Properties of Realized Variance for Pure Jump Processes: Calendar Time Sampling versus Business Time Sampling

Symposium on "Econometric Forecasting and High-Frequency Data Analysis"

Jointly Organized by NUS & SMU

Singapore, 8 May 2004

Roel Oomen

Warwick Business School
The University of Warwick
Coventry, UK

roel.oomen@wbs.ac.uk

www.warwick.ac.uk/staff/R.C.A.Oomen

In This Talk . . .

[1] Properties of *realized variance* in the presence of market microstructure

In This Talk . . .

[1] Properties of *realized variance* in the presence of market microstructure

pure jump process extending Press (1967) model
closed form expressions for the bias and mean squared error
determine optimal sampling frequency
analyze alternative sampling schemes: BTS versus CTS

In This Talk . . .

[1] Properties of *realized variance* in the presence of market microstructure

pure jump process extending Press (1967) model
closed form expressions for the bias and mean squared error
determine optimal sampling frequency
analyze alternative sampling schemes: BTS versus CTS

IBM sampling frequency: 3 minutes, decay over time, noise ratio
IBM sampling scheme: BTS always outperforms CTS (!)

In This Talk . . .

$[1]$ Properties of *realized variance* in the presence of market microstructure

pure jump process extending Press (1967) model
closed form expressions for the bias and mean squared error
determine optimal sampling frequency
analyze alternative sampling schemes: BTS versus CTS

IBM sampling frequency: 3 minutes, decay over time, noise ratio
IBM sampling scheme: BTS always outperforms CTS (!)

$[2]$ Properties of *bias corrected realized variance* in the presence of market microstructure

In This Talk . . .

[1] Properties of *realized variance* in the presence of market microstructure

pure jump process extending Press (1967) model
closed form expressions for the bias and mean squared error
determine optimal sampling frequency
analyze alternative sampling schemes: BTS versus CTS

IBM sampling frequency: 3 minutes, decay over time, noise ratio
IBM sampling scheme: BTS always outperforms CTS (!)

[2] Properties of *bias corrected realized variance* in the presence of market microstructure

build on French, Schwert, and Stambaugh (1987), Hansen and Lunde (2004)
different (semi-parametric) price process, analysis of different sampling schemes

In This Talk . . .

[1] Properties of *realized variance* in the presence of market microstructure

pure jump process extending Press (1967) model
closed form expressions for the bias and mean squared error
determine optimal sampling frequency
analyze alternative sampling schemes: BTS versus CTS

IBM sampling frequency: 3 minutes, decay over time, noise ratio
IBM sampling scheme: BTS always outperforms CTS (!)

[2] Properties of *bias corrected realized variance* in the presence of market microstructure

build on French, Schwert, and Stambaugh (1987), Hansen and Lunde (2004)
different (semi-parametric) price process, analysis of different sampling schemes

IBM sampling frequency: 12 seconds (bias and MSE reduction over 65%!)
IBM sampling scheme: BTS always outperforms CTS

Motivation & Related Literature

- Realized variance (RV) defined as the sum of squared intra-period returns as a "feasible" or "discretized" version of the quadratic variation process

  see e.g. Andersen and Bollerslev (1998), Andersen, Bollerslev, Diebold, and Labys (2003), Barndorff-Nielsen and Shephard (2003), French, Schwert, and Stambaugh (1987), Hsieh (1988), Meddahi (2002), Merton (1980)

Motivation & Related Literature

- Realized variance (RV) defined as the sum of squared intra-period returns as a "feasible" or "discretized" version of the quadratic variation process
  see e.g. Andersen and Bollerslev (1998), Andersen, Bollerslev, Diebold, and Labys (2003), Barndorff-Nielsen and Shephard (2003), French, Schwert, and Stambaugh (1987), Hsieh (1988), Meddahi (2002), Merton (1980)

- Microstructure induced serial correlation renders realized variance biased

bias $\Downarrow$ by sampling *less* frequent $\Leftrightarrow$ variance $\Downarrow$ by sampling *more* frequent

Motivation & Related Literature

- Realized variance (RV) defined as the sum of squared intra-period returns as a "feasible" or "discretized" version of the quadratic variation process
  see e.g. Andersen and Bollerslev (1998), Andersen, Bollerslev, Diebold, and Labys (2003), Barndorff-Nielsen and Shephard (2003), French, Schwert, and Stambaugh (1987), Hsieh (1988), Meddahi (2002), Merton (1980)

- Microstructure induced serial correlation renders realized variance biased

  bias $\Downarrow$ by sampling *less* frequent $\Leftrightarrow$ variance $\Downarrow$ by sampling *more* frequent

- This tradeoff motives search for optimal sampling frequency
  see e.g. Andersen, Bollerslev, Diebold, and Labys (2000), Andreou and Ghysels (2001), Bai, Russell, and Tiao (2001), Bandi and Russell (2003), Oomen (2002a)

Motivation & Related Literature

- Realized variance (RV) defined as the sum of squared intra-period returns as a "feasible" or "discretized" version of the quadratic variation process
  see e.g. Andersen and Bollerslev (1998), Andersen, Bollerslev, Diebold, and Labys (2003), Barndorff-Nielsen and Shephard (2003), French, Schwert, and Stambaugh (1987), Hsieh (1988), Meddahi (2002), Merton (1980)

- Microstructure induced serial correlation renders realized variance biased

  bias $\Downarrow$ by sampling *less* frequent $\Leftrightarrow$ variance $\Downarrow$ by sampling *more* frequent

- This tradeoff motives search for optimal sampling frequency
  see e.g. Andersen, Bollerslev, Diebold, and Labys (2000), Andreou and Ghysels (2001), Bai, Russell, and Tiao (2001), Bandi and Russell (2003), Oomen (2002a)

✓ What I do here is (i) to characterize bias and MSE of RV and (ii) provide a flexible semi-parametric framework to determine optimal sampling frequency (iii) do all this under alternative sampling schemes

Motivation & Related Literature

- Aggregation (to optimal samp freq) may be "model-free" but is inefficient.

Motivation & Related Literature

- Aggregation (to optimal samp freq) may be "model-free" but is inefficient.

(1) filter (full use of information but model risk and measurement error)
(Andersen, Bollerslev, Diebold, and Ebens 2001)

Motivation & Related Literature

- Aggregation (to optimal samp freq) may be "model-free" but is inefficient.

(1) filter (full use of information but model risk and measurement error)
(Andersen, Bollerslev, Diebold, and Ebens 2001)

(2) model and bias correct (full use of information but model-risk)
(Corsi, Zumbach, Müller, and Dacorogna 2001)

Motivation & Related Literature

- Aggregation (to optimal samp freq) may be “model-free” but is inefficient.

(1) filter (full use of information but model risk and measurement error)
(Andersen, Bollerslev, Diebold, and Ebens 2001)

(2) model and bias correct (full use of information but model-risk)
(Corsi, Zumbach, Müller, and Dacorogna 2001)

(3) subsampling (full use of information but only asymptotic results)
(Zhang, Mykland, and Ait-Sahalia 2003)

Motivation & Related Literature

- Aggregation (to optimal samp freq) may be "model-free" but is inefficient.

(1) filter (full use of information but model risk and measurement error)
(Andersen, Bollerslev, Diebold, and Ebens 2001)

(2) model and bias correct (full use of information but model-risk)
(Corsi, Zumbach, Müller, and Dacorogna 2001)

(3) subsampling (full use of information but only asymptotic results)
(Zhang, Mykland, and Ait-Sahalia 2003)

(4) Newey-West style bias correction (some inefficiency, concentrated on bias)
(Hansen and Lunde 2004)

Motivation & Related Literature

- Aggregation (to optimal samp freq) may be "model-free" but is inefficient.

(1) filter (full use of information but model risk and measurement error)
(Andersen, Bollerslev, Diebold, and Ebens 2001)

(2) model and bias correct (full use of information but model-risk)
(Corsi, Zumbach, Müller, and Dacorogna 2001)

(3) subsampling (full use of information but only asymptotic results)
(Zhang, Mykland, and Ait-Sahalia 2003)

(4) Newey-West style bias correction (some inefficiency, concentrated on bias)
(Hansen and Lunde 2004)

✓ What I do here is (i) to characterize bias and MSE of bias corrected RV and (ii) provide a flexible semi-parametric framework to determine optimal sampling frequency (iii) do all this under alternative sampling schemes

A Pure Jump Process For High Frequency Financial Data

- Let the logarithmic price at time t, $P(t)$, follow CPP-MA(s), i.e.

$$P(t) = P(0) + \sum_{j=1}^{M(t)} (\varepsilon_j + \eta_j) \qquad \text{where} \qquad \eta_j = \rho_0 \nu_j + \rho_1 \nu_{j-1} + \ldots + \rho_s \nu_{j-s}$$

where $\varepsilon_j \sim$iid $\mathcal{N}\left(\mu_\varepsilon, \sigma_\varepsilon^2\right)$, $\nu_j \sim$ iid $\mathcal{N}\left(\mu_\nu, \sigma_\nu^2\right)$, and $M(t)$ is a Poisson process with instantaneous intensity $\lambda(t) > 0$

A Pure Jump Process For High Frequency Financial Data

- Let the logarithmic price at time t, $P(t)$, follow CPP-MA(s), i.e.

$$P(t) = P(0) + \sum_{j=1}^{M(t)} (\varepsilon_j + \eta_j) \quad \text{where} \quad \eta_j = \rho_0 \nu_j + \rho_1 \nu_{j-1} + \ldots + \rho_s \nu_{j-s}$$

  where $\varepsilon_j \sim$iid $\mathcal{N}\left(\mu_\varepsilon, \sigma_\varepsilon^2\right)$, $\nu_j \sim$ iid $\mathcal{N}\left(\mu_\nu, \sigma_\nu^2\right)$, and $M(t)$ is a Poisson process with instantaneous intensity $\lambda(t) > 0$

- Pure jump processes not widely used in finance (that is relative to for example SV and ARCH) . . . notable exceptions include:

Low Frequency:	Press (1967), Maheu and McCurdy (2003, 2004), Piazzesi (2004)
High Frequency:	Bowsher (2002), Oomen (2002b), Rogers and Zane (1998) Rydberg and Shephard (2003)
Option Pricing:	Barndorff-Nielsen and Shephard (2004), Carr and Wu (2003) Geman, Madan, and Yor (2001), Mürmann (2001)

Market Microstructure Noise Interpretation of the Model

- Let P_k denote the logarithmic price after the k^{th} transaction, i.e.

$$P_k = P_0 + \sum_{j=1}^{k} \varepsilon_j + \sum_{j=1}^{k} \eta_j = P_k^e + \sum_{j=1}^{k} \eta_j$$

"CONTAMINATED PRICE" = "EFFICIENT PRICE" + "ACCUMULATED NOISE"

Market Microstructure Noise Interpretation of the Model

- Let P_k denote the logarithmic price after the k^{th} transaction, i.e.

$$P_k = P_0 + \sum_{j=1}^{k} \varepsilon_j + \sum_{j=1}^{k} \eta_j = P_k^e + \sum_{j=1}^{k} \eta_j$$

“CONTAMINATED PRICE” = “EFFICIENT PRICE” + “ACCUMULATED NOISE”

- Restrict MA(s) parameters to avoid accumulation of noise. For example, for MA(1) impose $\rho_0 = -\rho_1 = 1$

$$\eta_k = (\nu_k - \nu_{k-1}) \Rightarrow \sum_{j=1}^{k} \eta_j = (\nu_k - \nu_0)$$

Market Microstructure Noise Interpretation of the Model

- Let P_k denote the logarithmic price after the k^{th} transaction, i.e.

$$P_k = P_0 + \sum_{j=1}^{k} \varepsilon_j + \sum_{j=1}^{k} \eta_j = P_k^e + \sum_{j=1}^{k} \eta_j$$

"CONTAMINATED PRICE" = "EFFICIENT PRICE" + "ACCUMULATED NOISE"

- Restrict MA(s) parameters to avoid accumulation of noise. For example, for MA(1) impose $\rho_0 = -\rho_1 = 1$

$$\eta_k = (\nu_k - \nu_{k-1}) \Rightarrow \sum_{j=1}^{k} \eta_j = (\nu_k - \nu_0)$$

✓ Negative serial correlation of returns (OK for transaction data)

✓ Higher order MA(s) and different restrictions can lead to positive serial correlation

Properties of the Price Process

- In transaction time returns are normal (Ané and Geman 2000).
- In calendar time returns are (highly) non-normal

Properties of the Price Process

- In transaction time returns are normal (Ané and Geman 2000).
- In calendar time returns are (highly) non-normal
- The joint characteristic function of $\{R(t_1|\tau_1), R(t_2|\tau_2), R(t_3|\tau_3), R(t_4|\tau_4)\}$ *conditional on intensity process* for the restricted CPP-MA(1) can be derived as:

$$
\begin{aligned}
& e^{-\lambda_4-\lambda_3-\lambda_2-\lambda_1} + e^{-\lambda_2-\lambda_3}\Upsilon(1)\,\Lambda_1\left(e^{c_1}\lambda_1\right)\Upsilon(4)\,\Lambda_1\left(e^{c_4}\lambda_4\right)\exp\left\{\xi_1\xi_4\sigma_\nu^2-\lambda_{1,2}-\lambda_{2,3}-\lambda_{3,4}\right\} \\
& + e^{-\lambda_2-\lambda_4}\Upsilon(1)\,\Lambda_1\left(e^{c_1}\lambda_1\right)\Upsilon(3)\,\Lambda_1\left(e^{c_3}\lambda_3\right)\exp\left\{\xi_1\xi_3\sigma_\nu^2-\lambda_{1,2}-\lambda_{2,3}\right\} + e^{-\lambda_4-\lambda_3-\lambda_2}\Upsilon(1)\,\Lambda_1\left(e^{c_1}\lambda_1\right) \\
& + e^{-\lambda_1-\lambda_3}\Upsilon(2)\,\Lambda_1\left(e^{c_2}\lambda_2\right)\Upsilon(4)\,\Lambda_1\left(e^{c_4}\lambda_4\right)\exp\left\{\xi_2\xi_4\sigma_\nu^2-\lambda_{2,3}-\lambda_{3,4}\right\} + e^{-\lambda_4-\lambda_3-\lambda_1}\Upsilon(2)\,\Lambda_1\left(e^{c_2}\lambda_2\right) \\
& + e^{-\lambda_3}\Upsilon(1)\,\Lambda_1\left(e^{c_1}\lambda_1\right)\Upsilon(2)\,\Lambda_1\left(e^{c_2}\lambda_2\right)\Upsilon(4)\,\Lambda_1\left(e^{c_4}\lambda_4\right)\Psi_{12}\Psi_{24} + e^{-\lambda_1-\lambda_2}\Upsilon(3)\,\Lambda_1\left(e^{c_3}\lambda_3\right)\Upsilon(4)\,\Lambda_1\left(e^{c_4}\lambda_4\right)\Psi_{34} \\
& + e^{-\lambda_2}\Upsilon(1)\,\Lambda_1\left(e^{c_1}\lambda_1\right)\Upsilon(3)\,\Lambda_1\left(e^{c_3}\lambda_3\right)\Upsilon(4)\,\Lambda_1\left(e^{c_4}\lambda_4\right)\Psi_{13}\Psi_{34} + e^{-\lambda_3-\lambda_4}\Upsilon(1)\,\Lambda_1\left(e^{c_1}\lambda_1\right)\Upsilon(2)\,\Lambda_1\left(e^{c_2}\lambda_2\right)\Psi_{12} \\
& + e^{-\lambda_1}\Upsilon(2)\,\Lambda_1\left(e^{c_2}\lambda_2\right)\Upsilon(3)\,\Lambda_1\left(e^{c_3}\lambda_3\right)\Upsilon(4)\,\Lambda_1\left(e^{c_4}\lambda_4\right)\Psi_{23}\Psi_{34} + e^{-\lambda_1-\lambda_4}\Upsilon(2)\,\Lambda_1\left(e^{c_2}\lambda_2\right)\Upsilon(3)\,\Lambda_1\left(e^{c_3}\lambda_3\right)\Psi_{23} \\
& + e^{-\lambda_4}\Upsilon(1)\,\Lambda_1\left(e^{c_1}\lambda_1\right)\Upsilon(2)\,\Lambda_1\left(e^{c_2}\lambda_2\right)\Upsilon(3)\,\Lambda_1\left(e^{c_3}\lambda_3\right)\Psi_{12}\Psi_{23} + e^{-\lambda_4-\lambda_2-\lambda_1}\Upsilon(3)\,\Lambda_1\left(e^{c_3}\lambda_3\right) \\
& + \Upsilon(1)\,\Lambda_1\left(e^{c_1}\lambda_1\right)\Upsilon(2)\,\Lambda_1\left(e^{c_2}\lambda_2\right)\Upsilon(3)\,\Lambda_1\left(e^{c_3}\lambda_3\right)\Upsilon(4)\,\Lambda_1\left(e^{c_4}\lambda_4\right)\Psi_{12}\Psi_{23}\Psi_{34} + e^{-\lambda_3-\lambda_2-\lambda_1}\Upsilon(4)\,\Lambda_1\left(e^{c_4}\lambda_4\right)
\end{aligned}
$$

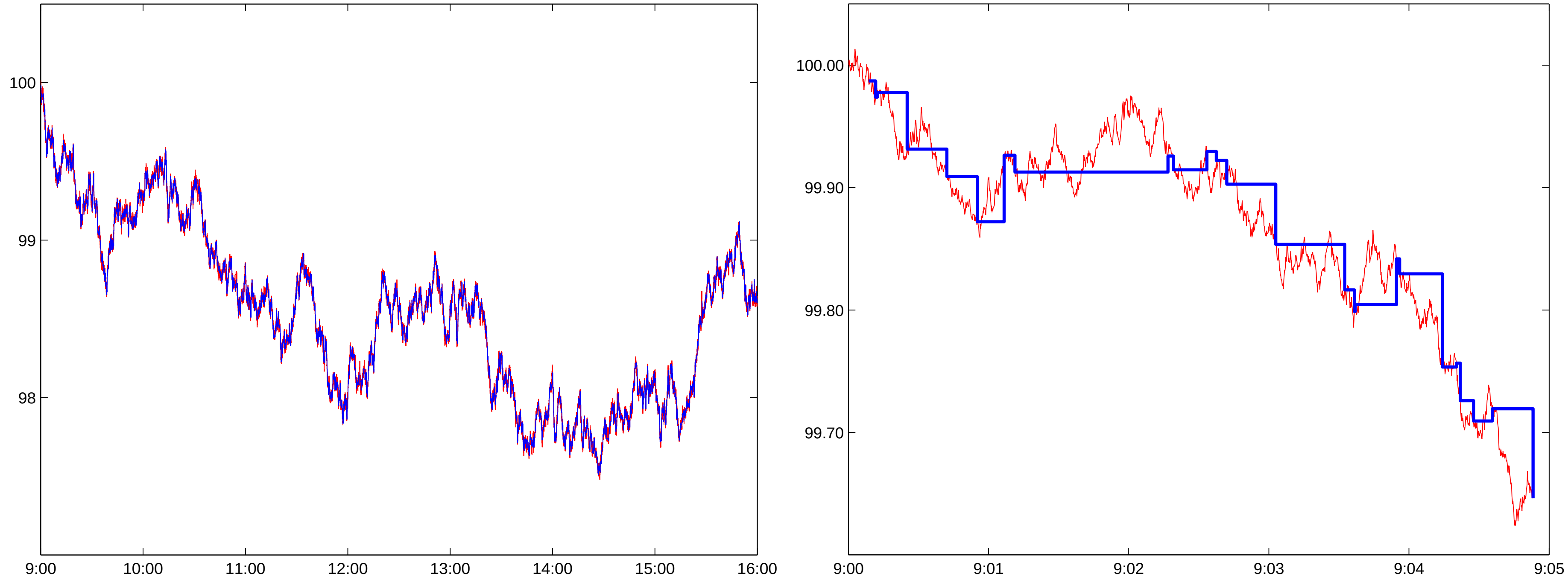
where $\Psi_{ij} = \left(\exp\left\{\xi_i\xi_j\sigma_\nu^2-\lambda_{i,j}\right\}+\Lambda_1\left(\lambda_{i,j}\right)\right)$, $\Upsilon(i) = \exp\left(\xi_i^2 c_0 + \left(e^{c_i}-1\right)\lambda_i\right)$, and $\Lambda_q(x) = 1-\frac{\Gamma(q,x)}{\Gamma(q)}$.

Properties of the Price Process

- Stochastic intensity for the CPP-MA(s) $\Leftrightarrow$ stochastic volatility for SV
- CPP-MA(s) can capture seasonals, ACD & ARCH effects, serial correlation, fat tails

Properties of the Price Process

- Stochastic intensity for the CPP-MA(s) ⇔ stochastic volatility for SV
- CPP-MA(s) can capture seasonals, ACD & ARCH effects, serial correlation, fat tails
- A fundamental difference with diffusive process: CPP-MA(s) is of finite variation



Alternative Sampling Schemes

- ⋆ General Time Sampling: Under GTS_N, the price process is sampled at time points $\{t_0^g, \ldots, t_N^g\}$ over the interval $[t_0, t_0 + T]$ such that $t_0^g = t_0$, $t_N^g = t_0 + T$, and $t_i^g < t_{i+1}^g$.

Alternative Sampling Schemes

- General Time Sampling: Under GTS_N, the price process is sampled at time points $\{t_0^g, \ldots, t_N^g\}$ over the interval $[t_0, t_0 + T]$ such that $t_0^g = t_0$, $t_N^g = t_0 + T$, and $t_i^g < t_{i+1}^g$.

- Calendar Time Sampling: Under CTS_N, the price process is sampled at equidistantly spaced points in calendar time over the interval $[t_0, t_0 + T]$, i.e. $t_i^p = t_0 + i\delta$ for $i = \{0, \ldots, NT\}$ where $N = \delta^{-1}$.

Alternative Sampling Schemes

- ⋆ General Time Sampling: Under GTS_N, the price process is sampled at time points $\{t_0^g, \ldots, t_N^g\}$ over the interval $[t_0, t_0 + T]$ such that $t_0^g = t_0$, $t_N^g = t_0 + T$, and $t_i^g < t_{i+1}^g$.

- ⋆ Calendar Time Sampling: Under CTS_N, the price process is sampled at equidistantly spaced points in calendar time over the interval $[t_0, t_0 + T]$, i.e. $t_i^p = t_0 + i\delta$ for $i = \{0, \ldots, NT\}$ where $N = \delta^{-1}$.

- ⋆ Business Time Sampling: Under BTS_N, the price process is sampled at equidistantly spaced points in business time over the interval $[t_0, t_0 + T]$, i.e. t_i^b for $i = \{0, \ldots, NT\}$ such that $t_0^b = t_0$, $t_N^b = t_0 + T$ and

$$\int_{t_i^b}^{t_{i+1}^b} \lambda(u)du = \frac{1}{NT} \int_{t_0}^{t_0+T} \lambda(u)du \equiv \lambda_N$$

✓ Throughout I translate N to corresponding sampling frequency in minutes

RV in Absence of Market Microstructure Noise

- To set the stage I first consider the CPP-MA(0) (RV is unbiased)

$$MSE\left(GTS_N\right) = \sum_{i=1}^{N}\left(3\sigma_\varepsilon^4\lambda_i^2 + 3\sigma_\varepsilon^4\lambda_i\right) + 2\sigma_\varepsilon^4\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}\lambda_i\lambda_j - \lambda_{(0,1)}^2\sigma_\varepsilon^4$$

⋆ Notation: $\lambda_i = \int_{t_i-\tau_i}^{t_i}\lambda(u)du$ and $\lambda_{(0,1)} = \int_0^1\lambda(u)du$

RV in Absence of Market Microstructure Noise

- To set the stage I first consider the CPP-MA(0) (RV is unbiased)

$$MSE\left(GTS_N\right) = \sum_{i=1}^{N} \left(3\sigma_\varepsilon^4 \lambda_i^2 + 3\sigma_\varepsilon^4 \lambda_i\right) + 2\sigma_\varepsilon^4 \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \lambda_i \lambda_j - \lambda_{(0,1)}^2 \sigma_\varepsilon^4$$

…which simplifies under BTS to:

$$MSE\left(BTS_N\right) = 2N^{-1} \left(\lambda_{(0,1)}\sigma_\varepsilon^2\right)^2 + 3\sigma_\varepsilon^2 \left(\lambda_{(0,1)}\sigma_\varepsilon^2\right)$$

⋆ Notation: $\lambda_i = \int_{t_i-\tau_i}^{t_i} \lambda(u)du$ and $\lambda_{(0,1)} = \int_0^1 \lambda(u)du$

RV in Absence of Market Microstructure Noise

- To set the stage I first consider the CPP-MA(0) (RV is unbiased)

$$MSE\left(GTS_N\right) = \sum_{i=1}^{N} \left(3\sigma_\varepsilon^4 \lambda_i^2 + 3\sigma_\varepsilon^4 \lambda_i\right) + 2\sigma_\varepsilon^4 \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \lambda_i \lambda_j - \lambda_{(0,1)}^2 \sigma_\varepsilon^4$$

  ...which simplifies under BTS to:

$$MSE\left(BTS_N\right) = 2N^{-1}\left(\lambda_{(0,1)}\sigma_\varepsilon^2\right)^2 + 3\sigma_\varepsilon^2\left(\lambda_{(0,1)}\sigma_\varepsilon^2\right)$$

- RV is inconsistent under pure jump process
- Consistency in "diffusion" limit where $\lambda \rightarrow \infty$ while $\lambda\sigma_\varepsilon^2$ constant

⋆ Notation: $\lambda_i = \int_{t_i - \tau_i}^{t_i} \lambda(u)du$ and $\lambda_{(0,1)} = \int_0^1 \lambda(u)du$

Absence of Market Microstructure Noise

- The difference in MSE among different sampling schemes can be derived as:

$$MSE\left(GTS_N\right) - MSE\left(BTS_N\right) = 3\sigma_\varepsilon^4 \sum_{i=1}^{N} \left(\lambda_i^2 - \lambda_N^2\right) + 2\sigma_\varepsilon^4 \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \left(\lambda_i \lambda_j - \lambda_N^2\right)$$
$$= 2\sigma_\varepsilon^4 \sum_{i=1}^{N} \vartheta_i^2 > 0$$

where $\vartheta_i = \int_{t_i-\tau_i}^{t_i} \lambda\left(u\right) du - \lambda_N$

Absence of Market Microstructure Noise

- The difference in MSE among different sampling schemes can be derived as:

$$
\begin{aligned}
MSE\left(GTS_N\right) - MSE\left(BTS_N\right) &= 3\sigma_\varepsilon^4 \sum_{i=1}^{N} \left(\lambda_i^2 - \lambda_N^2\right) + 2\sigma_\varepsilon^4 \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \left(\lambda_i \lambda_j - \lambda_N^2\right) \\
&= 2\sigma_\varepsilon^4 \sum_{i=1}^{N} \vartheta_i^2 > 0
\end{aligned}
$$

  where $\vartheta_i = \int_{t_i - \tau_i}^{t_i} \lambda\left(u\right) du - \lambda_N$

- *In the absence of market microstructure noise, the realized variance measure under BTS is more efficient than under any other conceivable sampling scheme*

Absence of Market Microstructure Noise

- The difference in MSE among different sampling schemes can be derived as:

$$MSE\left(GTS_N\right) - MSE\left(BTS_N\right) = 3\sigma_\varepsilon^4 \sum_{i=1}^{N} \left(\lambda_i^2 - \lambda_N^2\right) + 2\sigma_\varepsilon^4 \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \left(\lambda_i \lambda_j - \lambda_N^2\right)$$
$$= 2\sigma_\varepsilon^4 \sum_{i=1}^{N} \vartheta_i^2 > 0$$

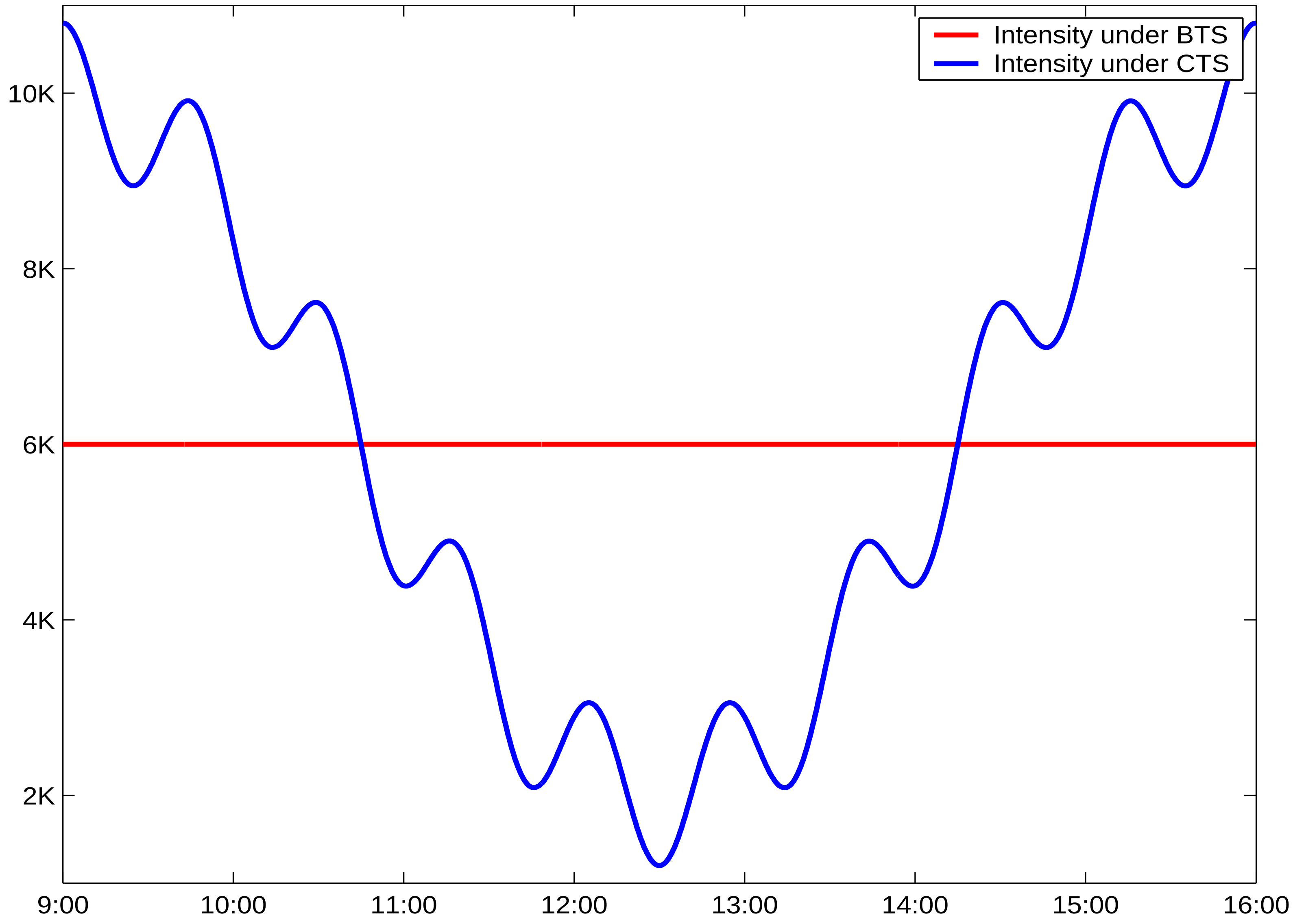
where $\vartheta_i = \int_{t_i - \tau_i}^{t_i} \lambda\left(u\right) du - \lambda_N$

- *In the absence of market microstructure noise, the realized variance measure under BTS is more efficient than under any other conceivable sampling scheme*

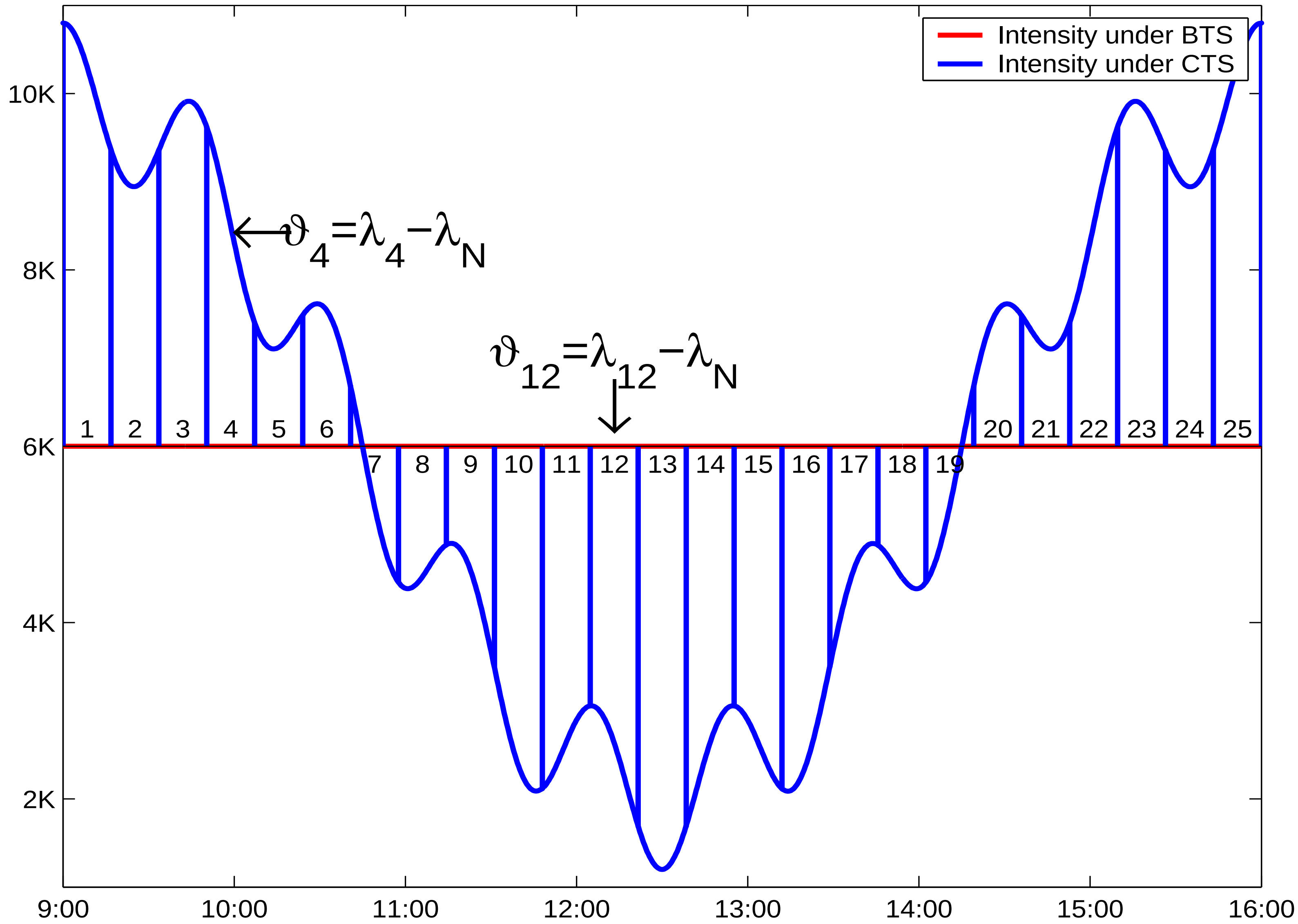
- The efficiency gain associated with BTS, relative to CTS, increases with

  (i) an increase in the variability of trade intensity
  (ii) an increase in the variance of the price innovations

Integrated Incremental Intensity under BTS and CTS: ϑ



Integrated Incremental Intensity under BTS and CTS: ϑ



Presence of Market Microstructure Noise – The Bias . . .

- Now move on to the CPP-MA(1), i.e. first order dependence in noise component

$$Bias\left(GTS_N\right) = E_\lambda \left[\sum_{i=1}^{N} R\left(t_i|\tau_i\right)^2\right] - \lambda_{(0,1)}\sigma_\varepsilon^2 = 2\sigma_\nu^2 \sum_{i=1}^{N} \left(1 - e^{-\lambda_i}\right) > 0$$

Presence of Market Microstructure Noise – The Bias ...

- Now move on to the CPP-MA(1), i.e. first order dependence in noise component

$$Bias\left(GTS_N\right) = E_\lambda\left[\sum_{i=1}^{N} R\left(t_i|\tau_i\right)^2\right] - \lambda_{(0,1)}\sigma_\varepsilon^2 = 2\sigma_\nu^2\sum_{i=1}^{N}\left(1 - e^{-\lambda_i}\right) > 0$$

...and the difference in bias between two sampling schemes is:

$$Bias\left(GTS_N\right) - Bias\left(BTS_N\right) = 2e^{-\lambda_N}\sigma_\nu^2\sum_{i=1}^{N}\left(1 - e^{-\vartheta_i}\right) < 0$$

- *In the presence of first order market microstructure noise, the bias of the realized variance measure under BTS is larger than under any other sampling scheme*

Presence of Market Microstructure Noise – The Bias ...

- Now move on to the CPP-MA(1), i.e. first order dependence in noise component

$$Bias\left(GTS_N\right) = E_\lambda\left[\sum_{i=1}^{N} R\left(t_i|\tau_i\right)^2\right] - \lambda_{(0,1)}\sigma_\varepsilon^2 = 2\sigma_\nu^2 \sum_{i=1}^{N}\left(1 - e^{-\lambda_i}\right) > 0$$

...and the difference in bias between two sampling schemes is:

$$Bias\left(GTS_N\right) - Bias\left(BTS_N\right) = 2e^{-\lambda_N}\sigma_\nu^2 \sum_{i=1}^{N}\left(1 - e^{-\vartheta_i}\right) < 0$$

- *In the presence of first order market microstructure noise, the bias of the realized variance measure under BTS is larger than under any other sampling scheme*

- When $N = 1$ or $N \to \infty$ all sampling schemes are equivalent.

$$\lim_{N\to\infty} Bias\left(GTS_N\right) = 2\sigma_\nu^2\lambda_{(0,1)}$$

Presence of Market Microstructure Noise – The MSE …

- For CPP-MA(1), CTS may perform better than BTS in terms of MSE…

  …compute %-loss in MSE under CTS relative to BTS (i.e. “CTS loss”)

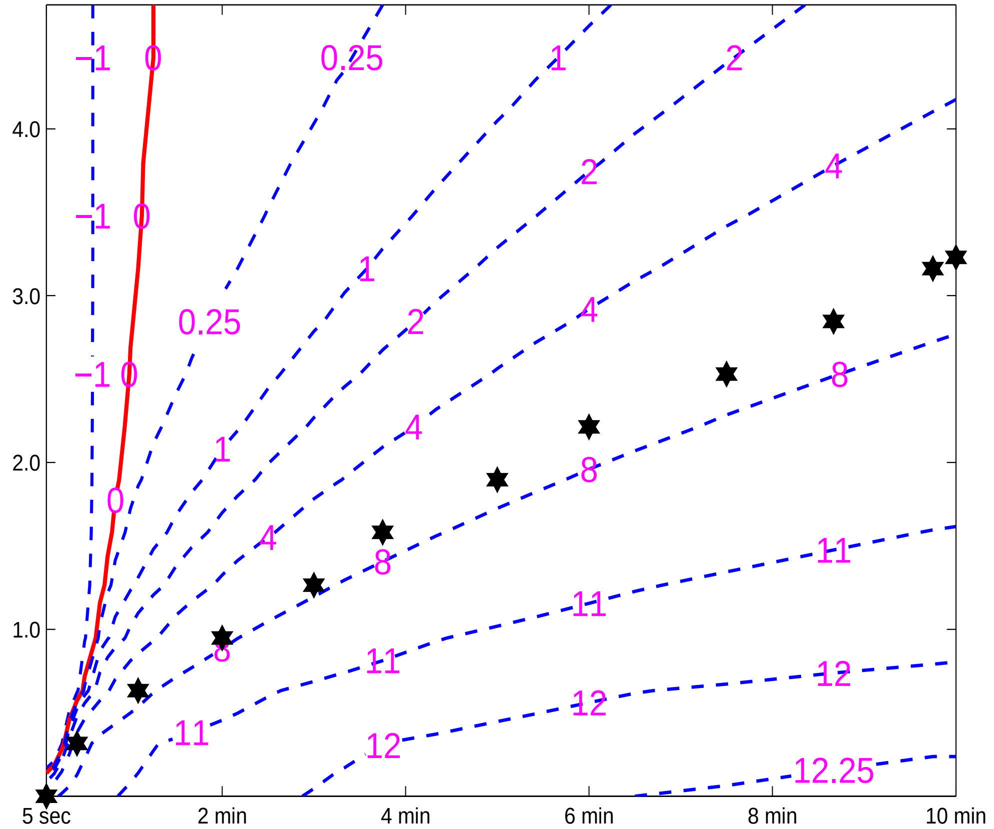
Presence of Market Microstructure Noise – The MSE . . .

- For CPP-MA(1), CTS may perform better than BTS in terms of MSE. . .

  . . . compute %-loss in MSE under CTS relative to BTS (i.e. "CTS loss")

- CTS loss at different sampling frequencies and noise ratios ($\sigma_\nu/\sigma_\varepsilon$)

- Asterisks indicate optimal sampling frequency



Presence of Market Microstructure Noise – The MSE . . .

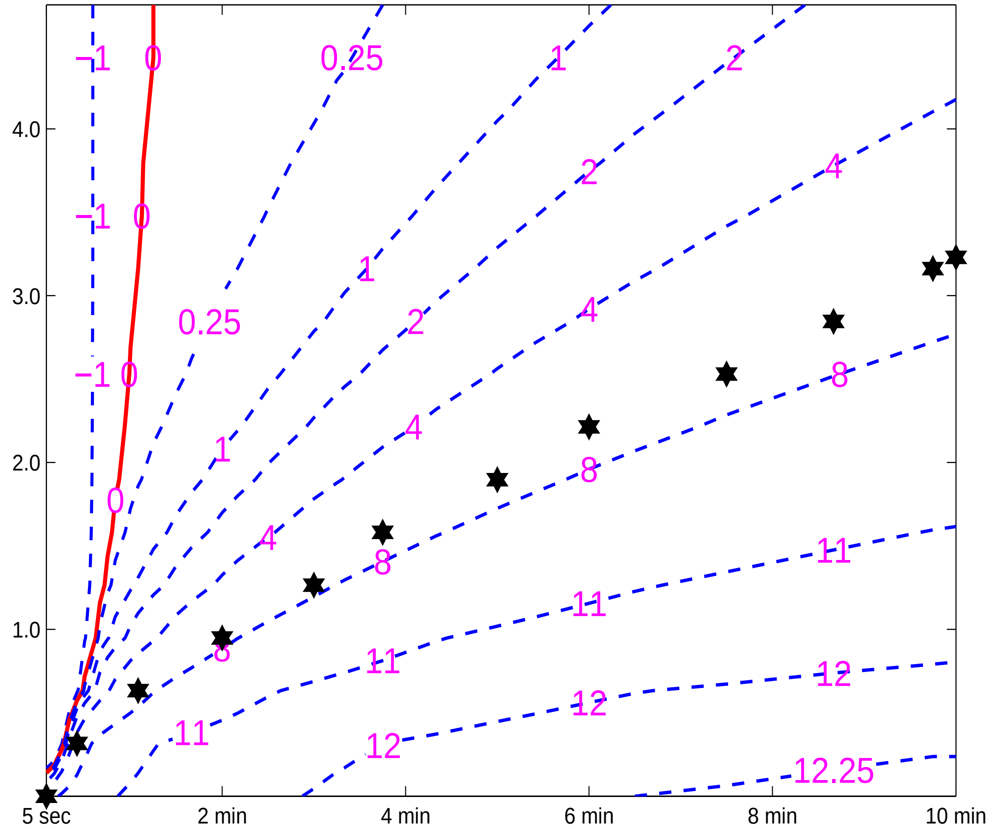
- For CPP-MA(1), CTS may perform better than BTS in terms of MSE. . .

  . . . compute %-loss in MSE under CTS relative to BTS (i.e. "CTS loss")

- CTS loss at different sampling frequencies and noise ratios ($\sigma_\nu/\sigma_\varepsilon$)

- Asterisks indicate optimal sampling frequency

  BTS always outperforms CTS along "optimal sampling frequency"!



Presence of Market Microstructure Noise – The MSE . . .

- For CPP-MA(1), CTS may perform better than BTS in terms of MSE. . .

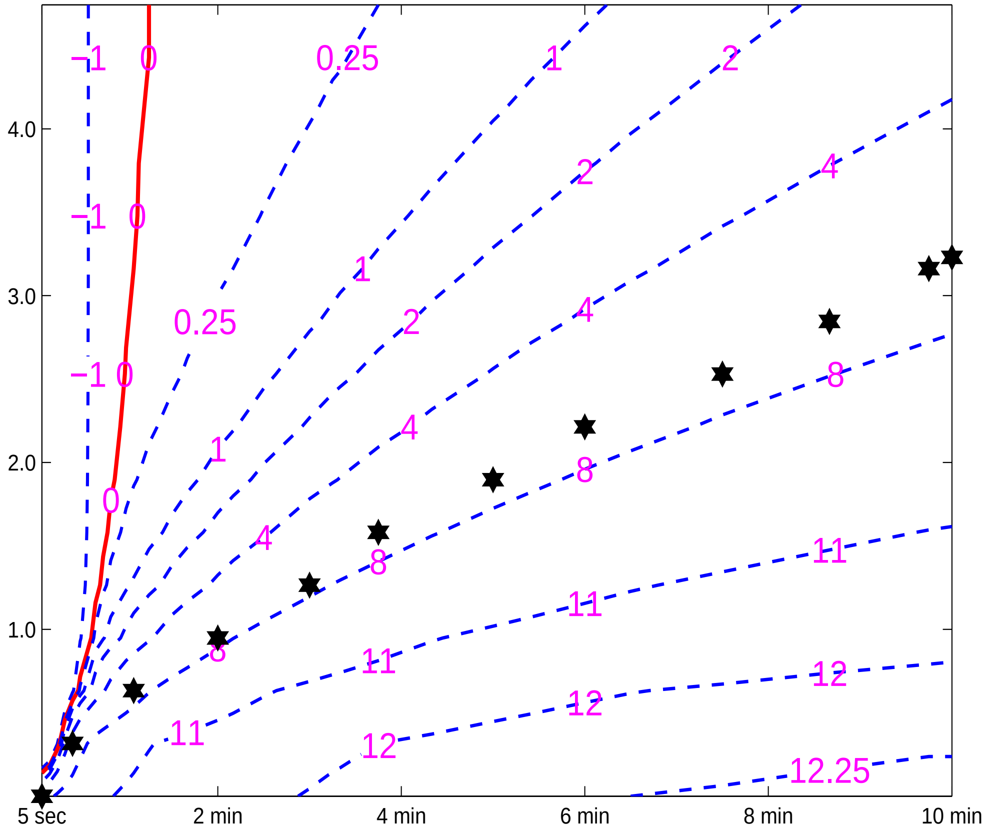
  . . . compute %-loss in MSE under CTS relative to BTS (i.e. “CTS loss”)

- CTS loss at different sampling frequencies and noise ratios ($\sigma_\nu/\sigma_\varepsilon$)

- Asterisks indicate optimal sampling frequency

  BTS always outperforms CTS along “optimal sampling frequency”!

- Similar results for higher order dependence, i.e. CPP-MA(q)



Empirical Analysis

- Estimate the model parameters of the restricted CPP-MA(1) and CPP-MA(2) using IBM transaction data

  (i) determine the optimal sampling frequency

  (ii) measure the improvement in MSE resulting from BTS relative to CTS.

Empirical Analysis

- Estimate the model parameters of the restricted CPP-MA(1) and CPP-MA(2) using IBM transaction data

  (i) determine the optimal sampling frequency

  (ii) measure the improvement in MSE resulting from BTS relative to CTS.

- Data available through Trade and Quote (TAQ) database from NYSE.

  January 1, 2000 until 31 August 2003 (917 days)

  Transactions from all exchanges between 9.45 and 16.00

  Filter for instantaneous price reversals (detect 1358)

  Total number of transactions is 5,522,929

Empirical Analysis – Parameter Estimation

- σ_ε, σ_ν, and ρ estimated in "business time" using all transactions

$$Cov\left(R_k, R_{k-1}\right) = -\left(1-\rho\right)^2 \sigma_\nu^2$$

$$Cov\left(R_k, R_{k-2}\right) = -\rho\sigma_\nu^2$$

$$Var\left(R_k\right) = \sigma_\varepsilon^2 + \left(2 + 2\rho^2 - 2\rho\right)\sigma_\nu^2$$

- Solve system of equations (choose solution with $\sigma_\nu^2 > 0$ and $|\rho| < 1$)

Empirical Analysis – Parameter Estimation

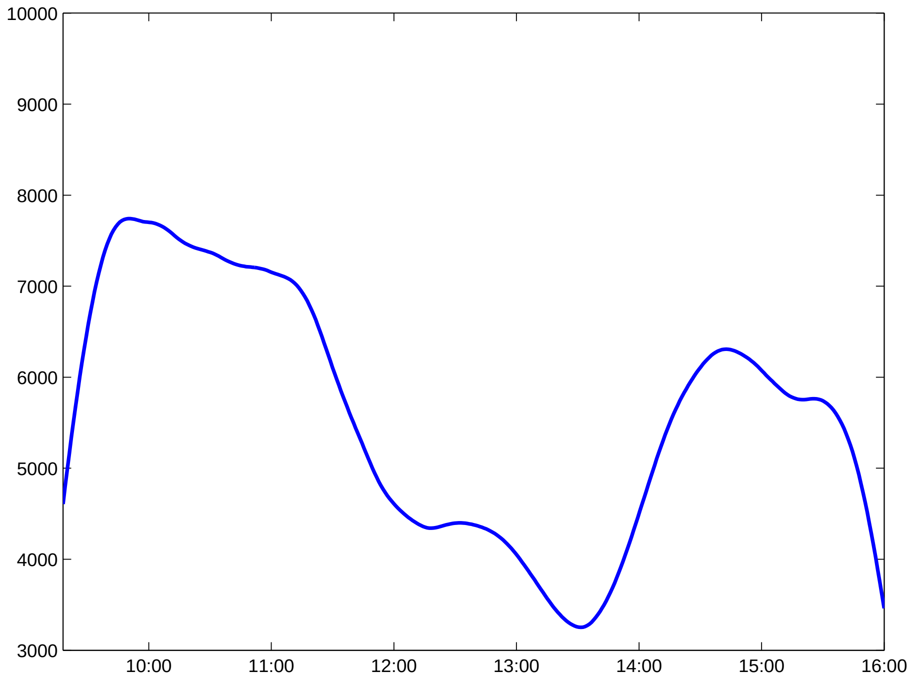
- σ_ε, σ_ν, and ρ estimated in "business time" using all transactions

$$Cov\left(R_k, R_{k-1}\right) = -\left(1-\rho\right)^2 \sigma_\nu^2$$

$$Cov\left(R_k, R_{k-2}\right) = -\rho\sigma_\nu^2$$

$$Var\left(R_k\right) = \sigma_\varepsilon^2 + \left(2 + 2\rho^2 - 2\rho\right)\sigma_\nu^2$$

- Solve system of equations (choose solution with $\sigma_\nu^2 > 0$ and $|\rho| < 1$)



- $\lambda\left(t\right)$ is estimated using non-parametric smoothing techniques (Cowling, Hall, and Phillips 1996) similar to density estimation

Empirical Analysis – Parameter Estimation

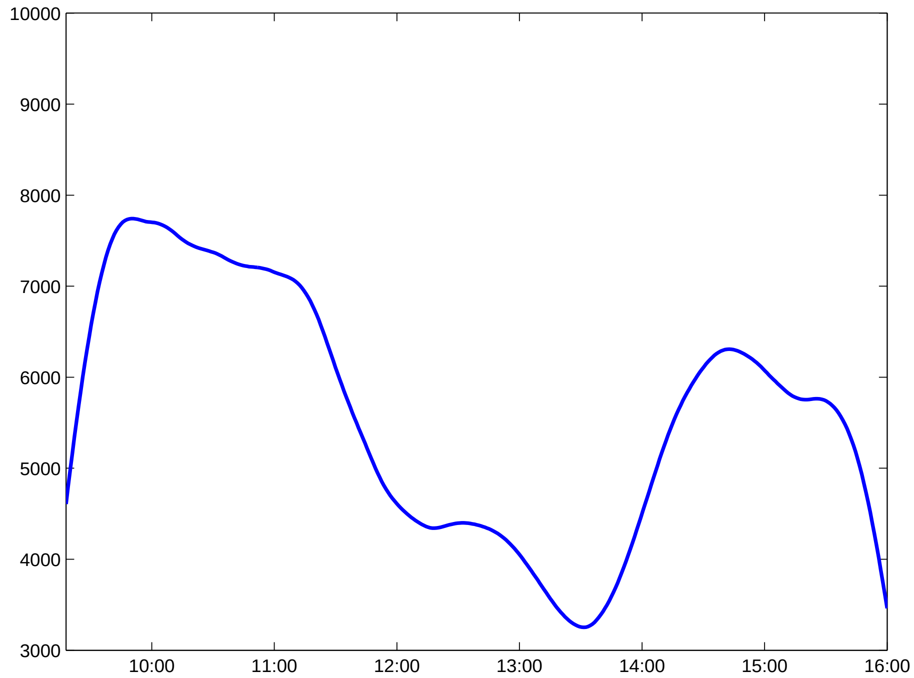
- σ_ε, σ_ν, and ρ estimated in "business time" using all transactions

$$Cov\left(R_k, R_{k-1}\right) = -\left(1-\rho\right)^2 \sigma_\nu^2$$

$$Cov\left(R_k, R_{k-2}\right) = -\rho\sigma_\nu^2$$

$$Var\left(R_k\right) = \sigma_\varepsilon^2 + \left(2 + 2\rho^2 - 2\rho\right)\sigma_\nu^2$$

- Solve system of equations (choose solution with $\sigma_\nu^2 > 0$ and $|\rho| < 1$)



- $\lambda\left(t\right)$ is estimated using non-parametric smoothing techniques (Cowling, Hall, and Phillips 1996) similar to density estimation

(i) Bias at the edges!
(ii) How significant?

Empirical Analysis – Parameter Estimation

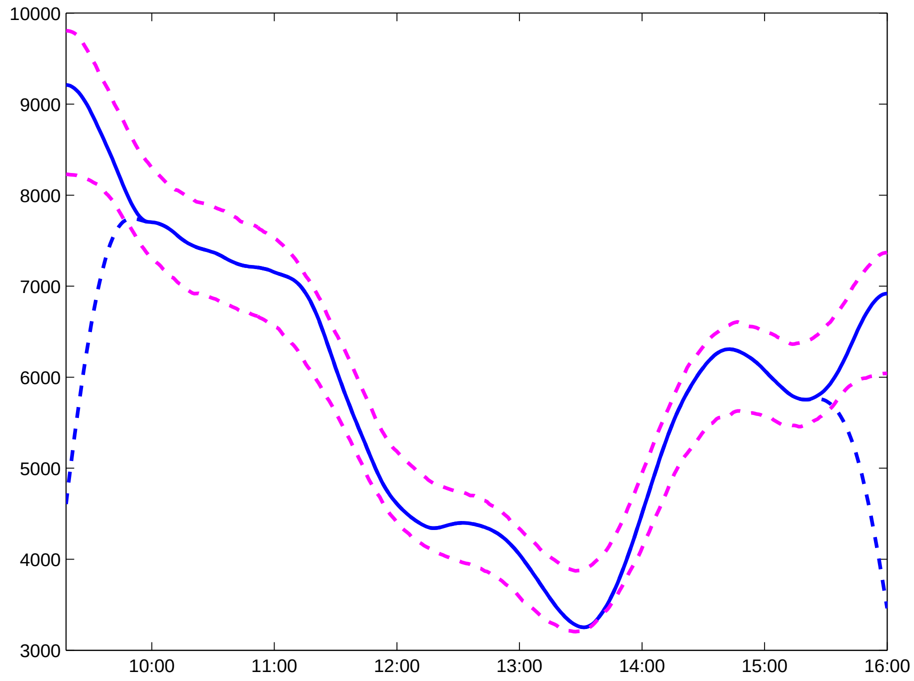
- σ_ε, σ_ν, and ρ estimated in "business time" using all transactions

$$Cov\left(R_k, R_{k-1}\right) = -\left(1-\rho\right)^2 \sigma_\nu^2$$

$$Cov\left(R_k, R_{k-2}\right) = -\rho\sigma_\nu^2$$

$$Var\left(R_k\right) = \sigma_\varepsilon^2 + \left(2 + 2\rho^2 - 2\rho\right)\sigma_\nu^2$$

- Solve system of equations (choose solution with $\sigma_\nu^2 > 0$ and $|\rho| < 1$)



- $\lambda\left(t\right)$ is estimated using non-parametric smoothing techniques (Cowling, Hall, and Phillips 1996) similar to density estimation

(i) Bias at the edges! ("mirror image" correction Diggle and Marron (1988))
(ii) How significant? (bootstrap based on Cowling, Hall, and Phillips (1996))

Impact of Measurement Error on Optimal Sampling Frequency

1 Simulate transaction data based on the CPP-MA(2) model ($\rho = 0.6$, $\lambda_{(0,1)} = 5000$, $\sigma_\nu/\sigma_\varepsilon = 1.1$, annualized return volatility of 25%)

2 Estimate model parameters as outlined above

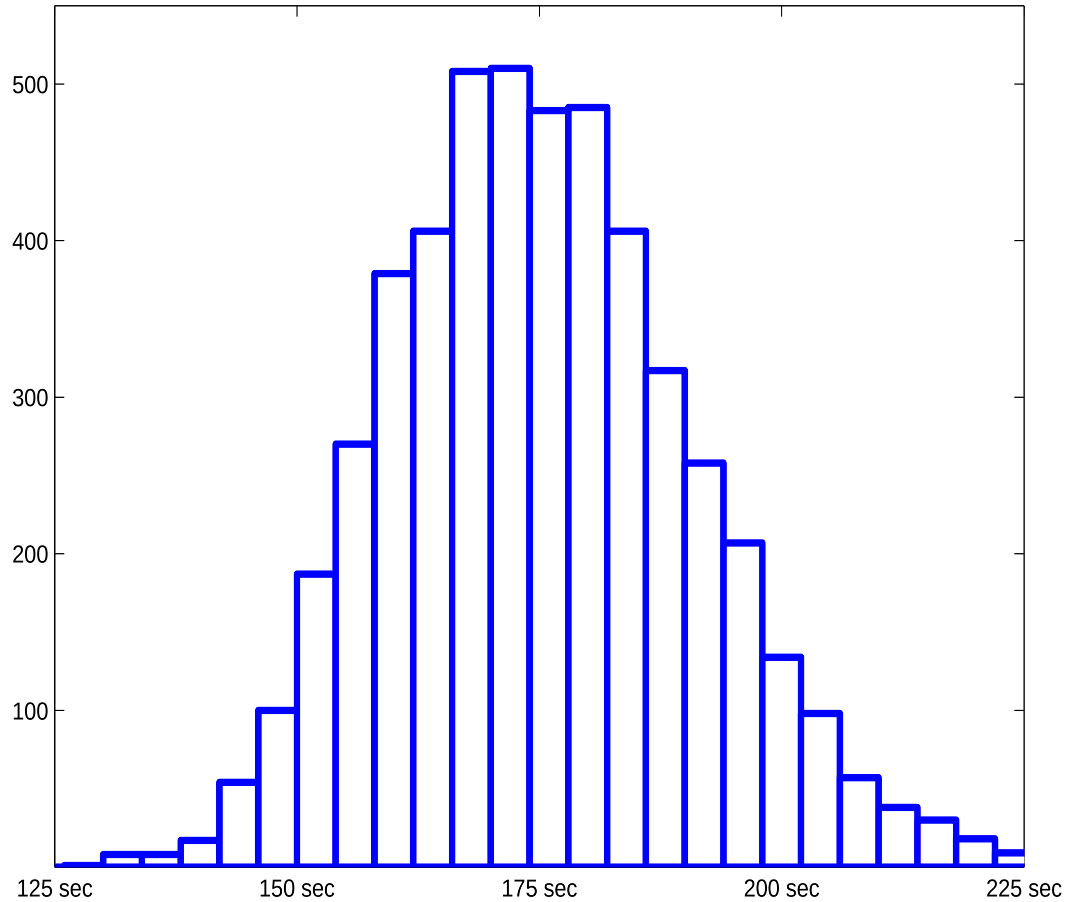
3 Determine optimal sampling frequency under BTS by minimizing the MSE over N

Impact of Measurement Error on Optimal Sampling Frequency

1 Simulate transaction data based on the CPP-MA(2) model ($\rho = 0.6$, $\lambda_{(0,1)} = 5000$, $\sigma_\nu/\sigma_\varepsilon = 1.1$, annualized return volatility of 25%)

2 Estimate model parameters as outlined above

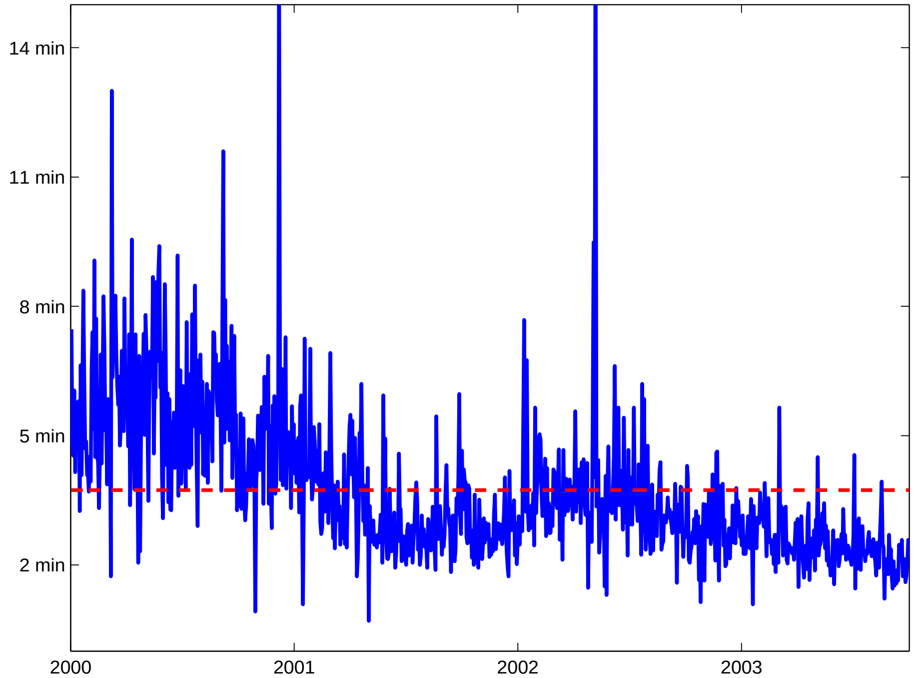
3 Determine optimal sampling frequency under BTS by minimizing the MSE over N



- No bias due to measurement error! (this is particularly important when analyzing illiquid securities)

Optimal Sampling Frequency and Sampling Scheme Efficiency

Optimal Sampling Frequency

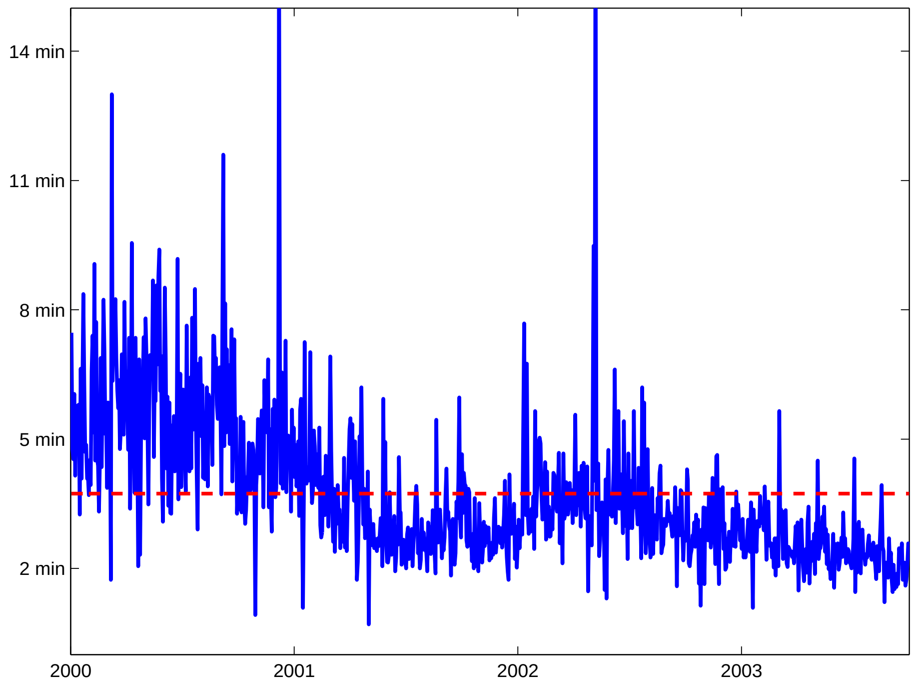


- Considerable day-to-day variation

- Downward trend

Optimal Sampling Frequency and Sampling Scheme Efficiency

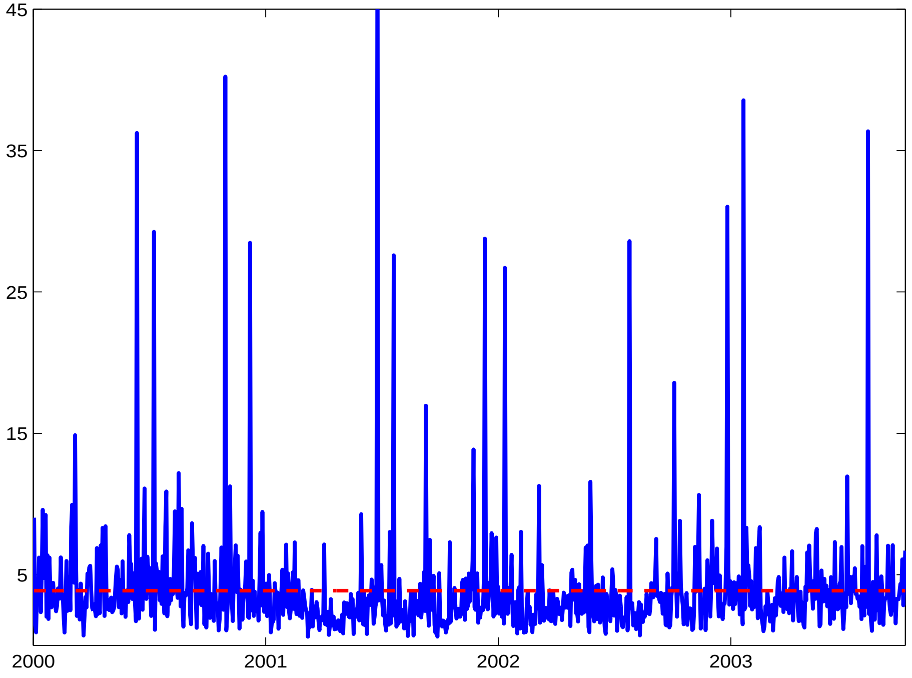
Optimal Sampling Frequency



- Considerable day-to-day variation

- Downward trend

CTS loss



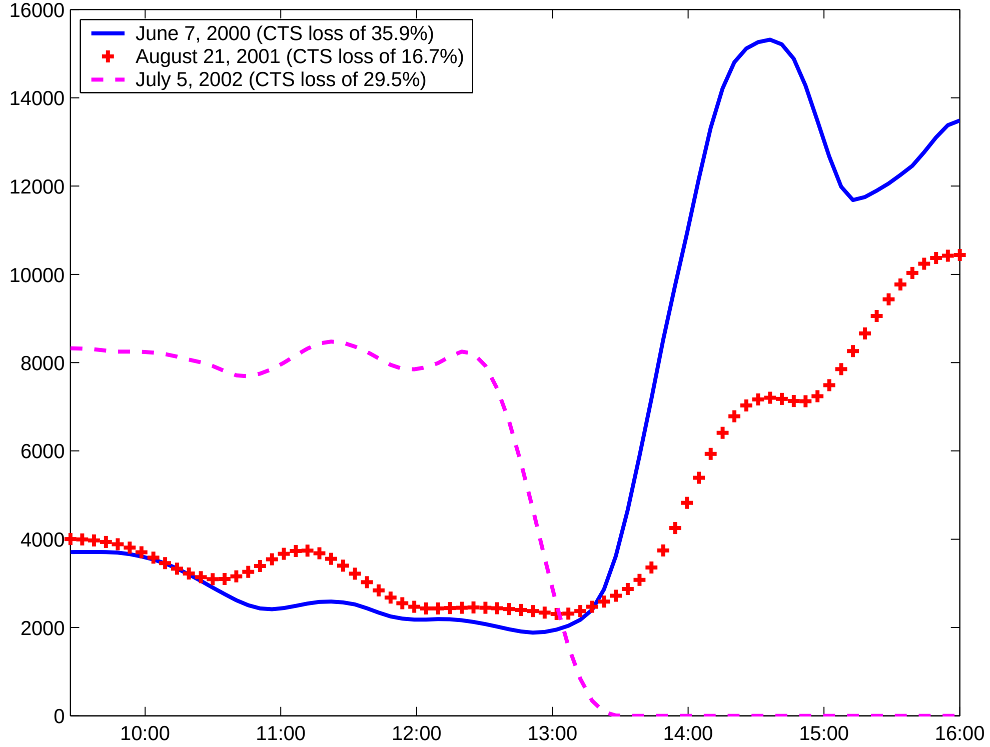
- Largest on days with highly irregular trading patterns, early market closures, or sudden moves in market activity

CTS loss on Irregular Trading Days

- On June 7, 2000 Dow Jones Business News headlined:

"*Wall Street Closes Higher, Paced By IBM Rebound On Goldman Sachs Comments*"

"…A late-day rally in IBM shares helped push stocks higher Wednesday…International Business Machines (IBM) jumped 8 3/8 to 120 3/4 after Goldman Sachs analyst Laura Conigliaro told CNBC that the computer maker should see revenue improvements in the second half of the year"

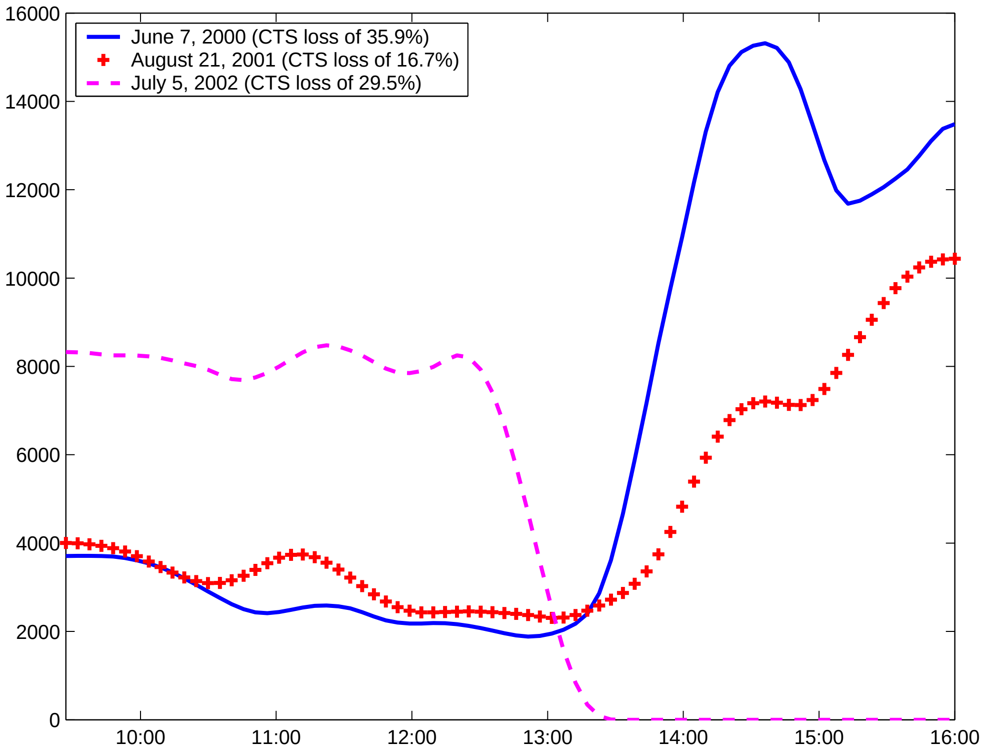


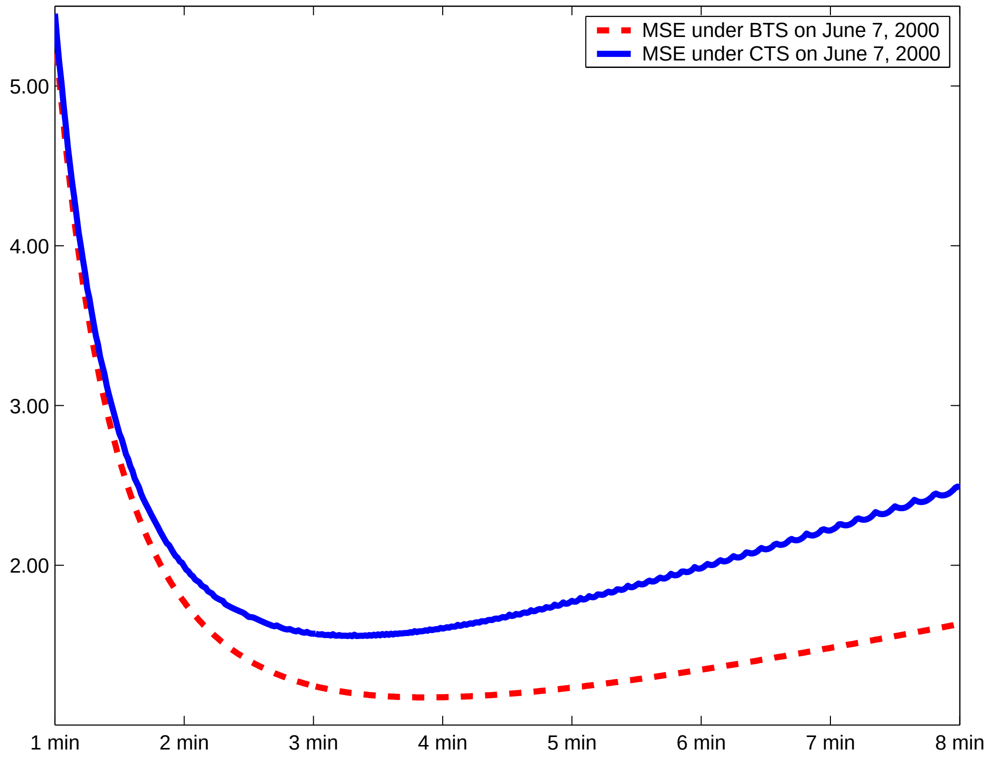
CTS loss on Irregular Trading Days

- On June 7, 2000 Dow Jones Business News headlined:

"*Wall Street Closes Higher, Paced By IBM Rebound On Goldman Sachs Comments*"

"…A late-day rally in IBM shares helped push stocks higher Wednesday…International Business Machines (IBM) jumped 8 3/8 to 120 3/4 after Goldman Sachs analyst Laura Conigliaro told CNBC that the computer maker should see revenue improvements in the second half of the year"





Summary

- So far I have analyzed RV under alternative sampling schemes

  Bias and MSE in closed form $\Rightarrow$ day-to-day optimal sampling frequency

Summary

- So far I have analyzed RV under alternative sampling schemes

  Bias and MSE in closed form $\Rightarrow$ day-to-day optimal sampling frequency

  BTS is superior to CTS along optimal sampling frequency (theory and practice)

Summary

- So far I have analyzed RV under alternative sampling schemes

  Bias and MSE in closed form $\Rightarrow$ day-to-day optimal sampling frequency

  BTS is superior to CTS along optimal sampling frequency (theory and practice)

  CTS loss particularly large on days with irregular trading patterns

Summary

- So far I have analyzed RV under alternative sampling schemes

  Bias and MSE in closed form $\Rightarrow$ day-to-day optimal sampling frequency

  BTS is superior to CTS along optimal sampling frequency (theory and practice)

  CTS loss particularly large on days with irregular trading patterns

  Downward trend in optimal sampling frequency

Summary

- So far I have analyzed RV under alternative sampling schemes

  Bias and MSE in closed form $\Rightarrow$ day-to-day optimal sampling frequency

  BTS is superior to CTS along optimal sampling frequency (theory and practice)

  CTS loss particularly large on days with irregular trading patterns

  Downward trend in optimal sampling frequency

  Close relation of optimal sampling frequency to noise ratio

Summary

- So far I have analyzed RV under alternative sampling schemes

  Bias and MSE in closed form $\Rightarrow$ day-to-day optimal sampling frequency

  BTS is superior to CTS along optimal sampling frequency (theory and practice)

  CTS loss particularly large on days with irregular trading patterns

  Downward trend in optimal sampling frequency

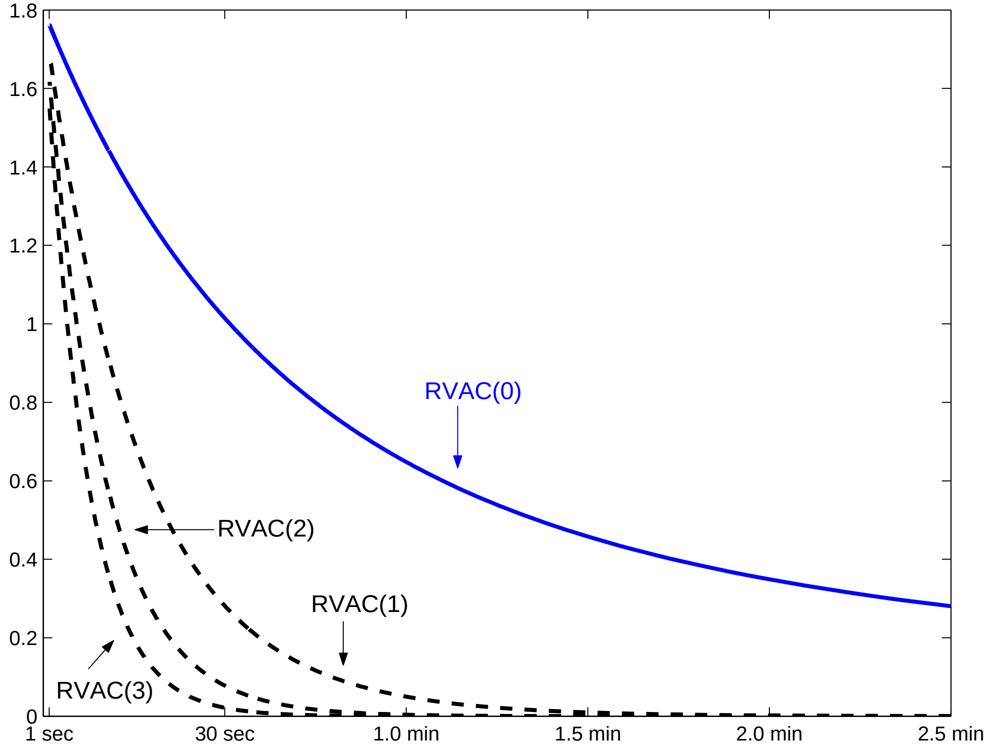
  Close relation of optimal sampling frequency to noise ratio

- Now turn to Newey-West type bias corrected realized variance (Hansen and Lunde 2004):

$$RVAC_{N,q} = \sum_{i=1}^{N} R(t_i|\tau_i)^2 + \sum_{i=1}^{N} R(t_i|\tau_i) \sum_{k=1}^{q} \left(R(t_{i-k}|\tau_{i-k}) + R(t_{i+k}|\tau_{i+k})\right)$$

Bias and MSE of RVAC(q) for CPP-MA(1)

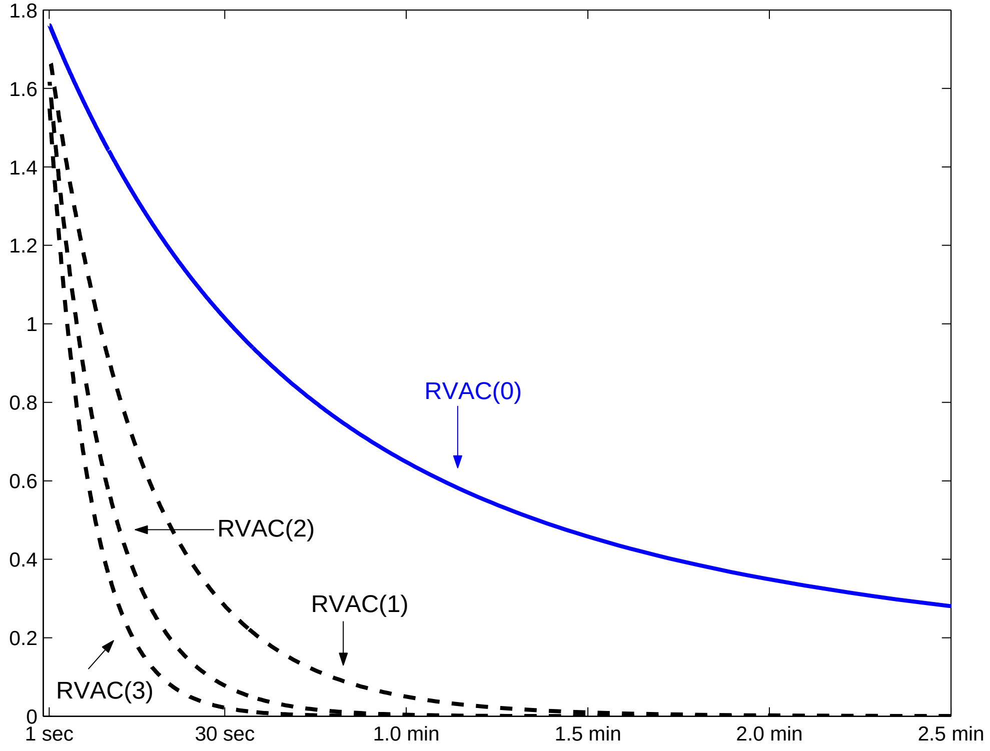
Bias of RVAC(q)



$$Bias = 2N\sigma_\nu^2 \left(1 - e^{-\lambda N}\right) e^{-q\lambda N}$$

Bias and MSE of RVAC(q) for CPP-MA(1)

Bias of RVAC(q)

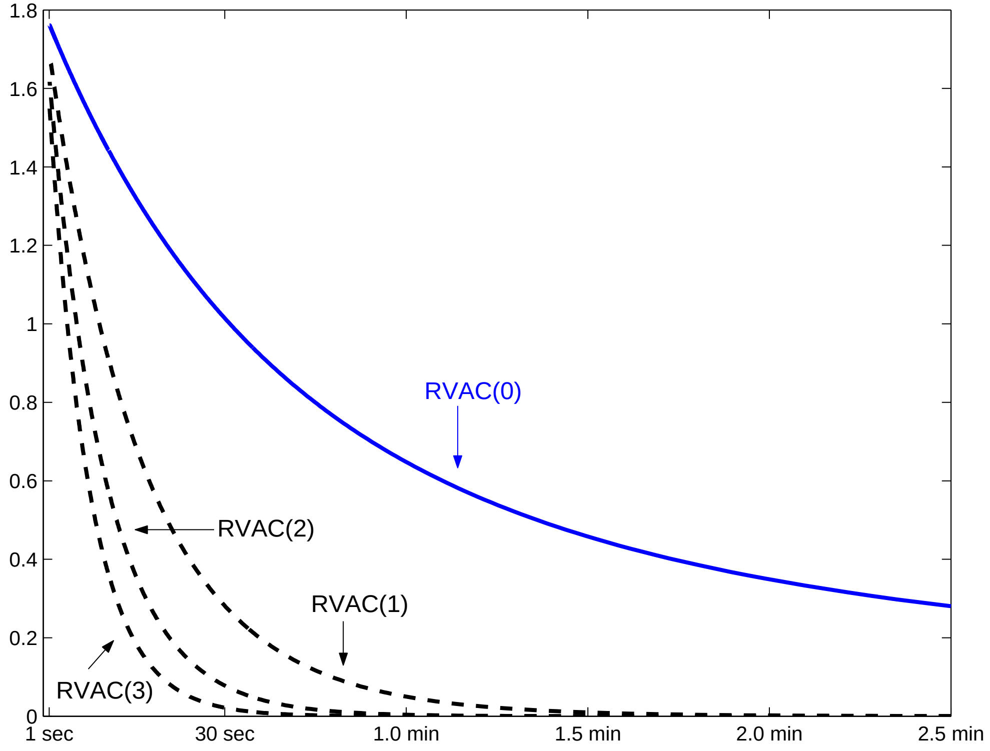


$$Bias = 2N\sigma_\nu^2 \left(1 - e^{-\lambda_N}\right) e^{-q\lambda_N}$$

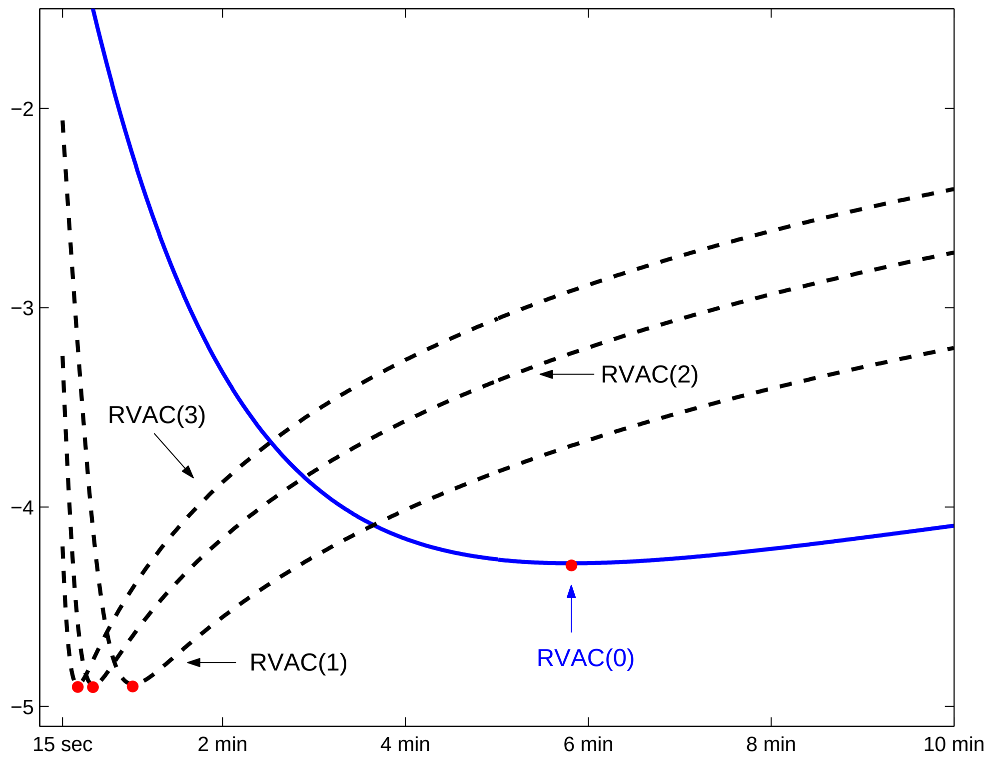
$$\lim_{N\to\infty}\left(\frac{Bias}{\lambda_{(0,1)}\sigma_\varepsilon^2}\right) = 2(\sigma_\nu/\sigma_\varepsilon)^2$$

Bias and MSE of RVAC(q) for CPP-MA(1)

Bias of RVAC(q)



MSE of RVAC(q)



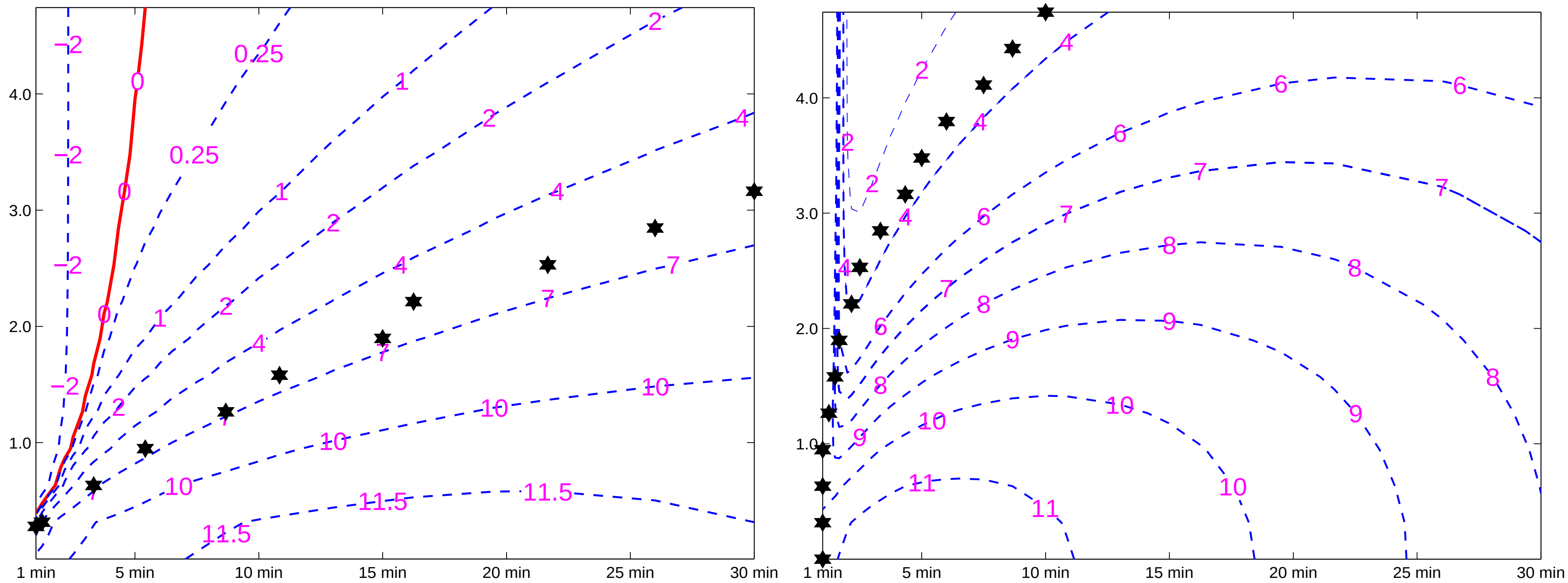
$$Bias = 2N\sigma_\nu^2 \left(1 - e^{-\lambda_N}\right) e^{-q\lambda_N}$$

$$\lim_{N\to\infty} \left(\frac{Bias}{\lambda_{(0,1)}\sigma_\varepsilon^2}\right) = 2(\sigma_\nu/\sigma_\varepsilon)^2$$

At given frequency, bias decreases but MSE increases with increase in q

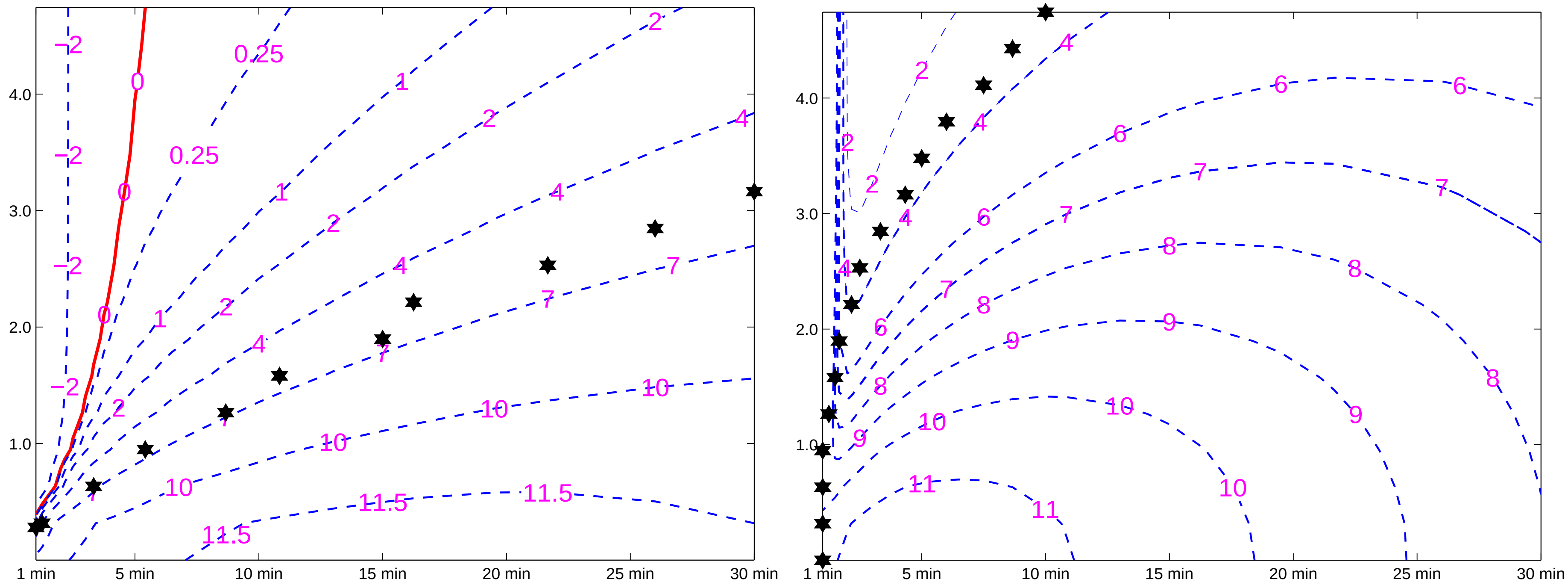
CTS loss of RVAC(1) for CPP-MA(1)

- Compare MSE under BTS and CTS for RV (left graph) and RVAC(1) (right graph)



CTS loss of RVAC(1) for CPP-MA(1)

- Compare MSE under BTS and CTS for RV (left graph) and RVAC(1) (right graph)



✓ BTS superior to CTS along optimal sampling frequency

✓ Optimal sampling frequency much higher for RVAC(1) than for RV

Empirical Results for IBM

	CPP-MA(1)			RVAC(0)				RVAC(1)			
IBM	$\sigma_\nu/\sigma_\varepsilon$	σ_ε	$\lambda_{(0,1)}$	Freq.	Bias	MSE	CTSloss	Freq.	Bias	MSE	CTSloss
Jan 03	1.66	1.27	8583	170	8.61	5.14	2.51	12	1.23	1.59	2.63
Feb 03											
Mar 03											
Apr 03											
May 03											
Jun 03											
Jul 03											
Aug 03											
Jan 03 - Aug 03											

- 166 days with 1,224,127 transactions. Relative bias and MSE in percentage points.

Empirical Results for IBM

	CPP-MA(1)			RVAC(0)				RVAC(1)			
IBM	$\sigma_\nu/\sigma_\varepsilon$	σ_ε	$\lambda_{(0,1)}$	Freq.	Bias	MSE	CTSloss	Freq.	Bias	MSE	CTSloss
Jan 03	1.66	1.27	8583	170	8.61	5.14	2.51	12	1.23	1.59	2.63
Feb 03	1.51	1.38	7698	162	8.35	5.57	1.99	13	1.41	1.76	2.71
Mar 03											
Apr 03											
May 03											
Jun 03											
Jul 03											
Aug 03											
Jan 03 - Aug 03											

- 166 days with 1,224,127 transactions. Relative bias and MSE in percentage points.

Empirical Results for IBM

IBM	CPP-MA(1) $\sigma_\nu/\sigma_\varepsilon$	σ_ε	$\lambda_{(0,1)}$	RVAC(0) Freq.	Bias	MSE	CTSloss	RVAC(1) Freq.	Bias	MSE	CTSloss
Jan 03	1.66	1.27	8583	170	8.61	5.14	2.51	12	1.23	1.59	2.63
Feb 03	1.51	1.38	7698	162	8.35	5.57	1.99	13	1.41	1.76	2.71
Mar 03	1.46	1.42	8408	146	7.95	7.41	2.16	11	1.42	2.27	2.85
Apr 03											
May 03											
Jun 03											
Jul 03											
Aug 03											
Jan 03 - Aug 03											

- 166 days with 1,224,127 transactions. Relative bias and MSE in percentage points.

Empirical Results for IBM

IBM	CPP-MA(1) $\sigma_\nu/\sigma_\varepsilon$	σ_ε	$\lambda_{(0,1)}$	RVAC(0) Freq.	Bias	MSE	CTSloss	RVAC(1) Freq.	Bias	MSE	CTSloss
Jan 03	1.66	1.27	8583	170	8.61	5.14	2.51	12	1.23	1.59	2.63
Feb 03	1.51	1.38	7698	162	8.35	5.57	1.99	13	1.41	1.76	2.71
Mar 03	1.46	1.42	8408	146	7.95	7.41	2.16	11	1.42	2.27	2.85
Apr 03	1.48	1.25	7772	156	8.22	3.43	3.09	12	1.43	1.09	4.14
May 03											
Jun 03											
Jul 03											
Aug 03											
Jan 03 - Aug 03											

- 166 days with 1,224,127 transactions. Relative bias and MSE in percentage points.

Empirical Results for IBM

IBM	CPP-MA(1) $\sigma_\nu/\sigma_\varepsilon$	σ_ε	$\lambda_{(0,1)}$	RVAC(0) Freq.	Bias	MSE	CTSloss	RVAC(1) Freq.	Bias	MSE	CTSloss
Jan 03	1.66	1.27	8583	170	8.61	5.14	2.51	12	1.23	1.59	2.63
Feb 03	1.51	1.38	7698	162	8.35	5.57	1.99	13	1.41	1.76	2.71
Mar 03	1.46	1.42	8408	146	7.95	7.41	2.16	11	1.42	2.27	2.85
Apr 03	1.48	1.25	7772	156	8.22	3.43	3.09	12	1.43	1.09	4.14
May 03	1.35	1.14	7391	143	7.90	2.09	2.30	12	1.59	0.67	3.28
Jun 03											
Jul 03											
Aug 03											
Jan 03 - Aug 03											

- 166 days with 1,224,127 transactions. Relative bias and MSE in percentage points.

Empirical Results for IBM

IBM	CPP-MA(1) $\sigma_\nu/\sigma_\varepsilon$	σ_ε	$\lambda_{(0,1)}$	RVAC(0) Freq.	Bias	MSE	CTSloss	RVAC(1) Freq.	Bias	MSE	CTSloss
Jan 03	1.66	1.27	8583	170	8.61	5.14	2.51	12	1.23	1.59	2.63
Feb 03	1.51	1.38	7698	162	8.35	5.57	1.99	13	1.41	1.76	2.71
Mar 03	1.46	1.42	8408	146	7.95	7.41	2.16	11	1.42	2.27	2.85
Apr 03	1.48	1.25	7772	156	8.22	3.43	3.09	12	1.43	1.09	4.14
May 03	1.35	1.14	7391	143	7.90	2.09	2.30	12	1.59	0.67	3.28
Jun 03	1.37	1.25	7053	150	8.05	3.08	3.24	13	1.60	1.02	4.15
Jul 03											
Aug 03											
Jan 03 - Aug 03											

- 166 days with 1,224,127 transactions. Relative bias and MSE in percentage points.

Empirical Results for IBM

IBM	CPP-MA(1)			RVAC(0)				RVAC(1)			
	$\sigma_\nu/\sigma_\varepsilon$	σ_ε	$\lambda_{(0,1)}$	Freq.	Bias	MSE	CTSloss	Freq.	Bias	MSE	CTSloss
Jan 03	1.66	1.27	8583	170	8.61	5.14	2.51	12	1.23	1.59	2.63
Feb 03	1.51	1.38	7698	162	8.35	5.57	1.99	13	1.41	1.76	2.71
Mar 03	1.46	1.42	8408	146	7.95	7.41	2.16	11	1.42	2.27	2.85
Apr 03	1.48	1.25	7772	156	8.22	3.43	3.09	12	1.43	1.09	4.14
May 03	1.35	1.14	7391	143	7.90	2.09	2.30	12	1.59	0.67	3.28
Jun 03	1.37	1.25	7053	150	8.05	3.08	3.24	13	1.60	1.02	4.15
Jul 03	1.22	1.34	6203	140	7.78	2.67	2.19	14	1.79	0.96	3.55
Aug 03											
Jan 03 - Aug 03											

- 166 days with 1,224,127 transactions. Relative bias and MSE in percentage points.

Empirical Results for IBM

IBM	CPP-MA(1) $\sigma_\nu/\sigma_\varepsilon$	σ_ε	$\lambda_{(0,1)}$	RVAC(0) Freq.	Bias	MSE	CTSloss	RVAC(1) Freq.	Bias	MSE	CTSloss
Jan 03	1.66	1.27	8583	170	8.61	5.14	2.51	12	1.23	1.59	2.63
Feb 03	1.51	1.38	7698	162	8.35	5.57	1.99	13	1.41	1.76	2.71
Mar 03	1.46	1.42	8408	146	7.95	7.41	2.16	11	1.42	2.27	2.85
Apr 03	1.48	1.25	7772	156	8.22	3.43	3.09	12	1.43	1.09	4.14
May 03	1.35	1.14	7391	143	7.90	2.09	2.30	12	1.59	0.67	3.28
Jun 03	1.37	1.25	7053	150	8.05	3.08	3.24	13	1.60	1.02	4.15
Jul 03	1.22	1.34	6203	140	7.78	2.67	2.19	14	1.79	0.96	3.55
Aug 03	1.05	1.15	5907	119	7.20	1.25	2.66	14	1.93	0.44	4.71
Jan 03 - Aug 03	**1.39**	**1.28**	**7377**	**148**	**8.01**	**3.83**	**2.52**	**13**	**1.55**	**1.22**	**3.50**

- 166 days with 1,224,127 transactions. Relative bias and MSE in percentage points.
- First order correction $\Rightarrow$ bias $\downarrow$, MSE $\downarrow$, optimal sampling frequency $\uparrow$
- BTS superior to CTS for each month in sample

Empirical Results for S&P500 Spiders

SPY	CPP-MA(1)			RVAC(0)				RVAC(1)			
	$\sigma_\nu/\sigma_\varepsilon$	σ_ε	$\lambda_{(0,1)}$	Freq.	Bias	MSE	CTSloss	Freq.	Bias	MSE	CTSloss
Jan 03	2.24	0.74	19666	147	8.01	2.94	5.72	7	0.41	0.76	3.64
Feb 03	2.23	0.78	23454	130	7.51	4.47	5.64	6	0.38	1.08	3.67
Mar 03	2.15	0.83	27747	112	6.98	9.13	5.24	5	0.41	2.00	3.56
Apr 03	2.13	0.70	24087	120	7.24	2.51	6.95	6	0.47	0.61	4.12
May 03	2.18	0.60	22819	128	7.48	1.37	6.56	6	0.43	0.34	4.37
Jun 03	1.99	0.59	24467	108	6.87	1.21	6.59	5	0.57	0.29	5.20
Jul 03	1.72	0.75	28043	82	5.95	8.29	4.52	4	0.69	2.76	3.76
Aug 03	1.78	0.55	24736	95	6.42	1.03	4.77	5	0.72	0.23	3.99
Jan 03 - Aug 03	**2.05**	**0.69**	**24389**	**115**	**7.05**	**3.86**	**5.75**	**6**	**0.51**	**1.01**	**4.04**

- 166 days with 4,048,665 transactions. 25K transaction per day!
- Higher noise ratio than for IBM but downward trend (market efficiency improved?)
- Bias correction leads to 60%-80% reduction in MSE!

The End…

- Flexible and easy-to-implement framework for studying properties of RV and bias corrected RV under alternative sampling schemes

The End...

- Flexible and easy-to-implement framework for studying properties of RV and bias corrected RV under alternative sampling schemes

- Allows for straightforward analysis of optimal sampling frequency on a day-to-day basis

The End...

- Flexible and easy-to-implement framework for studying properties of RV and bias corrected RV under alternative sampling schemes

- Allows for straightforward analysis of optimal sampling frequency on a day-to-day basis

- BTS superior to commonly used CTS although gains in MSE are modest

The End...

- Flexible and easy-to-implement framework for studying properties of RV and bias corrected RV under alternative sampling schemes

- Allows for straightforward analysis of optimal sampling frequency on a day-to-day basis

- BTS superior to commonly used CTS although gains in MSE are modest

- Substantial gains associated with bias correction

References

Andersen, T. G., and T. Bollerslev, 1998, “Answering the Skeptics: Yes, Standard Volatility Models Do Provide Accurate Forecasts,” *International Economic Review*, 39(4), 885–905.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and H. Ebens, 2001, “The Distribution of Stock Return Volatility,” *Journal of Financial Economics*, 61, 43–76.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys, 2000, “Great Realizations,” *Risk*, pp. 105–108.

——— , 2003, “Modeling and Forecasting Realized Volatility,” *Econometrica*, 71 (2), 579–625.

Andreou, E., and E. Ghysels, 2001, “When Does Microstructure Noise Affect the Estimation of Asset Pricing Models?,” Manuscript University of Manchester and North Carolina.

Ané, T., and H. Geman, 2000, “Order Flow, Transaction Clock, and Normality of Asset Returns,” *Journal of Finance*, 55(5), 2259–2284.

Bai, X., J. R. Russell, and G. C. Tiao, 2001, “Beyond Merton’s Utopia (I): Effects of Non-Normality and Dependence on the Precision of Variance Estimates using High Frequency Data,” Manuscript University of Chicago.

Bandi, F. M., and J. R. Russell, 2003, “Microstructure Noise, Realized Volatility, and Optimal Sampling,” manuscript GSB, The University of Chicago.

Barndorff-Nielsen, O. E., and N. Shephard, 2003, “Realised Power Variation and Stochastic Volatility,” *Bernoulli*, 9, 243–265.

Barndorff-Nielsen, O. E., and N. Shephard, 2004, *Continuous Time Approach to Financial Volatility*. Cambridge University Press, forthcoming.

Bowsher, C. G., 2002, “Modelling Security Market Events in Continuous Time: Intensity-Based, Multivariate Point Process Models,” Manuscript Nuffield College, University of Oxford.

Carr, P., and L. Wu, 2003, “Time-Changed Lévy Processes and Option Pricing,” forthcoming *Journal of Financial Economics*.

Corsi, F., G. Zumbach, U. A. Müller, and M. Dacorogna, 2001, “Consistent High-Precision Volatility from High-Frequency Data,” Olsen Group Working Paper.

Cowling, A., P. Hall, and M. J. Phillips, 1996, “Bootstrap Confidence Regions for the Intensity of a Poisson Point Process,” *JASA*, 91 (436), 1516–1524.

Diggle, P., and J. Marron, 1988, “Equivalence of Smoothing Parameter Selectors in Density and Intensity Estimation,” *JASA*, 83 (403), 793–800.

French, K. R., G. W. Schwert, and R. Stambaugh, 1987, “Expected Stock Returns and Volatility,” *Journal of Financial Economics*, 19, 3–29.

Geman, H., D. B. Madan, and M. Yor, 2001, “Time Changes for Levy Processes,” *Mathematical Finance*, 11 (1), 79–96.

Hansen, P. R., and A. Lunde, 2004, “An Unbiased Measure of Realized Variance,” manuscript Department of Economics, Brown University.

Hsieh, D., 1988, “The statistical properties of daily foreign exchange rates: 1974-1983,” *Journal of International Economics*, 24, 129–145.

Maheu, J. M., and T. H. McCurdy, 2003, “Modeling Foreign Exchange Rates with Jumps,” manuscript University of Toronto.

——— , 2004, “News Arrival, Jump Dynamics and Volatility Components for Individual Stock Returns,” *Journal of Finance*, 59 (2).

Meddahi, N., 2002, “A Theorical Comparison Between Integrated and Realized Volatility,” *Journal of Applied Econometrics*, 17, 479–508.

Merton, R. C., 1980, “On Estimating the Expected Return on the Market; An Exploratory Investigation,” *Journal of Financial Economics*, 8, 323–361.

Mürmann, A., 2001, “Pricing Catastrophe Insurance Derivatives,” Manuscript Insurance and Risk Management Department, The Wharton School.

Oomen, R. C., 2002a, “Modeling Realized Variance when Returns are Serially Correlated,” Manuscript Warwick Business School.

——— , 2002b, “Statistical Models for High Frequency Security Prices,” Manuscript Warwick Business School.

Piazzesi, M., 2004, “Bond Yields and the Federal Reserve,” *forthcoming Journal of Political Economy*.

Press, S. J., 1967, “A Compound Events Model for Security Prices,” *Journal of Business*, 40(3), 317–335.

Rogers, L., and O. Zane, 1998, “Designing and Estimating Models of High-Frequency Data,” Manuscript University of Bath.

Rydberg, T. H., and N. Shephard, 2003, “Dynamics of Trade-by-Trade Price Movements: Decomposition and Models,” *Journal of Financial Econometrics*, 1 (1), 2–25.

Zhang, L., P. A. Mykland, and Y. Ait-Sahalia, 2003, “A Tale of Two Time Scales: Determining Integrated Volatility with Noisy High Frequency Data,” manuscript Princeton University.