Stochastic Volatility and Jumps: Exponentially Affine Yes or No? An Empirical Analysis of S&P500 Dynamics

Katja Ignatieva, Paulo Rodrigues, Norman Seeger

Far East and South Asia Meeting of the Econometric Society  
3-5 August 2009 Tokyo
Motivation

**Research Topic:**

- Which model should be used to model **dynamics of equity indices or stock**

- Capturing **stylized facts** in the data:
  - Non-normality
  - Heavy tails
  - Skewness
  - Volatility clustering
  - Leverage effect

- What has been done:
  - GBM: Black, Scholes, Merton Model (1973)
  - Jumps in returns: Merton (1976)
  - Stochastic volatility (SV): Heston (1993)
Objective of the Paper

Empirical findings:

1. Heston model is misspecified
2. Jump components in returns reduce misspecification

Two approaches to better capture stock return properties:
- Eraker et al. (2003): affine SV structure plus jumps in returns and volatility (based on Duffie, Pan, Singleton (2000))
- Christofferson et al. (2007): non-affine structure of SV process (extending Heston (1993))

Objective of the paper:
- compare the two approaches
- combine the two approaches
- consider SV, SVJ, and SVCJ model classes
- estimate parameters via Markov Chain Monte Carlo (MCMC)
- compare model performance
Research Question

Objective of the paper / Research Questions:

1. Does the performance of non-affine SV models improve by including jumps (in general)?

2. Do we still have to leave the class of affine models after including jumps?
Contributions

Overall result:

1. Jump models are clearly preferred by test statistics
   - results hold for affine and non-affine model specifications

2. Non-affine models exhibit a good fit to the data and are worth investigating
   - mathematical and economic properties are unknown

3. Affine models with jumps have similar performance to the non-affine models
   - we tend to prefer affine models since they are well understood
     (closed form solution, mathematical properties)
Agenda:

1. Motivation
2. Model Setup and Estimation
3. Model Selection Criteria and Model Testing
4. Data Set
5. Results
6. Conclusion
SVCJ model specification

- We assume that the logarithm of the stock price solves

\[
\begin{align*}
    dY_t &= \mu dt + \sqrt{V_t} dW^Y_t + \xi^Y dN_t \\
    dV_t &= \kappa V_t^a (\theta - V_t) dt + V_t^b \sigma_V dW^V_t + \xi^V dN_t
\end{align*}
\]

- Assumptions

  - \(dW^Y_t, dW^V_t\) are Brownian motions with correlation \(\rho\)
  - \(N_t\) is a Poisson process with intensity \(\lambda\)
  - SVCJ: \(\xi^Y_t \sim \text{Exp}(\mu_V); \quad \xi^Y_t | \xi^V_t \sim \mathcal{N}(\mu_y + \rho \xi^V_t, \sigma_y)\)
  - \(a \in [0; 1]\) and \(b \in [1/2; 1; 3/2]\)

- Stochastic volatility (SV) and stochastic volatility with jumps in returns (SVJ) are special cases
Model Setup (cont’d)

Model specifications for each model class

\[
dV_t = \kappa V_t^a (\theta - V_t) dt + V_t^b \sigma V \, dW_t^V
\]

<table>
<thead>
<tr>
<th>a</th>
<th>b</th>
<th>Name</th>
<th>Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0</td>
<td>0.5</td>
<td>SQR</td>
<td>variance drift is linear in variance square root diffusion</td>
</tr>
<tr>
<td>1.0</td>
<td>0.5</td>
<td>SQRN</td>
<td>variance drift is nonlinear in variance square root diffusion</td>
</tr>
<tr>
<td>0.0</td>
<td>1.0</td>
<td>VAR</td>
<td>variance drift is linear in variance linear diffusion</td>
</tr>
<tr>
<td>1.0</td>
<td>1.0</td>
<td>VARN</td>
<td>variance drift is nonlinear in variance linear diffusion</td>
</tr>
<tr>
<td>0.0</td>
<td>1.5</td>
<td>3/2</td>
<td>variance drift is linear in variance 3/2 diffusion</td>
</tr>
<tr>
<td>1.0</td>
<td>1.5</td>
<td>3/2N</td>
<td>variance drift is nonlinear in variance 3/2 diffusion</td>
</tr>
</tbody>
</table>

- Model classes: SV, SVJ, SVCJ (overall 18 different models)
Estimation

Estimation is based on

- Euler discretization yields

\[
R_{t+1} = \mu + \sqrt{V_t} \xi^y_{t+1} + \xi^y_{t+1} J_{t+1} \\
V_{t+1} = V_t + \kappa V_t (\theta - V_t) + \sigma_v V_t \xi^v_{t+1} + \xi^v_{t+1} J_{t+1}.
\]

where

- \( R_{t+1} = Y_{t+1} - Y_t \)
- \( J_{t+1} = N_{t+1} - N_t \)
- \( \xi^i_{t+1} = W^i_{t+1} - W^i_t \) for \( i = y, v \)

Aim is to estimate

- Parameters: \( \Theta = (\rho, \kappa, \theta, \sigma_v, \mu, \mu_y, \sigma_y, \lambda, \mu_v, \rho_j) \)
- Latent variables: \( X = \{ V_t, J_t, \xi^y_t, \xi^v_t \}_{t=1}^T \)
**Bayesian Framework**

**Bayesian framework used for estimation**

- Posterior distribution (PD) is given by Bayes’ Theorem

\[
p(\theta|y) = \frac{p(y|\theta)p(\theta)}{p(y)} \propto p(y|\theta)p(\theta)
\]

- PD combines information in model and prices
- likelihood \( p(y|\theta) \) is given by the model
- prior \( p(\theta) \) exogenously given (uninformative)

- As point estimator for parameters from posterior we use

\[
E(\theta|y) = \int \theta p(\theta|y)d\theta
\]
Problems for estimation procedure:

- Posterior distribution does not take the form of a well known density
  - no closed form solution
- Posterior distribution is high dimensional
- Simultaneous estimation of parameters and latent variables

Solution: Using Markov Chain Monte Carlo (MCMC)
MCMC in a nutshell:

- We want to sample from PD $p(\theta|y)$

**MC:** Construct Markov Chain which converges to the PD

- Given initial values $\theta^{(0)}$ draw a sequence

$$
\theta_{1}^{(1)} \sim p(\theta_{1}|\text{all other parameters}, Y) \\
\vdots \\
\theta_{K}^{(1)} \sim p(\theta_{K}|\text{all other parameters}, Y)
$$

- The resulting sequence $\{\theta^{(g)}\}_{g=1}^{G}$ converges to PD

**MC:** Calculate point estimators by approximating

$$
E(\theta|y) = \int \theta p(\theta|y) d\theta \approx \frac{1}{N} \sum_{n=G+1}^{N} \theta^{(n)}
$$
Specify the prior distributions:
use conjugate priors proposed by Eraker et al. (2003):

- $\mu \sim \mathcal{N}(1, 25)$
- $\kappa \theta \sim \mathcal{N}(0, 1)$, $\kappa \sim \mathcal{N}(0, 1)$
- $\lambda \sim \mathcal{B}(2, 40)$
- $\mu_y \sim \mathcal{N}(0, 100)$, $\sigma_y^2 \sim \mathcal{IG}(5, 20)$
- $\mu_v \sim \mathcal{G}(20, 10)$, $\sigma_v^2 \sim \mathcal{IG}(2.5, 0.1)$
- $\rho_j \sim \mathcal{N}(0, 4)$,
- $\rho \sim \mathcal{U}(-1, 1)$ (not a conjugate)
Agenda:

1. Motivation
2. Model Setup and Estimation
3. Model Selection Criteria and Model Testing
4. Data Set
5. Results
6. Conclusion
Criteria of Model Fit

Model Choice is based on using

- **Quantile to quantile plots (QQ-Plots)**
  - plot quantiles of estimated errors of return equation against standard normal distribution

\[ \varepsilon_{t+1}^y = \left( R_{t+1} - \mu - \xi_{t+1}^y J_{t+1} \right) / \sqrt{V_t} \]

- **Deviance Information Criterion (DIC)**
  - like any other information criterion combines a term for model fit and model complexity

\[ DIC = \bar{D} + p_D \]

- **Bayes Factors**
  - ratio of probabilities of two models given the data

\[ \frac{p(M_1|data)}{p(M_2|data)} \]
Agenda:

1. Motivation
2. Model Setup and Estimation
3. Model Selection Criteria and Model Testing
4. Data Set
5. Results
6. Conclusion
Data used

- Time series data
  - daily returns of S&P500
- Sample period from January 2, 1986 to July 31, 2008
- MCMC procedure
  - Number of draws 500,000; burn-in period of 200,000
Agenda:

1. Motivation
2. Model Setup and Estimation
3. Model Selection Criteria and Model Testing
4. Data Set
5. Results
6. Conclusion
Comparing models via QQ-Plots:

- Heston (SV-SQR) is misspecified
- Non-affine SV model, and affine jump diffusion model have a good fit
- Performance in the tails is much better
## Comparing models via DIC-Statistics:

<table>
<thead>
<tr>
<th></th>
<th>Model</th>
<th>DIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>SVCJ-VARN</td>
<td>14023.51</td>
</tr>
<tr>
<td>2</td>
<td>SVJ-3/2N</td>
<td>14062.19</td>
</tr>
<tr>
<td>3</td>
<td>SVCJ-3/2N</td>
<td>14091.67</td>
</tr>
<tr>
<td>4</td>
<td>SVJ-VAR</td>
<td>14103.33</td>
</tr>
<tr>
<td>5</td>
<td>SVJ-VARN</td>
<td>14125.84</td>
</tr>
<tr>
<td>6</td>
<td>SVCJ-SQR</td>
<td>14144.14</td>
</tr>
<tr>
<td>7</td>
<td>SVCJ-VAR</td>
<td>14177.38</td>
</tr>
<tr>
<td>8</td>
<td>SVJ-SQR</td>
<td>14177.90</td>
</tr>
<tr>
<td>9</td>
<td>SV-3/2N</td>
<td>14199.65</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Model</th>
<th>DIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>SVJ-SQRN</td>
<td>14212.51</td>
</tr>
<tr>
<td>11</td>
<td>SVCJ-SQRN</td>
<td>14222.77</td>
</tr>
<tr>
<td>12</td>
<td>SV-VARN</td>
<td>14240.15</td>
</tr>
<tr>
<td>13</td>
<td>SV-VAR</td>
<td>14263.10</td>
</tr>
<tr>
<td>14</td>
<td>SV-SQR</td>
<td>14281.59</td>
</tr>
<tr>
<td>15</td>
<td>SV-SQRN</td>
<td>14401.61</td>
</tr>
<tr>
<td>16</td>
<td>SV-3/2</td>
<td>15355.37</td>
</tr>
<tr>
<td>17</td>
<td>SVCJ-3/2</td>
<td>15373.22</td>
</tr>
<tr>
<td>18</td>
<td>SVJ-3/2</td>
<td>15474.34</td>
</tr>
</tbody>
</table>

- **SV models are outperformed** by jump diffusion models (affine as well as non-affine)
- **Non-affine models perform best**
- **Of all models with linear drift SVCJ-SQR is second best**
Comment on non-linear drift

Problems with non-linear drift specification

- Non-linear drift specification means $a = 1$

$$dV_t = \kappa V_t^a(\theta - V_t)dt + V_t^b \sigma_v dW^V_t$$

$$dV_t = \kappa V_t \theta dt - \kappa V_t^2 dt + V_t^b \sigma_v dW^V_t$$

- Drift and diffusion term vanishes when variance hits 0
  - In this case long run mean of variance is 0
- Is this specification economically questionable!?
- Further research needed
  (solution: condition process on not hitting 0)
## Bayes Factors

### Comparing models via Bayes Factors:

<table>
<thead>
<tr>
<th>(a; b)</th>
<th>SVJ vs. SV</th>
<th>SVCJ vs. SV</th>
<th>SVCJ vs. SVJ</th>
</tr>
</thead>
<tbody>
<tr>
<td>(0.0;0.5)</td>
<td>19.74</td>
<td>4.57</td>
<td>-15.17</td>
</tr>
<tr>
<td>(1.0;0.5)</td>
<td>19.58</td>
<td>41.36</td>
<td>21.78</td>
</tr>
<tr>
<td>(0.0;1.0)</td>
<td>27.82</td>
<td>25.66</td>
<td>-2.15</td>
</tr>
<tr>
<td>(1.0;1.0)</td>
<td>25.05</td>
<td>26.50</td>
<td>1.45</td>
</tr>
<tr>
<td>(0.0;1.5)</td>
<td>40.46</td>
<td>38.70</td>
<td>-1.76</td>
</tr>
<tr>
<td>(1.0;1.5)</td>
<td>34.03</td>
<td>43.30</td>
<td>9.26</td>
</tr>
</tbody>
</table>

- Computation of Bayes factors only for nested models
- Ratio from 6 to 10: strong evidence for model in nominator
- **SV models are outperformed** by jump diffusion models (affine as well as non-affine)
- Mixed results for SVCJ vs. SVJ
Summerize Results: 

1. In terms of QQ plots affine jump diffusion models similar to non-affine models
2. Affine jump diffusion model second best DIC statistic of models with linear drift
3. Jump diffusion models are clearly preferred by DIC statistic / Bayes factors
4. We suggest further investigation of non-affine models, due to good statistical properties
5. We tend to prefer affine model class
   - performance of affine models similar as non-affine models
   - mathematical and statistical properties of affine models are well known
Agenda:

1. Motivation
2. Model Setup and Estimation
3. Data Set
4. Model Selection Criteria and Model Testing
5. Results
6. Conclusion
Contributions and Findings

Contributions:

- **Combine** two approaches to overcome model misspecification of Heston (1993) model
- **Estimate** model parameters via MCMC
- **Compare** different model specifications by several test statistics

Findings:

- **Jump models outperform** SV models
- Affine and non-affine models have **similar performance**
- We **prefer affine models** since better understanding of mathematical and economical properties
Future Research:

- Consider out of sample test
- Comparison of models via capability of capturing the smile
- Comparison to Levy processes
- Using high frequency data
- Different drift specifications (regime switching models)
Thank you very much for your attention!