High-frequency Stochastic Volatility Models

Toshiaki Watanabe (Hitotsubashi University)
Jouchi Nakajima (Bank of Japan)

The 5th Hitotsubashi Summer Institute (HSI2019)
Macro- and Financial Econometrics
Hitotsubashi University
August 3, 2019
The views expressed in this research are those of the authors and do not necessarily reflect the official views of the Bank of Japan.
Financial volatility

- Financial volatility is the standard deviation or variance of financial returns.
- It plays an important role in financial risk management.
- It is time-varying.
- How to estimate time-varying volatility.
  1. GARCH and stochastic volatility (SV) models
  2. Realized volatility (RV)
Realized volatility (RV)

- RV is the estimator of volatility using intraday high-frequency returns.
- The plain RV is the sum of squared intraday returns.
- Since the plain RV is subject to the bias caused by microstructure noise, several methods for mitigating the bias have been proposed.
- Such methods can’t beat the RV calculated using 5-min returns (Liu, Patton and Sheppard 2015).
- RV does not require us to model the dynamics of daily volatility.
- RV does not require us to model the dynamics of intraday volatility.
Realized volatility (RV)

- RV is the estimator of volatility using intraday high-frequency returns.
- The plain RV is the sum of squared intraday returns.
- Since the plain RV is subject to the bias caused by microstructure noise, several methods for mitigating the bias have been proposed.
- Such methods can’t beat the RV calculated using 5-min returns (Liu, Patton and Sheppard 2015).
- RV does not require us to model the dynamics of daily volatility.
- RV does not require us to model the dynamics of intraday volatility.
Introduction

- Can we beat RV by modelling the dynamics of intraday volatility?

- It is not straightforward to model the dynamics of intraday volatility.
  
  (1) Intraday seasonality
  
  (2) Announcement effects
Introduction

This research

- extends daily SV models to intraday SV models.
- develops a Bayesian method using MCMC for the analysis of intraday SV models.
- applies them to 5-min returns of Nikkei 225 stock index.
- shows
  1. Intraday SV models fit the data better than intraday GARCH-type models.
  2. Intraday SV models perform better than daily RV models such as HAR and realized EGARCH models in one-day-ahead volatility forecasting.
- Stroud and Johaness (2014) conduct a similar analysis.
## Difference of this research from Stroud and Johaness (2014)

<table>
<thead>
<tr>
<th></th>
<th>SJ</th>
<th>This research</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data</strong></td>
<td>S&amp;P500 E-mini futures (24-hour trading)</td>
<td>Nikkei 225 (Overnight and lunch-time breaks)</td>
</tr>
<tr>
<td><strong>Intraday seasonality</strong></td>
<td>Cubic smoothing spline</td>
<td>Random-walk</td>
</tr>
<tr>
<td><strong>Sampling SV</strong></td>
<td>Mixture sampler (Approximate)</td>
<td>Block sampler (Exact)</td>
</tr>
</tbody>
</table>
Intraday SV models

Intraday high-frequency return

\[ y_t = V_t \varepsilon_t + J_t Z_t^y, \quad \varepsilon_t \sim N(0, 1) \text{ or standardized } t(\nu). \]

- \( V_t \) = total volatility.
- \( J_t Z_t^y \) = jump component.

Standardized \( t(\nu) \)

\[ \varepsilon_t = \sqrt{\lambda_t} z_t, \quad \frac{\nu - 2}{\lambda_t} \sim \chi^2(\nu), \quad z_t \sim N(0, 1). \]

- We assume \( \nu > 2 \) for a finite variance.
Intraday SV models

Total volatility

\[ V_t = X_t S_t A_t \quad \text{or} \quad h_t = x_t + s_t + a_t. \]

- \( X_t \) = SV part.
- \( S_t \) = intraday seasonality.
- \( A_t \) = announcement effects.
- \( h_t = \log(V_t^2), x_t = \log(X_t^2), s_t = \log(S_t^2), a_t = \log(A_t^2) \).

SV part

\[ x_{t+1} = \mu + \phi(x_t - \mu) + J_t Z_t^\nu + \eta_t, \]

\[
\begin{pmatrix}
\varepsilon_t \\
\eta_t
\end{pmatrix}
\sim \mathcal{N}(0, \mathbf{S}), \quad \mathbf{S} = \begin{pmatrix}
1 & \rho \sigma \\
\rho \sigma & \sigma^2
\end{pmatrix}.\]
Intraday SV models

Jump components

- Jump components \( (J_t Z_t^y, J_t Z_t^v) \) are assumed to be coincident in returns and the volatility with the common jump indicator variable, \( J_t \in \{0, 1\} \).

- Jump occurs with the probability \( \Pr[J_t = 1] = \kappa \).

- We assume

\[
Z_t^y \sim N(\mu_y, \sigma_y^2), \quad Z_t^v \sim N(\mu_v, \sigma_v^2).
\]
Intraday seasonality

- \( K \) = number of intraday returns during a day
- \( N \) = number of days in the sample periods
- \( T = K \times N \): total sample size
- Tokyo stock exchange is open for 9:00-11:30 and for 12:30-15:00.
- We use 5-min returns when the market is open.
- We also use overnight (15:00-9:00) and lunch-time (11:30-12:30) returns.
- Then, \( K = 62 \).
Intraday SV models

Average intraday Nikkei 225 absolute 5-min returns

(a) 2015/Apr/1 - 2016/Mar/31
(b) 2016/Apr/3 - 2017/Jul/14
Intraday SV models

- $\beta_k =$intraday seasonality at intraday period $k$ 
  \((k = 1, \ldots, K)\).
- $H_{tk} =$ intraday-period indicator

\[
H_{tk} = \begin{cases} 
1 & \text{if time } t \text{ corresponds to intraday period } k \\
0 & \text{otherwise}
\end{cases}
\]

- Then, $s_t$ is represented by

\[
s_t = \sum_{k=1}^{K} H_{tk} \beta_k, \quad t = 1, \ldots, T
\]
Intraday SV models

- We assume that $\beta_k$ follows the following random walk.

$$
\beta_{k+1} = \beta_k + w_k, \quad w_k \sim N(0, c_k v_\beta^2), \quad k = 1, \ldots, K - 1,
\beta_1 \sim N(0, 100).
$$

- We set $c_k = 100$ for $k = 1, 31, 32, 61$ and $c_k = 1$ for other $k$.

- Identifying restriction: $\frac{1}{K} \sum_{k=1}^{K} \beta_k = 0$. 
Intraday SV models

**Announcement effects**

\[ a_t = \sum_{j=1}^{J} \sum_{\ell=1}^{L} I_{jt\ell} \alpha_{j,\ell}, \]

\[ \alpha_{j,\ell+1} = \psi_j \alpha_{j,\ell} + \zeta_{j,\ell}, \quad \zeta_{j,\ell} \sim N(0, \nu_{\alpha_j}^2), \quad |\psi_j| < 1. \]

- \( J = 4 \) (GDP, IP, CPI, MPM)
- \( t_j^* \) = announcement time of the \( j \)th variable
- \( I_{jt\ell} \) = announcement-period indicator:

\[
I_{jt\ell} = \begin{cases} 
1 & \text{if } t \in \{t_j^*, t_j^* + 1, \ldots, t_j^* + L - 1\} \\
0 & \text{otherwise}
\end{cases}
\]

- \( L = 18 \) (i.e., 90 minutes).
Intraday SV models

Parameters

1. Parameters for SV: \( \theta_X = (\phi, \sigma, \rho, \mu) \).

2. Jump parameters: \( \theta_J = (\kappa, \mu_y, \sigma_y, \mu_v, \sigma_v) \).

3. Parameters for intraday seasonality and announcement effects: \( \nu = (\nu_\beta, \nu_{\alpha 1}, \ldots, \nu_{\alpha J}), \psi = (\psi_1, \ldots, \psi_J) \).

4. The degree of freedom for \( t \) distribution: \( \nu \).
Intraday SV models

Latent variables

1. SV process: \( x = (x_1, \ldots, x_T) \).

2. Jump components:
   \[
   J = (J_1, \ldots, J_{T-1}), \\
   Z^\varphi = (Z_1^\varphi, \ldots, Z_{T-1}^\varphi) \text{ for } \varphi = y, v.
   \]

3. Intraday seasonality and announcement effects:
   \[
   \beta = (\beta_1, \ldots, \beta_K), \\
   \alpha = \{\alpha_1, \ldots, \alpha_J\} \text{ where } \alpha_j = (\alpha_{j1}, \ldots, \alpha_{jL}).
   \]

4. Latent variables for \( t \) distribution: \( \lambda = (\lambda_1, \ldots, \lambda_T) \).
We develop a Bayesian estimation using MCMC.

- Parameters: $\theta = \{\theta_X, \theta_J, \nu, \psi, \nu\}$.

- Latent variables: $\Theta = \{x, J, Z^y, Z^v, \beta, \alpha, \lambda\}$.

- We sample them from the joint posterior density $\pi(\theta, \Theta | y)$ using the Gibbs sampler.
Bayesian analysis and computation

1. Parameters for SV: straightforward.
3. Parameters for the intraday seasonal and announcement effects: straightforward.
Empirical analysis

Data

- 5-min intraday returns of Nikkei 225 stock index.

Priors

\[
\begin{align*}
(\phi + 1)/2 & \sim B(20, 1.5), \quad \sigma^2 \sim IG(40, 0.2), \quad (\rho + 1)/2 \sim B(1, 1), \\
\mu & \sim N(-5, 4), \quad \nu \sim G(16, 0.8)I[\nu > 2], \quad \kappa \sim B(1, 500), \\
\mu_y & \sim N(0, 1), \quad \mu_v \sim N(1, 1), \quad \sigma_i^2 \sim IG(20, 4) \quad (i = y, v), \\
\nu^2_\beta & \sim IG(10, 1), \quad \nu^2_{\alpha,j} \sim IG(10, 1), \quad (\psi_j + 1)/2 \sim B(20, 1.5).
\end{align*}
\]
Empirical analysis

1. **SV**: normal distribution, no jumps.

2. **SVt**: Student $t$-distribution, no jumps.

3. **SVJ**: normal distribution, with jumps in return and volatility.

4. **SVJt**: Student $t$-distribution, with jumps in return and volatility.
Empirical analysis

Posterior estimates of the selected parameters for the high-frequency SV models (2015/Apr – 2016/Mar)

<table>
<thead>
<tr>
<th></th>
<th>SV</th>
<th>SVt</th>
<th>SVJ</th>
<th>SVJt</th>
</tr>
</thead>
<tbody>
<tr>
<td>μ</td>
<td>0.9898 (0.0015)</td>
<td>0.9935 (0.0011)</td>
<td>0.9918 (0.0011)</td>
<td>0.9935 (0.0011)</td>
</tr>
<tr>
<td>φ</td>
<td>[0.9868, 0.9926]</td>
<td>[0.9911, 0.9956]</td>
<td>[0.9894, 0.9941]</td>
<td>[0.9913, 0.9955]</td>
</tr>
<tr>
<td></td>
<td>5.0</td>
<td>16.5</td>
<td>12.6</td>
<td>20.1</td>
</tr>
<tr>
<td>σ</td>
<td>0.1558 (0.0070)</td>
<td>0.1213 (0.0069)</td>
<td>0.1082 (0.0068)</td>
<td>0.1090 (0.0068)</td>
</tr>
<tr>
<td></td>
<td>[0.1411, 0.1694]</td>
<td>[0.1094, 0.1358]</td>
<td>[0.0956, 0.1224]</td>
<td>[0.0954, 0.1221]</td>
</tr>
<tr>
<td></td>
<td>15.5</td>
<td>44.5</td>
<td>36.1</td>
<td>53.2</td>
</tr>
<tr>
<td>ρ</td>
<td>-0.2279 (0.0301)</td>
<td>-0.2711 (0.0348)</td>
<td>-0.2989 (0.0384)</td>
<td>-0.2917 (0.0405)</td>
</tr>
<tr>
<td></td>
<td>[-0.2872, -0.1685]</td>
<td>[-0.3392, -0.2008]</td>
<td>[-0.3673, -0.2197]</td>
<td>[-0.3701, -0.2123]</td>
</tr>
<tr>
<td></td>
<td>10.7</td>
<td>9.4</td>
<td>16.1</td>
<td>23.3</td>
</tr>
</tbody>
</table>

- The first row: posterior mean and standard deviation in parentheses.
- The second row: 95% credible interval in square brackets.
- The third row: inefficiency factor.
### Empirical analysis

<table>
<thead>
<tr>
<th></th>
<th>SV</th>
<th>SVt</th>
<th>SVJ</th>
<th>SVJt</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu$</td>
<td>-4.6803 (0.1228)</td>
<td>-4.6400 (0.1542)</td>
<td>-5.4439 (0.1762)</td>
<td>-4.9691 (0.1757)</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>1.0</td>
<td>22.0</td>
<td>30.3</td>
</tr>
<tr>
<td>$\nu_\beta$</td>
<td>0.2325 (0.0201)</td>
<td>0.2328 (0.0212)</td>
<td>0.2430 (0.0217)</td>
<td>0.2317 (0.0209)</td>
</tr>
<tr>
<td></td>
<td>[0.1967, 0.2745]</td>
<td>[0.1946, 0.2806]</td>
<td>[0.2046, 0.2880]</td>
<td>[0.1944, 0.2764]</td>
</tr>
<tr>
<td></td>
<td>0.9</td>
<td>0.9</td>
<td>1.4</td>
<td>1.0</td>
</tr>
<tr>
<td>$\nu_{\alpha 1}$</td>
<td>0.3241 (0.0521)</td>
<td>0.3225 (0.0522)</td>
<td>0.3240 (0.0512)</td>
<td>0.3222 (0.0515)</td>
</tr>
<tr>
<td></td>
<td>[0.2417, 0.4470]</td>
<td>[0.2404, 0.4467]</td>
<td>[0.2423, 0.4414]</td>
<td>[0.2385, 0.4395]</td>
</tr>
<tr>
<td></td>
<td>0.9</td>
<td>0.6</td>
<td>1.1</td>
<td>1.7</td>
</tr>
<tr>
<td>$\psi_1$</td>
<td>0.7074 (0.1414)</td>
<td>0.7056 (0.1389)</td>
<td>0.7082 (0.1401)</td>
<td>0.7071 (0.1436)</td>
</tr>
<tr>
<td></td>
<td>[0.4223, 0.9430]</td>
<td>[0.4138, 0.9518]</td>
<td>[0.4231, 0.9563]</td>
<td>[0.4150, 0.9512]</td>
</tr>
<tr>
<td></td>
<td>0.8</td>
<td>0.7</td>
<td>0.7</td>
<td>0.4</td>
</tr>
<tr>
<td>$\nu$</td>
<td>10.889 (0.9258)</td>
<td>12.851 (1.3137)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[9.2183, 12.784]</td>
<td>[10.759, 15.874]</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>43.1</td>
<td>57.2</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## Empirical analysis

<table>
<thead>
<tr>
<th></th>
<th>SVJ</th>
<th>SVJt</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\kappa)</td>
<td>0.0091 (0.0018)</td>
<td>0.0017 (0.0008)</td>
</tr>
<tr>
<td></td>
<td>[0.0057, 0.0130]</td>
<td>[0.0005, 0.0034]</td>
</tr>
<tr>
<td></td>
<td>24.5</td>
<td>69.8</td>
</tr>
<tr>
<td>(\mu_y)</td>
<td>-0.0134 (0.0407)</td>
<td>-0.1258 (0.1187)</td>
</tr>
<tr>
<td></td>
<td>[-0.0970, 0.0632]</td>
<td>[-0.3744, 0.1048]</td>
</tr>
<tr>
<td></td>
<td>9.4</td>
<td>31.7</td>
</tr>
<tr>
<td>(\mu_v)</td>
<td>0.6391 (0.1198)</td>
<td>1.2841 (0.2916)</td>
</tr>
<tr>
<td></td>
<td>[0.4256, 0.8921]</td>
<td>[0.7553, 1.8865]</td>
</tr>
<tr>
<td></td>
<td>15.4</td>
<td>47.3</td>
</tr>
<tr>
<td>(\sigma_y)</td>
<td>0.3301 (0.0255)</td>
<td>0.4155 (0.0410)</td>
</tr>
<tr>
<td></td>
<td>[0.2853, 0.3821]</td>
<td>[0.3445, 0.5039]</td>
</tr>
<tr>
<td></td>
<td>17.7</td>
<td>10.0</td>
</tr>
<tr>
<td>(\sigma_v)</td>
<td>0.4898 (0.0540)</td>
<td>0.4753 (0.0562)</td>
</tr>
<tr>
<td></td>
<td>[0.3927, 0.6068]</td>
<td>[0.3787, 0.5988]</td>
</tr>
<tr>
<td></td>
<td>17.7</td>
<td>7.3</td>
</tr>
</tbody>
</table>
Empirical analysis

Posterior mean of total volatility $V_t^2$ (5-min in the first 10 days)
Empirical analysis

Posterior mean of total volatility $V_t^2$ (daily)
Empirical analysis

Posterior probability of jump (5-min)
Empirical analysis

Posterior mean of parameters \((\beta_1, \ldots, \beta_K)\) for the intraday seasonality

Posterior means (solid) and 95% credible intervals (dashed).
Empirical analysis

Announcement effects $\alpha_j$ for GDP, industrial production (IP), consumer price index (CPI) and the Bank of Japan’s monetary policy meeting (MPM)

- Posterior means (solid) and 95% credible intervals (dashed).
Empirical analysis

- This result is consistent with


## Model comparison

### BIC for intraday SV and GARCH models

(2015/Apr/1 – 2016/Mar/31)

<table>
<thead>
<tr>
<th>Model</th>
<th>BIC</th>
<th>Ranking</th>
</tr>
</thead>
<tbody>
<tr>
<td>SV</td>
<td>-30090</td>
<td>3</td>
</tr>
<tr>
<td>SVJ</td>
<td>-29723</td>
<td>4</td>
</tr>
<tr>
<td>SVt</td>
<td>-31016</td>
<td>1</td>
</tr>
<tr>
<td>SVJt</td>
<td>-30327</td>
<td>2</td>
</tr>
<tr>
<td>GARCH</td>
<td>-26275</td>
<td>9</td>
</tr>
<tr>
<td>GJR</td>
<td>-26341</td>
<td>8</td>
</tr>
<tr>
<td>EGARCH</td>
<td>-26264</td>
<td>10</td>
</tr>
<tr>
<td>GARCHt</td>
<td>-27315</td>
<td>7</td>
</tr>
<tr>
<td>GJ Rt</td>
<td>-27379</td>
<td>5</td>
</tr>
<tr>
<td>EGARCHt</td>
<td>-27371</td>
<td>6</td>
</tr>
</tbody>
</table>
Models for daily RV

- HAR model (Corsi 2009)

\[
\log RV_t = c + b_d \log RV_{t-1} + b_w \log RV_{t-1}^{(w)} + b_m \log RV_{t-1}^{(m)} + u_t,
\]

where

\[
RV_{t-1}^{(w)} = \frac{1}{5} \sum_{i=1}^{5} RV_{t-i}, \quad RV_{t-1}^{(m)} = \frac{1}{22} \sum_{i=1}^{22} RV_{t-i}.
\]
Model comparison

- Realized EGARCH model (Hansen and Huang 2016)

\[ R_t = \sigma_t z_t, \quad z_t \sim N(0, 1), \]

\[ \log \sigma_t^2 = \omega + \beta \ln(\sigma_{t-1}^2) + \tau_1 z_t + \tau_2(z_t^2 - 1) + \gamma \nu_{t-1}, \]

\[ \log \text{RV}_t = \xi + \log \sigma_t^2 + \delta_1 z_t + \delta_2(z_t^2 - 1) + \nu_t, \]

\[ \nu_t \sim N(0, \sigma_{\nu}^2). \]

- Realized SV model (Takahashi, Omori and Watanabe 2009, Takahashi, Watanabe and Omori 2016)
Out-of-sample forecast performance

- We compare the predictive ability of one-day-ahead volatility of our model with those of the HAR and realized EGARCH models.

- Since the true volatility is unobserved, we must use the proxy.

- Patton (2011) shows that the MSE and QLIKE are robust loss functions in the sense that they lead to the same ranking as the one when the true volatility is used if the proxy is the unbiased estimator of the true volatility.
Model comparison

\[
\text{MSE} = \frac{1}{N_1} \sum_{\tau = N_0 + 1}^{N_0 + N_1} (\hat{\sigma}_\tau^2 - \sigma^2_\tau)^2 ,
\]

\[
\text{QLIKE} = \frac{1}{N_1} \sum_{\tau = N_0 + 1}^{N_0 + N_1} \left( \frac{\sigma^2_\tau}{\hat{\sigma}^2_\tau} + \log \hat{\sigma}^2_\tau \right) ,
\]

where \( \sigma^2_\tau \) and \( \hat{\sigma}^2_\tau \) denote the true volatility and the volatility forecast for day \( \tau \), respectively.
Model comparison

- We calculate the realized kernel (Barndorff-Nielsen et al. 2008, 2009) using 1-min returns when the market is open.

- We convert it to the volatility estimates for 24 hours including non-trading hours such as lunch-time and overnight following Hansen and Lunde (2005).
Model comparison

One-day ahead volatility forecasting performance (2016/Apr/1 – 2017/July/14)

<table>
<thead>
<tr>
<th>Model</th>
<th>MSE</th>
<th>QLIKE</th>
</tr>
</thead>
<tbody>
<tr>
<td>SV</td>
<td>10.85</td>
<td>1.381</td>
</tr>
<tr>
<td>SVJ</td>
<td>10.80</td>
<td>1.352</td>
</tr>
<tr>
<td>SVt</td>
<td>11.04</td>
<td>1.402</td>
</tr>
<tr>
<td>SVJt</td>
<td>10.32</td>
<td>1.305</td>
</tr>
<tr>
<td>HAR</td>
<td>11.34</td>
<td>1.309</td>
</tr>
<tr>
<td>REGARCH</td>
<td>11.60</td>
<td>1.417</td>
</tr>
</tbody>
</table>

- Ranking is in brackets.
Conclusion

1. We extend daily SV models to intraday high-frequency SV models.
2. We develop an MCMC Bayesian method for the analysis of intraday high-frequency SV models.
3. We apply them to 5-min returns of Nikkei 225 stock index.
4. We show
   (1) Intraday SV models fit the data better than intraday GARCH-type models.
   (2) Intraday SV models perform better than daily RV models such as HAR and realized EGARCH models in one-day-ahead volatility forecasting.


References


