

Joint analysis and estimation of stock prices and trading volume in Barndorff-Nielsen and Shephard stochastic volatility models

Petra Posedel

Department of Mathematics, University of Zagreb, Croatia

Joint work with Friedrich Hubalek, Technische Universität Vienna

Workshop on Statistical Inference for Lévy Processes

EURANDOM

Eindhoven, July 15-17, 2009

Outline

- 1 Barndorff-Nielsen Shepard (BNS) stochastic volatility models
 - Discretely observed continuous time model
- 2 The simple explicit estimator
 - Estimating functions
 - The explicit solution
 - Consistency and asymptotic normality
 - Numerical illustrations
- 3 Application of the results to real data
- 4 Future work

Inference for BNS stochastic volatility models

Outline of the talk – part 1

- In the discretely observed Barndorff-Nielsen Shepard (BNS) setting we explore the joint distribution of spot prices X and the instantaneous trading volume τ since both quantities are observable
- inference by the ***martingale estimating function approach*** leads to an explicit estimator
- we show the strong law of large numbers for $(\tau_i^p, i \geq 0)$ and $(\tau_{i-1}^r X_i^p \tau_i^q, i \geq 1)$ without ergodicity arguments

Inference for BNS stochastic volatility models

Outline of the talk – part 2

- **consistency** and **asymptotic normality** of the simple explicit estimator
- we evaluate the performance of our estimator in a simulation study for a concrete specification, the Γ -OU
- application of the results to *daily* data
- further issues: superposition of OU-processes; comparison with the GMM estimation procedure

The model

Continuous time model

$$dX(t) = (\mu + \beta\tau(t-))dt + \sigma\sqrt{\tau(t-)}dW(t) + \rho dZ_\lambda(t), \quad X(0) = 0.$$

and

$$d\tau(t) = -\lambda\tau(t-)dt + dZ_\lambda(t), \quad \tau(0) = \tau_0,$$

where $\mu, \beta, \rho, \lambda \in \mathbb{R}$ with $\lambda > 0$. $Z = (Z_t)$ is the BDLP
 $Z_\lambda(t) = Z(\lambda t)$.

τ_0 has a self-decomposable distribution corresponding to the BDLP s. t. the process τ is strictly stationary and

$$E[\tau_0] = \zeta, \quad \text{Var}[\tau_0] = \eta.$$

Assumptions: $E[\tau_0^n] < \infty, \forall n \in \mathbb{N}$.

True for Γ -OU, IG-OU,

For our analysis we will assume that the instantaneous variance process V is a constant time the trading volume/number of trades τ . That is,

$$dV(t) = \sigma^2 \cdot d\tau(t), \quad (1)$$

with $\sigma > 0$.

τ_0 has a self-decomposable distribution corresponding to the BDLP s. t. the process τ is strictly stationary and

$$E[\tau_0] = \zeta, \quad \text{Var}[\tau_0] = \eta.$$

Assumptions: $E[\tau_0^n] < \infty, \forall n \in \mathbb{N}$.

True for Γ -OU, IG-OU,

For our analysis we will assume that the instantaneous variance process V is a constant time the trading volume/number of trades τ . That is,

$$dV(t) = \sigma^2 \cdot d\tau(t), \quad (1)$$

with $\sigma > 0$.

Outline

- 1 Barndorff-Nielsen Shepard (BNS) stochastic volatility models
 - Discretely observed continuous time model
- 2 The simple explicit estimator
 - Estimating functions
 - The explicit solution
 - Consistency and asymptotic normality
 - Numerical illustrations
- 3 Application of the results to real data
- 4 Future work

Discretely observed continuous time model

Observing X and τ on a discrete grid of points in time, $0 = t_0 < t_1 < \dots < t_n$, we obtain:

$$\tau(t_i) = \tau(t_{i-1})e^{-\lambda(t_i - t_{i-1})} + \int_{t_{i-1}}^{t_i} e^{-\lambda(t_i - s)} dZ_\lambda(s)$$

and

$$\begin{aligned} X(t_i) - X(t_{i-1}) &= \mu(t_i - t_{i-1}) + \beta(Y(t_i) - Y(t_{i-1})) + \sigma \int_{t_{i-1}}^{t_i} \sqrt{\tau(s-)} dW(s) \\ &\quad + \rho(Z_\lambda(t_i) - Z_\lambda(t_{i-1})), \end{aligned}$$

where $Y(t) = \int_0^t \tau(s-) ds$ is the integrated trading volume process.

Discretely observed continuous time model

Observing X and τ on a discrete grid of points in time, $0 = t_0 < t_1 < \dots < t_n$, we obtain:

$$\tau(t_i) = \tau(t_{i-1})e^{-\lambda(t_i - t_{i-1})} + \int_{t_{i-1}}^{t_i} e^{-\lambda(t_i - s)} dZ_\lambda(s)$$

and

$$\begin{aligned} X(t_i) - X(t_{i-1}) &= \mu(t_i - t_{i-1}) + \beta(Y(t_i) - Y(t_{i-1})) + \sigma \int_{t_{i-1}}^{t_i} \sqrt{\tau(s-)} dW(s) \\ &+ \rho(Z_\lambda(t_i) - Z_\lambda(t_{i-1})), \end{aligned}$$

where $Y(t) = \int_0^t \tau(s-) ds$ is the integrated trading volume process.

Discretely observed continuous time model

Introducing

$$X_i = X(t_i) - X(t_{i-1}), \quad Y_i = Y(t_i) - Y(t_{i-1}), \quad Z_i = Z_\lambda(t_i) - Z_\lambda(t_{i-1}),$$

$$W_i = \frac{1}{\sqrt{Y_i}} \int_{t_{i-1}}^{t_i} \sqrt{\tau(s-)} dW(s) \stackrel{i.i.d.}{\sim} N(0, 1),$$

and an auxiliary quantity $U(t) = \int_0^t e^{-\lambda(t-s)} dZ_\lambda(s)$,

using $t_k = \Delta k$, $\Delta > 0$ fixed and $\tau_i = \tau(t_i)$ we have

$$\tau_i = \gamma \tau_{i-1} + U_i, \quad Y_i = \epsilon \tau_{i-1} + S_i, \quad S_i = \frac{1}{\lambda} (Z_i - U_i)$$

$$X_i = \mu \Delta + \beta Y_i + \sigma \sqrt{Y_i} W_i + \rho Z_i,$$

where $\gamma = e^{-\lambda}$, $U_i = U(t_i) - U(t_{i-1})$, (U_i, Z_i) i.i.d

Discretely observed continuous time model

Introducing

$$X_i = X(t_i) - X(t_{i-1}), \quad Y_i = Y(t_i) - Y(t_{i-1}), \quad Z_i = Z_\lambda(t_i) - Z_\lambda(t_{i-1}),$$

$$W_i = \frac{1}{\sqrt{Y_i}} \int_{t_{i-1}}^{t_i} \sqrt{\tau(s-)} dW(s) \stackrel{i.i.d.}{\sim} N(0, 1),$$

and an auxiliary quantity $U(t) = \int_0^t e^{-\lambda(t-s)} dZ_\lambda(s)$,

using $\mathbf{t}_k = \Delta k$, $\Delta > 0$ fixed and $\tau_i = \tau(t_i)$ we have

$$\tau_i = \gamma \tau_{i-1} + U_i, \quad Y_i = \epsilon \tau_{i-1} + S_i, \quad S_i = \frac{1}{\lambda} (Z_i - U_i)$$

$$X_i = \mu \Delta + \beta Y_i + \sigma \sqrt{Y_i} W_i + \rho Z_i,$$

where $\gamma = e^{-\lambda}$, $U_i = U(t_i) - U(t_{i-1})$, (U_i, Z_i) i.i.d

Outline

- 1 Barndorff-Nielsen Shepard (BNS) stochastic volatility models
 - Discretely observed continuous time model
- 2 The simple explicit estimator
 - Estimating functions
 - The explicit solution
 - Consistency and asymptotic normality
 - Numerical illustrations
- 3 Application of the results to real data
- 4 Future work

Martingale estimating functions

- Suppose that $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n$ are observations from a model with a d -dimensional parameter $\theta \in \Theta$.
- an estimator $\hat{\theta}_n$ is obtained solving the equation

$$G_n(\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n; \theta) = 0, \quad (2)$$

where $G_n(\theta)$ is a d -dim estimating function of the parameter $\theta \in \mathbb{R}^d$.

- Among the class of unbiased or Fisher consistent estimating functions, we will analyze those estimating functions that are *martingales*.

Martingale estimating functions

- Suppose that $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n$ are observations from a model with a d -dimensional parameter $\theta \in \Theta$.
- an estimator $\hat{\theta}_n$ is obtained solving the equation

$$G_n(\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n; \theta) = 0, \quad (2)$$

where $G_n(\theta)$ is a d -dim estimating function of the parameter $\theta \in \mathbb{R}^d$.

- Among the class of unbiased or Fisher consistent estimating functions, we will analyze those estimating functions that are ***martingales***.

Estimating equations

Let $\theta = (\nu, \alpha, \lambda, \mu, \beta, \sigma, \rho)^\top$, $\mathbf{X} = (X, \tau)$.

We consider the following 7 martingale estimating functions:

$$G_n^1(\theta) = \sum_{i=1}^n [\tau_i - E(\tau_i | \tau_{i-1})], \quad G_n^2(\theta) = \sum_{i=1}^n [\tau_i \tau_{i-1} - \tau_{i-1} E(\tau_i | \tau_{i-1})],$$

$$G_n^3(\theta) = \sum_{i=1}^n [\tau_i^2 - E(\tau_i^2 | \tau_{i-1})], \quad G_n^4(\theta) = \sum_{i=1}^n [X_i - E(X_i | \tau_{i-1})],$$

$$G_n^5(\theta) = \sum_{i=1}^n [X_i \tau_{i-1} - \tau_{i-1} E(X_i | \tau_{i-1})], \quad G_n^6(\theta) = \sum_{i=1}^n [X_i \tau_i - E(X_i \tau_i | \tau_{i-1})],$$

$$G_n^7(\theta) = \sum_{i=1}^n [X_i^2 - E(X_i^2 | \tau_{i-1})].$$

The specifications for $G_n^1 - G_n^7$ belong to the more general class of martingale estimating functions of the form

$$G_n^j(\theta) = \sum_{i=1}^n \alpha^j(\tau_{i-1}; \theta) [X_i^{r_j} \cdot \tau_i^{s_j} - f^j(\tau_{i-1}; \theta)], \quad j = 1, \dots, d,$$

where $\alpha^j(V_{\tau_{i-1}})$ is some \mathcal{F}_{i-1} -adapted random variable,

$$f^j(\iota, \theta) = E[X_1^{r_j} \tau_1^{s_j} | \tau_0 = \iota]$$

and we have

$$f^j(\iota; \theta) = \sum_{l=0}^{r_j+s_j} \phi_l^j(\theta) \cdot \iota^l.$$

Remark

- for the simple explicit estimator $\alpha^j(\iota) = \iota$ or $\alpha^j(\iota) = 1$ which gives explicit $G_n(\theta)$, the explicit solution of $G_n(\theta) = 0$ and the explicit $Cov(\hat{\theta}_n)$.
- in this setting the problem of finding the resulting estimator explicitly amounts to **solving d explicitly given equations**.

Remark

- for the simple explicit estimator $\alpha^j(\iota) = \iota$ or $\alpha^j(\iota) = 1$ which gives explicit $G_n(\theta)$, the explicit solution of $G_n(\theta) = 0$ and the explicit $Cov(\hat{\theta}_n)$.
- in this setting the problem of finding the resulting estimator explicitly amounts to **solving d explicitly given equations.**

Outline

- 1 Barndorff-Nielsen Shepard (BNS) stochastic volatility models
 - Discretely observed continuous time model
- 2 The simple explicit estimator
 - Estimating functions
 - **The explicit solution**
 - Consistency and asymptotic normality
 - Numerical illustrations
- 3 Application of the results to real data
- 4 Future work

The simple estimator $\hat{\theta}_n = (\nu, \alpha, \lambda, \mu, \beta, \sigma, \rho)$ is given by

$$\gamma_n = (\xi_n^2 - \xi_n^1 v_n^1) / (v_n^2 - (v_n^1)^2), \quad \zeta_n = \frac{\gamma_n v_n^1 - \xi_n^1}{-1 + \gamma_n},$$

$$\eta_n = -\frac{(-1 + \gamma_n^2)(v_n^1)^2 - \gamma_n^2 v_n^2 + \xi_n^3}{-1 + \gamma_n^2}, \quad \lambda_n = -\log(\gamma_n) / \Delta,$$

$$\epsilon_n = (1 - \gamma_n) / \lambda_n, \quad \beta_n = \frac{(\xi_n^5 - v_n^1 \xi_n^4)}{\epsilon_n (v_n^2 - (v_n^1)^2)},$$

$$\rho_n = (-\beta_n \epsilon_n (-(v_n^1)^2 + \epsilon_n \lambda_n (\eta_n + (v_n^1)^2 - v_n^2) + v_n^2) - \xi_n^1 \xi_n^4 + \xi_n^6) / (2 \epsilon_n \eta_n \lambda_n),$$

$$\mu_n = (-\Delta \lambda_n \rho_n \zeta_n - \beta_n (\Delta \zeta_n + \epsilon_n (-\zeta_n + v_n^1)) + \xi_n^4) / \Delta, \quad \sigma_n = \sqrt{a_n / b_n};$$

$$a_n = \frac{4\beta_n(-\Delta + \epsilon_n)\eta_n\lambda_n\rho_n + \beta_n^2(-2\Delta\eta_n + \epsilon_n(\eta_n(2 + \epsilon_n\lambda_n) + \epsilon_n\lambda_n((v_n^1)^2 - v_n^2)))}{\lambda_n} + \frac{\lambda_n(-2\Delta\eta_n\lambda_n\rho_n^2 - (\xi_n^4)^2 + \xi_n^7)}{\lambda_n},$$

$$b_n = \Delta \zeta_n + \epsilon_n (-\zeta_n + v_n^1),$$

where

$$\begin{aligned}\xi_n^1 &= \frac{1}{n} \sum_{i=1}^n \tau_i, & \xi_n^2 &= \frac{1}{n} \sum_{i=1}^n \tau_i \tau_{i-1}, & \xi_n^3 &= \frac{1}{n} \sum_{i=1}^n \tau_i^2, \\ \xi_n^4 &= \frac{1}{n} \sum_{i=1}^n X_i, & \xi_n^5 &= \frac{1}{n} \sum_{i=1}^n X_i \tau_{i-1}, & \xi_n^6 &= \frac{1}{n} \sum_{i=1}^n X_i \tau_i, \\ \xi_n^7 &= \frac{1}{n} \sum_{i=1}^n X_i^2, & v_n^1 &= \frac{1}{n} \sum_{i=1}^n \tau_{i-1}, & v_n^2 &= \frac{1}{n} \sum_{i=1}^n \tau_{i-1}^2.\end{aligned}$$

Outline

- 1 Barndorff-Nielsen Shepard (BNS) stochastic volatility models
 - Discretely observed continuous time model
- 2 The simple explicit estimator
 - Estimating functions
 - The explicit solution
 - **Consistency and asymptotic normality**
 - Numerical illustrations
- 3 Application of the results to real data
- 4 Future work

Consistency

Theorem

$P(C_n) \rightarrow 1$ as $n \rightarrow \infty$, where $C_n = \{\xi_n^2 - \xi_n^1 v_n^1 > 0\}$ and the estimator $\hat{\theta}_n = (\nu_n, \alpha_n, \lambda_n, \mu_n, \beta_n, \sigma_n, \rho_n)$ is **consistent** on C_n , namely

$$\hat{\theta}_n \xrightarrow{\text{a.s.}} \theta_0 \quad \text{on } C_n \quad \text{as } n \rightarrow \infty.$$

Asymptotic normality

Theorem

The estimator $\hat{\theta}_n = (\nu_n, \alpha_n, \lambda_n, \mu_n, \beta_n, \sigma_n, \rho)$ is **asymptotically normal**, namely

$$\sqrt{n}[\hat{\theta}_n - \theta_0] \xrightarrow{\mathcal{D}} N(0, T) \quad \text{as } n \rightarrow \infty,$$

where

- $T = A(\theta)^{-1} \Upsilon (A(\theta)^{-1})^T, \quad \Upsilon_{ij} = E[\text{Cov}(\Xi_i^j, \Xi_1^j | \tau_0)]$

- $\Xi_k = (\tau_k, \tau_k \tau_{k-1}, \tau_k^2, X_k, X_k \tau_{k-1}, X_k \tau_k, X_k^2)^T$

- $A(\theta) \stackrel{\text{a.s.}}{=} \lim_{n \rightarrow \infty} \frac{1}{n} J_n(\theta), \quad J_n^{j,k}(\theta^{(1)}, \dots, \theta^{(d)}) = \frac{\partial G_n^j(\theta^{(l)})}{\partial \theta_k}, \quad k = 1, \dots, d.$

Asymptotic normality

Theorem

The estimator $\hat{\theta}_n = (\nu_n, \alpha_n, \lambda_n, \mu_n, \beta_n, \sigma_n, \rho)$ is **asymptotically normal**, namely

$$\sqrt{n}[\hat{\theta}_n - \theta_0] \xrightarrow{\mathcal{D}} N(0, T) \quad \text{as } n \rightarrow \infty,$$

where

- $T = A(\theta)^{-1} \Upsilon (A(\theta)^{-1})^T, \quad \Upsilon_{ij} = E[\text{Cov}(\Xi_1^i, \Xi_1^j | \tau_0)]$

- $\Xi_k = (\tau_k, \tau_k \tau_{k-1}, \tau_k^2, X_k, X_k \tau_{k-1}, X_k \tau_k, X_k^2)^T$



$$A(\theta) \stackrel{\text{a.s.}}{=} \lim_{n \rightarrow \infty} \frac{1}{n} J_n(\theta), \quad J_n^{j,k}(\theta^{(1)}, \dots, \theta^{(d)}) = \frac{\partial G_n^j(\theta^{(0)})}{\partial \theta_k}, \quad k = 1, \dots, d.$$

Asymptotic normality

Theorem

The estimator $\hat{\theta}_n = (\nu_n, \alpha_n, \lambda_n, \mu_n, \beta_n, \sigma_n, \rho)$ is **asymptotically normal**, namely

$$\sqrt{n}[\hat{\theta}_n - \theta_0] \xrightarrow{\mathcal{D}} N(0, T) \quad \text{as } n \rightarrow \infty,$$

where

- $T = A(\theta)^{-1} \Upsilon (A(\theta)^{-1})^T, \quad \Upsilon_{ij} = E[\text{Cov}(\Xi_1^i, \Xi_1^j | \tau_0)]$
- $\Xi_k = (\tau_k, \tau_k \tau_{k-1}, \tau_k^2, X_k, X_k \tau_{k-1}, X_k \tau_k, X_k^2)^T$
- $A(\theta) \stackrel{\text{a.s.}}{=} \lim_{n \rightarrow \infty} \frac{1}{n} J_n(\theta), \quad J_n^{j,k}(\theta^{(1)}, \dots, \theta^{(d)}) = \frac{\partial G_n^j(\theta^{(i)})}{\partial \theta_k}, \quad k = 1, \dots, d.$

Outline

- 1 Barndorff-Nielsen Shepard (BNS) stochastic volatility models
 - Discretely observed continuous time model
- 2 **The simple explicit estimator**
 - Estimating functions
 - The explicit solution
 - Consistency and asymptotic normality
 - **Numerical illustrations**
- 3 Application of the results to real data
- 4 Future work

BNS-ΓOU

We consider the Γ -OU model, for which the trading volume τ has a stationary $\Gamma(\nu, \alpha)$ distribution.

- The corresponding BDLP Z is a compound Poisson process with intensity ν and jumps from the exponential distribution with parameter α
- We use as time unit one year consisting of 250 trading days.

The true parameters are:

$$\nu = 6.17, \quad \alpha = 1.42, \quad \lambda = 177.95, \quad \beta = -0.015, \quad \rho = -0.00056, \quad \mu = 0.435, \quad \sigma = 0.087.$$

BNS-ΓOU

We consider the Γ -OU model, for which the trading volume τ has a stationary $\Gamma(\nu, \alpha)$ distribution.

- The corresponding BDLP Z is a compound Poisson process with intensity ν and jumps from the exponential distribution with parameter α
- We use as time unit one year consisting of 250 trading days.

The true parameters are:

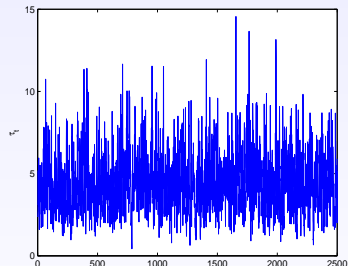
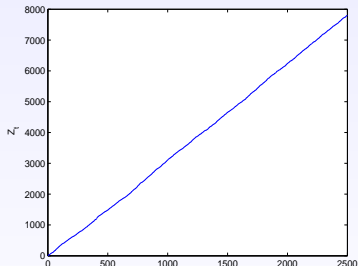
$$\nu = 6.17, \quad \alpha = 1.42, \quad \lambda = 177.95, \quad \beta = -0.015, \quad \rho = -0.00056, \quad \mu = 0.435, \quad \sigma = 0.087.$$

Interpretation

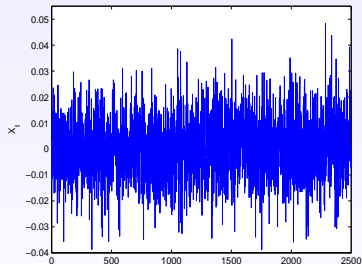
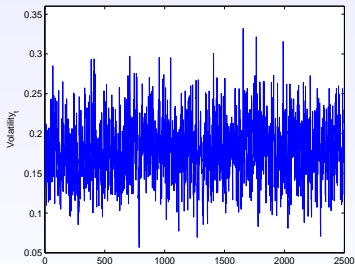
There are on average 4.4 jumps per day and the jumps in the BDLP and in the trading volume are exponentially distributed with mean and standard deviation 0.704.

- typically every day 4 or 5 new pieces of information arrive and make the trading volume process jump
- the stationary mean of the trading volume is 4.35, and of the variance is 0.033
- if we define instantaneous volatility to be the square root of the variance, it will fluctuate around 18% in our example
- The half-life of the autocorrelation of the variance process is about a day
- Annual log returns have (unconditional) mean -6.5% and annual volatility 18.2%.

The resulting time series: BDLP and trading volume



The resulting time series: volatility and log returns



The asymptotic covariance matrix of the estimator

We do not estimate the asymptotic covariance, but evaluate the explicit expression using the true parameters. Denoting the vector of asymptotic standard deviations of the estimates and the correlation matrix by s/\sqrt{n} resp. r we have

$$s = [12.0257, 2.7878, 443.85, 9.0211, 2.5536, 0.0657, 0.007]^T$$

$$r = \begin{bmatrix} 1 & 0.938 & 0.578 & 0.007 & 0.051 & 0.006 & -0.003 \\ 0.938 & 1 & 0.574 & 0.008 & 0.051 & 0.017 & -0.004 \\ 0.578 & 0.574 & 1 & 0.011 & 0.088 & -0.0006 & -6.2 \cdot 10^{-17} \\ 0.007 & 0.008 & 0.011 & 1 & -0.827 & -0.013 & 0.03 \\ 0.0511 & 0.051 & 0.088 & -0.827 & 1 & 0.012 & -0.515 \\ 0.006 & 0.013 & -0.0006 & -0.013 & 0.012 & 1 & -0.005 \\ -0.003 & -0.004 & -6.2 \cdot 10^{-17} & 0.03 & -0.515 & -0.005 & 1 \end{bmatrix}$$

Finite sample performance of the estimator

We will perform the estimation procedure for two different sample sizes, namely 2500 and 8000, corresponding to 10 years and 32 years respectively, with 250 daily observation per year.

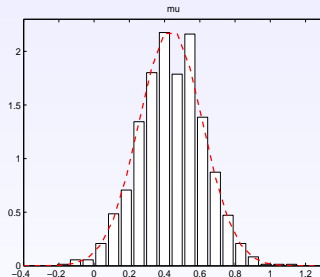
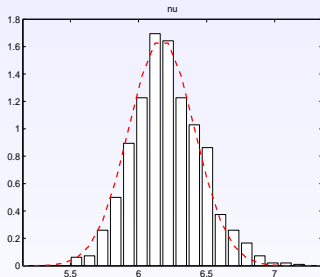
The empirical distribution of the simple estimators for the Γ -OU model is illustrated. The histograms are produced from $m = 1000$ replications consisting of $n = 2500$ observations each, corresponding to 10 years. We compare to the AN distribution

Finite sample performance of the estimator

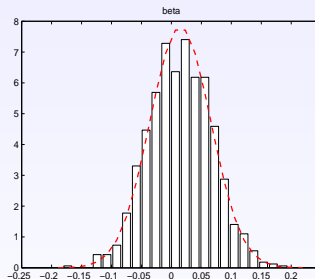
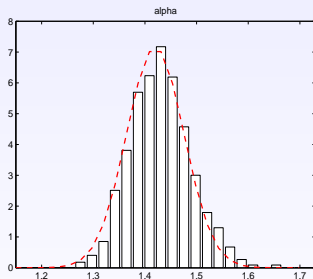
We will perform the estimation procedure for two different sample sizes, namely 2500 and 8000, corresponding to 10 years and 32 years respectively, with 250 daily observation per year.

The empirical distribution of the simple estimators for the Γ -OU model is illustrated. The histograms are produced from $m = 1000$ replications consisting of $n = 2500$ observations each, corresponding to 10 years. We compare to the AN distribution

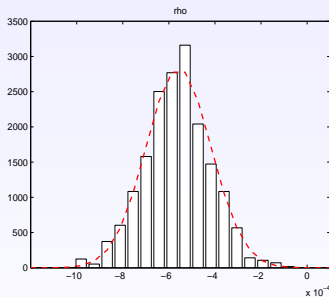
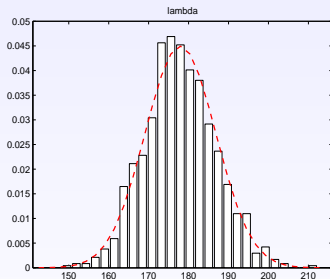
The distribution of the estimate: ν , μ



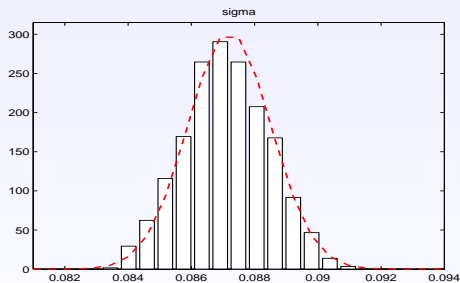
The distribution of the estimate α, β



The distribution of the estimate λ, ρ



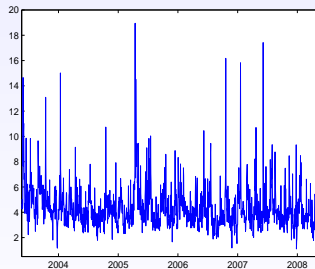
The distribution of the estimate σ



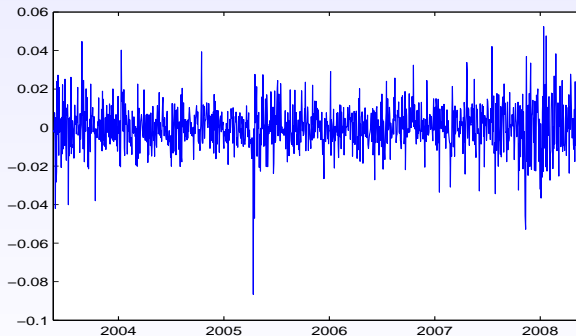
Application to daily data: the IBM stock

- The BNS model will be fitted to daily log returns of the International Business Machines Corporation (IBM) stock at the New York Stock Exchange (NYSE)
- The data spans over 5 years starting in March 23, 2003 to March 23, 2008
- There were 1259 observations of daily closing prices and trading volumes. Data on trading volumes are expressed in millions
- Sample measures of skewness and kurtosis of the returns are -0.35 and 7.42

The resulting time series: closing prices and trading volume



The resulting time series: log returns



Results and interpretation

Parameter	Value	St.dev.
$\hat{\nu}$	6.17	0.339
$\hat{\alpha}$	1.42	0.079
$\hat{\lambda}$	177.95	12.509
$\hat{\mu}$	0.435	0.254
$\hat{\beta}$	-0.015	0.072
$\hat{\sigma}$	0.087	0.002
$\hat{\rho}$	-0.00056	0.0002

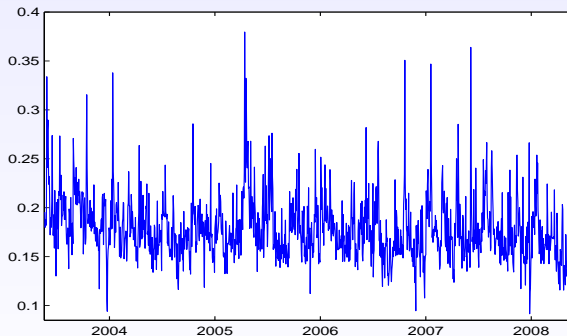
Results and interpretation

Unconditional moments	Value
$E[X]$	-0.027%
$St.dev[X]$	1.15%
$E[V]$	3.3%
$St.dev[V]$	1.32%

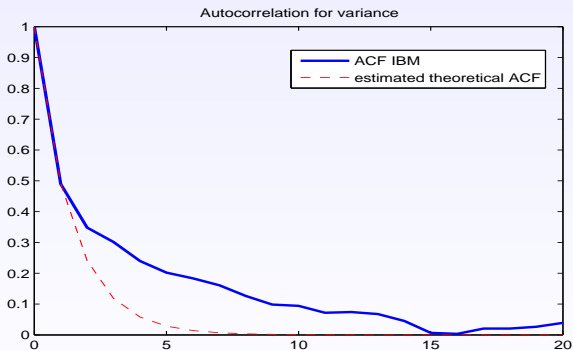
Results and interpretation

- There are on average 4.4 jumps per day each with mean and standard deviation 0.704
- Typical volatility is 0.18 with standard deviation 0.11
- The proportionality of trading volume and the instantaneous variance is given by $\sigma^2 = 0.0076$
- The leverage ρ is very small. For instance, the mean values of daily log-returns including or not a leverage effect in the model equal -0.027% , and 0.146% respectively. If the trading volume process jumps by a typical size, the returns jump by 0.0004.

Estimation of the volatility



The autocorrelation function for the variance process and the estimated theoretical autocorrelation



Model fit investigation

- We performed a Ljung-Box test for squared residuals of the data set
- We test the normality of the *residuals*
- What are exactly the *residuals* in the BNS setting?

Model fit investigation

- We performed a Ljung-Box test for squared residuals of the data set
- **We test the normality of the *residuals***
- What are exactly the *residuals* in the BNS setting?

Model fit investigation

- We performed a Ljung-Box test for squared residuals of the data set
- We test the normality of the *residuals*
- **What are exactly the *residuals* in the BNS setting?**

Estimated residuals

From $X_i = \mu\Delta + \beta Y_i + \sigma\sqrt{Y_i}W_i + \rho Z_i$, it follows that

$$\hat{W}_i = (X_i - \hat{\mu}\Delta - \hat{\beta}Y_i - \hat{\rho}Z_i) / \hat{\sigma}\sqrt{Y_i}, \quad i \in \mathbb{N}$$

Since, Z_i is usually not observable, we have to approximate it. For the integral we use simple Euler approximations

$$Y_i = \int_{t_{i-1}}^{t_i} \tau(s-) ds \approx \tau_i \Delta \quad \text{and} \quad \int_{t_{i-1}}^{t_i} \sqrt{\tau(s)} dW(s) \approx \sqrt{\tau_i \cdot \Delta} \varepsilon, \quad (3)$$

with $\varepsilon \sim N(0, 1)$.

The residuals

$$\hat{\varepsilon}_i = (X_i - \hat{\mu}\Delta - \hat{\beta}\tau_i\Delta - \hat{\rho}Z_i) / \hat{\sigma}\sqrt{\tau_i \cdot \Delta}$$

where

$$Z_i = (\lambda\Delta + 1)\tau_i - \tau_{i-1}.$$

The estimated mean, standard deviation, skewness and kurtosis of the residuals:

	mean($\hat{\varepsilon}$)	std($\hat{\varepsilon}$)	skew($\hat{\varepsilon}$)	kurt($\hat{\varepsilon}$)
IBM	-0.01568	1.03142	-0.17053	5.40659

The residuals

$$\hat{\varepsilon}_i = (X_i - \hat{\mu}\Delta - \hat{\beta}\tau_i\Delta - \hat{\rho}Z_i) / \hat{\sigma}\sqrt{\tau_i \cdot \Delta}$$

where

$$Z_i = (\lambda\Delta + 1)\tau_i - \tau_{i-1}.$$

The estimated mean, standard deviation, skewness and kurtosis of the residuals:

	mean($\hat{\varepsilon}$)	std($\hat{\varepsilon}$)	skew($\hat{\varepsilon}$)	kurt($\hat{\varepsilon}$)
IBM	-0.01568	1.03142	-0.17053	5.40659

Model fit - continued

- Ljung-Box test: the test statistic used 35 lags of the corresponding empirical autocorrelation function
- The test statistic for the IBM squared residuals was equal to 451.61, which led to a rejection of the null hypothesis, since the test had a critical value of 113.15 at the 0.05 level
- The IBM residuals pass the Kolmogorov-Smirnov test of normality, for example, with p -value 0.0886,

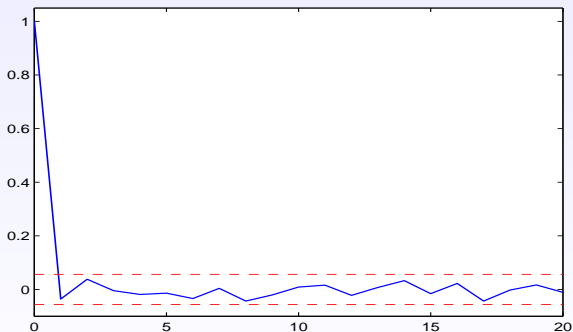
Model fit - continued

- Ljung-Box test: the test statistic used 35 lags of the corresponding empirical autocorrelation function
- The test statistic for the IBM squared residuals was equal to 451.61, which led to a rejection of the null hypothesis, since the test had a critical value of 113.15 at the 0.05 level
- The IBM residuals pass the Kolmogorov-Smirnov test of normality, for example, with p -value 0.0886,

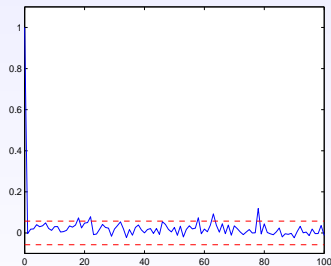
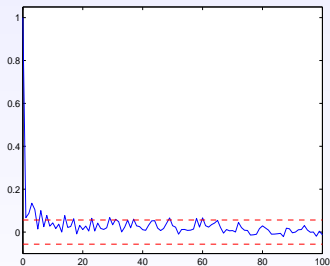
Model fit - continued

- Ljung-Box test: the test statistic used 35 lags of the corresponding empirical autocorrelation function
- The test statistic for the IBM squared residuals was equal to 451.61, which led to a rejection of the null hypothesis, since the test had a critical value of 113.15 at the 0.05 level
- The IBM residuals pass the Kolmogorov-Smirnov test of normality, for example, with p -value 0.0886,

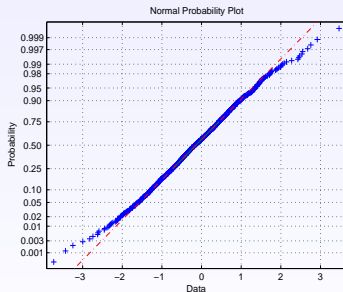
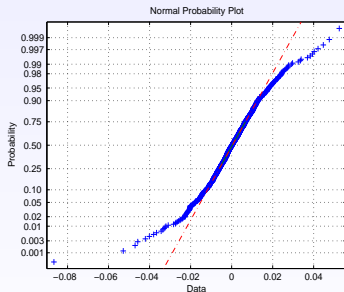
The empirical autocorrelation function for the residuals



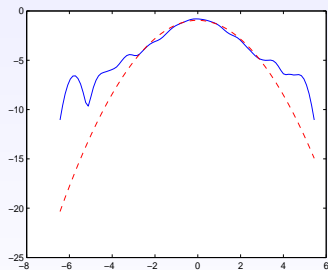
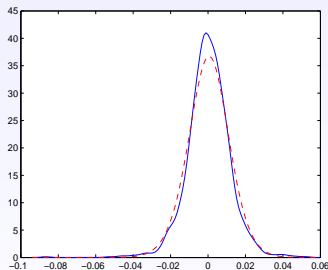
The empirical autocorrelation function for the squared log returns and squared residuals for IBM



The normal probability plot of log returns and *residuals* for IBM



Kernel estimates of the density and log density with the theoretical one for IBM



Estimated mean, mean square error (MSE) and mean absolute error (MAE) of all the estimated parameter values, with the empirical standard deviations in brackets.

$n = 2500$	ν_n	α_n	λ_n
Mean	6.2145 (0.2552)	1.435 (0.0588)	177.865 (8.9257)
MSE	0.0672 (0.1046)	0.0036 (0.0055)	79.5956 (115.6766)
MAE	0.2016 (0.1629)	0.047 (0.0369)	7.0692 (5.4454)
$n = 8000$	ν_n	α_n	λ_n
Mean	6.1642 (0.1424)	1.4186 (0.0329)	177.1342 (5.208)
MSE	0.0203 (0.0283)	0.0011 (0.0016)	27.7663 (39.1414)
MAE	0.1135 (0.0862)	0.0259 (0.0203)	4.2191 (3.1584)

Estimated mean, mean square error (MSE) and mean absolute error (MAE) of all the estimated parameter values, with the empirical standard deviations in brackets.

μ_n	β_n	σ_n	ρ_n
0.4402 (0.1849)	0.0148 (0.053)	0.0871 (0.001)	$-6 \cdot 10^{-4}$ ($1 \cdot 10^{-4}$)
0.0342 (0.0492)	0.0028 (0.0039)	$2 \cdot 10^{-6}$ ($2 \cdot 10^{-6}$)	$2 \cdot 10^{-8}$ ($3 \cdot 10^{-8}$)
0.1483 (0.1105)	0.0428 (0.0313)	0.0011 (0.001)	$1 \cdot 10^{-4}$ ($9 \cdot 10^{-5}$)
μ_n	β_n	σ_n	ρ_n
0.4388 (0.1002)	0.0129 (0.03)	0.1 ($7 \cdot 10^{-4}$)	$-6 \cdot 10^{-4}$ ($8.07 \cdot 10^{-5}$)
0.01 (0.0138)	0.0008 (0.001)	$6 \cdot 10^{-7}$ ($8 \cdot 10^{-7}$)	$7 \cdot 10^{-9}$ ($1 \cdot 10^{-8}$)
0.08 (0.0604)	0.0222 (0.02)	$6 \cdot 10^{-4}$ ($5 \cdot 10^{-4}$)	$6 \cdot 10^{-5}$ ($5.04 \cdot 10^{-5}$)

Further issues

- Instantaneous variance is not an observable quantity in discrete time and various quantities are suggested as substitutes for the variance
- the resulting estimating function will not be a martingale estimating function anymore and the bias has to be accounted for in a rigorous analysis
- superposition of OU-processes
- comparison with the GMM estimation procedure (or combination of estimation procedures)

Further issues

- Instantaneous variance is not an observable quantity in discrete time and various quantities are suggested as substitutes for the variance
- the resulting estimating function will not be a martingale estimating function anymore and the bias has to be accounted for in a rigorous analysis
- superposition of OU-processes
- comparison with the GMM estimation procedure (or combination of estimation procedures)

Further issues

- Instantaneous variance is not an observable quantity in discrete time and various quantities are suggested as substitutes for the variance
- the resulting estimating function will not be a martingale estimating function anymore and the bias has to be accounted for in a rigorous analysis
- superposition of OU-processes
- comparison with the GMM estimation procedure (or combination of estimation procedures)

Further issues

- Instantaneous variance is not an observable quantity in discrete time and various quantities are suggested as substitutes for the variance
- the resulting estimating function will not be a martingale estimating function anymore and the bias has to be accounted for in a rigorous analysis
- superposition of OU-processes
- comparison with the GMM estimation procedure (or combination of estimation procedures)

Bibliography



Barndorff-Nielsen, Ole E. and Shephard, Neil,
Non-Gaussian Ornstein-Uhlenbeck-based models and some of their
uses in financial economics,
J. R. Stat. Soc. Ser. B Stat. Methodol., 63(2):167–241, 2001,



Jun Pan,
The jump-risk premia implicit in options: evidence from an integrated
time-series study,
J. Financial Economics, 63 : 3–50, 2002



Sørensen, Michael
On asymptotics of estimating functions
Brazilian J. Prob. and Stat., 13: 111-136, 1999,

Bibliography



Barndorff-Nielsen, Ole E. and Shephard, Neil,
Non-Gaussian Ornstein-Uhlenbeck-based models and some of their
uses in financial economics,
J. R. Stat. Soc. Ser. B Stat. Methodol., 63(2):167–241, 2001,



Jun Pan,
The jump-risk premia implicit in options: evidence from an integrated
time-series study,
J. Financial Economics, 63 : 3–50, 2002



Sørensen, Michael
On asymptotics of estimating functions
Brazilian J. Prob. and Stat., 13: 111-136, 1999,