Finite sample results of Range-based integrated volatility estimation

Eduardo Rossi[1], Filippo Spazzini[2]

[1]University of Pavia

[2]Edison Trading S.p.A.

FEEM Seminars

Milan

November 3, 2011

The availability of high-frequency financial prices allows the use of realized variance, or realized volatility, *RV* to estimate the integrated variance *IV*.

The Realized Range, *RRG*, is a rivaling approach to the estimation of *IV*, based on the aggregation of intra-daily ranges. Martens and Van Dijk (2007), and Christensen and Podolskij (2007) and Christensen, Podolskij and Vetter (2009).

This paper:

- New version of RRG estimator for irregular sampling (with and without MN).
- Finite-sample properties of *RRG* estimators of *IV* are considered and compared to *RV* estimators when high-frequency data are observed.

The availability of high-frequency financial prices allows the use of realized variance, or realized volatility, *RV* to estimate the integrated variance *IV*.

The Realized Range, *RRG*, is a rivaling approach to the estimation of *IV*, based on the aggregation of intra-daily ranges. Martens and Van Dijk (2007), and Christensen and Podolskij (2007) and Christensen, Podolskij and Vetter (2009).

This paper:

- New version of RRG estimator for irregular sampling (with and without MN).
- Finite-sample properties of *RRG* estimators of *IV* are considered and compared to *RV* estimators when high-frequency data are observed.

Outline

Introduction

The market microstructure (MN) dynamics generate a divergence between the observed price and the true or *noiseless* price process, whose quadratic variation is the object of interest.

- With microstructure noise the *RV* is both biased and inconsistent. Zhou (1996), Hansen and Lunde (2006), Bandi and Russell (2006) Bandi and Russell (2008).
- Robust Estimators (e.g. TSRV estimator Zhang Ait-Sahalia Mykalnd (2005)).
- Simple solution is to sample at moderate frequencies, e.g., every $5-$, $10-$, or $30-$ minutes.

The *RRG* is appealing when MN prevents from the use of the whole record of high-frequency prices to compute *RV* since it exploits a larger amount of information and it is in principle able to attain a higher precision.

Realized variance

Consider an equidistant partition $0 = t_0 < t_1 < \ldots < t_n = 1$, where $t_i = i/n$, and $\Delta = 1/n$. The *RV* at sampling frequency n is

$$RV^{\Delta} = \sum_{i=1}^{n} r^2_{i\Delta,\Delta}.$$

From the theory of stochastic integration, when $n \to \infty$ the quadratic variation:

$$\langle p \rangle = p \lim_{n \to \infty} \sum_{i=1}^{n} r^2_{i\Delta,\Delta} = \int_0^1 \sigma(u)^2 \mathrm{d}u = IV.$$

Then $RV^{\Delta} \xrightarrow{p} IV$. Barndorff-Nielsen and Shephard (2002) derived a distribution theory for RV^{Δ}.

RV and microstructure noise

Suppose the price is contaminated by MN, $\eta_t \sim i.i.d.(0, \omega^2)$:

$$\widetilde{p}_t = p_t + \eta_t$$

with η independent of p_t.

This noise introduces spurious volatility and a negative serial correlation in $\mathrm{d}\widetilde{p}_t$.

Independent market MN leads to a bias that diverges to infinity.

Hansen and Lunde (2006) consider the following estimator:

$$RV_{HL}^{\Delta} = \sum_{i=1}^{n} r_{i\Delta,\Delta}^2 + 2\frac{M}{M-1}\sum_{i=2}^{M} r_{i\Delta,\Delta} r_{(i-1)\Delta,\Delta}.$$

which is robust to the first order serial correlation in returns.

RV with jumps in volatility

Andersen and Bollerslev (1998) and Barndorff-Nielsen and Shephard (2001), under very general processes for the instantaneous volatility:

- Jumps
- Long memory or non-stationarity

The quadratic variation of the price process coincides with the IV.

Nielsen and Frederiksen (2008):

- Bias and RMSE of RV increase with the jump size.
- Increasing the sensitivity of the arrival intensity of jumps towards the level of the volatility implies an increase in the relative biases of the realized volatility estimators.
- Pronounced impact on the variability of the estimators.

Parkinson estimator

The daily range estimator of Parkinson (1980) developed under the assumption that the log-price follows a scaled Brownian motion:

$$\mathrm{d}p(t) = \sigma \mathrm{d}W(t)$$

the daily range is defined as

$$s_p = \sup_{0 \leq t,s \leq 1} \{p_t - p_s\}.$$

On the $[0, 1]$ interval:

$$E[s_p^r] = \lambda_r \sigma^r.$$

Parkinson estimator

Under the assumption of a fully observed continuous time log-price path, Parkinson's estimator of the daily integrated volatility:

$$RG_P = \frac{1}{\lambda_2} s_p^2 = \frac{s_p^2}{4\log(2)}.$$

where

$$\lambda_r = E[s_W^r] = E[\sup\{W_1 - W_0\}^r]$$

Continuous-time price process

Objective: Estimation of a suitable measure of the return variation over the trading day which is normalized to be $[0, 1]$.

Stochastic volatility model

Log-price of an asset, $p(t)$, follows a stochastic volatility model (SV):

$$\mathrm{d}p(t) = \mu(t)\mathrm{d}t + \sigma(t)\mathrm{d}W(t), \quad t \geq 0$$

$\mu(t)$ is locally bounded and predictable, $\sigma(t)$ is independent of sBm $W(\cdot)$ and càdlàg.

Common set up in the literature on RV. Barndorff-Nielsen and Shephard (2008), Andersen, Bollerslev and Diebold (2002).

Realized range

The intraday range at sampling times t_{i-1} and t_i ($i = 1, 2, \ldots, n$) is

$$s_{p_{i\Delta,\Delta}} = \sup_{t_{i-1} \leq t, s \leq t_i} \{p_t - p_s\}.$$

The *RRG* estimator for the interval $[0, 1]$ is defined as:

$$RRG^{\Delta} = \frac{1}{\lambda_2} \sum_{i=1}^{n} s^2_{p_{i\Delta,\Delta}}$$

Christensen and Podolskij (2007) show that

$$RRG^{\Delta} \xrightarrow{p} IV.$$

Result obtained for very general continuous time processes, including models with leverage, long-memory, diurnal effects or jumps (in $\sigma(t)$).

RRG Asymptotic distribution

RRG^{Δ} converges in law to a mixed normal with σ governing the mixture, i.e.:

$$\sqrt{n}(RRG^{\Delta} - IV) \xrightarrow{d} MN(0, \Lambda IQ)$$

with $IQ = \int_0^1 \sigma(u)^4 \mathrm{d}u$.

For RRG Λ is approximately 0.4 while for RV is 2.
The sampling error of RRG^{Δ} are about one-fifth of those based on realized volatility.

RRG^{Δ} uses all the data, whereas RV is based on high-frequency returns sampled at fixed points in time. It follows that

$$\frac{\sqrt{n}(RRG^{\Delta} - IV)}{\sqrt{\Lambda RRQ^{\Delta}}} \xrightarrow{d} N(0, 1)$$

where $RRQ^{\Delta} = \frac{n}{\lambda_4} \sum_{i=1}^{n} s^4_{p_{i\Delta},\Delta}$ consistently estimates the IQ.

RRG with discretely sampled data

When the inference is based on a finite sample, the intraday high-low statistic will be progressively more downward biased as n gets larger, since the number of prices in each Δ decreases.

The true range is not observed!

The source of bias is λ_2, which is constructed on the presumption that p is fully observed.

- Christensen and Podolskij (2007) assume that $mn + 1$ equally spaced price observations are available.
- n intervals each with m returns.
- The log-price for each time in the interval $(0, 1)$ is $p_{\frac{i-1}{n}+\frac{t}{mn}}$, $i = 1 \ldots, n$ and $t = 0, \ldots, m$.
- The observed range over the i-th interval is:

$$s_{p_{i\Delta,\Delta},m} = \max_{0 \leq s,t \leq m} \left\{ p_{\frac{i-1}{n}+\frac{t}{mn}} - p_{\frac{i-1}{n}+\frac{s}{mn}} \right\}.$$

RRG with discretely sampled data

$$s_{W,m} = \max_{0 \leq s,t \leq m} \left\{ W_{t/m} - W_{s/m} \right\}$$

and

$$\lambda_{r,m} = E[s^r_{W,m}].$$

$\lambda_{r,m}$ is the r-th moment of the range of a sBm over $[0, 1]$ when only m increments of the underlying continuous time process are observed.

Numerical simulation to compute $\lambda_{r,m}$

The *RRG* estimator based on discrete observations is

$$RRG^{\Delta}_m = \frac{1}{\lambda_{2,m}} \sum_{i=1}^{n} s^2_{p_{i\Delta,\Delta},m}.$$

RRG^{Δ}_m is a consistent estimator of IV as $n \to \infty$.

RRG with discretely sampled data

RRG^{Δ}_m is a consistent estimator of IV as $n \to \infty$. If we assume that the log-price follows the SV process and $m \to c \in \mathbb{N} \cup \infty$:

$$\frac{\sqrt{n}(RRG^{\Delta}_m - IV)}{\sqrt{\Lambda_m RRQ^{\Delta}_m}} \stackrel{d}{\to} N(0,1)$$

with $\Lambda_m = \frac{\lambda_{4,m} - \lambda^2_{2,m}}{\lambda^2_{2,m}}$, and

$$RRQ^{\Delta}_m = \frac{n}{\lambda_{4,m}} \sum_{i=1}^{n} s^4_{p_{i\Delta,\Delta},m}.$$

RRG with microstructure noise

Christensen, Podolskij and Vetter (2009): With an i.i.d. MN, the *RRG* estimator of *IV* is

$$RRG^{\Delta}_{m,BC} = \frac{1}{\widetilde{\lambda}_{2,m}} \sum_{i=1}^{n} (s_{\widetilde{p}_{i\Delta,\Delta},m} - 2\widehat{\omega}_N)^2$$

where

$$\widetilde{\lambda}_{r,m} = E\left[\left|\max_{t:\eta\frac{t}{m}=\omega,\, s:\eta\frac{s}{m}=-\omega}\left(W_{\frac{t}{m}} - W_{\frac{s}{m}}\right)\right|^r\right].$$

$Var[\omega^2]$ can be consistently estimated with

$$\widehat{\omega}^2_N = \frac{RV^N}{2N},$$

$N = nm$ is the total number of log-returns.

RRG with irregularly sampled data

Christensen and Podolskij (2007): equally spaced observations.

We analyze the effect that MN and irregularly spaced observations have on the bias and the variance of realized-range estimators.

Two different time scales

- transaction time where prices are sampled with every transaction
- tick time where prices are sampled with every price change.

RRG with irregularly sampled data

With irregular sampling, the range:

$$s_{p_{i\Delta,\Delta},m_i} = \max_{0\leq s,t\leq m_i} \left\{ p_{\frac{i-1}{n}+\frac{t}{m_i n}} - p_{\frac{i-1}{n}+\frac{s}{m_i n}} \right\}$$

$$RRG^{\Delta}_{m_i} = \sum_{i=1}^{n} \frac{s^2_{p_{i\Delta,\Delta},m_i}}{\lambda_{2,m_i}}$$

Each λ_{2,m_i} is computed by simulating the sBm with m_i steps, assuming that the discrete realizations of the price are fully observed while they are not.

Irregular sampling implies that some of the prices are missed.

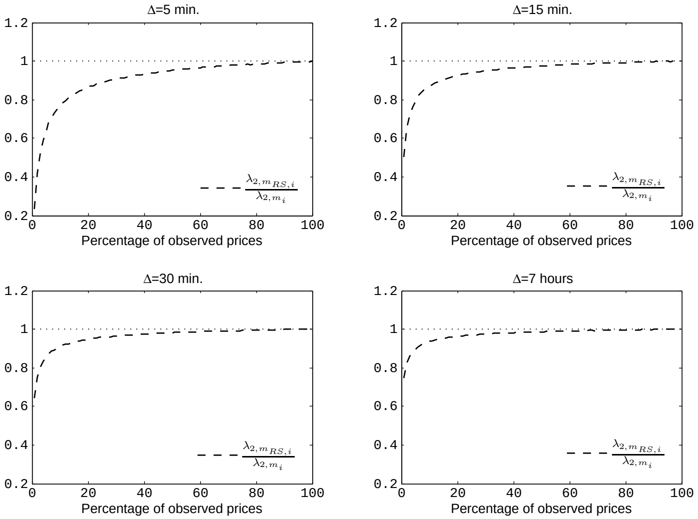
This produces a necessarily downward bias because the observed range is always smaller or equal to the actual underlying range.

RRG with irregularly sampled data

Reducing this bias by modifying the normalizing factor λ_{2,m_i}.
To compute a new $\lambda_{2,m_{RS,i}}$

1. Assume a frequency of 1 second for the underlying discrete process. M be the maximum number of price increments in an interval of length Δ. When Δ increases, M gets larger.
2. Fix the fraction of prices actually observed, namely m_i/M.
3. Simulate 1 million paths of discrete sBm of length $M+1$ and steps $1/M$.
4. Randomly select m_i prices out of $M+1$.
5. Average these squared ranges over the Monte Carlo replications.

RRG with irregularly sampled data



RRG with irregularly sampled data

We modify the RRG estimator by using $\lambda_{2,m_{RS,i}}$ for each intraday range:

$$RRG^{\Delta}_{m_{RS}} = \sum_{i=1}^{n} \frac{s^2_{p_{i\Delta,\Delta},m_i}}{\lambda_{2,m_{RS,i}}}.$$

The numerical stability of the simulation of $\lambda_{r,m_{RS}}$ is quite poor when we observe a percentage of data points too low (m_{RS}).

RRG with irregularly sampled data and MN

With irregularly sampled observations and MN, the RRG estimator can be modified as:

$$RRG^{\Delta}_{m_{RS},BC} = \sum_{i=1}^{n} \frac{(s_{\widetilde{p}_{i\Delta,\Delta},m_i} - 2\widehat{\omega}_N)^2}{\widetilde{\lambda}_{2,m_{RS,i}}}.$$

To compute $\widetilde{\lambda}_{2,m_{RS,i}}$ we coarse with the bid-ask effect the simulated realizations of sBm.

Simulation Processes

For the Monte Carlo simulations we generate the log-price $p(t)$ as:

$$\mathrm{d}p(t) = \sigma(t)\mathrm{d}W_1(t)$$

the instantaneous volatility process $\sigma(t)$ follows, alternatively, two distinct processes:

Long memory in volatility (Model A)

$$\mathrm{d}\log\sigma^2(t) = \alpha(\beta - \log\sigma^2(t))\mathrm{d}t + \nu\mathrm{d}W_d(t)$$

Jumps in volatility (Model B)

$$\mathrm{d}\sigma^2(t) = \alpha(\beta - \sigma^2(t))\mathrm{d}t + \nu\mathrm{d}W_2(t) + \kappa(t)\mathrm{d}q(t)$$

where $W_1(t)$ and $W_2(t)$ are independent Wiener processes.

Simulation Processes

- Generation of $\sigma^2(t)$ with a population long memory parameter of d, resulting in a *IV* with long memory parameter of $1 + d$.
- The arrival of jumps is assumed to follow a Poisson process with intensity given by $\lambda_1 + \lambda_2\sigma^2(t)$.
- The size of jumps is exponentially distributed with mean μ. Eraker, Johannes and Polson (2003) and Eraker (2004).

Discretization

- Euler discrete approximation of Model A and B.
- Discrete trajectories with a time step of 1 second for 7 hours per day.
- Total $60 \times 60 \times 7 = 25.200$ log-prices per day for 10.000 days.
- The generated series are used to compute the IV for each day and to calculate the relative bias and variance for each of the estimators considered.

Irregularly spaced sampling

Randomly sampled series:

- Draw from the generated log-price series at random times.
- The interval between two successive available log-prices is modeled as an random variable exponentially distributed of mean τ, which is set to 5 seconds.

- The day is split into n intervals of constant length $\Delta = 1/n$
- A random number of prices in each interval, m_i, distributed according to a Poisson law with intensity parameter Δ/τ.
- Each observed path is constituted by a random number of $p(t)$: $\sum_{i=1}^{n} m_i$, on average $25200/\tau$.
- Each interval i contains a random number of observed log-prices, which is a decreasing function of τ.

Microstructure noise - Bid-ask bounce

Bid-ask bounce

$$\widetilde{p}(t) = p(t) + \frac{\xi}{2} \mathbb{1}(t)$$

- ξ is the percentage spread.
- the order-driven indicator variables $\mathbb{1}(t)$ are independently across p and t and identically distributed with $Pr\{\mathbb{1}(t) = 1\} = Pr\{\mathbb{1}(t) = -1\} = \frac{1}{2}$.
- This variable takes value 1 when the transaction is buyer-initiated, and -1 when it is seller-initiated.

Evaluation criteria

Relative error statistic:

$$\nu_s = \frac{\widehat{IV}_t - \sum_{t=1}^{T}(p_s(t) - p_s(t-1))^2}{\sum_{t=1}^{T}(p_s(t) - p_s(t-1))^2} \quad s = 1, \ldots, S,$$

$\sum_{t=1}^{T}(p_s(t) - p_s(t-1))^2$ is a discrete-time approximation of the quadratic variation of $p(t)$.

The **Daily-average percentage relative bias** is:

$$\text{bias} = \overline{\nu}_S = \frac{100}{S}\sum_{s=1}^{S}\nu_s.$$

The **Root Mean Square Error**

$$RMSE = \left(\frac{1}{S}\sum_{s=1}^{S}\nu_s^2\right)^{1/2} = \sqrt{\overline{\nu}_S^2 + g_\nu^2}$$

where $g_\nu^2 = S^{-1}\sum_{s=1}^{S}(\nu_s - \overline{\nu}_S)^2$ is the sample variance of ν_s.

Model A with no MN - Relative Percentage Bias

α	d	Realized Volatility			Daily Range			Realized Range			Realized Range RS		
		30 min	15 min	5 min	Park.	Real.	RS	30 min	15 min	5 min	30 min	15 min	5 min
0.000	0.00	-0.116	0.195	**0.021**	-2.360	-0.929	0.073	-1.626	-2.613	-5.873	**-0.021**	-0.114	-0.960
	0.15	0.254	0.504	0.299	-1.390	**0.056**	-0.396	-1.407	-2.487	-5.770	0.032	-0.096	-1.028
	0.30	0.469	0.492	0.155	-1.611	-0.169	0.440	-1.383	-2.448	-5.842	0.147	**-0.039**	-0.921
	0.45	0.252	**0.081**	0.229	-2.792	-1.367	-1.192	-1.524	-2.557	-5.770	-0.149	-0.245	-1.031
0.006	0.00	**-0.067**	0.080	0.140	-2.584	-1.156	0.802	-1.682	-2.614	-5.856	-0.048	-0.224	-1.041
	0.15	0.517	0.394	0.191	-2.616	-1.188	1.250	-1.381	-2.477	-5.799	0.204	**0.068**	-0.929
	0.30	-0.552	**-0.001**	-0.062	-2.247	-0.814	0.428	-1.831	-2.598	-5.916	-0.086	-0.242	-1.009
	0.45	-0.157	0.083	**0.005**	-2.606	-1.178	0.417	-1.719	-2.620	-5.943	0.082	-0.022	-0.940
0.012	0.00	0.350	0.297	**0.029**	-1.977	-0.540	0.193	-1.567	-2.636	-5.879	-0.244	-0.232	-1.027
	0.15	0.123	0.178	0.329	-2.146	-0.712	-0.666	-1.627	-2.592	-5.777	-0.182	**-0.091**	-1.028
	0.30	0.582	0.245	0.175	-1.816	-0.377	0.793	-1.361	-2.501	-5.804	**-0.008**	-0.247	-0.997
	0.45	-0.377	-0.395	-0.114	-3.713	-2.302	**0.006**	-2.088	-2.936	-5.969	-0.183	-0.169	-1.044

Model A with no MN - RMSE

α	d	Realized Volatility			Daily Range			Realized Range			Realized Range RS		
		30 min	15 min	5 min	Park.	Real.	RS	30 min	15 min	5 min	30 min	15 min	5 min
0.000	0.00	0.378	0.268	0.153	0.633	0.642	0.653	0.177	0.128	0.094	0.178	0.129	0.079
	0.15	0.376	0.270	0.155	0.641	0.651	0.632	0.176	0.129	0.094	0.178	0.129	0.078
	0.30	0.380	0.269	0.155	0.636	0.645	0.645	0.176	0.128	0.094	0.178	0.128	0.078
	0.45	0.382	0.269	0.154	0.638	0.646	0.623	0.178	0.129	0.093	0.177	0.128	0.078
0.006	0.00	0.375	0.267	0.154	0.635	0.643	0.660	0.177	0.129	0.094	0.179	0.128	0.078
	0.15	0.382	0.268	0.155	0.624	0.633	0.661	0.178	0.128	0.094	0.178	0.129	0.077
	0.30	0.378	0.268	0.154	0.644	0.653	0.648	0.176	0.128	0.094	0.179	0.128	0.078
	0.45	0.378	0.267	0.156	0.621	0.630	0.653	0.177	0.128	0.095	0.179	0.129	0.078
0.012	0.00	0.380	0.269	0.154	0.643	0.652	0.644	0.179	0.128	0.094	0.176	0.127	0.077
	0.15	0.376	0.269	0.154	0.647	0.657	0.634	0.175	0.129	0.093	0.179	0.130	0.079
	0.30	0.386	0.267	0.156	0.643	0.652	0.654	0.178	0.129	0.094	0.178	0.129	0.078
	0.45	0.378	0.264	0.153	0.614	0.623	0.657	0.178	0.128	0.095	0.179	0.130	0.078

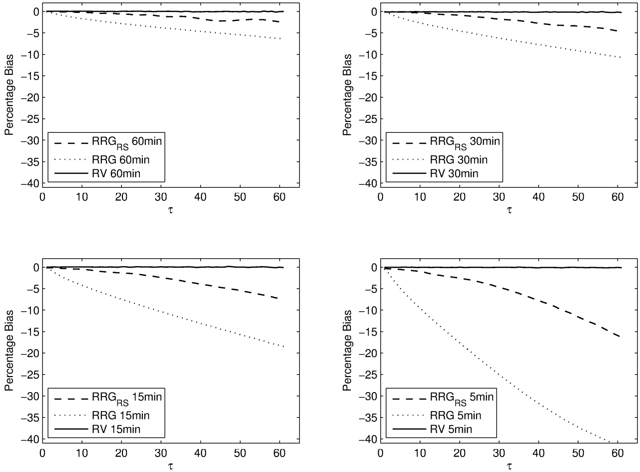
Model B with no MN - Bias

μ	λ_0	λ_1	Realized Volatility			Daily Range			Realized Range			Realized Range RS		
			30 min	15 min	5 min	Park.	Real.	RS	30 min	15 min	5 min	30 min	15 min	5 min
0.751	0.002	0.000	**0.002**	-0.184	0.080	-1.239	0.209	-0.531	-1.803	-2.769	-5.879	-0.127	-0.397	-1.084
		1.298	0.395	0.103	0.229	-2.377	-0.946	0.338	-1.560	-2.643	-5.824	**-0.099**	-0.276	-1.053
		2.596	-0.303	-0.255	0.206	-3.196	-1.777	-1.057	-1.629	-2.646	-5.915	**-0.131**	-0.155	-0.940
	0.010	0.000	-0.260	0.037	-0.009	-2.035	-0.599	-0.788	-1.758	-2.728	-5.939	**0.008**	-0.093	-0.879
		1.298	-0.117	-0.010	**0.006**	-1.153	0.296	-0.203	-1.868	-2.696	-5.882	0.151	-0.087	-0.838
		2.596	0.239	0.113	**0.000**	-2.380	-0.949	-0.092	-1.789	-2.720	-5.936	0.031	-0.102	-1.062
1.530	0.002	0.000	0.372	-0.126	**-0.043**	-1.732	-0.291	0.397	-1.594	-2.763	-5.920	0.090	-0.145	-1.081
		1.298	-0.251	0.070	0.108	-1.849	-0.410	-0.720	-1.657	-2.636	-5.837	**-0.062**	-0.257	-0.991
		2.596	-0.159	**-0.073**	0.322	-1.122	0.327	0.614	-1.696	-2.647	-5.807	0.433	0.128	-0.814
	0.010	0.000	0.077	0.316	0.152	-2.193	-0.759	0.834	-1.541	-2.592	-5.864	0.358	**-0.054**	-0.920
		1.298	-0.235	**-0.125**	0.195	-1.797	-0.358	0.382	-1.652	-2.646	-5.826	-0.127	-0.307	-1.147
		2.596	0.013	-0.171	**0.003**	-1.780	-0.340	-0.569	-1.872	-2.836	-5.967	0.097	-0.021	-0.909

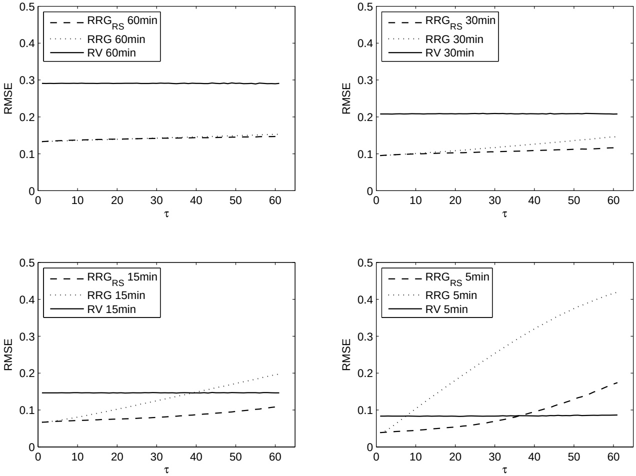
Model B with no MN - RMSE

μ	λ_0	λ_1	Realized Volatility			Daily Range			Realized Range			Realized Range RS		
			30 min	15 min	5 min	Park.	Real.	RS	30 min	15 min	5 min	30 min	15 min	5 min
0.751	0.000	0.000	0.384	0.268	0.153	0.649	0.658	0.650	0.177	0.129	0.094	0.179	0.129	0.079
		1.298	0.376	0.263	0.155	0.620	0.628	0.644	0.176	0.127	0.094	0.176	0.126	0.078
		2.596	0.375	0.267	0.156	0.620	0.628	0.633	0.176	0.129	0.095	0.179	0.129	0.077
	0.010	0.000	0.371	0.268	0.154	0.638	0.647	0.646	0.176	0.129	0.094	0.178	0.128	0.078
		1.298	0.381	0.268	0.155	0.653	0.662	0.652	0.179	0.129	0.095	0.179	0.128	0.078
		2.596	0.381	0.270	0.157	0.631	0.640	0.641	0.178	0.130	0.096	0.177	0.127	0.078
1.530	0.000	0.000	0.376	0.263	0.153	0.641	0.650	0.653	0.175	0.127	0.094	0.181	0.130	0.079
		1.298	0.373	0.264	0.153	0.634	0.643	0.650	0.174	0.127	0.094	0.179	0.128	0.078
		2.596	0.379	0.270	0.156	0.643	0.652	0.655	0.178	0.129	0.094	0.180	0.129	0.078
	0.010	0.000	0.379	0.271	0.156	0.647	0.656	0.657	0.178	0.130	0.095	0.179	0.127	0.078
		1.298	0.377	0.268	0.156	0.640	0.650	0.658	0.177	0.128	0.095	0.179	0.128	0.078
		2.596	0.380	0.268	0.155	0.631	0.640	0.639	0.178	0.129	0.095	0.179	0.129	0.078

Bias of *IV* estimators as a function of τ



RMSE of IV estimators as a function of τ



Model A with MN - Bias

d	ξ	ν	RV^{Δ}			HL			Daily Range				
			30min	15min	5min	30min	15min	5min	RG^D	RRG^D	RRG^D_{RS}	RRG^D_{BC}	$RRG^D_{RS,BC}$
0.0	0.001	0.05	51.40	102.31	308.74	-0.59	0.19	0.48	29.84	31.74	32.77	-5.60	-4.47
		0.10	51.18	102.32	308.41	-0.05	-0.35	0.01	30.06	31.97	33.00	-5.50	-4.37
	0.003	0.05	459.96	923.23	2774.08	-1.07	0.97	2.48	110.80	113.89	115.56	-6.47	-5.35
		0.10	461.29	922.07	2775.90	-4.64	-0.04	0.42	110.72	113.81	115.48	-6.49	-5.37
	0.005	0.05	1286.42	2576.52	7703.77	1.56	-4.30	18.35	214.06	218.67	221.16	-6.15	-5.03
		0.10	1294.52	2578.99	7723.93	-15.01	-0.37	-15.54	216.52	221.16	223.67	-4.83	-3.69
0.3	0.001	0.05	50.73	102.38	308.44	0.12	0.10	-0.58	30.46	32.37	33.40	-5.16	-4.02
		0.10	50.44	102.04	308.18	-0.98	-1.48	-0.22	29.47	31.37	32.40	-5.94	-4.81
	0.003	0.05	459.95	927.33	2778.02	-0.18	-5.44	-0.60	113.06	116.18	117.87	-4.83	-3.69
		0.10	462.74	929.73	2774.91	1.55	-5.29	-2.24	112.45	115.56	117.25	-5.27	-4.13
	0.005	0.05	1286.65	2572.55	7724.44	-1.26	-3.76	6.83	217.14	221.79	224.30	-4.53	-3.39
		0.10	1278.82	2579.09	7713.39	0.27	-7.48	3.36	215.26	219.88	222.38	-5.43	-4.30

d	ξ	ν	RRG^{Δ}_{m}			$RRG^{\Delta}_{m_{RS}}$			$RRG^{\Delta}_{m,BC}$			$RRG^{\Delta}_{m_{RS},BC}$		
			30min	15min	5min	30min	15min	5min	30min	15min	5min	30min	15min	5min
0.0	0.001	0.05	150.76	238.13	528.63	154.77	246.39	564.08	-4.40	-4.47	-5.54	-3.55	-3.08	-2.82
		0.10	150.79	238.20	529.29	154.80	246.45	564.78	-4.63	-4.79	-5.98	-3.78	-3.40	-3.28
	0.003	0.05	703.53	1226.18	3219.80	716.38	1258.56	3407.16	-3.51	-2.67	-0.58	-2.65	-1.25	2.27
		0.10	703.42	1226.30	3221.93	716.26	1258.67	3409.41	-3.59	-2.76	-0.46	-2.74	-1.34	2.40
	0.005	0.05	1575.84	2872.55	8047.04	1602.64	2945.13	8506.97	-1.49	0.79	9.70	-0.61	2.27	12.85
		0.10	1576.98	2874.41	8049.21	1603.80	2947.04	8509.24	-1.67	0.47	8.47	-0.79	1.94	11.59
0.3	0.001	0.05	150.73	237.72	528.55	154.74	245.96	564.01	-4.58	-4.94	-6.00	-3.73	-3.55	-3.30
		0.10	150.41	237.85	528.59	154.41	246.10	564.04	-4.82	-4.88	-6.06	-3.97	-3.49	-3.37
	0.003	0.05	703.81	1226.26	3219.99	716.67	1258.65	3407.34	-3.11	-2.27	0.01	-2.25	-0.84	2.88
		0.10	704.53	1226.73	3220.72	717.39	1259.13	3408.16	-2.95	-2.30	-0.15	-2.09	-0.87	2.72
	0.005	0.05	1577.10	2874.04	8046.66	1603.91	2946.66	8506.56	-1.45	0.66	8.85	-0.57	2.14	11.98
		0.10	1577.55	2874.84	8048.96	1604.37	2947.47	8508.96	-1.00	1.39	10.11	-0.12	2.87	13.28

Model A with MN - RMSE

d	ξ	ν	RV^{Δ}			HL			Daily Range				
			30min	15min	5min	30min	15min	5min	RG^D	RRG^D	RRG^D_{RS}	RRG^D_{BC}	$RRG^D_{RS,BC}$
0.0	0.001	0.05	0.75	1.14	3.13	0.82	0.67	0.68	0.78	0.80	0.81	0.62	0.63
		0.10	0.76	1.14	3.14	0.81	0.68	0.67	0.78	0.80	0.81	0.62	0.62
	0.003	0.05	4.92	9.47	27.94	2.29	2.83	4.46	1.42	1.45	1.47	0.60	0.61
		0.10	4.94	9.47	28.00	2.29	2.86	4.57	1.42	1.46	1.47	0.61	0.62
	0.005	0.05	13.45	26.30	77.56	5.53	7.26	12.19	2.39	2.44	2.46	0.60	0.61
		0.10	13.58	26.36	77.85	5.60	7.31	12.33	2.43	2.48	2.50	0.63	0.63
0.3	0.001	0.05	0.75	1.14	3.13	0.81	0.68	0.67	0.77	0.79	0.80	0.61	0.62
		0.10	0.75	1.14	3.13	0.81	0.67	0.68	0.79	0.80	0.81	0.63	0.63
	0.003	0.05	4.92	9.51	27.97	2.32	2.84	4.45	1.45	1.48	1.50	0.62	0.63
		0.10	4.95	9.55	27.97	2.33	2.86	4.51	1.44	1.48	1.49	0.62	0.62
	0.001	0.05	13.47	26.26	77.75	5.61	7.33	12.10	2.43	2.47	2.50	0.62	0.62
		0.10	13.42	26.36	77.73	5.55	7.41	12.29	2.41	2.46	2.49	0.62	0.62

d	ξ	ν	RRG^{Δ}_{m}			$RRG^{\Delta}_{m_{RS}}$			$RRG^{\Delta}_{m,BC}$			$RRG^{\Delta}_{m_{RS},BC}$		
			30min	15min	5min	30min	15min	5min	30min	15min	5min	30min	15min	5min
0.0	0.001	0.05	1.53	2.39	5.29	1.57	2.48	5.65	0.19	0.15	0.16	0.19	0.15	0.16
		0.50	1.53	2.39	5.30	1.57	2.48	5.66	0.18	0.15	0.16	0.18	0.15	0.15
	0.003	0.05	7.05	12.27	32.21	7.18	12.60	34.09	0.23	0.24	0.39	0.23	0.24	0.40
		0.50	7.06	12.28	32.26	7.19	12.61	34.14	0.23	0.24	0.39	0.23	0.24	0.40
	0.005	0.05	15.78	28.74	80.50	16.05	29.47	85.10	0.29	0.36	0.66	0.29	0.37	0.68
		0.50	15.80	28.79	80.61	16.07	29.51	85.22	0.29	0.36	0.65	0.29	0.37	0.67
0.3	0.001	0.05	1.53	2.39	5.29	1.57	2.47	5.64	0.19	0.15	0.16	0.18	0.15	0.16
		0.50	1.53	2.39	5.29	1.57	2.47	5.65	0.19	0.15	0.16	0.19	0.15	0.15
	0.003	0.05	7.06	12.27	32.21	7.18	12.60	34.09	0.23	0.24	0.38	0.23	0.24	0.39
		0.50	7.07	12.29	32.24	7.20	12.61	34.12	0.23	0.24	0.38	0.23	0.24	0.40
	0.005	0.05	15.79	28.76	80.49	16.06	29.48	85.10	0.29	0.36	0.66	0.29	0.37	0.69
		0.50	15.80	28.79	80.59	16.07	29.51	85.19	0.29	0.37	0.67	0.30	0.37	0.69

Model B with MN - Bias

ξ	λ_1	RV^Δ			HL			Daily Range				
		30min	15min	5min	30min	15min	5min	RG^D	RRG^D	RRG^D_{RS}	RRG^D_{BC}	$RRG^D_{RS,BC}$
0.001	0.000	8.52	15.88	47.21	0.80	0.49	-0.02	8.82	10.41	11.27	-5.88	-4.75
	1.298	7.98	16.11	47.46	-0.33	0.41	0.02	9.36	10.96	11.83	-5.35	-4.22
	2.596	72.10	142.33	425.87	-0.68	0.36	1.84	36.25	38.25	39.33	-5.62	-4.49
0.003	0.000	71.39	143.25	428.02	-0.13	-0.46	-0.54	36.75	38.76	39.84	-5.22	-4.09
	1.298	197.33	395.38	1187.21	2.34	0.04	2.55	66.63	69.07	70.39	-5.62	-4.49
	2.596	197.55	393.97	1187.13	0.59	2.22	0.75	67.00	69.44	70.77	-5.37	-4.24
0.005	0.000	8.06	16.09	47.34	0.08	0.17	0.36	9.75	11.36	12.23	-4.99	-3.85
	1.298	7.96	15.86	47.35	-0.28	0.19	0.13	9.54	11.14	12.01	-5.20	-4.06
	2.596	71.29	143.21	428.30	0.42	0.32	1.56	36.76	38.77	39.85	-5.28	-4.15

ξ	ν	RRG^Δ_m			$RRG^\Delta_{m_{RS}}$			$RRG^\Delta_{m,BC}$			$RRG^\Delta_{m_{RS},BC}$		
		30min	15min	5min	30min	15min	5min	30min	15min	5min	30min	15min	5min
0.001	0.000	46.35	68.55	133.35	48.69	72.66	146.50	-4.50	-4.96	-6.78	-3.65	-3.57	-4.11
	1.298	46.14	68.57	133.41	48.48	72.69	146.57	-4.61	-4.84	-6.59	-3.76	-3.45	-3.91
	2.596	186.84	298.86	680.40	191.42	308.60	724.41	-4.26	-4.38	-5.42	-3.41	-2.99	-2.71
0.003	0.000	186.92	298.88	680.62	191.51	308.61	724.64	-4.35	-4.53	-5.59	-3.49	-3.13	-2.88
	1.298	376.39	630.55	1556.32	384.01	648.38	1649.77	-3.89	-3.57	-3.12	-3.03	-2.16	-0.34
	2.596	376.50	630.51	1556.11	384.12	648.35	1649.59	-4.01	-3.80	-3.71	-3.16	-2.39	-0.95
0.005	0.000	46.09	68.38	133.16	48.43	72.49	146.30	-4.62	-4.97	-6.71	-3.77	-3.58	-4.03
	1.298	45.90	68.20	133.28	48.23	72.31	146.43	-4.83	-5.19	-6.79	-3.98	-3.81	-4.12
	2.596	187.30	299.49	682.26	191.90	309.24	726.38	-4.30	-4.51	-5.54	-3.45	-3.12	-2.83

Model B with MN - RMSE

ξ	λ_1	RV^{Δ}			HL			Daily Range				
		30min	15min	5min	30min	15min	5min	RG^D	RRG^D	RRG^D_{RS}	RRG^D_{BC}	$RRG^D_{RS,BC}$
0.001	0.000	0.42	0.35	0.52	0.70	0.50	0.32	0.67	0.68	0.68	0.61	0.62
	1.298	0.41	0.35	0.53	0.69	0.50	0.32	0.67	0.69	0.69	0.62	0.63
	2.596	0.96	1.54	4.33	0.89	0.78	0.85	0.82	0.84	0.85	0.62	0.63
0.003	0.000	0.95	1.56	4.35	0.88	0.77	0.84	0.83	0.85	0.86	0.62	0.63
	1.298	2.21	4.12	12.01	1.30	1.42	2.03	1.04	1.07	1.08	0.61	0.62
	2.596	2.22	4.10	12.01	1.31	1.42	2.01	1.06	1.08	1.09	0.62	0.63
0.005	0.000	0.41	0.35	0.52	0.69	0.50	0.32	0.68	0.69	0.70	0.62	0.63
	1.298	0.41	0.35	0.52	0.69	0.49	0.32	0.68	0.70	0.70	0.63	0.63
	2.596	0.95	1.56	4.35	0.88	0.77	0.87	0.82	0.84	0.85	0.61	0.62

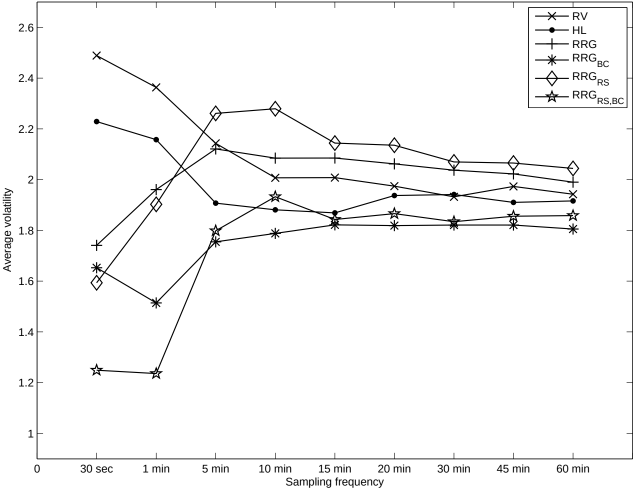
ξ	ν	RRG^{Δ}_m			$RRG^{\Delta}_{m_{RS}}$			$RRG^{\Delta}_{m,BC}$			$RRG^{\Delta}_{m_{RS},BC}$		
		30min	15min	5min	30min	15min	5min	30min	15min	5min	30min	15min	5min
0.001	0.000	0.51	0.71	1.34	0.53	0.75	1.47	0.18	0.14	0.12	0.18	0.14	0.11
	1.298	0.51	0.71	1.34	0.53	0.75	1.47	0.18	0.14	0.12	0.18	0.14	0.11
	2.596	1.89	3.00	6.82	1.94	3.10	7.26	0.19	0.16	0.18	0.19	0.16	0.17
0.003	0.000	1.89	3.00	6.82	1.94	3.10	7.26	0.19	0.15	0.18	0.19	0.15	0.17
	1.298	3.79	6.33	15.60	3.86	6.51	16.54	0.20	0.19	0.26	0.20	0.19	0.27
	2.596	3.79	6.33	15.60	3.87	6.50	16.54	0.20	0.19	0.25	0.20	0.19	0.26
0.005	0.000	0.51	0.70	1.34	0.53	0.74	1.47	0.18	0.14	0.12	0.18	0.14	0.11
	1.298	0.51	0.70	1.34	0.53	0.74	1.47	0.18	0.14	0.12	0.18	0.14	0.11
	2.596	1.90	3.01	6.84	1.94	3.11	7.28	0.19	0.16	0.18	0.19	0.15	0.17

Empirical Application

Tick-by-tick data for the IBM stock over the period from January 1, 2004 to December 31, 2006 for a total of 755 trading days. On average 5000 prices per day.

	Mean	Var	Skew	Kurt	Min	Max
RV^{30}	1.932	2.309	2.806	16.555	0.121	15.832
RV^{15}	2.008	2.338	3.510	23.762	0.230	16.730
RV^{5}	2.143	1.625	2.614	16.092	0.383	12.461
HL^{30}	1.941	3.656	2.345	10.699	0.263	13.855
HL^{15}	1.869	2.550	2.476	12.344	0.224	14.069
HL^{5}	1.907	2.036	3.336	21.582	0.303	13.399
$Park$	1.841	2.916	5.451	62.982	0.122	26.195
RRG^{Daily}	1.869	3.003	5.450	62.965	0.123	26.579
RRG^{Daily}_{RS}	1.755	2.676	5.406	62.480	0.111	25.073
RRG^{Daily}_{BC}	1.735	2.613	5.404	62.444	0.110	24.776
$RRG^{Daily}_{BC,RS}$	1.755	2.676	5.406	62.480	0.111	25.073
RRG^{30}	2.037	1.418	2.576	14.704	0.317	11.176
RRG^{15}	2.085	1.350	2.894	19.211	0.340	11.941
RRG^{5}	2.121	1.079	2.081	12.052	0.343	10.042
RRG^{30}_{RS}	2.070	1.461	2.582	14.756	0.321	11.443
RRG^{15}_{RS}	2.144	1.415	2.914	19.428	0.349	12.167
RRG^{5}_{RS}	2.261	1.204	2.052	11.835	0.370	10.624
RRG^{30}_{BC}	1.821	1.122	2.443	13.542	0.278	9.609
RRG^{15}_{BC}	1.822	1.011	2.735	18.201	0.284	10.681
RRG^{5}_{BC}	1.755	0.732	1.583	7.124	0.237	6.667
$RRG^{30}_{BC,RS}$	1.834	1.135	2.443	13.553	0.281	9.689
$RRG^{15}_{BC,RS}$	1.843	1.031	2.736	18.252	0.288	10.825
$RRG^{5}_{BC,RS}$	1.798	0.762	1.571	7.069	0.243	6.799

Empirical Application: Volatility signature plot



Conclusions

- With no MN
  1. underestimation of the true range leads to bias-RMSE trade-off in *RRG*.
  2. RRG_{RS} eliminates this trade-off.
- With MN.
  1. Bias: Hansen and Lunde is to be preferred, with the exception of strong LM in $\sigma(t)^2$.
  2. RRG_{BC} and $RRG_{RS,BC}$ are steadily downward biased.
  3. RMSE: RRG_{BC} and $RRG_{RS,BC}$ are three times more efficient than the Hansen and Lunde.
- In general *RRG* estimators provide an interesting alternative to RV-based estimators in presence of microstructure noise, in particular thanks to their higher precision