Estimation of Jump Variation in a Bayesian Model of High-Frequency Asset Returns

Brian Donhauser

Department of Economics
University of Washington

June 14, 2012

Outline

Outline

Emergence of High-Frequency Financial Data

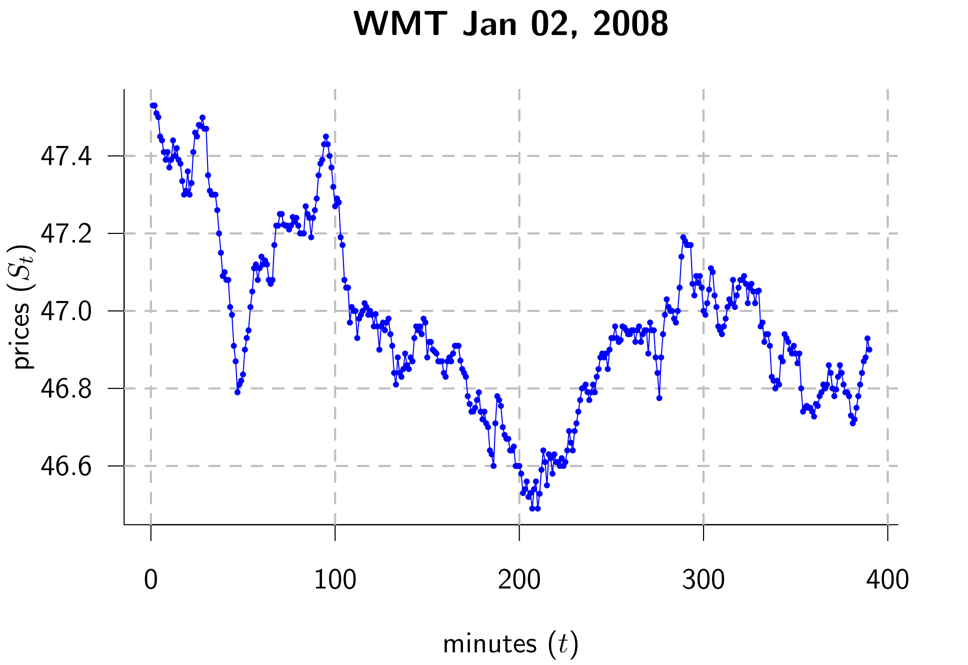
In the early 1990's, developments in

- the storage and vending of high-frequency financial data and
- computing power

laid the groundwork for research in high-frequency finance (Wood, 2000).

What to Do With High-Frequency Financial Data?

- financial economics: test theoretical market microstructure models
- financial econometrics: model high-frequency prices, volumes, and durations
- estimate lower-frequency quantities (e.g., *conditional daily return variance*)



WMT Jan 02, 2008
prices (S_t)
47.4
47.2
47.0
46.8
46.6
0
100
200
300
400
minutes (t)

Simple Returns vs. Log Returns

If S_t is the spot price of an asset at minute t, then the *simple return* over the interval $[t-1, t]$ is

$$R_t = \frac{S_t - S_{t-1}}{S_{t-1}}$$

and the *logarithmic return* is

$$\begin{aligned} r_t &= \log(S_t) - \log(S_{t-1}) \\ &= \log(1 + R_t) \\ &\approx R_t, \qquad \text{for small } R_t \end{aligned}$$

We prefer the latter because of additivity and infinite support.

A Mathematical Model for an Asset X

$$X \equiv (X_t)_{t \in [0,\infty)}$$

- an *adapted, càdlàg stochastic process*
- on a *filtered complete probability space* $(\Omega, \mathcal{F}, \mathbb{F}, P)$, satisfying the *usual hypothesis*
- will represent an asset with log-price X_t at time t.

Toward a Simple Jump-Diffusion Model for X

Under mild regularity conditions:

$$r_t \equiv X_t - X_0 = \int_0^t a_u \mathrm{d}u + \int_0^t \sigma_u \, \mathrm{d}W_u + J_t, \qquad 0 \le t < \infty$$

- r_t is the *log-return* over the trading time interval $[0, t]$
- $W \in \mathcal{BM}$
- $a, \sigma \in \mathcal{PRE}$ are the *spot drift* and *spot volatility* of returns
- $J_t = \sum_{i=1}^{N_t} C_i$ is the *cumulative jump process*
- N_t is a *simple counting process* of the number of jumps in the time interval $[0, t]$
- C_i is the size of the ith jump.

Revisiting the Motivating Question

Q: What is the return variance for WMT on January 2nd, 2008?

Q:

$$\mathrm{Var}\{r_t \mid \mathcal{F}_t\} = ?$$

where

$$\mathcal{F}_t \equiv \mathcal{F}\{a_u, \sigma_u\}_{u \in [0,t]}$$

$\mathrm{Var}\{r \mid \mathcal{F}\} = IV$

From Andersen et al. (2003, Thm. 2), if $(a, \sigma) \perp\!\!\!\perp W$ and $J \equiv 0$,

$$r_t \mid \mathcal{F}_t \sim \mathcal{N}\left(\int_0^t a_u \, \mathrm{d}u, \int_0^t \sigma_u^2 \, \mathrm{d}u \right)$$

so

$$\begin{aligned} \mathrm{E}\{r_t \mid \mathcal{F}_t\} &= \int_0^t a_u \, \mathrm{d}u \\ \mathrm{Var}\{r_t \mid \mathcal{F}_t\} &= IV_t \end{aligned}$$

where

$$IV_t \equiv \int_0^t \sigma_u^2 \, \mathrm{d}t$$

Realized Variance: RV

For

- an asset X
- on a time interval $[0, t]$
- with $(n+1)$-element *partition*

$$\mathcal{P}_n \equiv \{\tau_0, \tau_1, \ldots, \tau_n\}, \qquad 0 = \tau_0 \leq \tau_1 \leq \cdots \leq \tau_n = t < \infty,$$

the *realized variance*

$$RV_t^{(n)} \equiv \sum_{i=1}^{n} r_i^2, \qquad 0 \leq t < \infty$$

where

$$r_i \equiv X_{\tau_i} - X_{\tau_{i-1}}, \qquad i = 1, \ldots, n$$

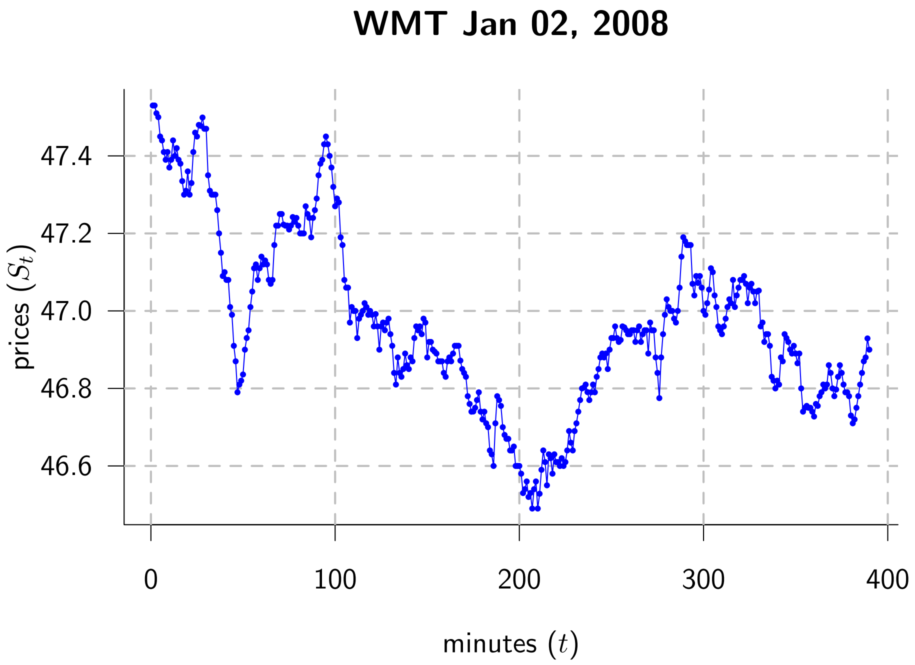
is the i-th log-return on $\mathcal{P}_n$.

$RV \to IV$

From Andersen et al. (2003); Barndorff-Nielsen and Shephard (2002),

$$RV_t^{(n)} \to IV_t$$

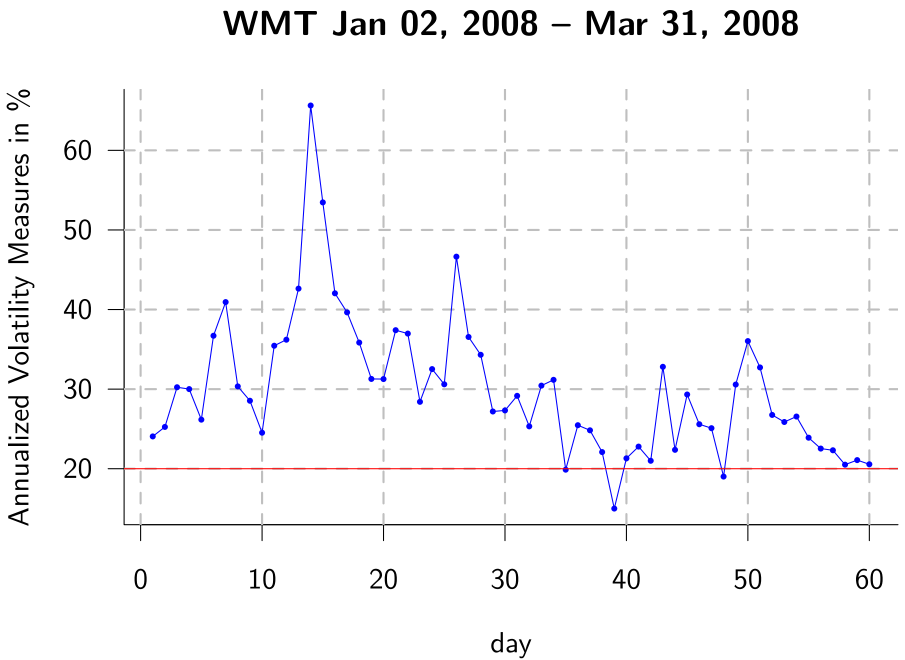
- convergence *ucp* on $[0, t]$
- limit taken over all $\mathcal{P}_n$ with $lim_{n\to\infty} \|\mathcal{P}_n\| = 0$.



WMT Jan 02, 2008
prices (S_t)
47.4
47.2
47.0
46.8
46.6
0
100
200
300
400
minutes (t)

RV for WMT on Jan 2, 2008

	$100\sqrt{\cdot}$	$100\sqrt{252}\sqrt{\cdot}$
RV	1.5	24.1


WMT Jan 02, 2008 – Mar 31, 2008
Annualized Volatility Measures in %
20
30
40
50
60
0
10
20
30
40
50
60
day

Estimating $\mathrm{Var}\{r_t \mid \mathcal{F}_t\}$: Other Methods

- The realized quantity r_t^2
- GARCH (Bollerslev, 1986)
- Stochastic volatility (Melino and Turnbull, 1990; Taylor, 1994; Harvey et al., 1994; Jacquier et al., 1994)

$RV \to IV + JV$

Relaxing $J \equiv 0$, from Barndorff-Nielsen and Shephard (2004),

$$RV_t^{(n)} \to IV_t + JV_t \approx \mathrm{Var}\{r_t \mid \mathcal{F}_t\}$$

where

$$JV_t \equiv \sum_{i=1}^{N_t} C_i^2.$$

Realized Bipower Variation: $RBPV$

For

- an asset X
- on a time interval $[0, t]$
- with $(n+1)$-element *partition* $\mathcal{P}_n$,

the *realized bipower variation*

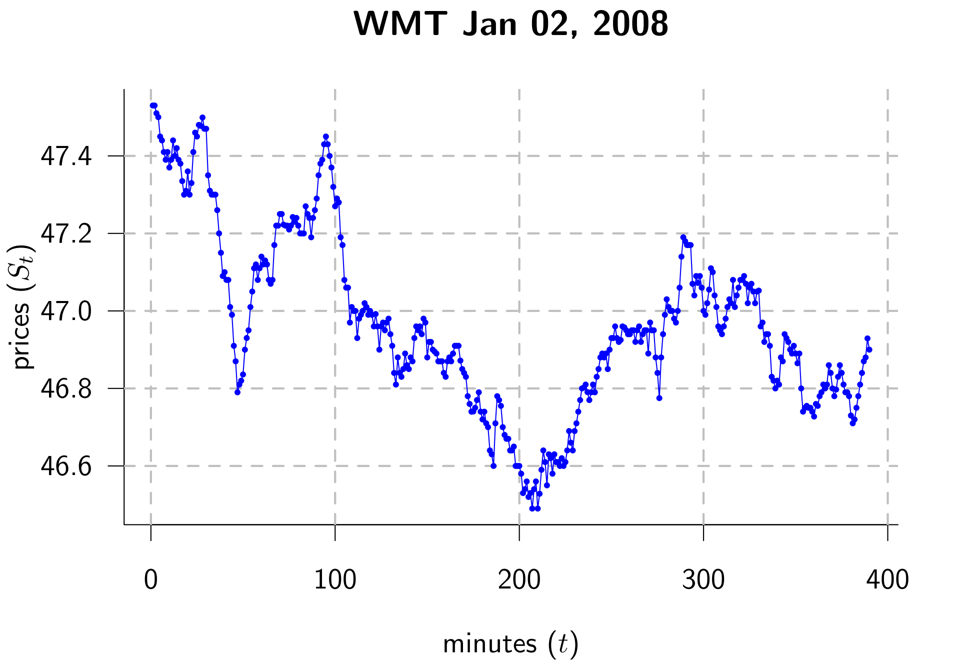
$$RBPV_t^{(n)} \equiv \sum_{i=2}^{n} |r_i||r_{i-1}|, \qquad 0 \le t < \infty.$$

$\widehat{IV}_{\mathrm{BNS04}}, \widehat{JV}_{\mathrm{BNS04}}$

From Barndorff-Nielsen and Shephard (2004), under mild regularity conditions

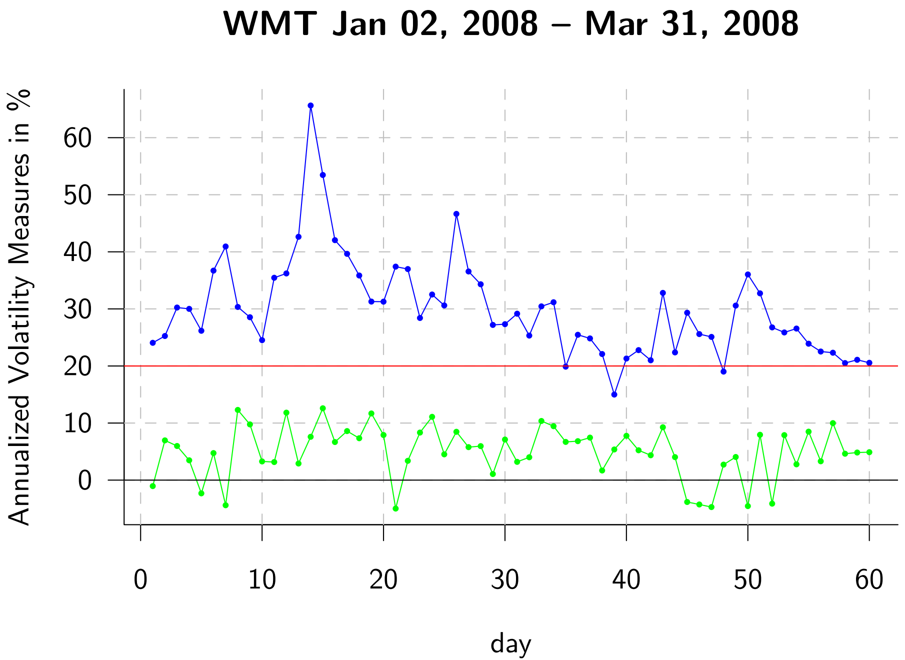
$$\widehat{IV}^{(n)}_{\mathrm{BNS04},t} \equiv \frac{\pi}{2} RBPV_t^{(n)} \to IV_t$$

$$\widehat{JV}^{(n)}_{\mathrm{BNS04},t} \equiv RV_t^{(n)} - \widehat{IV}^{(n)}_{\mathrm{BNS04},t} \to JV_t.$$


WMT Jan 02, 2008
prices (S_t)
47.4
47.2
47.0
46.8
46.6
0
100
200
300
400
minutes (t)

$\widehat{IV}_{\text{BNS04}}, \widehat{JV}_{\text{BNS04}}$ for WMT on Jan 2, 2008

	$100\sqrt{\cdot}$	$100\sqrt{252}\sqrt{\cdot}$	$\%RV$
RV	1.5	24.1	100.0
$\widehat{JV}_{\text{BNS04}}$	–0.1	–1.0	–0.2
$\widehat{IV}_{\text{BNS04}}$	1.5	24.1	100.2


WMT Jan 02, 2008 – Mar 31, 2008
Annualized Volatility Measures in %
0
10
20
30
40
50
60
day
0
10
20
30
40
50
60

Summary of $\widehat{JV}_{\mathrm{BNS04}}$

WMT: Jan 2, 2008 – Mar 31, 2008

	$100 * \left(\frac{\widehat{JV}_{\mathrm{BNS04}}}{RV}\right)$
Min.	-3.5
1st Qu.	1.2
Median	3.6
Mean	4.8
3rd Qu.	8.1
Max.	20.0

Simulation Results for $\widehat{JV}_{\mathrm{BNS04}}$: MPE and MAPE

	# Jumps			
Est.	0	3	10	30
$\widehat{JV}_{\mathrm{BNS04}}$	0.3	–4.5	–12.7	–26.6
$\widehat{IV}_{\mathrm{BNS04}}$	–0.4	4.1	12.7	26.4

	# Jumps			
Est.	0	3	10	30
$\widehat{JV}_{\mathrm{BNS04}}$	3.3	5.4	12.7	26.6
$\widehat{IV}_{\mathrm{BNS04}}$	6.7	7.3	12.9	26.4

Estimating Jump Locations: LM Stat

Extending Barndorff-Nielsen and Shephard, Lee and Mykland (2008) define their jump statistic

$$LM_{l,\tau_i} \equiv \frac{r_i}{\widehat{\sigma_i}},$$

where

$$\widehat{\sigma_i}^2 \equiv \frac{1}{n}\widehat{IV}_{\mathrm{BNS04},t}$$

and show that jumps are identified with perfect classification accuracy asymptotically.

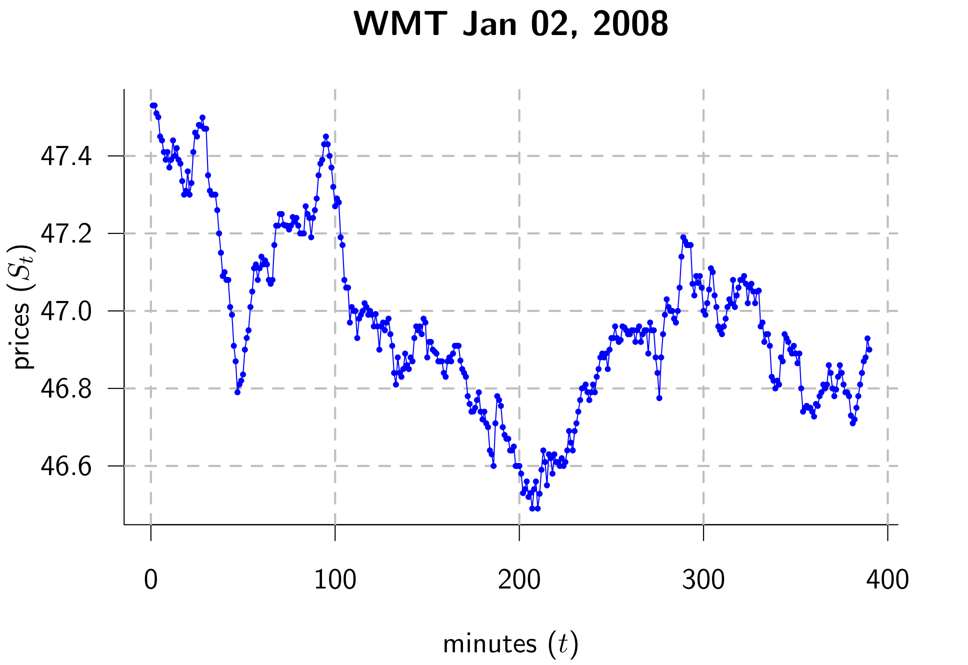
Naïve Shrinkage Estimators: $\widehat{JV}_{\text{NS}}, \widehat{IV}_{\text{NS}}$

In previous work, I extended Lee and Mykland (2008) and defined *naïve shrinkage estimators* of JV, IV:

$$\widehat{JV}^{(n)}_{\text{NS},t} \equiv \sum_{i=1}^{n} \left[F_{\xi}(LM_{g,\tau_i})\, r_i \right]^2$$

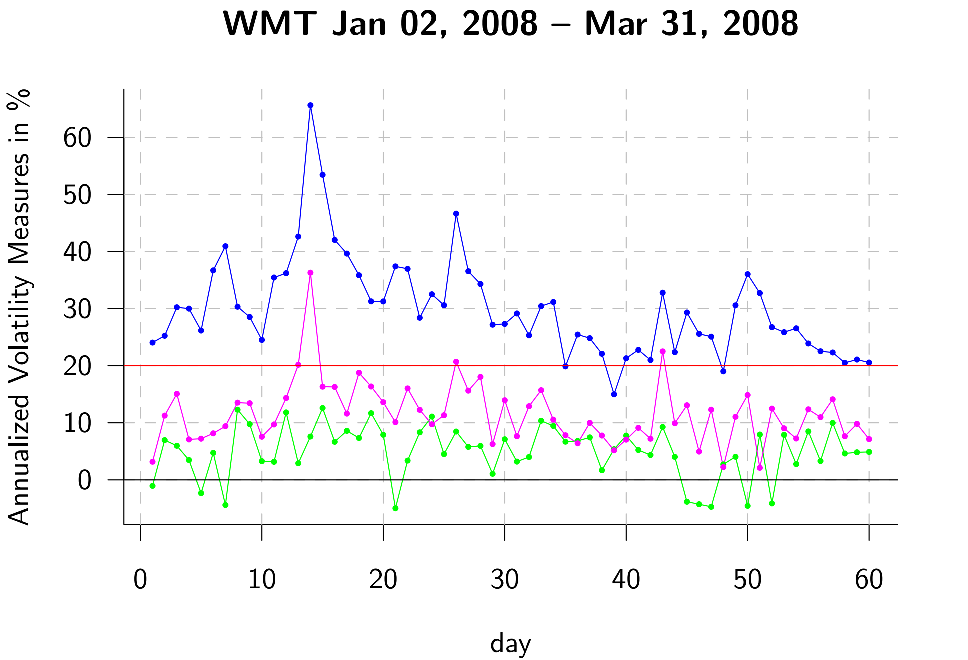
$$\widehat{IV}^{(n)}_{\text{NS},t} \equiv RV^{(n)}_t - \widehat{JV}_{\text{NS},t}.$$

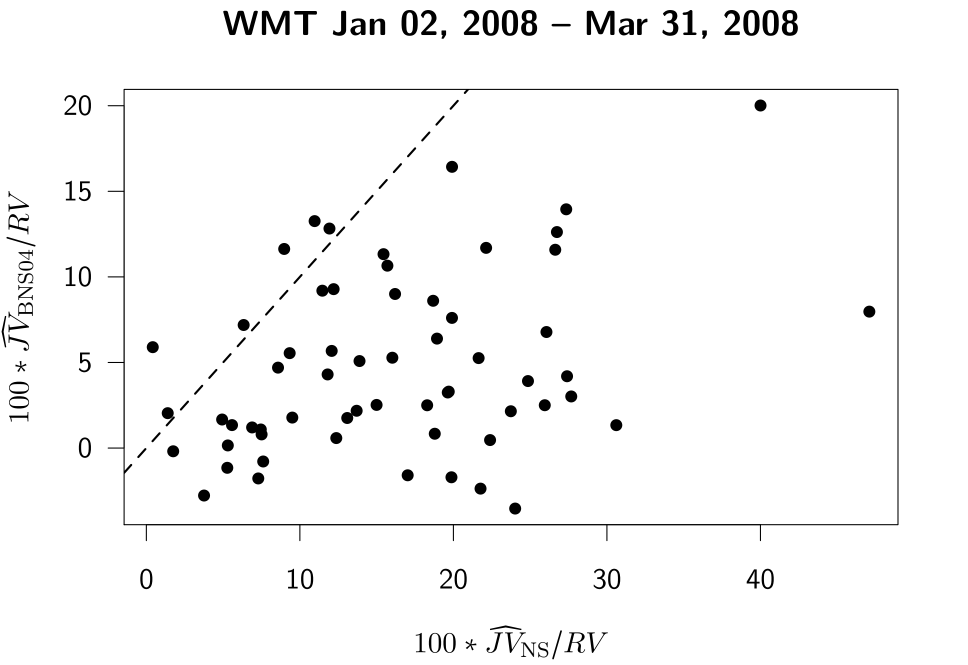
and showed their consistency.


WMT Jan 02, 2008
prices (S_t)
47.4
47.2
47.0
46.8
46.6
0
100
200
300
400
minutes (t)

$\widehat{IV}_{\mathrm{NS}}, \widehat{JV}_{\mathrm{NS}}$ for WMT on Jan 2, 2008

	$100\sqrt{\cdot}$	$100\sqrt{252}\sqrt{\cdot}$	$\% RV$
RV	1.5	24.1	100.0
$\widehat{JV}_{\mathrm{NS}}$	0.2	3.2	1.7
$\widehat{JV}_{\mathrm{BNS04}}$	−0.1	−1.0	−0.2
$\widehat{IV}_{\mathrm{NS}}$	1.5	23.8	98.3
$\widehat{IV}_{\mathrm{BNS04}}$	1.5	24.1	100.2


WMT Jan 02, 2008 – Mar 31, 2008
Annualized Volatility Measures in %
0
10
20
30
40
50
60
day


WMT Jan 02, 2008 – Mar 31, 2008
20
15
10
5
0
$100 * \widehat{JV}_{\mathrm{BNS04}}/RV$
0
10
20
30
40
$100 * \widehat{JV}_{\mathrm{NS}}/RV$

Summary of $\widehat{JV}_{\text{NS}}$

WMT: Jan 2, 2008 – Mar 31, 2008

	$100 * \left(\frac{\widehat{JV}_{\text{BNS04}}}{RV}\right)$	$100 * \left(\frac{\widehat{JV}_{\text{NS}}}{RV}\right)$
Min.	-3.5	0.4
1st Qu.	1.2	8.9
Median	3.6	15.6
Mean	4.8	16.1
3rd Qu.	8.1	21.9
Max.	20.0	47.1

Simulation Results for $\widehat{JV}_{\mathrm{NS}}$: MPE and MAPE

	# Jumps			
Est.	0	3	10	30
$\widehat{JV}_{\mathrm{NS}}$	1.3	−0.0	−2.1	−10.3
$\widehat{JV}_{\mathrm{BNS04}}$	0.3	−4.5	−12.7	−26.6
$\widehat{IV}_{\mathrm{NS}}$	−1.4	−0.4	2.1	10.0
$\widehat{IV}_{\mathrm{BNS04}}$	−0.4	4.1	12.7	26.4

	# Jumps			
Est.	0	3	10	30
$\widehat{JV}_{\mathrm{NS}}$	1.3	2.7	4.4	10.6
$\widehat{JV}_{\mathrm{BNS04}}$	3.3	5.4	12.7	26.6
$\widehat{IV}_{\mathrm{NS}}$	5.9	5.3	4.7	10.1
$\widehat{IV}_{\mathrm{BNS04}}$	6.7	7.3	12.9	26.4

Conclusion: Naïve Shrinkage Estimators

- Simulations demonstrate superiority over $\widehat{JV}_{\mathrm{BNS04}}$:)
- Bounded above zero and below RV :)
- Non-model-based :(

Outline

Continuous → Discrete: The Reduced Model

Under assumptions and definitions to follow, the continuous jump-diffusion model reduces to a discrete model where we observe $r_i = \sigma_i x_i$, with

$$x_i = \mu_i + \epsilon_i, \qquad i = 1, \ldots, n,$$

and

$$\epsilon_i \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1) \quad \text{with pdf } \phi$$

where μ_i and σ_i are unspecified jump and stochastic volatility processes.

Continuous → Discrete: Assumptions and Definitions

Assume,

(i) (Zero Drift). $a_u = 0,\ \forall u \in [0, t]$,

(ii) (Homogenous Sampling). An $(n+1)$-element homogenous sampling of X_u over the time interval $[0, t]$, $\{X_0, X_\delta, X_{2\delta}, \dots, X_{(n-1)\delta}, X_t\}$, where $\delta = t/n$ is the width of the sampling interval,

and define for $i = 1, \dots, n$,

(i) (Interval Volatility). $\sigma_i \equiv \sigma_{[\delta(i-1),\, \delta i]}$.

(ii) (Scaled Interval Jump). $\mu_i \equiv \frac{1}{\sigma_i}(J_{\delta i} - J_{\delta(i-1)})$.

(iii) (Log-Returns). $r_i \equiv X_{\delta i} - X_{\delta(i-1)}$.

(iv) (Scaled Log-Returns). $x_i \equiv \frac{1}{\sigma_i} r_i$.

JV and IV in the Discrete, Reduced Model

$$
\begin{aligned}
JV &\approx \sum_{i=1}^{n} \sigma_i^2 \mu_i^2 \\
IV &\approx \sum_{i=1}^{n} \sigma_i^2 \\
QV &\approx \sum_{i=1}^{n} \sigma_i^2 + \sum_{i=1}^{n} \sigma_i^2 \mu_i^2 .
\end{aligned}
$$

We take these to *define* JV and IV in the discrete model.

The Discrete Bayesian Model

From Johnstone and Silverman (2004, 2005),

$$r_i = \sigma_i x_i \qquad i = 1, \ldots, n, \qquad \text{(Observations)}$$

$$\mathcal{L}(x_i \mid \mu_i, \sigma_i) = \phi(x_i - \mu_i) \qquad i = 1, \ldots, n, \qquad \text{(Likelihood)}$$

$$\pi(\mu_i) = \begin{cases} 1 - w & \text{for } \mu_i = 0 \\ w\gamma(\mu_i) & \text{for } \mu_i \neq 0, \end{cases} \qquad i = 1, \ldots, n,$$

(Prior Jump Density)

$$w = \pi\left\{\mu \neq 0\right\}, \qquad \text{(Prior Jump Probability)}$$

$$\gamma(\mu) = \frac{1}{2} a \exp(-a|\mu|), \qquad \text{(Prior Density of } \mu \mid \{\mu \neq 0\}\text{)}$$

$$\sigma_i, w, a \qquad \text{(Hyperparemeters)}$$

Bayesian Model Estimation: Initial Steps

1. Set $a = 0.5$ a priori
2. Assume $\sigma_i = \sigma$ a constant and take $\sigma = 1.48\,\mathrm{MAD}\{r_1, \ldots, r_n\}$
3. Calculate $x_i = \frac{r_i}{\sigma}$ for $i = 1, \ldots, n$
4. Calculate w by marginal maximum likelihood of $\mathcal{L}(\breve{w})$. I.e., take

   $$w = \operatorname*{argmax}_{0 \le \breve{w} \le 1} \sum_{i=1}^{n} \log\{(1-\breve{w})\phi(x_i) + \breve{w}g(x_i)\}$$

   where

   $$g(x) \equiv (\phi \star \gamma)(x) \equiv \int \phi(x-\mu)\gamma(\mu)\,\mathrm{d}\mu$$

   is interpreted as the *marginal density of* $x \mid \{\mu \neq 0\}$
5. For $i = 1, \ldots, n$, solve for the posterior density of $\mu_i \mid x_i \ldots$

Posterior Density of $\mu \mid x$

$\pi(\mu \mid x)$

After much algebra, we can write the posterior density of $\mu \mid x$ as

$$\pi(\mu \mid x) = \begin{cases} 1 - w(x) & \text{for } \mu = 0 \\ w(x)\gamma(\mu \mid x) & \text{for } \mu \neq 0, \end{cases}$$

where $\gamma(\mu \mid x)$ is the *posterior density of* $\mu \mid \{x, \mu \neq 0\}$ and $w(x) \equiv \pi(\mu \neq 0 \mid x)$ is the *posterior non-zero jump probability*

Posterior Density of $\mu \mid \{x, \mu \neq 0\}$ for Laplace prior

$\gamma(\mu \mid x)$

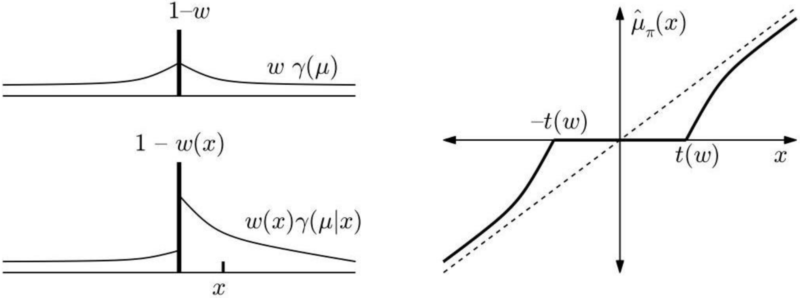
For the Laplace prior,

$$\gamma(\mu \mid x) = \begin{cases} \dfrac{e^{-ax}}{D}\phi(\mu - x + a) & \text{for } \mu > 0 \\ \dfrac{e^{ax}}{D}\phi(\mu - x - a) & \text{for } \mu \leq 0, \end{cases}$$

where $D = e^{-ax}\Phi(x - a) + e^{ax}\tilde{\Phi}(x + a)$.

Posterior Density Illustration

From Johnstone (2011):



Posterior Mean of $\mu \mid x$ for Laplace Prior

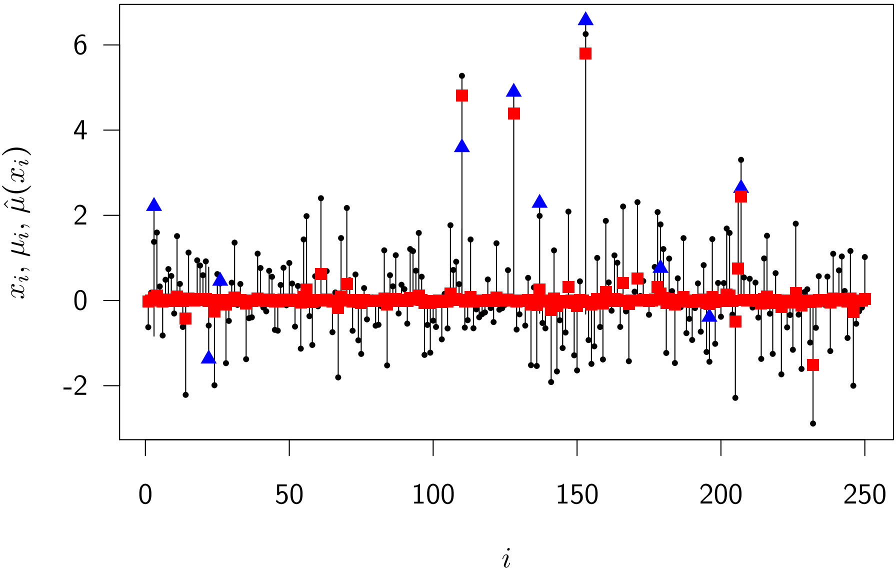
Recall the posterior density of $\mu \mid x$ is given by

$$\pi(\mu \mid x) = \begin{cases} 1 - w(x) & \text{for } \mu = 0 \\ w(x)\gamma(\mu \mid x) & \text{for } \mu \neq 0. \end{cases}$$

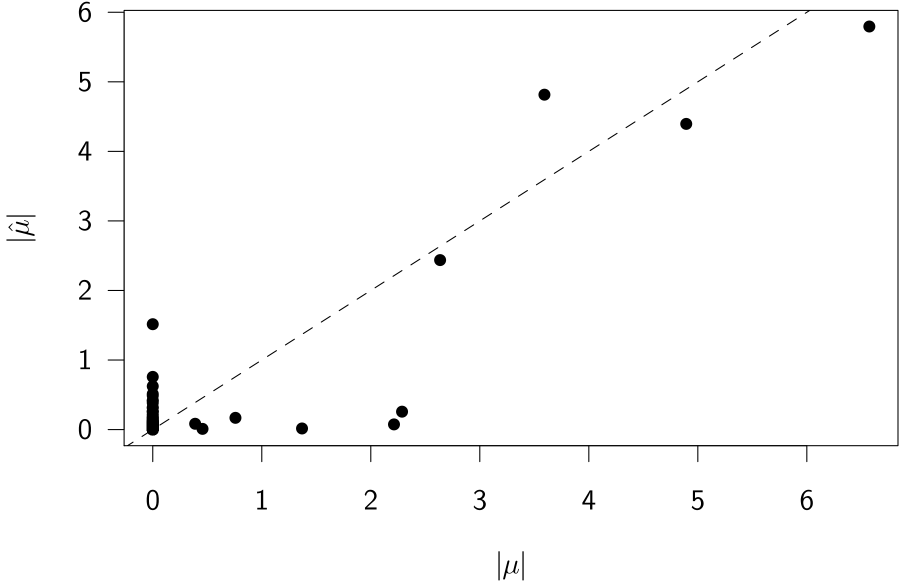
Then the posterior mean of $\mu \mid x$ is

$$\begin{aligned} \hat{\mu}_\pi(x) &= w(x)\hat{\mu}_\gamma(x) \\ &= \underbrace{w(x)}_{\to 1} \underbrace{\left(x - a \frac{\{e^{-ax}\Phi(x-a) - e^{ax}\tilde{\Phi}(x+a)\}}{e^{-ax}\Phi(x-a) + e^{ax}\tilde{\Phi}(x+a)} \right)}_{\to (x-a)} \\ &\to x - a \quad \text{for large } x \end{aligned}$$

in the case of Laplace prior density $\gamma(\mu)$



6
4
2
0
-2
0
50
100
150
200
250
$x_i,\ \mu_i,\ \hat{\mu}(x_i)$
i



6
5
4
3
2
1
0
$|\hat{\mu}|$
0
1
2
3
4
5
6
$|\mu|$

Outline

Posterior Mean of $JV \mid \mathbf{x}$: Naïve Approach

$\widehat{JV}_{\text{NEB},\ell_2}(\mathbf{x})$

Recall that

$$JV \equiv \sum_{i=1}^{n} \sigma_i^2 \mu_i^2.$$

Then, a naïve estimation approach gives

$$\widehat{JV}_{\text{NEB},\ell_2}(\mathbf{x}) \equiv \sum_{i=1}^{n} \sigma_i^2 \hat{\mu}_{\pi,i}(x_i)^2.$$

From Jensen's inequality we expect this to under-represent the actual posterior mean of JV: $\widehat{JV}_{\text{EB},\ell_2}(\mathbf{x})$. But, very easy to calculate!

Posterior Mean of $JV \mid \mathbf{x}$: Naïve Relation

$\widehat{JV}_{\mathrm{EB},\ell_2}(\mathbf{x})$

$$\begin{aligned}
\widehat{JV}_{\mathrm{EB},\ell_2}(\mathbf{x}) &\equiv \mathrm{E}\left[JV \mid \mathbf{x}\right] \\
&= \mathrm{E}\left[\sum_{i=1}^{n} \sigma_i^2 \mu_i^2 \mid \mathbf{x}\right] \\
&= \sum_{i=1}^{n} \sigma_i^2\, \mathrm{E}\left[\mu_i^2 \mid x_i\right] \\
&= \sum_{i=1}^{n} \sigma_i^2 \hat{\mu}_{\pi,i}(x_i)^2 + \sigma_i^2\, \mathrm{Var}[\mu_i \mid x_i] \\
&= \widehat{JV}_{\mathrm{NEB},\ell_2}(\mathbf{x}) + \sum_{i=1}^{n} \sigma_i^2\, \mathrm{Var}[\mu_i \mid x_i].
\end{aligned}$$

Posterior Mean of $JV \mid \mathbf{x}$ for Laplace Prior

$\widehat{JV}_{\mathrm{EB},\ell_2}(\mathbf{x})$

So then,

$$\widehat{JV}_{\mathrm{EB},\ell_2}(\mathbf{x}) = \sum_{i=1}^{n} \sigma_i^2 \widehat{\mu_{\pi,i}^2}(x_i)$$

where, after skipping a large amount of algebra,

$$\begin{aligned}
\widehat{\mu_\pi^2}(x) &= w(x)\widehat{\mu_\gamma^2}(x) \\
&= w(x)\left(x^2 + a^2 + 1 - 2ax\left\{\frac{e^{-ax}\Phi(x-a) - e^{ax}\tilde{\Phi}(x+a)}{D}\right\}\right. \\
&\qquad \left. -2a\left\{\frac{e^{-ax}\phi(x-a)}{D}\right\}\right) \\
&\to (x-a)^2 + 1 \quad \text{for large } x \\
&= \hat{\mu}_\pi(x)^2 + 1 \qquad \text{for large } x
\end{aligned}$$

for the Laplace prior. Closed form!

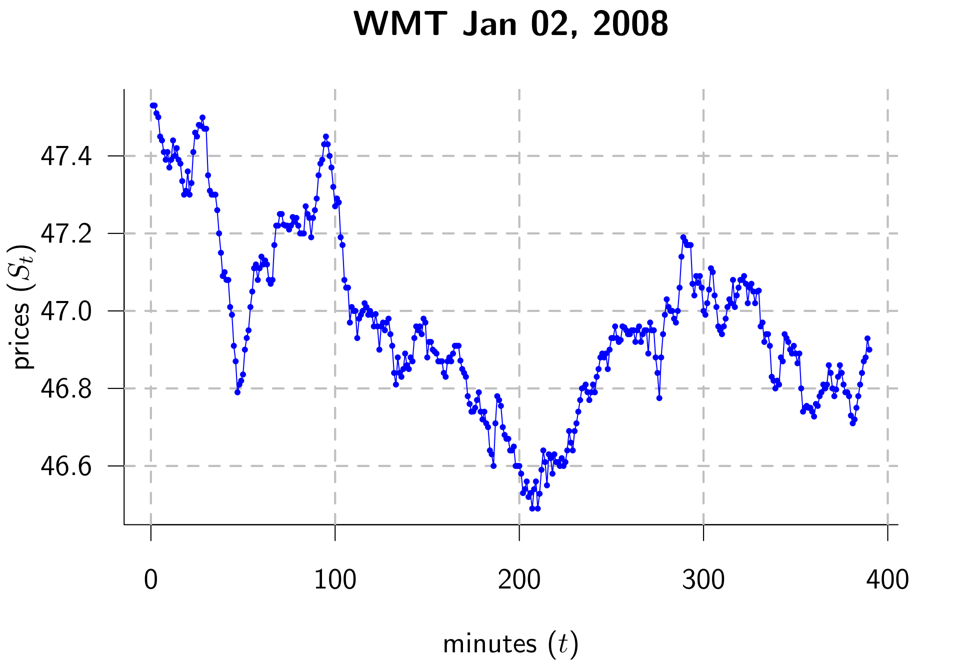
Posterior Median of $JV \mid \mathbf{x}$ for Laplace Prior

$\widehat{JV}_{\mathrm{EB},\ell_1}(\mathbf{x})$

No simple closed form result. Can still get a numerical result via the following procedure:

1. Simulate one value $\mu_{i,1}$ from each of the n posterior jump densities $\pi_i(\mu_i \mid x_i)$
2. Calculate $JV_1 = \sum_{i=1}^{n} \sigma_i^2 \mu_{i,1}^2$
3. Repeat steps 1-2 k-times, returning $\{JV_1, \ldots, JV_k\}$
4. Then, for large k, $\widehat{JV}_{\mathrm{EB},\ell_1}(\mathbf{x}) \approx \mathrm{Median}\{JV_1, \ldots, JV_k\}$

Outline



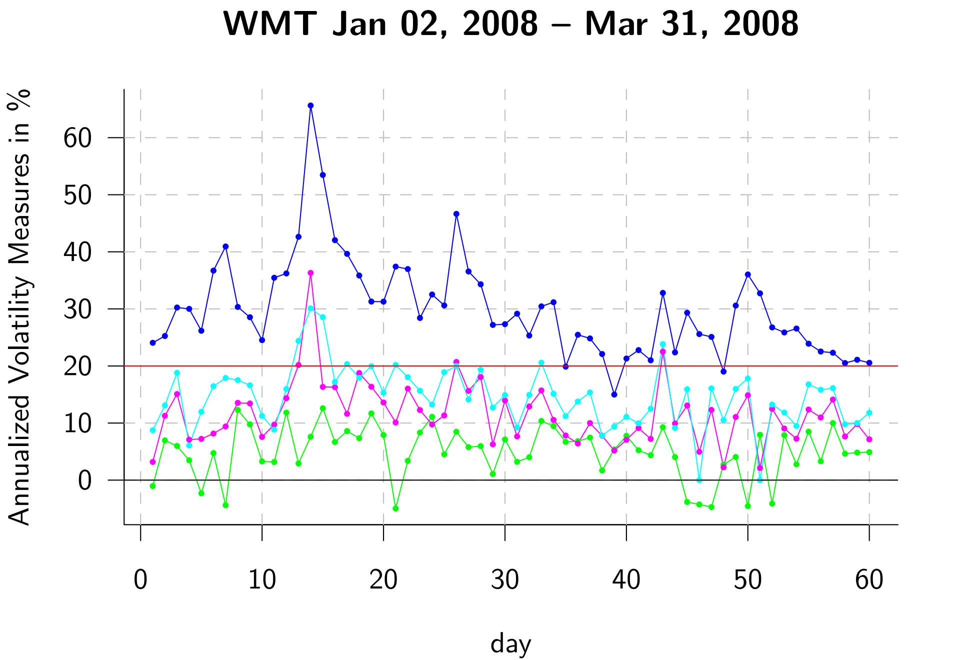
WMT Jan 02, 2008
prices (S_t)
47.4
47.2
47.0
46.8
46.6
0
100
200
300
400
minutes (t)

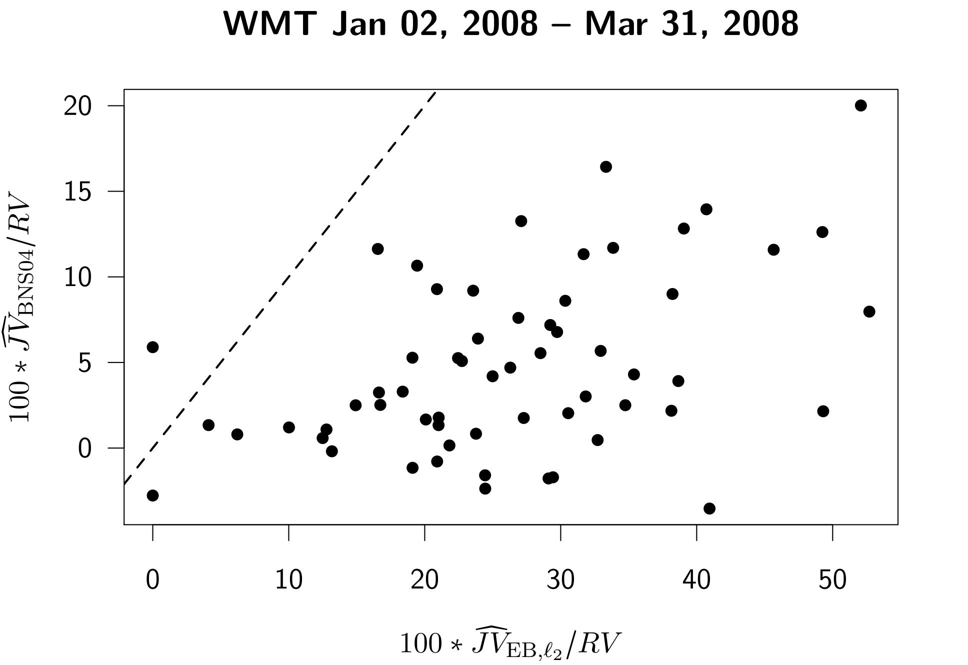
$\widehat{JV}_{\mathrm{EB},\ell_2}$ for WMT on Jan 2, 2008

	Window	$100\sqrt{\cdot}$	$100\sqrt{252}\sqrt{\cdot}$	$\%RV$
$\widehat{JV}_{\mathrm{EB},\ell_2}$	full	0.8	12.1	25.4
	60min	0.7	11.7	23.8
	30min	0.6	8.7	13.2
	15min	0.5	8.4	12.1
	10min	0.6	9.4	15.3
	05min	0.7	11.6	23.3
$\widehat{IV}_{\mathrm{EB},\ell_2}$	full	1.3	20.8	74.6
	60min	1.3	21.0	76.2
	30min	1.4	22.4	86.8
	15min	1.4	22.6	87.9
	10min	1.4	22.1	84.7
	05min	1.3	21.1	76.7

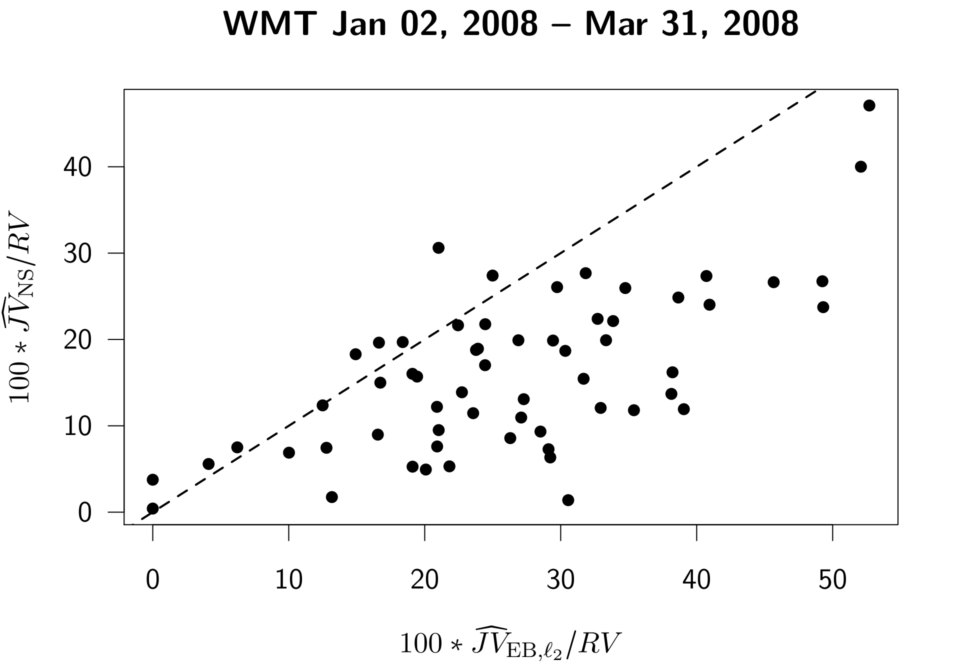
$\widehat{IV}_{\mathrm{EB}}, \widehat{JV}_{\mathrm{EB}}$ for WMT on Jan 2, 2008

	$100\sqrt{\cdot}$	$100\sqrt{252}\sqrt{\cdot}$	$\%RV$
RV	1.5	24.1	100.0
$\widehat{JV}_{\mathrm{EB},\ell_1}$	0.5	8.5	12.5
$\widehat{JV}_{\mathrm{EB},\ell_2}$	0.6	8.7	13.2
$\widehat{JV}_{\mathrm{NS}}$	0.2	3.2	1.7
$\widehat{JV}_{\mathrm{BNS04}}$	−0.1	−1.0	−0.2
$\widehat{IV}_{\mathrm{EB},\ell_1}$	1.4	22.5	87.5
$\widehat{IV}_{\mathrm{EB},\ell_2}$	1.4	22.4	86.8
$\widehat{IV}_{\mathrm{NS}}$	1.5	23.8	98.3
$\widehat{IV}_{\mathrm{BNS04}}$	1.5	24.1	100.2


WMT Jan 02, 2008 – Mar 31, 2008
Annualized Volatility Measures in %
0
10
20
30
40
50
60
day
0
10
20
30
40
50
60



WMT Jan 02, 2008 – Mar 31, 2008
20
15
10
5
0
0
10
20
30
40
50
$100 * \widehat{JV}_{\mathrm{BNS04}}/RV$
$100 * \widehat{JV}_{\mathrm{EB},\ell_2}/RV$


WMT Jan 02, 2008 – Mar 31, 2008
0
10
20
30
40
50
$100 * \widehat{JV}_{\mathrm{EB},\ell_2}/RV$
$100 * \widehat{JV}_{\mathrm{NS}}/RV$

Summary of $\widehat{JV}_{\mathrm{EB},\ell_2}$

WMT: Jan 2, 2008 – Mar 31, 2008

	$100 * \left(\frac{\widehat{JV}_{\mathrm{BNS04}}}{RV}\right)$	$100 * \left(\frac{\widehat{JV}_{\mathrm{NS}}}{RV}\right)$	$100 * \left(\frac{\widehat{JV}_{\mathrm{EB},\ell_2}}{RV}\right)$
Min.	–3.5	0.4	0.0
1st Qu.	1.2	8.9	19.4
Median	3.6	15.6	25.6
Mean	4.8	16.1	26.3
3rd Qu.	8.1	21.9	33.0
Max.	20.0	47.1	52.7

Outline

Simulation Model

(**Number of Observations**). $n = 390$.

(**Observations**). For $i = 1, \ldots, n$,

$$r_i = \sigma_i x_i,$$

where

$$x_i = \mu_i + \epsilon_i, \qquad \epsilon_i \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1).$$

(**Interval Volatility**). $\sigma_i = \sigma = 0.000638 \approx 0.20/\sqrt{252 * 390}$.

(**Scaled Jumps**). $\mu_i \mid \mu_i \neq 0 \sim \mathcal{U}(-s, s)$, with s deterministically set equal to 4, 7, 10, and 15.

(**Number of Jumps**). Deterministically set equal to 0, 3, 10, and 30.

(**Jump Locations**). Each uniformly chosen from $i = 1, \ldots, n$

(**Number of Simulations**). $m = 5000$

Estimator Evaluation Criteria: MPE, MAPE

We define the *mean percentage error* and *mean absolute percentage error* for an estimator $\widehat{JV}$ of JV as

$$\mathrm{MPE}(\widehat{JV}) = 100 * \mathrm{E}\left[\frac{\widehat{JV} - JV}{QV}\right]$$

$$\mathrm{MAPE}(\widehat{JV}) = 100 * \mathrm{E}\left[\frac{|\widehat{JV} - JV|}{QV}\right]$$

and similarly for an estimator $\widehat{IV}$ of IV

$$\mathrm{MPE}(\widehat{IV}) = 100 * \mathrm{E}\left[\frac{\widehat{IV} - IV}{QV}\right]$$

$$\mathrm{MAPE}(\widehat{IV}) = 100 * \mathrm{E}\left[\frac{|\widehat{IV} - IV|}{QV}\right]$$

Estimator Evaluation Criteria: Sample MPE and MAPE

We approximate MPE and MAPE by their sample counterparts

$$\mathrm{MPE}(\widehat{JV}) \approx \frac{100}{m} \sum_{j=1}^{m} \frac{\widehat{JV_j} - JV_j}{QV_j}$$

$$\mathrm{MAPE}(\widehat{JV}) \approx \frac{100}{m} \sum_{j=1}^{m} \frac{|\widehat{JV_j} - JV_j|}{QV_j}$$

$$\mathrm{MPE}(\widehat{IV}) \approx \frac{100}{m} \sum_{j=1}^{m} \frac{\widehat{IV_j} - IV_j}{QV_j}$$

$$\mathrm{MAPE}(\widehat{IV}) \approx \frac{100}{m} \sum_{j=1}^{m} \frac{|\widehat{IV_j} - IV_j|}{QV_j}$$

MPE for $s = 4$

	# Jumps			
Est.	0	3	10	30
$\widehat{JV}_{\mathrm{NEB},\ell_2}$	0.3	−2.1	−6.5	−15.5
$\widehat{JV}_{\mathrm{EB},\ell_2}$	0.8	−0.4	−2.1	−6.3
$\widehat{JV}_{\mathrm{NS}}$	1.3	−0.5	−4.0	−15.1
$\widehat{JV}_{\mathrm{BNS04}}$	0.2	−2.0	−7.3	−19.1
$\widehat{IV}_{\mathrm{NEB},\ell_2}$	−0.2	2.0	6.9	15.6
$\widehat{IV}_{\mathrm{EB},\ell_2}$	−0.7	0.3	2.5	6.4
$\widehat{IV}_{\mathrm{NS}}$	−1.4	0.1	3.9	14.8
$\widehat{IV}_{\mathrm{BNS04}}$	−0.1	1.9	7.7	19.1

MAPE for $s = 4$

	# Jumps			
Est.	0	3	10	30
$\widehat{JV}_{\mathrm{NEB},\ell_2}$	0.3	2.7	6.8	15.5
$\widehat{JV}_{\mathrm{EB},\ell_2}$	0.8	3.3	5.6	8.5
$\widehat{JV}_{\mathrm{NS}}$	1.3	2.1	4.8	15.1
$\widehat{JV}_{\mathrm{BNS04}}$	3.3	3.7	7.6	19.1
$\widehat{IV}_{\mathrm{NEB},\ell_2}$	5.8	5.9	7.9	15.7
$\widehat{IV}_{\mathrm{EB},\ell_2}$	6.0	6.2	6.7	8.5
$\widehat{IV}_{\mathrm{NS}}$	5.9	5.7	6.4	14.9
$\widehat{IV}_{\mathrm{BNS04}}$	6.7	6.4	8.7	19.2

MPE for $s = 7$

	# Jumps			
Est.	0	3	10	30
$\widehat{JV}_{\mathrm{NEB},\ell_2}$	0.3	−2.1	−5.6	−11.4
$\widehat{JV}_{\mathrm{EB},\ell_2}$	0.8	1.4	1.2	−1.7
$\widehat{JV}_{\mathrm{NS}}$	1.3	−0.0	−2.1	−10.3
$\widehat{JV}_{\mathrm{BNS04}}$	0.2	−4.3	−12.6	−26.1
$\widehat{IV}_{\mathrm{NEB},\ell_2}$	−0.2	2.2	5.4	11.5
$\widehat{IV}_{\mathrm{EB},\ell_2}$	−0.7	−1.3	−1.4	1.8
$\widehat{IV}_{\mathrm{NS}}$	−1.4	−0.4	2.1	10.0
$\widehat{IV}_{\mathrm{BNS04}}$	−0.1	4.4	12.3	26.2

MAPE for $s = 7$

Est.	# Jumps 0	3	10	30
$\widehat{JV}_{\mathrm{NEB},\ell_2}$	0.3	3.3	6.3	11.4
$\widehat{JV}_{\mathrm{EB},\ell_2}$	0.8	3.9	4.8	4.7
$\widehat{JV}_{\mathrm{NS}}$	1.3	2.7	4.4	10.6
$\widehat{JV}_{\mathrm{BNS04}}$	3.3	5.3	12.6	26.1
$\widehat{IV}_{\mathrm{NEB},\ell_2}$	5.8	5.8	6.5	11.5
$\widehat{IV}_{\mathrm{EB},\ell_2}$	6.0	6.1	5.1	3.7
$\widehat{IV}_{\mathrm{NS}}$	5.9	5.3	4.7	10.1
$\widehat{IV}_{\mathrm{BNS04}}$	6.7	7.3	12.6	26.2

MPE for $s = 10$

	# Jumps			
Est.	0	3	10	30
$\widehat{JV}_{\mathrm{NEB},\ell_2}$	0.3	−2.3	−6.0	−9.6
$\widehat{JV}_{\mathrm{EB},\ell_2}$	0.8	1.4	−0.1	−2.7
$\widehat{JV}_{\mathrm{NS}}$	1.3	0.2	−1.0	−6.4
$\widehat{JV}_{\mathrm{BNS04}}$	0.2	−6.3	−15.6	−27.4
$\widehat{IV}_{\mathrm{NEB},\ell_2}$	−0.2	2.3	6.0	9.8
$\widehat{IV}_{\mathrm{EB},\ell_2}$	−0.7	−1.4	0.1	2.9
$\widehat{IV}_{\mathrm{NS}}$	−1.4	−0.6	1.0	6.1
$\widehat{IV}_{\mathrm{BNS04}}$	−0.1	6.3	15.5	27.7

MAPE for $s = 10$

	# Jumps			
Est.	0	3	10	30
$\widehat{JV}_{\text{NEB},\ell_2}$	0.3	3.9	6.5	9.7
$\widehat{JV}_{\text{EB},\ell_2}$	0.8	4.0	4.2	4.3
$\widehat{JV}_{\text{NS}}$	1.3	3.1	4.2	6.9
$\widehat{JV}_{\text{BNS04}}$	3.3	7.1	15.6	27.4
$\widehat{IV}_{\text{NEB},\ell_2}$	5.8	5.0	6.3	9.8
$\widehat{IV}_{\text{EB},\ell_2}$	6.0	5.1	3.6	3.3
$\widehat{IV}_{\text{NS}}$	5.9	4.8	3.5	6.2
$\widehat{IV}_{\text{BNS04}}$	6.7	7.8	15.6	27.7

MPE for $s = 15$

	# Jumps			
Est.	0	3	10	30
$\widehat{JV}_{\mathrm{NEB},\ell_2}$	0.3	−2.7	−5.4	−7.9
$\widehat{JV}_{\mathrm{EB},\ell_2}$	0.8	0.6	−1.3	−4.0
$\widehat{JV}_{\mathrm{NS}}$	1.3	0.2	−0.4	−3.7
$\widehat{JV}_{\mathrm{BNS04}}$	0.2	−7.8	−16.0	−25.0
$\widehat{IV}_{\mathrm{NEB},\ell_2}$	−0.2	2.7	5.4	7.9
$\widehat{IV}_{\mathrm{EB},\ell_2}$	−0.7	−0.6	1.3	4.0
$\widehat{IV}_{\mathrm{NS}}$	−1.4	−0.6	0.4	3.5
$\widehat{IV}_{\mathrm{BNS04}}$	−0.1	7.8	16.0	25.0

MAPE for $s = 15$

	# Jumps			
Est.	0	3	10	30
$\widehat{JV}_{\mathrm{NEB},\ell_2}$	0.3	4.3	6.0	7.9
$\widehat{JV}_{\mathrm{EB},\ell_2}$	0.8	4.1	4.0	4.5
$\widehat{JV}_{\mathrm{NS}}$	1.3	3.6	3.8	4.5
$\widehat{JV}_{\mathrm{BNS04}}$	3.3	8.4	16.0	25.0
$\widehat{IV}_{\mathrm{NEB},\ell_2}$	5.8	4.5	5.5	7.9
$\widehat{IV}_{\mathrm{EB},\ell_2}$	6.0	4.1	2.6	4.0
$\widehat{IV}_{\mathrm{NS}}$	5.9	3.9	2.2	3.6
$\widehat{IV}_{\mathrm{BNS04}}$	6.7	8.8	16.0	25.0

Conclusion: Empirical Bayesian Estimators

- Simulations demonstrate superiority over $\widehat{JV}_{\mathrm{BNS04}}$ and $\widehat{JV}_{\mathrm{NS}}$:)
- Model-based :)
- Bounded above zero :)
- Not necessarily bounded below RV :(
- May depend on window of volatility estimation :(

Bibliography I

Andersen, T., T. Bollerslev, F. Diebold, and P. Labys (2003). Modeling and forecasting realized volatility. *Econometrica 71*(2), 579–625.

Barndorff-Nielsen, O. and N. Shephard (2002). Econometric analysis of realized volatility and its use in estimating stochastic volatility models. *Journal of the Royal Statistical Society. Series B (Statistical Methodology) 64*(2), 253–280.

Barndorff-Nielsen, O. E. and N. Shephard (2004). Power and bipower variation with stochastic volatility and jumps. *Journal of Financial Econometrics 2*(1), 1–37.

Barndorff-Nielsen, O. E. and N. Shephard (2006). Econometrics of testing for jumps in financial economics using bipower variation. *Journal of Financial Econometrics 4*(1), 1–30.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics 31*(3), 307–327.

Bibliography II

Harvey, A., E. Ruiz, and N. Shephard (1994). Multivariate stochastic variance models. *The Review of Economic Studies 61*(2), 247–264.

Jacquier, E., N. Polson, and P. Rossi (1994). Bayesian analysis of stochastic volatility models. *Journal of Business & Economic Statistics 12*(4), 371–389.

Johnstone, I. and B. Silverman (2005). Ebayesthresh: R programs for empirical bayes thresholding. *Journal of Statistical Software 12*(8), 1–38.

Johnstone, I. M. (2011). *Gaussian estimation: Sequence and multiresolution models*.

Johnstone, I. M. and B. W. Silverman (2004). Needles and straw in haystacks: Empirical bayes estimates of possibly sparse sequences. *The Annals of Statistics 32*(4), 1594–1649.

Bibliography III

Lee, S. S. and P. A. Mykland (2008). Jumps in financial markets: A new nonparametric test and jump dynamics. *Review of Financial studies 21*(6), 2535–2563.

Melino, A. and S. Turnbull (1990). Pricing foreign currency options with stochastic volatility. *Journal of Econometrics 45*(1-2), 239–265.

Taylor, S. (1994). Modeling stochastic volatility: A review and comparative study. *Mathematical finance 4*(2), 183–204.

Wood, R. (2000). Market microstructure research databases: History and projections. *Journal of Business & Economic Statistics*, 140–145.