:00 GMT. We excluded weekends and trading days with less than 1000 quotes for the DM-$ and less than 320 quotes for the Y-$. Moreover we applied the filter described in Dacorogna et al. (1993) which removes roughly 0.36 % of the quotes. We end up with $1{\cdot}466{\cdot}944$ quotes for the DM-$ and $567{\cdot}758$ quotes for Y-$, distributed over 258 trading days.

When applying the Fourier estimator to a high frequency time series, we compute $a_0(\sigma^2)$ through the expansion (2.27) stopped at some frequency N. In the previous Section we used the Nyquist frequency, $N = n/2$, where n is the number of data, see Priestley (1979). With real data we encounter a severe difficulty: the diffusion model (3.1) does not hold for small time steps because of microstructure effects such as price discreteness or bid-ask bounce effect[3]. Microstructure effects jeopardize the computation of the Fourier coefficients

[3]See Madhavan (2000) for a survey. See also the recent papers of Bandi and Russell (2003); Ait-Sahalia et al. (2003); Andreou and Ghysels (2004).

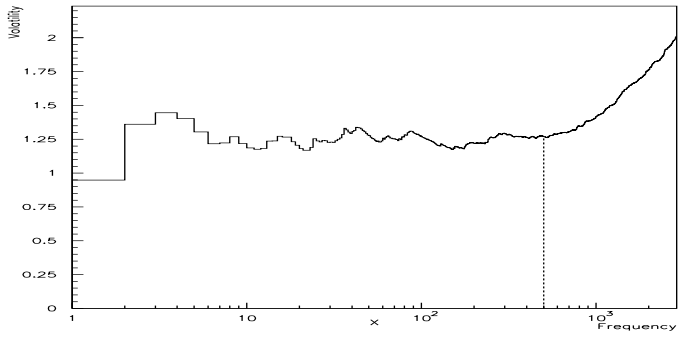

Figure 3.4: $\sqrt{2\pi a_0(\sigma^2)}$ computed according to (2.27) as a function of n for the DM-\$ exchange rate (October 1^{st}, 1992). The dashed line indicates the cut-off frequency that we adopt to compute integrated volatility.

at high frequencies. This fact is shown in Figure 3.4 where the square root of the integrated volatility, computed according to (2.27), is plotted as a function of the highest frequency n employed in the sum (2.27), with n ranging from $n_0 = 1$ to $N/2$. The plot in Figure 3.4 can be interpreted as the mean of the power spectrum of the exchange rate from frequency 0 to n. If $dp(t)$ is Normally distributed, as in a model like (3.1), then the spectrum of p should be flat. Figure 3.4 shows that the power spectrum of p is not flat[4]; for a frequency larger than a certain value (denoted by N_{cut}) the Fourier coefficients become considerably higher than the lower frequency coefficients. In our setting, it turns out that $N_{cut} \simeq 500$ for the DM-\$ exchange rate and $N_{cut} \simeq 160$ for the Y-\$. These frequencies correspond roughly to a time step, computed as $\frac{86400}{2 \cdot N_{cut}}$ seconds, of 1.5 and 4.5 minutes respectively. We conclude that the price process cannot be modeled by (3.1) for time steps smaller than two (five) minutes for the DM-\$ (Y-\$) exchange rate.

This behavior of the exchange rate spectrum can be motivated by the fact that high frequency returns are negatively correlated, a phenomenon that has been documented by Andersen and Bollerslev (1997); Bollerslev and Domowitz (1993) among others. We can show this by simulating the process

$$\begin{cases} \sigma^2_{t+1} = \psi + \alpha \cdot p_t^2 + \beta \cdot \sigma_t^2 \\ \epsilon_{t+1} = \rho \epsilon_t + \sqrt{1-\rho^2}\eta_t \\ p_{t+1} = \sigma_{t+1}\epsilon_t, \end{cases} \tag{3.10}$$

where ρ is the first-order serial correlation coefficient and η_t is a sequence of i.i.d. Normal distributed random variables with mean 0 and variance 1. Figure 3.5 shows the analogous of Figure 3.4 for a time series simulated according to (3.10) with $\rho = -0.985$ and a time step of

[4]A plot analogous to Figure 3.4 can be found in Andersen et al. (2000b).

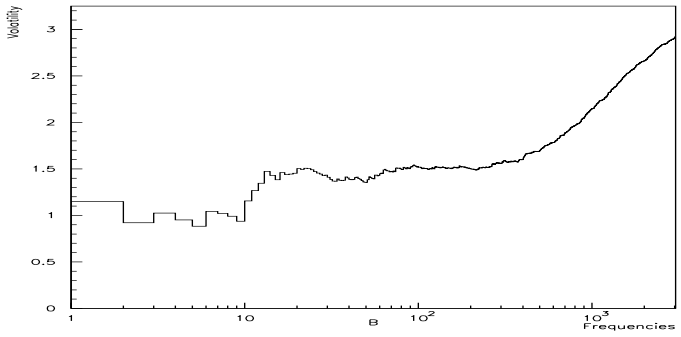

Figure 3.5: $\sqrt{2\pi a_0(\sigma^2)}$ for the simulated process (3.10) as a function of the frequency n in (2.27).

one second: the phenomenon illustrated in Figure 3.4 occurs. According to our simulations, smaller values of $|\rho|$ would lead to a larger cut-off frequency. Such a negative correlation can be linked to non-synchronous trading (Lo and MacKinlay, 1990), to the management of inventory positions by market makers (Andersen and Bollerslev, 1997) and to the bid-ask bounce effect (Madhavan, 2000). Oomen (2002) studies the impact of serial correlation on realized volatility.

In what follows, we cut the highest frequencies in the computation of the integrated volatility, i.e. we evaluate the Fourier coefficient (2.5) for $n = \min(N/2, N_{cut})$. In the literature on volatility computation with high frequency data, microstructure effects are attenuated by aggregating the data through linear interpolation or an imputation scheme, building a five minute return time series. With our method we do not interpolate nor aggregate the original time series and we use all the data in order to compute the Fourier coefficients of dp, we only stop expansion (2.27) properly.

We evaluate the GARCH(1,1) model forecasting performance, by looking at the one-step-ahead daily forecasts, when the integrated volatility is computed according to the Fourier method. The parameters of the model are those estimated in Andersen and Bollerslev (1998a) on the same time series. In Figure 3.6, we show the forecasts of the GARCH model together with the integrated volatility of the DM-$ exchange rate computed according to the Fourier method. Table 3.4 shows the corresponding R^2. We observe that the GARCH model is evaluated to perform quite well in forecasting when the Fourier method is employed to compute the integrated volatility. Its performance is better than that associated with the cumulative squared intraday returns (with linear interpolation of observations) as an integrated volatility measure. The poor performance on the Y-$ time series, when compared

Table 3.4: R^2 for the two foreign exchange rate time series. The table reports also the estimates $\hat{a}, \hat{b}$ of the regression (3.7), together with their standard errors (s.e.), the Mean Square Error (MSE) and the Ljung-Box test on residuals at lag 10, LB(10).

DM-$							
Estimator	R^2	$\hat{a}$	s.e.	$\hat{b}$	s.e.	MSE	LB(10)
$\sum_{i=2}^{288}\left[p\left(\frac{i}{288}\right)-p\left(\frac{i-1}{288}\right)\right]^2$	0.400	0.006	0.047	0.99	0.07	0.082	25.72
Fourier	0.470	0.098	0.042	1.00	0.06	0.081	40.64
Y-$							
Estimator	R^2	$\hat{a}$	s.e.	$\hat{b}$	s.e.	MSE	LB(10)
$\sum_{i=2}^{288}\left[p\left(\frac{i}{288}\right)-p\left(\frac{i-1}{288}\right)\right]^2$	0.128	0.149	0.072	0.86	0.11	0.220	44.43
Fourier	0.143	0.070	0.070	0.89	0.11	0.210	41.72

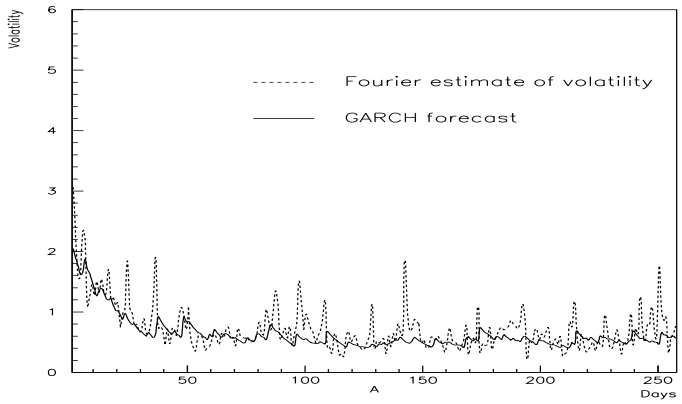

Figure 3.6: Comparison between GARCH model forecasts and realized volatilities for the DM-$ exchange rate, from October, 2^{th} 1992 to September, 30^{th} 1993. Realized volatility is measured with the Fourier estimator.

to that obtained in simulated data in Table 3.3, can be explained by the presence of few days with very high volatility, see also Andersen and Bollerslev (1998a).

Table 3.5: Coefficients of the GARCH(1,1) model at different frequencies, obtained (according to Drost and Werker (1996)) from the continuous-time coefficients $\theta = 0.035, \omega = 0.636, \lambda = 0.296$ for the DM-$ exchange rate and $\theta = 0.054, \omega = 0.476, \lambda = 0.480$ for the Y-$ exchange rate as estimated in Andersen and Bollerslev (1998a), $\psi_m = \sigma_m^2(1 - \alpha_m - \beta_m)$.

DM-$				Y-$			
m	σ_m^2	α_m	β_m	m	σ_m^2	α_m	β_m
1	0.6365	0.0679	0.8978	1	0.4760	0.1043	0.8431
2	0.3180	0.0541	0.9285	2	0.2380	0.0840	0.8893
3	0.2122	0.0466	0.9418	3	0.1587	0.0726	0.9095
4	0.1590	0.0417	0.9496	4	0.1190	0.0651	0.9215
6	0.1061	0.0353	0.9589	6	0.0793	0.0553	0.9358
12	0.0530	0.0262	0.9709	12	0.0397	0.0411	0.9544
24	0.0265	0.0192	0.9794	24	0.0198	0.0302	0.9676
48	0.0133	0.0139	0.9854	48	0.0099	0.0219	0.9770
96	0.0066	0.0100	0.9896	96	0.0050	0.0157	0.9837
144	0.0044	0.0082	0.9915	144	0.0033	0.0130	0.9867
288	0.0022	0.0059	0.9940	288	0.0016	0.0093	0.9906

3.3.1 Forecasting daily exchange rate volatility using intraday returns

Since temporal aggregation of the continuous-time GARCH process (3.1) holds, we can discretize it at any frequency in a straightforward manner. Denote

$$r_m(t) = p(t) - p\left(t - \frac{1}{m}\right),$$

then we can write:

$$\zeta_m^2(t) = \psi_m + \alpha_m \cdot r_m^2\left(t - \frac{1}{m}\right) + \beta_m \cdot \zeta_m^2\left(t - \frac{1}{m}\right), \tag{3.11}$$

where $\zeta_m^2(t)$ is the best linear predictor of $r_m^2(t)$ expressed as a linear combination of lagged squared intraday returns. The relation between $(\psi_m, \alpha_m, \beta_m)$ in equation (3.11) and $(\omega, \theta, \lambda)$ in equation (3.1) can be obtained for every m in closed form following Drost and Werker (1996); (3.3) corresponds to (3.11) with $m = 1$. Table 3.5 reports the GARCH(1,1) coefficients at different frequencies for the DM-$ and Y-$ exchange rate time series.

Following Andersen et al. (1999), the forecast for the integrated volatility $\left(\int_t^{t+h}\sigma^2(s)ds\right)$ using returns spaced by $1/m$ is given by

$$F_{m,h}(t) = mh\sigma_m^2 + \frac{\alpha_m+\beta_m}{1-\alpha_m-\beta_m}[1-(\alpha_m+\beta_m)^{m\cdot h}](\zeta_m^2(t)-\sigma_m^2), \tag{3.12}$$

where $\sigma_m^2 = \psi_m \cdot (1-\alpha_m-\beta_m)^{-1}$. The realized ex-post integrated volatility will be denoted by $\hat{\sigma}_h^2(t)$. In what follows, we will concentrate only on daily forecast evaluation ($h=1$).

We choose to evaluate the GARCH(1,1) model forecasting performance with R^2 and the following statistics:

$$RMSE = \mathbb{E}[(\hat{\sigma}_h^2(t) - F_{m,h}(t))^2]^{\frac{1}{2}},$$

$$HRMSE = \mathbb{E}[(1 - F_{m,h}(t)/\hat{\sigma}_h^2(t))^2]^{\frac{1}{2}}.$$

We analyze the GARCH(1,1) model forecasting performance using intraday returns both on simulated time series and on the DM-$ and Y-$ exchange rate time series.

Using the technique described in Section 3.2.3, we simulate the time series with the parameters of the DM-$ exchange rate, setting the mean duration to 14 seconds and $N_{cut} = 500$. The results, presented in Table 3.6, are in agreement with those in Andersen et al. (1999); Martens (2001): increasing the sampling of intraday returns, an improvement of the forecasting performance is observed. We remark that the forecasting performance associated with the Fourier estimator is better than that associated with the 5-minute estimator, with the exception of $RMSE$ which is slightly larger for $m = 1, 2, 3$. However, at frequencies higher than $m = 3$ the GARCH model is evaluated to perform better when volatility is estimated with the Fourier method, than when it is estimated with the cumulative squared intraday returns. These results are confirmed for all the frequencies by adjusting the RMSE for heteroskedasticity. In Table 3.6 we also report the results when the largest frequencies are included in the computation. As expected, on simulated data this inclusion leads to a performance improvement.

The above statistics are reported in Table 3.7 for the DM-$, and in Table 3.8 for the Y-$ time series, with $\hat{\sigma}_h^2(t)$ computed as the sum of 5-minute squared intraday returns and with the Fourier method. We stress that we are completely neglecting intraday patterns and macro-economic announcement effects, which have been documented to be important at the intraday level, see Andersen and Bollerslev (1998a); Martens (2001). We also stress that we are neglecting the fact that temporal aggregation of the continuous time GARCH process (3.1) has not been confirmed empirically, see for example Andersen and Bollerslev (1998c); Zumbach (2000). Our results on the forecasting performance of the GARCH model as a function of the sampling frequency are substantially in agreement with those reported in

Table 3.6: Summary statistics of the GARCH(1,1) model forecasts for the simulated time series of daily volatility when returns are spaced by $1/m$ days. Between parenthesis we report the values when frequencies larger than N_{cut} are included in the computation. Results are computed with 10,000 "daily" replications.

5 minute returns				Fourier Estimator			
m	R^2	RMSE	HRMSE	m	R^2	RMSE	HRMSE
1	0.374	0.293	0.542	1	0.376 (0.378)	0.296 (0.296)	0.371 (0.357)
2	0.540	0.250	0.461	2	0.544 (0.545)	0.251 (0.251)	0.423 (0.405)
3	0.584	0.238	0.425	3	0.589 (0.590)	0.239 (0.239)	0.390 (0.372)
4	0.640	0.222	0.391	4	0.646 (0.647)	0.222 (0.221)	0.357 (0.341)
6	0.696	0.203	0.356	6	0.704 (0.706)	0.202 (0.202)	0.323 (0.308)
12	0.766	0.179	0.304	12	0.777 (0.780)	0.176 (0.175)	0.272 (0.258)
24	0.815	0.158	0.263	24	0.827 (0.830)	0.154 (0.153)	0.233 (0.221)
48	0.853	0.141	0.228	48	0.866 (0.868)	0.136 (0.135)	0.199 (0.190)
96	0.877	0.128	0.201	96	0.891 (0.894)	0.122 (0.121)	0.173 (0.165)
144	0.893	0.120	0.186	144	0.906 (0.909)	0.113 (0.112)	0.159 (0.153)
288	0.906	0.111	0.165	288	0.921 (0.924)	0.105 (0.105)	0.141 (0.137)

Andersen et al. (1999); Martens (2001). Volatility forecasting improves when the GARCH model is discretized at intraday frequencies, but this effect has an intrinsic limit due to the fact that beyond a certain time scale intraday features and microstructure effects become prominent. As in previous studies, we confirm the observation that such a time scale is around few hours. This turns out to be true also for the Y-$ time series, which has not yet been analyzed from this perspective. For the DM-$ time series the best R^2 and $HRMSE$ are obtained at $m = 4$ (six hour returns) for both estimators, while on the Y-$ time series the best R^2 is obtained at $m = 48$ (two hours returns) and the best $HRMSE$ is obtained at $m = 4$ (six hours returns) for both estimators.

In general, the GARCH model forecasting performance associated with the Fourier estimator is better than that associated with the cumulative squared intraday returns. For the DM-$ time series, the R^2 associated with the Fourier estimator is higher than that of the five minute returns at any m, while the $HRMSE$ is largely lower. The $RMSE$ reports similar results for the two methods, this confirms what has been shown in Table 3.6 with simulated time series. For the Y-$ exchange rate and $m = 1, 2, 3, 6, 144$, we find that the Fourier estimator performs better than the cumulative squared intraday returns with all the statistics. For

Table 3.7: Summary statistics of the GARCH(1,1) forecasts for the DM-$ daily volatility when returns are spaced by $1/m$ days.

5 minute returns				Fourier Estimator			
m	R^2	RMSE	HRMSE	m	R^2	RMSE	HRMSE
1	0.400	0.299	0.619	1	0.470	0.292	0.377
2	0.413	0.296	0.545	2	0.491	0.300	0.362
3	0.414	0.307	0.516	3	0.490	0.311	0.362
4	0.445	0.293	0.469	4	0.513	0.308	0.355
6	0.434	0.308	0.522	6	0.502	0.313	0.384
12	0.424	0.311	0.486	12	0.494	0.323	0.389
24	0.422	0.314	0.549	24	0.484	0.326	0.430
48	0.415	0.324	0.565	48	0.473	0.329	0.431
96	0.403	0.331	0.653	96	0.453	0.326	0.473
144	0.406	0.336	0.678	144	0.449	0.320	0.469
288	0.404	0.380	0.798	288	0.436	0.356	0.558

Table 3.8: Summary statistics of the GARCH(1,1) forecasts for the Y-$ daily volatility when returns are spaced by $1/m$ days.

5 minute returns				Fourier Estimator			
m	R^2	RMSE	HRMSE	m	R^2	RMSE	HRMSE
1	0.128	0.503	0.588	1	0.143	0.493	0.562
2	0.129	0.521	0.535	2	0.138	0.514	0.531
3	0.169	0.536	0.662	3	0.171	0.532	0.632
4	0.237	0.479	0.454	4	0.235	0.478	0.461
6	0.219	0.520	0.559	6	0.221	0.517	0.548
12	0.263	0.479	0.484	12	0.261	0.477	0.477
24	0.275	0.474	0.507	24	0.270	0.473	0.504
48	0.290	0.466	0.513	48	0.283	0.466	0.513
96	0.266	0.478	0.581	96	0.255	0.480	0.595
144	0.233	0.511	0.664	144	0.236	0.507	0.660
288	0.220	0.520	0.765	288	0.219	0.517	0.758

$m = 4, 12, 24$ the results are similar; with $m = 96$ the 5-minute estimator performs better. At high frequency the results are not clear cut; this can be due to intraday patterns, low liquidity of the time series and the breakdown of the GARCH temporal aggregation properties.

3.4 A linear model for volatility

The role played by volatility in most financial applications is crucial, especially in risk management, where Value at Risk (VaR) estimates are mandatory for regulatory reasons and asset allocation decisions.

In the recent years, literature focused on the role of integrated volatility; the importance of integrated daily volatility in VaR applications relies on two facts. First of all, it has been shown that it is possible to measure it by using intra-day data with very good precision, see Andersen and Bollerslev (1998a); Barucci and Renò (2002a,b), paralleling the use of daily returns in computing monthly volatility, see e.g. French et al. (1987); Schwert (1989, 1990, 1998). Second, empirical studies (Andersen et al., 2000a; Bollen and Inder, 2002) showed that the distribution of returns divided by the square root of the integrated volatility can be well approximated by a Gaussian distribution with zero mean and variance equal to one. This fact means that VaR estimates are linked to integrated volatility forecasting, since the quantiles of the return distribution can be extracted by a Gaussian distribution with zero mean and variance given by the integrated volatility.

On the other hand, persistence properties displayed by volatility suggest that daily volatility can be forecasted with reliable precision. Typically, volatility models regard it as a latent factor which drives asset prices-returns (ARCH, GARCH models). This is the approach followed also in Beltratti and Morana (1999), where high frequency data are used to estimate the dynamical model for latent volatility (FIGARCH model) and to compute VaR.

In this Section, we will model directly the integrated volatility as an observable quantity through a simple $AR(n)$ model. A similar approach has been proposed by Corsi et al. (2001) and Andersen et al. (2003) in a multivariate setting. The main difference is provided by the computation method of the integrated volatility. In the above papers, integrated volatility is computed by using an equally spaced high frequency time series (typically five minute returns) as the sum of squared intraday logarithmic returns. Our procedure instead is based on the Fourier method.

We will compare the performance of this model to that of the GARCH(1,1) model or Riskmetrics, which is very popular among practitioners. We show that, though the $AR(n)$ is

quite simple, it performs better than traditional models in forecasting daily volatility. Our findings suggest that constructing directly a model for volatility based on the measurements of daily integrated volatility, instead of modeling volatility as a latent factor, can be a good modeling tool.

The model and results

Being able to measure daily volatility with good precision, we try to model integrated volatility as an observed quantity, instead of a latent factor as in GARCH models. We use the simplest possible model for the time evolution of the integrated volatility, i.e. an $AR(n)$ model:

$$\hat{\sigma}_t^2 = \sigma_0^2 + \sum_{i=1}^{n} \alpha_i \hat{\sigma}_{t-i}^2 + \varepsilon_t \tag{3.13}$$

with $\mathbb{E}[\varepsilon_t] = 0,\ \mathbb{E}[\varepsilon_t^2] = \Sigma^2$. The parameters $\sigma_0^2, \Sigma^2, \alpha_1, \ldots, \alpha_n$ can be estimated by ordinary least squares (OLS) and $\hat{\sigma}_t^2$ is measured by the integrated volatility estimator in (2.27). We will use (3.13) to forecast volatility.

We compare this very simple volatility forecasting model with two models largely used by practitioners, the GARCH(1,1) model, where future volatility is estimated as:

$$\hat{\sigma}_{t+1}^2 = \psi + \alpha \cdot r_t^2 + \beta \cdot \hat{\sigma}_t^2, \tag{3.14}$$

where $r(t) = p(t) - p(t-1)$ is the daily return, and the model used by RiskMetrics, which estimates the future volatility as a sum of past realizations with exponentially declining weights:

$$\hat{\sigma}_{t+1}^2 = \frac{\displaystyle\sum_{i=0}^{M-1} \lambda^i \left(r_{t-i} - \frac{1}{M} \sum_{i=0}^{M-1} r_{t-i} \right)^2}{\displaystyle\sum_{i=0}^{M-1} \lambda^i} \tag{3.15}$$

where $\lambda = 0.94$.

Other authors try to model the integrated volatility, measuring it via the sum of squared intraday returns. In Corsi et al. (2001) it is found that a EMA-HAR model for the integrated volatility performs better than Riskmetrics. In Andersen et al. (2003) a tri-variate vector auto-regression (VAR), which incorporates long memory effects, is fitted on the DM-$ and Y-$ foreign exchange time-series. These authors choose a polynomial of lag 5, and find that this model performs largely better than GARCH(1,1) and Riskmetrics.

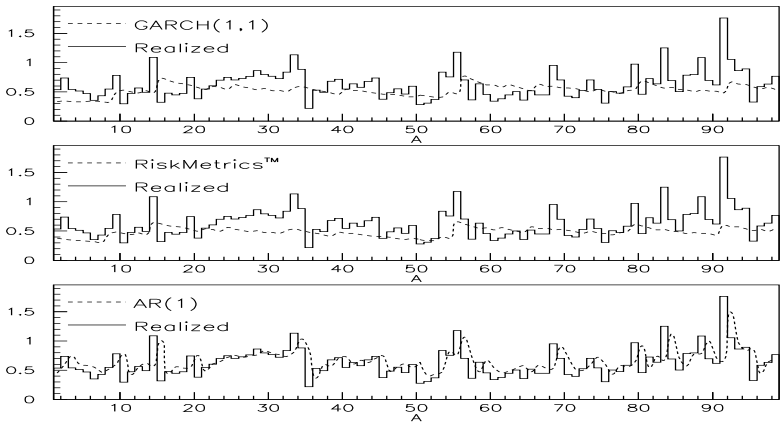

Figure 3.7: Out-of-sample measured integrated volatility for the DM-$ time series (solid line) together with the forecast (dashed line) of the GARCH(1,1) model (top), Riskmetrics (center) and $AR(1)$ model (bottom).

The data set under study in this Section is again the one-year collection of bid-ask quotes of the Deutshche Mark-U.S. Dollar and Japanese Yen-U.S. Dollar exchange rates, as they appeared on the Reuters screen from October, 1^{st} 1992 to September, 30^{th} 1993. In the expansion (2.27) we use N as in Barucci and Renò (2002a), i.e. $N = 500$ for the DM-$ time series and $N = 160$ for the Y-$ time series. We discard days in which the observations are less than 1000 and 320 respectively; we end up with 258 daily volatility measurements for the DM-$ time series and 259 for the Y-$ time series. We divide our samples in 160 days for in-sample model estimate and the remainder for out-of-sample comparison.

Table 3.9 shows the OLS in-sample estimates of model (3.13), together with standard errors, the in-sample R^2 and the R^2 adjusted for degrees of freedom. For the model (3.3) we use the estimates given in Andersen and Bollerslev (1998a), i.e. $\psi = 0.022, \alpha = 0.068, \beta = 0.898$ for the DM-$ time series and $\psi = 0.026, \alpha = 0.104, \beta = 0.844$ for the Y-$ time series. For the model (3.15) we will use $M = 160$, i.e. the largest M at our disposal.

Table 3.10 compares the forecasting performance. In spite of its simplicity, model (3.13) performs considerably better than (3.3) and (3.15). For the DM-$ time series, it is already true with $n = 1$; however, by increasing the order of the auto-regressive model, we find better results. We interpret this finding as an evidence of long-memory effects in the volatility evolution. For the Y-$ time series it is necessary to employ $n = 2$, while the best result is obtained with $n = 5$. Figure 3.7 and 3.8 show the comparison between the integrated out-of-sample volatilities and the forecasts of the three models for the two exchange rate

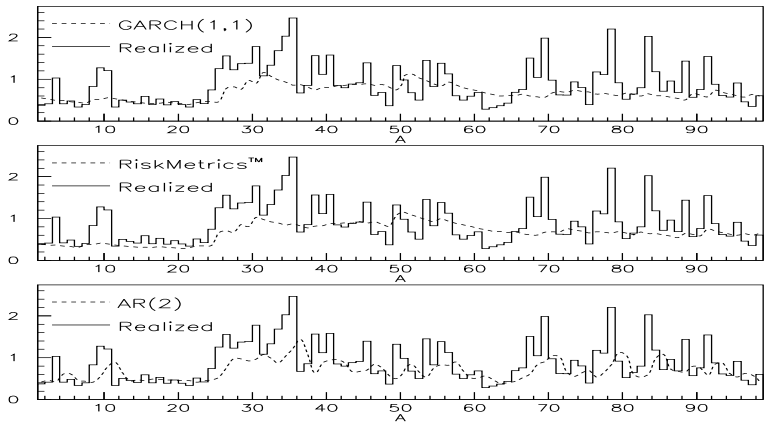

Figure 3.8: Out-of-sample measured integrated volatility for the Y-$ time series (solid line) together with the forecast (dashed line) of the GARCH(1,1) model (top), Riskmetrics (center) and $AR(2)$ model (bottom).

time series. Also visual inspection confirms that the simple $AR(1), AR(2)$ model does a good job in tracking the volatility time series. We interpret the results of our simple exercise as a confirmation that the use of high frequency data in measuring volatility, and directly modeling the integrated volatility dynamics, can substantially improve volatility forecasting, thus Value at Risk estimates.

3.5 The volatility of the Italian overnight market

What are the driving forces of short term interest rate movements?

This Section tries to answer this question in two steps. First, the so-called *martingale hypothesis* is tested, which predicts that interest rate changes should not be predictable within the so-called maintenance period[5]. Then, it is investigated the relationship between interest rate movements and several market activity measures: trading volume, number of contracts, payments values and order flow. The analysis concerns the Italian overnight interest rate time series since the start of the EMU.

[5]During the maintenance period banks have to keep the required reserve of the accounts they hold with the central bank. The requirement is determined as a percentage of the previous calendar period deposits (a month in the EMU). The required balance can be mobilized by banks on a daily basis, but must be kept on average over the maintenance period.

Table 3.9: OLS in-sample estimates and R^2 of model (3.13) for the two time series. In parenthesis, we report the standard error and the R^2 adjusted by the degrees of freedom.

Model	FX	σ_0^2	α_1	α_2	α_3	α_4	α_5	R^2 (R^2_{adj})
AR(1)	DM-$	0.209	0.728					0.245
		(0.048)	(0.051)					(0.245)
AR(2)	DM-$	0.225	0.575	0.122				0.246
		(0.047)	(0.075)	(0.072)				(0.242)
AR(3)	DM-$	0.194	0.541	-0.018	0.209			0.240
		(0.050)	(0.079)	(0.086)	(0.072)			(0.230)
AR(4)	DM-$	0.159	0.487	0.007	0.064	0.215		0.248
		(0.051)	(0.079)	(0.088)	(0.084)	(0.072)		(0.233)
AR(5)	DM-$	0.152	0.461	-0.004	0.068	0.158	0.096	0.250
		(0.053)	(0.082)	(0.088)	(0.088)	(0.084)	(0.074)	(0.231)
AR(1)	Y-$	0.299	0.381					0.150
		(0.043)	(0.074)					(0.150)
AR(2)	Y-$	0.216	0.276	0.267				0.217
		(0.048)	(0.077)	(0.077)				(0.212)
AR(3)	Y-$	0.204	0.272	0.252	0.047			0.219
		(0.051)	(0.081)	(0.080)	(0.080)			(0.209)
AR(4)	Y-$	0.192	0.264	0.250	0.032	0.052		0.221
		(0.054)	(0.081)	(0.084)	(0.083)	(0.080)		(0.206)
AR(5)	Y-$	0.188	0.262	0.248	0.030	0.046	0.023	0.222
		(0.056)	(0.082)	(0.085)	(0.087)	(0.084)	(0.081)	(0.201)

Table 3.10: Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) for the $AR(n)$ model (till the best n plus one), the GARCH(1,1) model and the Riskmetrics model, for the two time series

	DM-$ time series		Y-$ time series	
Model	RMSE	MAE	RMSE	MAE
GARCH(1,1)	0.272	0.199	0.508	0.346
RiskMetrics	0.267	0.205	0.486	0.344
AR(1)	0.256	0.179	0.497	0.344
AR(2)	0.245	0.170	0.472	0.340
AR(3)	0.241	0.168	0.467	0.339
AR(4)	0.242	0.167	0.461	0.335
AR(5)	0.234	0.160	0.457	0.332
AR(6)	0.229	0.155	0.458	0.334
AR(7)	0.230	0.156		

The reference model on the day-to-day behavior of the overnight interest rate is the martingale model, which exploits standard no arbitrage arguments. According to it, banks consider funds with different maturities as perfect substitutes; therefore no patterns should be observed in the interest rate time series within a reserve maintenance period. The argument is simple: if interest rate changes were predictable within the maintenance period, then banks would exploit this feature to minimize reserve requirement costs. The implicit cost of reserve requirement is represented by the differential between the money market rate and the reserves remuneration rate. Thus, banks would like to detain reserves to satisfy their requirements when the interest rate is low compared to other days.

The literature on the interbank market is wide, but most of it concerns the US federal funds market; the analysis of the Euro money market is still limited. Several empirical investigations on the US market have shown that the interest rate does not satisfy the restrictions imposed by no arbitrage-equilibrium conditions. In particular, the rate follows regular patterns both over the maintenance period and on an intra-week basis. Among others, Campbell (1987); Hamilton (1996); Bartolini et al. (2001, 2002) have shown that the federal funds rate rises at the end of the maintenance period (EOM) with a slight decline over the first days of each period [6]. Models explaining these results through transaction costs and credit line limits have been proposed in Campbell (1987); Hamilton (1996); Bartolini et al. (2001, 2002). On an intraweek basis, the rate tends to fall slightly on Friday and to rise on Monday, see Hamilton (1996). Interest rate volatility patterns are also observed: it is significantly higher on EOM days and at the end of the business day (Spindt and Hoffmeister, 1988; Griffiths and Winters, 1995).

Not all of the above results are confirmed all over the world[7]. In particular, interest rate patterns over the maintenance period exhibit wide differences, but the martingale hypothesis is in general rejected and higher volatility on EOM days is usually confirmed.

Turning to the Euro-area money market, Angelini (2000) analyzes the Italian screen-based overnight market (e-MID) considering hourly observations in the period $1993 - 1996$. He does not detect any significant pattern for the interest rate during the maintenance period and during the day. Quiros and Mendizabal (2000) compare the behavior of the German money market rate before the start of the EMU and that of EONIA rate afterward. They find that till 1998 EOM days were characterized by higher rates and volatility than other days. Since 1999, volatility on EOM days is lower than before and the interest rate does

[6] Bartolini et al. (2001) assess an increase of 18 basis point. According to Taylor (2000), this tendency has been reverted since 1998, with lower rates at EOM days.

[7] Prati et al. (2003) perform an empirical analysis on the overnight interest rate in G-7 countries and the Eurozone.

not increase. After the start of EMU the martingale hypothesis is still rejected, though the interest rate pattern has become closer to it; authors claim that this effect may be due to the stabilizing role played by deposit facilities. Angelini and Silipo (2001) focus on the EONIA rate, discovering the presence of weak seasonality. There is some evidence against the martingale hypothesis but it is weaker than in the US market; no effect associated with the EOM period is observed. Barucci et al. (2003) find evidence against the martingale hypothesis in the e-MID, too, and show intramaintenance patterns for the Italian overnight rate, with certain settlement (e.g. for Treasury operations) and calendar (e.g end of the year) dates being significant.

Given the evidence generally against the martingale hypothesis, it is worthwhile to investigate the relation between interest rate movements and market activity. Five measures of market activity are considered in our analysis: large-value payments value, number of contracts, average size of contracts, trading volume, order flow.

There is a large literature on the relationship between price changes-volatility and market activity, mainly analyzing the stock market. A well established regularity has shown that trading volume and volatility are positively correlated, with price changes leading trading volume, see Gallant et al. (1993); Karpoff (1987). In this perspective, some authors argue that the number of contracts, and not trading volume, is associated with volatility, see Jones et al. (1994). An analysis of exchange rate changes and signed volume (order flow) is provided in Evans and Lyons (2002). No results are available on the money market. The main reason for this lack is that in the US and in the majority of European countries the interbank market is of the OTC kind (exchanges occur through brokers) and therefore no official data on trading volume are available (Cyree and Winters, 2001).

In a stock market there is a hedging or a speculative motivation for trading. In the first case agents trade to cope with non optimal risk-portfolio allocations, while in the second case they trade to exploit private information. This is not the case in screen-based markets (like the Italian e-MID, see Barucci et al. (2003)), where by definition all the relevant information is publicly available and banks trade to meet reserve targets by minimizing reserve requirement costs. Banks trade for two main reasons, none of which is speculative in nature: execution of large value payments in the real-time-gross-settlement set up, and compliance with the compulsory reserves requirement. The market is mainly driven by aggregate liquidity conditions and by the monetary policy of the ECB, it is almost impossible for a bank to detain private information on these items. Note that all the theoretical literature on stock markets explains the relation between price changes and volume through (private and/or public) information arrival (Harris and Raviv, 1993; Kim and Verrecchia, 1991; Kandel and Pearson, 1995; Wang, 1994; He and Wang, 1995). To our knowledge this is the first contri-

bution dealing with the relationship between market price movements and trading activity in absence of a speculative component.

As far as the martingale hypothesis is concerned, similarities with the US Federal Funds market emerge in our analysis, together with some peculiarities. The overnight interest rate follows a clear pattern during the maintenance period, with a decline during the last days and a peak at the end of the calendar month. Only the second effect is statistically significant. Other calendar regularities are detected. Previous evidence (Angelini and Silipo, 2001; Angelini, 2000; Quiros and Mendizabal, 2000) is only partially confirmed; overall, results against the martingale hypothesis are much stronger.

As far as the relationship between interest rate movements and e-MID activity is concerned, results are similar to those reported in (Jones et al., 1994). It is definitely shown that volatility is affected by the number of contracts, rather than by trading volume. The number of contracts turns out to be the relevant market activity measure also when including the order flow or the large-value payments value.

3.5.1 The data set

The empirical analysis is based on daily and high frequency data. Data span the four-year period $1999 - 2002$, thus from the starting of the EMU, for a total of $1,022$ working days. Intraday data start on April 1^{st}, 1999. The data set includes overnight exchanges and average interest rates; daily data on e-MID are recorded by S.I.A. (*Società Italiana per l'Automazione*) and are splitted into exchanges generated by applying either a "bid" or an "ask" price. Furthermore, daily data include large-value payment flows, recorded by Bank of Italy. Two broad kinds of operations are considered and separately handled: BI-REL domestic debit flows and cross border debit and credit payments. The first time series includes payments stemming from customers' orders and interbank transfers of various kinds. The second one encompasses cross border inflows and outflows channeled through the European system TARGET; only the interbank component is considered, the most reactive to interest rate differentials and to shocks.

Figure 3.9 shows the time series of the overnight rate. The average sample value is 3.6248. The daily rate shows the usual spikes due to liquidity effect at the EOM period. Figure 3.10 shows the time series of daily trading volume: the average trading volume per day amounts to 11.814 billions of Euro, and the time series displays a slight increase with time. As explained later, intraday data are used to compute daily volatility. Figure 3.11 shows daily volatility in the data sample. Table 3.11 reports summary statistics on trading volume, volatility and

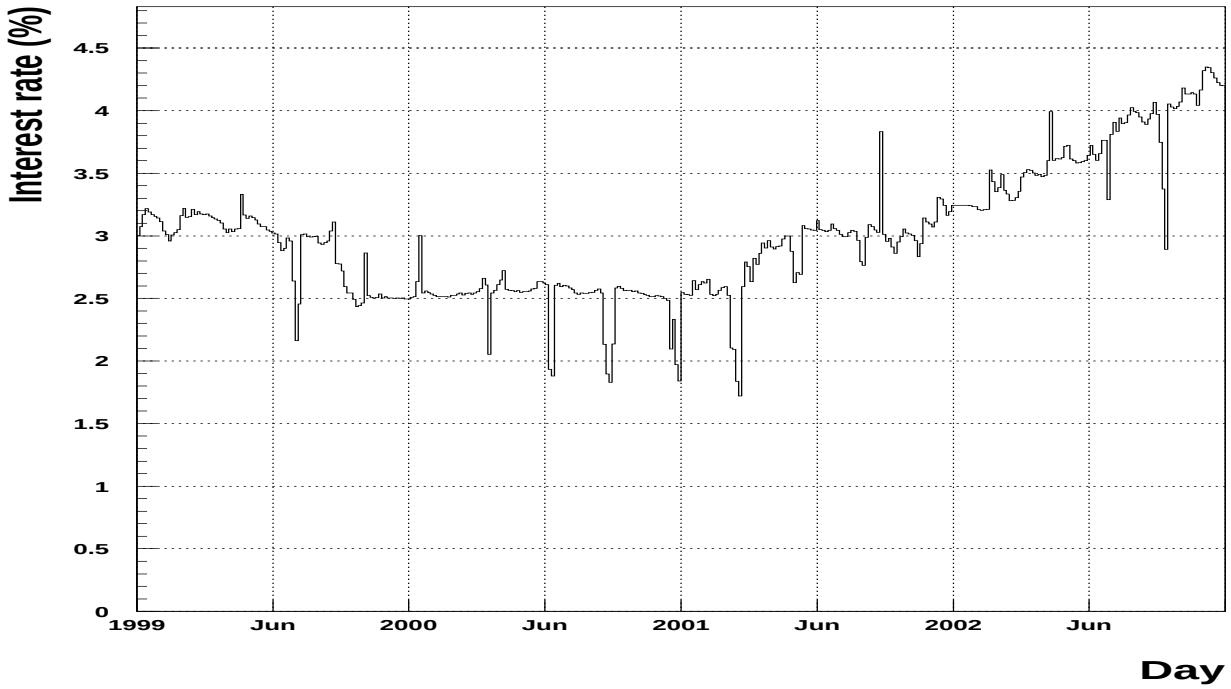

Figure 3.9: Daily mean overnight e-MID rate from January, 4^{th} 1999 to December, 30^{th} 2002.

interest rate level, divided by day of the week. Figure 3.12 shows the patterns of the relevant variables over the maintenance period. It has to be noted the deep difference among EOM and other trading days. The interest rate declines during the last day of a maintenance period (as it is also visible in Figure 3.9). Volatility and bid-ask spread suddenly increase during EOM days. Trading volume displays a slight decline on EOM days, while the number of contracts slightly increases on last maintenance days. Payment activity presents a double-peak pattern, end of the month and EOM days, respectively. For further discussion and a thorough description of the institutional setting, see Barucci et al. (2003, 2004).

3.5.2 Testing the martingale hypothesis

This section focuses on the dynamics of overnight interest rates in the Italian money market. To this aim, it is necessary to define reliable volatility measures.

Measuring volatility is an awkward task. Estimating it via daily observations (typically, by the squared interest rate difference or its absolute value) provides an unbiased but very noisy

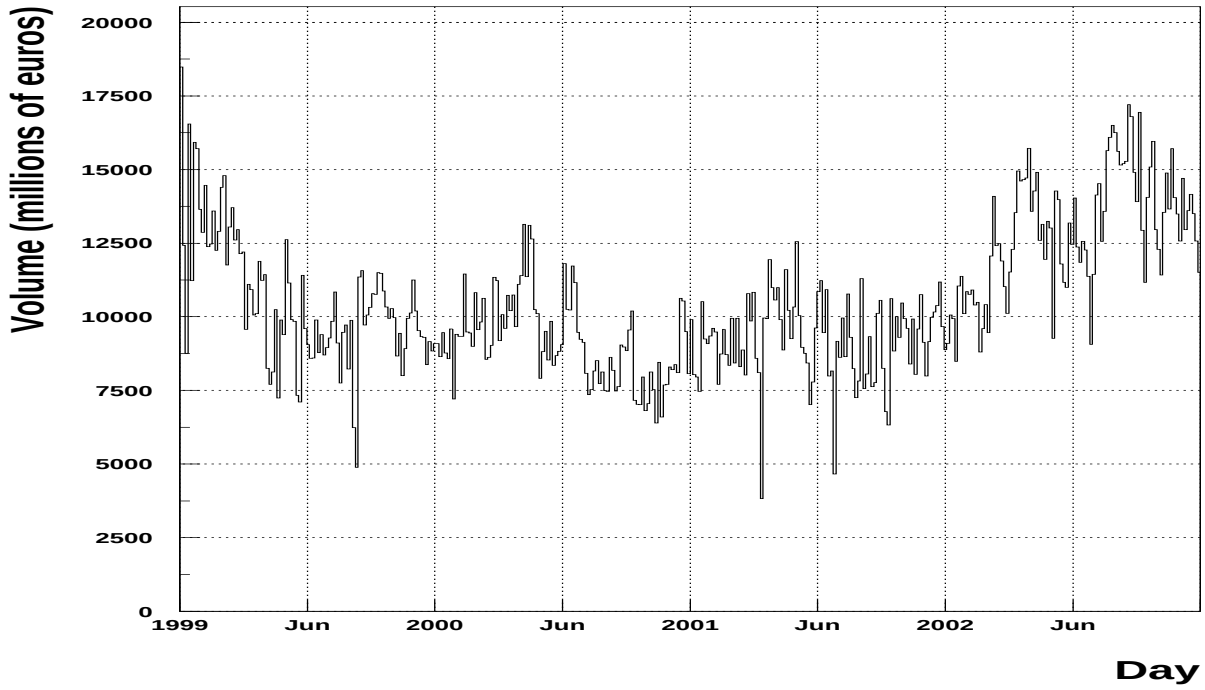

Figure 3.10: Daily overnight exchanges in e-MID from January, 4^{th} 1999 to December, 30^{th} 2002.

estimate. Intraday observations can be used to overcome these problems, as suggested in Andersen and Bollerslev (1998a), who propose to estimate the daily (integrated) volatility as the cumulative squared intraday returns. By exploiting this procedure, it is possible to estimate more precisely daily volatility, and to handle it as an observable variable instead of a latent one. Andersen et al. (2003); Barucci and Renò (2002c) follow this route to forecast volatility through a simple autoregressive model; their results show that, in spite of its simplicity, the forecasting performance of such a model is better than that of alternative, classical models, like GARCH(1,1) or Riskmetrics.

In this Section we follow this strand of literature, adopting the method described in Chapter 2 to compute daily volatility. As already discussed, the method turns out to be particularly well suited to estimate volatility through high-frequency data using all the observations with no aggregation. The dynamics of daily volatility is shown in Figure 3.11.

To test the martingale hypothesis, the following equation has been estimated

$$i_t - i_{t-1} = \sum_{k=1}^{n_i} \beta_k \left(i_{t-k} - i_{t-k-1} \right) + m_t + \eta_t \tag{3.16}$$

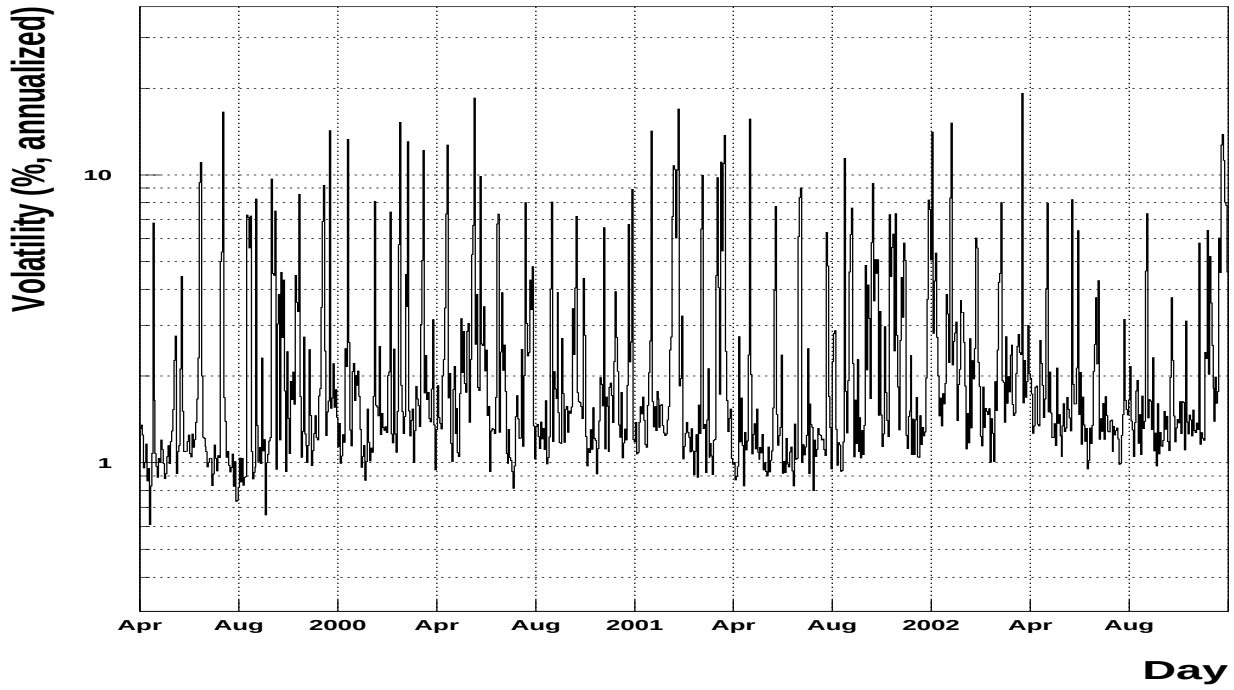

Figure 3.11: Daily overnight e-MID rate volatility from April, 1^{st} 1999 to December, 30^{th} 2002.

where i_t is the overnight rate at day t, n_i is the number of lags, and m_t is of the form:

$$m_t = \sum_{k=1}^{n_{di}} \alpha_k X_k \tag{3.17}$$

where X_k are dummy variables listed in Table 3.12 (n_{di} denotes the number of dummy variables), which are inserted to incorporate calendar effects, η_t are i.i.d. random variables with zero mean.

We first estimate equation (3.16) via OLS (Table 3.12).[8] This provides consistent estimates, which however are inefficient because of the pronounced heteroskedasticity of the overnight rate (Figure 3.11). Indeed, OLS estimation residuals are clearly heteroskedastic: both Ljung-Box and ARCH effects tests on squared residuals significantly reject the null hypothesis of zero serial correlation. This observation has induced many authors, e.g. Hamilton (1996), to model the rate via a GARCH-like model. We follow a different approach. If it were possible

[8]In all the estimations, standard errors are computed using Newey-West corrections for heteroskedasticity and residual autocorrelation. Specification tests are run in the form of Ljung-Box portmanteau tests on residuals and on squared residuals, as well as with the ARCH effect test.

Week day	Number of days	Volume	Rate	Volatility
Mon	178	12002.2	3.61056	3.11666
Tue	185	11858.0	3.62361	3.28784
Wed	184	11796.6	3.63581	3.19133
Thu	186	11540.8	3.62838	3.60023
Fri	185	11881.3	3.62557	4.16813

Table 3.11: Summary statistics for the days of the week (full sample average)

to observe the variance of η_t, efficiency could be achieved via GLS. Daily volatility, measured as described in Malliavin and Mancino (2002), can be used as a proxy of the variance. Then, WLS estimation of equation (3.16) has been performed, using as weights the inverse of the estimate of σ_t, where σ_t is the daily volatility. With WLS estimation, the Ljung-Box test on squared residuals decreases from 89.63 to 8.38 and is no longer significant, as well as the ARCH effect test, which decreases from 76.53 to 4.60. Results are shown in Table 3.12; now the values of the tests are compatible with the absence of autocorrelation in residuals and in squared residuals.

This analysis provides some interesting results. The overnight rate shows remarkable patterns over the maintenance period, see Table 3.12 and Figure 3.12. The rate is low on the first day of a maintenance period, increases significantly until the end of the month (day 8 in the figure), then behaves approximately like a martingale, with a decline on the EOM day. While the decrease during EOM days is not statistically significant, the increase during the last day of the month is significant. There are other calendar effects: the rate rises after and before 3-4 holiday days and declines significantly on the first day of the year. Note that the martingale hypothesis does not provide any insight on interest rate movements during the first day of a maintenance period. As in Angelini and Silipo (2001), there is a clear break in the auto-regressive pattern: during the last day of a period and the first day of the subsequent one the overnight rate recovers fully all the decline occurred one and two days before EOM days. Note that coefficients associated with lagged interest rate movements are statistically significant. The above pattern contrasts in part the evidence detected for the EONIA and other Eurozone rates by several authors, who did not discover any significant pattern in the interest rate over the maintenance period, see Angelini (2000); Angelini and Silipo (2001); Quiros and Mendizabal (2000); Prati et al. (2003). Overall the evidence against of the martingale hypothesis is strong.

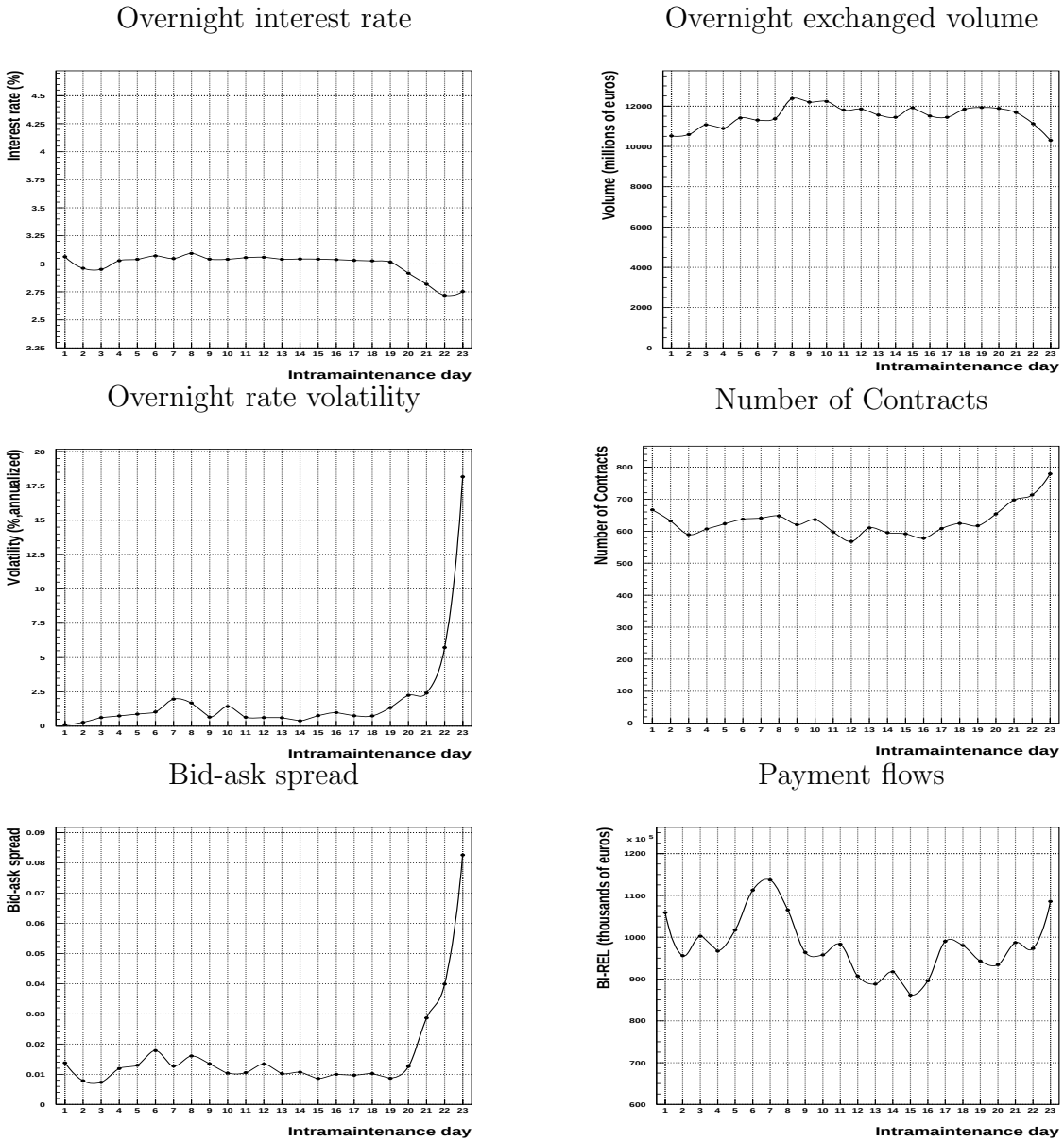

Figure 3.12: Patterns in the maintenance period; averages over the full sample

3.5.3 Interest rate volatility and market activity

Let σ_t be the integrated volatility over day t; to include calendar effects on volatility evolution the following autoregressive model is estimated:

$$\log \sigma_t = \sum_{k=1}^{n_r} \alpha_k \log \sigma_{t-k} + \mu_t + \varepsilon_t \tag{3.18}$$

Regressor	OLS		WLS	
	Coefficient	Standard Error	Coefficient	Standard Error
$i_{t-1} - i_{t-2}$	-0.10051***	0.03042	-0.20601***	0.02135
$i_{t-2} - i_{t-3}$	-0.03968	0.02930	-0.10317***	0.01930
$i_{t-3} - i_{t-4}$	-0.04269	0.02926	-0.05952***	0.01752
$i_{t-4} - i_{t-5}$	-0.01455	0.02911	-0.03087*	0.01802
$i_{t-5} - i_{t-6}$	-0.01282	0.02732	-0.03607**	0.01528
constant	-0.00363	0.00864	0.02305**	0.01011
EOM t- 0	-0.03485*	0.01788	-0.02162	0.06025
EOM t- 1	0.00567	0.01748	0.01128	0.01947
EOM t- 2	-0.01971	0.01764	-0.00423	0.01232
$(i_{t-1} - i_{t-2})$*Dummy first maint. day	-0.96620***	0.08333	-0.91629***	0.04957
$(i_{t-2} - i_{t-3})$*Dummy first maint. day	-0.68810***	0.13212	-0.59036***	0.09547
$(i_{t-3} - i_{t-4})$*Dummy first maint. day	-0.79251***	0.13889	-0.84677***	0.09124
$(i_{t-4} - i_{t-5})$*Dummy first maint. day	-0.67077***	0.15332	-0.80833***	0.09163
$(i_{t-5} - i_{t-6})$*Dummy first maint. day	-1.83120***	0.25549	-1.45441***	0.14770
End of month	0.05971***	0.01906	0.04254***	0.01282
End of quarter	0.02965	0.02259	0.02685*	0.01486
End of year	0.55069***	0.08612	0.53969*	0.30789
First day month	0.01215	0.01807	0.00641	0.00900
First day of the year	-0.66312***	0.08576	-0.57763***	0.04353
Before 3-4 holiday	0.07452**	0.02969	-0.00045***	0.00014
After 3-4 Holiday	-0.01327	0.02971	0.00000	0.00000
Tuesday	-0.00029	0.01190	0.02987*	0.01617
Wednesday	0.00755	0.01184	0.01754	0.01784
Thursday	-0.01220	0.01189	0.00201	0.00609
Friday	0.00048	0.01203	-0.00009	0.00606

Table 3.12: Estimates of the model (3.16) (one star, 90%, two stars, 95 %, three stars 99% significance). OLS estimation: $R^2 = 41.74\%$, Ljung-Box on residuals: $L(10) = 18.80$, Ljung-Box on squared residuals: $L(10) = 89.63$, ARCH effect test: 76.53. WLS estimation: $R^2 = 70.42\%$. Ljung-Box on residuals: $L(10) = 28.61$, Ljung-Box on squared residuals: $L(10) = 8.38$, ARCH effect test: 4.60

where n_r is the number of lags in the auto-regressive setting,

$$\mu_t = \sum_{k=1}^{n_d} \beta_k X_k \tag{3.19}$$

and $X_k,\ k = 1, \ldots, n_d$, is a set of n_d variables reported in Table 3.13, ε_t are random variables with zero mean.

Regressor	Coefficient	Standard Error
$\log(\sigma_{t-1})$	0.47290***	0.03172
$\log(\sigma_{t-2})$	0.08005***	0.02278
constant	-1.15417***	0.07407
EOM t- 0	1.64744***	0.07444
EOM t- 1	0.84872***	0.06741
EOM t- 2	0.39937***	0.06697
EOM t- 3	0.31659***	0.06646
EOM t- 4	0.18463***	0.06794
First maintenance day	-0.86765***	0.09173
Bce (THURSDAY)	0.53430***	0.12920
Bce (FRIDAY)	-0.06664	0.13064
End of quarter	0.43128***	0.09730
First day of a new quarter	0.10983	0.10496
End of month	0.28605***	0.07139
First day month	-0.24701***	0.07429
End of year	0.11726	0.32859
First day of the year	-1.06540***	0.32479
Before 3-4 holiday	0.06818	0.11020
After 3-4 holiday	0.02383	0.11088
Tuesday	0.12143***	0.04446
Wednesday	0.09768**	0.04421
Thursday	0.12890***	0.04491
Friday	0.10430**	0.04518

Table 3.13: Overnight rate volatility fit, equation (3.18) (one star, 90%, two stars, 95 %, three stars 99% significance); $R^2 = 67.37\%$; Ljung-Box on residuals: $L(10) = 8.22$; Ljung-Box on squared residuals: $L(10) = 27.78$.

An advantage of handling volatility as an observable variable is that equation (3.18) can be simply estimated by OLS, thus avoiding all numerical problems of maximum likelihood estimation associated with standard volatility models. When ε_t are i.i.d., OLS also provides efficient estimates. Regression results are shown in Table 3.13, while daily volatility estimates for the full sample are reported in Figure 3.11. Ljung-Box tests on residuals and squared residuals show that the autocorrelation has been removed.

Estimation results can be summarized as follows. Volatility changes remarkably during the month, with a high degree of persistence; the autoregressive component is highly significant. The main regularities on volatility are shown in Figure 3.12: the last four days of a maintenance period are characterized by a sharp increase; volatility is also high on the last day of a month and of a year, while it is substantially low on the first day of the year and of the

$\log(\sigma_{t-1})$	0.458***	0.469***	0.475***	0.456***	0.457***	0.462***	0.473***	0.452***	0.449***
$\log(\sigma_{t-2})$	0.075***	0.080***	0.081***	0.075***	0.075***	0.077***	0.080***	0.074***	0.070***
constant	-1.599***	-1.234***	-1.014***	-1.630***	-1.621***	-1.372***	-1.142***	-1.714***	-1.776***
EOM t- 0	1.538***	1.654***	1.596***	1.545***	1.539***	1.630***	1.640***	1.536***	1.487***
EOM t- 1	0.790***	0.843***	0.815***	0.793***	0.790***	0.836***	0.838***	0.790***	0.768***
EOM t- 2	0.352***	0.390***	0.377***	0.352***	0.352***	0.386***	0.389***	0.351***	0.331***
EOM t- 3	0.282***	0.306***	0.304***	0.281***	0.282***	0.308***	0.309***	0.282***	0.268***
EOM t- 4	0.166**	0.174**	0.178***	0.164**	0.165**	0.177***	0.177***	0.166**	0.158**
First maintenance day	-0.846***	-0.865***	-0.885***	-0.842***	-0.845***	-0.873***	-0.876***	-0.845***	-0.832***
Bce (THURSDAY)	0.459***	0.499***	0.504***	0.456***	0.458***	0.493***	0.512***	0.452***	0.460***
First day of a new quarter	0.316***	0.365***	0.351***	0.317***	0.316***	0.342***	0.368***	0.305***	0.314***
End of month	0.370***	0.387***	0.374***	0.372***	0.371***	0.321***	0.382***	0.334***	0.357***
First day month	-0.314***	-0.330***	-0.318***	-0.315***	-0.315***	-0.347***	-0.329***	-0.326***	-0.312***
First day of the year	-1.190***	-1.220***	-1.225***	-1.189***	-1.189***	-1.176***	-1.221***	-1.163***	-1.157***
Tuesday	0.128***	0.127***	0.127***	0.129***	0.128***	0.137***	0.127***	0.135***	0.132***
Wednesday	0.112**	0.106**	0.109**	0.112**	0.112**	0.100**	0.105**	0.108**	0.114***
Thursday	0.154***	0.146***	0.146***	0.155***	0.154***	0.150***	0.143***	0.158***	0.157***
Friday	0.096**	0.108**	0.108**	0.096**	0.095**	0.112**	0.109**	0.098**	0.089**
Number of Contracts	0.745***			0.735***	0.763***			0.716***	1.089***
	(0.153)			(0.155)	(0.182)			(0.155)	(0.177)
Volume		0.006		0.003					
		(0.006)		(0.006)					
Average size			-5.747**		0.478				
			(2.323)		(2.741)				
Payment volume						2.021*		1.195	
						(1.061)		(1.064)	
Order flow							-0.020		-0.115***
							(0.027)		(0.030)

Table 3.14: Overnight rate volatility fit, equation (3.18) with the inclusion of activity variables (one star, 90%, two stars, 95 %, three stars 99% significance); R^2 values are between 66.81% (min) and 67.53% (max). Ljung-Box on residuals at lag 10 is between 6.52 and 11.16. On squared residuals, same lag, is between 19.35 and 24.16.

maintenance period.

The data-set allows us to consider several variables as proxy of market activity on a daily basis:

- trading volume
- number of contracts
- average size of contracts
- large-value domestic payments volume
- order flow, in absolute value (defined as $|q_{ask} - q_{bid}|$, where q_{ask}, q_{bid} are the quantities exchanged at the ask and the bid price, respectively)

Jones et al. (1994) tested the influence of the number of contracts and trading volume on stock prices volatility, concluding that it is the number of contracts which mainly influences it. On foreign exchange markets, Evans and Lyons (2002) show a close correspondence between

exchange rate movements and order flow. Finally, on the U.S. money market Furfine (2000) provides evidence of a significant relation between interest rate volatility and both large value payments value and volatility.

Equation (3.18) has then been estimated again adding the selected indicators of market activity as regressors. In this new set of regressions, we only include dummy variables which resulted significant in the previous analysis. When including contemporaneous variables, OLS is not any more efficient, but t-statistics are still valid and can be used for inference.

Results of a first set of regressions, with market activity indicators included one by one, are displayed in Table 3.14. Number of contracts and payments value positively and significantly affect interest rate volatility. No relationship between interest rate and trading volume or order flow is detected. To disentangle the effect of the number of contracts and volume, Jones et al. (1994) consider both the number of contracts and the average size of contracts during a day as exogenous variables. Confirming the results obtained above, when volatility is regressed on number of contracts and average size we find that the number of contracts turns out to be significant and the average size does not. Moreover note that the average size alone negatively affects the interest rate volatility. The result in Furfine (2000) about a positive significant relation between payments value and interest rate volatility is confirmed. However, this relations fades when payments are regressed with the number of contracts; the correlation between the two quantities is 0.25.

The number of contracts still retains its explicative power even if it is regressed jointly with trading volume and average size of contracts. The result for the average size of contracts, which is largely significant and negative alone and not significant when it is regressed jointly with the number of contracts, can be explained by the strong anti-collinearity; indeed, the linear correlation coefficient between the number and the average size of contracts is -0.58.

This apparently striking result can be properly understood considering the difference between speculative and pure hedging markets. Stock and exchange markets have both features, while the interbank market is a pure hedging one. As a matter of fact, since it is almost impossible for banks to detain private information on either future system liquidity conditions or interest rates policy decisions, we can assess that the interbank market is almost free of information asymmetry-heterogeneity. Speculative markets are characterized by agents endowed with private information. As discussed in the Introduction, the literature has shown that information arrival (either private or public) creates a link between trading volume-number of contracts and volatility: agents trade both for speculative and hedging reasons and therefore trading volume and price changes are linked. On the other hand, money market exchanges satisfy the need of funds replenishment or storage in view of two kinds of needs,

none of which is speculative in nature: the execution of large value payments in the real-time-gross-settlement set up, and the compliance with the compulsory reserves requirements. These needs crucially depend on technical dates (e.g., the EOM or the payments to the State Treasury). Banks who are affected by major liquidity shocks (either positive or negative) and/or make mistakes in their liquidity forecasts within the maintenance period come to negotiate (positive or negative) excess funds in proximity of those dates, with much lower demand elasticity. Often, in particular during EOM days, excess funds and/or forecast errors are small and so we have many contracts for a small amount of funds. This feature makes compatible low trading volume, high number of contracts and high volatility. Also the positive coefficient associated with payments value confirms that liquidity management drives interest rate volatility.

Thus, the number of contracts play a prior role as an indicator of market activity. In order to check if its effect is different during EOM days, when the interest rate dynamics is substantially different from that of other days, the following experiment has been run. Consider the market activity indicator N_t (say the number of contracts), interest rate volatility is regressed on the two orthogonal variables $N_t(1-Y_t)$ and N_tY_t, where Y_t is a dummy variable whose value is 1 in the last two maintenance days and 0 otherwise. The results are reported in Table 3.15. They confirm those obtained on the full data set; moreover all the variables have a stronger influence on EOM days, as the interpretation provided above would suggest. Note that payments value is not significant in non EOM days. Number and average size of contracts were regressed jointly, but separately for EOM and non-EOM days. Again, when considered jointly, the number of contracts is significant over all the period, while the average size is never significant.

To further investigate the relationship between market activity and price volatility, we estimate equation (3.18) measuring daily volatility as it is done in Jones et al. (1994), i.e., as the absolute value of the residuals of equation (3.16).

Results of the regression with market activity variables are reproduced in Table 3.16. They are not qualitatively different from those obtained measuring volatility with high frequency data: the number of contracts still play the main role. The only notable difference is that payment volume retains its significance when regressed with the number of contracts.

$\log(\sigma_{t-1})$	0.461***	0.472***	0.469***	0.454***	0.469***	0.454***
$\log(\sigma_{t-2})$	0.078***	0.079***	0.081***	0.077***	0.079***	0.077***
constant	-1.341***	-1.047***	-1.225***	-1.524***	-1.148***	-1.521***
EOM t- 0	1.257***	1.981***	1.558***	0.472	1.591***	0.539
EOM t- 1	0.489	1.293***	0.740***	-0.145	0.781***	-0.074
EOM t- 2	0.389***	0.380***	0.390***	0.362***	0.390***	0.362***
EOM t- 3	0.308***	0.305***	0.307***	0.289***	0.309***	0.289***
EOM t- 4	0.177***	0.177***	0.174**	0.168**	0.176***	0.168**
First maintenance day	-0.866***	-0.876***	-0.867***	-0.836***	-0.868***	-0.836***
Bce (THURSDAY)	0.495***	0.507***	0.499***	0.469***	0.517***	0.469***
First day of a new quarter	0.348***	0.355***	0.365***	0.329***	0.373***	0.329***
End of month	0.333***	0.376***	0.387***	0.376***	0.381***	0.375***
First day month	-0.341***	-0.318***	-0.330***	-0.314***	-0.328***	-0.314***
First day of the year	-1.182***	-1.219***	-1.221***	-1.189***	-1.213***	-1.189***
Tuesday	0.138***	0.127***	0.127***	0.128***	0.125***	0.128***
Wednesday	0.101**	0.108**	0.106**	0.111**	0.103**	0.111**
Thursday	0.151***	0.145***	0.146***	0.154***	0.141***	0.154***
Friday	0.114**	0.109**	0.108**	0.097**	0.111**	0.098**
Payment volume	1.663					
	(1.111)					
Average size		-4.789**				-0.070
		(2.365)				(2.792)
Volume			0.006			
			(0.006)			
Number of Contracts				0.594***		0.592***
				(0.160)		(0.191)
Order flow					-0.036	
					(0.028)	
Payment volume EOM	5.257*					
	(3.175)					
Average size EOM		-29.565**				-2.269
		(11.922)				(13.983)
Volume EOM			0.015			
			(0.023)			
Number of Contracts EOM				2.138***		2.089***
				(0.478)		(0.567)
Order flow EOM					0.142	
					(0.089)	

Table 3.15: Overnight rate volatility fit, equation (3.18) with the inclusion of activity variables divided in last two maintenance days and others (one star, 90%, two stars, 95 %, three stars 99% significance); R^2 values are between 66.81% (min) and 67.53% (max). Ljung-Box on residuals at lag 10 is between 6.52 and 11.16. On squared residuals, same lag, it is between 19.35 and 24.16.

3.6 Conclusions

Recently, a large literature has been devoted to compute-forecast volatility for financial time series. In this field, the importance of high frequency data has been stressed, in particular to evaluate the forecasting performance of GARCH models.

In this Chapter we use a new method to compute volatility, that is the method described in Chapter 2. The main feature of this method is that it is based upon integration instead of differentiation of the time series, so that it naturally exploits the time structure of high frequency data by including all the observations in the volatility computation. Using simulated time series, we illustrated that, when realized volatility is computed using linear interpolation, the Fourier method performs better than realized volatility in measuring integrated

sigma t- 1	0.303***	0.311***	0.306***	0.303***	0.303***	0.294***	0.306***	0.290***	0.302***
sigma t- 2	0.031	0.027	0.033	0.031	0.030	0.021	0.023	0.024	0.029
constant	-0.042**	0.003	0.032**	-0.044**	-0.043	-0.077***	0.010	-0.108***	-0.037*
EOM t- 0	0.104***	0.122***	0.115***	0.105***	0.104***	0.110***	0.121***	0.099***	0.106***
EOM t- 1	0.041***	0.049***	0.045***	0.041***	0.041***	0.045***	0.048***	0.039***	0.041***
EOM t- 2	0.040***	0.046***	0.045***	0.040***	0.040***	0.043***	0.046***	0.039***	0.041***
EOM t- 3	0.024*	0.028**	0.028**	0.024*	0.024*	0.027**	0.028**	0.024*	0.025*
EOM t- 4	0.007	0.008	0.009	0.007	0.007	0.009	0.009	0.008	0.007
First maintenance day	0.031**	0.031**	0.031**	0.031**	0.031**	0.022	0.032**	0.022	0.031**
Bce (THURSDAY)	-0.026	-0.021	-0.019	-0.026	-0.026	-0.027	-0.021	-0.031	-0.026
First day of a new quarter	0.065***	0.072***	0.071***	0.065***	0.065***	0.060***	0.070***	0.055***	0.065***
End of month	0.010	0.013	0.011	0.010	0.010	-0.019	0.013	-0.018	0.010
First day month	0.052***	0.050***	0.053***	0.052***	0.052***	0.039***	0.051***	0.042***	0.052***
First day of the year	-0.094	-0.096	-0.096	-0.094	-0.094	-0.080	-0.096	-0.080	-0.095
Tuesday	0.011	0.011	0.011	0.011	0.011	0.017*	0.011	0.017*	0.011
Wednesday	0.005	0.004	0.004	0.005	0.005	0.002	0.004	0.003	0.005
Thursday	0.015	0.013	0.013	0.015	0.015	0.017*	0.013	0.017*	0.015
Friday	0.001	0.002	0.002	0.001	0.001	0.004	0.003	0.003	0.001
Number of Contracts	0.102***			0.101***	0.103***			0.076**	0.092***
	(0.030)			(0.031)	(0.036)			(0.031)	(0.035)
Volume		0.001							
		(0.001)		(0.001)					
Volume average			-0.807*		0.030				
			(0.470)		(0.551)				
Payment volume						0.966***		0.860***	
						(0.204)		(0.208)	
Order flow							0.012**		0.004
							(0.005)		(0.006)

Table 3.16: Overnight rate volatility fit, equation (3.18), with volatility measured as absolute value of residuals as in Jones et al. (1994), with the inclusion of activity variables (one star, 90%, two stars, 95 %, three stars 99% significance)

volatility and that, by adopting it, the forecasting performance of the GARCH model can be better evaluated.

We showed that linear interpolation of the time series induces a downward bias in the volatility estimate, and this effect is avoided by assuming the price process to be piecewise constant. On the other hand, relying upon Monte Carlo simulations of models belonging to the SR-SARV(1) class, which includes the GARCH(1,1), we have shown that the Fourier estimator is unbiased and its variance is smaller than that of the cumulative squared intraday returns, when the latter is chosen with a reasonable bias in mean. Moreover, when the Fourier method is employed to evaluate the forecasting performance of the models, their performance is better than that obtained by employing the cumulative intraday squared returns.

We applied this method to two exchange rate time series. On real data one has to deal with microstructure effects, which become dominant when the return sampling frequency becomes comparable to the frequency of tick-by-tick quotes. We gave a precise estimate of the time step above which these effects can be neglected, and we showed how to remove microstructure distortions: since the Fourier estimator is given by an expansion of the Fourier coefficients, it is enough to cut the highest frequencies in a suitable way. When employing the Fourier

method, GARCH forecasts turn out to be more accurate than those associated with the sum of squared intraday returns.

We used the Fourier method to evaluate the forecasting performance of the GARCH(1,1) model when it is discretized at intraday frequencies; the results obtained in the recent literature are confirmed, moreover the forecasting properties of the GARCH model are evaluated to be better if the Fourier estimator is employed, instead of the cumulative squared intraday returns, to measure integrated volatility.

The importance of volatility measuring in risk management is increasing. Providing a good forecasting model for daily integrated volatility is essential in calculating reliable VaR estimates. Recently, it has been shown that the use of high frequency data allows to measure daily integrated volatility with high precision. This suggests that, instead of using models in which the volatility is a latent factor, like for example GARCH(1,1), one can try to model directly the dynamics of integrated volatility. In this Chapter we build a simple forecasting $AR(n)$ model for integrated volatility, estimated by OLS, and we show that it performs considerably better than traditional models. We conclude that modeling directly integrated volatility, measured by high frequency data, can be a promising direction for risk management.

Finally we evaluate the relationship between volatility and market activity in a market with liquidity-hedging trading, that is the Italian money market. As a matter of fact, banks trade in the interbank money market only for liquidity reasons because it is difficult for a bank to detain private information on variables affecting future interest rate (liquidity conditions, ECB decisions). We have shown that the number of contracts and payment volume are associated with interest rate volatility, while trading volume is not. Our result can be explained by the fact that banks trade to cope with liquidity needs, that may be small in particular when the end of the maintenance period approaches. The fact that liquidity management lies behind the relation between market activity and volatility is confirmed by the fact that the relation is stronger during end of maintenance days.

Chapter 4

Multivariate applications

4.1 Introduction

Volatility computation is one of the most challenging problems in financial mathematics. Indeed, unlike other quantities of interest in financial markets (price, volume), volatility cannot be directly observed and it has to be inferred from price evolution. On the other hand, volatility plays a crucial role in a variety of applications.

The literature on the estimate of volatility is mainly based on the quadratic variation formula. Nevertheless, the use of this formula has some disadvantages in the implementation because, since it is based on a differentiation procedure so it needs equally and regularly spaced data. This fact becomes more evident when considering cross volatilities, since the construction of an evenly sampled time series automatically biases the correlation estimates toward zero. Typically, results on volatility modeling are still substantially low-dimensional, if not univariate (Andersen et al., 2003; Bollerslev et al., 1994; Ghysels et al., 1996).

In this Chapter, we start from the Fourier series method of Malliavin and Mancino (2002) developing the analysis in the multivariate case. In Barucci et al. (2003) the Fourier method for computing cross-volatilities has been applied to determine a time dependent *elasticity matrix*, which is an indicator of market stability. In this Chapter we first study the performance of the method in the bivariate case using Monte Carlo simulations of high frequency asset prices, extending the framework of Barucci and Renò (2002a) to a bi-variate case.

In the second part, we investigate the so-called Epps effect. Epps (1979) reported empirical evidence that stock correlations decrease when sampling frequency increases. This phenomenon has been observed in several markets. In our analysis, the dynamics underlying

the Epps effect are investigated. Using Monte Carlo simulations and the analysis of high frequency foreign exchange rate and stock price data, it is shown that the Epps effect can largely be explained by two factors: the non-synchronicity of price observations and the existing lead-lag relationship between asset prices. In order to compute co-volatilities, the method based upon the Fourier analysis is adopted, since it performs well in estimating correlations precisely, as illustrated by simulated experiments. Being naturally embedded in the frequency domain, this estimator is well suited to the study of the Epps effect.

In the third part of this chapter we proceed to a dynamic principal component analysis. Making mild assumptions on the multivariate price behavior (bounded quadratic variation), we get by the Fourier method multivariate volatilities as a time dependent quadratic form. We suggest a method to construct an abstract curve which contains the essential information coming from multivariate volatilities. Analyzing the time evolution of this curve can be important to decipher the interaction of an asset with the whole market. We illustrate these ideas via the analysis of two months of high-frequency data for 98 stocks; our results point out the need of a time-dependent analysis versus a static one, since the variance-covariance eigenvalues structure turns out to be deeply time-varying. In particular, we find some days in which correlations are widely distributed, thus less factors are needed to explain the variance-covariance structure. We then introduce the concept of reference assets, an asset whose volatility is mainly due to the market volatility instead of its idiosyncratic noise.

This Chapter is structured in the following way: Section 4.2 illustrates, via Monte Carlo experiments, how the Fourier estimator performs better than classical estimators. Section 4.3 investigates the Epps effect both on simulated and real data. Section 4.4 depicts a geometrical interpretation of the multivariate analysis and outlines some empirical results; Section 4.5 concludes.

4.2 Performance on simulated data

As in the previous Chapter, we start with Monte Carlo simulation of high frequency asset prices, extending the framework of Barucci and Renò (2002a) to a bi-variate analysis.

We simulate two correlated asset price diffusions with the bi-variate continuous GARCH(1,1)

model:

$$\begin{aligned} dp_1(t) &= \sigma_1(t)dW_1(t), \\ dp_2(t) &= \sigma_2(t)dW_2(t), \\ d\sigma_1^2(t) &= \lambda_1[\omega_1 - \sigma_1^2(t)]dt + \sqrt{2\lambda_1\theta_1}\sigma_1^2(t)dW_3(t), \\ d\sigma_2^2(t) &= \lambda_2[\omega_2 - \sigma_2^2(t)]dt + \sqrt{2\lambda_2\theta_2}\sigma_2^2(t)dW_4(t), \\ corr(dW_1, dW_2) &= \rho, \end{aligned} \tag{4.1}$$

and all other correlations between the Brownian motions W_1, W_2, W_3, W_4 set to zero. The choice of this particular model comes from the fact that it is the continuous time limit of the very popular GARCH(1,1) model, and it has been studied extensively in the literature, e.g. see Kroner and Ng (1998). We will use the parameter values estimated by Andersen and Bollerslev (1998a) on foreign exchange rates, i.e. $\theta_1 = 0.035, \omega_1 = 0.636, \lambda_1 = 0.296, \theta_2 = 0.054, \omega_2 = 0.476, \lambda_2 = 0.480$. We will instead analyze two mirror cases for the correlation coefficient: $\rho = 0.35$ and $\rho = -0.35$.

To get a representation of high-frequency tick-by-tick data, after discretizing (4.1) by a first-order Euler discretization scheme with a time step of one second, we extract observation times drawing the durations from an exponential distribution with mean 60 seconds. Observation times are the same for the two time series, thus we avoid the problems linked to the Epps effect, which appear when increasing frequency due to non-synchronous quotes and which will be discussed later.

After simulating the process (4.1), we compute daily (86400 seconds, corresponding to 24 hours of trading, as for currencies) variance-covariance matrix according to the Fourier method of Chapter 2, and according to the realized volatility measure of Andersen et al. (2000a, 2003) given by

$$\sigma_{ij}^2 = \sum_{j=1}^{m-1} \left[p_i\left(t + \frac{k+1}{m}\right) - p_i\left(t + \frac{k}{m}\right) \right] \left[p_j\left(t + \frac{k+1}{m}\right) - p_j\left(t + \frac{k}{m}\right) \right]. \tag{4.2}$$

Barndorff-Nielsen and Shephard (2004a) develop a full asymptotic theory for this estimator. The choice of m in (4.2) comes from a tradeoff between increasing precision and cutting out microstructure distortions. A typical value is $m = 288$, corresponding to five minute returns. Since the time series $p(t)$ is not observed in continuous time, one has to resort to interpolation techniques to obtain the values $p\left(t + \frac{k}{m}\right)$ in (4.2) at equally spaced times, when these values are not observed directly. Two alternatives have been followed in the literature: linear interpolation between adjacent observations, and previous-tick interpolation, i.e. the price at time t is equal to the price of the last observation.

We implement both these interpolation schemes, for $m = 288, 96, 48$ (corresponding respectively to 5,15,30 minute returns), when measuring correlations on Monte Carlo experiments.

Table 4.1: Average correlation on 10,000 Monte Carlo replications of the model (4.1). Two generated values of the correlation are considered, $\rho = 0.35$ and $\rho = -0.35$. We compute the variance-covariance matrix via the Fourier estimator and via the realized volatility estimator (4.2). L.I. means Linear Interpolation, while P.T. means Previous Tick interpolation. Standard deviations of in-sample measurements are reported in the columns named Std. The standard error on the mean is the standard deviation divided by 100.

Estimator	Generated correlation $\rho = 0.35$		Generated correlation $\rho = -0.35$	
	Measured	Std	Measured	Std
Fourier	0.350	0.039	-0.349	0.039
Realized 5′, L.I.	0.204	0.058	-0.203	0.055
Realized 5′, P.T.	0.181	0.060	-0.180	0.058
Realized 15′, L.I.	0.338	0.090	-0.337	0.090
Realized 15′, P.T.	0.329	0.091	-0.328	0.092
Realized 30′, L.I.	0.345	0.127	-0.344	0.126
Realized 30′, P.T.	0.342	0.127	-0.341	0.126

Table 4.1 shows the results. First of all we notice that the Fourier estimator performs considerably better than realized volatility, which is biased toward zero. The bias in the correlation measurement of realized volatility is more and more severe as the sampling frequency increases. For the five-minute estimator with the previous-tick interpolation, we get a mean value of 0.181 (-0.180), which is quite far from the true value of 0.35 (-0.35); this bias is completely due to the use of interpolated prices on the evenly spaced grid. Realized volatility with linearly interpolated returns is closer to the right value, but this is because of the downward bias in the volatility measurement due to the linear interpolation documented in Barucci and Renò (2002a,b). In these papers, the authors show that the spurious positive serial correlation induced by the linear interpolation technique lowers the volatility estimates. Since variances are spuriously measured to be lower, correlations turn out to be spuriously higher, thus compensating in some way the bias due to non-synchronicity. This is also true, but to a much lower extent, for the 15-minute and 30-minute realized volatility estimator. The precision of the Fourier estimator, as measured by the standard deviation of measurements across Monte Carlo replications, is always better than that obtained with the realized volatility estimator. We implemented the Fourier estimator with $N = 500$ coefficients for the first time series, $N = 160$ coefficients for the second and $N = 160$ coefficients for the computation of covariance. Increasing the number of coefficients would increase the measurement precision. On the other hand, even the gain in precision of the realized covari-

ance measurement obtained when increasing the sampling frequently is canceled out by the bias.

Thus, even if the Fourier estimator is comparable to realized volatility when using the previous-tick interpolation scheme, it is remarkably better in computing covariances.

4.3 Investigating the Epps effect

In his 1979 paper, Epps reported empirical evidence of a dramatic drop in correlations between stocks when decreasing the sampling time horizon. This phenomenon has been observed across different markets, see for example Bonanno et al. (2001); Zebedee (2001) for stock prices, Lundin et al. (1999); Muthuswamy et al. (2001) for foreign exchange rates. On the other hand Andersen et al. (2001) and Andersen et al. (2001), with regard to stock prices and foreign exchange rates respectively, report correlations significantly different from zero when computed using five minute returns.

In this Section, as the title suggests, there is an attempt to understand the dynamics underlying the Epps effect better. For this purpose, we adopt the Fourier variance-covariance estimator set out in Chapter 2. This estimator has two appealing features: first, it uses all the tick-by-tick observations with no need to change their time structure. Second, all the other estimators developed so far (Andersen and Bollerslev, 1998a; de Jong and Nijman, 1997; Lundin et al., 1999; Ball and Torous, 2000) are constructed in the time domain, while the Fourier estimator is naturally embedded in the frequency domain. The Epps effect deals with the behavior of correlations as a function of the sampling frequency, so the Fourier estimator is well suited for its inspection.

The main idea pursued in this Section is that correlation measurements are always biased toward zero when observations are not perfectly synchronous. Two main statistical features of the data may produce this effect: asynchronous trading and lead-lag relationships. The impact of asynchronous data on covariance measurement has been widely studied: two examples are Lo and MacKinlay (1990) and, previously, Scholes and Williams (1977). In addition to this, it is sometimes suggested that the Epps effect may depend upon the fact that correlations are lagged[1], so that when reducing the sampling frequency under time scales comparable to this lag, the correlation measurements turn out to be lower.

The motivation of this work is to assess the following questions:

[1]It is worth to remark that non-synchronous trading itself could be a source of spurious lead-lag relations, see Chan (1992, 1993).

- what is the relative magnitude of the impact of asynchronous trading and lead-lag relationships?
- are these two features sufficient to explain the Epps effect?

In the literature, we have found mixed answers to these two questions. For example, Zebedee (2001) argues that the Epps effect is mainly due to the lead-lag relationship, maintaining that as frequency increases, correlation is shifting to other nearby time intervals. On the other hand, Lundin et al. (1999) claim that different actors play different roles at different frequencies, so that it is not possible to recover the same correlation at different time scales. However, it is not clear which kind of price formation process could lead to the Epps effect. Moreover, Lundin et al. (1999) find a significant inverse relation between correlation and activity: the more an asset is traded, less evident is the Epps effect. Again this seems to enforce the importance of synchronicity in explaining the correlation decrease at higher frequencies.

In our analysis, we will primarily make use of Monte Carlo simulation, showing that, if price observations of two traded assets are synchronous, and if there is no lead-lag relationship, no frequency effects should be observed in the correlation measurements. Non-synchronicity and lags are then introduced, and the results show that both have a substantial effect (although it is mostly due to synchronicity), and thus are good candidates to explain the Epps effect. We then turn to the analysis of foreign exchange rates and stock prices, showing that, when applying adjustment techniques based only on these two factors, a significant attenuation of the drop in correlation measurement at high frequencies is observed.

4.3.1 Monte Carlo experiments

The purpose of this study is to analyze the behavior of correlations between high frequency asset prices, as a function of the sampling frequency. We first start with studying Monte Carlo experiments. The simulation of two correlated asset price diffusions is accomplished via the bivariate continuous-time GARCH(1,1) model (4.1) with the parameters as in Andersen and Bollerslev (1998a). The value of the correlation is set to $\rho = 0.35$.

If we could observe the process (4.1) continuously, we would not have problems in computing the variance-covariance matrix and recover the exact value of the correlation. Here we want to study the impact of two main features of high-frequency data:

- intraday asset prices are recorded in form of tick-by-tick transactions or quotes, which

are unevenly spaced and whose frequency depends on the liquidity of the asset;

- correlation may be lagged, due to different liquidity, economic significance or recording effects.

Using the Monte Carlo simulation, we should be able to disentangle the impact of these two effects on correlation measurements. We will proceed as follows: first, the bivariate process (4.1) is discretized using a first-order Euler discretization scheme with a time step of one second. This mimics continuous time synchronous record. Then, two kind of samples are simulated: in the first one, tick-by-tick durations are extracted from an exponential distribution with mean equal to 15 seconds for the first time series and to 45 seconds for the second time series. The choice of the average duration again comes from the average durations of DEM-USD and JPY-USD respectively. In order to illustrate the effect of non-synchronicity of high frequency data, we use a second kind of sample in which the durations of both the time series are extracted from an exponential distribution with mean equal to 45 seconds, thus forcing the first time series to be observed exactly at the same times as the second time series. These two simulated time series are labeled asynchronous and synchronous respectively.

In order to introduce lagging, we use the sample described above, but shift all the observation times of the second time series backward by 8 seconds, a choice motivated by subsequent data analysis. These two new samples are labeled synchronous-lagged and asynchronous-lagged respectively.

In these Monte Carlo experiments, daily (86400 seconds) covariance matrices are computed according to (2.44), as a function of frequency M. Figure 4.1 shows the resulting average daily realized correlation as a function of the sampling frequency M used in the computation, for the four different samples. Let us look at the asynchronous-lagged sample first, which is thought to be closer to actual data. Even if we do not go in the deep high frequency regime, for further comparison with FX data, the Epps effect is clearly displayed. Correlation begins to drop above a certain frequency, going far from the "true" generated value. If we increase the sampling frequency, then correlation goes to zero (not displayed). If we look at asynchronous (not lagged) data, we see that the Epps effect is still present, but slightly less relevant. It is evident that, in this Monte Carlo setting, microstructure effects are not present, so we can see that asynchronous quoting can be a very important factor explaining the reduction of the correlation measurements by itself. The effect of non-synchronicity on the correlation estimates of daily stock prices, which are typically recorded at slightly different closing hours, has been analyzed by Burns et al. (1998); Martens and Poon (2001) among others; these studies also show that this effect, which may look negligible at first

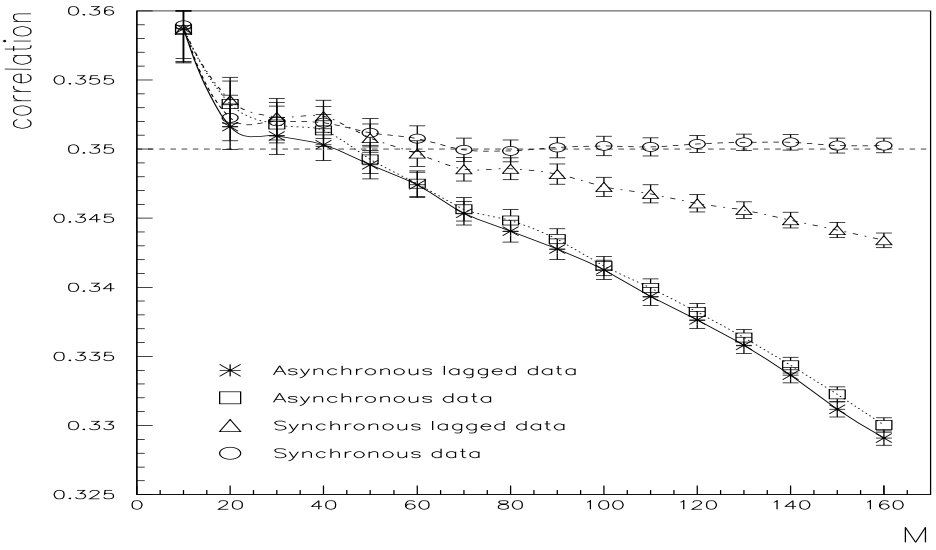

Figure 4.1: Average correlation between two Monte Carlo simulations of asset prices, according to (4.1), as a function of the sampling frequency M in (2.44). Boxes: simulated observation times are drawn independently. Crosses: simulated observation times are drawn independently and the correlation is lagged by 8 seconds. Triangles: simulated observation times are forced to be the same and the correlation is lagged by 8 seconds. Circles: simulated observation times are forced to be the same, no lag. The generated value of the correlation is $\rho = 0.35$, corresponding to the dashed line in the figure. Error bars are computed according to the Normal distribution. These results are obtained with 10,000 replications.

glance, can seriously affect the correlation estimates.

We now turn to the analysis of synchronous data. From Figure 4.1, we observe that the Epps effect is dramatically reduced. However, we find another relevant source of the Epps effect. When we consider lagged data, we still observe a substantial drop in correlation, even if it appears to be smaller than that caused by non-synchronicity. The effect is visible even if the introduced lag, i.e. 8 seconds, is very small, at least when compared with the frequency $M = 160$, which corresponds to 3.5 minutes in the time domain.

In contrast, if data are synchronous and not lagged, we clearly observe that correlations do not drop at higher frequencies. Moreover, as expected, by increasing the sampling frequency we increase the measurement precision. To give a quantitative idea of the magnitude of this effect, we should link the correlation drop to the level of non-synchronicity or lag. Monte Carlo experiments were repeated keeping fixed the average duration of the second time series,

Table 4.2: Correlation estimates on Monte Carlo experiments with asynchronous data. The second time series observations are extracted with an average duration of 45 seconds, while we report estimates at different values of the average duration τ_1 of the first time series, at three different frequencies. All estimates are obtained with 1,000 replications.

$\tau_1 \rightarrow$	15	25	35	45	55	65	75
$M = 100$	0.3452	0.3452	0.3437	0.3417	0.3384	0.3345	0.3297
$M = 130$	0.3373	0.3371	0.3354	0.3323	0.3279	0.3227	0.3162
$M = 160$	0.3311	0.3312	0.3288	0.3244	0.3184	0.3110	0.3028

Table 4.3: Correlation estimates on Monte Carlo experiments with synchronous lagged data. Both the time series observations are extracted with an average duration of 45 seconds, while we report estimates at different values of backward lag (in seconds) for the second time series, at three different frequencies. All estimates are obtained with 1,000 replications.

LAG $\rightarrow$	4	8	12	16	20	24	28
$M = 100$	0.3512	0.3496	0.3485	0.3476	0.3479	0.3456	0.3441
$M = 130$	0.3488	0.3465	0.3442	0.3435	0.3431	0.3399	0.3382
$M = 160$	0.3487	0.3449	0.3421	0.3398	0.3383	0.3343	0.3329

set to $\tau_2 = 45$ seconds, and varying the average duration of the first, named τ_1. Clearly, increasing τ_1 will decrease the activity of the first time series, and consequently the number of synchronous quotes. Table 4.2 shows the results for the frequencies $M = 100, 130, 160$. Not surprisingly, we confirm the inverse relation between activity and correlation drop found in Lundin et al. (1999). We repeat these experiments by keeping fixed the durations of the two time series to 45 seconds, and backward the second time series backward by a variable lag. Table 4.3 shows the results for different frequencies M, as above. As expected, we observe a larger drop in correspondence with a larger lag.

It is important to determine the significance of these results. If we look at the standard deviation of measurements across the Monte Carlo sample at $M = 160$, which provides a reasonable estimate of the error, we find that it is around 0.05 for all four data sample. With 10,000 replications, this corresponds to a standard deviation of the mean a hundred times smaller. Thus, on the basis of a simple variance ratio test for the mean at 95%

level, when comparing the correlations between the synchronous (not lagged) sample and the two asynchronous (lagged and not lagged) samples, we reject equality for $M \geq 60$, while when comparing the synchronous sample with the synchronous lagged, we reject equality for $M \geq 100$. Kolmogorov-Smirnoff tests of equality among the distributions were also performed[2], and they confirmed these results.

Finally, it seems quite interesting that the variance of the measurements is nearly the same[3] across the four samples. This finding indicates that synchronicity plays an important role also in the precision of the Fourier estimator. Indeed, switching from asynchronous to synchronous data, we reduce the number of observations of the first time series by one third, but the standard deviation of measurements remains nearly the same.

The question now becomes whether the Epps effect observed in the literature can be explained only by means of the non-synchronicity of quotes and lead-lag relationship. The following Section tries to address this point by looking at market data.

4.3.2 Data analysis

The first data sample under study is the very well known collection of DEM-USD and JPY-USD exchange rate quotes as they appeared on the Reuters screen from October, 1^{st} 1992 to September 30^{th} 1993. The price is defined as the bid-ask midpoint. This data set has been collected and distributed by Olsen & Associates, and it has been extensively studied in the high frequency data literature.

The analysis in Barucci and Renò (2002a) suggests that the highest M that can be used in (2.27), in order to prevent microstructure effects from distorting our results, is given by $M = 500$ for the DEM-USD time series and $M = 160$ for the JPY-USD time series. These values will be used when computing variances, while when computing covariances, since the two time series are analyzed jointly, $M \leq 160$ will be always used. This frequency corresponds to roughly 3.5 minutes in the time domain. Figure 4.2 shows the average daily[4] correlation as a function of the sampling frequency. Clearly the Epps effect is again in action, shifting the correlations downward above a given frequency. In order to investigate the effect of non-synchronicity, our measurements are repeated using only synchronous data, i.e. quotes which have the same time stamp. This kind of data are about 16% of DEM-USD quotes and

[2]Detailed results on the tests are available from the author upon request.

[3]The standard deviation at $M = 160$ is 0.0519 for asynchronous data, 0.0525 for synchronous data, 0.0522 for asynchronous lagged data, 0.0530 for synchronous lagged data.

[4]We define one day starts and ends at 21:00 GMT.

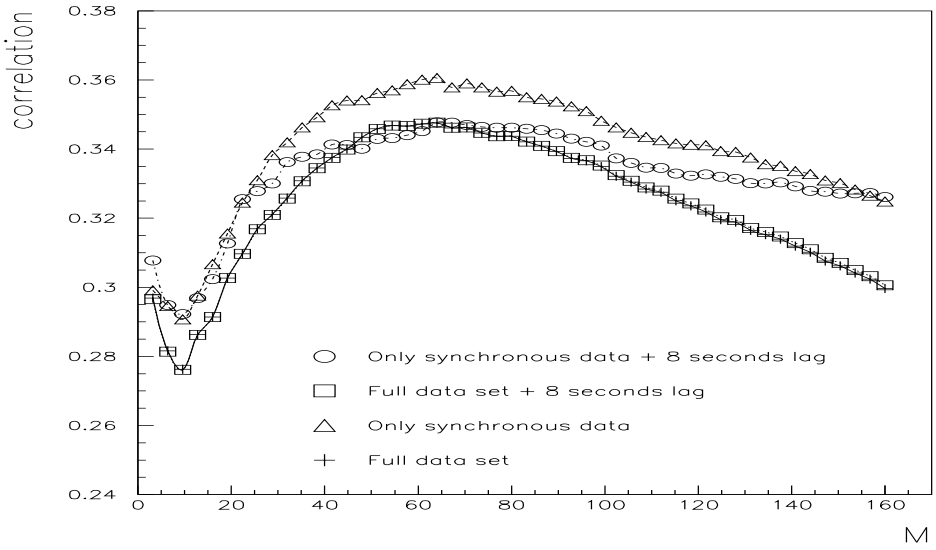

Figure 4.2: Average realized daily correlation between DEM-USD and JPY-USD foreign exchange rate, from October 1^st 1992 to September 30^{th} 1993, as a function of the sampling frequency M in (2.44). Crosses: all observations are included in the computation. Triangles: Only synchronous observations are included in the computation. Boxes: all observations are included, JPY-USD time series is shifted backward by 8 seconds. Circles: Only synchronous observations are included in the computation after shifting the JPY-USD time series backward by 8 seconds.

42% of JPY-USD quotes, so we have a substantial reduction of our data sample[5]. Results are again plotted in Figure 4.2. The reader may be confused by the different levels of correlation measurements in the different cases. These are due to the statistical fluctuations induced by the scarcity of observations, especially when using only synchronous data. The standard deviation of measurements ranges from 0.14 for asynchronous data to 0.18 for synchronous lagged data, thus, with 256 data points, standard error estimates are around 0.01. Thus the observed rise of correlation measurements when only synchronous observations are taken in account is not statistically significant. Figure 4.2 is redrawn in Figure 4.3 in a different fashion: while Figure 4.2 plots $\rho(M)$, Figure 4.3 plots $[\rho(M) - \rho(64)]/\rho(64)$, where M is the sampling frequency. We choose $\rho(64)$ to normalize since it is the highest measured correlation

[5]Actually, the number of synchronous quotes is surprisingly high. Indeed, we have on average 5653 DEM-USD quotes per day, and 2186 JPY-USD quotes per day. Taking into account that in our data sample quote times are rounded to the nearest even second, and assuming independent quoting, we expect on average 286 synchronous quotes per day, while we find 909! Clearly the assumption of independence is violated. For example, it is plausible that market makers post all their quotes contemporaneously.

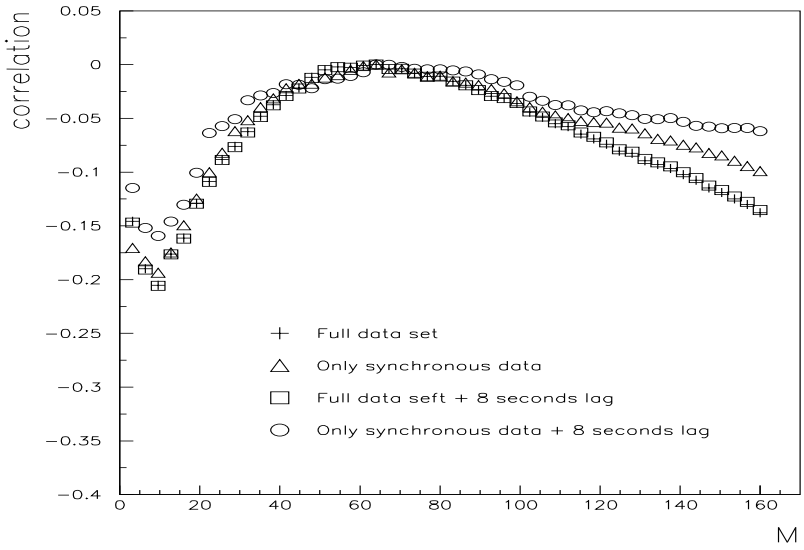

Figure 4.3: Shows $[\rho(M) - \rho(64)]/\rho(64)$ as a function of M, where M is the sampling frequency in (2.44). Crosses: all observations are included in the computation. Triangles: Only synchronous observations are included in the computation. Boxes: all observations are included, JPY-USD time series is shifted backward by 8 seconds. Circles: Only synchronous observations are included in the computation after shifting the JPY-USD time series backward by 8 seconds.

in all the four cases.

We observe that the effect of correlation reduction when the sampling frequency increases is less strong. However, this effect is not completely canceled out. This result shows that the non-synchronicity of quotes plays a substantial role in the Epps effect, but other effects must be taken into account when trying to explain it.

The presence of lead-lag effect in our exchange rate data sample is checked by computing the daily lagged correlation:

$$\rho(\tau) = \int_0^{2\pi} corr\left[p_1(t)p_2(t+\tau)\right] dt, \tag{4.3}$$

as a function of τ, where p_1 and p_2 denote DEM-USD and JPY-USD exchange rate respectively, and the interval $[0, 2\pi]$ denotes one day. Similar studies on lead-lag correlation have been conducted, for example, by Chan (1992); Ballocchi et al. (1999); Muller et al. (1997), using a much longer time scales than that analyzed in this Section. In the present analysis, again the Fourier method was used when computing (4.3), by computing the correlation after

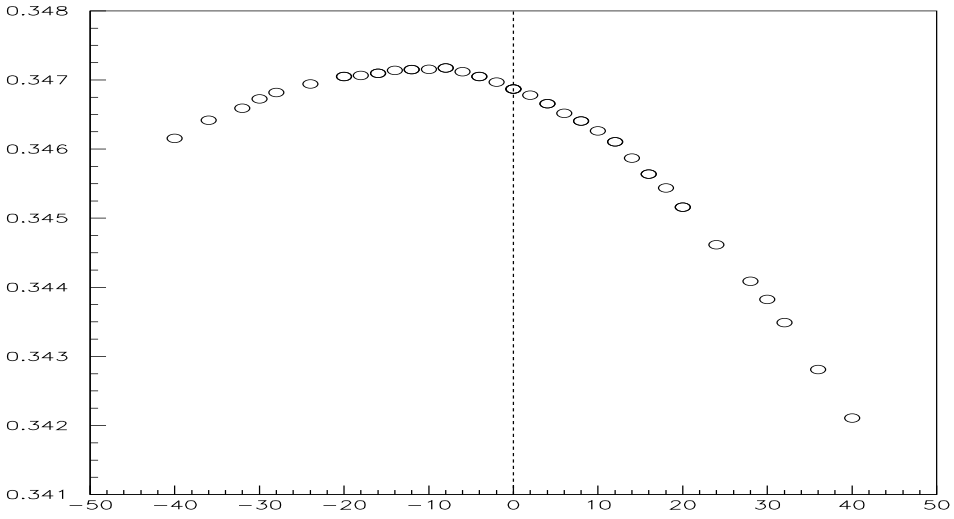

Figure 4.4: Average daily lagged correlation (4.3) between JPY-USD and DEM-USD exchange rates, as a function of τ; a negative τ indicates that JPY-USD leads DEM-USD.

shifting the second time series temporally by τ. Covariance is computed at $M = 64$. Results are shown in Figure 4.4. In our data sample the JPY-USD exchange rate leads DEM-USD by nearly 8 seconds. This lead-lag relationship looks very small when compared to the time scale we are investigating ($M = 160$, corresponding to nearly 4 minutes), but, as shown with Monte Carlo experiments, this could be enough to explain the drop in correlation measurements. Given that the two time series are both very liquid, and that the lag is found to be of the same order of magnitude of the average durations, we argue that this lag is likely to be due to non-synchronicity effects only.

The measurement of daily correlation between DEM-USD and JPY-USD is then repeated after shifting the JPY-USD time series backward by 8 seconds. These measurements are performed using all the observations and synchronous observations only (after shifting), as in the previous case. Results are shown in Figure 4.2: by removing the joint effect of lead-lag relationship and non-synchronicity (white circles) a strong reduction of the Epps effect is obtained. This is in agreement with what we found on Monte Carlo experiments (see Figure 4.1) on simulated time series. The lead-lag relation causes the Epps effect, even if we repeat the experiment forcing the two time series to be observed at the same time.

Figure 4.3 shows more clearly that using synchronous quotes, and adding a lag of 8 seconds, the Epps effect is strongly reduced. Comparing the measurements at $M = 160$ with those

at $M = 64$, we obtain a 12.9% drop if we do not apply any correction; a 12.7% drop when lagging backward the DEM-USD time series; a 9.0% drop when using only synchronous observations and a 5.3% drop when using only synchronous observations after lagging backward by 8 seconds the DEM-USD time series. This drop is found to be statistically significant. Equivalence tests on the mean of $[\rho(M) - \rho(64)]$ are performed across the samples: at 95% level, we reject the equality of the measurements on the synchronous-lagged sample for $M \geq 124$ with the full sample, for $M \geq 131$ with the lagged sample, and for $M \geq 156$ with the synchronous (not lagged) sample. In order to further explain that 5.3% drop, it seems that other factors should be taken into account. The failure of continuous-time models to describe high-frequency data is the first factor under suspicion. However, a larger data sample should be used to understand these second-order effects.

All these results can be checked by repeating all the measurements on two much less frequently traded asset prices. The second data set under study consists of the high frequency trades of the stock prices of Exxon and Mobil, from January 1995 to April 1995 for a total of 82 trading days. Only trades from 9:30 to 16:00 at the NYSE time were used. We have on average 397 Mobil trades and 724 Exxon trades per day. These two stocks belong to the same economic sector (oil) and were eventually merged into the same firm on November 30^{th} 1999, so there are good reasons to suppose that a substantial correlation should exist between the two time series considered. We start by checking the presence of a lead-lag effect of the same kind of that observed on FX rates. We use $M = 20$ when computing the variances of both stock prices, and $M = 10$ when computing the covariance.

Figure 4.5 shows that the lead-lag effect is again present, and in a more substantial way: the Mobil price leads Exxon by roughly a couple of minutes. On the basis of these results, we performed the correlation measurements on the raw data sample, on synchronous data only and on the data sample and only synchronous data after shifting the Exxon time series backward by 70 seconds, corresponding to the largest correlation in figure 4.5. Figure 4.6 presents the results. We notice immediately that large fluctuations among measurements are present, due to the fact that the number of trades (and especially of synchronous trades) is considerably smaller than exchange rate quotes. We have on average 20 synchronous trades per day, and 12 synchronous trades per day after shifting. The standard deviation of the means is estimated around 0.02 for asynchronous data, and 0.04 for synchronous data, both lagged and not lagged. The Epps effect is again evident when computing correlations with all the trades, and again using only synchronous trades drastically reduces the effect. When using synchronous trades after shifting backward the Exxon time series, the Epps effect is no longer observed, and correlations remain stable in the range $10 \leq M \leq 20$. Thus, the analysis of stock prices confirms the results obtained with FX rates.

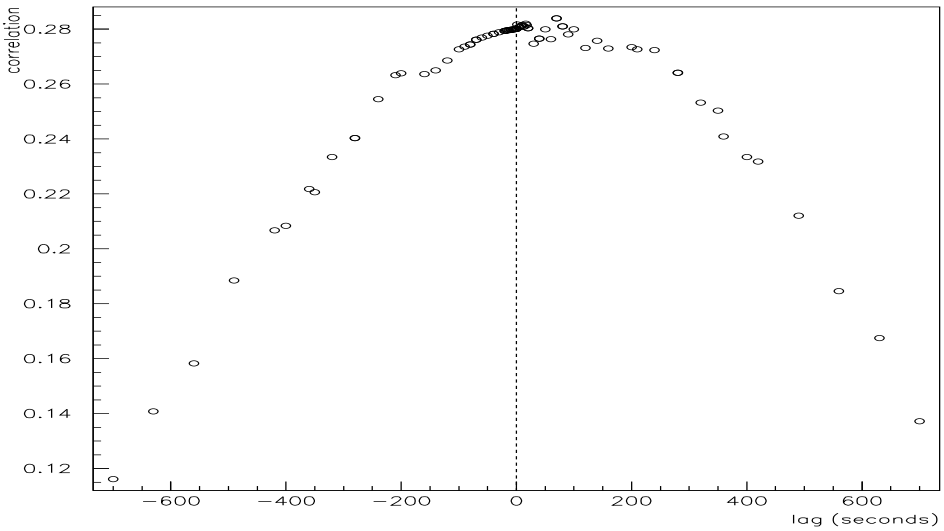

Figure 4.5: Average daily lagged correlation (4.3) between Mobil and Exxon stock prices, as a function of τ; a positive τ indicates that Mobil leads Exxon.

These results can be very important from the risk management point of view. It is clear that a correct assessment of risk management strategies has to take care of correlations, and especially of the stochastic nature of correlations (Ball and Torous, 2000). Our results suggest that, even if continuous-time multivariate models fail to describe the dynamics of asset prices at the very high frequencies, it is still possible to use them to compute variances and correlations precisely with high frequency data, just taking into account non-synchronicity and lead-lag relationships. From this perspective, the Fourier estimator turns out to be a promising tool.

4.4 Dynamic principal component analysis

In the previous Sections multivariate volatilities provided us a time dependent quadratic form on which we want to proceed to a time dependent principal component analysis.

Principal component analysis has been proposed as a natural statistical tool to analyze stock prices in the framework of factor models, as the APT of Ross (1976), or approximate factor models, as in Chamberlain and Rothschild (1983). Mostly, it has been used for the highly correlated bond of different maturities (Litterman and Scheinkman, 1991); in a recent application, Scherer and Avellaneda (2002) also analyzed the term structure via PCA in a

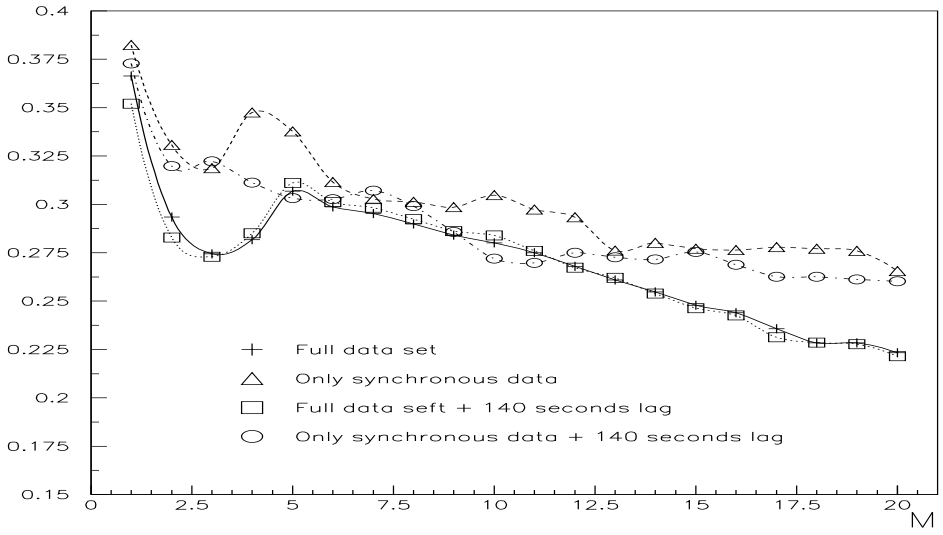

Figure 4.6: Average realized daily correlation between Exxon and Mobil stock prices, from January 1^{st} 1995 to April 30^{th} 1995, as a function of the sampling frequency M in (2.44). Crosses: all observations are included in the computation. Triangles: Only synchronous observations are included in the computation. Boxes: all observations are included, Exxon time series is shifted backward by 70 seconds. Circles: Only synchronous observations are included in the computation after shifting the Exxon time series backward by 70 seconds.

dynamic way. See also the work of Laloux et al. (1999); Plerou et al. (1999), who analyze via PCA a large covariance matrix of stock returns to separate noisy eigenvectors from significant eigenvectors.

Our analysis is model free and it is oriented toward dynamical description. Classical PCA provides a linear sub-manifold of smaller dimension which carries the essential information coming from the data. Dynamic PCA will produce an abstract curve, which we call the *Core*, which allows us to determine the eigenvectors of multivariate volatility in continuous time.

Consider the curve $p(t)$ describing the assets in $\mathbb{R}^n$. The geometric construction of the Core is the following: we consider the $n \times n$ multivariate volatility matrix $\Sigma(t)$ on the time window $[0, 2\pi]$. Denote by $G_{ij} = \Sigma_{ij}$ the inverse matrix of Σ^{ij}. It is known (Levi-Civita, 1925) that $G_{ij}(t)$ defines a Riemannian metric, because $\Sigma^{ij}(t)$ transforms as a contravariant 2-tensor. We fix an integer q with $q \ll n$, independent of t. Consider the q eigenvectors of $\Sigma(t)$ associated to the q largest eigenvalues. Denote by $V(t)$ the q-dimensional space spanned by

these q eigenvectors. We emphasize that, because $\Sigma^{ij}(t)$ is a contravariant 2-tensor, then $V(t)$ is a subspace of the cotangent vectors. Therefore $V(t)$ has to be considered as the cotangent space (that is the dual space to the tangent space) to the curve of interest, which we call the *Core* at time t. The restriction to $V(t)$ of the quadratic form defined by $\Sigma(t)$ defines a Riemannian metric on the Core.

In order to analyze the contribution of a given asset to volatility, we consider its projection on the Core. We choose an orthonormal basis of $V(t)$:

$$d\phi_s = \sum_{j=1}^{n} \gamma_s^j(t) dp_j \quad 1 \leq s \leq q. \tag{4.4}$$

Then $(d\phi_s)_{1\leq s\leq q}$ is a basis for $V(t)$ which, we recall, is a space of cotangent vectors. The projection of the asset p_1 on the Core is given:

$$d\psi = \sum_{s=1}^{q} \theta_s d\phi_s \tag{4.5}$$

where θ_s is the scalar product $du_1 * d\phi_s$. Finally consider

$$dU_1 = dp_1 - d\psi,$$

and analogous construction for $U_i,\ i = 1, \ldots, n$. Then U_i can be interpreted as the genuine fluctuation of the asset with price p_i. The value of U_i provides a measure of the idiosyncratic noise of the i-th asset, since it is the noise which cannot be accounted by the principal component eigenvalues. We say that the i-th asset is a "reference asset" if the function U_i is relatively small, according to a given criterion. We define in this way a basket of reference assets. In our dynamic setting it could be important to decipher if an asset suddenly gets outside the basket of reference assets.

4.4.1 Empirical Results

The data set at our disposal consists of the whole set of stock transactions, as recorded by NYSE and collected in the TAQ database, for the months of April and May, 2001. Given the huge quantity of data, we restrict our attention to 98 stocks, chosen among the most liquid ones; the selected stocks all belong to the $S\&P$ 100 index. In our sample, the most heavily traded stock is Cisco Systems, with one transaction every 1.21 seconds; the less frequently traded is Allegheny Technologies, with one transaction every 95 seconds.

On the whole, we analyze 42 days; for each day, we compute the variance-covariance matrix using the Fourier method described in Chapter 2. It is important to choose the frequency at

which to stop the expansion (2.44); as already remarked, we cannot choose it arbitrarily large because of microstructure effects. We adopt a unique value for all stocks; when computing variances, we use $M = 150$, corresponding to a time scale of a couple of minutes. When computing covariances, as mentioned earlier, we encounter the additional problem that, since stocks do not trade simultaneously, the covariance drops at larger frequency (Renò, 2003). Thus we have to use an *ad hoc* smaller frequency; guided by the data, we choose $M = 50$.[6] Figure 4.7 shows the time series of correlations among some selected stocks, and it illustrates the variety of patterns which can be observed in a stock market. Stocks in the same economic sector usually display larger positive correlation; sometimes, but rarely, negative correlation arises between stocks. Let us remark the peculiar behavior around the 12^{th} and the 32^{th} day, corresponding to April 18^{th} and May 16^{th}, 2001. In both these days, the Federal Reserve cutted the federal funds rate. The cut on April 18^{th} was particurarly unexpected.

The thick line in Figure 4.8 shows the distribution of correlations across all stocks for all days in our sample; since we have $n = 98$ stocks, in total we have $\frac{n^2}{2} - n = 4753$ correlations per day. The plot shows that correlations among stocks are rather weak; the average value is 0.1732, thus it is not possible to conjecture that few factors could explain the behavior of the market, as in the case of interest rates with different maturities (Litterman and Scheinkman, 1991). However, the correlation distribution changes substantially from day to day; in particular, there are two days (those corresponding to the FED cut) in our sample in which correlations are sparse, indicating that most of the stocks are more strongly correlated, but some other become more negatively correlated.

We start by performing PCA for every different day, after normalizing the variance-covariance matrix in order to have the variance of every stock equal to 1. Figure 4.9 shows the percentage of the variance explained by $1, 5, 10, 20, 30$ factors. We see that the first factor explains, in average, 25.79% of the movements; this is not that much, and it is in line with the small average correlation coefficient. Anyway in some days the first factor's weight can be as large as 56.09% (April, 18^{th}) and 73.26% (May, 16^{th}). Moreover, and more interestingly for any financial application, this phenomenon seems to present some degree of persistence.

In the second step of our analysis, we define the Core of the market as the vector subspace spanned by the first 30 eigenvectors, and we divide our 42 market days into 6 periods of 7 market days each. In each subperiod, we perform principal component analysis on the aggregate variance-covariance matrix, and we obtain the coordinates of the Core. In each

[6]The use of different frequencies for variances and covariances could, in principle, result in measured correlations larger than one. It is instead straightforward to show that, choosing the same frequency for all the elements of the variance-covariance matrix, the Fourier estimator for the correlation lies between -1 and 1.

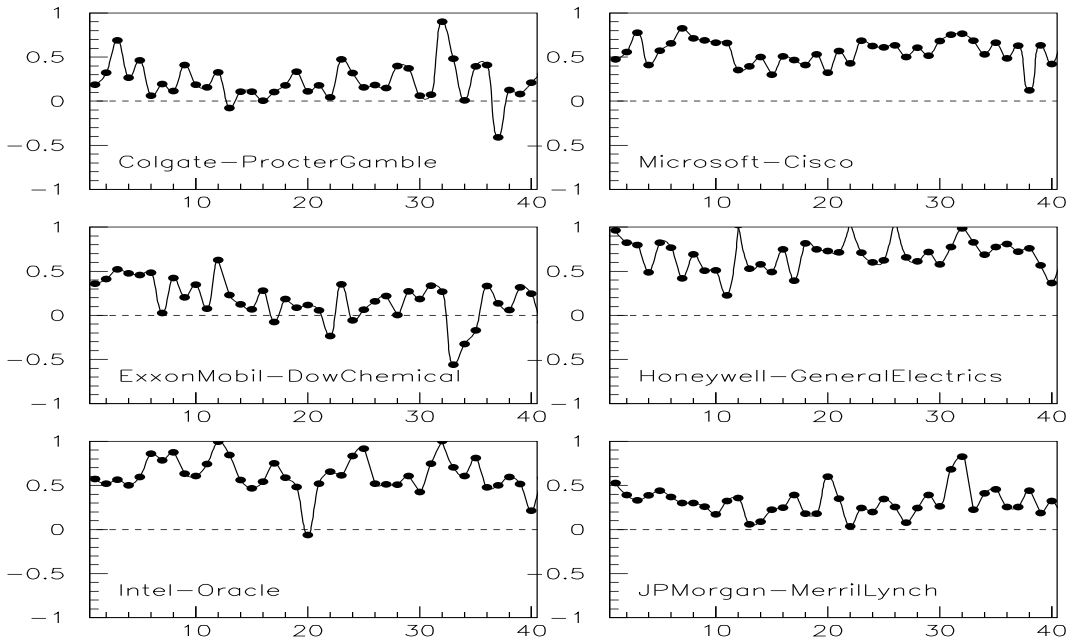

Figure 4.7: Time series of correlation (42 daily measurements) for different stocks; stocks in the same economic sector (Microsoft-Cisco or Honeywell-General Electric) display stronger positive correlation.

subperiod, we define 15 reference assets as those who have the largest projection onto the Core. We interpret these assets as those who are more correlated with the market itself, or alternatively as the basic constituents of the market. The list of the reference assets in any subperiod is shown in Table 4.4. Given the low value of the average correlation, we expect that the basket of reference assets is quite variable, given the unalienable noise in the correlation measurements. This is indeed the case. The month of April shows more persistence: 4 reference assets in the first period, out of 15, are in the second period too; and 6 of the second are still reference assets in the third. The month of May shows much more variability, or less "market integration"; only two stocks are reference assets in the third and fourth period, only one in the fourth and fifth period and none out of 15 in the fifth and sixth period. Looking at individual stocks, AES Corporation is a reference asset in the whole sample, with the exception of the sixth period, and in three periods it is the asset with the largest absolute projection on the Core. Table 4.4 also shows the percentage projection on the Core, and the five assets, in any period, with the lowest projection onto it. For example, in period 2, 93.3% of the variance of AES can be considered to be driven by

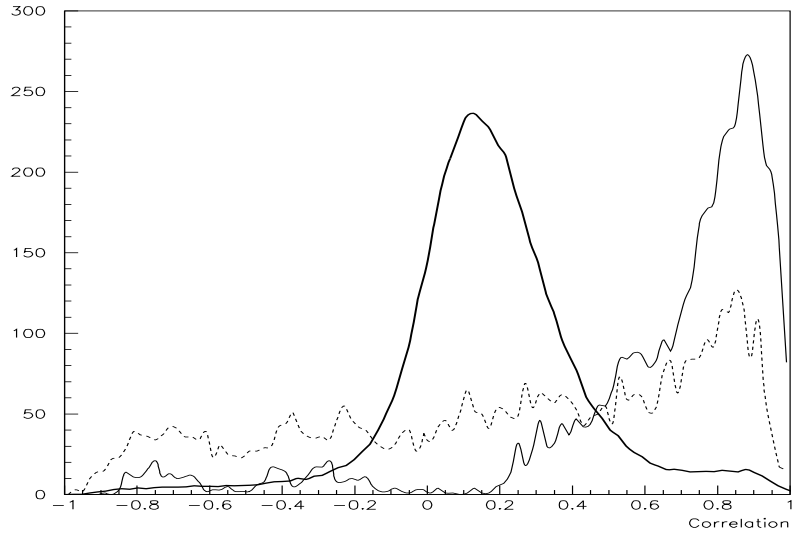

Figure 4.8: The tick solid line shows the cumulative distribution of correlations across the 98 stocks considered. The narrow solid line shows the distribution on April 18^{th}, 2001, while the dashed line shows the distribution on May 16^{th}, 2001. In these two days, the FED lowered the federal funds rate.

the market, and 6.7% is explained by idiosyncratic fluctuation; thus AES essentially lies on the Core, which is the subspace which explains most of the variance of the whole market; on the other hand, in the same period for Coca-Cola (KO), only 35.1% of its variance is driven by the market, while 64.9% is independent fluctuation.

Summarizing, out of the 98 assets, 32 are never reference assets; 46 are only once; 18 are twice, Johnson & Johnson (JNJ) is thrice and AES is five times. Then our analysis identifies nearly 20 assets which had a major role in market integration. In the set of this 20 assets, 8 are among the 20 most capitalized; thus capitalization plays an important role in defining leading assets, but it is not the only factor to be taken in account. For example, in our analysis AES turned out to be the most important stock, but its capitalization (measured as market value) is only about 0.5 % of the capitalization of Microsoft, the most capitalized stock in our sample.

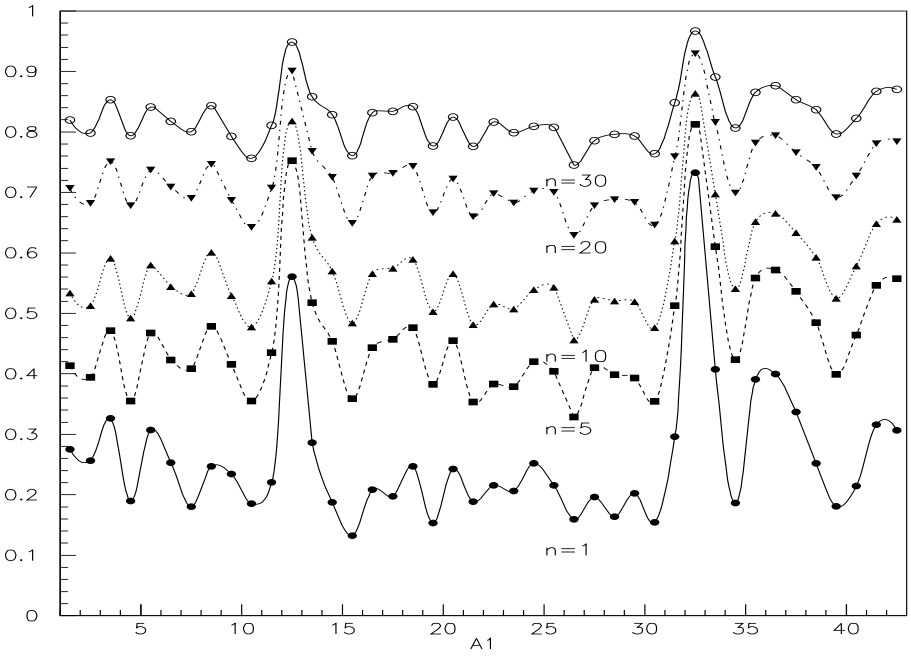

Figure 4.9: Percentage of eigenfactors variance explanations, for $n = 1, 5, 10, 20, 30$ factors, for the 42 days considered.

4.5 Conclusions

In this Chapter, we outline a new procedure, stemming from the contribution in Malliavin and Mancino (2002), to compute pointwise variance-covariance matrices among assets. Monte Carlo experiments illustrate how this procedure performs better than other commonly used estimators in the literature.

An investigation of the so-called Epps effect (Epps, 1979) is attempted. To this aim, we adopt the Fourier estimator to compute cross-correlations. This estimator is well suited to the time structure of high frequency data and to our frequency analysis. When tested on Monte Carlo bivariate experiments, the Fourier estimator proves to be a good candidate for computing correlations in a precise way. In our Monte Carlo experiments, price diffusions are simulated by a continuous-time model, namely the GARCH(1,1) continuous-time model. The results show that the Epps effect may be explained by the non-synchronicity of quotes and by lead-lag relationships.

Further evidence for the Epps effect is derived from foreign exchange data. When considering

Table 4.4: Lists the reference assets by ticker name in the six subperiods considered. They are ranked according to their projection on the core, which is reported in brackets. Also the five assets with the smallest projection on the Core are listed. Each subperiod is composed of 7 market days. In bold face, we indicate those stocks who remain reference assets in the subsequent period

Periods					
1	2	3	4	5	6
F (74.4%)	**AES** (93.3%)	**AES** (84.0%)	WY (72.2%)	AES (88.9%)	HON (86.4%)
CCU (72.2%)	AOL (86.1%)	**WFC** (77.1%)	TOY (72.1%)	HIG (78.7%)	GM (74.7%)
AOL (71.1%)	INTC(84.1%)	ATI (73.9%)	**AES** (72.0%)	HAL (76.7%)	DOW (72.1%)
INTC(70.0%)	AA (82.0%)	LU (73.4%)	WMB (68.2%)	WMT (71.8%)	IBM (70.4%)
AES (69.5%)	JPM (80.0%)	AEP (72.5%)	CSC (67.7%)	HWP (71.7%)	GE (70.4%)
PHA (68.8%)	IBM (79.9%)	AVP (72.2%)	XOM (67.2%)	TYC (70.9%)	LTD (70.4%)
IP (64.6%)	**JNJ** (78.5%)	TXN (70.3%)	MSFT(66.3%)	ONE (68.8%)	BUD (69.7%)
UTX (64.1%)	**AEP** (77.5%)	BUD (69.8%)	PFE (65.8%)	HNZ (68.3%)	FDX (68.4%)
CPB (64.0%)	**ATI** (74.8%)	AIG (67.2%)	PEP (65.1%)	F (68.1%)	JNJ (67.6%)
GE (63.7%)	**MRK** (70.5%)	MRK (64.9%)	CSCO(65.1%)	UTX (67.3%)	MAY (67.1%)
TYC (63.6%)	IP (70.0%)	LTD (64.8%)	BAX (64.4%)	HCA (66.8%)	MSFT(66.4%)
EK (63.2%)	T (69.6%)	AMGN(64.7%)	BMY (64.3%)	ORCL(66.1%)	AIG (65.5%)
VZ (62.0%)	MCD (67.7%)	MDT (63.5%)	SLB (63.1%)	HET (65.5%)	ETR (65.3%)
BNI (62.0%)	**LU** (67.5%)	JNJ (61.7%)	WFC (63.1%)	DIS (64.9%)	MMM (64.6%)
UIS (61.6%)	MEDI(66.6%)	EMC (61.5%)	ETR (62.8%)	VIAB(64.1%)	AXP (63.3%)
BMY (43.0%)	USB (35.9%)	BNI (41.4%)	IBM (40.4%)	MER (37.4%)	HCA (37.4%)
AXP (42.8%)	HWP (35.2%)	HON (39.7%)	IP (40.3%)	LU (36.3%)	PFE (36.9%)
MMM (42.0%)	S (35.2%)	DIS (38.1%)	LEH (38.3%)	NT (34.4%)	SLE (36.7%)
MSFT(39.1%)	KO (35.1%)	IBM (37.9%)	AIG (37.6%)	AA (34.3%)	VIAB(35.4%)
MAY (37.4%)	GD (33.1%)	DD (37.5%)	HD (31.0%)	WFC (30.9%)	CL (34.6%)
BAX (37.0%)	VZ (28.5%)	T (29.1%)	T (30.3%)	HON (28.0%)	AES (30.2%)

only synchronous quotes, the effect is reduced but not eliminated. It is shown that in our sample there is a lead-lag relationship of 8 seconds. However, this very short time scale is enough to generate a frequency effect on correlations. If the measurements are performed using synchronous quotes after shifting one time series by 8 seconds, the Epps effect is drastically reduced. All these results are confirmed by the analysis of stock price data, where the lag is found to be around 70 seconds.

We conclude that even if other factors, apart from non-synchronicity and lead-lag relationships, which clearly play the main role, concur in the Epps effect, their effect is negligible. It remains questionable if these factors can be explained in the framework of continuous-time models.

It would be also interesting to find precise laws which relate the magnitude of the Epps

effect to, say, correlation, average durations, lag, cut-off frequency. A Monte Carlo approach would turn out to be too time-consuming, given the number of parameters which affect the correlation estimates, and an analytic answer, as in Scholes and Williams (1977), would be desirable. We leave this topic for future research.

Finally, we provide a geometrical interpretation of our results in term of dynamic principal component analysis. We illustrate the reliability and the potential of the method, by analyzing two months of high-frequency data. Our results show that, since correlations are time-varying, also principal component analysis can be addressed in a time-varying fashion. In particular, the number of factors needed to explain the co-movements in the market varies dramatically with time and it is influenced by macroeconomic announcements, such as the FED decisions.

We introduce the definition of reference assets, as those which are more integrated with the market. Our analysis, though preliminary, shows that capitalization plays an important role in defining a reference asset, but other factors must be taken into account.

In this Chapter, we concentrated on contemporaneous correlation, while a structural feature of the market is to exhibit delayed or lead-lag correlations. It could be important to construct an estimator which takes into account this delay: we leave this problem to further research.

Chapter 5

Nonparametric estimation

In this Chapter a new fully nonparametric estimator of the diffusion coefficient is introduced, based on Fourier analysis of the observed trajectory. The proposed estimator is proved to be consistent and asymptotically normally distributed. After testing the estimator on Monte Carlo simulations, we use it to estimate an univariate model of the short rate with available interest rate data. Data analysis helps shedding new light on the functional form of the diffusion coefficient.

5.1 Introduction

The last decade witnessed a growing literature in the field of the diffusion coefficient estimation. The main motivation underlying this strand of research is that the diffusion coefficient, which is called volatility, plays a fundamental role in practically every financial application. We concentrate on univariate models of the kind:

$$dX_t = \mu(X_t)dt + \sigma(X_t)dW_t \tag{5.1}$$

where W_t is a standard real Brownian motion and the real functions $\mu(x)$ and $\sigma(x)$ are such that a unique solution X_t of the stochastic differential equation (5.1) exists. X_t can be any variable; however, in this Chapter we concentrate on short rate modeling, that is $X_t = r_t$. Our specific problem is then to estimate the diffusion term $\sigma(r)$ when we observe a discrete realization of the process r, namely n observation $\hat{r}_{t_1}, \ldots, \hat{r}_{t_n}$ in the interval $[0, T]$.

The methods for measuring volatility can be coarsely divided into parametric and nonparametric. The parametric approach consists in specifying the function $\sigma(r) = \sigma(r; \vec{\theta})$, with $\vec{\theta}$

being a vector of real parameters. As a popular example, a large parametric class has been explored by Chan et al. (1992), who study the following model:

$$dr_t = \beta(\alpha - r_t)dt + \sigma r_t^{\gamma} dW_t,$$

where $\beta, \alpha, \sigma, \gamma$ are real numbers. This specification nests many popular one-factor models, like the constant variance model of Vasicek (1977), for $\gamma = 0$ or the square-root diffusion of Cox et al. (1985), for $\gamma = 0.5$. The methodology is to estimate $\vec{\theta}$ through point estimation. For interest rates diffusions, this can be done via maximum likelihood (Duffie et al., 2002) or GMM, direct (Chan et al., 1992) or simulated (Gallant and Tauchen, 1996; Dai and Singleton, 2000). The advantage of parametric models is that closed form solutions exist for bond and derivative pricing. On the other hand, the advantage of nonparametric specification is clearly its flexibility. For example, Jiang (1998) shows that the nonparametric specification provides more accurate prices for bonds and derivatives. However, for bond and derivative pricing one has to resort to Monte Carlo simulations since closed form solutions are out of reach.

One important example of nonparametric estimator for the diffusion coefficient is that proposed by Ait-Sahalia (1996a). The Ait-Sahalia estimator is based on the fact that, if the spot rate r_t is driven by equation (5.1), then we have:

$$\sigma^2(r) = \frac{2\displaystyle\int_{-\infty}^{r} \mu(y)\pi(y)dy}{\pi(r)} \tag{5.2}$$

where $\pi(r)$ is the unconditional distribution of r. Equation (5.2) allows, given two out of the three functions $\mu(r), \sigma(r), \pi(r)$, to obtain the third after integration or derivation. This estimator is not fully nonparametric, since to get an estimate of the variance a specification of the drift term is needed. Ait-Sahalia (1996a) suggests to specify the drift $\mu(r)$ as an affine function of r, then to estimate the conditional variance $\sigma(r)$. Given the drift $\mu(r)$, the estimator (5.2) still depends on the unconditional distribution $\pi(r)$. However, we can obtain an estimate of $\pi(r)$ with a nonparametric technique (Scott, 1992) and replace $\pi(r)$ in (5.2) with its estimate. Suppose our observations are equally spaced, and denote them by $\hat{r}_i, i = 1, \ldots, n$. Then the nonparametric estimator of the density is given by:

$$\hat{\pi}(r) = \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{r - \hat{r}_i}{h}\right) \tag{5.3}$$

where $K(\cdot)$ is the kernel function and h a bandwidth parameter which depends on n. One popular way of estimating densities through formula (5.3) is the histogram, where the kernel function is an indicator function of a compact real interval centered around zero.

Ait-Sahalia (1996a) estimation results on interest rate data show that a departure from classical univariate models (e.g. CIR and Vasicek) is observed. In a related paper, Ait-Sahalia (1996b) rejected almost all the most popular one-factor model used for modeling interest rates. However this rejection is very controversial. The problem is that estimating (5.3) in finite samples can be problematic if the mean reversion is low. For example, Pritsker (1998) shows that the rejection of the most popular parametric models comes from severe underestimation of confidence interval for testing the null in small samples.

The most remarkable example of a fully nonparametric estimator of the diffusion coefficient can be found in Florens-Zmirou (1993). She introduces an estimator which is conceptually different from that used in Ait-Sahalia (1996a), since it does not need any assumption on the drift. Florens-Zmirou (1993) shows that, given discretely sampled data, the diffusion coefficients in (5.1) may be estimated by:

$$\hat{\sigma}^2(r) = \frac{n\sum_{i=1}^{n-1}(\hat{r}_{i+1} - \hat{r}_i)^2 K\left(\frac{r-\hat{r}_i}{h}\right)}{T\sum_{i=1}^{n} K\left(\frac{r-\hat{r}_i}{h}\right)}. \tag{5.4}$$

The variance estimator (5.4) looks more appealing since there is the same kernel in the numerator and in the denominator, so biases in finite samples coming from nonparametric estimation of the density could cancel out. The estimator (5.4) has been used by Jiang and Knight (1997) on Canadian interest rates, and by Stanton (1997) on U.S. interest rates. In both those papers, the Authors conclude in favor of a departure from standard models, and they suggest a strong mean reversion for values of the spot rate r less than 3% and larger than 15%, see Chapman and Pearson (2000) for a discussion on these results.

An estimator similar to that of Florens-Zmirou (1993) has been proposed by Bandi and Phillips (2003), and studied in Bandi (2002). The estimator proposed in Bandi and Phillips (2003) is the following:

$$\hat{\sigma}^2(r) = \frac{n\sum_{i=1}^{n} K\left(\frac{r-\hat{r}_i}{h}\right)\left(\frac{1}{m_i}\sum_{j=0}^{m_i}\left[\hat{r}_{t_{i,j}+1} - \hat{r}_{t_{i,j}}\right]^2\right)}{T\sum_{i=1}^{n} K\left(\frac{r-\hat{r}_i}{h}\right)} \tag{5.5}$$

where $t_{i,j}$ is a subset of indexes such that

$$t_{i,0} = \inf\left\{t \geq 0 : |\hat{r}_t - \hat{r}_i| \leq \varepsilon_s\right\},$$

and

$$t_{i,j+1} = \inf\left\{t \geq t_{i,j} + \Delta t : |\hat{r}_t - \hat{r}_i| \leq \varepsilon_s\right\},$$

m_i is the number of times that $|\hat{r}_t - \hat{r}_i| \leq \varepsilon_s$, ε_s is a parameter to be selected and Δt is the time step between adjacent observations. We refer the reader to the cited paper for details. It is important to remark that Bandi and Phillips do not require the process (5.1) to be stationary, but only the weaker condition to be recurrent. This condition can be important theoretically, since Bandi (2002) and many others, e.g. Ball and Torous (1996), show that there is no strong support to the assumption of stationarity of interest rate data. Watching carefully expressions (5.4) and (5.5), we can see that the difference between Florens-Zmirou and Bandi-Phillips estimators is that, while the Florens-Zmirou estimator weights the observation r_t with the quadratic variation at time t, the Bandi-Phillips estimator weights the observation r_t with the average quadratic variation at all observations which are "close" to r_t.

Other estimators have been proposed, in a more general framework, by Jacod (1999). Hoffmann (1999) proposes a wavelet estimator which is consistent in $\mathcal{L}^p$. Both these authors study convergence rate properties for their estimator, showing that the Florens-Zmirou (1993) estimator is not optimal. However, asymptotic distribution can only be assessed for the estimator (5.4).

In this Chapter we introduce a new fully nonparametric estimator of the diffusion coefficient of an univariate stochastic differential equation. The estimator is fully nonparametric in the sense that we do not impose any restriction on the functional form of the drift term. Moreover, it is developed under mild regularity conditions for the stochastic differential equation (5.1). As for the estimator in Bandi and Phillips (2003), the stationarity assumption is not strictly required, being substituted by the milder assumption of recurrency. The estimator is proved to be consistent in the $\mathcal{L}^2$ sense, and asymptotically normally distributed. In order to assess the asymptotic properties, we borrow from the limit theory for semimartingales, and in particular of the convergence of a semimartingale to a process with independent increments. The asymptotic distribution turns out to be identical to that of Florens-Zmirou (1993). However, the estimator is basically different, and this allows us to re-examine the estimation of the diffusion coefficient with the available interest rate data.

This Chapter is structured as follows. Section 5.2 presents the estimator, and shows its consistency and asymptotic normality, using limit theory for semimartingales. In Section 5.3 the estimator is implemented on Monte Carlo experiments of the Vasicek model, and compared with the estimator of Florens-Zmirou (1993). In Section 5.4 we implement the estimator for measuring the variance of the diffusion of short interest rates, and compare our results to those in the literature, and in particular with the results of Stanton (1997). Finally, Section 5.5 concludes.

5.2 Nonparametric estimation of the diffusion coefficient

We consider the SDE:

$$\begin{cases} dX_t = \mu(X_t)dt + \sigma(X_t)dW_t \\ X_0 = x_0 \end{cases} \tag{5.6}$$

in the filtered probability space $(\Omega, (\mathcal{F}_t)_{0\leq t\leq T}, P)$ satisfying the usual conditions. We will write $X_t(\omega)$ to explicit the dependence of X from $t \in [0,T]$ and $\omega \in \Omega$. We assume the following:

Assumption 5.1 *Given the SDE (2.1), we have:*

1. $x_0 \in \mathcal{L}^2(\Omega)$ *is independent of* $W_t,\ t \in [0,T]$ *and measurable with respect to* $\mathcal{F}_0$.
2. $\mu(x)$ *and* $\sigma(x)$ *are defined on a compact interval* I. $\mu(x)$ *is once continuously differentiable,* $\sigma(x)$ *is thrice continuously differentiable.*
3. *There exists a constant* K *such that* $0 < \sigma(x) \leq K$ *and* $|\mu(x)| \leq K$.
4. *(Feller condition for non-explosion). Given:*

$$S(\alpha) = \int_0^\alpha e^{\int_0^y -\frac{2\mu(x)}{\sigma^2(x)}dx}dy, \tag{5.7}$$

$$V(\alpha) = \int_0^\alpha S'(y)\int_0^y \frac{2}{S'(x)\sigma^2(x)}dxdy, \tag{5.8}$$

then $V(\alpha)$ *diverges at the boundaries of* I.

Assumption 5.1 insures existence and uniqueness of a strong solution. Asking for the Feller condition allows to deal with models which, as noted by Ait-Sahalia (1996a), do not satisfy global Lipschitz and growth condition (e.g. CIR). Moreover, Feller condition is necessary and sufficient for recurrence in I, see the discussion in Bandi and Phillips (2003). Alternatively, one can ask for global Lipschitz and growth conditions on μ and σ (Karatzas and Shreve, 1988).

Asking for a bounded volatility (and drift) is harmless from an econometric point of view, since we always observe a finite, thus bounded, set of observations. For example, in the CIR model the variance is unbounded and proportional to $\sqrt{X_t}$. However, since estimation is on a finite sample, the observations $\hat{X}_i$ are bounded and $\sqrt{X_t}$ is indistinguishable from $\min(\sqrt{X_t}, \sqrt{\max \hat{X}_i})$.

Moreover we will consider a kernel function for nonparametric estimation with the following properties:

Assumption 5.2 *We define a kernel $K(\cdot)$ a bounded function in $\mathcal{L}^2(\mathbb{R})$ which is continuously differentiable, with bounded first derivative, positive, with $\int_{\mathbb{R}} K(s)ds = 1$ and such that $\lim_{s\to\pm\infty} K(s) = 0$ faster than any inverse polynomial.*

A typical choice is the Gaussian kernel:

$$K(s) = \frac{1}{\sqrt{2\pi}} e^{-\frac{s^2}{2}} \tag{5.9}$$

We will moreover consider a sequence h_n such that:

Assumption 5.3 *$(h_n)_{n\in\mathbb{N}}$ is a real sequence such that, as $n \to \infty$, we have $h_n \to 0$ and $nh_n \to \infty$.*

An example which is very popular in applications (Scott, 1992) is the following:

$$h_n = h_s \hat{\sigma} n^{-\frac{1}{5}} \tag{5.10}$$

where h_s is a real constant to be tuned, and $\hat{\sigma}$ is the sample standard deviation. We will assume 5.1,5.2,5.3 holding throughout all the analysis.

Consider the solution process X_t with $t \in [0, T]$. We now consider the fact that the process X_t is usually recorded at equally spaced times. When subdividing the interval in n steps of equal length, we use, as shorthand notation, $X_i = X_{iT/n}$, that is $X_0 = x_0$, $X_n = X_T$. Moreover, we set $t_i = iT/n$. Assumption 5.1 implies that X_t is a continuous semimartingale, thus we can define its local time (Revuz and Yor, 1998) as:

$$L_t(x) = \lim_{\varepsilon\to 0} \frac{1}{2\varepsilon} \int_0^t I_{]x-\varepsilon, x+\varepsilon[}(X_\tau) d\tau \tag{5.11}$$

We can estimate the local time of a diffusion via the following approximation:

$$L_t^n(x) = \frac{1}{nh_n} \sum_{i=1}^{[nt]} K\left(\frac{X_i - x}{h_n}\right) \tag{5.12}$$

where $[x]$ is the integer part of x. We have indeed:

Proposition 5.4 *If, as $n \to \infty$ we have $nh_n^4 \to 0$, then $L_t^n(x) \to L_t(x)$ in the $\mathcal{L}^2$ sense. The convergence is almost sure if $\frac{\log n}{nh_n^2} \to 0$.*

Proof. These are Propositions 1 and 2 in Florens-Zmirou (1993). □

The estimator of the diffusion coefficient proposed in Florens-Zmirou (1993); Stanton (1997); Jiang and Knight (1997) is based on the following quantity:

$$V_t^n(x) = \frac{1}{Th_n} \sum_{i=0}^{n-1} K\left(\frac{X_i - x}{h_n}\right) (X_{i+1} - X_i)^2 \tag{5.13}$$

We have indeed,

Proposition 5.5 *If $nh_n^4 \to 0$ as $n \to \infty$, then $V_t^n(x)$ converges to $\sigma^2(x)L_t(x)$ in the $\mathcal{L}^2$ sense.*

Proof. This is Proposition 3 in Florens-Zmirou (1993). □

Thus dividing $V_t^n(x)$ by $L_t^n(x)$ we get a consistent estimator of $\sigma^2(x)$. In our analysis, relying on the result of Chapter 2, we want to substitute the quantity (5.13) with the following:

$$U_t^n(x) = \frac{1}{Tnh_n} \sum_{i=0}^{n-1} K\left(\frac{X_i - x}{h_n}\right) \hat{\sigma}^2(t_i) \tag{5.14}$$

where $\hat{\sigma}^2(t_i)$ is computed via (2.4) on the observed trajectory of X_t.

We then define the estimator:

$$S^n(x) = \frac{\sum_{i=1}^{n} K\left(\frac{X_i - x}{h_n}\right) \hat{\sigma}^2(t_i)}{T \sum_{i=1}^{n} K\left(\frac{X_i - x}{h_n}\right)} \tag{5.15}$$

We now prove that $S^n(x)$ is a consistent estimator of $\sigma^2(x)$.

Theorem 5.6 *If $nh_n^4 \to 0$ as $n \to \infty$, then $S^n(x)$ is a consistent estimator of $\sigma^2(x)$ in the $\mathcal{L}^2$ sense.*

Proof. Suppose $\mu(x) = 0$ in (2.1). For every $\omega \in \Omega$, consider the solution $X_t(\omega)$ of equation (2.1), and define the process Y_t^ω defined by:

$$dY_t^\omega = \sigma(X_t(\omega))dW_t' \tag{5.16}$$

where W'_t is a standard Brownian in an auxiliary probability space $(\Omega', (\mathcal{F}'_t)_{0\leq t\leq T}, P')$, and $\sigma(X_t(\omega))$ is the realization of $\sigma(X_t)$. Assumption (5.1) guarantees the existence of the solution of (5.16). Then we can construct random variables in Ω by taking expectations in Ω'. We denote by E' the expected value in Ω' and by $\mathbb{E}$ the expected value in Ω.

Using equation (2.36), we get for $0 \leq t < s \leq T$ and almost surely in Ω:

$$\begin{aligned}
\frac{1}{s-t}E'[(Y_t^\omega - Y_s^\omega)^2] &= \frac{1}{s-t}\int_t^s \sigma^2(X_u(\omega))du = \\
&= \frac{1}{s-t}\int_t^s \sum_{q=-\infty}^{+\infty} A_q(\sigma^2)e^{iqu}du = \\
&= \sum_{q=-\infty}^{+\infty} A_q(\sigma^2)e^{iqt}\frac{e^{iq(s-t)}-1}{iq(s-t)} = \\
&= \sigma^2(X_t(\omega)) + \sum_{q=-\infty}^{+\infty} A_q(\sigma^2)e^{iqt}\left(\frac{e^{iq(s-t)}-1}{iq(s-t)} - 1\right) = \\
&= \sigma^2(X_t(\omega)) + F(s,t)
\end{aligned} \tag{5.17}$$

since the Fourier series can be integrated term by term. Moreover, since the integral of a Fourier series converges uniformly, we have almost surely in Ω:

$$\lim_{s\to t} F(s,t) = 0 \tag{5.18}$$

Now it is simple to prove that, almost surely:

$$\mathbb{E}\left[E'[(Y_t^\omega - Y_s^\omega)^2]|\mathcal{F}_s\right] = \mathbb{E}[(X_t - X_s)^2|\mathcal{F}_s], \tag{5.19}$$

where X is the solution of (2.1), since both are equal to $\int_s^t \mathbb{E}[\sigma^2(X_u)|\mathcal{F}_s]du$. Moreover, using $\mathbb{E}[(X_t - X_s)^4|\mathcal{F}_s] = 3\mathbb{E}^2[\int_s^t \sigma^2(X_u)du|\mathcal{F}_s]$ and Cauchy-Schwartz inequality, we get almost surely:

$$\mathbb{E}\left[E'^2[(Y_t^\omega - Y_s^\omega)^2]|\mathcal{F}_s\right] \leq \mathbb{E}[(X_t - X_s)^4|\mathcal{F}_s], \tag{5.20}$$

Now, let us denote the $\mathcal{L}^2(\Omega)$ norm of X by $\|X\|^2 = \mathbb{E}[X^2]$. We then use almost sure identity (5.17) and get:

$$\begin{aligned}
&\left\|U_t^n(x) - \sigma^2(x)L_t^n(x)\right\| = \left\|\frac{1}{Th_n}\sum_{i=1}^{[nt]-1} K\left(\frac{X_i - x}{h_n}\right)E'\left[\left(Y_{i+1}^\omega - Y_i^\omega\right)^2\right] + \right. \\
&\left. -\sigma^2(x)L_t^n(x) - \frac{1}{Tnh_n}\sum_{i=1}^{[nt]-1} K\left(\frac{X_i - x}{h_n}\right)F(t_i, t_{i+1})\right\|
\end{aligned} \tag{5.21}$$

When expanding the square in the norm (5.21), we can use the fact that $\mathbb{E}[X] = \mathbb{E}[\mathbb{E}[X|\mathcal{F}_{t_i}]]$, then use (5.19) and (5.20), to get:

$$\begin{aligned}
&\left\| U^n(x) - \sigma^2(x)L^n(x) \right\| \leq \\
&\leq \left\| V^n(x) - \sigma^2(x)L^n(x) - \frac{1}{Tnh_n} \sum_{i=1}^{[nt]-1} K\left(\frac{X_i - x}{h_n}\right) F(t_i, t_{i+1}) \right\| \leq \\
&\leq \left\| V^n(x) - \sigma^2(x)L_t^n(x) \right\| + \left\| \frac{1}{Tnh_n} \sum_{i=1}^{[nt]-1} K\left(\frac{X_i - x}{h_n}\right) F(t_i, t_{i+1}) \right\|.
\end{aligned} \tag{5.22}$$

Both terms converge to zero: the first, because of Proposition 5.5; the second given Proposition 5.4 and since $F(t_i, t_{i+1}) = o\left(\frac{1}{n}\right)$.

If $\mu(x) \neq 0$, then (5.19) becomes:

$$\mathbb{E}\left[E'[(Y_t^\omega - Y_s^\omega)^2]\right] = \mathbb{E}[(X_t - X_s)^2] - \mathbb{E}\left[\left(\int_s^t \mu(X_u)du\right)^2\right], \tag{5.23}$$

and the second term of the r.h.s vanishes as $s \to t$ i.e. $n \to \infty$. □

We want now to assess the asymptotic normality of $S^n(x)$. We first state two lemmas.

Lemma 5.7 *If $nh_n^3 \to 0$ when $n \to \infty$, then,*

$$\sum_{i=1}^{[nt]} \mathbb{E}\left[\frac{n}{h_n}\left(K\left(\frac{X_i - x}{h_n}\right)[(X_{i+1} - X_i)^2 - \sigma^2(x)/n]\right)^2 \middle| \mathcal{F}_{i-1}\right] \to \sigma^4(x)L_t(x) \tag{5.24}$$

where the above convergence is in probability.

Proof. This is Lemma 2(b) in Florens-Zmirou (1993). □

Lemma 5.8 *Let $g(x) : \mathbb{R} \to \mathbb{R}$ be a continuously differentiable bounded function, with bounded first derivative. Let $nh_n^3 \to 0$ when $n \to \infty$. Consider:*

$$G_t(x) = \frac{1}{\sqrt{nh_n}} \sum_{i=1}^{[nt]} K\left(\frac{X_i - x}{h_n}\right)[g(X_i) - g(x)] \tag{5.25}$$

then, as $n \to \infty$, $G_t(x) \to 0$ in the $\mathcal{L}^1$ sense, and thus in probability.

Proof. We have:

$$\mathbb{E}[|G_t(x)|] \leq \mathbb{E}\left[\frac{1}{\sqrt{nh_n}}\sum_{i=1}^{[nt]} K\left(\frac{X_i - x}{h_n}\right)|g(X_i) - g(x)|\right] \tag{5.26}$$

We now divide the sum in terms such that $|X_i - x| \leq n^{-\frac{1}{3}}$ and their complementary. Then:

$$\begin{aligned}
&\mathbb{E}[|G(x)|] \leq \mathbb{E}\left[\frac{1}{\sqrt{nh_n}}\sup_{|X_i-x|\leq n^{-\frac{1}{3}}}|g(X_i)-g(x)|\sum_{|X_i-x|\leq n^{-\frac{1}{3}}} K\left(\frac{X_i-x}{h_n}\right)+\right.\\
&\left.+\frac{1}{\sqrt{nh_n}}\sum_{|X_i-x|>n^{-\frac{1}{3}}} K\left(\frac{X_i-x}{h_n}\right)|g(X_i)-g(x)|\right] \leq\\
&\leq \mathbb{E}\left[\sqrt{nh_n}\sup_{|X_i-x|\leq n^{-\frac{1}{3}}}|g(X_i)-g(x)|\sum_{i=1}^{[nt]}\frac{1}{nh_n}K\left(\frac{X_i-x}{h_n}\right)+\right.\\
&\left.+\frac{1}{\sqrt{nh_n}}\sup_{|X_i-x|>n^{-\frac{1}{3}}}|g(X_i)-g(x)|\sum_{|X_i-x|>n^{-\frac{1}{3}}} K\left(\frac{X_i-x}{h_n}\right)\right]
\end{aligned} \tag{5.27}$$

Now, using Taylor's rule, and given the boundedness of the first derivative of $g(\cdot)$, we get that $\sup_{|X_i-x|\leq n^{-\frac{1}{3}}}|g(X_i) - g(x)| = o(n^{-\frac{1}{3}})$. Then, using Proposition 5.4, we have that the first term goes as $(nh_n^3)^{\frac{1}{6}}$, then it goes to zero as $n \to \infty$. The second term goes to zero given the boundedness of g, the fact that the argument of the kernel is greater than $1/n^{\frac{1}{3}}h_n = 1/(nh_n^3)^{\frac{1}{3}}$, and the fact that the kernel goes to zero faster than inverse polynomials when its argument goes to infinity, as in this case.

□

We finally state the main result of this Chapter. The idea is still to substitute $n(X_{i+1} - X_i)^2$ with $\hat{\sigma}^2(t_i)$ in (5.24), with the remainder vanishing in probability.

Theorem 5.9 *If $nh_n^3 \to 0$ then*

$$\sqrt{nh_n}\left(\frac{S_n(x)}{\sigma^2(x)} - 1\right) \xrightarrow{\mathcal{L}} \frac{1}{\sqrt{L_T(x)}}\mathcal{N}(0,1), \tag{5.28}$$

where the above convergence is in law, and $\mathcal{N}(0,1)$ is a standard Normal variable.

Proof. Consider the discrete filtration $\mathcal{F}_i = \mathcal{F}_{t_i},\ i = 0, \ldots, n$. Define the following:

$$\Theta_{i+1} = \sqrt{\frac{1}{nh_n}}K\left(\frac{X_i - x}{h_n}\right)\left[\sigma^2(X_i) - \sigma^2(x)\right] \tag{5.29}$$

Since $\sigma^2(\cdot)$ is bounded, Θ_i is bounded, and it is adapted to $\mathcal{F}_{i-1}$ (it is actually a constant with respect to $\mathcal{F}_{i-1}$). We want now to verify the conditions of Theorem 1.51. In the limit $n \to \infty$, we have the following:

1. $\sum_{i=0}^{[nt]-1} \mathbb{E}[\Theta_{i+1}|\mathcal{F}_i] = \sum_{i=0}^{[nt]-1} \Theta_{i+1}$ which tends to 0 in probability, given Lemma 5.8.

2. We have to prove that $\sum_{i=0}^{[nt]-1} \mathbb{E}[\Theta_{i+1}^2|\mathcal{F}_i] \to 0$. We can use formula (5.17) and use the same reasoning leading to the proof of Theorem 5.6 getting:

$$\begin{aligned}
&\sum_{i=0}^{[nt]-1} \mathbb{E}\left[\Theta_{i+1}^2|\mathcal{F}_i\right] = \\
&\sum_{i=0}^{[nt]-1} \mathbb{E}\left[\frac{1}{Tnh_n}\left\{K\left(\frac{X_i - x}{h_n}\right)\left[nE'\left[(Y_{i+1}^{\omega} - Y_i^{\omega})^2\right] - F(t_i, t_{i+1}) - \sigma^2(x)\right]\right\}^2 \middle| \mathcal{F}_i\right] \leq \\
&\leq \sum_{i=0}^{[nt]-1} \mathbb{E}\left[\frac{1}{Tnh_n}\left\{K\left(\frac{X_i - x}{h_n}\right)\left[n(X_{i+1} - X_i)^2 - F(t_i, t_{i+1}) - \sigma^2(x)\right]\right\}^2 \middle| \mathcal{F}_i\right]
\end{aligned} \tag{5.30}$$

   From this inequality and Lemma 5.7 we get the result.

3. We have to prove conditional Lindeberg condition. We have:

$$\mathbb{E}\left[\left|\mathbb{E}\left[\Theta_i^2|\mathcal{F}_{i-1}\right]\right|\right] = \sum_{|\Theta_i| > \varepsilon} \Theta_i^2. \tag{5.31}$$

   Now, the sum (5.31) is bounded by $\sigma^4(x)L_t(x)$; moreover, we can rewrite $|\Theta_i| > \varepsilon$ as:

$$K\left(\frac{X_i - x}{h_n}\right)|\sigma^2(X_i) - \sigma^2(x)| > \varepsilon\sqrt{nh_n} \tag{5.32}$$

   The l.h.s of equation (5.32) is bounded, thus as $n \to \infty$ we have $\varepsilon\sqrt{nh_n} \to \infty$ and the sum (5.31) vanishes in probability.

Thus, we fulfill the assumptions of Theorem 1.51, then if we define $Y_t^n(x)$ as:

$$Y_t^n(x) = \sum_{i=0}^{[nt]-1} \Theta_{i+1}(x) \tag{5.33}$$

than we have that $Y_t^n(x)$ converges in law to the continuous martingale M_t with quadratic variation $[M, M]_t = \sigma^4(x)L_t(x)$. We then set $M_t = B(\sigma^4(x)L_t(x))$, where $B(t)$ is a standard Brownian motion. Now consider:

$$Z_t^n(x) = \sum_{i=0}^{[nt]-1} \left(W_{t_{i+1}} - W_{t_i}\right) \tag{5.34}$$

where W_t is the standard Brownian motion in (2.1). It is clear that $Z_t^n(x)$ converges in law to the standard Brownian motion W_t. Moreover we have:

$$\sum_{i=0}^{[nt]-1} \mathbb{E}\left[\Theta_{i+1}\left(W_{t_{i+1}} - W_{t_i}\right)|\mathcal{F}_i\right] = 0 \tag{5.35}$$

By equation (5.35), we get that M_t and W_t are orthogonal. We can also write $B(t) = M_{T(t)}$, where $T(t) = \inf_s\left(\frac{s}{\sigma^4(x)L_s(x)}\right)$. Then, by Knight's Theorem 1.41 we get that $B(t)$ and W_t are independent Brownian motions. Then $B(t)$ and $L_t(x)$ are independent, since the filtration generated by X_t is included in the filtration generated by W_t. We then have that $Y_t^n(x) \to \sqrt{L_t(x)}\sigma^2(x)\mathcal{N}(0,1)$, where $\mathcal{N}(0,1)$ is a standard normal random variable independent of $L(x)$. Since $L_t^n(x)$ converges in probability to $L_t(x)$, we have the desired result. □

The above result can be easily generalized to a multivariate framework, and, more importantly, if the observations X_i are not equally spaced, but if they are such that, as $n \to \infty$, then we have $sup|t_i - t_{i-1}| \to 0$.

5.3 Small sample properties

In the previous Section we assessed the asymptotic properties of the estimator (5.15). We now turn to the analysis of small sample properties, analyzing them via Monte Carlo simulations.

The selected model for simulations is the very popular Vasicek (1977) model:

$$dr_t = k(\alpha - r)dt + \sigma dW_t, \tag{5.36}$$

which displays mean reversion and constant variance. As it is well known, the unconditional distribution of r_t implied in equation (5.36) is Normal with mean α and variance $\sigma^2/2k$. While it is universally known that the Vasicek model does not provide a satisfactory description of real world interest rates, it is still useful in our framework as a first step. Our purpose is to estimate the variance, and starting with the simpler model in this sense, that is constant variance, we should be able to disentangle small sample peculiarities coming from the estimator from those which are a statistical product of the model.

We simulate the model with parameters resembling actual interest rates distribution. For example, the annualized variance of 3-months T-bill is around 3%, so we keep this value throughout all our simulations. From previous literature, it is well known that two parameters play a crucial role: the choice of the bandwidth parameter h_s in (5.10), and the mean reversion parameter k.

The choice of the bandwidth parameter has been long debated in the literature. While consistency of nonparametric estimators is independent of h_s, convergence rates and small sample properties depend crucially on it. Scott (1992) suggests the choice of $h_s = 1.06$, as it is optimal with respect to the mean integrated standard error criterion. Anyway, typical choices in the literature were larger than this value. For example, Stanton (1997) uses $h_s = 4$, while Ait-Sahalia (1996a) uses $h_s = 5$.

The role of mean-reversion is even more debated. The strength of mean reversion is measured by the parameter k. Estimates of k in the literature range around 0.1, which is a very low value. Thus many studies argue that in the available interest rate data there is no mean reversion at all, if not for extreme values (Ait-Sahalia, 1996a; Bandi, 2002). Chapman and Pearson (2000) show, via Monte Carlo evidence, that nonparametric methods could be biased toward finding non-linearities in the drift even if the drift is linear. Jones (2003) concludes in a similar way, and he argues that mean reversion in available interest rate data is so weak, if present, that its detection is very difficult with any statistical method.

We then select a grid of (h_s, k) ranging across the values of interest. We will select $h_s = 1.06, 3, 5$ and $k = 0.05, 0.5, 5.0$. In all the replications of the model (5.36), we draw the starting value from the marginal distribution, and we use sample of $n = 8,000$ observations, for comparison with the data set in the next Section. For every value of h_s and k, we use $5,000$ replications.

Figure 5.1 shows the average measurements on Monte Carlo replications, together with estimated confidence intervals, obtained with the estimator (5.15). Simulations show that the Fourier estimator is unbiased in small samples for the selected parameters. As expected, confidence intervals are broader for smaller mean reversion and smaller bandwidth parameter, and broader for large and small interest rates, which are less frequent.

We now turn to the comparison with the classical estimator (5.4). We choose to compare with this instead of the Ait-Sahalia estimator or the estimator (5.5) since Renò et al. (2004) show that the first one performs very badly for low levels of mean reversion, while the second one produces results which are almost identical to those obtained with the Florens-Zmirou estimator.

When comparing the estimators in small sample, we have to check not only unbiasedness and precision of the estimator, but the reliability of the asymptotic confidence intervals, since those are the ones actually used to draw inference. Thus it is also important to check the normality of the distribution of the estimators in finite samples. We will then compute the

standard deviations, skewness and kurtosis[1] on the estimates on replications.

In the case of the Vasicek model, asymptotic standard errors for the estimator (5.15), which are the same for the classical estimator, are easily computed and, denoting the asymptotic standard deviation by $s(r)$, it is given by:

$$s^2(r) = \frac{\sqrt{2\pi}\sigma^4}{h_s n^{\frac{4}{5}} \sqrt{2k} e^{-k\frac{(r-\alpha)^2}{\sigma^2}}} \tag{5.39}$$

where $n = 8,000$ is the length of every simulation. We will then plot standardized estimates, skewness and kurtosis. Under normality of the distributions we would get standardized variance equal to 1, skewness equal to 0 and kurtosis equal to 3. Figure 5.2 shows the results for the standardized variance. In small samples, we observe deviations from the asymptotic values. In the case of high mean reversion ($k = 5.0$) the two estimators performs nearly in the same way, overestimating the asymptotic variance. The same conclusion can be drawn in the case of average mean reversion $k = 0.5$, but in this case the two estimator underestimate the asymptotic variance. The most interesting case, which is the closer to actual data, is the case of small mean reversion ($k = 0.05$). The estimator of Florens-Zmirou (1993) considerably underestimates the asymptotic variance, thus leading to wrongly narrower confidence intervals. The Fourier estimator performs in a completely different way: it overestimates the variance for large and small values of r, while performs like the classical estimator in the center of the distribution. These results seem to suggest a choice of h_s in the range $1 - 3$ when implementing the Fourier estimator.

The reasons for this underestimation can be guessed looking at the small sample skewness and kurtosis, plotted in Figures 5.3 and 5.4 respectively. While the skewness for both the estimators is close to the normal value of zero for high and average mean reversion, in the case of small mean reversion the skewness of the classical estimator is always positive, while that of the Fourier estimator oscillates around zero. On the other hand, the kurtosis of the two estimators is comparable, with the Fourier estimator being more leptokurtic in the case of high and average mean reversion in the tails, but less leptokurtic in the center in the case

[1] Given N observations $x_1, \dots x_n$, whose sample mean is μ, define the j-th moment ($j > 2$) as:

$$\mu_j = \frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^j \tag{5.37}$$

Then we define the skewness ν and the kurtosis κ as:

$$\nu = \frac{\mu_3}{(\mu_2)^{\frac{3}{2}}}, \qquad \kappa = \frac{\mu_4}{(\mu_2)^2}. \tag{5.38}$$

Under the assumption of normality, standard errors for skewness and kurtosis are $\sqrt{6/N}$ and $\sqrt{24/N}$ respectively.

of small mean reversion.

5.4 Data Analysis

In this Section, we turn to the analysis of interest rate data. Our aim is to estimate the diffusion coefficient of the univariate model:

$$dr_t = \mu(r_t)dt + \sigma(r_t)dW_t \tag{5.40}$$

when discretely observing the short rate r_t in a time interval $[0, T]$. Since the spot rate is inherently unobservable, we use proxies for it, typically the three-months rate as it is common in the literature, see Duffee (1996); Chapman et al. (1999). Alternatively, one can regard the model (5.40) as a model for the three-months rate itself.

We first test the methodology on the same data set in Jiang (1998), that is the daily time series of the annualized yields of the three-months U.S. Treasury Bill, from January 1962 to January 1996, for a total of $8,503$ observations. The minimum and the maximum of the yield in this time span are 2.61 % and 17.14 %, thus the estimates of the diffusion coefficient outside of this interval are an artifact of the nonparametric estimation procedure. Figure 5.5 shows the estimation results. The estimate obtained with the Fourier estimator (5.15), using as bandwidth parameter (5.10) with $h_s = 4$, is the solid line. Confidence intervals are computed according to (5.28), using estimated local times via equation (5.12). The Fourier estimator is implemented with the maximal $N = n/2$ and with $M = n/4$. The Fourier estimator confirms the departure from standard parametric models, such as the Vasicek variance $\sigma^2(r) = k$, or the CIR variance $\sigma^2(r) = kr$. In order to better clarify this point, we consider the parametric model of Chan et al. (1992) which nests many popular one-factor models including CIR and Vasicek:

$$dr_t = \mu(\alpha - r_t)dt + \sigma r_t^{\gamma} dW_t. \tag{5.41}$$

This model has been estimated in Jiang (1998) on the same data set using indirect inference. Parameter estimates are $\alpha = 0.079(0.044),\ \mu = 0.093(0.100),\ \gamma = 1.474(0.008),\ \sigma = 0.794(0.019)$, where standard errors are in brackets. Figure 5.5 shows the function σr^{γ} for comparison with the nonparametric estimate. While the shape of the two estimates is increasing in both case, and the two estimates are compatible for r around 8%, we get a significantly higher estimate at low interest rates, and a significantly lower estimate at high interest rates when using the nonparametric method. It is clear that we are exploiting the flexibility provided by the nonparametric methodology.

We then compare the estimates with those obtained with other nonparametric estimators. To this purpose, we use the same data set used in Stanton (1997), that is the daily time series of the annualized yields of the three-months U.S. Treasury Bill, from January 1965 to July 1995, for a total of $n = 7,975$ observations (minimum 2.61 %, maximum 17.15 %). Thus, we can directly compare our results with those obtained in Stanton (1997). From this perspective, we use the same value $h_s = 4$ used by Stanton. Figure 5.6 shows that the two estimates are quite different. With respect to the estimate obtained in Stanton (1997), the Fourier estimate coincides only in the central part of the distribution, i.e. $r \simeq 11\%$, while it is higher for smaller values of r and considerably smaller for larger values of r. For larger values of r confidence intervals are wider, since the local time is small for the paucity of observations in that zone. For further comparison purposes, we also estimate the variance with the nonparametric estimator proposed in Bandi and Phillips (2003), with the same bandwidth parameter $h_s = 4.0$ and $\varepsilon_s = 1.5\%$ in (5.5). The result obtained with the estimator proposed in Bandi and Phillips (2003) is almost identical to that obtained with the Stanton estimator, confirming that the empirical performance of the two estimators is nearly the same. We do not report the estimate obtained with the Ait-Sahalia method, since it is very different from those obtained here, and it is very unstable at the level of mean reversion displayed by three-months interest rates, see Renò et al. (2004).

Finally, we estimate the diffusion coefficient on the full data set at our disposal, that is the daily yields on the three-months Treasury Bill from from 4 February 1960 to 11 December 2003, for a total of $10,944$ observation (minimum 0.79% on 19 June 2003, maximum 17.14% on 11 December 1980) and the daily yields on the ten-years Treasury Note from 2 February 1962 to 11 December 2003, for a total of $10,447$ observations (minimum 3.10% on 13 June 2003, maximum 15.51% on 4 September 1981). Figure 5.7 shows the results which are in line with the previous findings. We also find that the volatility of the longer maturity contract is less than the shorter one, as it is well known. We leave the reader the judgment on the opportunity for using such a long data set when estimating the model, given the heterogeneity of the economic conditions which drove the interest rate evolution: it is quite clear that the answer to this question depends on the specific application.

5.5 Conclusions

In this Chapter a new nonparametric estimator for the diffusion coefficient based on discrete observations is introduced. This estimator is based on a result on volatility estimation derived in Malliavin and Mancino (2002). The estimator is proved to be asymptotically consistent and normally distributed, and asymptotic confidence intervals are provided. The estimator

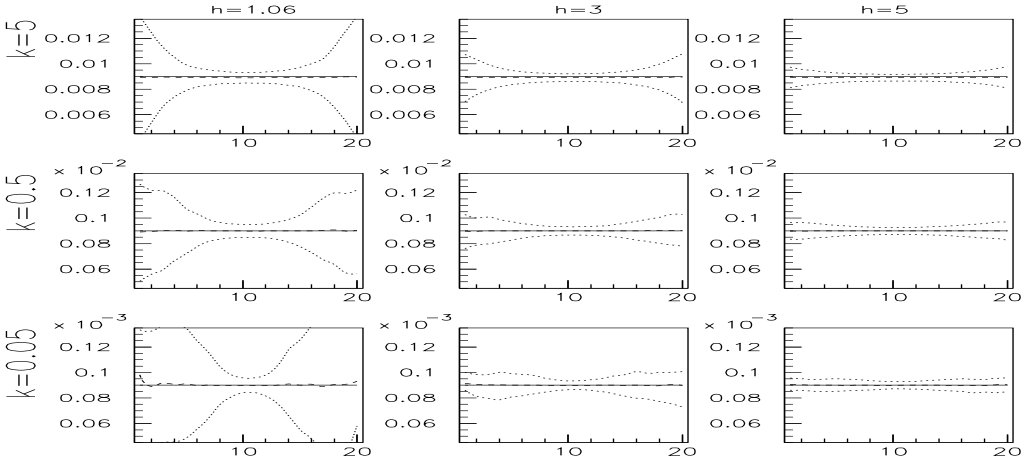

Figure 5.1: Variance $\hat{\sigma}^2(r)$ obtained with the Fourier estimator (5.15) on 5,000 replications of 8,000 observations of the Vasicek model (5.36), with different values of h and k, and $\alpha = 10.5\%$, $\sqrt{\sigma^2/2k} = 3\%$. Solid line: the generated variance σ^2. Dashed line: average estimates. Dotted lines: 5% and 95% confidence intervals.

is then compared to the estimator proposed in Florens-Zmirou (1993) and discussed in Jiang and Knight (1997) and Stanton (1997), via Monte Carlo simulations of the Vasicek model. Our results are quite mixed, indicating the difficulty of assessing small sample properties of nonparametric estimators. Anyway, our Monte Carlo evidence shows that the Fourier estimator is more symmetric than the classical one, and this could lead to better inference. This nonparametric estimators, as well others, can be used in a variety of applications. We used it to compute the diffusion coefficient for a daily time series of short interest rates. Our results are in line with those in the literature, but with some peculiarity. We show that our nonparametric estimates is quite different from standard parametric specifications. Moreover, the estimate with the Fourier estimator provides larger variances for interest rates smaller than 9% and smaller variances for interest rates larger than 12% than the variances obtained on the same data set by Stanton (1997). We conclude that the estimator proposed here can be a very useful tool in the problem of correctly specifying the diffusion term of interest rate stochastic models.

On the other hand, a serious limitation of this approach is that the assumption of a univariate process (5.1) for interest rate modelling, and in general for asset price modelling, is too restrictive. It is well known that stochastic volatility is a salient feature of asset price diffusions. It would be then desiderable to extend the results presented in this Chapter

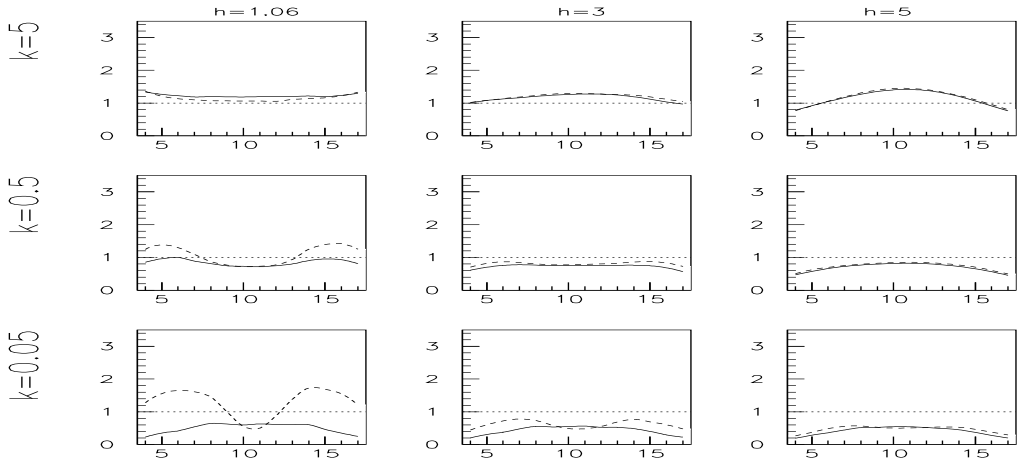

Figure 5.2: Standardized standard deviation, with the asymptotic standard deviation (5.39), of the estimates $\hat{\sigma}^2(r)$ obtained with the Fourier estimator (5.15) (dashed line) and the classical estimator (5.4) (solid line) on $5,000$ replications of $8,000$ observations of the Vasicek model (5.36), with different values of h and k, and $\alpha = 10.5\%$, $\sqrt{\sigma^2/2k} = 3\%$. Solid line: the classical estimator. We also plot the asymptotic value of 1.

to multivariate diffusions in which volatility is a latent factor. This topic is now under investigation.

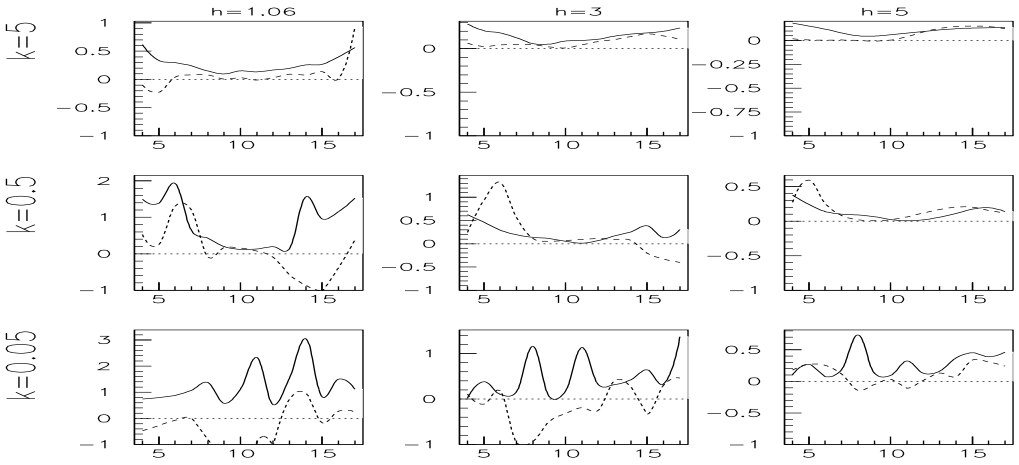

Figure 5.3: Skewness of the estimates $\hat{\sigma}^2(r)$ obtained with the Fourier estimator (5.15) (dashed line) and the classical estimator (5.4) (solid line) on $5,000$ replications of $8,000$ observations of the Vasicek model (5.36), with different values of h and k, and $\alpha = 10.5\%$, $\sqrt{\sigma^2/2k} = 3\%$. Solid line: the classical estimator. We also plot the asymptotic value of 0.

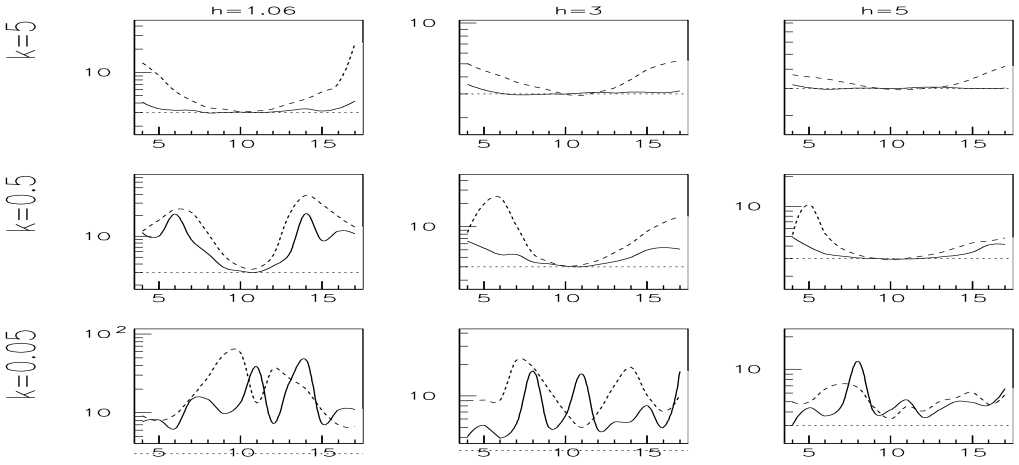

Figure 5.4: Kurtosis of the estimates $\hat{\sigma}^2(r)$ obtained with the Fourier estimator (5.15) (dashed line) and the classical estimator (5.4) (solid line) on $5,000$ replications of $8,000$ observations of the Vasicek model (5.36), with different values of h and k, and $\alpha = 10.5\%$, $\sqrt{\sigma^2/2k} = 3\%$. Solid line: the classical estimator. We also plot the asymptotic value of 3.

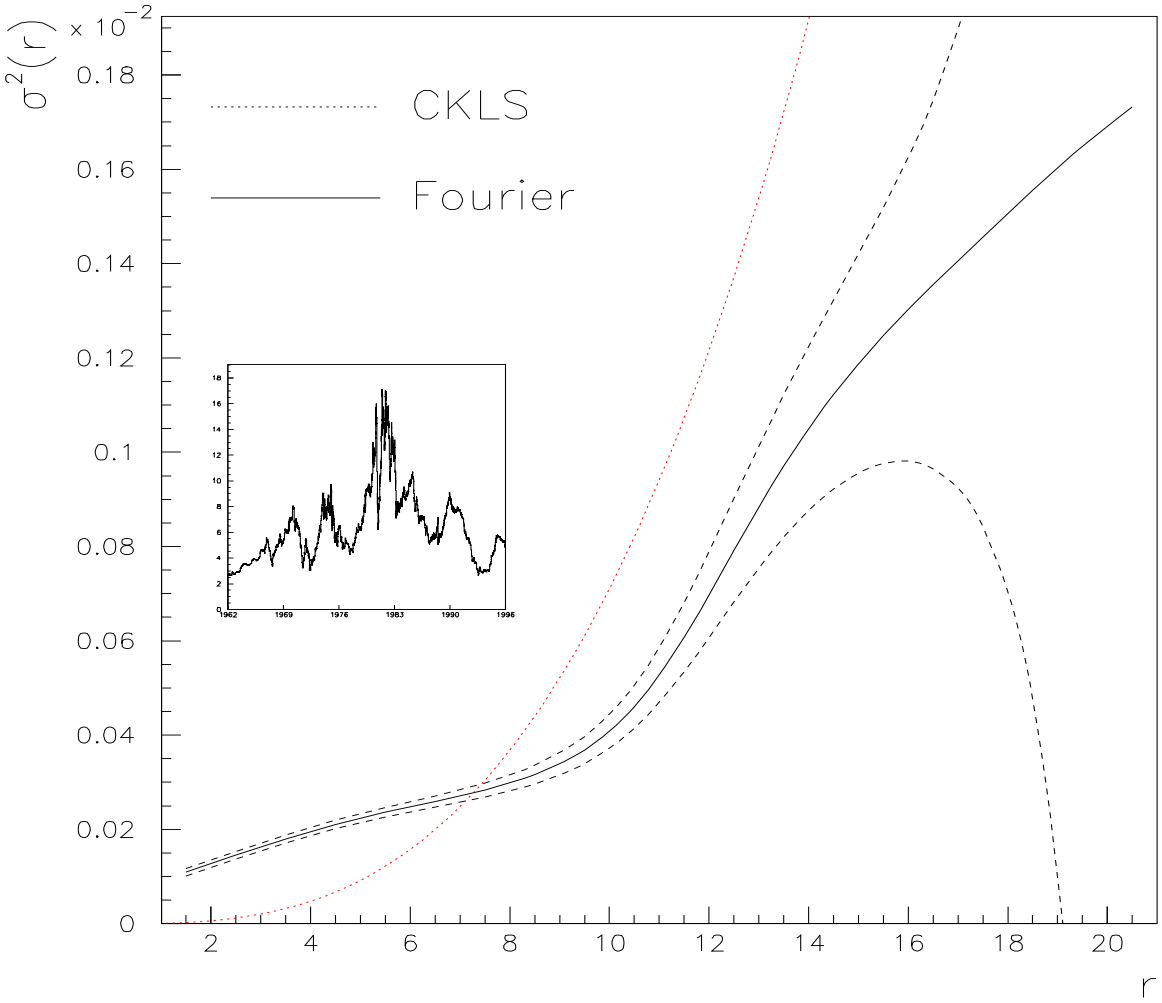

Figure 5.5: Estimate of the diffusion coefficient $\sigma^2(r)$ on Jiang (1998) data set. Solid line: Fourier estimate. Dashed lines: 5% and 95% confidence intervals. Dotted line: estimate obtained with the parametric model (5.41). In the inset, the time series of the yields on three-months T-bill under study, from January 1962 to January 1996.

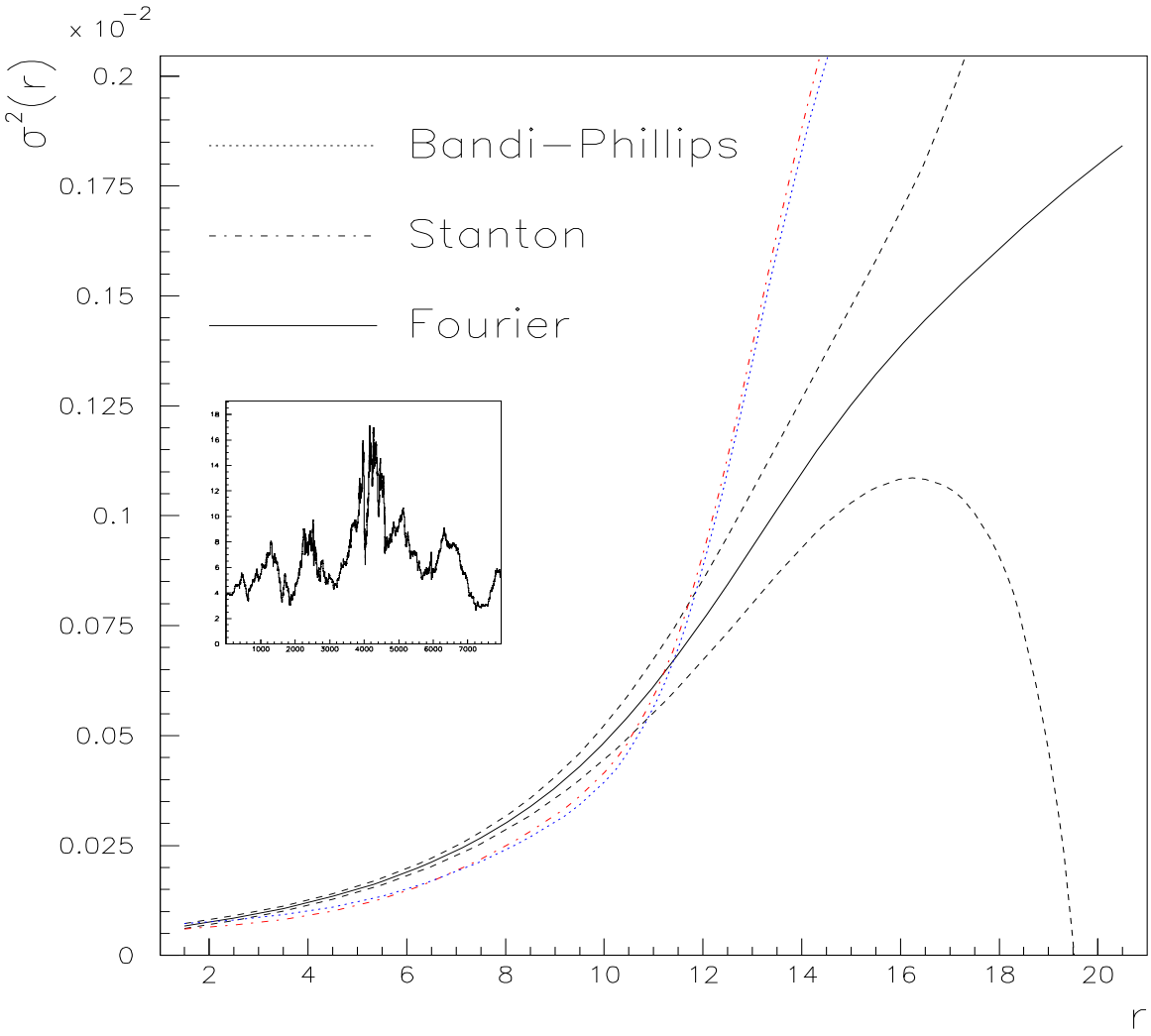

Figure 5.6: Estimate of the diffusion coefficient $\sigma^2(r)$ on Stanton data set. Solid line: Fourier estimate. Dashed lines: 5% and 95% confidence intervals. Dashed-dotted line: Stanton estimate. Dotted line: estimate obtained with the estimator (5.5) of Bandi and Phillips. In the inset, the time series of the yields on three-months T-bill under study, from January 1965 to July 1995.

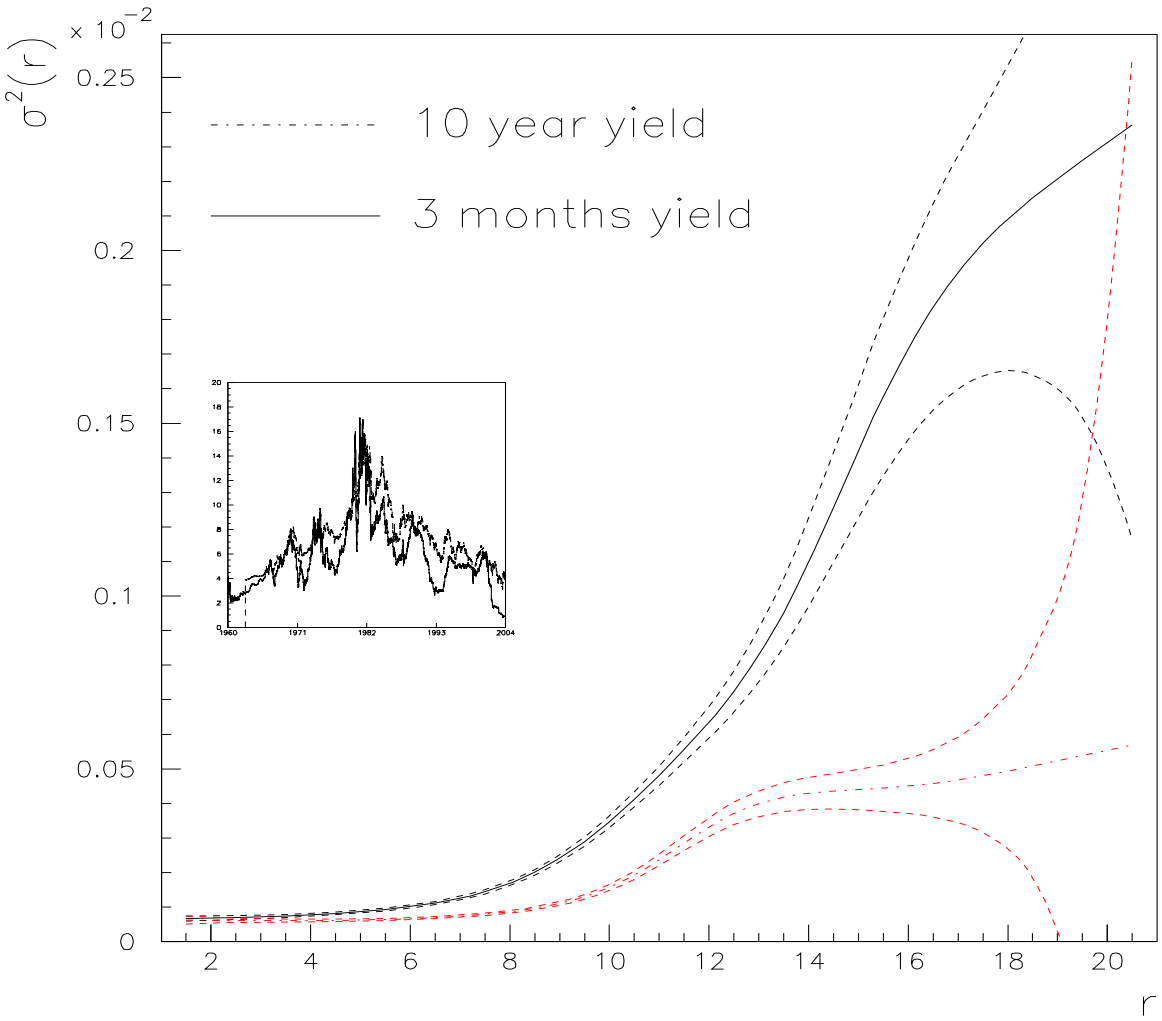

Figure 5.7: Estimate of the diffusion coefficient $\sigma^2(r)$ on the full 1960-2003 data set. Solid line: diffusion coefficient of the yield of the 3-months Treasury Bill. Dashed-dotted line: diffusion coefficient of the yield of the 10-year Treasury Note (from 1962). Dashed lines: 5% and 95% confidence intervals. In the inset, the time series under study.

Bibliography

Ait-Sahalia, Y. (1996a). Nonparametric pricing of interest rate derivative securities. *Econometrica 64*(3), 527–560.

Ait-Sahalia, Y. (1996b). Testing continuous time models of the spot interest rate. *Review of Financial Studies 9*(2), 385–426.

Ait-Sahalia, Y., P. Mykland, and L. Zhang (2003). How often to sample a continuous-time process in the presence of microstructure noise. *Review of Financial Studies*. Forthcoming.

Alizadeh, S., M. Brandt, and F. Diebold (2002). High and low frequency exchange rate volatility dynamics: Range-based estimation of stochastic volatility models. *Journal of Finance 57*, 1047–1092.

Amihud, Y. and H. Mendelson (1987). Trading mechanism and stock retuns: An empirical investigation. *The Journal of Finance 42*(3), 533–553.

Andersen, T. (1994). Stochastic autoregressive volatility: A framework for volatility modeling. *Mathematical Finance 4*, 75–102.

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

Andersen, T. and T. Bollerslev (1998a). Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *International Economic Review 39*, 885–905.

Andersen, T. and T. Bollerslev (1998b). Deutsche mark-dollar volatility: Intraday activity patterns, macroeconomic announcements, and longer run dependencies. *Journal of Finance 53*, 219–265.

Andersen, T. and T. Bollerslev (1998c). Towards a unified framework for high and low frequency return volatility modeling. *Statistica Neerlandica 52*(3), 273–302.

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

Andersen, T., T. Bollerslev, and F. Diebold (2004). Some like it smooth, and some like it rough: untangling continuous and jump components in measuring, modeling, and forecasting asset return volatility. Manuscript.

Andersen, T., T. Bollerslev, F. Diebold, and H. Ebens (2001). The distribution of realized stock volatility. *Journal of Financial Economics 61*, 43–76.

Andersen, T., T. Bollerslev, F. Diebold, and P. Labys (2000a). Exchange rate returns standardized by realized volatility are (nearly) Gaussian. *Multinational Finance Journal 4*, 159–179.

Andersen, T., T. Bollerslev, F. Diebold, and P. Labys (2000b). Great realizations. *Risk 13*, 105–108.

Andersen, T., T. Bollerslev, F. Diebold, and P. Labys (2001). The distribution of realized exchange rate volatility. *Journal of the American Statistical Association 96*, 42–55.

Andersen, T., T. Bollerslev, F. Diebold, and P. Labys (2003). Modeling and forecasting realized volatility. *Econometrica 71*, 579–625.

Andersen, T., T. Bollerslev, and F. X. Diebold (2003). Parametric and nonparametric volatility measurement. In L. P. Hansen and Y. Ait-Sahalia (Eds.), *Handbook of Financial Econometrics*. Amsterdam: North-Holland.

Andersen, T., T. Bollerslev, and S. Lange (1999). Forecasting financial market volatility: Sample frequency vis-à-vis forecast horizon. *Journal of Empirical Finance 6*, 457–477.

Andreou, E. and E. Ghysels (2004). The impact of sampling frequency and volatility estimators on change-point tests. *Journal of Financial Econometrics 2*, 290–318.

Angelini, P. (2000). Are banks risk averse? Intraday timing of operations in the interbank market. *Journal of Money, Credit and Banking 32*, 54–73.

Angelini, P. and P. Silipo (2001). The demand for settlement balances in the Euro area. Manuscript, Bank of Italy.

Areal, N. and S. J. Taylor (2002). The realized volatility of FTSE-100 futures prices. *Journal of Futures Markets 22*, 627–648.

Ball, C. and W. Torous (1984). The maximum likelihood estimation of security price volatility: Theory, evidence, and application to option pricing. *Journal of business 57*, 97–112.

Ball, C. and W. Torous (1996). Unit roots and the estimation of interest rate dynamics. *Journal of Empirical Finance 3*, 215–238.

Ball, C. and W. Torous (2000). Stochastic correlation across international stock markets. *Journal of Empirical Finance 7*, 373–388.

Ballocchi, G., M. Dacorogna, C. Hopman, U. Muller, and R. Olsen (1999). The intraday multivariate structure of the Eurofutures markets. *Journal of Empirical Finance 6*, 479–513.

Bandi, F. (2002). Short-term interest rate dynamics: a spatial approach. *Journal of Financial Economics 65*, 73–110.

Bandi, F. and P. Phillips (2003). Fully nonparametric estimation of scalar diffusion models. *Econometrica 71*(1), 241–283.

Bandi, F. and J. Russell (2003). Microstructure noise, realized volatility and optimal sampling. Working paper.

Barndorff-Nielsen, O. E. and N. Shephard (2002a). Econometric analysis of realised volatility and its use in estimating stochastic volatility models. *Journal of the Royal Statistical Society, Series B 64*, 253–280.

Barndorff-Nielsen, O. E. and N. Shephard (2002b). Estimating quadratic variation using realized variance. *Journal of Applied Econometrics 17*, 457–478.

Barndorff-Nielsen, O. E. and N. Shephard (2004a). Econometric analysis of realised covariation: high frequency based covariance, regression and correlation in financial economics. *Econometrica 72*, 885–925.

Barndorff-Nielsen, O. E. and N. Shephard (2004b). Econometrics of testing for jumps in financial economics using bipower variation. *Journal of Financial Econometrics*. Forthcoming.

Barndorff-Nielsen, O. E. and N. Shephard (2004c). Power and bipower variation with stochastic volatility and jumps. *Journal of Financial Econometrics 2*, 1–48.

Bartolini, L., G. Bertola, and A. Prati (2001). Banks' reserve management, transaction costs and the timing of the federal reserve intervention. *Journal of Banking and Finance 25*(1), 1287–1317.

Bartolini, L., G. Bertola, and A. Prati (2002). Day-to-day monetary policy and the volatility of the federal funds interest rate. *Journal of Money, Credit and Banking 25*(1), 1287–1317.

Barucci, E., C. Impenna, and R. Renò (2003). The Italian overnight market: microstructure effects, the martingale hyphotesis and the payment system. *Temi di Discussione, Bank of Italy*. N. 475.

Barucci, E., C. Impenna, and R. Renò (2004). Interest rate dynamics and market activity in the Italian interbank overnight market. Working paper, Università di Pisa.

Barucci, E., P. Malliavin, M. E. Mancino, R. Renò, and A. Thalmaier (2003). The price-volatility feedback rate: an implementable mathematical indicator of market stability. *Mathematical Finance 13*, 17–35.

Barucci, E., M. Mancino, and R. Renò (2000). Volatility estimate via Fourier analysis. In *Finanza Computazionale, Atti della scuola estiva 2000*, Venezia. Università Ca' Foscari.

Barucci, E. and R. Renò (2002a). On measuring volatility and the GARCH forecasting performance. *Journal of International Financial Markets, Institutions and Money 12*, 183–200.

Barucci, E. and R. Renò (2002b). On measuring volatility of diffusion processes with high frequency data. *Economics Letters 74*, 371–378.

Barucci, E. and R. Renò (2002c). Value at risk with high frequency data. In C. Zopounidis (Ed.), *New Trends in Banking Management*, pp. 223–232. Physica-Verlag.

Bekaert, G. and G. Wu (2000). Asymmetric volatility and risk in equity markets. *Review of Financial Studies 13*(1), 1–42.

Beltratti, A. and C. Morana (1999). Computing value at risk with high frequency data. *Journal of Empirical Finance 6*, 431–455.

Bianco, S. and R. Renò (2005). Dynamics of intraday serial correlation in the Italian futures market. *Journal of Futures Markets*. Forthcoming.

Black, F. (1976). Studies of stock price volatility changes. In *Proceedings if the 1976 meetings of the american statistical association, business and economical statistics section*, pp. 177–181.

Black, F. and M. Scholes (1973). The pricing of options and corporate liabilities. *Journal of Political Economy 81*, 637–659.

Blair, B., S.-H. Poon, and S. Taylor (2001). Forecasting S&P100 volatility: the incremental information content of implied volatilities and high frequency returns. *Journal of Econometrics 7*, 5–26.

Bollen, B. and B. Inder (2002). Estimating daily volatility in financial markets utilizing intraday data. *Journal of Empirical Finance 9*, 551–562.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics 31*, 307–327.

Bollerslev, T., J. Cai, and F. Song (2000). Intraday periodicity, long memory volatility and macroeconomic announcement effects in the US Treasury bond market. *Journal of Empirical Finance 7*, 37–55.

Bollerslev, T., R. Y. Chou, and K. F. Kroner (1992). ARCH modelling in finance: a selective review of the theory and empirical evidence. *Journal of Econometrics 52*, 5–59.

Bollerslev, T. and I. Domowitz (1993). Trading patterns and prices in the interbank foreign exchange market. *Journal of Finance 48*, 1421–1443.

Bollerslev, T., R. Engle, and D. Nelson (1994). Arch models. In R. Engle and D. McFadden (Eds.), *Handbook of Econometrics*, Volume IV. Amsterdam: North-Holland.

Bollerslev, T. and H. Zhou (2002). Estimating stochastic volatility diffusion using conditional moments of integrated volatility. *Journal of Econometrics 109*, 33–65.

Bonanno, G., F. Lillo, and R. Mantegna (2001). High-frequency cross-correlation in a set of stocks. *Quantitative Finance 1*, 1–9.

Brandt, M. and C. Jones (2002). Volatility forecasting with range-based EGARCH models. Manuscript, Duke University.

Brandt, M. W. and F. X. Diebold (2004). A no-arbitrage approach to range-based estimation of return covariances and correlations. *Journal of Business*. Forthcoming.

Brunetti, C. and P. M. Lildholdt (2002a). Return-based and range-based (co)variance estimation - with an application to foreign exchange markets. CAF Working Paper N. 127.

Brunetti, C. and P. M. Lildholdt (2002b). Time series modelling of daily log-price ranges for SF/USD and USD/GBP. CAF Working Paper N. 126.

Burns, P., R. Engle, and J. Mezrich (1998). Correlations and volatilities of asynchronous data. *Journal of Derivatives 5*, 7–18.

Campbell, J. (1987). Money announcements, the demand for bank reserves, and the behavior of the federal funds rate within the statement week. *Journal of Money, Credit and Banking 19*, 56–67.

Campbell, J., M. Lettau, B. Malkiel, and Y. Xu (2001). Have individual stocks become more volatile? An empirical exploration of idiosyncratic risk. *Journal of Finance 56*(1), 1–42.

Chamberlain, G. and M. Rothschild (1983). Arbitrage, factor structure, and mean-variance analysis on large asset markets. *Econometrica 51*(5), 1281–1304.

Chan, K. (1992). A further analysis of the lead-lag relationships between the cash market and the stock index futures market. *Review of financial studies 5*, 123–152.

Chan, K. (1993). Imperfect information and cross-autocorrelation among stock prices. *Journal of Finance 48*(4), 1211–1230.

Chan, K., A. Karolyi, F. Longstaff, and A. Sanders (1992). An empirical comparison of alternative models of the short-term interest rate. *Journal of Finance 47*(3), 1209–1227.

Chapman, D., J. Long, and N. Pearson (1999). Using proxies for the short rate: When are three months like an instant? *Review of Financial Studies 12*(4), 763–806.

Chapman, D. and N. Pearson (2000). Is the short rate drift actually nonlinear? *Journal of Finance 55*(1), 355–388.

Christensen, B. J. and N. R. Prabhala (1998). The relation between implied and realized volatility. *Journal of Financial Economics 37*, 125–150.

Christie, A. (1982). The stochastic behavior of common stock variances: value, leverage and interest rate effects. *Journal of Financial Economics 10*, 407–432.

Christoffersen, P. and F. X. Diebold (2000). How relevant is volatility forecasting for financial risk management? *Review of Economics and Statistics 82*, 1–11.

Corradi, V. (2000). Reconsidering the continuous time limit of the GARCH(1,1) process. *Journal of Econometrics 96*, 145–153.

Corsi, F. (2003). A simple long memory model of realized volatility. Manuscript.

Corsi, F., G. Zumbach, U. Muller, and M. Dacorogna (2001). Consistent high-precision volatility from high-frequency data. *Economic Notes 30*(2), 183–204.

Cox, J., J. Ingersoll, and S. Ross (1985). A theory of the term structure of interest rates. *Econometrica 53*, 385–406.

Cox, J. C. and S. A. Ross (1976). The valuation of options for alternative stochastic processes. *Journal of Financial Economics 3*, 145–166.

Curci, G. and F. Corsi (2003). A discrete sine transform approach for realized volatility measurement. Manuscript, University of Pisa.

Cyree, K. and D. Winters (2001). An intraday examination of the federal funds market: implications for the theories of the reverse-J pattern. *The Journal of Business 74*(4), 535–556.

Da Prato, G. (1998). *Introduction to differential stochastic equations*. Appunti. Scuola Normale Superiore, Pisa.

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

Dai, Q. and K. Singleton (2000). Specification analysis of affine term-structure models. *Journal of Finance 55*(5), 1943–1978.

de Jong, F. and T. Nijman (1997). High frequency analysis of lead-lag relationships between financial markets. *Journal of Empirical Finance 4*, 259–277.

Dellacherie, C. and P. A. Meyer (1976). *Probabilités and Potentiel*. Paris: Hermann.

Deo, R., C. Hurvich, and Y. Lu (2003). Forecasting realised volatility using a long memory stochastic volatility model: Estimation, prediction and seasonal adjustment. Manuscript.

Drost, F. and T. Nijman (1993). Temporal aggregation of GARCH processes. *Econometrica 61*(4), 909–927.

Drost, F. and B. Werker (1996). Closing the GARCH gap: Continuous time GARCH modeling. *Journal of Econometrics 74*, 31–57.

Duan, J.-C. (1997). Augmented GARCH(p,q) process and its diffusion limit. *Journal of Econometrics 79*, 97–127.

Duffee, G. (1995). Stock returns and volatility a firm-level analysis. *Journal of Financial Economics 37*(3), 399–420.

Duffee, G. (1996). Idiosyncratic variation of Treasury Bill yield spread. *Journal of Finance 51*, 527–552.

Duffie, D., J. Pan, and K. Singleton (2002). Transform analysis and asset pricing for affine junp-diffusions. *Econometrica 68*, 1343–1376.

Engle, R. (1982). Autoregressive conditional heteroskedasticity with estimates of the variance of UK inflation. *Econometrica 50*, 987–1008.

Epps, T. (1979). Comovements in stock prices in the very short run. *Journal of the American Statistical Association 74*, 291–298.

Evans, M. and R. Lyons (2002). Order flow and exchange rate dynamics. *Journal of Political Economy 110*, 170–180.

Fama, E. (1965). The behavior of stock market prices. *Journal of Business 38*, 34–105.

Figlewski, S. (1997). Forecasting volatility. *Financial Markets, Institutions and Instruments 6*, 1–88.

Fleming, J. and C. Kirby (2003). A closer look at the relation between GARCH and stochastic autoregressive volatility. *Journal of Financial Econometrics 1*(3), 365–419.

Fleming, J., C. Kirby, and B. Ostdiek (2001). The economic value of volatility timing. *Journal of Finance 56*(1), 329–352.

Fleming, J., C. Kirby, and B. Ostdiek (2003). The economic value of volatility timing using "realized" volatility. *Journal of Financial Economics 67*, 473–509.

Florens-Zmirou, D. (1993). On estimating the diffusion coefficient from discrete observations. *Journal of Applied Probability 30*, 790–804.

French, K. and R. Roll (1986). Stock return variances: the arrival of information and the reaction of traders. *Journal of Financial Economics 17*, 5–26.

French, K., G. W. Schwert, and R. Stambaugh (1987). Expected stock returns and volatility. *Journal of Financial Economics 19*, 3–29.

Furfine, C. (2000). Interbank payments and the daily federal funds rate. *Journal of Monetary Economics 46*, 535–553.

Galbraith, J. and V. Zinde-Walsh (2000). Properties od estimates of daily GARCH parameters based on intra-day observations. Manuscript, McGill University.

Gallant, P., P. Rossi, and G. Tauchen (1993). Nonlinear dynamics structures. *Econometrica 61*, 871–907.

Gallant, R., C.-T. Hsu, and G. Tauchen (1999). Using daily range to calibrate volatility diffusions and extract the forward integrated variance. *Review of Economics and Statistics 81*, 617–631.

Gallant, R. and G. Tauchen (1996). Which moments to match? *Econometric Theory 12*(4), 657–681.

Garman, M. B. and M. J. Klass (1980). On the estimation of security price volatilities from historical data. *Journal of Business 53*, 61–65.

Genon-Catalot, V., C. Laredo, and D. Picard (1992). Non-paramteric estimation of the diffusion coefficient by wavelets methods. *Scandinavian Journal of Statistics 19*, 317–335.

Ghysels, E., A. Harvey, and E. Renault (1996). Stochastic volatility. In G. Maddala (Ed.), *Handbook of Statistics*, Volume 14, pp. 119–191. Amsterdam: North-Holland.

Goodhart, C. and M. O'Hara (1997). High frequency data in financial markets: Issues and applications. *Journal of Empirical Finance 4*, 73–114.

Granger, C. (1980). Long memory relationships and the aggregation of dynamic models. *Journal of Econometrics 14*, 227–238.

Griffiths, M. and D. Winters (1995). Day of the week effects in the federal funds rates: further empirical findings. *Journal of Banking and Finance 19*, 1265–1284.

Guillaume, D., M. Dacorogna, R. Dave, U. Mueller, R. Olsen, and O. Pictet (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.

Hamilton, J. (1996). The daily market for federal funds. *Journal of Political Economy 104*, 26–56.

Hansen, P. and A. Lunde (2004). A comparison of alternative volatility model. Does anything beat a GARCH(1,1) model? *Journal of Applied Econometrics*. Forthcoming.

Harris and Raviv (1993). Difference of opinion make a horse race. *Review of Financial Studies 6*, 473–506.

He, H. and J. Wang (1995). Differential information and dynamic behavior of stock trading volume. *Review of Financial Studies 8*, 919–972.

Hoffmann, M. (1999). L_p estimation of the diffusion coefficient. *Bernoulli 5*(3), 447–481.

Jacod, J. (1999). Non-parametric kernel estimation of the coefficient of a diffusion. *Scandinavian Journal of Statistics 27*, 83–96.

Jacod, J. and A. N. Shiryaev (1987). *Limit Theorems for Stochastic Processes*. Springer.

Jiang, G. and J. Knight (1997). A nonparametric approach to the estimation of diffusion processes, with an application to a short-term interest rate model. *Econometric Theory 13*, 615–645.

Jiang, G. J. (1998). Nonparametric modeling of U.S. interest rate term structure dynamics and implications on the prices of derivative securities. *Journal of Financial and Quantitative Analysis 33*(4), 465–497.

Jones, C. (2003). Nonlinear mean reversion in the short-term interest rate. *Review of Financial Studies 16*, 765–791.

Jones, C., G. Kaul, and M. Lipson (1994). Transactions, volume and volatility. *Review of Financial Studies 7*, 631–651.

Jorion, P. (1995). Predicting volatility in the foreign exchange market. *Journal of Finance 50*, 507–528.

Kandel, S. and N. Pearson (1995). Differential interpretation of public signals and trade in speculative markets. *Journal of Political Economy 103*, 831–872.

Karatzas, I. and E. Shreve (1988). *Brownian Motion and Stochastic Calculus*. Springer-Verlag.

Karpoff, J. (1987). The relation between price changes and trading volume: a survey. *Journal of Financial and Quantitative Analysis 22*, 109–126.

Kim, O. and R. Verrecchia (1991). Market reaction to anticipated anouncements. *Journal of Financial Economics 30*, 273–309.

Kloeden, P. and E. Platen (1992). *Numerical Solutions of Stochastic Differential Equations*. Springer.

Knight, F. B. (1971). A reduction of continuous square integrable martingale to brownian motion. In *Lecture Notes in Mathematics*, Volume 190, pp. 19–31. Springer-Verlag.

Kroner, K. and V. Ng (1998). Modeling asymmetric comovements of asset returns. *Review of Financial Studies 11*(4), 817–844.

Laloux, L., P. Cizeau, J.-P. Bouchaud, and M. Potters (1999). Noise dressing of financial correlation matrices. *Physical Review Letters 83*(7), 1467–1470.

Levi-Civita, T. (1925). *Lezioni di calcolo differenziale ed assoluto*. Roma: Stock Editor.

Litterman, R. and J. Scheinkman (1991). Common factors affecting bond returns. *Journal of Fixed Income 1*(1), 54–61.

Lo, A. W. and A. C. MacKinlay (1990). An econometric analysis of nonsynchronous trading. *Journal of Econometrics 45*, 181–211.

Lockwood, L. and S. Linn (1990). An examination of stock market return volatility during overnight and intraday periods, 1964-1989. *The Journal of Finance 45*(2), 591–601.

Lundin, M., M. Dacorogna, and U. Muller (1999). Correlation of high-frequency financial time series. In P. Lequeux (Ed.), *Financial Markets Tick by Tick*. Wiley & Sons.

Madhavan, A. (2000). Market microstructure: A survey. *Journal of Financial Markets 3*(3), 205–258.

Maheu, J. and T. McCurdy (2002). Nonlinear features of realized FX volatility. *Review of Economics and Statistics 84*(3), 345–372.

Maheu, J. and T. McCurdy (2004). News arrival, jump dynamics and volatility components for individual stock returns. *Journal of Finance*. Forthcoming.

Malliavin, P. and M. Mancino (2002). Fourier series method for measurement of multivariate volatilities. *Finance & Stochastics 6*(1), 49–61.

Malliavin, P. and M. Mancino (2005). Harmonic analysis methods for nonparametric estimation of volatility. Working paper, Università di Firenze.

Mancino, M. and R. Renò (2002). Dynamic principal component analysis of multivariate volatility via Fourier analysis. *Applied Mathematical Finance*. Forthcoming.

Mandelbrot, B. (1963). The variation of certain speculative prices. *Journal of Business 36*, 349–419.

Martens, M. (2001). Forecasting daily exchange rate volatility using intraday returns. *Journal of International Money and Finance 20*(1), 1–23.

Martens, M. (2002). Measuring and forecasting S&P500 index futures volatility using high-frequency data. *Journal of Futures Markets 22*, 497–518.

Martens, M. and S. H. Poon (2001). Returns synchronization and daily correlation dynamics between international stock markets. *Journal of Banking and Finance 25*(10), 1805–1827.

Meddahi, N. (2002). A theoretical comparison between integrated and realized volatility. *Journal of Applied Econometrics 17*, 479–508.

Meddahi, N. and E. Renault (2004). Temporal aggregation of volatility models. *Journal of Econometrics*. Forthcoming.

Merton, R. (1973). Theory of rational option pricing. *Bell Journal of Economics and Management Science 4*(1), 141–183.

Merton, R. (1980). On estimating the expected return on the market: an exploratory investigation. *Journal of Financial Economics 8*, 323–361.

Muller, U., M. Dacorogna, R. Davé, R. Olsen, O. Pictet, and J. von Weizsacker (1997). Volatilities of different time resolutions - analyzing the dynamics of market components. *Journal of Empirical Finance 4*, 213–239.

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

Muthuswamy, J., S. Sarkar, A. Low, and E. Terry (2001). Time variation in the correlation structure of exchange rates: high frequency analysis. *Journal of Futures Markets 21*(2), 127–144.

Nelson, D. (1991). Conditional heteroskedasticity in asset returns: a new approach. *Econometrica 59*, 347–370.

Nelson, D. (1992). Filtering and forecasting with misspecified ARCH models I: Getting the right variance with the wrong model. *Journal of Econometrics 52*, 61–90.

Oomen, R. (2002). Modelling realized variance when returns are serially correlated. Manuscript, Wharton Business School.

Pagan, A. and G. W. Schwert (1990). Alternative models for conditional stock volatility. *Journal of Econometrics 45*, 267–290.

Pan, J. (2002). The jump-risk premia implicit in options: Evidence from an integrated time series study. *Journal of Financial Economics 63*, 3–50.

Parkinson, M. (1980). The extreme value method for estimating the variance of the rate of return. *Journal of Business 53*(1), 61–65.

Pasquale, M. and R. Renò (2005). Statistical properties of trading volume depending on size. *Physica A 346*, 518–528.

Plerou, V., P. Gopikrishnan, B. Rosenow, L. Nunes Amaral, and E. Stanley (1999). Universal and nonuniversal properties of cross correlations in financial time series. *Physical Review Letters 83*(7), 1471–1474.

Poon, S.-H. and C. Granger (2003). Forecasting volatility in financial markets: A review. *Journal of Economic Literature 41*(2), 478–539.

Poterba, J. and L. Summers (1986). The persistence of volatility and stock market fluctuation. *American Economic Review 76*, 1142–1151.

Prati, A., L. Bartolini, and G. Bertola (2003). The overnight interbank market: Lessons from the G7. *Journal of Banking and Finance 27*(10), 2045–2083.

Priestley, M. (1979). *Spectral Time Series Analysis*. Wiley.

Pritsker, M. (1998). Nonparametric density estimation and tests of continuous time interest rate models. *Review of Financial Studies 11*(3), 449–487.

Protter, P. (1990). *Stochastic Integration and Differential Equations*. Springer-Verlag.

Quiros, G. and H. Mendizabal (2000). The daily market for funds in europ: has something changed with EMU? European Central Bank working paper series, n. 67.

Renò, R. (2003). A closer look at the Epps effect. *International Journal of Theoretical and Applied Finance 6*(1), 87–102.

Renò, R. (2004). Nonparametric estimation of the diffusion coefficient via Fourier analysis, with an application to short interest rates. Working Paper, Università di Siena.

Renò, R. and R. Rizza (2003). Is volatility lognormal? Evidence from Italian futures. *Physica A 322*, 620–628.

Renò, R., A. Roma, and S. Schaefer (2004). A comparison of alternative nonparametric estimators of the diffusion coefficient. Working Paper.

Revuz, D. and M. Yor (1998). *Continuous Martingales and Brownian Motion*. Springer-Verlag.

Rogers, C. and S. Satchell (1991). Estimating variance from high, low and closing prices. *Annals of Applied Probability 1*(4), 504–512.

Ross, S. A. (1976). The arbitrage theory of capital asset pricing. *Journal of Economic Theory 13*, 341–360.

Scherer, K. P. and M. Avellaneda (2002). All for one... one for all? A principal component analysis of Latin American brady bond debt from 1994 to 2000. *International Journal of Theoretical and Applied Finance 5*(1), 79–106.

Scholes, M. and J. Williams (1977). Estimating betas from non-synchronous data. *Journal of Financial Economics 5*, 309–327.

Schwert, G. W. (1989). Why does stock market volatility change over time? *Journal of Finance 44*(5), 1115–1153.

Schwert, G. W. (1990). Stock market volatility. *Financial Analyst Journal 46*, 23–34.

Schwert, G. W. (1998). Stock market volatility: ten years after the crash. *Brooking-Wharton Papers on Financial Services 1*, 65–114.

Schwert, G. W. and P. Seguin (1990). Heteroskedasticity in stock returns. *Journal of Finance 45*, 1129–1155.

Scott, D. (1992). *Multivariate Density Estimation*. John Wiley & Sons.

Spindt, P. and R. Hoffmeister (1988). The micromechanics of the federal funds market: implications for the day of the week effects in funds rate variability. *Journal of Financial and Quantitative Analysis 23*(4), 401–416.

Stanton, R. (1997). A nonparametric model of term structure dynamics and the market price of interest rate risk. *Journal of Finance 52*, 1973–2002.

Taylor, J. (2000). Expectations, open market operations and changes in the federal funds rate. Manuscript, Stanford University.

Taylor, S. J. and X. Xu (1997). The incremental volatility information in one million foreign exchange quotations. *Journal of Empirical Finance 4*, 317–340.

Thomakos, D. and T. Wang (2003). Realized volatility in the futures market. *Journal of Empirical Finance 10*, 321–353.

Thomakos, D., T. Wang, and L. Wille (2002). Modeling daily realized futures volatility using singular spectrum analysis. *Physica A 312*, 505–519.

Vasicek, O. (1977). An equilibrium characterization of the term structure. *Journal of Financial Economics 5*, 177–188.

Wang, J. (1994). A model of competitive stock trading volume. *Journal of Political Economy 102*, 127–168.

Wu, G. (2001). The determinants of asymmetric volatility. *Review of Financial Studies 14*(3), 837–859.

Yang, D. and Q. Zhang (2000). Drift-independent volatility estimation based on high, low, open and close price. *Journal of Business 73*(3), 477–491.

Zebedee, A. (2001). A closer look at co-movements among stock returns. San Diego State University.

Zhou, B. (1996). High-frequency data and volatility in foreign-exchange rates. *Journal of Business and Economics Statistics 14*(1), 45–52.

Zumbach, G. (2000). The pitfalls in fitting GARCH processes. In C. Dunis (Ed.), *Advances in Quantitative Asset Management*. Kluver Academic Publisher.