Investigating Impacts of Self-Exciting Jumps in Returns and Volatility: A Bayesian Learning Approach

Andras Fulop[a], Junye Li[b], and Jun Yu[c]

[a] *ESSEC Business School and CREST*
[b] *ESSEC Business School*
[c] *Singapore Management University*

November 2012

Outline

Jump episodes

Table: S&P 500 Index Returns in Four Turbulent Periods

Period 1		Period 2		Period 3		Period 4	
Date	Ret	Date	Ret	Date	Ret	Date	Ret
10/16/87	-5.30	7/18/02	-2.74	9/15/08	-4.83	8/04/11	-4.90
10/19/87	-22.9	7/19/02	-3.91	9/16/08	1.74	8/05/11	-0.06
10/20/87	5.20	7/22/02	-3.35	9/17/08	-4.83	8/08/11	-6.90
10/21/87	8.71	7/23/02	-2.74	9/18/08	4.24	8/09/11	4.63
10/22/87	4.00	7/24/02	5.57	9/19/08	3.95	8/10/11	-4.52
10/23/87	-0.01	7/25/02	-0.56	9/22/08	-3.90	8/11/11	-4.53
10/26/87	-8.64	7/26/02	1.67	9/23/08	-1.58	8/12/11	0.52

- Jump clustering: Stylized fact during the financial crisis
- Existing Literature:
  - Price and diffusion volatility jump at the same time; Eraker, Johannes, and Polson (2003), Eraker (2004)
  - Jumps are self-exciting; Yu (2004), McCurdy and Maheu (2004), Aït-Sahalia, Cacho-Diaz, and Laeven (2010) and Carr and Wu (2010)
- Jump clustering is trigged by negative jumps.

This Paper

- Set up a class of jump-diffusion models where negative jumps in asset prices can feedback both to diffusion volatility (Channel 1) and to jump intensity (Channel 2)
  - Channel 1: A negative jump in price happens at the same time as a jump in diffusion volatility. Since diffusion volatility is persistent, another large volatility value is expected in the next period. Consequently, another extreme movement in asset price is highly likely to be followed, even if there is no jump arrival.
  - Channel 2: A negative jump in price increases the likelihood of extreme events in future price movements.

This Paper

- Develop an econometric toolbox to perform sequential Bayesian inference over the hidden states and fixed parameters
- Estimate a set of models on S&P 500 stock returns between 1980-2011
  - Find strong evidence of Channel 1 (jump to diffusion volatility) throughout the sample
  - Find weak evidence of Channel 2 after 2008 financial crisis
  - Strong implications for VaR, option pricing and volatility forecasting

Outline

Model

- Equity Price follow a jump-diffusion with variance-gamma jumps

$$\ln S_t/S_0 = \int_0^t \mu_s ds + \left(W_{T_{1,t}} - k_W(1)T_{1,t}\right) + \left(J_{T_{2,t}} - k_J(1)T_{2,t}\right),$$

- Diffusion variance, $V_{1,t}$, and jump intensity, $V_{2,t}$ follow

$$\begin{aligned} dV_{1,t} &= \kappa_1(\theta_1 - V_{1,t})dt + \sigma_{11}\sqrt{V_{1,t}}dZ_t - \sigma_{12}dJ^-_{T_{2,t}}, \\ dV_{2,t} &= \kappa_2(\theta_2 - V_{2,t})dt - \sigma_2 dJ^-_{T_{2,t}}, \end{aligned}$$

Model Specifications we investigate

SE-M1: Full model, jump clustering generated by jump to diffusion volatility and time varying jump intensity

SE-M2: $\sigma_{12} = 0$, jump clustering only through jump intensity

SE-M3: $\sigma_2 = 0$, jump clustering only through jump to diffusion volatility

SE-M4: $\sigma_{12} = \sigma_2 = 0$, no jump clustering

Outline

By an Euler approximation, we have a state space model

- Measurement Equation

$$\begin{aligned} \ln S_t &= \ln S_{t-\tau} + \left(\mu - \frac{1}{2}V_{1,t-\tau} - k(1)V_{2,t-\tau}\right)\tau + \sqrt{\tau V_{1,t-\tau}}w_t \\ &\quad + J_{u,t} + J_{d,t}, \end{aligned}$$

- Transition Equations

$$\begin{aligned} V_{1,t} &= \kappa_1\theta_1\tau + (1-\kappa_1\tau)V_{1,t-\tau} + \sigma_{11}\sqrt{\tau V_{1,t-\tau}}z_t - \sigma_{12}J_{d,t}, \\ V_{2,t} &= \kappa_2\theta_2\tau + (1-\kappa_2\tau)V_{2,t-\tau} - \sigma_2 J_{d,t}, \\ J_{u,t} &= \Gamma(\tau V_{2,t-\tau}; \mu_u, v_u), \qquad (1) \\ J_{d,t} &= -\Gamma(\tau V_{2,t-\tau}; \mu_d, v_d). \qquad (2) \end{aligned}$$

Filtering the hidden states: Particle Filters

- Use a sequential Monte Carlo routine with M interacting particles to recursively approximate the filtering distributions of the hidden states

$$\hat{p}(x_{1:t}|y_{1:t},\Theta)=\sum_{i=1}^{M}\tilde{w}_t^{(i)}1_{\{x_{1:t}=x_{1:t}^{(i)}\}},$$

- We need to design an efficient proposal to take care of the jumps. Use a mixture proposal that performs well both for large and small returns.
- The algorithm also provide an estimate of the individual marginal likelihoods

$$\hat{p}(y_l|y_{1:l-1},\Theta)=\frac{1}{M}\sum_{i=1}^{M}w_l^{(i)}.$$

Mixture Particle Filter I

If $R_t = \ln S_t - \ln S_{t-\tau} > 0$,

- draw $J_{d,t}^{(i)}$ from its transition law (2);
- draw $J_{u,t}^{(i)}$ both from its transition law (1) and its conditional posterior distribution $J_{u,t} = \ln S_t - \ln S_{t-\tau} - (\mu - \frac{1}{2}V_{1,t-\tau} - k(1)V_{2,t-\tau})\tau - J_{d,t} - \sqrt{\tau V_{1,t-\tau}}w_t$, which is normally distributed. Equal weights are attached to particles obtained from the transition law and the conditional posterior;

Mixture Particle Filter II

- compute the particle weight by

$$w_t^{(i)} = \frac{p(\ln S_t | J_{u,t}^{(i)}, J_{d,t}^{(i)}, V_{1,t-\tau}^{(i)}, V_{2,t-\tau}^{(i)}) p(J_{u,t}^{(i)} | V_{2,t-\tau}^{(i)})}{0.5 p(J_{u,t}^{(i)} | V_{2,t-\tau}^{(i)}) + 0.5 \phi(\bar{\mu}, \bar{\sigma})},$$

  where $\phi(\cdot, \cdot)$ represents the normal density with mean $\bar{\mu} = \ln S_t - \ln S_{t-\tau} - (\mu - \frac{1}{2} V_{1,t-\tau}^{(i)} - k(1) V_{2,t-\tau}^{(i)})\tau - J_{d,t}^{(i)}$ and standard deviation $\bar{\sigma} = \sqrt{\tau V_{1,t-\tau}^{(i)}}$;

If $R_t = \ln S_t - \ln S_{t-\tau} < 0$ switch and use the mixture sampler on $J_{d,t}^{(i)}$

GPU Implementation

- Later in the parameter learning algorithm we need to evaluate the PF for thousands of sets of fixed parameters simultaneously
- To make the task manageable we implemented a parallel PF in CUDA where each thread in essence runs a separate filter for a set of fixed parameters

GPU Implementation

- In implementation, when drawing Gamma random numbers, we sample from an approximate Gamma distribution using proposals from the rejection sampling algorithm in Ahrens and Dieter (1974) and Marsaglia and Tsang (2000). However, to keep the algorithm parallel, instead of rejection sampling, we attach importance weights to account for the difference between the proposal and the target gamma.
- Tested both on an NVIDIA GTX 590 and a NVIDIA FERMI 2050 card, got roughly the same speed.

Parameter Learning: Marginalized Resample-Move/SMC^2

- Consider the following decomposition of the target

$$p(\Theta, x_{1:t} \mid y_{1:t}) = p(x_{1:t}|y_{1:t}, \Theta)p(\Theta \mid y_{1:t})$$

  Notice that we can approximate the RHS by the particle filter!

- Run a resample-move algorithm (SMC+MCMC) with N particles over the fixed parameters over the sequence of estimated densities

$$\hat{p}(\Theta|y_{1:t}) \propto \prod_{l=1}^{t} \hat{p}(y_l|y_{1:l-1}, \Theta)p(\Theta)$$

- For further details see Fulop and Li (2011), Chopin et al. (2011)

Properties of Marginalized Resample-Move

- The algorithm delivers exact draws from $p(\Theta, x_{1:t} \mid y_{1:t})$ for any given M as $N \to \infty$.
- Provides estimates of the individual marginal likelihoods

$$f(y_t|y_{1:t-1}) \equiv \int p(y_t|y_{1:t-1},\Theta)p(\Theta|y_{1:t-1})d\Theta$$

These can be used to obtain sequential Bayes factors across different models

$$\mathcal{BF}_t \equiv \frac{p(y_{1:t}|M_1)}{p(y_{1:t}|M_2)} = \frac{p(y_t|y_{1:t-1},M_1)}{p(y_t|y_{1:t-1},M_2)}\mathcal{BF}_{t-1}.$$

- Still M needs to be increased linearly with sample size T to stabilize acceptance rates in parameter move steps.
- Computationally intensive, $M \times N$ particles overall. But massively parallel, so we can use a graphical processing units (GPU) to speed up.

Outline

Data and Parameter settings

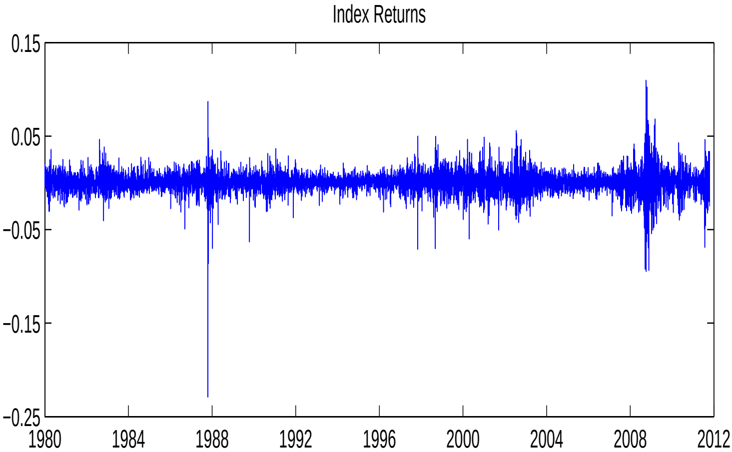
- Daily S& P returns between 1980-2010
- Set particle sizes to $N = 2000, M = 8000$
- Use independent M-H proposals in parameter move sets with multivariate normal proposals calibrated to previous posteriors

Data and Parameter settings

Summary Statistics of S&P500 Returns between 2/1/80-31/10/11

Returns	Mean	Std	Skew	Kur	Min	Max
	.078	.184	-1.193	29.73	-.229	.110
ACF	ρ_1	ρ_2	ρ_3	ρ_4	ρ_5	ρ_6
	-.028	-.044	-.004	-.015	-.016	.008

Data and Parameter settings



Results

Log Bayes Factors at T: Column model against row model

	SE-M1	SE-M2	SE-M3	SE-M4
SE-M1	0.000	—	—	—
SE-M2	12.58	0.000	—	—
SE-M3	0.842	-11.74	0.000	—
SE-M4	13.33	0.758	12.49	0.000

Results

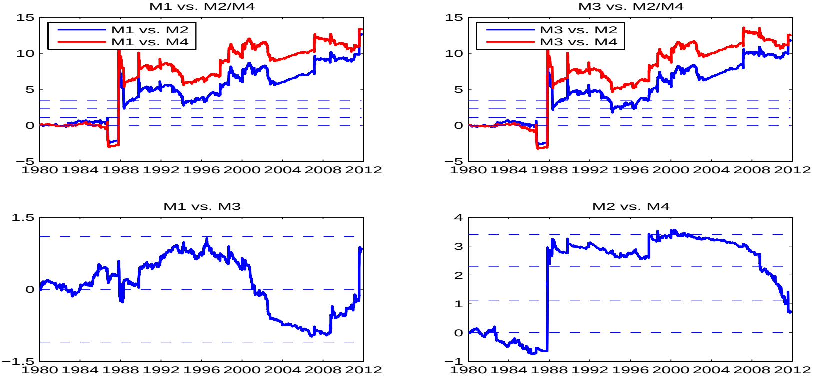


Figure: Sequential Model Comparison

Results

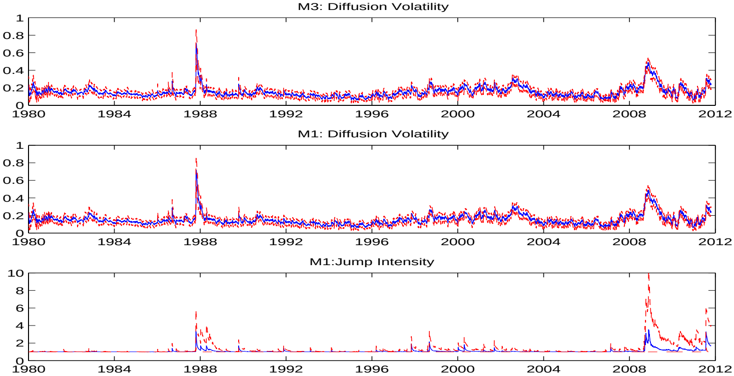


Figure: Filtered Diffusion Volatility and the Jump Intensity

Results

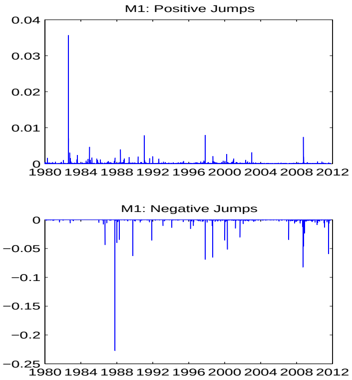


Figure: Filtered Jumps

Results

Value-at-Risk Implied by the Models

	One Day			
	SE-M1	SE-M2	SE-M3	SE-M4
A. 1% VaR				
Average, FS	-0.024	-0.024	-0.024	-0.024
Minimum, FS	-0.110	-0.076	-0.110	-0.076
Average, AL	-0.034	-0.033	-0.033	-0.033
Minimum, AL	-0.079	-0.076	-0.073	-0.076
B. 0.1% VaR				
Average, FS	-0.072	-0.070	-0.067	-0.067
Minimum, FS	-0.295	-0.302	-0.160	-0.148
Average, AL	-0.079	-0.069	-0.068	-0.060
Minimum, AL	-0.149	-0.149	-0.105	-0.106

Results

Value-at-Risk Implied by the Models

	One Week			
	SE-M1	SE-M2	SE-M3	SE-M4
A. 1% VaR				
Average, FS	-0.077	-0.076	-0.076	-0.075
Minimum, FS	-0.315	-0.223	-0.290	-0.202
Average, AL	-0.101	-0.098	-0.097	-0.096
Minimum, AL	-0.214	-0.207	-0.198	-0.202
B. 0.1% VaR				
Average, FS	-0.255	-0.255	-0.254	-0.257
Minimum, FS	-0.834	-0.789	-0.569	-0.572
Average, AL	-0.197	-0.202	-0.185	-0.173
Minimum, AL	-0.342	-0.311	-0.283	-0.282

Results

Impacts of Learning on Option Pricing

	SE-M1			SE-M3		
K/S	7 Days	30 Days	90 Days	7 Days	30 Days	90 Days
0.85	1.113	1.093	1.042	0.999	1.118	1.058
0.90	1.077	1.043	1.020	0.978	1.060	1.031
0.95	1.023	1.004	1.010	0.957	1.012	1.016
1.00	0.997	1.008	1.015	0.999	1.010	1.018
1.05	0.998	1.049	1.035	1.007	1.046	1.036
1.10	1.059	1.091	1.068	1.088	1.084	1.065
1.15	1.168	1.143	1.104	1.478	1.137	1.097

Results

Volatility Forecasting

	SE-M1	SE-M2	SE-M3	SE-M4
2001.01-2011.10	**6.549**	6.846	6.565	6.773
2001.01-2007.12	**5.129**	5.380	5.159	5.366
2008.01-2011.10	**8.825**	9.202	8.826	9.049

Results

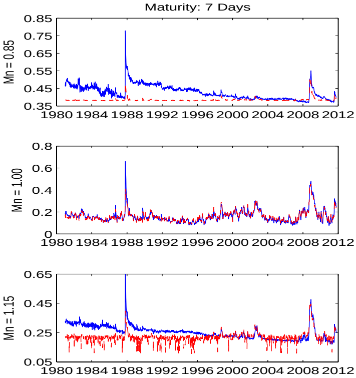


Figure: Implied Volatility for Call Options with Maturity 7 and 30 Days. Blue: learning

Conclusions

- Evidence of jumps feeding back to diffusion volatility is strong throughout the sample
- We find evidence for jumps feeding back to jump intensity, concentrated to the financial crisis period
- Important Risk Management Implications on VaR numbers deep in the tail
- Important implication for option pricing
- Improvement in volatility forecasting