On Trades, Volume, and the Martingale Estimating Function Approach for Stochastic Volatility Models with Jumps

Friedrich Hubalek (Joint work with Petra Posedel)

Johann Radon Institute for Computational and Applied Mathematics (RICAM)
Special Semester on Stochastics with Emphasis on Finance
Concluding Workshop
Linz, December 2–4, 2008.

Our papers

- Friedrich Hubalek and Petra Posedel, Joint analysis and estimation of stock prices and trading volume in Barndorff-Nielsen and Shephard stochastic volatility models, arXiv:0807.3464 (July 2008)
- Friedrich Hubalek and Petra Posedel, Asymptotic analysis for a simple explicit estimator in Barndorff-Nielsen and Shephard stochastic volatility models, arXiv:0807.3479 (July 2008)
- Friedrich Hubalek and Petra Posedel, Asymptotic analysis for an optimal estimating function for Barndorff-Nielsen-Shephard stochastic volatility models, Work in progress.

The Barndorff-Nielsen and Shephard stochastic volatility models with jumps

- Logarithmic returns (discounted)

$$dX(t) = (\mu + \beta V(t-))dt + \sqrt{V(t-)}dW(t) + \rho dZ_\lambda(t)$$

- Instantaneous variance

$$dV(t) = -\lambda V(t-)dt + dZ_\lambda(t)$$

  $W\ldots$ Brownian motion, $Z\ldots$ subordinator,
  $Z_\lambda(t) = Z(\lambda t)$ [...]

- Parameters: $\mu \in \mathbb{R}\ldots$ linear drift, $\beta \in \mathbb{R}\ldots$ Itô drift, $\rho \in \mathbb{R}\ldots$ leverage, $\lambda > 0\ldots$ acf parameter.

Analytical tractability

- $(X(t), V(t), t \geq 0)$... Markov, affine model (in continuous time)
- simple Riccati-type equations for characteristic resp. moment generating function
- general solution (up to one integral)
- Γ-OU and IG-OU completely explicitly in terms of elementary functions

Exploited in

- Option pricing (Nicolato and Venardos)
- Portfolio optimization (Benth et al.)
- Minimum entropy martingale measure (Benth et al., Rheinländer and Steiger)
- Semimartingal Esscher transform (Hubalek and Sgarra)
- ...

But statistical inference seems difficult! Bayesian, MCMC — computer intensive approaches!

- Barndorff-Nielsen O.E., Shephard N. (2001), Non-Gaussian Ornstein-Uhlenbeck-based models and some of their uses in financial economics.
- Roberts G.O., Papaspiliopoulos O., Dellaportas P. (2004), Bayesian inference for non-Gaussian Ornstein-Uhlenbeck stochastic volatility processes,
- J.E. Griffin, M.F.J. Steel (2006), Inference with non-Gaussian Ornstein-Uhlenbeck processes for stochastic volatility
- Matthew P.S. Gandera and David A. Stephens (2007), Stochastic volatility modelling in continuous time with general marginal distributions: Inference, prediction and model selection
- Sylvia Frühwirth-Schnatter and Leopold Sögner (2007?), Bayesian estimation of stochastic volatility models based on OU processes with marginal Gamma laws.

Discrete observations

Grid $t_i = i\delta$, $i \geq 0$, fixed width $\Delta > 0$, discrete time observations

$$X_i = X(t_i) - X(t_{i-1}), \qquad V_i = V(t_i)$$

Discrete dynamics

$$X_i = \mu\Delta + \beta Y_i + \sqrt{Y_i} W_i + \rho Z_i, \qquad V_i = e^{\lambda\Delta} V_{i-1} + U_i$$

Auxiliary quantities (no discretization error!)

$$Z_i = Z_\lambda(t_i) - Z_\lambda(t_{i-1}), \quad U_i = \int_{t_{i-1}}^{t_i} e^{-\lambda(t_i - s)} dZ_\lambda(s)$$

and

$$Y_i = \int_{t_{i-1}}^{t_i} V(s-) ds, \qquad W_i = \frac{1}{\sqrt{Y_i}} \int_{t_{i-1}}^{t_i} \sqrt{V(s-)} dW(s).$$

$(X_i, V_i, i \in \mathbb{N})\ldots$ Markov affine model (in discrete time)

Construction and moments

Two starting points

- L ... infinitely divisible distribution on $\mathbb{R}_+ \Rightarrow$ subordinator Z with $Z(1) \stackrel{d}{=} L \Rightarrow$ (OU-L)
- D ... self-decomposable distribution on $\mathbb{R}_+ \Rightarrow$ stationary Ornstein-Uhlenbeck process V with $V(t) \stackrel{d}{=} D \Rightarrow$ (D-OU)

Moments of D resp. $L \rightarrow$ all (mixed, conditional, unconditional) integer moments by simple algebra (multivariate Faa di Bruno formula resp. Bell polynomials, practical calculations best by recursions!)

$$E[X_i^n], E[V_i^n], E[X_i^m V_i^n], E[X_i^\ell V_i^m V_{i-1}^n],$$
$$E[X_i^n | V_{i-1}], E[V_i^n | V_{i-1}], E[X_i^m V_i^n | V_{i-1}], \ldots$$

$\Rightarrow$ method of moments estimation

Various methods of moments

- Method of moments — MM (Pearson 1893)
- Generalized method of moments — GMM (Hansen 1982)
- Simulated method of moments — SMM (. . .)
- Efficient method of moments — EMM (Gallant and Tauchen 1996),
- . . .
- [Methods of moments for weak convergence]

Estimation: Setting and problems

Grid, fixed width, horizon (number of observations) going to Infinity for asymptotics! (Several other possibilities...)

- Rich, well-informed financial institutions and traders observe and trade in continous-time
- Poor, academic statisticians and econometers do inference with daily (or less frequent!) observations
- [But: High-frequence analyses ...]

Discrete time observations $\Rightarrow V_i$ not observed, BNS becomes non-Markovian, (a hidden Markov model)!

Remedies

- Substitute unobserved $V_i \mapsto$ model-implied $\hat{V}_i$ from option data, i.e., joint analysis of P and Q. Cf.
  - Jun Pan, The Jump-Risk Premia Implicit in Options: Evidence from an Integrated Time-Series Study (2002).

  (GMM, realistic, complicated, many assumptions.)
  Also our long term goal!
- Ignore the problem. Purely theoretical study, exhibits methodology, provides an upper bound for the accuracy for this type of methods. See our first paper!
- NOW: Substitute unobserved V_i by an observable proxy, volume or number of trades.

Prices, volatility, trading intensity

Our incentive

- Carl Lindberg, The estimation of the Barndorff-Nielsen and Shephard model from daily data based on measures of trading intensity. Applied Stochastic Models in Business and Industry 24 (4), 2008.

Some earlier/classical references

- J. M. Karpoff, The relation between price changes and trading volume: a survey. JFQA 22, 1987.
- R.P.E. Gallant, A.R. and G. Tauchen, Stock prices and volume, Rev.Fin.Stud. 5:199–242, 1992.
- K.G. Jones, C. and M.L. Lipson, Transactions, volume and volatility. Rev.Fin.Stud. 7:631–651, 1994.
- G.E. Tauchen and M.Pitts, The Price Variability-Volume Relationship on Speculative Markets Econometrica 51,(1983).

The new variant/interpretation of the BNS models

Bold simplification/assumption: Instantaneous variance IS a (constant) multiple of the trading volume resp. number of trades. Introduce a proportionality parameter $\sigma > 0$. [...]

- Logarithmic returns

$$dX(t) = (\mu + \beta V(t-))dt + \sigma \sqrt{V(t-)} dW(t) + \rho dZ_\lambda(t)$$

- Trading volume (or number of trades)

$$dV(t) = -\lambda V(t-)dt + dZ_\lambda(t)$$

  W ... Brownian motion, Z ... subordinator,
  $Z_\lambda(t) = Z(\lambda t)$ [...]

- Parameters: $\mu \in \mathbb{R}$ ... linear drift, $\beta \in \mathbb{R}$ ... Itô drift, $\sigma > 0$ ... proportionality, $\rho \in \mathbb{R}$ ... leverage, $\lambda > 0$ ... acf parameter.

What about maximum likelihood ?

- Practical issue: Bivariate Markov, known transition probability (in terms of characteristic resp. cumulant function) $\Rightarrow$ invert for each observation in each iterations [Possible remedies, approximate inversions, LeCam's trick,...]
- Theoretical issue: For infinite activity BDLP (e.g., IG-OU) fine, for finite activity (e.g., Γ-OU with exponential compound Poisson BDLP)

$$P_\lambda[V_1 = ve^{-\lambda\Delta} | V_0 = v] = e^{-\lambda\Delta} \quad \text{(no jump)}$$

  $\Rightarrow$ No dominating sigma-finite measure! $\Rightarrow$ Usual ML framework does not apply!
  - Generalized ML (Kiefer and Wolfowitz 1956) [...]
  - Much better than $\sqrt{n}$ by ad hoc (?) methods! [...]

Martingale estimating functions

E.g., $\Gamma(\nu,\alpha)$-OU: Parameter vector $(3+4=7)$

$$\theta = (\lambda, \nu, \alpha, \mu, \beta, \sigma, \rho)$$

Moments

$$\Xi_i = (V_i, V_i V_{i-1}, V_i^2, X_i, X_i V_{i-1}, X_i V_i, X_i^2), \qquad \Upsilon_i = (V_{i-1}, V_{i-1}^2)$$

Martingale estimating function

$$G_n(\theta) = \frac{1}{n}\sum_{i=1}^{n} \left[\Xi_i - f(V_{i-1}, \theta)\right], \qquad f(v,\theta) = E_\theta[\Xi_1 | V_0 = v]$$

Estimator: Solve $G_n(\theta) = 0$! Sample moments

$$\xi_n = \frac{1}{n}\sum_{i=1}^{n} \Xi_i, \qquad \upsilon_n = \frac{1}{n}\sum_{i=1}^{n} \Upsilon_i,$$

The explicit estimator

Unique solution exists on $C_n = \left\{\xi_n^2 - \xi_n^1 \upsilon_n^1 > 0, \upsilon_n^2 - (\upsilon_n^1)^2 > 0\right\}$ and is given by

$$\gamma_n = (\xi_n^2 - \xi_n^1 \upsilon_n^1)/(\upsilon_n^2 - (\upsilon_n^1)^2); \quad \zeta_n = \frac{\gamma_n \upsilon_n^1 - \xi_n^1}{-1 + \gamma_n} \quad \lambda_n = -\log(\gamma_n)/\Delta;$$

$$\eta_n = -\frac{(-1+\gamma_n^2)(\upsilon_n^1)^2 - \gamma_n^2 \upsilon_n^2 + \xi_n^3}{-1+\gamma_n^2} \quad \epsilon_n = (1-\gamma_n)/\lambda_n;$$

$$\beta_n = \frac{(\xi_n^5 - \upsilon_n^1 \xi_n^4)}{\epsilon_n(\upsilon_n^2 - (\upsilon_n^1)^2)};$$

$$\rho_n = \big(-\beta_n \epsilon_n(-(\upsilon_n^1)^2 + \epsilon_n \lambda_n(\eta_n + (\upsilon_n^1)^2 - \upsilon_n^2) + \upsilon_n^2) - \xi_n^1 \xi_n^4 + \xi_n^6\big)/(2\epsilon_n \eta_n \lambda_n);$$

$$\mu_n = \big(-\Delta \lambda_n \rho_n \zeta_n - \beta_n(\Delta \zeta_n + \epsilon_n(-\zeta_n + \upsilon_n^1)) + \xi_n^4\big)/\Delta;$$

$$\sigma_n = \sqrt{a_n/b_n}; \quad b_n = \Delta \zeta_n + \epsilon_n(-\zeta_n + \upsilon_n^1);$$

$$a_n = \lambda_n^{-1}\Big[4\beta_n(-\Delta + \epsilon_n)\eta_n \lambda_n \rho_n + \beta_n^2(-2\Delta \eta_n + \epsilon_n(\eta_n(2 + \epsilon_n \lambda_n) + \epsilon_n \lambda_n((\upsilon_n^1)^2 - \upsilon_n^2))) + \lambda_n(-2\Delta \eta_n \lambda_n \rho_n^2 - (\xi_n^4)^2 + \xi_n^7)\Big];$$

Structure: First $\lambda_n, \nu_n, \alpha_n$ are simple AR(1) estimators, then μ_n, β_n, ρ_n from a simple linear system, finally σ_n from a quadratic equation.

Consistency

The basic (and only!) assumption: V_0 self-decomposable rv on $\mathbb{R}_+$ with

$$E[V_0^n] < \infty \quad \forall n \in \mathbb{N}.$$

The basic convergence result

$$\frac{1}{n}\sum_{i=1}^{n} X_i^p V_i^q V_{i-1}^r \xrightarrow{a.s.} E[X_1^p V_1^q V_0^r] \quad \forall p, q, r \in \mathbb{N}.$$

Remark: Ergodicity vs. simple proof. Martingale differences $\Rightarrow$ uncorreclated $\Rightarrow$ elementary convergence result.

Theorem

We have $P(C_n) \to 1$ and the estimator $\hat{\theta}_n$ is consistent on C_n, namely

$$\hat{\theta}_n I_{C_n} \xrightarrow{a.s.} \theta_0$$

as $n \to \infty$.

Asymptotic normality — delta method

- Explicit estimator ⇒ Delta-Method applies: Sample moments

$$(\xi_n, \upsilon_n) \xrightarrow{\mathcal{D}} N(0, \Sigma)$$

  estimator

$$\theta_n = h(\xi_n, \upsilon_n)$$

  result

$$\sqrt{n}(\theta_n - \theta_0) \xrightarrow{\mathcal{D}} N(0, T) \qquad T = J\Sigma J^\top$$

  Jacobian $J = \nabla h$. Messy.
- Better: General framework (implicit function theorem)
  - Michael Sørensen, Statistical inference for discretely observed diffusions, Lecture Notes, Berlin, 1997.
  - Michael Sørensen, On asymptotics of estimating functions, Brazil. J. Prob. Stat. (1999).

  Also when estimating functions $G_n(\theta)$ explicit, but estimator θ_n is not [...optimal estimating functions].

Asymptotic normality — general framework

Basic result: asymptotic normality of estimating functions

$$\frac{1}{\sqrt{n}} G_n(\theta_0) \xrightarrow{\mathcal{D}} N(0, \Lambda), \qquad \Lambda = E[\mathrm{Var}[\Xi_1] | V_0]$$

Proof by multivariate martingale central limit theorem.

Theorem

The estimator $\theta_n I_{C_n}$ is asymptotically normal, namely

$$\sqrt{n}\big[\hat{\theta}_n - \theta_0\big] \xrightarrow{\mathcal{D}} N(0, T), \qquad T = A^{-1}\Lambda(A^{-1})^\top$$

as $n \to \infty$, with Jacobian $A = E[\nabla f(V_0, \theta_0)]$.

- Recall $f(v, \theta) = E_\theta[\Xi_1 | V_0 = v]$ and $E = E_{\theta_0}$.
- Matrices A and Λ simple, explicit, (slightly lengthy).

Finite sample performance — the controlled simulation experiment

Γ-OU: Volume $V(t) \sim \Gamma(\nu, \alpha)$ stationary, BDLP Z compound Poisson, intensity λ exponential jumps with mean $1/\alpha$.

- Parameter values (annual, 250 trading days)

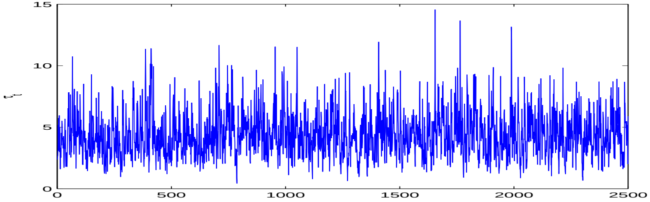
$$\nu = 6.17, \quad \alpha = 1.42, \quad \lambda = 177.95,$$

$$\beta = -0.015, \quad \rho = -0.00056, \quad \mu = 0.435, \quad \sigma = 0.087.$$

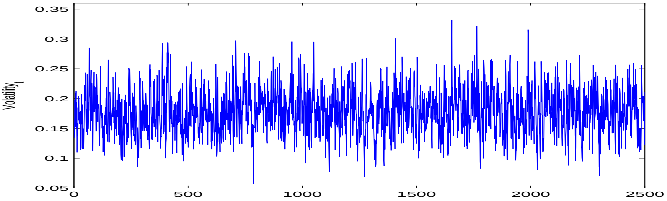
- BDLP: 4.4 jumps per day (interesting pieces of news arriving?), each jump with mean and stddev 0.704.
- Volume (in Mio): Stationary mean 4.35, variance 0.033 $\Rightarrow$ Volatility $\approx$ 18%. ACF half-life $\approx$1 day.
- Log returns: Mean -6.5%, volatility 18%.
- Experiments: n=2500 (10 years), $n = 8000$ (32 years, theoretical check).

Simulated paths 1

Volume

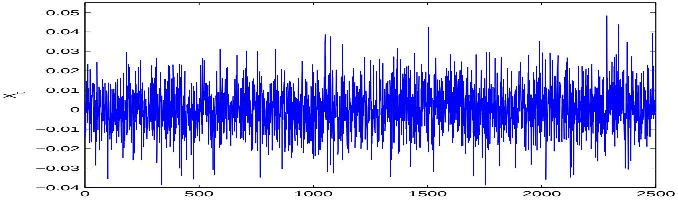


Volatility



Simulated paths 2

Returns



Asymptotic performance

- True values $\theta = (\nu, \alpha, \lambda, \beta, \rho, \mu, \sigma)$

$$\theta = (6.17, 1.42, 177.95, -0.015, -0.00056, 0.435, 0.087)$$

- Asymptotic stddev $s/\sqrt{n}$

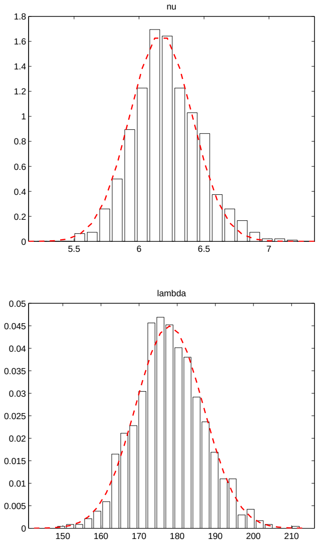
$$s = (12.0, 2.8, 440, 9.0, 2.6, 0.066, 0.007)$$

- Asymptotic correlation r

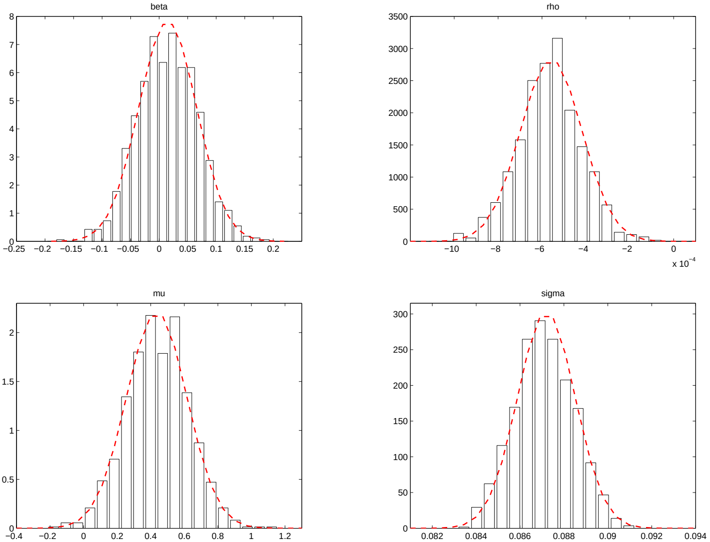
$$r = \begin{bmatrix} 1 & 0.9 & 0.6 & 0.007 & 0.05 & 0.006 & -0.003 \\ 0.9 & 1 & 0.6 & 0.007 & 0.05 & 0.01 & -0.004 \\ 0.6 & 0.6 & 1 & 0.01 & 0.09 & -0.0006 & 0.00 \\ 0.007 & 0.008 & 0.01 & 1 & -0.8 & -0.01 & 0.03 \\ 0.05 & 0.05 & 0.09 & -0.8 & 1 & 0.01 & -0.5 \\ 0.006 & 0.01 & -0.0006 & -0.01 & 0.01 & 1 & -0.005 \\ -0.003 & -0.004 & 0.00 & 0.03 & -0.5 & -0.005 & 1 \end{bmatrix}$$

- Big r in AR(1)-part! $\Rightarrow$ Optimal estimating function.

Histograms: $m = 1000$ replications, each $n = 2500$ observations, volume parameters



Histograms : $m = 1000$ replications, each $n = 2500$ observations, return parameters



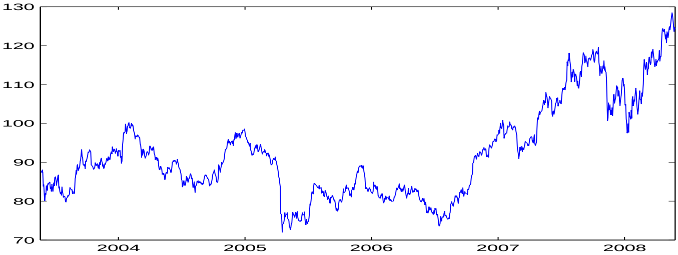
A first empirical analysis — data

Closing price and volume

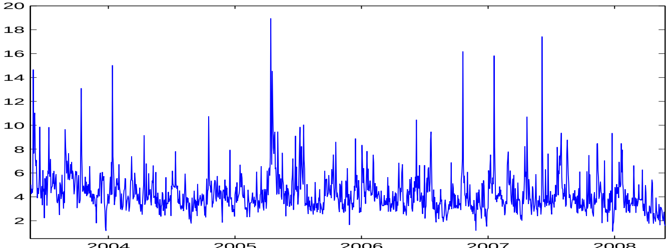
- IBM: March 23, 2003 – March 23, 2008 [NYSE], 1259 observations
- MSFT: April 11, 2003 – Feb 4, 2008 [Nasdaq], 1212 observations

IBM data

Price

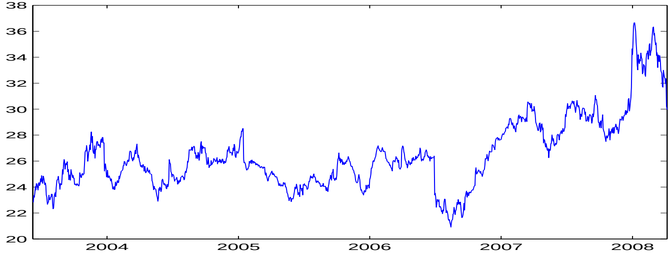


Volume

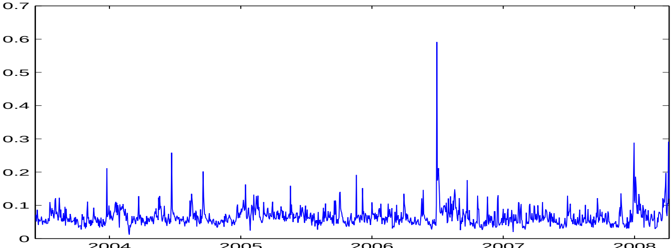


MSFT data

Price



Volume



Estimation results

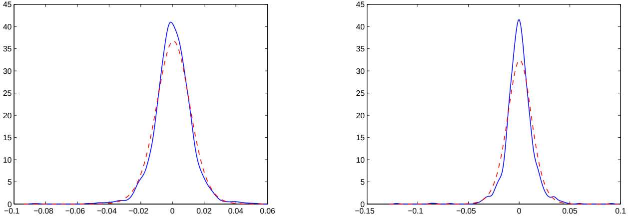
	IBM	stddev
$\hat{\nu}$	6.17	0.339
$\hat{\alpha}$	1.42	0.079
$\hat{\lambda}$	177.95	12.509
$\hat{\mu}$	0.435	0.254
$\hat{\beta}$	-0.015	0.072
$\hat{\sigma}$	0.087	0.002
$\hat{\rho}$	-0.00056	0.0002

	MSFT	stddev
$\hat{\nu}$	4.496	0.247
$\hat{\alpha}$	67.895	3.773
$\hat{\lambda}$	201.99	14.420
$\hat{\mu}$	0.4162	0.265
$\hat{\beta}$	-0.464	5.059
$\hat{\sigma}$	0.81	0.018
$\hat{\rho}$	-0.025	0.013

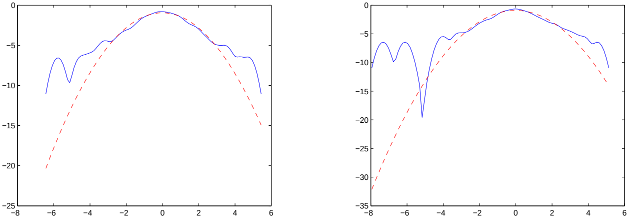
Interpretation: [. . .]

Unconditional return distributions

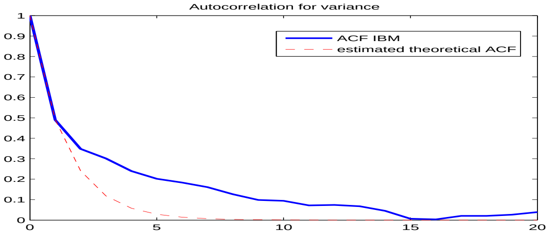
Theoretical BNS (dashed) versus kernel estimates (solid)

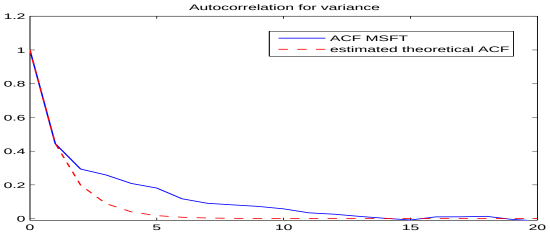


Log densities



Autocorrelation function (volume)





$\Rightarrow$ BNS with Superposition of OU-processes [...]

Model fit — residual analysis

- Volume: Usual (and exact) AR(1) analysis, though with funny innovations (U_i) iid,

$$V_i - e^{-\lambda\Delta} = U_i, \qquad U_i = \int_{t_{i-1}}^{t_i} e^{-\lambda(t_i - s)} dZ_\lambda(s)$$

- Returns: Not exact (?), Euler approximation

$$\dots$$

Further developments and directions 1

Superposition

$$V(t) = w_1 V_1(t) + \cdots + w_m V_m(t), \quad dV_i(t) = -\lambda_i V_i(t-)dt + dZ_i(\lambda_i t)$$

$(X, V_1, \ldots, V_m)$ Markov affine $\Rightarrow$ Observations?

- $V_1 \ldots$ common factor (market volume,...)
- $V_2 \ldots$ idiosyncratic factor (asset volume,...)
- $V_3 \ldots$? (similar asset? ...?)

Further developments and directions 2

- Number of trades (Lindberg!)
- Optimal martingale estimating functions

$$G_n(\theta) = \sum_{i=1}^{n} B(\theta, V_{i-1}) \left[\Xi_i - f(\theta, V_{i-1})\right) \quad f(\theta, v) = E_\theta[\Xi_i | V_{i-1} = v]$$

- Comparison with ML and related methods (for infinite activity)
- Comparison with GMM
- Hybrid approaches
- Other moments (trigonometric, c.f., Singleton, ...)
- Other time-scales (!!!)
- Integrated analysis for asset and derivatives