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 \textit{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: \textit{3 minutes}, decay over time, noise ratio
IBM sampling scheme: BTS always outperforms CTS (!!)

[2] Properties of \textit{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

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

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
Motivation & Related Literature


- Microstructure induced serial correlation renders realized variance biased.

  \[
  \text{bias } \downarrow \text{ by sampling less frequent } \Leftrightarrow \text{variance } \downarrow \text{ by sampling more frequent}
  \]
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

- 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

- Microstructure induced **serial correlation** renders realized variance **biased**
  
  bias $\downarrow$ by sampling *less* frequent $\iff$ 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.

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

where $\varepsilon_j \sim \text{iid } \mathcal{N}(\mu_\varepsilon, \sigma_\varepsilon^2)$, $\nu_j \sim \text{iid } \mathcal{N}(\mu_\nu, \sigma_\nu^2)$, 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.

$$P(t) = P(0) + \sum_{j=1}^{M(t)} (\varepsilon_j + \eta_j)$$

where $\eta_j = \rho_0 \nu_j + \rho_1 \nu_{j-1} + \ldots + \rho_s \nu_{j-s}$

where $\varepsilon_j \sim \text{iid } \mathcal{N}(\mu_\varepsilon, \sigma_\varepsilon^2)$, $\nu_j \sim \text{iid } \mathcal{N}(\mu_\nu, \sigma_\nu^2)$, 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:

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.

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

"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.

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

“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:

\[
e^{-\lambda_4-\lambda_3-\lambda_2-\lambda_1} + e^{-\lambda_2-\lambda_3}\gamma (1) \Lambda_1 \left( e^{c_1\lambda_1} \right) \gamma (4) \Lambda_1 \left( e^{c_4\lambda_4} \right) \exp \left\{ \xi_1\xi_4\sigma^2_{\nu} - \lambda_{1,2} - \lambda_{2,3} - \lambda_{3,4} \right\}
\]

\[
+ e^{-\lambda_2}\gamma (1) \Lambda_1 \left( e^{c_1\lambda_1} \right) \gamma (3) \Lambda_1 \left( e^{c_3\lambda_3} \right) \exp \left\{ \xi_1\xi_3\sigma^2_{\nu} - \lambda_{1,2} - \lambda_{2,3} \right\} + e^{-\lambda_4-\lambda_3-\lambda_2}\gamma (1) \Lambda_1 \left( e^{c_1\lambda_1} \right)
\]

\[
+ e^{-\lambda_1-\lambda_3}\gamma (2) \Lambda_1 \left( e^{c_2\lambda_2} \right) \gamma (4) \Lambda_1 \left( e^{c_4\lambda_4} \right) \exp \left\{ \xi_2\xi_4\sigma^2_{\nu} - \lambda_{2,3} - \lambda_{3,4} \right\} + e^{-\lambda_4-\lambda_3-\lambda_1}\gamma (2) \Lambda_1 \left( e^{c_2\lambda_2} \right)
\]

\[
+ e^{-\lambda_3}\gamma (1) \Lambda_1 \left( e^{c_1\lambda_1} \right) \gamma (2) \Lambda_1 \left( e^{c_2\lambda_2} \right) \gamma (4) \Lambda_1 \left( e^{c_4\lambda_4} \right) \Psi_{12}\Psi_{24} + e^{-\lambda_1-\lambda_2}\gamma (3) \Lambda_1 \left( e^{c_3\lambda_3} \right) \gamma (4) \Lambda_1 \left( e^{c_4\lambda_4} \right) \Psi_{34}
\]

\[
+ e^{-\lambda_2}\gamma (1) \Lambda_1 \left( e^{c_1\lambda_1} \right) \gamma (3) \Lambda_1 \left( e^{c_3\lambda_3} \right) \gamma (4) \Lambda_1 \left( e^{c_4\lambda_4} \right) \Psi_{13}\Psi_{34} + e^{-\lambda_3-\lambda_4}\gamma (1) \Lambda_1 \left( e^{c_1\lambda_1} \right) \Lambda_1 \left( e^{c_2\lambda_2} \right) \Psi_{12}\Psi_{12}
\]

\[
+ e^{-\lambda_1}\gamma (2) \Lambda_1 \left( e^{c_2\lambda_2} \right) \gamma (3) \Lambda_1 \left( e^{c_3\lambda_3} \right) \gamma (4) \Lambda_1 \left( e^{c_4\lambda_4} \right) \Psi_{23}\Psi_{34} + e^{-\lambda_1-\lambda_4}\gamma (2) \Lambda_1 \left( e^{c_2\lambda_2} \right) \gamma (3) \Lambda_1 \left( e^{c_3\lambda_3} \right) \Psi_{23}
\]

\[
+ e^{-\lambda_4}\gamma (1) \Lambda_1 \left( e^{c_1\lambda_1} \right) \gamma (2) \Lambda_1 \left( e^{c_2\lambda_2} \right) \gamma (3) \Lambda_1 \left( e^{c_3\lambda_3} \right) \gamma (4) \Lambda_1 \left( e^{c_4\lambda_4} \right) \Psi_{12}\Psi_{34}\Psi_{34} + e^{-\lambda_3-\lambda_2-\lambda_1}\gamma (4) \Lambda_1 \left( e^{c_4\lambda_4} \right)
\]

where \( \Psi_{ij} = \left( \exp \{ \xi_i\xi_j\sigma^2_{\nu} - \lambda_{i,j} \} \right) + \Lambda_1 \left( \lambda_{i,j} \right) \), \( \gamma (i) = \exp \left( \xi_i^2 c_0 + (e^{ci} - 1) \lambda_i \right) \), 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 \( \{ t^g_0, \ldots, t^g_N \} \) over the interval \( [t_0, t_0 + T] \) such that \( t^g_0 = t_0 \), \( t^g_N = t_0 + T \), and \( t^g_i < t^g_{i+1} \).
Alternative Sampling Schemes

⋆ General Time Sampling: Under $GTS_N$, the price process is sampled at time points \{t^g_0, \ldots, t^g_N\} over the interval $[t_0, t_0 + T]$ such that $t^g_0 = t_0$, $t^g_N = t_0 + T$, and $t^g_i < t^g_{i+1}$.

⋆ 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. $t^p_i = t_0 + i\delta$ 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 $\{t^g_0, \ldots, t^g_N\}$ over the interval $[t_0, t_0 + T]$ such that $t^g_0 = t_0$, $t^g_N = t_0 + T$, and $t^g_i < t^g_{i+1}$.

★ 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. $t^p_i = t_0 + i\delta$ 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. $t^b_i$ for $i = \{0, \ldots, NT\}$ such that $t^b_0 = t_0$, $t^b_N = t_0 + T$ and

$$\int_{t^b_i}^{t^b_{i+1}} \lambda(u)du = \frac{1}{NT} \int_{t_0}^{t_0 + T} \lambda(u)du \equiv \lambda_N$$

✓ 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)

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

★ 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)

\[
MSE (GTS_N) = \sum_{i=1}^{N} \left( 3\sigma^4 \lambda_i^2 + 3\sigma^4 \lambda_i \right) + 2\sigma^4 \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \lambda_i \lambda_j - \lambda^2_{(0,1)} \sigma^4
\]

…which simplifies under BTS to:

\[
MSE (BTS_N) = 2N^{-1} \left( \lambda_{(0,1)} \sigma^2 \right)^2 + 3\sigma^2 \left( \lambda_{(0,1)} \sigma^2 \right)
\]

★ Notation: \( \lambda_i = \int_{t_{i-1}}^{t_i} \lambda(u) \, du \) 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)

\[ MSE(\text{GTS}_N) = \sum_{i=1}^{N} (3\sigma^4\lambda_i^2 + 3\sigma^4\lambda_i) + 2\sigma^4 \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \lambda_i \lambda_j - \lambda^2_{(0,1)}\sigma^4 \]

…which simplifies under BTS to:

\[ MSE(\text{BTS}_N) = 2N^{-1} (\lambda_{(0,1)}\sigma^2)^2 + 3\sigma^2 (\lambda_{(0,1)}\sigma^2) \]

• RV is inconsistent under pure jump process

• Consistency in “diffusion” limit where \( \lambda \to \infty \) while \( \lambda \sigma^2 \) 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:

\[
MSE(GTS_N) - MSE(BTS_N) = 3\sigma_\varepsilon^4 \sum_{i=1}^{N} (\lambda_i^2 - \lambda_N^2) + 2\sigma_\varepsilon^4 \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} (\lambda_i \lambda_j - \lambda_N^2) \\
= 2\sigma_\varepsilon^4 \sum_{i=1}^{N} \psi_i^2 > 0
\]

where \( \psi_i = \int_{t_i - \tau_i}^{t_i} \lambda(u) \, du - \lambda_N \)
Absence of Market Microstructure Noise

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

\[
MSE(GTS_N) - MSE(BTS_N) = 3\sigma_\varepsilon^4 \sum_{i=1}^{N} (\lambda_i^2 - \lambda_N^2) + 2\sigma_\varepsilon^4 \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} (\lambda_i \lambda_j - \lambda_N^2)
\]

\[
= 2\sigma_\varepsilon^4 \sum_{i=1}^{N} \vartheta_i^2 > 0
\]

where \(\vartheta_i = \int_{t_{i-1}}^{t_i} \lambda(u) du - \lambda_N\)

- 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:

\[
MSE (GTS_N) - MSE (BTS_N) = 3\sigma^4 \sum_{i=1}^{N} (\lambda_i^2 - \lambda_N^2) + 2\sigma^4 \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} (\lambda_i \lambda_j - \lambda_N^2)
\]

\[
= 2\sigma^4 \sum_{i=1}^{N} \vartheta_i^2 > 0
\]

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

- 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: $\vartheta$
Integrated Incremental Intensity under BTS and CTS: $\vartheta$
Presence of Market Microstructure Noise – The Bias …

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

\[
Bias (GTS_N) = E_{\lambda} \left[ \sum_{i=1}^{N} R(t_i|\tau_i)^2 \right] - \lambda_{(0,1)} \sigma^2_{\epsilon} = 2\sigma^2_{\nu} \sum_{i=1}^{N} (1 - e^{-\lambda_i}) > 0
\]
Presence of Market Microstructure Noise – The Bias …

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

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

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

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

- 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

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

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

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

\[
\text{Bias} (GTS_N) - \text{Bias} (BTS_N) = 2e^{-\lambda N} \sigma^2_\nu \sum_{i=1}^{N} (1 - e^{-\vartheta_i}) < 0
\]

• 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.

\[
\lim_{N \to \infty} \text{Bias} (GTS_N) = 2\sigma^2_\nu \lambda_{(0,1)}
\]
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

- $\sigma_\varepsilon$, $\sigma_\nu$, and $\rho$ estimated in “business time” using all transactions

\[
Cov(R_k, R_{k-1}) = -(1 - \rho)^2 \sigma_\nu^2
\]

\[
Cov(R_k, R_{k-2}) = -\rho \sigma_\nu^2
\]

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

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

- $\sigma_\varepsilon$, $\sigma_\nu$, and $\rho$ estimated in “business time” using all transactions

\[
\begin{align*}
Cov(R_k, R_{k-1}) &= -(1 - \rho)^2 \sigma_\nu^2 \\
Cov(R_k, R_{k-2}) &= -\rho \sigma_\nu^2 \\
Var(R_k) &= \sigma_\varepsilon^2 + (2 + 2\rho^2 - 2\rho) \sigma_\nu^2
\end{align*}
\]

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

- $\lambda(t)$ is estimated using non-parametric smoothing techniques (Cowling, Hall, and Phillips 1996) similar to density estimation
Empirical Analysis – Parameter Estimation

- $\sigma_\varepsilon$, $\sigma_\nu$, and $\rho$ estimated in “business time” using all transactions

\[
Cov(R_k, R_{k-1}) = -(1 - \rho)^2 \sigma_\nu^2
\]

\[
Cov(R_k, R_{k-2}) = -\rho \sigma_\nu^2
\]

\[
Var(R_k) = \sigma_\varepsilon^2 + (2 + 2\rho^2 - 2\rho) \sigma_\nu^2
\]

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

- $\lambda(t)$ 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

- $\sigma_{\varepsilon}, \sigma_\nu,$ and $\rho$ estimated in “business time” using all transactions

\[
Cov(R_k, R_{k-1}) = -(1 - \rho)^2 \sigma_\nu^2
\]
\[
Cov(R_k, R_{k-2}) = -\rho \sigma_\nu^2
\]
\[
Var(R_k) = \sigma_{\varepsilon}^2 + (2 + 2\rho^2 - 2\rho) \sigma_\nu^2
\]

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

- $\lambda(t)$ 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 ($\rho = 0.6$, $\lambda_{(0,1)} = 5000$, $\sigma_\nu/\sigma_\epsilon = 1.1$, 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

---

**ROEL OOMEN**

*Market Microstructure Noise & Properties of Realized Variance*
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 ⇒ 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 ⇒ 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):

  \[ RV \ AC_{N,q} = \sum_{i=1}^{N} R(t_i|\tau_i)^2 + \sum_{i=1}^{N} R(t_i|\tau_i) \sum_{k=1}^{q} (R(t_{i-k}|\tau_{i-k}) + R(t_{i+k}|\tau_{i+k})) \]
Bias and MSE of RVAC(q) for CPP-MA(1)

Bias of RVAC(q)

\[
\text{Bias} = 2N \sigma^2 \nu \left( 1 - e^{-\lambda N} \right) e^{-q\lambda N}
\]
Bias and MSE of RVAC(q) for CPP-MA(1)

Bias of RVAC(q)

\[
\text{Bias} = 2N\sigma^2_\nu \left(1 - e^{-\lambda N}\right) e^{-q\lambda N}
\]

\[
\lim_{N \to \infty} \left( \frac{\text{Bias}}{\lambda_{(0,1)}^2} \right) = 2\left(\frac{\sigma_\nu}{\sigma_\varepsilon}\right)^2
\]
Bias and MSE of RVAC(q) for CPP-MA(1)

Bias of RVAC(q)

\[ Bias = 2N\sigma^2_\nu \left(1 - e^{-\lambda N}\right) e^{-q\lambda N} \]
\[ \lim_{N \to \infty} \frac{Bias}{\lambda_{(0,1)}\sigma^2_\varepsilon} = 2\left(\sigma_\nu/\sigma_\varepsilon\right)^2 \]

MSE of RVAC(q)

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

<table>
<thead>
<tr>
<th>IBM</th>
<th>CPP-MA(1) $\sigma_\nu/\sigma_\epsilon$ $\sigma_\epsilon$ $\lambda_{(0,1)}$</th>
<th>RVAC(0) Freq. Bias MSE CTSloss</th>
<th>RVAC(1) Freq. Bias MSE CTSloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jan 03</td>
<td>1.66 1.27 8583</td>
<td>170 8.61 5.14 2.51</td>
<td>12 1.23 1.59 2.63</td>
</tr>
<tr>
<td>Feb 03</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mar 03</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Apr 03</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>May 03</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jun 03</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jul 03</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aug 03</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jan 03 - Aug 03</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- 166 days with 1,224,127 transactions. Relative bias and MSE in percentage points.
### Empirical Results for IBM

<table>
<thead>
<tr>
<th>IBM</th>
<th>CPP-MA(1)</th>
<th>RVAC(0)</th>
<th>RVAC(1)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\sigma_\nu/\sigma_\epsilon$</td>
<td>$\sigma_\epsilon$</td>
<td>$\lambda_{(0,1)}$</td>
</tr>
<tr>
<td>Jan 03</td>
<td>1.66</td>
<td>1.27</td>
<td>8583</td>
</tr>
<tr>
<td>Feb 03</td>
<td>1.51</td>
<td>1.38</td>
<td>7698</td>
</tr>
<tr>
<td>Mar 03</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Apr 03</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>May 03</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jun 03</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jul 03</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aug 03</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jan 03 - Aug 03</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- 166 days with 1,224,127 transactions. Relative bias and MSE in percentage points.
### Empirical Results for IBM

<table>
<thead>
<tr>
<th>IBM</th>
<th>CPP-MA(1)</th>
<th>RVAC(0)</th>
<th>RVAC(1)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\sigma_\nu/\sigma_\epsilon$</td>
<td>$\sigma_\epsilon$</td>
<td>$\lambda_{(0,1)}$</td>
</tr>
<tr>
<td>Jan 03</td>
<td>1.66</td>
<td>1.27</td>
<td>8583</td>
</tr>
<tr>
<td>Feb 03</td>
<td>1.51</td>
<td>1.38</td>
<td>7698</td>
</tr>
<tr>
<td>Mar 03</td>
<td>1.46</td>
<td>1.42</td>
<td>8408</td>
</tr>
<tr>
<td>Apr 03</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>May 03</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jun 03</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jul 03</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aug 03</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jan 03 - Aug 03</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- 166 days with 1,224,127 transactions. Relative bias and MSE in percentage points.
## Empirical Results for IBM

<table>
<thead>
<tr>
<th>IBM</th>
<th>CPP-MA(1)</th>
<th>RVAC(0)</th>
<th>RVAC(1)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\sigma_\nu / \sigma_\varepsilon$</td>
<td>$\sigma_\varepsilon$</td>
<td>$\lambda_{(0,1)}$</td>
</tr>
<tr>
<td>Jan 03</td>
<td>1.66</td>
<td>1.27</td>
<td>8583</td>
</tr>
<tr>
<td>Feb 03</td>
<td>1.51</td>
<td>1.38</td>
<td>7698</td>
</tr>
<tr>
<td>Mar 03</td>
<td>1.46</td>
<td>1.42</td>
<td>8408</td>
</tr>
<tr>
<td>Apr 03</td>
<td>1.48</td>
<td>1.25</td>
<td>7772</td>
</tr>
<tr>
<td>May 03</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jun 03</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jul 03</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aug 03</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jan 03 - Aug 03</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th></th>
<th>CPP-MA(1)</th>
<th>RVAC(0)</th>
<th>RVAC(1)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\sigma_\nu / \sigma_\varepsilon$</td>
<td>$\sigma_\varepsilon$</td>
<td>$\lambda_{(0,1)}$</td>
</tr>
<tr>
<td>Jan 03</td>
<td>1.66</td>
<td>1.27</td>
<td>8583</td>
</tr>
<tr>
<td>Feb 03</td>
<td>1.51</td>
<td>1.38</td>
<td>7698</td>
</tr>
<tr>
<td>Mar 03</td>
<td>1.46</td>
<td>1.42</td>
<td>8408</td>
</tr>
<tr>
<td>Apr 03</td>
<td>1.48</td>
<td>1.25</td>
<td>7772</td>
</tr>
<tr>
<td>May 03</td>
<td>1.35</td>
<td>1.14</td>
<td>7391</td>
</tr>
<tr>
<td>Jun 03</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jul 03</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aug 03</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jan 03 - Aug 03</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- 166 days with 1,224,127 transactions. Relative bias and MSE in percentage points.
## Empirical Results for IBM

<table>
<thead>
<tr>
<th>IBM</th>
<th>CPP-MA(1)</th>
<th>RVAC(0)</th>
<th>RVAC(1)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\sigma_\nu / \sigma_\varepsilon$</td>
<td>$\sigma_\varepsilon$</td>
<td>$\lambda_{(0,1)}$</td>
</tr>
<tr>
<td>Jan 03</td>
<td>1.66</td>
<td>1.27</td>
<td>8583</td>
</tr>
<tr>
<td>Feb 03</td>
<td>1.51</td>
<td>1.38</td>
<td>7698</td>
</tr>
<tr>
<td>Mar 03</td>
<td>1.46</td>
<td>1.42</td>
<td>8408</td>
</tr>
<tr>
<td>Apr 03</td>
<td>1.48</td>
<td>1.25</td>
<td>7772</td>
</tr>
<tr>
<td>May 03</td>
<td>1.35</td>
<td>1.14</td>
<td>7391</td>
</tr>
<tr>
<td>Jun 03</td>
<td>1.37</td>
<td>1.25</td>
<td>7053</td>
</tr>
<tr>
<td>Jul 03</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aug 03</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jan 03 - Aug 03</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- 166 days with 1,224,127 transactions. Relative bias and MSE in percentage points.
## Empirical Results for IBM

<table>
<thead>
<tr>
<th>IBM</th>
<th>CPP-MA(1)</th>
<th>RVAC(0)</th>
<th>RVAC(1)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\sigma_\nu / \sigma_\epsilon$</td>
<td>$\sigma_\epsilon$</td>
<td>$\lambda_{(0,1)}$</td>
</tr>
<tr>
<td>Jan 03</td>
<td>1.66</td>
<td>1.27</td>
<td>8583</td>
</tr>
<tr>
<td>Feb 03</td>
<td>1.51</td>
<td>1.38</td>
<td>7698</td>
</tr>
<tr>
<td>Mar 03</td>
<td>1.46</td>
<td>1.42</td>
<td>8408</td>
</tr>
<tr>
<td>Apr 03</td>
<td>1.48</td>
<td>1.25</td>
<td>7772</td>
</tr>
<tr>
<td>May 03</td>
<td>1.35</td>
<td>1.14</td>
<td>7391</td>
</tr>
<tr>
<td>Jun 03</td>
<td>1.37</td>
<td>1.25</td>
<td>7053</td>
</tr>
<tr>
<td>Jul 03</td>
<td>1.22</td>
<td>1.34</td>
<td>6203</td>
</tr>
<tr>
<td>Aug 03</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jan 03 - Aug 03</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>IBM</th>
<th>CPP-MA(1)</th>
<th>RVAC(0)</th>
<th>RVAC(1)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\sigma_{\nu}/\sigma_{\varepsilon}$</td>
<td>$\sigma_{\varepsilon}$</td>
<td>$\lambda_{(0,1)}$</td>
</tr>
<tr>
<td>Jan 03</td>
<td>1.66</td>
<td>1.27</td>
<td>8583</td>
</tr>
<tr>
<td>Feb 03</td>
<td>1.51</td>
<td>1.38</td>
<td>7698</td>
</tr>
<tr>
<td>Mar 03</td>
<td>1.46</td>
<td>1.42</td>
<td>8408</td>
</tr>
<tr>
<td>Apr 03</td>
<td>1.48</td>
<td>1.25</td>
<td>7772</td>
</tr>
<tr>
<td>May 03</td>
<td>1.35</td>
<td>1.14</td>
<td>7391</td>
</tr>
<tr>
<td>Jun 03</td>
<td>1.37</td>
<td>1.25</td>
<td>7053</td>
</tr>
<tr>
<td>Jul 03</td>
<td>1.22</td>
<td>1.34</td>
<td>6203</td>
</tr>
<tr>
<td>Aug 03</td>
<td>1.05</td>
<td>1.15</td>
<td>5907</td>
</tr>
<tr>
<td>Jan 03 - Aug 03</td>
<td>1.39</td>
<td>1.28</td>
<td>7377</td>
</tr>
</tbody>
</table>

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

- First order correction $\Rightarrow$ bias ↓, MSE ↓, optimal sampling frequency ↑

- BTS superior to CTS for each month in sample
Empirical Results for S&P500 Spiders

<table>
<thead>
<tr>
<th>SPY</th>
<th>CPP-MA(1)</th>
<th>RVAC(0)</th>
<th>RVAC(1)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\sigma_\nu/\sigma_\varepsilon$</td>
<td>$\sigma_\varepsilon\lambda_{(0,1)}$</td>
<td>Freq.</td>
</tr>
<tr>
<td>Jan 03</td>
<td>2.24</td>
<td>0.74</td>
<td>19666</td>
</tr>
<tr>
<td>Feb 03</td>
<td>2.23</td>
<td>0.78</td>
<td>23454</td>
</tr>
<tr>
<td>Mar 03</td>
<td>2.15</td>
<td>0.83</td>
<td>27747</td>
</tr>
<tr>
<td>Apr 03</td>
<td>2.13</td>
<td>0.70</td>
<td>24087</td>
</tr>
<tr>
<td>May 03</td>
<td>2.18</td>
<td>0.60</td>
<td>22819</td>
</tr>
<tr>
<td>Jun 03</td>
<td>1.99</td>
<td>0.59</td>
<td>24467</td>
</tr>
<tr>
<td>Jul 03</td>
<td>1.72</td>
<td>0.75</td>
<td>28043</td>
</tr>
<tr>
<td>Aug 03</td>
<td>1.78</td>
<td>0.55</td>
<td>24736</td>
</tr>
<tr>
<td>Jan 03 - Aug 03</td>
<td>2.05</td>
<td>0.69</td>
<td>24389</td>
</tr>
</tbody>
</table>

- 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


