Financial Econometrics and Volatility Models
Estimating Realized Variance

Eric Zivot

June 2, 2010
Outline

- Volatility Signature Plots
- Realized Variance and Market Microstructure Noise
- Unbiased Estimation of Realized Variance
- Empirical Applications
- Realized Covariance Estimation
- Bipower Variation and Tests for Jumps
Reading


Volatility Signature Plots

- Observed log price (transaction price or mid-quote) for day \( t = 1, \ldots, n \) is denoted \( \tilde{p}_t \)

- Divide each day into \( M \) subperiods and define \( \delta = 1/M \). This creates a regularly spaced time clock

- Align observed prices to time clock
  - Previous tick method
  - Linear interpolation between adjacent ticks
• Define the $j$th inter-daily return for day $t$

$$
\tilde{r}_{j,t} = \tilde{p}(t-1)+j\delta - \tilde{p}(t-1)+(j-1)\delta, \ j = 1, \ldots, M
$$

• Define realized variance (RV) for day $t$ at frequency $M$

$$RV_t^{(M)} = \sum_{j=1}^{M} \tilde{r}_{j,t}^2$$

Result: If returns are free of microstructure noise and prices follow a continuous time diffusion then

$$RV_t^{(M)} \xrightarrow{p} IV_t = \int_{t-1}^{t} \sigma^2(s)ds$$
Result: If \( M \) is chosen too large (\( \delta \) is too small) then \( RV_t^{(M)} \) becomes biased due to microstructure noise.

- Lack of liquidity could cause observed price to differ from true price (e.g. large trades or short time periods)
- Bid-Ask spread and discrete nature of price price data that implies rounding errors
- Econometric method to construct price data (infer prices from transaction data or mid-quotes; impute prices at times when no prices are observed)
- Data recording errors
Variance Signature Plots can be used to uncover biases due to microstructure noise

- Plot average realized variance, $\overline{RV}^{(M)} \equiv \frac{1}{n} \sum_{t=1}^{n} RV_t^{(M)}$, against the sampling frequency $M$, where the average is taken over $n$ days


5 years of intra-day data for Alcoa and Microsoft ($n = 1250$)

Compute $RV_t^{(M)}$ for $M = 1, \ldots, 3600$ seconds using transactions and mid-quotes and prices aligned using previous tick and linear interpolation methods.
Realized Variance and Market Microstructure Noise


Main Points

- Observed HF log price = log efficient price + microstructure noise
  - Variance of daily return = variance of efficient returns + variance of microstructure noise
• Both unobserved components of variance can be estimated using HF data sampled at different frequencies
  – High frequency sampling captures microstructure noise
  – Low frequency sampling captures efficient return variance

• Provide procedure to purge HF return data of microstructure components and extract information on efficient return variance by sampling at optimal frequencies
Price Formation Mechanism

- Consider \( t = 1, \ldots , n \) trading days.

- Observed log price at time \( t \)
  \[
  \tilde{p}_t = p_t + \eta_t \\
  p_t = \text{unobserved efficient price} \\
  \eta_t = \text{unobserved microstructure noise}
  \]

- Divide each day into \( M \) subperiods and define \( \delta = 1/M \). Define the \( j \)th inter-daily return for day \( t \)
  \[
  \tilde{r}_{j,t} = \tilde{p}(t-1)+j\delta - \tilde{p}(t-1)+(j-1)\delta, \quad j = 1, \ldots , M
  \]
• Return decomposition

\[ \tilde{r}_{j,t} = r_{j,t} + \varepsilon_{j,t} \]

\[ r_{j,t} = p(t-1)+j\delta - p(t-1)+(j-1)\delta : \text{efficient return} \]

\[ \varepsilon_{j,t} = \eta(t-1)+j\delta - \eta(t-1)+(j-1)\delta : \text{microstructure noise} \]
Assumption 1 (Efficient Price Process)

1. $p_t$ is is a continuous stochastic volatility local martingale. Specifically,
   \[ p_t = m_t = \int_0^t \sigma_s dW_s, \quad W_t = \text{Wiener process} \]

2. Spot volatility $\sigma_t$ is cadlag and bounded away from zero

3. $\sigma_t$ is independent of $W_t$ for all $t$

4. Integrated volatility: \( IV_t = \int_0^t \sigma_s^2 ds \)

5. Quarticity \( Q_t = \int_0^t \sigma_s^4 ds < M \leq \infty \)
Assumption 2 (Microstructure Noise)

1. The random shocks $\eta_t$ are iid, $E[\eta_t] = 0$, $E[\eta_t^2] = \sigma_\eta^2$, $E[\eta_t^8] < \infty$

2. True return process $r_{j,t}$ is independent of $\eta_{j,t}$ for all $t$ and $j$

Remarks

- $\sigma_s$ can display jumps, diurnal effects, high persistence (long memory) and nonstationarities
• $\varepsilon_{j,t}$ follows an MA(1) process independent of $\delta$, with negative first order autocovariance

$$E[\varepsilon^2_{j,t}] = 2E[\eta^2_{j,t}] = 2\sigma^2$$
$$E[\varepsilon_{j,t}\varepsilon_{j-1,t}] = -E[\eta^2_{j,t}] = -\sigma^2, \ E[\varepsilon_{j,t}\varepsilon_{j-k,t}] = 0, \ k > 1$$

justified by Roll’s bid-ask bounce model

• Efficient returns $r_{j,t}$ are of order $O_p(\sqrt{\delta})$ over periods of size $\delta$

• Microstructure noise returns $\varepsilon_{j,t}$ are $O_p(1)$ over any time period
  
  – Price discreteness, bid-ask spread

• Longer period intra-day returns are less contaminated by noise than shorter period returns
Identification at High Frequencies: Volatility of the Unobserved Microstructure Noise

Squared Return Decomposition

\[
\sum_{j=1}^{M} \tilde{\sigma}_{j,t}^2 = \sum_{j=1}^{M} \sigma_{j,t}^2 + \sum_{j=1}^{M} \epsilon_{j,t}^2 + 2 \sum_{j=1}^{M} r_{j,t} \epsilon_{j,t} \\
= O_p(\sqrt{\delta}) + O_p(1) + O_p\left(\frac{1}{2}\right)
\]

Result (Bandi and Russell, 2004). As \( M \to \infty \)

\[
\frac{1}{M} \sum_{j=1}^{M} \tilde{\sigma}_{j,t}^2 \xrightarrow{p} E[\epsilon^2], \quad \frac{2}{M} \sum_{j=1}^{M} \tilde{\sigma}_{j,t}^2 \xrightarrow{p} E[\eta^2]
\]
Note: if $\eta$ are iid across days ($t = 1, \ldots, n$)

$$\frac{1}{nM} \sum_{t=1}^{n} \sum_{j=1}^{M} \tilde{r}_{j,t}^2 \xrightarrow{p} E[\varepsilon^2]$$

Use highest possible sample frequency to construct estimates
Identification at Low Frequencies: Volatility of the Unobserved Efficient Return

Result (Bandi and Russell, 2004). As $M \to \infty$

$$\sum_{j=1}^{M} r_{j,t}^2 \xrightarrow{p} \int_{t-1}^{t} \sigma_s^2 ds = IV_t, \quad \sum_{j=1}^{M} \epsilon_{j,t}^2 \xrightarrow{p} \infty$$

$$\sum_{j=1}^{M} r_{j,t} \epsilon_{j,t} = O_p(1)$$

Consequently,

$$\sum_{j=1}^{M} \tilde{r}_{j,t}^2 \to \infty$$

Hence, traditional RV estimator is inconsistent in the presence of microstructure noise!
Optimal Sampling: Balancing Bias-Variance Tradeoff

Intuition: RV estimator is expected to be less biased when sampled at low frequencies, since noise plays less of a roll when $\delta$ is large, but considerably more volatile. The optimal sampling frequency minimizes the MSE.

Result (Bandi and Russell, 2004)

\[
\text{MSE} \left( \sum_{j=1}^{M} \hat{r}_{j,t}^2, IV_t \right) = E_{\sigma} \left[ \left( \sum_{j=1}^{M} \hat{r}_{j,t}^2 - IV_t \right)^2 \right]
\]

\[
= 2 \frac{1}{M} (Q_t + o(1)) + M\beta + M^2\alpha + \gamma
\]

\[
\alpha = \left( E[\varepsilon^2] \right)^2, \quad \beta = 2E[\varepsilon^4] - 3\alpha
\]

\[
\gamma = 4E[\varepsilon^2]IV_t - E[\varepsilon^4] + 2\alpha
\]
Remarks:

- The necessary ingredients to compute the minimum of the MSE are $E[\varepsilon^2]$, $E[\varepsilon^4]$ and $Q_t$

- Consistent estimators for $E[\varepsilon^2]$ and $E[\varepsilon^4]$ (as $M \to \infty$)

\[
\frac{1}{nM} \sum_{t=1}^{n} \sum_{j=1}^{M} \tilde{r}_{j,t}^2 \xrightarrow{p} E[\varepsilon^2], \quad \frac{1}{nM} \sum_{t=1}^{n} \sum_{j=1}^{M} \tilde{r}_{j,t}^4 \xrightarrow{p} E[\varepsilon^4]
\]

- Under microstructure noise, the BNS estimator of $Q_t$ is inconsistent

\[
\hat{Q}_t = \frac{M}{3} \sum_{j=1}^{M} \tilde{r}_{j,t}^4 \to \infty \text{ as } M \to \infty
\]
• Bandi and Russell suggest to estimate $Q_t$ using $\hat{Q}_t$ with a low sampling frequency (e.g. 15 minutes)

$$\hat{Q}_t = \frac{M^{\text{low}}}{3} \sum_{j=1}^{M^{\text{low}}} \tilde{r}_{j,t}^4$$
Result (Approximate optimal sampling frequency). The approximate optimal sampling frequency is chosen as the value \( \delta_t^* = 1/M_t^* \) with

\[
M_t^* = \left( \frac{\hat{Q}_t}{\hat{\alpha}} \right)^{1/3}
\]

\[
\hat{\alpha} = \left( \frac{1}{nM_{\text{high}}} \sum_{t=1}^{n} \sum_{j=1}^{M_{\text{high}}} \tilde{r}_{j,t}^2 \right)^2, \quad M \text{ is highest frequency}
\]

\[
\hat{Q}_t = \frac{M_{\text{low}}^{\text{low}}}{3} \sum_{j=1}^{M_{\text{low}}} \tilde{r}_{j,t}^4, \quad M_{\text{low}} \text{ is low frequency (15 mins)}
\]

Optimal sampling frequency RV estimate

\[
RV_t^{(M_t^*)} = \sum_{j=1}^{M_t^*} \tilde{r}_{j,t}^2
\]
Empirical Applications

- S&P 100 Stocks: 1993 - 2003
- Use mid-quotes as observed prices
- Compute optimal sampling frequencies for 100 stocks
  - Mean value of 4 minutes
  - Vary considerably over time

Main points:

- Characterize how RV is affected by market microstructure noise under a general specification for the noise that allows for various forms of stochastic dependencies

- Market microstructure noise is time-dependent and correlated with efficient returns

- For Dow 30 stocks, noise may be ignored when returns are sampled at low frequencies (e.g. 20 mins)
Ugly Facts about Market Microstructure Noise

- Noise is correlated with efficient price

- Noise is time dependent

- Noise is quite small in Dow Jones 30 stocks

- Properties of noise have changed substantially over time
Notation and Assumptions

• \( p^*(t) = p(t) + u(t) = \) efficient price + noise

• \( dp^*(t) = \sigma(t)dW(t) \)

• Data are observed on interval \([a, b]\) (e.g. trading day)

• \( IV = \int_a^b \sigma^2(t)dt \)
• Partition $[a, b]$ into $m$ subintervals

• $i$th subinterval is $[t_{i-1,m}, t_{i,m}]$

\[
[a = t_{0,m} < t_{1,m} < \cdots < t_{m,m} = b] \\
\delta_{i,m} = t_{i,m} - t_{i-1,m}
\]

• Intra-day returns

\[
y_{i,m} = p(t_{i,m}) - p(t_{i-1,m}), \ i = 1, \ldots, m \\
= y_{im}^* + e_{i,m} \\
y_{i,m}^* = p^*(t_{i,m}) - p^*(t_{i-1,m}) \\
e_{i,m} = u(t_{i,m}) - u(t_{i-1,m})
\]
• IV over \([t_{i-1,m}, t_{i,m}]\)

\[
\sigma^2_{i,m} = \int_{t_{i-1,m}}^{t_{i,m}} \sigma^2(s) ds = \text{var}(y^*_i,m)
\]

• RV of efficient price

\[
RV^*_m = \sum_{i=1}^{m} y^*_{i,m}^2
\]

• RV of observed price

\[
RV^m = \sum_{i=1}^{m} y^2_{i,m}
\]
Sampling Schemes

• Calendar time sampling (CTS)
  
  – Align prices to common regularly spaced time clock associated with
  
    \[a = t_{0,m} < t_{1,m} < \cdots < t_{m,m} = b\]

• Tick time sampling (TTS)
  
  – \(t_{i,m}\) denotes actual transaction time
  
  – e.g. sample every fifth transaction
Characterizing the Bias of RV

Assumption 2: The noise process, \( u \), is covariance stationary with mean 0, such that its autocovariance function is defined by \( \pi(s) = E[u(t)u(t + s)] \)

Remark: Assumption 2 allows for dependence between \( p^* \) and \( u \)

Decomposition of \( RV(m) \) when \( y_{i,m} = y_{im}^* + e_{i,m} \)

\[
RV(m) = \sum_{i=1}^{m} y_{i,m}^2 + 2 \sum_{i=1}^{m} e_{i,m} y_{i,m}^* + \sum_{i=1}^{m} e_{i,m}^2
\]
Theorem 1. Given Assumptions 1 and 2, the bias of $RV^{(m)}$ under CTS is given by

$$E[RV^{(m)} - IV] = 2\rho_m + 2m \left[ \pi(0) - \pi \left( \frac{b-a}{m} \right) \right]$$

$$\rho_m = E \left[ \sum_{i=1}^{m} e_{i,m} y_{i,m}^* \right]$$

Remarks

- Bias always positive when $\text{cov}(y_{i,m}^*, e_{i,m}) = 0$
- Bias can be negative when $\text{cov}(y_{i,m}^*, e_{i,m}) < 0$ and large
Bias Corrected RV

Assumption 4: The noise process has finite dependence in the sense that $\pi(s) = 0$ for all $s > \theta_0$ for some $\theta_0 > 0$, and $E[u(t)|p^*(s)] = 0$ for all $|t - s| > \theta_0$

Theorem 2. Suppose Assumptions 1, 2, and 4 hold and let $q_m$ be such that $q_m/m > \theta_0$. Then under CTS,

$$E[RV_{ACq_m}^{(m)} - IV] = 0$$

where

$$RV_{ACq_m}^{(m)} = \sum_{i=1}^{m} y_{i,m}^2 + 2 \sum_{h=1}^{q_m} \tilde{\gamma}_h$$

$$\tilde{\gamma}_h = \frac{m}{m-h} \sum_{i=1}^{m-h} y_{i,m} y_{i+h,m}$$
Remarks

- $RV_{AC_{qm}}^{(m)}$ may be negative because $\tilde{\gamma}_h$ is not scaled downward in a way that would guarantee positivity (e.g. as with the NW type long-run variance estimator)

- One could use various kernels (e.g. Bartlett) to ensure positivity but the resulting estimators may not be unbiased

- $RV_{AC_{qm}}^{(m)} \overset{p}{\rightarrow} IV$ as $m \rightarrow \infty$ because $q_m/m \rightarrow 0$ sufficiently fast
Realized Kernel Estimators


Idea: Create approximately unbiased RV estimators that have good finite sample properties
\[ RV_{\text{ker}_{qm}}^{(m)} = \sum_{i=1}^{m} y_{i,m}^2 + 2 \sum_{h=1}^{q_{m}} k \left( \frac{h - 1}{q_{m}} \right) \tilde{\gamma}_h \]

\[ \tilde{\gamma}_h = \frac{m}{m - h} \sum_{i=1}^{m-h} y_{i,m} y_{i+h,m} \]

\[ k(\cdot) = \text{kernel weight function} \]

\[ q_{m} = \text{lag truncation parameter} \]
Realized Covariance Estimation


Notation

- Observed vector of log prices for $k$ assets, aligned to a common time clock equally spaced by $\delta = 1/M$, for day $t = 1, \ldots, n$ is denoted $\mathbf{p}_t = (p_{t1}, \ldots, p_{t2})'$.  

- Define the $j$th inter-daily return vector for day $t$

$$\mathbf{r}_{j,t} = \mathbf{p}_{(t-1) + j\delta} - \mathbf{p}_{(t-1) + (j-1)\delta}, \ j = 1, \ldots, M$$

- Define $k \times k$ realized covariance (RCOV) matrix for day $t$ at frequency $M$

$$\text{RCOV}_t^{(M)} = \sum_{j=1}^{M} \mathbf{r}_{j,t} \mathbf{r}_{j,t}'$$
For 2 assets with intra-day returns $r_{j,t}^1$ and $r_{j,t}^2$ define

$$RCOV_t^{(M)} = \sum_{j=1}^{M} \tilde{r}_{j,t}^1 \tilde{r}_{j,t}^2$$
Result: If returns are free of microstructure noise and prices follow a continuous time diffusion then

\[
\mathbf{RCOV}_t^{(M)} \xrightarrow{p} \mathbf{ICOV}_t = \int_{t-1}^{t} \mathbf{\Sigma}(t) dt
\]

Result: If \( M \) is chosen too large (\( \delta \) is too small) then \( \mathbf{RCOV}^{(M)}_t \) becomes biased due to microstructure noise. However, the bias of the off diagonal terms is different from the bias observed in realized variance due to non-synchronous trading.
Covariance Signature Plots

Average Plot: Plot average pairwise realized covariance variance, \( \overline{RCOV}^{(M)} \equiv \frac{1}{n} \sum_{t=1}^{n} RCOV_t^{(M)} \), against the sampling frequency \( M \), where the average is taken over \( n \) days.

Remark: Payseur (2008) argues that averaging over days masks the instability of \( RCOV_t^{(M)} \) for different \( M \), and advocates the use of 1-day pairwise covariance signature plots.
Kernel Estimators for Realized Covariance

\[
\text{RCOV}_{t, \text{ker}_{q_m}}^{(M)} = \sum_{j=1}^{M} r_{j,t} r'_{j,t} + 2 \sum_{h=1}^{q_m} k \left( \frac{h - 1}{q_m} \right) (\tilde{\Gamma}_h + \tilde{\Gamma}'_h)
\]

\[
\tilde{\Gamma}_h = \frac{M}{M - h} \sum_{j=1}^{M-h} r_{j,t} r'_{j+h,t}
\]

\[k(\cdot) = \text{kernel weight function}\]

\[q_m = \text{lag truncation parameter}\]
Realized Bipower Variation and Tests for Jumps


Main point: Show how realized bipower variation can be used to create test statistics to construct nonparametric tests for the presence of jumps in the price process.
Notation and Quadratic Variation

• $p_t = \text{continuous time log price process (semi-martingale)}$

  $p_t = p_t^c + p_t^d$, $t > 0$
  $p_t^c = \text{continuous part}$
  $p_t^d = \text{discontinuous (jump) part}$

• Quadratic variation defined

  $QV_t = p - \lim_{n \to \infty} \sum_{j=0}^{n-1} (p_{t_{j+1}} - p_{t_j})^2$

  $= QV_t^c + QV_t^d$

  $QV_t^d = \sum_{0 \leq u < t} \Delta Y_u^2$, $\Delta Y_t = Y_t - Y_{t-} = \text{jumps}$
• Intra-day returns for $\delta = 1/m$

\[ r_j = p_j \delta - p_{(j-1)} \delta, \ j = 1, \ldots, t/\delta \]

• Realized quadratic variation (realized variance)

\[ RV_t^{(m)} = \sum_{j=1}^{[t/\delta]} y_j^2 \]
Assume $p_t$ belongs to the Brownian semimartingale plus jump (BSMJ) process

$$p_t = \int_0^t a_s ds + \int_0^t \sigma_s dW_s + Z_t$$

$$Z_t = \sum_{j=1}^{N_t} c_j$$

$N_t$ = counting process and $c_j$ are non-zero random variables.

Result:

$$QV_t = QV_t^c + QV_t^d$$

$$QV_t^c = \int_0^t \sigma_s^2 ds$$

$$QV_t^d = \sum_{j=1}^{N_t} c_j^2$$
Bipower Variation

Defn: The 1,1-order BPV process is

\[
BPV_t^{1,1} = p - \lim_{\delta \to 0} \sum_{j=2}^{[t/\delta]} |y_{j-1}| |y_j|
\]

Result: If \( p_t \in \text{BSMJ} \) with \( a = 0 \) and \( \sigma \) independent of \( W \) then

\[
BPV_t^{1,1} = \mu_1^2 \int_0^t \sigma_s^2 ds = \mu_1^2 QV_t^c
\]

\[
\mu_1 = E[z] = \sqrt{2}/\sqrt{\pi}, \quad z \sim N(0, 1)
\]

Hence,

\[
\mu_1^{-2} BPV_t^{1,1} = QV_t^c
\]
Result:

\[ QV_t - \mu_1^{-2}BPV_{t}^{1,1} = QV_t^d = \sum_{j=1}^{N_t} c_j^2 \]

Jump component can be estimated using realized variance and realized bipower variation

\[ QV_t^d = RV_t^{(m)} - \mu_1^{-2}RBPV_{t}^{1,1(m)} \]

\[ RBPV_{t}^{1,1(m)} = \sum_{j=2}^{[t/\delta]} |y_{j-1}||y_j| \]
Testing for Jumps

Theorem 1. Let $p_t$ belong to the Brownian semimartingale (BSM) process without jumps, and suppose that $\sigma_s$ is independent of $W$. Then, as $\delta \to 0$

$$G = \frac{\delta^{-1/2} \left( \mu_1^{-2} RBPV_t^{1,1(m)} - RV_t^{(m)} \right)}{\left( \theta \int_0^t \sigma_u^4 du \right)^{1/2}} \to N(0, 1)$$

$$H = \frac{\delta^{-1/2} \left( \frac{\mu_1^{-2} RBPV_t^{1,1(m)}}{RV_t^{(m)}} - 1 \right)}{\left( \theta \frac{\int_0^t \sigma_u^4 du}{\left\{ \int_0^t \sigma_u^2 du \right\}^2} \right)^{1/2}} \to N(0, 1)$$

where $\theta = (\pi^2/4) + \pi - 5$. 


Feasible Tests

Problem: Need a consistent estimator of $\int_0^t \sigma_u^4 du$ under null and alternative

Solution: Use realized quadpower variation

$$RBPV_t^{1,1,1,1}(m) = \sum_{j=4}^{[t/\delta]} |y_{j-3}||y_{j-2}||y_{j-1}||y_j|$$

Result:

$$RBPV_t^{1,1,1,1}(m) \rightarrow \mu_1^4 \int_0^t \sigma_u^4 du$$
\[
\hat{G} = \frac{\delta^{-1/2} \left( \mu_1^{-2} RBPV_{t,1,1}^{1,1}(m) - RV_t^{(m)} \right)}{\left( \theta \mu_1^{-4} RBPV_{t,1,1,1}^{1,1,1,1}(m) \right)^{1/2}} \rightarrow N(0, 1)
\]

\[
\hat{H} = \frac{\delta^{-1/2} \left( \mu_1^{-2} RBPV_{t,1}^{1,1}(m) \right)}{\left( \theta \frac{RBPV_{t,1,1,1}^{1,1,1,1}(m)}{\left( RBPV_{t,1,1}(m) \right)^2} \right)^{1/2}} \rightarrow N(0, 1)
\]