

Discussion Papers No. 614, March 2010
Statistics Norway, Research Department

Arvid Raknerud and Øivind Skare

Multivariate stochastic volatility models based on non-Gaussian Ornstein-Uhlenbeck processes: A quasi-likelihood approach

Abstract:

This paper extends the ordinary quasi-likelihood estimator for stochastic volatility models based on non-Gaussian Ornstein-Uhlenbeck (OU) processes to vector processes. Despite the fact that multivariate modeling of asset returns is essential for portfolio optimization and risk management -- major areas of financial analysis -- the literature on multivariate modeling of asset prices in continuous time is sparse, both with regard to theoretical and applied results. This paper uses non-Gaussian OU-processes as building blocks for multivariate models for high frequency financial data. The OU framework allows exact discrete time transition equations that can be represented on a linear state space form. We show that a computationally feasible quasi-likelihood function can be constructed by means of the Kalman filter also in the case of high-dimensional vector processes. The framework is applied to Euro/NOK and US Dollar/NOK exchange rate data for the period 2.1.1989-4.2.2010.

Keywords: multivariate stochastic volatility, exchange rates, Ornstein-Uhlenbeck processes, quasi-likelihood, factor models, state space representation

JEL classification: C13, C22, C51, G10

Acknowledgement: We appreciate useful comments from Torbjørn Hægeland, Terje Skjerpen and Anders Rygh Swensen. Accompanying software written in C++ code (with R-interface) can be downloaded from <http://folk.uio.no/skare/SV/>. Financial support from the Norwegian Research Council ("Finansmarkedsfondet") is gratefully acknowledged.

Address: Arvid Raknerud, Statistics Norway, Research Department. E-mail: rak @ssb.no

Øivind Skare, Norwegian Institute of Public Health and University of Bergen, Department of Public Health and Primary Health Care. E-mail: oivind.skare@medisin.uio.no

Discussion Papers

comprise research papers intended for international journals or books. A preprint of a Discussion Paper may be longer and more elaborate than a standard journal article, as it may include intermediate calculations and background material etc.

Abstracts with downloadable Discussion Papers
in PDF are available on the Internet:

<http://www.ssb.no>

<http://ideas.repec.org/s/ssb/dispap.html>

For printed Discussion Papers contact:

Statistics Norway
Sales- and subscription service
NO-2225 Kongsvinger

Telephone: +47 62 88 55 00

Telefax: +47 62 88 55 95

E-mail: Salg-abonnement@ssb.no

1 Introduction

Several models for high frequency financial asset price data have been put forward during the last decade. Overviews of recent developments and some of the related methodological background are given in Barndorff-Nielsen et al. (2001), Harvey et al. (2004), Shephard (2005) and Andersen et al. (2009). This paper builds on the class of models proposed and analyzed in Barndorff-Nielsen and Shephard (2001, 2002, 2003) (hereafter BS), where volatility (stochastic variance) is modeled as positive Ornstein-Uhlenbeck processes driven by Levy jump processes. A main benefit of this class of *continuous time* models is that many closed form solutions can be derived, such as *exact* discrete time transition equations, that are useful both for fitting empirical models (using data collected in discrete time), and for applications in financial engineering, e.g. derivative pricing. However, when it comes to joint modeling of several asset prices in continuous time, the literature is sparse, both with regard to theoretical and applied results. Nevertheless, multivariate modeling is highly relevant for portfolio optimization and risk management, which are major areas of financial analysis. For example, to measure the value at risk, quantiles in the returns distribution of a portfolio consisting of a large number of assets and derivatives must be estimated, obviously requiring realistic and analytically tractable *multivariate* models. To address these issues, we here advocate using non-Gaussian Ornstein-Uhlenbeck (OU) processes as building blocks for multivariate stochastic volatility models. We implement a quasi-likelihood estimator for this class of models and also provide software written as a user friendly R-package that interfaces efficient C++ code.¹

Multivariate *discrete time* models have been available in the econometric literature for some time. Most of these models build on the ARCH/GARCH tradition; see Bollerslev et al. (1988), Diebold and Nerlove (1989), Engle et al. (1990) and King et al. (1994). In GARCH-type models the conditional variance of financial returns is

¹See <http://folk.uio.no/skare/SV/> for software and user documentation (“How to get started”).

modeled as a deterministic function of past returns. Recent surveys of multivariate GARCH models is given by Bauwens et al. (2006) and Silvennoinen and Teräsvirta (2009). There exists a parallel literature on multivariate discrete time stochastic volatility (MSV) models, where the conditional variance is modeled as a dynamic latent variable. See e.g. the surveys in Asai et al. (2006) and Chib et al. (2009). MSV models combine features of the classical univariate stochastic volatility model (see e.g. Taylor, 1986) with classical factor analysis: The co-movements in the volatility among several return series is captured by a few common dynamic factors. Stochastic volatility based models are more flexible than GARCH-type models and many empirical studies have confirmed that they give a better fit to high frequency asset returns data. The MSV literature goes back to Harvey et al. (1994), and have been extended, among others, by Aguilar and West (2000) and Chib et al. (2006). A main problem associated with many of these models is that estimation is complex, requiring simulation based (MCMC) methods (see Yu and Meyer, 2006, and Johannes and Polson, 2009, for overviews). The MCMC approach is cumbersome for large data sets and also relies on specific distributional assumptions, including prior distributions over fixed parameters. The latter feature may be unattractive to non-Bayesians.

In contrast to the GARCH and MSV literature, this paper builds on the tradition of continuous time models built out of Brownian motions (the Black-Scholes-Merton model), which is the cornerstone of modern mathematical finance. The class of analytically tractable continuous time models was substantially extended by BS (2001), who consider the variance of asset returns as determined by a (continuous time) OU process driven by Levy jump processes. Since asset returns data are not registered in continuous time, the observable counterpart of a continuous time model is typically a time series of integrated (or aggregate) returns over fixed time intervals with length Δ . The models we will consider have integrated returns that are uncorrelated, yet not independent. In particular, squared returns have autocorrelation functions that are exponentially decaying. Moreover, by adding independent OU processes with different parameters, great flexibility is achieved, leading to autocorrelation functions of squared returns that are weighted averages of exponential functions. Since

the distribution of aggregate returns is derived by integrating the underlying continuous time processes over fixed intervals, the parameters of the models are trivially invariant to the choice of interval width, Δ . This property is highly desirable and does not apply to discrete time GARCH or to stochastic volatility models (see Drost and Nijman, 1993).

In the univariate case, the statistical properties of OU processes as well as implications for derivative pricing have been examined by BS (2001), Nicolato and Venardos (2003), Benth et al. (2007), and others. However, to the best of our knowledge, none have applied these models to vector processes. Almost all issues regarding practical implementations and estimation in the multivariate case are therefore unresolved. This paper uses univariate Levy OU processes as building blocks for multivariate models. A parsimonious parametrization, which is necessary to avoid the curse of dimensionality when going from univariate to multivariate high frequency time series, is achieved by imposing a factor structure, similar to what is done in the discrete time MSV models mentioned above. We model the individual variables (each asset returns series) as univariate processes, augmented by common dynamic factors that account for the co-movements among them. These co-movements are assumed to be represented by a small, or moderate, number of independent, common shocks, which are increments of a standard Brownian motion multiplied by a stochastic scale coefficient. The stochastic, time-varying scale coefficient accounts for stochastic volatility. The common shocks enter one or more series at the same time, but with a different weight (“loading coefficient”). A flexible stochastic volatility structure is obtained by assuming that the stochastic scale coefficients are generated by sums of independent non-Gaussian OU processes.

As shown by BS (2001), the OU framework allows exact discrete time transition equations that can be represented on a linear state space form with the observation vector consisting of returns and squared returns. We extend these results to the multivariate case. A Gaussian quasi-likelihood function is derived by treating the one-step-ahead predictions of the returns and squared returns as if they were conditional expectations. Although the error terms of the state space representation are not Gaussian, they are white noise and strong mixing (see BS, 2001). Thus the

Kalman filter provides optimal linear predictors in a mean squared error sense (see Brockwell and Davis, 1993, Ch. 12). Moreover, the asymptotic theory of Dunsmuir (1979) applies, implying that the quasi-likelihood estimator based on the Gaussian approximation to the model provides consistent estimators.

The applied part of this paper considers a bivariate model for the Euro/NOK and US Dollar/NOK exchange rates using 5332 daily returns data for the period 2.1.1989-4.2.2010. While there exists a large literature on exchange rate dynamics, the well-known study of Meese and Rogoff (1983) demonstrate for a wide range of exchange rate models based on economic fundamentals that neither of these were able to outperform a simple random walk model in forecasting competitions. Modern econometric contributions to this research area focus mainly on the modeling of volatility. Another line of empirical research focus on market microstructure (see Lyons, 2001). Some influential contributions to the study of multivariate exchange rate dynamics during the last decade are Andersen et al. (2001), Barndorff-Nielsen and Shephard (2004), Jungbacker and Koopman (2006) and Chib et al. (2006).

The rest of this paper is organized as follows: Section 2 presents the formal modeling framework, Section 3 establishes the quasi-likelihood function and its derivatives, based on an approximative Gaussian state space model, Section 4 contains the empirical application, while Section 5 concludes.

2 Levy driven OU processes

Stochastic volatility models based on OU processes: Diffusion based models for asset returns with stochastic volatility usually takes as their starting point the stochastic differential equation

$$dy^*(t) = \mu dt + \sigma(t)dw(t), \quad (1)$$

where μ is the drift term, $w(t)$ is a standard Brownian motion and $\sigma^2(t)$ is a strictly positive stochastic process, usually called *spot volatility* in the econometric literature. To understand the implications of (1) for aggregate returns over fixed intervals of length Δ , set $t_n = n\Delta$, for some $\Delta > 0$ and $n = 1, 2, \dots, N$. Then the increments $w_n \equiv w(n\Delta) - w((n-1)\Delta)$ are identically and independently distributed with mean

0 and variance Δ . Moreover,

$$y_n \equiv y^*(n\Delta) - y^*((n-1)\Delta),$$

i.e., integrated returns over the interval $[(n-1)\Delta, n\Delta]$, can be written

$$y_n = \mu\Delta + u_{1n}, \quad (2)$$

where

$$u_{1n} = \int_{(n-1)\Delta}^{n\Delta} \sigma(u) dW(u).$$

It is well known that

$$u_{1n} \stackrel{D}{=} \sigma_n \varepsilon_n, \quad (3)$$

with

$$\varepsilon_n \sim \mathcal{IN}(0, 1) \text{ and } \sigma_n^2 \equiv \int_{(n-1)\Delta}^{n\Delta} \sigma^2(t) dt.$$

Thus u_{1n} is a white noise error term with a mixed Gaussian distribution and mixing parameter σ_n . In the econometric literature, σ_n^2 is called *actual volatility*. In the classical Black-Scholes-Merton model, $\sigma^2(t) = \sigma^2$. Hence $\sigma_n^2 = \sigma^2\Delta$ and there is no mixing. However, there is overwhelming evidence that this assumption leads to a poor fit to financial returns data over small to medium time intervals (see e.g. Jondeau et al., 2007, for an overview). Real time transaction data exhibit serious departure from normality and homoscedasticity: When Δ is small or moderate (corresponding to minutes, hours or days), the changes y_n are heavily tailed and squared changes, y_n^2 , are serially correlated ("volatility clustering"). Sometimes the distribution of y_n is also significantly skewed. Skewness is not considered here, but can be incorporated into (1) through so-called leverage effects, i.e., a negative correlation between returns and changes in actual volatility. As Δ increases, a central limit theorem seems to be at work, so that if $E(\sigma^2(t)) = \xi$, then

$$t^{-1/2}(y^*(t) - \mu t) \xrightarrow{D} \mathcal{N}(0, \xi). \quad (4)$$

In BS (2001) it is assumed that $\sigma^2(t)$ is a positive Ornstein-Uhlenbeck (OU) process:

$$d\sigma^2(t) = -\lambda\sigma^2(t)dt + dz(\lambda t), \quad z(0) = 0, \quad (5)$$

where $z(t)$ is a Levy jump process, i.e., a process with positive, stationary and independent increments. Some important features characterize this process:

First, $\sigma^2(t)$ moves up only by jumps in $z(t)$, and then tails off exponentially at the rate λ . Thus λ can be interpreted as a discount- or decay-parameter, which determines the memory of the process: a small λ implies a long-memory process $\sigma^2(t)$, while a large λ means that past jumps are quickly discounted. The parameter λ also determines the rate at which jumps in $z(\lambda t)$ occurs: A small λ yields a process with small, infrequent jumps that are slowly discounted. On the other hand, a high λ yields a process with frequent and large jumps, which are quickly discounted. Two simulated actual volatility series σ_n^2 (with $\lambda = 0.015$ and 0.6) are depicted in Figure 1 to illustrate these features. Second, $\sigma^2(t)$ has a stationary distribution which does not depend on λ – the latter result is obtained by the peculiar timing $z(\lambda t)$. Let $E(\sigma^2(t)) \equiv \xi$ and $Var(\sigma^2(t)) \equiv \omega^2$, then it is shown in BS (2001) that

$$\sigma^2(n\Delta) = e^{-\lambda\Delta}\sigma^2((n-1)\Delta) + \eta_n, \quad (6)$$

with

$$\eta_n \sim i.i.d. (\xi(1 - e^{(-\lambda\Delta)}), \omega^2(1 - e^{(-2\lambda\Delta)})).$$

Thus (6) can be interpreted as a continuous time autoregressive model, where $exp(-\lambda\Delta)$ is the autoregressive parameter in the corresponding (exact) discrete-time transition equation for $\sigma^2(n\Delta)$.

An advantage of this OU-based stochastic volatility model, compared both to more traditional discrete-time approaches and the (continuous time) constant elasticity of variance process advocated by Meddahi and Renault (2004), is that they generate many closed form solutions under temporal aggregation: BS (2001) characterize the marginal distribution of integrated (actual) volatility, σ_n^2 , and integrated price changes, y_n , not only their dynamic structure (see also BS, 2003). For example, using (5), it is easy to show that

$$\begin{aligned} E(\sigma_n^2) &= \xi\Delta \\ Var(\sigma_n^2) &= 2\omega^2\lambda^{-2}\{e^{-\lambda\Delta} - 1 + \lambda\Delta\}. \end{aligned} \quad (7)$$

Because estimation of the model at different time frequencies is just a matter of choosing a different Δ , the parameters of the model are (trivially) invariant under

temporal aggregation. In contrast, if we formulate a GARCH model for a given time frequency (e.g. daily) and then decide to estimate the model on another frequency (e.g. weekly), the latter model is no longer a GARCH model; GARCH processes are generally not closed under aggregation (see Drost and Nijman, 1993). The availability of closed form results regarding the distribution of integrated volatility has huge consequences for derivative pricing and risk analysis, as shown in Nicolato and Venardos (2003). Algorithms for exact simulation from some non-Gaussian OU processes are also available (see Zhang and Zhang, 2008).

The family of marginal distributions for $\sigma^2(t)$ which is consistent with the OU assumption, is the self-decomposable distributions on R_+ . In general, a random variable x (not necessarily restricted to R_+) is self-decomposable if for any $c \in (0, 1)$, there exists a random variable x_c , independent of x , such that

$$x \stackrel{\mathcal{D}}{=} cx + x_c.$$

The close relation between OU processes and self-decomposable distributions is not surprising in view of (6). An important class of self-decomposable distribution on R_+ , and hence candidate distributions for $\sigma^2(t)$, is the generalized inverse Gaussian distribution, which contains the inverse Gaussian, inverse χ^2 and Gamma distribution as special cases (see BS, 2001). Another example is the log normal distribution (see BS, 2003 and Bondesson, 2002), which has been advocated by Andersen et al. (2001) to model exchange rate volatility. In this paper, however, the particular choice of parametric family does not play any role, as we here only utilize the first- and second-order properties of integrated volatility and returns.

Multivariate extensions Multivariate diffusion based models with stochastic volatility have not been much studied, despite their obvious usefulness in financial applications. A natural generalization of (1), briefly discussed in BS (2001), is the vector OU process,

$$dy^*(t) = \mu dt + \Sigma(t)^{1/2} dw(t), \quad (8)$$

where $y^*(t) = [y_1^*(t), \dots, y_q^*(t)]'$ is a vector of q log-prices, $\mu = [\mu_1, \dots, \mu_q]'$, $\Sigma(t)$ is a $q \times q$ time-varying stochastic spot covariance matrix and $w(t) = [w_1(t), \dots, w_q(t)]'$ is

a vector of q independent standard Brownian motions. The main challenge when going from the univariate to the multivariate case is to specify a tractable model for $\Sigma(t)$. Our approach consists in specifying $\Sigma(t)$ indirectly through a particular factor structure. That is, for $i = 1, \dots, q$, we assume that

$$dy_i^*(t) = \mu_i dt + \sigma_i(t)dw_i(t) + \sum_{j=1}^p \phi_{ij}\sigma_{q+j}(t)dB_j(t); \quad p \leq q-1, \quad (9)$$

where $w_i(t)$ ($i = 1, \dots, q$) and $B_j(t)$ ($j = 1, \dots, p$) are mutually independent standard Brownian motions, whereas $\sigma_1^2(t), \dots, \sigma_{p+q}^2(t)$ are $p+q$ mutually independent Levy OU processes. The stochastic differential equation (9) implies a particular structure for the spot covariance matrix $\Sigma(t)$ in (8):

$$\Sigma(t) = \text{diag}(\sigma_1^2(t), \dots, \sigma_q^2(t)) + \Phi \text{diag}(\sigma_{q+1}^2(t), \dots, \sigma_{q+p}^2(t))\Phi',$$

where $\text{diag}(a_1, \dots, a_m)$ denotes the diagonal matrix with diagonal elements a_1, \dots, a_m , and

$$\Phi = \begin{bmatrix} \phi_{11} & \cdots & \phi_{1p} \\ \phi_{21} & \cdots & \phi_{2p} \\ \vdots & \vdots & \vdots \\ \phi_{q1} & \cdots & \phi_{qp} \end{bmatrix}, \quad \text{with } \phi_{ii} = 1 \text{ for } i = 1, \dots, q \text{ and } \phi_{ij} = 0 \text{ for } j > i.$$

The restrictions $\phi_{ij} = 0$ for $j > i$ are standard in factor models, while $\phi_{ii} = 1$ could be replaced by $\sum_{i=1}^q \phi_{ij}^2 = 1$; $j = 1, \dots, p$, implying that $\Phi\Phi'$ can be interpreted as a constant correlation matrix of the q innovations $\sum_{j=1}^p \phi_{ij}\sigma_{q+j}(t)dB_j(t)$, $i = 1, \dots, q$. Of course, other observationally equivalent parameter restrictions exist. It is easily seen that if $p = q-1$, a completely unrestricted spot covariance matrix $\Sigma(t)$ is allowed by this model specification, thus generalizing the constant conditional correlation matrix proposed by Bollerslev (1990) in the framework of multivariate GARCH and Harvey et al. (1994) in the context of discrete time MSV modeling. Clearly, when p is large, the model is not tractable due to the curse of dimensionality. But if the comovements across the q returns series, $dy_i^*(t)$ can be represented by a few common factors, $\sigma_{q+1}(t)dB_1(t), \dots, \sigma_{q+p}(t)dB_p(t)$, i.e., p is small, a computationally tractable model can be built, as we will demonstrate below. Note that because of the volatility factor $\sigma_i^2(t)$ which is specific to series i ($i = 1, \dots, q$), no linear combination of asset returns will have constant variance. Neither will any two linear combinations of asset

returns have constant correlation. These features of our model are highly desirable in practical modeling and represent improvements relative to most additive factor model hitherto considered in the literature; for instance by Aguilar and West (2000).

Analogously to actual volatility, σ_n^2 , we define the *actual covariance* as

$$\Sigma_n = \int_{(n-1)\Delta}^{n\Delta} \Sigma(t) dt. \quad (10)$$

Moreover, define

$$\sigma_{jn}^2 = \int_{(n-1)\Delta}^{n\Delta} \sigma_j^2(t) dt; \quad j = 1, \dots, q + p$$

and

$$y_{in} = \int_{(n-1)\Delta}^{n\Delta} dy_i^*(t); \quad i = 1, \dots, q.$$

Then, similarly to (2)-(3), we can write

$$y_{in} = \mu_i \Delta + u_{i1n}; \quad i = 1, \dots, q, \quad (11)$$

with

$$u_{i1n} \stackrel{D}{=} \sigma_{in} \varepsilon_{in} + \sum_{j=1}^p \phi_{ij} \sigma_{q+j,n} \eta_{jn}, \quad (12)$$

where $\varepsilon_{in} \sim \mathcal{IN}(0, 1)$, $\eta_{jn} \sim \mathcal{IN}(0, 1)$ and ε_{in} and η_{jm} are mutually independent for all i, j, n and m . Thus η_{jn} , $j = 1, \dots, p$, are common factors, while ε_{in} is an idiosyncratic error terms specific to series i ($i = 1, \dots, q$). Note that all the additive terms in (12) are mutually uncorrelated.

The main feature that separates (11)-(12) from a standard dynamic factor model is the presence of the time-varying, stochastic scale coefficients, $\sigma_{q+j,n}$. Our model does have similarities with Chib et al. (2006). An important difference, however, is that our approach leads to a computationally simple quasi-likelihood function based on a linear state space representation of the model. In contrast, Chib et al. (2006) use complicated MCMC methods in a Bayesian context to fit their model.

We will in Section 3 present a state space form of the model (11) in terms of the observation vector $Y_n = [y_{1n}, y_{1n}^2, \dots, y_{qn}, y_{qn}^2]'$, thereby extending the univariate case ($q = 1$) analyzed in BS (2001) and Raknerud and Skare (2009). To be able to do so, we first observe that, for $i = 1, \dots, q$,

$$y_{in}^2 = \mu_i^2 \Delta^2 + \sigma_{in}^2 + \sum_{j=1}^p \phi_{ij}^2 \sigma_{q+j,n}^2 + u_{i2n}, \quad (13)$$

where

$$\begin{aligned}
u_{i2n} &\stackrel{D}{=} \sigma_{in}^2 (\varepsilon_{in}^2 - 1) + \sum_{j=1}^p \phi_{ij}^2 \sigma_{q+j,n}^2 (\eta_{jn}^2 - 1) + 2\mu_i \Delta \sigma_{in} \varepsilon_{in} + 2\mu_i \Delta \sum_{j=1}^p \phi_{ij} \sigma_{q+j,n} \eta_{jn} \\
&+ 2 \sum_{j=1}^p \sigma_{in} \varepsilon_{in} \phi_{ij} \sigma_{q+j,n} \eta_{jn} + 2 \sum_{j>r} \phi_{ij} \phi_{ir} \sigma_{q+j,n} \sigma_{q+r,n} \eta_{jn} \eta_{rn}.
\end{aligned}$$

Note that the different additive terms in u_{i2n} are also uncorrelated (but not independent). Let $\sigma_n^2 = (\sigma_{1n}^2, \dots, \sigma_{q+p,n}^2)$. Then it is easily seen that

$$E(u_{i1n} | \sigma_n^2) = E(u_{i2n} | \sigma_n^2) = 0, \quad i = 1, \dots, q.$$

The conditional second order moments are also easy to derive. For $i = 1, \dots, q$,

$$\begin{aligned}
E(u_{i1n} u_{i2n} | \sigma_n^2) &= 2\mu_i \Delta \sigma_{in}^2 + 2\mu_i \Delta \sum_{j=1}^p \phi_{ij}^2 \sigma_{q+j,n}^2 \\
E(u_{i1n}^2 | \sigma_n^2) &= \sigma_{in}^2 + \sum_{j=1}^p \phi_{ij}^2 \sigma_{q+j,n}^2 \\
E(u_{2in}^2 | \sigma_n^2) &= 2\sigma_{in}^4 + 2 \sum_{j=1}^p \phi_{ij}^4 \sigma_{q+j,n}^4 + 4\mu_i^2 \Delta^2 \sigma_{in}^2 + 4\mu_i^2 \Delta^2 \sum_{j=1}^p \phi_{ij}^2 \sigma_{q+j,n}^2 + 4\sigma_{in}^2 \sum_{j=1}^p \phi_{ij}^2 \sigma_{q+j,n}^2 \\
&+ 4 \sum_{j>r} \phi_{ij}^2 \phi_{ir}^2 \sigma_{q+j,n}^2 \sigma_{q+r,n}^2. \tag{14}
\end{aligned}$$

Moreover, for $i, k = 1, \dots, q$ and $i \neq k$:

$$\begin{aligned}
E(u_{i1n} u_{k1n} | \sigma_n^2) &= \sum_{j=1}^p \phi_{ij} \phi_{kj} \sigma_{q+j,n}^2 \\
E(u_{i2n} u_{k2n} | \sigma_n^2) &= 2 \sum_{j=1}^p \phi_{ij}^2 \phi_{kj}^2 \sigma_{q+j,n}^4 + 4\mu_i \mu_k \Delta^2 \sum_{j=1}^p \phi_{ij} \phi_{kj} \sigma_{q+j,n}^2 + 4 \sum_{j>r} \phi_{ij} \phi_{ir} \phi_{kj} \phi_{kr} \sigma_{q+j,n}^2 \sigma_{q+r,n}^2 \\
E(u_{i1n} u_{k2n} | \sigma_n^2) &= 2\mu_k \Delta \sum_{j=1}^p \phi_{ij} \phi_{kj} \sigma_{q+j,n}^2. \tag{15}
\end{aligned}$$

Setting

$$u_n = [u_{11n}, u_{12n}, \dots, u_{q1n}, u_{q2n}]',$$

and applying the rule of iterated expectation to (14) and (15), we obtain

$$\begin{aligned}
\text{Var}(u_n) &\equiv \Sigma \\
&= \text{blockdiag}(\Sigma_1, \dots, \Sigma_q) + \tilde{\Phi} \Omega \tilde{\Phi}', \tag{16}
\end{aligned}$$

where $\text{blockdiag}(A_1, \dots, A_m)$, for general square matrices A_i , denotes the blockdiagonal matrix with the i 'th block equal to A_i . We have

$$\Sigma_i = \begin{bmatrix} E(\sigma_{in}^2) & 2\mu_i \Delta E(\sigma_{in}^2) \\ 2\mu_i \Delta E(\sigma_{in}^2) & 2E(\sigma_{in}^4) + 4\mu_i^2 \Delta^2 E(\sigma_{in}^2) + 4E(\sigma_{in}^2) \sum_{j=1}^p \phi_{ij}^2 E(\sigma_{q+j,n}^2) \end{bmatrix}, \quad i = 1, \dots, q.$$

Moreover,

$$\Omega = \begin{cases} \text{diag}(\omega_1) & \text{if } p = 1 \\ \text{diag}(\omega_1, \omega_2) & \text{if } p > 1 \end{cases},$$

with

$$\omega_1 = [E(\sigma_{q+1,n}^2), 2E(\sigma_{q+1,n}^4), E(\sigma_{q+2,n}^2), 2E(\sigma_{q+2,n}^4), \dots, E(\sigma_{q+p,n}^2), 2E(\sigma_{q+p,n}^4)] \quad (p \geq 1)$$

$$\omega_2 = [4E(\sigma_{q+2,n}^2 \sigma_{q+1,n}^2), 4E(\sigma_{q+3,n}^2 \sigma_{q+1,n}^2), 4E(\sigma_{q+3,n}^2 \sigma_{q+2,n}^2), \dots \\ \dots, 4E(\sigma_{q+p,n}^2 \sigma_{q+1,n}^2), 4E(\sigma_{q+p,n}^2 \sigma_{q+p-1,n}^2)] \quad (p \geq 2).$$

Furthermore,

$$\tilde{\Phi} = \begin{cases} \tilde{\Phi}_1 & \text{if } p = 1 \\ [\tilde{\Phi}_1 : \tilde{\Phi}_2] & \text{if } p > 1, \end{cases}$$

with

$$\tilde{\Phi}_1 = \begin{bmatrix} \phi_{11} & 0 & \cdots & \cdots & \phi_{1p} & 0 \\ 2\mu_1 \Delta \phi_{11} & \phi_{11}^2 & \cdots & \cdots & 2\mu_1 \Delta \phi_{1p} & \phi_{1p}^2 \\ \phi_{21} & 0 & \cdots & \cdots & \phi_{2p} & 0 \\ 2\mu_2 \Delta \phi_{21} & \phi_{21}^2 & \cdots & \cdots & 2\mu_2 \Delta \phi_{2p} & \phi_{2p}^2 \\ \vdots & \vdots & & & \vdots & \vdots \\ \vdots & \vdots & & & \vdots & \vdots \\ \phi_{q1} & 0 & & & \phi_{qp} & 0 \\ 2\mu_q \Delta \phi_{q1} & \phi_{q1}^2 & \cdots & \cdots & 2\mu_q \Delta \phi_{qp} & \phi_{qp}^2 \end{bmatrix} : 2q \times 2p$$

and

$$\tilde{\Phi}_2 = \begin{bmatrix} 0 & \cdots & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 \\ \phi_{12} \phi_{11} & \cdots & \phi_{1j} \phi_{11} & \cdots & \phi_{1j} \phi_{1,j-1} & \phi_{1,j+1} \phi_{11} & \cdots & \phi_{1,j+1} \phi_{1j} & \phi_{1p} \phi_{1,p-1} \\ 0 & \cdots & 0 & \cdots & 0 & 0 & \cdots & \vdots & 0 \\ \phi_{22} \phi_{21} & \cdots & \phi_{2j} \phi_{21} & \cdots & \phi_{2j} \phi_{2,j-1} & \phi_{2,j+1} \phi_{21} & \cdots & & \phi_{2p} \phi_{2,p-1} \\ 0 & \cdots & 0 & \cdots & 0 & 0 & \cdots & & 0 \\ \vdots & & \vdots & & \vdots & \vdots & & \vdots & \vdots \\ 0 & \cdots & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 \\ \phi_{q2} \phi_{q1} & \cdots & \phi_{qj} \phi_{q1} & \cdots & \phi_{qj} \phi_{q,j-1} & \phi_{p,j+1} \phi_{p1} & \cdots & \phi_{p,j+1} \phi_{pj} & \phi_{qp} \phi_{q,p-1} \end{bmatrix} : 2q \times \frac{p(p-1)}{2}.$$

Explicit expressions for $E(\sigma_{jn}^2)$ and $E(\sigma_{jn}^4)$ follow directly from (7):

$$\begin{aligned} E(\sigma_{jn}^2) &= \xi_j \Delta \\ E(\sigma_{jn}^4) &= \text{Var}(\sigma_{jn}^2) + \xi_j^2 \Delta^2 \\ &= 2\omega_j^2 \lambda_j^{-2} \{e^{-\lambda_j \Delta} - 1 + \lambda_j \Delta\} + \xi_j^2 \Delta^2. \end{aligned} \quad (17)$$

When $3p/2 + p^2/2 < 2q$, Σ has a sparse structure that can be utilized to simplify the calculation of Σ^{-1} needed in the estimation of the model (see Appendix B).

3 The quasi likelihood function based on an approximative Gaussian state space representation

A fully specified state space representation of the multivariate model considered in Section 2 is given below. The state space form allows us to formulate a quasi-likelihood function. We will combine features of the EM algorithm with an efficient quasi-Newton method. Let ψ denote the vector of unknown parameters to be estimated. In the EM algorithm, the log-likelihood function, $L(\psi)$ (which in our case will be a quasi log-likelihood function), is decomposed as:

$$L(\psi) = M(\psi|\psi') - H(\psi|\psi'), \quad (18)$$

where $M(\psi|\psi')$ is maximized iteratively with respect to ψ to update ψ' . Importantly, the function $M(\psi|\psi')$ has the following property:

$$\left. \frac{\partial L(\psi)}{\partial \psi} \right|_{\psi=\psi'} = \left. \frac{\partial M(\psi|\psi')}{\partial \psi} \right|_{\psi=\psi'}. \quad (19)$$

The relevance of this result in the present context is discussed in Raknerud and Skare (2009). Explicit expressions for $M(\psi|\psi')$ and $\partial M(\psi|\psi')/\partial \psi$ are given in Appendix B.

A state space form in the univariate case: $q = 1$ ($p = 0$), is given in BS (2001) with observation vector $Y_n = [y_1, y_1^2]'$ and state vector $\alpha_n = [\lambda\sigma_n^2 \quad \sigma^2(n\Delta)]$. Now consider the multivariate case, i.e., assuming that $\sigma_j^2(t)$ are independent OU processes for $j = 1, \dots, q + p$, with $E(\sigma_j^2(t)) = \xi_j$ and $Var(\sigma_j^2(t)) = \omega_j^2$. The autocorrelation function, $r_j(s)$, of $\sigma_j^2(t)$ is then given by $r_j(s) = \exp(-\lambda_j|s|)$. Our main result is stated in the following proposition:

Proposition 1 *Define as the observation vector*

$$Y_n = (y_{1n}, y_{1n}^2, \dots, y_{qn}, y_{qn}^2)'$$

and as the state vector

$$\tilde{\alpha}_n = [\sigma_{1n}^2 - \xi_1 \Delta, \sigma_1^2(n\Delta) - \xi_1, \dots, \sigma_{q+p,n}^2 - \xi_{q+p} \Delta, \sigma_{q+p}^2(n\Delta) - \xi_{q+p}]'.$$

Then

$$\begin{aligned} Y_n &= \tau + G\tilde{\alpha}_n + u_n, E(u_n) = 0, Var(u_n) = \Sigma \\ \tilde{\alpha}_n &= F\tilde{\alpha}_{n-1} + \tilde{\eta}_n, E(\tilde{\eta}_n) = 0, Var(\tilde{\eta}_n) = Q \end{aligned} \quad ; \quad n = 1, \dots, N,$$

where Σ is defined in (16); $\tilde{\eta}_n$ and u_m are uncorrelated vectors for all n, m ; and

$$\begin{aligned} \tau &= [\mu_1 \Delta, \mu_1^2 \Delta^2 + \xi_1 \Delta + \sum_{j=1}^p \phi_{1j} \xi_{q+j} \Delta, \dots, \mu_q \Delta, \mu_q^2 \Delta^2 + \xi_q \Delta + \sum_{j=1}^p \phi_{qj} \xi_{q+j} \Delta]' \\ G &= [I_q \quad \Phi] \otimes \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \\ F &= \begin{bmatrix} F_1 & 0 & 0 & 0 \\ 0 & F_2 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & F_{q+p} \end{bmatrix} \\ Q &= \begin{bmatrix} Q_1 & 0 & 0 & 0 \\ 0 & Q_2 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & Q_{q+p} \end{bmatrix}. \end{aligned}$$

F_j and Q_j are given by ($j = 1, \dots, q + p$):

$$\begin{aligned} F_j &= \begin{bmatrix} 0 & \frac{1 - e^{-\lambda_j \Delta}}{\lambda_j} \\ 0 & e^{-\lambda_j \Delta} \end{bmatrix} \\ Q_j &= 2\omega_j^2 \begin{bmatrix} \lambda_j^{-2} \left(-\frac{3}{2} - \frac{1}{2}e^{-2\lambda_j \Delta} + 2e^{-\lambda_j \Delta} + \lambda_j \Delta \right) & \lambda_j^{-1} \left(1 - e^{-\lambda_j \Delta} - \frac{1}{2}(1 - e^{-2\lambda_j \Delta}) \right) \\ \lambda_j^{-1} \left(1 - e^{-\lambda_j \Delta} - \frac{1}{2}(1 - e^{-2\lambda_j \Delta}) \right) & \frac{1}{2}(1 - e^{-2\lambda_j \Delta}) \end{bmatrix}. \end{aligned}$$

Having established (11)-(16) in Section 2, the proof of Proposition 1 is a straightforward extension of the proof of Proposition 2 in Raknerud and Skare (2009). The generalization to the case with superpositions is considered in Appendix A, i.e., where $\sigma_j^2(t)$ is a sum of m independent OU processes:

$$\sigma_j^2(t) = \sum_{k=1}^m \sigma_{jk}^2(t), \quad j = 1, \dots, q + p, \quad (20)$$

with $\sigma_{jk}^2(t)$ being independent OU processes with parameters $(\lambda_{jk}, \xi_{jk}, \omega_{jk})$ for $j = 1, \dots, p + q$ and $k = 1, \dots, m$. Then the autocorrelation function, $r_j(s)$, of $\sigma_j^2(t)$ is given by

$$r_j(s) = \sum_{k=1}^m w_{jk} e^{-\lambda_{jk}|s|}, \quad \text{with } w_{jk} = \omega_{jk}^2 / \sum_{k=1}^m \omega_{jk}^2. \quad (21)$$

As shown in Appendix A, only $E(\sigma_j^2(t)) \equiv \xi_j = \sum_{k=1}^m \xi_{jk}$ can be identified, not the individual parameters ξ_{jk} . As shown in Raknerud and Skare (2009), $m = 2$ gives a good fit to daily exchange rate data. This pattern is also confirmed by Griffin and Steel (2006) on daily US stock returns data.

To evaluate the log-likelihood, let $a_{t|s} = E(\tilde{\alpha}_t | z_1, \dots, z_s)$ and $V_{t|s} = Var(\tilde{\alpha}_t | z_1, \dots, z_s)$, which are easily computed by means of the Kalman filter and -smoother (see Appendix B) under the assumption of joint normality of all the random vectors. The (quasi) log-likelihood function based on the Gaussian state space model then takes the standard form:

$$L(\psi) = -\frac{1}{2} \sum_{n=1}^N \ln |D_n| + (Y_n - \tau - Ga_{n|n-1})' D_n^{-1} (Y_n - \tau - Ga_{n|n-1}).$$

See Appendix B for more details regarding the computation of D_n , $a_{n|n-1}$ and derivatives $\partial L(\psi)/\partial\psi = \partial M(\psi|\psi')/\partial\psi$.

The asymptotic theory for the quasi-likelihood estimator based on Gaussian approximations to models that can be formulated on a state space form with white noise error terms, is given in Dunsmuir (1979). In particular, the quasi-likelihood estimator is consistent, with

$$\sqrt{N}(\hat{\psi}_N - \psi^*) \xrightarrow{D} \mathcal{N}(0, J^{-1} I J^{-1}), \quad (22)$$

where

$$I = \lim_{N \rightarrow \infty} Var(N^{-1/2} \frac{\partial}{\partial\psi} L(\psi))$$

$$J = -\text{p} \lim_{N \rightarrow \infty} N^{-1} \frac{\partial^2}{\partial\psi \partial\psi'} L(\psi).$$

An estimator of J is a simple by-product of the quasi-Newton algorithm used to estimate ψ , while we estimate I using the formula in Gouriéroux et al. (1993).

4 Application: Exchange rates

The purpose of the application presented here is to demonstrate the potential use of the results established in Section 3. We consider bivariate data: the Euro/NOK and

US Dollar/NOK daily exchange rate data for the period 2.1.1989-4.2.2010. That is, y_{1n} and y_{2n} , for $n = 1, \dots, N$, are the daily changes in the log prices of Euro and US Dollar, respectively, measured in Norwegian kroner. Days where at least one of the exchange rates is missing from the data, are deleted, giving $N = 5332$. We will refer to y_{in} ($i = 1, 2$) as daily *returns*, i.e., the returns in NOK of the assets in a portfolio consisting of a unit Euro and a unit US Dollar. The exploration of the full potential of our methods in applications will be deferred to future research, as more work need to be done to optimize computer routines to handle high dimensional covariance matrices (utilizing sparse matrix structures in the programming codes, etc.). We first present results on univariate models, basically replicating Raknerud and Skare (2009), but on a more up-to-date data set (one more year of data is added here). Then we present the corresponding results for the bivariate model, i.e., with $q = 2$ equations and $p = 1$ common volatility factor. The state space equations for this special case of our model, with each volatility factor being a superposition of $m = 2$ OU processes, are given in Appendix A. In our application, $\Delta = 1$, which corresponds to one day.

The optimization of the quasi-likelihood $L(\psi)$ with respect to ψ is carried out by means of a quasi-Newton algorithm that incorporates Fletcher's line search sub-algorithm (Fletcher, 1987, p. 34). We consider the optimization as having converged when the gradient vector has no components exceeding 0.001 in absolute value. To take restrictions on the parameters into account, these are reparametrized as follows: For $j = 1, \dots, q$,

$$\begin{aligned}\lambda_{j1} &= \frac{\lambda_{\max}}{1 + e^{-c_{j1}}}, \quad \lambda_{jk} = \frac{\lambda_{j,k-1}}{1 + e^{-c_{jk}}}; \quad k = 2, \dots, m; \\ \omega_{jk} &= e^{c_{j,m+k}}; \quad k = 1, \dots, m; \quad \text{and } \xi_j = e^{c_{j,2m+1}}.\end{aligned}$$

λ_{\max} is a pre-specified upper bound on λ_{j1} , and $c_{j1}, \dots, c_{j,2m+1}$ are unrestricted parameters. Note that $\lambda_{j1} > \lambda_{j2} > \dots > \lambda_{jm}$.

The two returns series are depicted in Figure 2. The financial crisis that broke out in September 2008 is clearly visible in the form of a cluster of large spikes and dips in both returns series ("volatility clustering"). The presence of a positive correlation in daily returns between the Euro/NOK and US Dollar/NOK exchange rate is evident from the figure. The single largest spike in the *absolute* value of the

returns occurred from 22-23.10.2008 in both series, when both the Euro and Dollar rate dropped more than 4 per cent. High exchange rate volatility is also observed during the emerging market crisis of 1997-1998, characterized by high oil prices and wide NOK interest rate differentials against Euro and Dollar. In particular, both exchange rates decreased more than 3.3 per cent from 27.-28.08.1998. Related to the collapse of the ERM exchange rate system, we also observe considerable exchange rate volatility in the fourth quarter of 1992. In particular, from 9.-11.12.1992 the Euro increased by 3.9 percent, while the Dollar exchange rate increased by 3 per cent against the Norwegian krone. Descriptive statistics calculated from these empirical densities are shown in Table 1, including measures of skewness and kurtosis for daily returns, y_n , $n = 1, \dots, 5332$; and scaled 5-days returns: $\sqrt{5^{-1}} \sum_{i=1}^5 y_{5(m-1)+i}$, $m = 1, \dots, 1066$. Table 1 shows that the empirical coefficient of skewness is zero for all practical purposes, which is common for exchange rate data. For the daily returns, we find excess kurtosis (above 3) for both Euro and Dollar, but less so for Dollar (4.47) than for Euro (6.18). Both coefficients of kurtosis are closer to 3 for the 5-days returns than for the daily returns, which is as expected in view of the temporal aggregation result (4).

Univariate model ($q = 1, p = 0$) Results from the quasi-likelihood estimation of the model with superposition of $m = 2$ OU processes are shown in Table 2. When estimating models with $m = 3$, we obtain indistinguishable estimates of λ_2 and λ_3 , so we are not able to identify a third volatility component. Thus $m = 2$ seems to be adequate. Griffin and Steel (2006) came to the same conclusion using daily U.S. stock returns data. We see from Table 2 that the standard errors (SE) are quite large for the λ_i - and ω_i^2 -coefficients. For example, $SE(\hat{\lambda}_i)/\hat{\lambda}_i$ and $SE(\hat{\omega}_i^2)/\hat{\omega}_i^2$ are mostly between 1/2 and 1/3. As demonstrated in Raknerud and Skare (2009), a model without superposition is not able to pick up the slowly decaying empirical autocorrelation pattern for lags exceeding 5-10 days, cf. the empirical autocorrelation functions (ACFs) in chart (a) and (b) of Figure 4.

For both exchange rates, the smallest λ (λ_2) is estimated to around 0.01 and the largest λ (λ_1) to around 0.45. We also see from the estimates of ω_1^2 and ω_2^2 in Table 2, that the ACF for Dollar has more weight on λ_2 relative to λ_1 compared to

Euro, leading to a more slowly decaying pattern (cf. (21)).

We note that the estimated average spot volatility $E(\sigma^2(t)) = \xi$ is much higher for Dollar/NOK (0.44) than for Euro/NOK (0.12). The spot volatility of the Dollar/NOK rate also fluctuates much more over time: $Var(\sigma^2(t)) = \omega_1^2 + \omega_2^2$ is estimated to 0.37 for Dollar/NOK, but only to 0.12 for Euro/NOK. That the Dollar-volatility is much larger than the Euro-volatility is also evident from Figure 3, which shows the estimated (predicted) values of actual volatility obtained from the Kalman smoother. We see that the actual Dollar-volatility is almost uniformly higher than the Euro-volatility over the sample period.

Bivariate model ($q = 2, p = 1$) The results for the bivariate model are depicted in Table 3. The parameters are grouped into three classes: parameters specific to (i) the Euro/NOK equation, (ii) the Dollar/NOK equation and (iii) the common volatility factor ($\sigma_3^2(t)$). Guided by the univariate results in Table 2, a model specification where each $\sigma_j^2(t)$ ($j = 1, 2, 3$) is a superposition of $m = 2$ OU processes was estimated. The most striking aspect of the results is that no superposition of OU processes is necessary to represent $\sigma_1^2(t)$ and $\sigma_2^2(t)$, i.e., $\sigma_1^2(t) = \sigma_{11}^2(t)$ and $\sigma_2^2(t) = \sigma_{21}^2(t)$. This is seen from the fact that the maximum quasi-likelihood estimates of both ω_{12} and ω_{22} are zero (hence λ_{12} and λ_{22} are not identified). On the other hand, the estimated λ_{11} and λ_{21} are both very small (and smaller than the *smallest* λ , i.e., λ_2 , in the corresponding univariate models with superposition). This means that each volatility factor which is specific (idiosyncratic) to a particular exchange rate, is highly persistent, with few and small jumps (that are slowly discounted). On the other hand, the common volatility factor $\sigma_3^2(t)$ is a sum of a relatively short-memory component, with decay-parameter λ_{31} estimated to 0.28, and a long-memory component with decay-parameter λ_{32} estimated to 0.015. It is noteworthy that the standard error of $\hat{\lambda}_{31}$ (0.003) and $SE(\hat{\lambda}_{31})/\hat{\lambda}_{31}$ (0.07) is much smaller than the corresponding expressions for $\hat{\lambda}_1$ in both univariate models presented in Table 2 (around 0.15 and 0.3, respectively). Thus, the bivariate model seems to have substantially smaller estimation uncertainty with regard to key parameters, compared to the univariate models.

It follows from the results in Table 3 that the implied ACFs of the idiosyncratic

volatility factors $\sigma_1^2(t)$ and $\sigma_2^2(t)$ have a very slowly decaying pattern (small λ_{11} and λ_{21}). On the other hand, the ACF of $\sigma_3^2(t)$ is dominated by an OU process with a much higher decay-parameter λ_{31} (with estimated weight, w_{31} , equal to 6/7 (see (21)). The estimated loading coefficient ϕ_{21} of the common factor $\sigma_{3,n}\eta_{3n}$ in the Dollar/NOK equation (cf. (12)) is 1.27, whereas the loading coefficient in the Euro/NOK equation is $\phi_{11} = 1$ a priori. Thus both loading coefficients are of the same magnitude. We conclude that the periods of high-volatility in both series are determined by a common volatility factor characterized by large jumps that quickly die out. Note that all the co-movements in the returns depicted in Figure 4, are accounted for by the common factor.

Figure 4 depicts the empirical versus estimated (model-based) ACFs of squared returns for Euro/NOK (in chart (a)), Dollar/NOK (in chart (b)), and the cross-autocorrelation function (cross-ACF) between the two returns series (in chart (c)). The figure shows that the estimated model fits the auto- and cross-correlation patterns in the returns data well. At short lags, the estimated cross-ACF for the Euro/NOK and Dollar/NOK returns tail off quickly, and then much more slowly after 5–10 days lags. The unconditional correlation between the two returns (the cross-ACF at lag zero) is estimated to 0.39, but decreases rapidly towards zero. This holds both for the empirical and the estimated (model-based) cross-ACF. It appears from Figure 4 that the empirical cross-ACF pattern of the two exchange rates is picked up very well by the estimated model.

The strong dependence structure between the two exchange rate series is also evident from tables 4-6. Table 4 shows the simultaneous covariance matrix of the 4×1 vector $Y_n = [y_{1n}, y_{1n}^2, y_{2n}, y_{2n}^2]'$ of returns and squared returns. Table 5 displays the estimated covariance matrix Σ of the vector of error terms u_n (defined in (16)), whereas the corresponding correlation matrix is shown in Table 6. In particular, we see from Table 6 that the pairwise correlations between the error terms in the returns equations (i.e., the equations for y_{1n} and y_{2n} , see (11)) and the error terms of the two squared returns equations (i.e., the equations for y_{1n}^2 and y_{2n}^2 ; see (13)) are almost identical: 0.39. Figure 5 depicts the actual correlation for $n = 1, \dots, 5332$ between the Euro/NOK and Dollar/NOK returns, derived from the estimated actual

covariance Σ_n defined in (10). We see that the actual correlations fluctuate widely, from near zero to almost one. Periods with high actual correlation coincides with periods of high volatility, which is seen by comparing Figure 4 and 5. This result is consistent with the finding that the common factor $\sigma_{3n}\eta_{3n}$ is more volatile than the two idiosyncratic error components $\sigma_{in}\varepsilon_{in}$ ($i = 1, 2$), cf. (12).

5 Conclusions

In this paper we have developed and explored a quasi-likelihood estimator for a class of multivariate stochastic volatility models based on non-Gaussian Ornstein-Uhlenbeck (OU) processes. A parsimonious parametrization is achieved by imposing a factor structure, where individual asset returns are modelled as univariate processes, augmented with common dynamic factors that account for the co-movements among the returns series. These co-movements are assumed to consist of independent, common shocks, each of which are represented as the increment of a standard Brownian motion multiplied by a time-varying stochastic scale coefficient which accounts for volatility (stochastic variance). The common shocks enter one or more series at the same time, but with a different loading coefficient. We show that the OU framework allows exact discrete time transition equations that can be represented on a linear state space form. A quasi-likelihood function is constructed by means of the Kalman filter, assuming that the actual volatility process is a Gaussian latent variable. In an application using 5332 daily exchange rate observations for the period 2.1.1989-4.2.2010 for the Euro/NOK and US Dollar/NOK exchange rates, we find that our estimation algorithm is feasible with large data sets (large N) and have good convergence properties. The results show that periods of high volatility are mainly driven by one common volatility factor, and hence is a common characteristic of both series. This common factor can be represented as a superposition (sum) of two independent OU processes: one with a relatively high decay-parameter ($\lambda \approx 0.3$) and one with a small decay-parameter ($\lambda \approx 0.01$). On the other hand, shocks that are idiosyncratic to each of the series, have almost constant variance over time, i.e., they can be represented as OU-processes with small and infrequent jumps and corresponding decay-parameters that are less than 0.01.

6 References

- Aguilar, O. and M. West (2000). “Bayesian dynamic factor models and variance matrix discounting for portfolio allocation.” *Journal of Business and Economic Statistics*, 18, 338-357.
- Andersen, T., T. Bollerslev, F.X. Diebold and P. Labys (2001). “The distribution of realized exchange rate volatility.” *Journal of the American Statistical Association*, 96, 42-55.
- Andersen, T.G., R.A. Davis, J.-P. Kreiss and T. Mikosch (Eds.) (2009). *Handbook of Financial Time Series*. Berlin. Springer.
- Asai, M., M. McAleer and J. Yu (2006). “Multivariate stochastic volatility: a review.” *Econometric Reviews*, 25, 145-176.
- Barndorff-Nielsen O.E. and N. Shephard (2001). “Non-Gaussian Ornstein-Uhlenbeck-based models and some of their uses in financial economics (with discussion).” *Journal of the Royal Statistical Society Series B-Statistical Methodology*, 63, 167-241.
- Barndorff-Nielsen O.E. and N. Shephard (2002). “Econometric analysis of realized volatility and its use in estimating stochastic volatility models.” *Journal of the Royal Statistical Society Series B-Statistical Methodology*, 64, 253-280.
- Barndorff-Nielsen O.E. and N. Shephard (2003). “Integrated OU processes and non-Gaussian OU-based stochastic volatility models.” *Scandinavian Journal of Statistics*, 30, 277-295.
- Barndorff-Nielsen O.E. and N. Shephard (2004). “Econometric analysis of realized covariation: High frequency based covariance, regression, and correlation in financial economics.” *Econometrica*, 72, 885-925.
- Barndorff-Nielsen O.E, T. Mikosch and S.I. Resnick (Eds.) (2001). *Lévy processes : theory and applications*, Boston. Birkhäuser.
- Bauwens, I., S. Laurent and J.V.K. Rombouts (2006). “Multivariate GARCH: a survey.” *Journal of Applied Econometrics*, 21, 317-332.
- Benth F.E., M. Groth and R. Kufakunesu (2007). “Valuing volatility and variance swaps for a non-Gaussian Ornstein Uhlenbeck stochastic volatility model .” *Applied Mathematical Finance*, 14, 347-363.
- Bollerslev T., R.F. Engle and J.M. Wooldridge (1988). “A capital asset pricing model with time varying covariances.” *Journal of Political Economy*, 30, 116-131.
- Bollerslev T. (1990). “Modeling the coherence in short-run exchange rates: a multivariate ARCH approach.” *Review of Economics and Statistics*, 72, 498-505.
- Bondesson, L. (2002). “On the Levy measure of the Lognormal and the LogCauchy distributions.” *Methodology and Computing in Applied Probability*, 4, 243-256.
- Brockwell, P.J. and R.A. Davis (1993). *Time series: Theory and methods*. New York: Springer.
- Chib, S., F. Nardari and N. Shephard (2006). “Analysis of high dimensional multivariate stochastic volatility models.” *Journal of Econometrics*, 134, 341-371.

- Chib, S., Y. Omori and M. Asai (2009). "Multivariate Stochastic Volatility," in *Handbook of Financial Time Series*. (Eds: T.G. Andersen, R.A. Davis, J.-P. Kreiss and T. Mikosch), pp. 365-400. Berlin: Springer.
- Diebold, F.X. and M. Nerlove (1989). "The dynamics of exchange rate volatility: A multivariate latent factor ARCH model." *Journal of Applied Econometrics*, 4, 1-21.
- Drost, F.C. and T.E. Nijman (1993). "Temporal aggregation of GARCH processes." *Econometrica*, 61, 909-927.
- Dunsmuir, W. (1979). "A central limit theorem for parameter estimation in stationary time series and its application to models for a signal observed with noise." *Annals of Statistics*, 7, 490-506.
- Engle R.F., V.K. Ng and M. Rothschild (1990). "Asset pricing with a factor ARCH covariance structure." *Journal of Econometrics*, 45, 213-238.
- Fahrmeier, L. and G. Tutz (1994). *Multivariate statistical modelling based on generalized linear models*, Springer, New York.
- Fletcher, R. (1987). *Practical methods of optimization*. New York, John Wiley
- Gourieroux C., A. Monfort and E. Renault (1993). "Indirect inference." *Journal of Applied Econometrics*, 8, Supplement, S85-S118.
- Griffin, J.E. and M.F.J. Steel (2006). "Inference with non-Gaussian Ornstein-Uhlenbeck processes for stochastic volatility" *Journal of Econometrics*, 134, 605-644.
- Harvey, A.C., E. Ruiz and N. Shephard (1994). "Multivariate stochastic variance models." *Review of Economic Studies*, 61, 247-264.
- Harvey, A.C., S.J. Koopman and N. Shephard (Eds.) (2004). *State space and unobserved component models*. Cambridge. Cambridge University Press.
- Johannes, M. and N. Polson (2009). "Markov Chain Monte Carlo," in *Handbook of Financial Time Series*. (Eds: T.G. Andersen, R.A. Davis, J.-P. Kreiss and T. Mikosch). Berlin, Springer. pp. 1001-1013.
- Jondeau, E., S. Poon and M. Rockinger (2007). *Financial Modeling under Gaussian Distributions*. Springer.
- Jungbacker, B. and S.J. Koopman (2006). "Monte Carlo likelihood estimation for three multivariate stochastic volatility models." *Econometric Reviews*, 25, 385-408.
- King, M., E. Sentana and S. Wadhvani (1994). "Volatility and links between national stock markets." *Econometrica*, 62, 901-933.
- Lyons, R.K. (2001). *The Microstructure Approach to Exchange Rates*, MIT Press, Cambridge, MA.
- Meese, R. and K. Rogoff (1983). "The Out-of-Sample Failure of Empirical Exchange Rate Models: Sampling Error or Misspecification?" In *Exchange Rates and International Macroeconomics* (Ed.: J.A. Frenkel), National Bureau of Economic Research, pp. 67-112.
- Meddahi, N and E. Renault (2004). "Temporal aggregation of volatility models." *Journal of Econometrics*, 119, 355-379.

- Nicolato, E. and E. Venardos (2003). “Option pricing in stochastic volatility models of the Ornstein-Uhlenbeck type.” *Mathematical Finance*, 13, 445-466.
- Raknerud, A., Ø. Skare (2009). “Indirect inference methods for stochastic volatility models based on non-Gaussian Ornstein-Uhlenbeck processes”, Discussion Papers 601, Statistics Norway.
- Shephard, N. (Ed.) (2005). *Stochastic volatility: selected readings*, Oxford. Oxford University Press.
- Silvennoinen, A. and T. Teräsvirta (2009). “Multivariate GARCH models,” in *Handbook of Financial Time Series*. (Eds: T.G. Andersen, R.A. Davis, J.-P. Kreiss and T. Mikosch), pp. 201-229. Berlin: Springer.
- Taylor, S.J. (1986). *Modeling financial time series*. Chichester. John Wiley.
- Yu, J. and R. Meyer (2006). “Multivariate stochastic volatility models: Bayesian estimation and model comparison.” *Econometric Reviews*, 25, 361-384.
- Zhang, S. and X. Zhang (2008). “Exact simulation of IG-OU processes.” *Methodology and computing in applied probability*, 10, 337-355.

Appendix A: Superpositions

Assume the $\sigma_j^2(t)$ are given by (20), then

$$\begin{aligned} E(\sigma_{jn}^2) &= \xi_j \Delta, \quad \xi_j = \sum_{k=1}^m \xi_{jk}, \quad \xi_{jk} > 0 \\ E(\sigma_{jn}^4) &= \text{Var}(\sigma_{jn}^2) + \xi_j^2 \Delta^2 \\ &= \sum_{k=1}^m 2\omega_{jk}^2 \lambda_{jk}^{-2} \{e^{-\lambda_{jk}\Delta} - 1 + \lambda_{jk}\Delta\} + \xi_j^2 \Delta^2 \end{aligned}$$

and Σ is modified accordingly. The individual parameters ξ_{jk} are not identified, only $\xi_j = \sum_{k=1}^m \xi_{jk}$. The state space representation in Proposition 1 is modified in the following way:

$$\begin{aligned} \tilde{\alpha}_n &= \left[\sigma_{11n}^2 - \xi_{11}\Delta, \quad \sigma_{11}^2(n\Delta) - \xi_{11}, \quad \dots, \quad \sigma_{1mn}^2 - \xi_{1m}\Delta, \quad \sigma_{mn}^2(n\Delta) - \xi_{1m}, \right. \\ &\quad \dots, \quad \sigma_{j,1,n}^2 - \xi_{j,1}\Delta, \quad \sigma_{j1}^2(n\Delta) - \xi_{j1}, \quad \dots, \quad \sigma_{jmn}^2 - \xi_{jm}\Delta, \quad \sigma_{jm}^2(n\Delta) - \xi_{jm} \\ &\quad \left. \sigma_{q+p,1,n}^2 - \xi_{q+p,1}\Delta, \quad \sigma_{q+p,1}^2(n\Delta) - \xi_{q+p,1}, \quad \dots, \quad \sigma_{q+p,mn}^2 - \xi_{q+p,m}\Delta, \quad \sigma_{q+p,m}^2(n\Delta) - \xi_{q+p,m} \right] \\ G &= \left[I_q \quad \Phi \right] \otimes \left[G_1 \quad \dots \quad G_m \right], \quad \text{where } G_k = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \quad k = 1, \dots, m \\ F_i &= \text{blockdiag}[F_{i1}, \dots, F_{im}], \quad F_{ik} = \begin{bmatrix} 0 & \frac{1-e^{-\lambda_{ik}\Delta}}{\lambda_{ik}} \\ 0 & e^{-\lambda_{ik}\Delta} \end{bmatrix}, \quad k = 1, \dots, m, \\ Q_i &= \text{blockdiag}[Q_{i1}, \dots, Q_{im}], \quad \text{where} \\ Q_{ik} &= 2\omega_{ik}^2 \begin{bmatrix} \lambda_{ik}^{-2} \left(-\frac{3}{2} - \frac{1}{2}e^{-2\lambda_{ik}\Delta} + 2e^{-\lambda_{ik}\Delta} + \lambda_{ik}\Delta \right) & \lambda_{ik}^{-1} \left(1 - e^{-\lambda_{ik}\Delta} - \frac{1}{2}(1 - e^{-2\lambda_{ik}\Delta}) \right) \\ \lambda_{ik}^{-1} \left(1 - e^{-\lambda_{ik}\Delta} - \frac{1}{2}(1 - e^{-2\lambda_{ik}\Delta}) \right) & \frac{1}{2}(1 - e^{-2\lambda_{ik}\Delta}) \end{bmatrix}, \quad k = 1, \dots, m \end{aligned}$$

In a typical situation p will be much smaller than q , and so the following matrix inversion lemma is useful:

$$D_n^{-1} = \Sigma^{-1} - \Sigma^{-1}G(V_{n|n-1}^{-1} + G'\Sigma^{-1}G)^{-1}G'\Sigma^{-1}.$$

Moreover, if $3p/2 + p^2/2 < 2q$, Σ^{-1} can be calculated as (see (16)):

$$\begin{aligned} \Sigma^{-1} &= \text{blockdiag}(\Sigma_1^{-1}, \dots, \Sigma_q^{-1}) - \text{blockdiag}(\Sigma_1^{-1}, \dots, \Sigma_q^{-1})\tilde{\Phi} \\ &\quad \times \left(\Omega^{-1} + \tilde{\Phi}'\Omega^{-1}\tilde{\Phi} \right)^{-1} \\ &\quad \times \tilde{\Phi}'\text{blockdiag}(\Sigma_1^{-1}, \dots, \Sigma_q^{-1}) \end{aligned}$$

The required conditional expectations $a_{n|N}$ and variances $V_{n|N}$ are obtained in subsequent backward smoothing recursions (see Fahrmeir and Tutz, 1994, p. 265):

Kalman smoothing:

For $n = N, \dots, 2$:

$$a_{n-1|N} = a_{n-1|n-1} + B_n(a_{n|N} - a_{n|n-1})$$

$$V_{n-1|N} = V_{n-1|n-1} + B_n(V_{n|N} - V_{n|n-1})B_n'$$

where

$$B_n = V_{n-1|n-1}F'V_{n|n-1}^{-1}. \quad (23)$$

Expressions for $M(\psi|\psi^0)$ and $\frac{\partial M(\psi|\psi^0)}{\partial \psi}$ First

$$M(\psi|\psi^0) = E \{ \ln f(Y, \alpha; \psi) | Y; \psi^0 \}, \quad (24)$$

where $f(Y, \alpha; \psi)$ is generic notation for the joint normal density function of (Y, α) given the parameter vector ψ , where $Y = \{Y_n\}_{n=1}^N$ is the observed Y_n -vectors, $\alpha = \{\tilde{\alpha}_n\}_{n=1}^N$ are the latent variables and $E \{ \cdot | Y; \psi^0 \}$ denotes the condition expectation given Y evaluated at $\psi = \psi^0$. We need to evaluate

$$M(\psi|\psi^0) = M^{(1)}(\Sigma, \tau, G|\psi^0) + M^{(2)}(F, Q|\psi^0), \quad (25)$$

where

$$\begin{aligned} M^{(1)}(\Sigma, \tau, G|\psi^0) &= -\frac{N}{2} \ln |\Sigma| \\ &\quad - \frac{1}{2} \sum_{n=1}^N \left[\text{tr} \{ \Sigma^{-1}(Y_n - \tau - Ga_{n|N})(Y_n - \tau - Ga_{n|N})' \} + \text{tr} \{ \Sigma^{-1}GV_{n|N}G' \} \right] \end{aligned} \quad (26)$$

and

$$\begin{aligned}
M^{(2)}(F, Q|\psi^0) &= \sum_{i=1}^m \left(-\frac{N}{2} \ln |Q_i| \right. \\
&\quad \left. - \frac{1}{2} \sum_{n=1}^N \left[\text{tr} \left\{ Q_i^{-1} (a_{n|N}^{(i)} - F_i a_{n|N}^{(i)}) (a_{n|N}^{(i)} - F_i a_{n|N}^{(i)})' \right\} \right] \right. \\
&\quad \left. + \text{tr} \left\{ Q_i^{-1} \left(V_{n|N}^{(i,i)} - (V'_{n|N} B'_n)^{(i,i)} F'_i - F_i (B_n V_{n|N})^{(i,i)} \right) + F_i V_{n|N}^{(i,i)} F'_i \right\} \right), \tag{27}
\end{aligned}$$

where the $2m$ dimensional vector a is defined by the partition

$$a = \begin{bmatrix} a^{(1)} \\ \vdots \\ a^{(m)} \end{bmatrix},$$

with $a^{(i)}$ ($i = 1, 2, \dots, m$) being 2-dimensional vectors. Moreover, the $2m \times 2m$ matrix A , is defined by the partition

$$A = \begin{bmatrix} A^{(1,1)} & A^{(1,2)} & \dots & A^{(1,m)} \\ A^{(2,1)} & A^{(2,2)} & \dots & A^{(2,m)} \\ \vdots & \vdots & \ddots & \vdots \\ A^{(m,1)} & A^{(m,2)} & \dots & A^{(m,m)} \end{bmatrix},$$

with $A^{(i,j)}$ ($i, j = 1, 2, \dots, m$) being 2×2 matrices. Furthermore, B_n is defined in (23), and we have utilized that

$$E(\alpha_n \alpha'_{n-1} | Y; \psi) = a_{n|N} a_{n-1|N}' + V'_{n|N} B'_n$$

The partial derivatives are then given by:

$$\begin{aligned}
\frac{\partial M^{(1)}(\Sigma, \tau, G|\psi^0)}{\partial G} &= \sum_{n=1}^N \Sigma^{-1} \left((Y_n - \tau - G a_{n|N}) a'_{n|N} - G V_{n|N} \right) \\
\frac{\partial M^{(1)}(\Sigma, \tau, G|\psi^0)}{\partial \tau} &= \sum_{n=1}^N \Sigma^{-1} (Y_n - \tau - G a_{n|N}) \\
\frac{\partial M^{(1)}(\Sigma, \tau, G|\psi^0)}{\partial \text{vec}(\Sigma)} &= -\frac{N}{2} \text{vec}(\Sigma^{-1}) + \frac{1}{2} (\Sigma^{-1} \otimes \Sigma^{-1}) \sum_{n=1}^N \text{vec} \left[(Y_n - \tau - G a_{n|N}) (Y_n - \tau - G a_{n|N})' \right. \\
&\quad \left. + G V_{n|N} G' \right] \\
\frac{\partial M^{(2)}(F, Q|\psi^0)}{\partial \text{vec}(Q_i)} &= -\frac{N}{2} \text{vec}(Q_i^{-1}) + \frac{1}{2} (Q_i^{-1} \otimes Q_i^{-1}) \sum_{n=1}^N \text{vec} \left[(a_{n|N}^{(i)} - F_i a_{n-1|N}^{(i)}) (a_{n|N}^{(i)} - F_i a_{n-1|N}^{(i)})' \right. \\
&\quad \left. + V_{n|N}^{(i,i)} - F_i (B_n V_{n|N})^{(i,i)} - (V'_{n|N} B'_n)^{(i,i)} F'_i + F_i (V_{n-1|N})^{(i,i)} F'_i \right] \\
\frac{\partial M^{(2)}(F, Q|\psi^0)}{\partial F_i} &= Q_i^{-1} \left[\left(\sum_{n=1}^N a_{n|N}^{(i)} a_{n-1|N}^{(i)'} + (V'_{n|N} B'_n)^{(i,i)} \right) - \Phi \left(\sum_{n=1}^N a_{n-1|N}^{(i)} a_{n-1|N}^{(i)'} + V_{n-1|N}^{(i,i)} \right) \right] \tag{28}
\end{aligned}$$

Finally, (Σ, τ, G, F, Q) are functions of the free parameters ψ , and the partial derivatives with respect to ψ are trivially obtained using the chain rule on (28).

Figures and tables

Table 1: **Descriptive statistics.** Daily and scaled 5-day returns

	Euro/NOK		Dollar/NOK	
	$\Delta = 1$	$\Delta = 5$	$\Delta = 1$	$\Delta = 5$
Mean	0.001	0.006	0.0002	-0.0005
Variance	0.10	0.11	0.46	0.48
Skewness	0.17	0.29	0.15	0.07
Kurtosis	6.18	4.52	4.47	3.33

Table 2: **Quasi-likelihood estimates.** Univariate models for Euro/NOK and Dollar/NOK exchange rates (2.1.1989-4.2.2009). No. of superpositions: $m=2$

Parameter	Euro/NOK		Dollar/NOK	
	Estimate	95% CI	Estimate	95% CI
μ	.017 (.02)	[-.01, .03]	-.002 (.01)	[-.02, .02]
λ_1	.45 (.12)	[.27, .77]	.45 (.15)	[.23, .85]
λ_2	.015 (.008)	[.004, .04]	.010 (.002)	[.007, .015]
ξ_1	.12 (.02)	[.07, .18]	.51 (.02)	[.48, .55]
ω_1^2	.10 (.06)	[.03, .33]	.24 (.06)	[.15, .38]
ω_2^2	.02 (.01)	[.001, .07]	.13 (.04)	[.07, .22]

*Standard errors in parentheses based on (estimated) asymptotic covariance matrix $J^{-1}IJ$

**95 % Confidence Intervals (CI) are transformed CI of unrestricted parameters c_1, \dots, c_{2m+1}

Table 3: **Quasi-likelihood estimates of bivariate (Euro/NOK, Dollar/NOK) model.** Exchange rates from 2.1.1989-4.2.2009. No. of superpositions: $m=2$

Parameter	Estimates*	95% CI**
Euro/NOK specific parameters:		
μ_1	.022 (.005)	[.01, .03]
λ_{11}	.0004 (.0005)	[.000, .005]
λ_{12}	–	–
ξ_1	.06 (.006)	[.05, .08]
ω_{11}^2	.05 (.03)	[.02, .14]
ω_{12}^2	0 (–)	–
ϕ_{11}	1 (–)	–
Dollar/NOK specific parameters:		
μ_2	.007 (.01)	[–.01, .03]
λ_{21}	.003 (.001)	[.001, .006]
λ_{22}	–	–
ξ_2	.37 (.017)	[.34, .41]
ω_{21}^2	.20 (.07)	[.10, .40]
ω_{22}^2	0 (–)	–
ϕ_{21}	1.27 (.22)	[.81, 1.71]
Common parameters (corresponding to $\sigma_3^2(t)$):		
λ_{31}	.28 (.02)	[.21, .91]
λ_{32}	.015 (.003)	[.01, .02]
ξ_3	.08 (.007)	[.07, .10]
ω_{31}^2	.06 (.02)	[.03, .12]
ω_{32}^2	.01 (.006)	[.006, .033]

*Standard errors in parentheses based on (estimated) asymptotic covariance matrix $J^{-1}IJ$

**95 % Confidence Intervals (CI) are transformed CI of unrestricted parameters

$c_{j1}, \dots, c_{j,2m+1}$

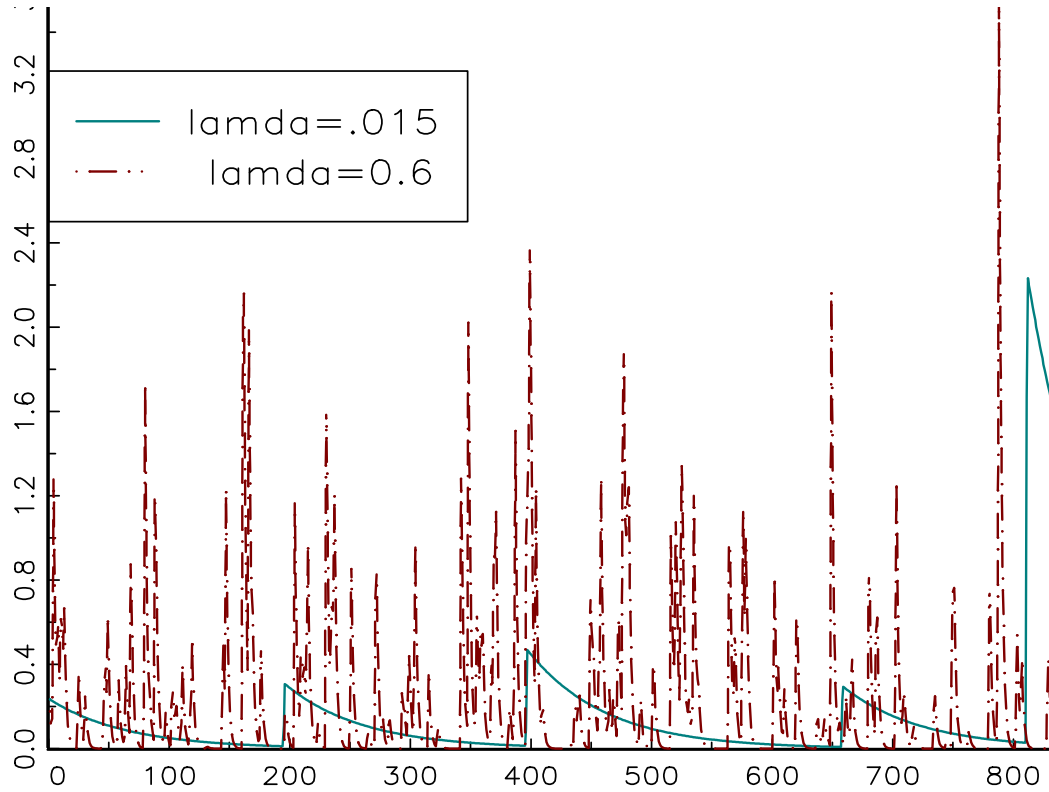


Figure 1: Two simulated actual volatility series σ_n^2

Table 4: **Empirical covariance matrix for returns and squared returns**

Euro/NOK		Dollar/NOK	
returns	squared returns	returns	squared returns
.14	.031	.11	.006
.	.39	.013	.40
.	.	.52	.016
.	.	.	1.62

Table 5: **Quasi-likelihood estimate of Σ** . No. of superpositions: $m=2$

Euro/NOK		Dollar/NOK	
returns	squared returns	returns	squared returns
.14	.006	.11	.002
.	.28	.005	.24
.	.	.52	.008
.	.	.	1.30

Table 6: Correlation matrix derived from estimate of Σ . No. of superpositions: $m=2$

	Euro/NOK		Dollar/NOK	
	returns	squared returns	returns	squared returns
1	1	.032	.399	.004
.		1	.013	.393
.			1	.010
.				1

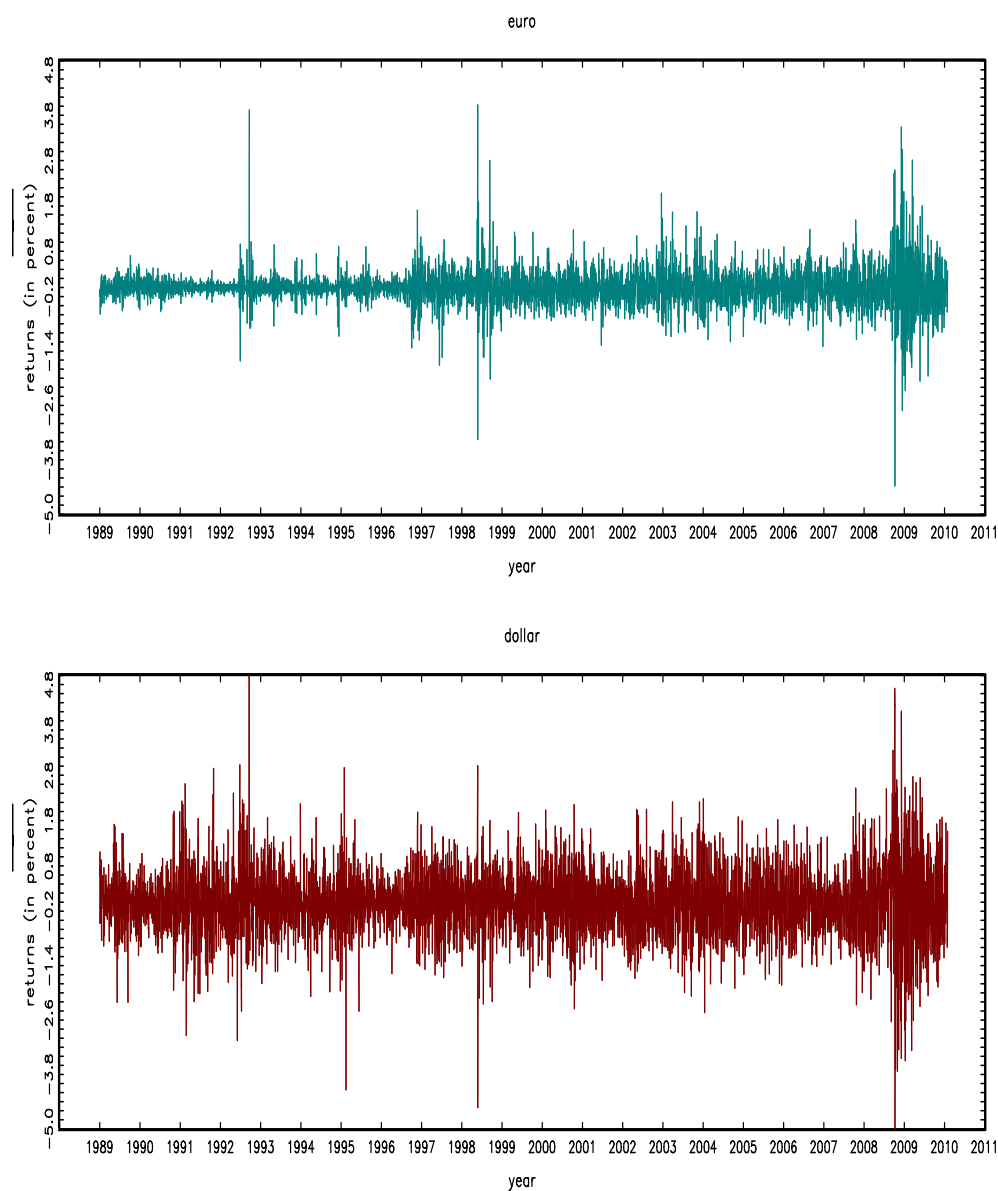


Figure 2: Daily relative changes in exchange rates (in percent): Euro/NOK and US Dollar/NOK

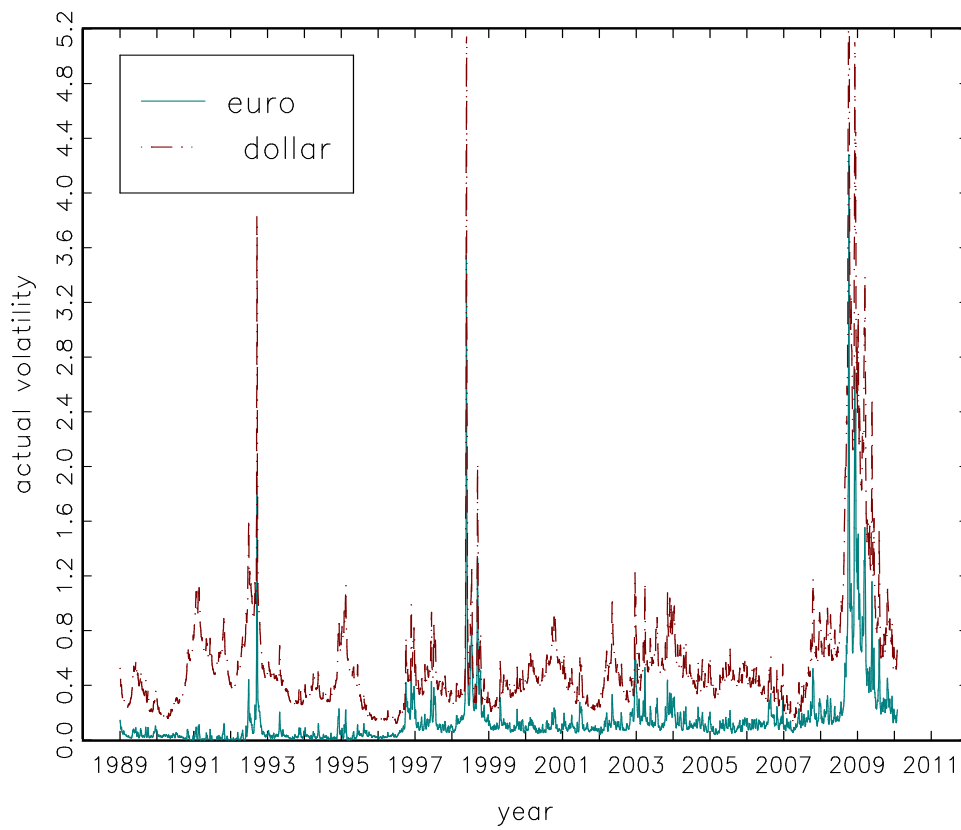


Figure 3: Actual volatility estimates

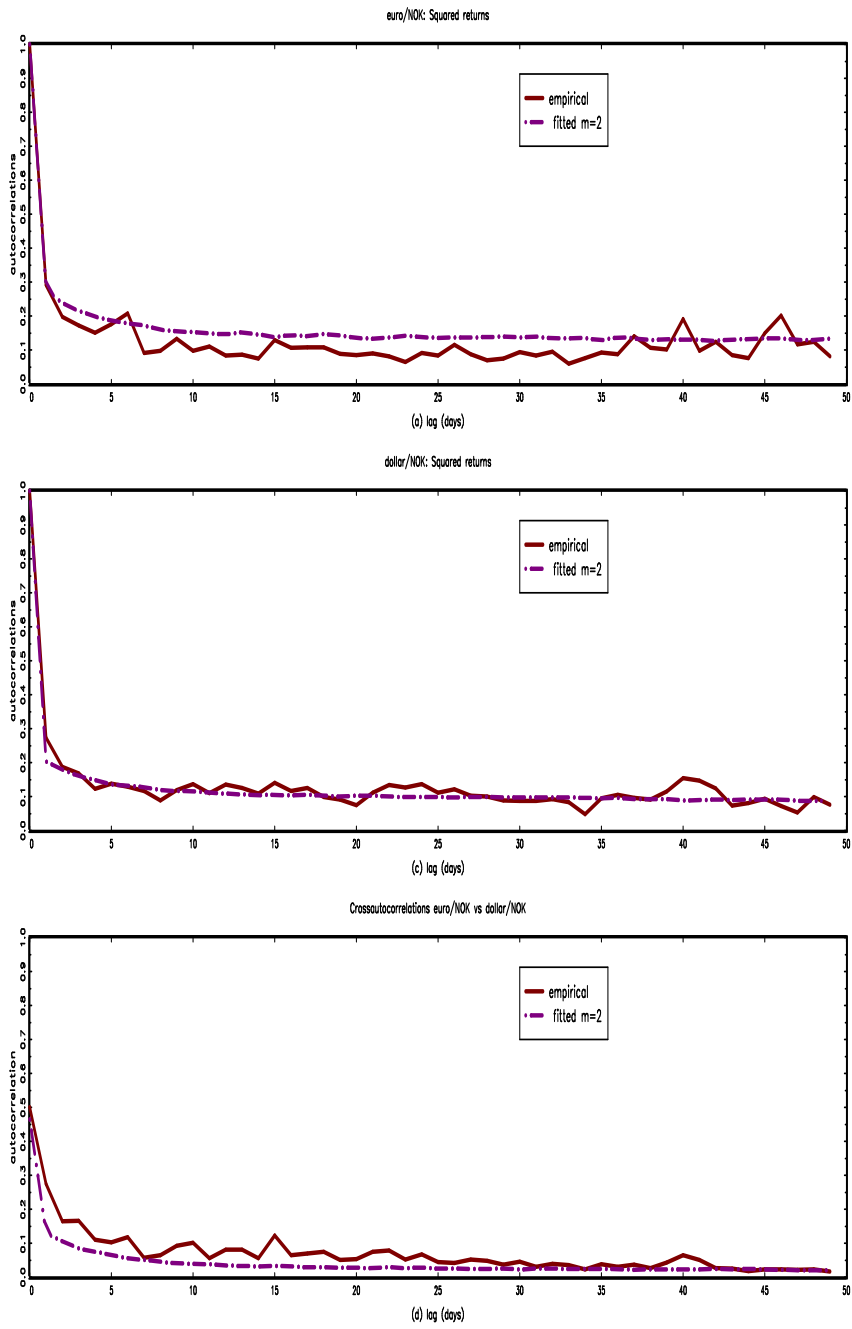


Figure 4: Autocorrelation functions of squared returns for Euro/NOK in chart (a), Dollar/NOK in chart (b) and the cross-autocorrelation function between the Euro/NOK and Dollar/NOK exchange rates in chart (c)

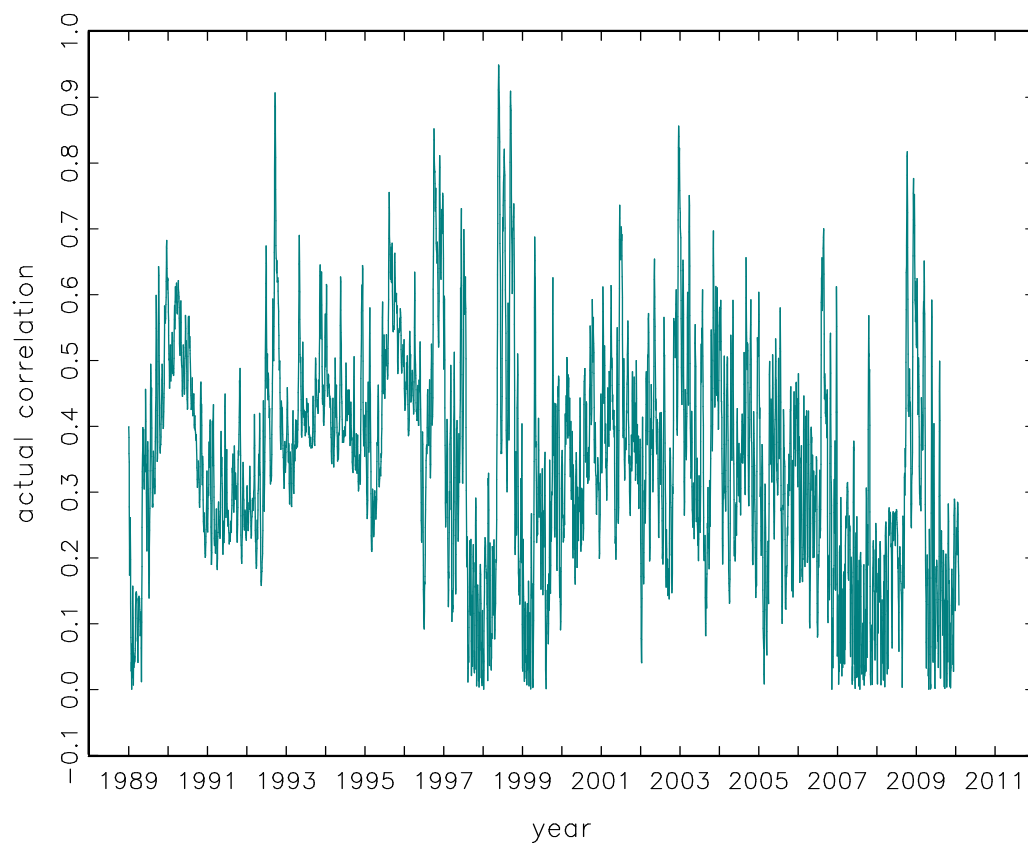


Figure 5: Actual correlation estimates