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

Katja Ignatieva, Paulo Rodrigues, Norman Seeger

GOETHE UNIVERSITÄT FRANKFURT AM MAIN

Swiss Institute of Banking and Finance
University of St.Gallen

6th World Congress of the Bachelier Finance Society
Toronto, June 24th

Motivation

Research Topic:

- Which model should be used to model dynamics of equity indices
- 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)
  - SV with jumps in returns: Bakshi et. al (1997), Pan (2002)

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 (statistically / economically)

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?

Agenda

Agenda:

1. Motivation
2. Model Setup and Estimation
3. Data Set
4. Results
5. Conclusion

Agenda

Agenda:

1. Motivation
2. Model Setup and Estimation
3. Data Set
4. Results
5. Conclusion

Model Setup

SVCJ model specification

- We assume that the logarithm of the stock price solves

$$\begin{aligned} dY_t &= \mu dt + \sqrt{V_t} dW_t^y + \xi^y dN_t \\ dV_t &= \kappa V_t^a (\theta - V_t) dt + V_t^b \sigma_v dW_t^v + \xi^v dN_t \end{aligned}$$

- Assumptions
  - dW_t^y, dW_t^v are Brownian motions with correlation ρ
  - N_t is a Poisson process with intensity λ
  - SVCJ: $\xi_t^v \sim Exp(\mu_v)$; $\quad \xi_t^y | \xi_t^v \sim N(\mu_y + \rho_j \xi_t^v, \sigma_y)$
  - 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

$$dV_t = \kappa V_t^{a}(\theta - V_t)dt + V_t^{b}\sigma_v dW_t^v$$

a	b	Name	Features
0.0	0.5	SQR	variance drift is linear in variance square root diffusion
1.0	0.5	SQRN	variance drift is nonlinear in variance square root diffusion
0.0	1.0	VAR	variance drift is linear in variance linear diffusion
1.0	1.0	VARN	variance drift is nonlinear in variance linear diffusion
0.0	1.5	3/2	variance drift is linear in variance 3/2 diffusion
1.0	1.5	3/2N	variance drift is nonlinear in variance 3/2 diffusion

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

Criteria of Model Fit

Statistical model choice is based on using

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

$$\varepsilon^y_{t+1} = (R_{t+1} - \mu - \xi^y_{t+1} J_{t+1})/V^b_t$$

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

$$\text{DIC} = \bar{D} + p_D$$

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

$$\frac{p(\mathcal{M}_1|data)}{p(\mathcal{M}_2|data)}$$

Agenda

Agenda:

1. Motivation
2. Model Setup and Estimation
3. Data Set
4. Results
5. Conclusion

Data Set

Data used

- Time series data
  - daily returns of S&P500 index
  - daily returns of NASDAQ index
- Sample period from January 2, 1986 to July 31, 2008
  - robustness check: using different sub samples
  - including data up to first half of 2009 (Lehman)
- MCMC procedure
  - number of draws 500,000; burn-in period of 200,000

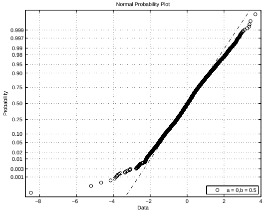
Agenda

Agenda:

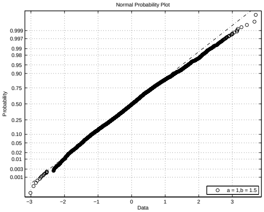
QQ-Plots

Comparing models via QQ-Plots:

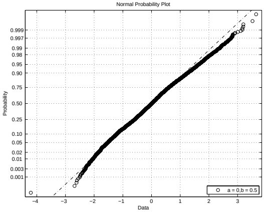
SV-SQR



SV-3/2N



SVCJ-SQR



- 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

Deviance Information Criterion

Comparing models via DIC-Statistics:

	Model	DIC
1	SVCJ-VARN	14023.51
2	SVJ-3/2N	14062.19
3	SVCJ-3/2N	14091.67
4	SVJ-VAR	14103.33
5	SVJ-VARN	14125.84
6	SVCJ-SQR	14144.14
7	SVCJ-VAR	14177.38
8	SVJ-SQR	14177.90
9	SV-3/2N	14199.65

	Model	DIC
10	SVJ-SQRN	14212.51
11	SVCJ-SQRN	14222.77
12	SV-VARN	14240.15
13	SV-VAR	14263.10
14	SV-SQR	14281.59
15	SV-SQRN	14401.61
16	SV-3/2	15355.37
17	SVCJ-3/2	15373.22
18	SVJ-3/2	15474.34

- 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

Bayes Factors

Comparing models via Bayes Factors:

(a; b)	SVJ vs. SV	SVCJ vs. SV	SVCJ vs. SVJ
(0.0;0.5)	19.74	4.57	-15.17
(1.0;0.5)	19.58	41.36	21.78
(0.0;1.0)	27.82	25.66	-2.15
(1.0;1.0)	25.05	26.50	1.45
(0.0;1.5)	40.46	38.70	-1.76
(1.0;1.5)	34.03	43.30	9.26

- 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

Agenda

Agenda:

1. Motivation
2. Model Setup and Estimation
3. Data Set
4. Results
5. Conclusion

Contributions and Findings

Contributions:

- Combine two approaches to overcome model misspecification
- Estimate model parameters via MCMC
- Compare different model specifications by several test statistics

Findings:

Overall result:

1. Jump models are clearly preferred by test statistics
   - results hold for affine and non-affine 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 similar performance than non-affine models
   - we tend to prefer affine models since they are well understood (closed form solution, mathematical properties)

Thank you very much!

Thank you very much!

Estimation

Estimation is based on

- Euler discretization yields

$$
\begin{aligned}
R_{t+1} &= \mu + \sqrt{V_t}\varepsilon^y_{t+1} + \xi^y_{t+1}J_{t+1} \\
V_{t+1} &= V_t + \kappa V_t^a(\theta - V_t) + \sigma_v V_t^b \varepsilon^v_{t+1} + \xi^v_{t+1}J_{t+1}.
\end{aligned}
$$

where

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

Aim is to estimate

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

Bayesian Framework

Baysian framework used for estimation

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

$$p(\boldsymbol{\theta}|\mathbf{y}) = \frac{p(\mathbf{y}|\boldsymbol{\theta})p(\boldsymbol{\theta})}{p(\mathbf{y})} \propto p(\mathbf{y}|\boldsymbol{\theta})p(\boldsymbol{\theta})$$

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

- As point estimator for parameters from posterior we use

$$\mathbf{E}(\boldsymbol{\theta}|\mathbf{y}) = \int \boldsymbol{\theta} p(\boldsymbol{\theta}|\mathbf{y}) d\boldsymbol{\theta}$$

Bayesian Framework

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

MCMC in a nutshell:

- We want to sample from PD $p(\boldsymbol{\theta}|\mathbf{y})$

MC: Construct Markov Chain which converges to the PD

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

$$\begin{aligned}\theta_1^{(1)} &\sim p(\theta_1|\text{all other parameters}, \mathbf{Y})\\ &\vdots\\ \theta_K^{(1)} &\sim p(\theta_K|\text{all other parameters}, \mathbf{Y})\end{aligned}$$

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

MC: Calculate point estimators by approximating

$$\mathbf{E}(\boldsymbol{\theta}|\mathbf{y}) = \int \boldsymbol{\theta} p(\boldsymbol{\theta}|\mathbf{y})d\boldsymbol{\theta} \approx \frac{1}{N}\sum_{n=G+1}^{N} \boldsymbol{\theta}^{(n)}$$

Comment on non-linear drift

Problems with non-linear drift specification

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

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

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

- 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)

Summerize Results

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

Future Research

Future Research:

- Comparison to Levy processes
- Using high frequency data
- Different drift specifications (regime switching models)
- Consider out-of-sample test
  - What about the conditional return distribution?

Criteria of Model Fit

Economic model choice is based on using

- Model capability of capturing the smile
  (a) calculation option prices via monte carlo (different strikes and maturities)
  (b) backing out the implied volatility
  (c) comparing model generated smiles with actual smile observed in the data (randomly chosen day)
  (d) consistent estimation of Q and P parameters by means of return and option data (Broadie et al.(2007))

$$(\hat{\Theta}^{\mathbb{Q}}, \hat{V}_t) = \arg\min \sum_{t,n} \left[IV_t(K_n, \tau_n, S_t, r) - IV(V_t, \Theta^{\mathbb{Q}}|\Theta^{\mathbb{P}}, K_n, \tau_n, S_t, r)\right]^2$$