QMF 2006

Sydney, 13-16 December 2006

Properties and Optimal Sampling of the Fourier Integrated Volatility Estimator under Microstructure Noise

Maria Elvira Mancino
University of Firenze - Italy
mariaelvira.mancino@dmd.unifi.it

Sanfelici Simona
University of Parma - Italy
simona.sanfelici@unipr.it

- The recent availability of high frequency financial data has motivated a growing literature devoted to the model free measurement of volatility.

- The recent availability of high frequency financial data has motivated a growing literature devoted to the model free measurement of volatility.
- *Instantaneous volatility* for a diffusion process p

$$\sigma^2(t) := \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} E[(p(t+\varepsilon) - p(t))^2 | \mathcal{F}_t].$$

Integrated variance or *volatility*

$$\int_0^t \sigma^2(s)\, ds.$$

- The recent availability of high frequency financial data has motivated a growing literature devoted to the model free measurement of volatility.
- *Instantaneous volatility* for a diffusion process p

$$\sigma^2(t) := \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} E[(p(t+\varepsilon) - p(t))^2 | \mathcal{F}_t].$$

Integrated variance or *volatility*

$$\int_0^t \sigma^2(s)\, ds.$$

- Market microstructure frictions $\Rightarrow$ observed asset prices deviate from their efficient values.

Introduction

- The recent availability of high frequency financial data has motivated a growing literature devoted to the model free measurement of volatility.
- *Instantaneous volatility* for a diffusion process p

$$\sigma^2(t) := \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} E[(p(t+\varepsilon) - p(t))^2 | \mathcal{F}_t].$$

Integrated variance or *volatility*

$$\int_0^t \sigma^2(s)\, ds.$$

- Market microstructure frictions $\Rightarrow$ observed asset prices deviate from their efficient values.
- Nonparametric integrated volatility estimators (*Realized volatility*, *Fourier estimator*, *Wavelet estimator* and *bias correction methods*).

References:

- *General works on realized volatility*: [French-Schwert-Stambaugh, JFE ’87]; [Andersen-Bollerslev-Diebold, ’04]; [Barndorff Nielsen-Shepard, JAE ’02].
- *Microstructure effects*: [Andersen-Bollerslev-Diebold-Labys, ’99, ’00]; [Bandi-Russel, ’05].
- *Filtering techniques and bias corrections*: [Hansen-Lunde, JFE ’05]; [Owens-Steigerwald, ’05].
- *Other estimators*: [Malliavin-Mancino, FS ’02]; [Hog-Lunde, ’03]; [Zhou, JBES ’96], [Genon Catalot-Larédo-Picard, SJS ’92]; [Barucci-Renó, JIFMIM ’02, EL ’02]; [Nielsen-Frederiksen, ’06].

Outline

- The aim of this paper is to fully understand the properties of the Fourier volatility estimator under microstructure noise.

- The aim of this paper is to fully understand the properties of the Fourier volatility estimator under microstructure noise.

- Simple price formation mechanism, where the logarithm of the *observed price* process is the sum of the *log-price process in equilibrium* and a component of *microstructure noise*

Outline

- The aim of this paper is to fully understand the properties of the Fourier volatility estimator under microstructure noise.
- Simple price formation mechanism, where the logarithm of the *observed price* process is the sum of the *log-price process in equilibrium* and a component of *microstructure noise*
- ASYMPTOTIC and FINITE SAMPLE properties for the Fourier volatility estimator under the presence of an MA(1) market microstructure noise process.

Outline

- The aim of this paper is to fully understand the properties of the Fourier volatility estimator under microstructure noise.
- Simple price formation mechanism, where the logarithm of the *observed price* process is the sum of the *log-price process in equilibrium* and a component of *microstructure noise*
- ASYMPTOTIC and FINITE SAMPLE properties for the Fourier volatility estimator under the presence of an MA(1) market microstructure noise process.
- Easily implementable bias and MSE estimate $\Rightarrow$ optimal sampling at higher frequency. Lower bias and MSE.

The price model with microstructure effects

- Observed log-price: $\widetilde{p}(t) = p(t) + \eta(t)$.

The price model with microstructure effects

- Observed log-price: $\widetilde{p}(t) = p(t) + \eta(t)$.
- ASSUMPTIONS:

  A.I The *efficient log-price* $p(t) = \sigma(t)\,dW(t) + b(t)\,d$ is a continuous semi-martingale such that $E[\int_0^T \sigma^4(t)dt] < \infty \;\;,\;\; E[\int_0^T b^2(t)dt] < \infty$.

  A.II We normalize the time window to $[0, 2\pi]$ and suppose the process is observed at a discrete grid $\{0 = t_0 \leq t_1 \leq \cdots \leq t_n \leq 2\pi\}$.

  A.III The *random shocks* $\eta_j = \eta(t_j)$ are i.i.d. with mean zero and bounded fourth moment.

  A.IV The true return process $\delta_j(p) := p(t_j) - p(t_{j-1})$ is independent of η_j for any j.

The finite sample estimators

- $$\boxed{\hat{V}_n := \sum_{j=1}^{n} (p(t_j) - p(t_{j-1}))^2 = \sum_{j=1}^{n} (\delta_j(p))^2}$$

REALIZED VOLATILITY

The finite sample estimators

- $$\hat{V}_n := \sum_{j=1}^{n} (p(t_j) - p(t_{j-1}))^2 = \sum_{j=1}^{n} (\delta_j(p))^2$$

REALIZED VOLATILITY

- $$\hat{\sigma}^2_{n,N} := \frac{(2\pi)^2}{2N+1} \sum_{|s| \leq N} \mathcal{F}(dp_n)(s)\mathcal{F}(dp_n)(-s)$$

FOURIER ESTIMATOR

where $\mathcal{F}(dp_n)(s) = \frac{1}{2\pi} \sum_{j=1}^{n} \exp(-ist_j)\delta_j(p).$

The finite sample estimators

- $$\hat{V}_n := \sum_{j=1}^{n} (p(t_j) - p(t_{j-1}))^2 = \sum_{j=1}^{n} (\delta_j(p))^2$$

REALIZED VOLATILITY

- $$\hat{\sigma}^2_{n,N} := \frac{(2\pi)^2}{2N+1} \sum_{|s| \leq N} \mathcal{F}(dp_n)(s)\mathcal{F}(dp_n)(-s)$$

FOURIER ESTIMATOR

where $\mathcal{F}(dp_n)(s) = \frac{1}{2\pi} \sum_{j=1}^{n} \exp(-ist_j)\delta_j(p).$

- Both consistent as $n, N \to \infty$ *in the absence of microstructure noise*.

Asymptotic results

- Asymptotic results for the *equilibrium price*:

$$\lim_{n,N\to\infty} \frac{(2\pi)^2}{2N+1} \sum_{|s|\leq N} \mathcal{F}(dp_n)(s)\mathcal{F}(dp_n)(-s) = \int_0^{2\pi} \sigma^2(t)\, dt$$

in probability.

Asymptotic results

- Asymptotic results for the *equilibrium price*:

$$\lim_{n,N\to\infty} \frac{(2\pi)^2}{2N+1} \sum_{|s|\le N} \mathcal{F}(dp_n)(s)\mathcal{F}(dp_n)(-s) = \int_0^{2\pi} \sigma^2(t)\, dt$$

  in probability.

- Asymptotic results for the *contaminated price*:

$$\lim_{n,N\to\infty} \sum_{|s|\le N} \mathcal{F}(d\tilde{p}_n)(s)\mathcal{F}(d\tilde{p}_n)(-s) = +\infty$$

  in probability.

Bias computation

- Recall: $\tilde{p}(t_j) = p(t_j) + \eta(t_j) \Rightarrow \delta_j(\tilde{p}) = \delta_j(p) + \varepsilon_j$, where $\varepsilon_j = \eta_j - \eta_{j-1}$.

Bias computation

- Recall: $\tilde{p}(t_j) = p(t_j) + \eta(t_j) \Rightarrow \delta_j(\tilde{p}) = \delta_j(p) + \varepsilon_j$, where $\varepsilon_j = \eta_j - \eta_{j-1}$.

- Realized Volatility estimate [B-R]:

$$bias := E[\hat{V}_n - V] = 2nE[\eta^2].$$

Bias computation

- Recall: $\tilde{p}(t_j) = p(t_j) + \eta(t_j) \Rightarrow \delta_j(\tilde{p}) = \delta_j(p) + \varepsilon_j$, where $\varepsilon_j = \eta_j - \eta_{j-1}$.
- Realized Volatility estimate [B-R]:

$$bias := E[\hat{V}_n - V] = 2nE[\eta^2].$$

- Fourier estimator:

$$E[\hat{\sigma}^2_{n,N} - V] = 2n\, E[\eta^2] \left(1 - \frac{1}{2N+1} \frac{\sin[(N+\frac{1}{2})\frac{2\pi}{n}]}{\sin(\frac{\pi}{n})}\right).$$

Bias computation

- Recall: $\tilde{p}(t_j) = p(t_j) + \eta(t_j) \Rightarrow \delta_j(\tilde{p}) = \delta_j(p) + \varepsilon_j$, where $\varepsilon_j = \eta_j - \eta_{j-1}$.
- Realized Volatility estimate [B-R]:

$$bias := E[\hat{V}_n - V] = 2nE[\eta^2].$$

- Fourier estimator:

$$E[\hat{\sigma}^2_{n,N} - V] = 2n\,E[\eta^2]\left(1 - \frac{1}{2N+1}\frac{\sin[(N+\frac{1}{2})\frac{2\pi}{n}]}{\sin(\frac{\pi}{n})}\right).$$

- Note that $\frac{1}{2N+1}\frac{\sin[(N+\frac{1}{2})\frac{2\pi}{n}]}{\sin(\frac{1}{2}\frac{2\pi}{n})} \longrightarrow 1$ as $\frac{N}{n} \longrightarrow 0$.

Bias computation

- Recall: $\tilde{p}(t_j) = p(t_j) + \eta(t_j) \Rightarrow \delta_j(\tilde{p}) = \delta_j(p) + \varepsilon_j$, where $\varepsilon_j = \eta_j - \eta_{j-1}$.
- Realized Volatility estimate [B-R]:

$$bias := E[\hat{V}_n - V] = 2nE[\eta^2].$$

- Fourier estimator:

$$E[\hat{\sigma}^2_{n,N} - V] = 2n\, E[\eta^2] \left(1 - \frac{1}{2N+1} \frac{\sin[(N+\frac{1}{2})\frac{2\pi}{n}]}{\sin(\frac{\pi}{n})}\right).$$

- Note that $\frac{1}{2N+1} \frac{\sin[(N+\frac{1}{2})\frac{2\pi}{n}]}{\sin(\frac{1}{2}\frac{2\pi}{n})} \longrightarrow 1$ as $\frac{N}{n} \longrightarrow 0$.

- Moreover $2n\, E[\eta^2] \left(1 - \frac{1}{2N+1} \frac{\sin[(N+\frac{1}{2})\frac{2\pi}{n}]}{\sin(\frac{1}{2}\frac{2\pi}{n})}\right) \longrightarrow 0$ as $\frac{N^2}{n} \longrightarrow 0$.

MSE computation

- Realized Volatility estimate [B-R]:

$$E[(\hat{V}_n - V)^2] = 2\frac{h}{n}(Q + o(1)) + \Lambda_n,$$

where $Q = \int_0^h \sigma^4(s)ds$ and

$$\Lambda_n := n^2\alpha + n\beta + \gamma,$$

with

$$\alpha = (E[\varepsilon^2])^2, \quad \beta = E[\varepsilon^4] + 2E[\varepsilon^2\varepsilon_{-1}^2] - 3(E[\varepsilon^2])^2$$

$$\gamma = 4E[\varepsilon^2]V - 2E[\varepsilon^2\varepsilon_{-1}^2] + 2(E[\varepsilon^2])^2.$$

MSE computation

◇ Fourier estimator:

$$E[(\hat{\sigma}^2_{n,N} - V)^2] = 2\frac{2\pi}{n}(Q + o(1)) + n^2\hat{\alpha} + n\hat{\beta} + \hat{\gamma}$$

where

$$D_N(t) := \frac{1}{2N+1}\sum_{|s|\leq N} e^{ist},$$

$$\hat{\alpha} = \alpha\left(1 + D_N^2(\frac{2\pi}{n}) - 2D_N(\frac{2\pi}{n})\right)$$

$$\hat{\beta} = \beta\left(1 + D_N^2(\frac{2\pi}{n}) - 2D_N(\frac{2\pi}{n})\right) + 8E[\eta^2]^2\left(D_N^2(\frac{2\pi}{n}) - D_N(\frac{2\pi}{n})\right)$$

$$\hat{\gamma} = \gamma + f\left(E[\eta^2], E[\eta^4], D_N(\frac{2\pi}{n})\right) + O\left(\frac{1}{2N+1}\right)$$

MSE computation

◇ Fourier estimator:

$$E[(\hat{\sigma}^2_{n,N} - V)^2] = 2\frac{2\pi}{n}(Q + o(1)) + n^2\hat{\alpha} + n\hat{\beta} + \hat{\gamma}$$

as $n, N \to \infty$, with $\frac{N^2}{n} \to 0$:

where

$$D_N(t) := \frac{1}{2N+1}\sum_{|s|\leq N} e^{ist},$$

$$\hat{\alpha} = \alpha\left(1 + D_N^2(\frac{2\pi}{n}) - 2D_N(\frac{2\pi}{n})\right)$$

$$\hat{\beta} = \beta\left(1 + D_N^2(\frac{2\pi}{n}) - 2D_N(\frac{2\pi}{n})\right) + 8E[\eta^2]^2\left(D_N^2(\frac{2\pi}{n}) - D_N(\frac{2\pi}{n})\right)$$

$$\hat{\gamma} = \gamma + f\left(E[\eta^2], E[\eta^4], D_N(\frac{2\pi}{n})\right) + O\left(\frac{1}{2N+1}\right)$$

MSE computation

◇ Fourier estimator:

$$E[(\hat{\sigma}^2_{n,N} - V)^2] = 2\frac{2\pi}{n}(Q + o(1)) + n^2\hat{\alpha} + n\hat{\beta} + \hat{\gamma}$$

as $n, N \to \infty$, with $\frac{N^2}{n} \to 0$: $o(1)$

where

$$D_N(t) := \frac{1}{2N+1}\sum_{|s|\leq N} e^{ist},$$

$$\hat{\alpha} = \alpha\left(1 + D_N^2(\frac{2\pi}{n}) - 2D_N(\frac{2\pi}{n})\right)$$

$$\hat{\beta} = \beta\left(1 + D_N^2(\frac{2\pi}{n}) - 2D_N(\frac{2\pi}{n})\right) + 8E[\eta^2]^2\left(D_N^2(\frac{2\pi}{n}) - D_N(\frac{2\pi}{n})\right)$$

$$\hat{\gamma} = \gamma + f\left(E[\eta^2], E[\eta^4], D_N(\frac{2\pi}{n})\right) + O\left(\frac{1}{2N+1}\right)$$

MSE computation

◇ Fourier estimator:

$$E[(\hat{\sigma}^2_{n,N} - V)^2] = 2\frac{2\pi}{n}(Q + o(1)) + n^2\hat{\alpha} + n\hat{\beta} + \hat{\gamma}$$

as $n, N \to \infty$, with $\frac{N^2}{n} \to 0$: $\quad o(1) \quad o(1)$

where

$$D_N(t) := \frac{1}{2N+1}\sum_{|s|\leq N} e^{ist},$$

$$\hat{\alpha} = \alpha\left(1 + D_N^2(\frac{2\pi}{n}) - 2D_N(\frac{2\pi}{n})\right)$$

$$\hat{\beta} = \beta\left(1 + D_N^2(\frac{2\pi}{n}) - 2D_N(\frac{2\pi}{n})\right) + 8E[\eta^2]^2\left(D_N^2(\frac{2\pi}{n}) - D_N(\frac{2\pi}{n})\right)$$

$$\hat{\gamma} = \gamma + f\left(E[\eta^2], E[\eta^4], D_N(\frac{2\pi}{n})\right) + O\left(\frac{1}{2N+1}\right)$$

MSE computation

◇ Fourier estimator:

$$E[(\hat{\sigma}^2_{n,N} - V)^2] = 2\frac{2\pi}{n}(Q + o(1)) + n^2\hat{\alpha} + n\hat{\beta} + \hat{\gamma}$$

as $n, N \to \infty$, with $\frac{N^2}{n} \to 0$: $o(1)$ $o(1)$ $O(1)$

where

$$D_N(t) := \frac{1}{2N+1}\sum_{|s|\leq N} e^{ist},$$

$$\hat{\alpha} = \alpha\left(1 + D_N^2(\frac{2\pi}{n}) - 2D_N(\frac{2\pi}{n})\right)$$

$$\hat{\beta} = \beta\left(1 + D_N^2(\frac{2\pi}{n}) - 2D_N(\frac{2\pi}{n})\right) + 8E[\eta^2]^2\left(D_N^2(\frac{2\pi}{n}) - D_N(\frac{2\pi}{n})\right)$$

$$\hat{\gamma} = \gamma + f\left(E[\eta^2], E[\eta^4], D_N(\frac{2\pi}{n})\right) + O\left(\frac{1}{2N+1}\right)$$

Monte Carlo setup

◇ Efficient log-price process and spot volatility:

$$dp(t) = \sigma(t)\, dW_1(t)$$

$$d\sigma^2(t) = \alpha(\beta - \sigma^2(t))dt + \nu\sigma(t)\, dW_2(t),$$

Monte Carlo setup

◇ Efficient log-price process and spot volatility:

$$dp(t) = \sigma(t)\, dW_1(t)$$
$$d\sigma^2(t) = \alpha(\beta - \sigma^2(t))dt + \nu\sigma(t)\, dW_2(t),$$

◇ Observed log-price process:

$$\tilde{p}(t_j) = p(t_j) + \eta(t_j), \qquad \eta_j \sim N(0, \xi^2)$$

Monte Carlo setup

◇ Efficient log-price process and spot volatility:

$$dp(t) = \sigma(t)\, dW_1(t)$$
$$d\sigma^2(t) = \alpha(\beta - \sigma^2(t))dt + \nu\sigma(t)\, dW_2(t),$$

◇ Observed log-price process:

$$\tilde{p}(t_j) = p(t_j) + \eta(t_j), \qquad \eta_j \sim N(0, \xi^2)$$

◇ Therefore: $\delta_j(\tilde{p}) = \delta_j(p) + \varepsilon_j$, where the $\varepsilon_j = \eta(t_j) - \eta(t_{j-1})$ follow a MA(1) process with negative first order autocorrelation.

Monte Carlo setup

◇ Efficient log-price process and spot volatility:

$$dp(t) = \sigma(t)\, dW_1(t)$$
$$d\sigma^2(t) = \alpha(\beta - \sigma^2(t))dt + \nu\sigma(t)\, dW_2(t),$$

◇ Observed log-price process:

$$\tilde{p}(t_j) = p(t_j) + \eta(t_j), \qquad \eta_j \sim N(0, \xi^2)$$

◇ Therefore: $\delta_j(\tilde{p}) = \delta_j(p) + \varepsilon_j$, where the $\varepsilon_j = \eta(t_j) - \eta(t_{j-1})$ follow a MA(1) process with negative first order autocorrelation.

◇ Moreover:

$$var(\delta_j(\tilde{p})) = \int_{t_{j-1}}^{t_j} \sigma^2(s)ds + 2\xi^2, \quad cov(\delta_j(\tilde{p}), \delta_{j-1}(\tilde{p})) = -\xi^2.$$

Monte Carlo setup

◇ We simulate second-by-second return and variance paths over a daily trading period of $h = 6$ hours, for a total of 21600 observation per day.

Monte Carlo setup

◇ We simulate second-by-second return and variance paths over a daily trading period of $h = 6$ hours, for a total of 21600 observation per day.

◇ We sample the observations for different choices of the uniform sampling interval, so that we obtain different data sets $(t_j,\ \tilde{p}(t_j),\ j = 0, 1 \ldots n)$ with σ recorded at every t.

Monte Carlo setup

◇ We simulate second-by-second return and variance paths over a daily trading period of $h = 6$ hours, for a total of 21600 observation per day.

◇ We sample the observations for different choices of the uniform sampling interval, so that we obtain different data sets $(t_j,\ \tilde{p}(t_j),\ j = 0, 1 \ldots n)$ with σ recorded at every t.

◇ In implementing the fourier estimator $\hat{\sigma}^2_{n,N}$, the smallest wavelength that can be evaluated is twice the smallest distance between two consecutive prices, which yields $N \leq n/2$ (*Nyquist frequency*).

Monte Carlo setup

◇ We simulate second-by-second return and variance paths over a daily trading period of $h = 6$ hours, for a total of 21600 observation per day.

◇ We sample the observations for different choices of the uniform sampling interval, so that we obtain different data sets $(t_j,\ \tilde{p}(t_j),\ j = 0, 1 \ldots n)$ with σ recorded at every t.

◇ In implementing the fourier estimator $\hat{\sigma}^2_{n,N}$, the smallest wavelength that can be evaluated is twice the smallest distance between two consecutive prices, which yields $N \leq n/2$ (*Nyquist frequency*).

◇ In order to estimate the relevant population moments of the noise components, we use sample moments constructed using quote-to-quote return data [B-R].

Results

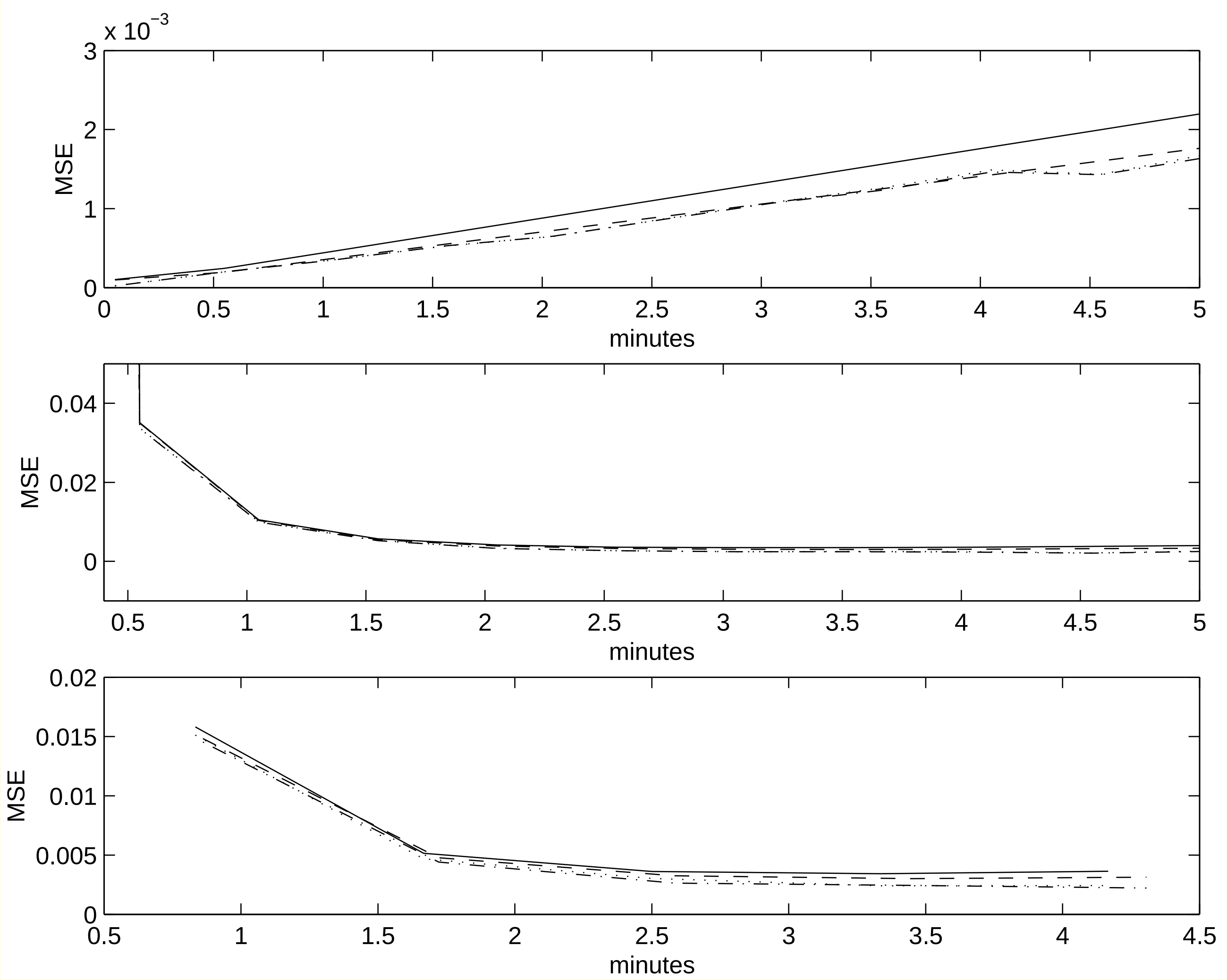


Figure 1: Fourier estimator and realized volatility MSE as a function of the sampling period.

Results

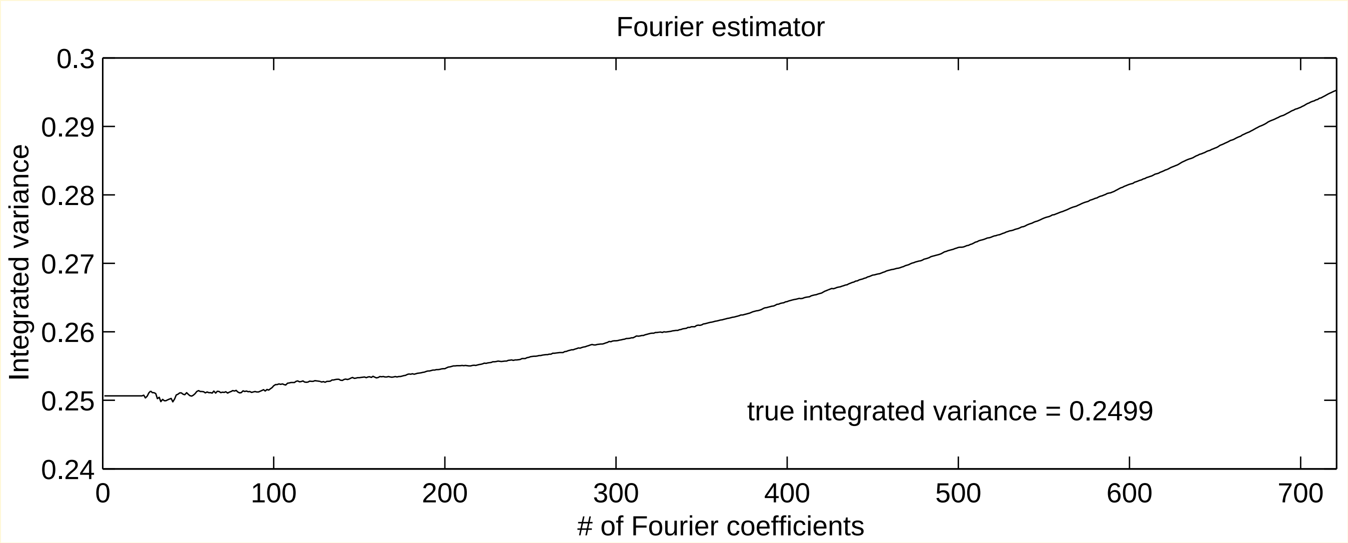


Figure 2: Average $\hat{\sigma}^2_{n,N}$ as a function of the highest frequency N employed in the Fourier expansion and $n = 21600$ (quote-to-quote returns). Parameter values: $\alpha = 0.01$, $\beta = 1.0$, $\nu = 0.05$, $\xi = 0.0142\%$. The simulations are run for 500 daily replications.

Results

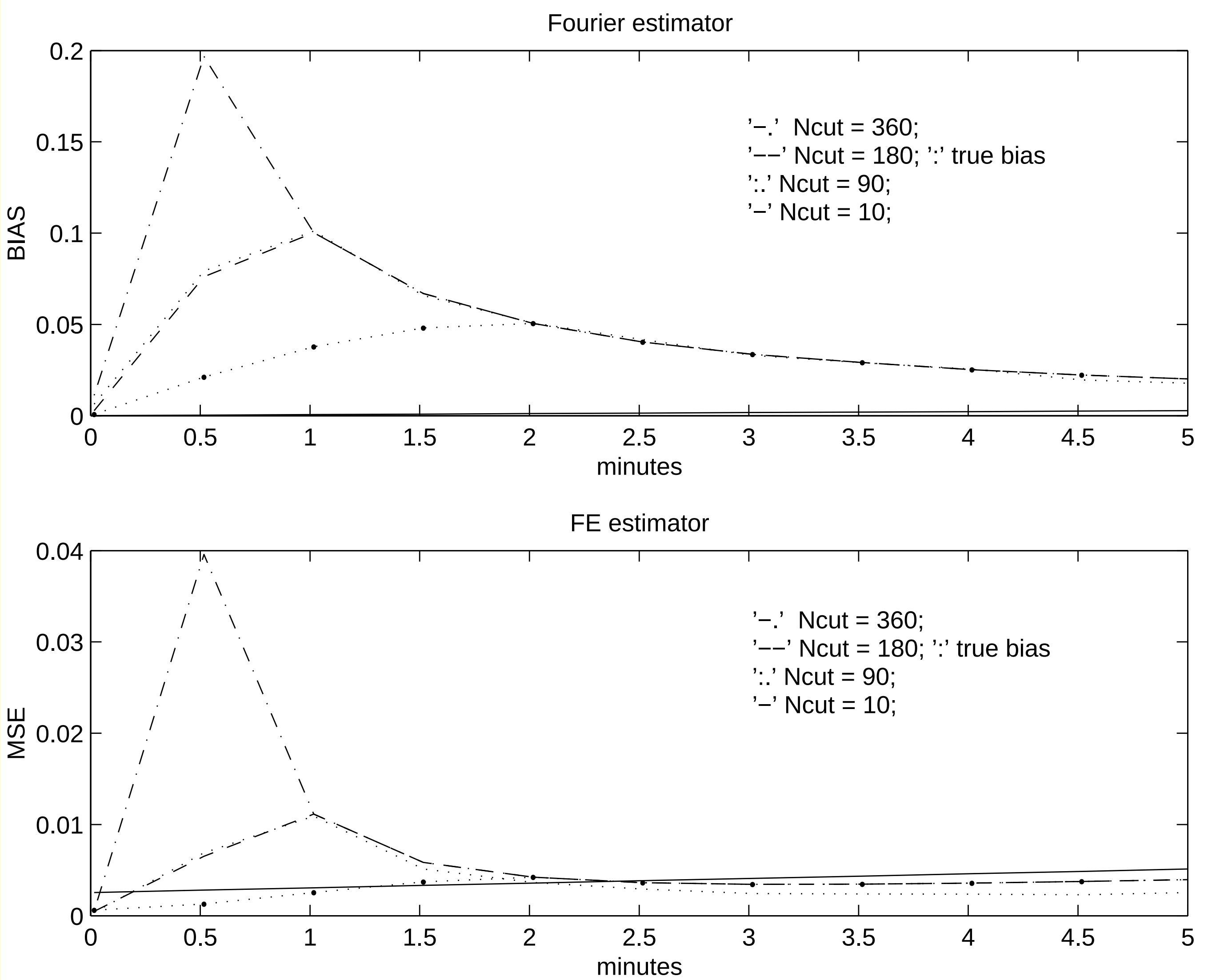


Figure 3:

Results

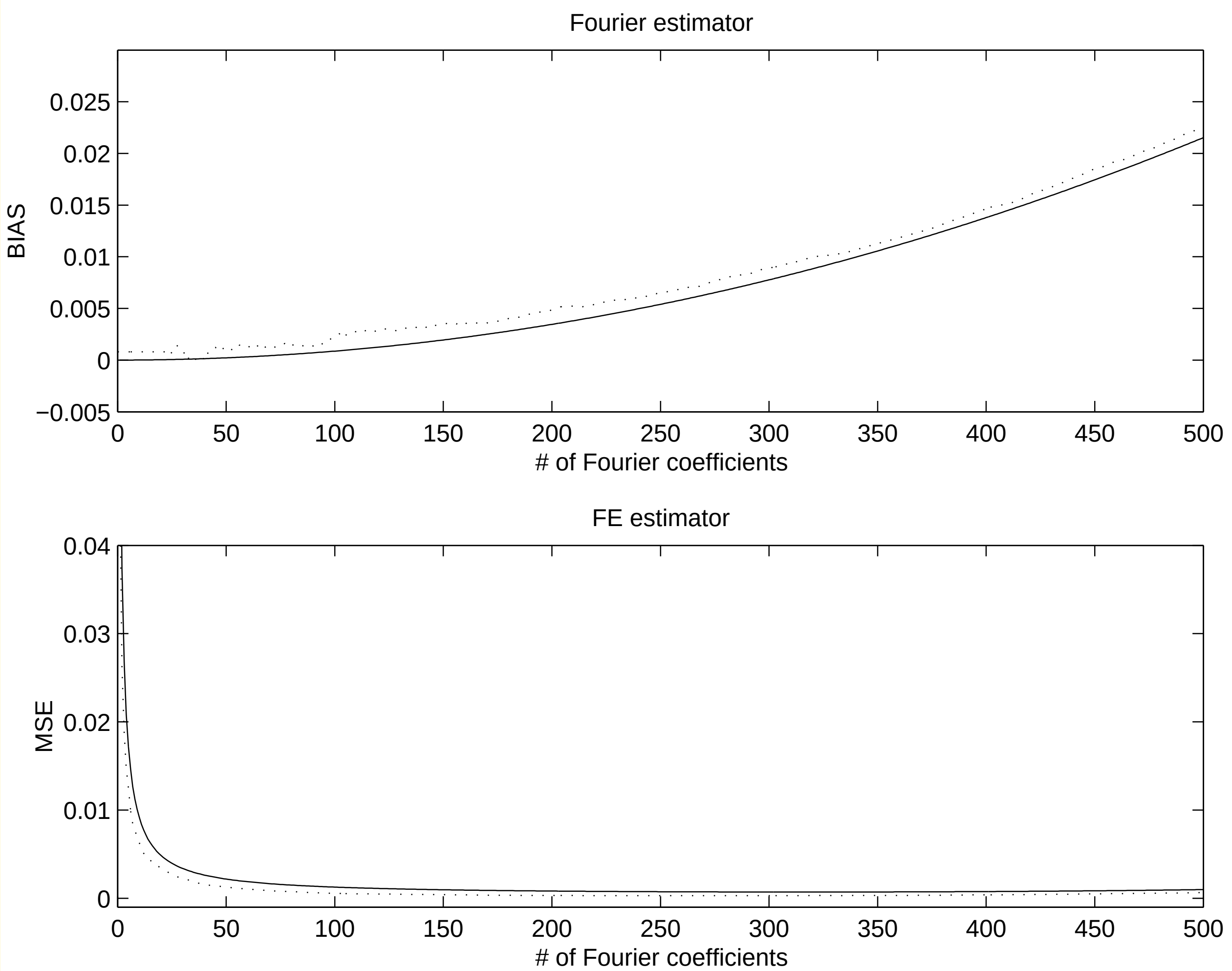


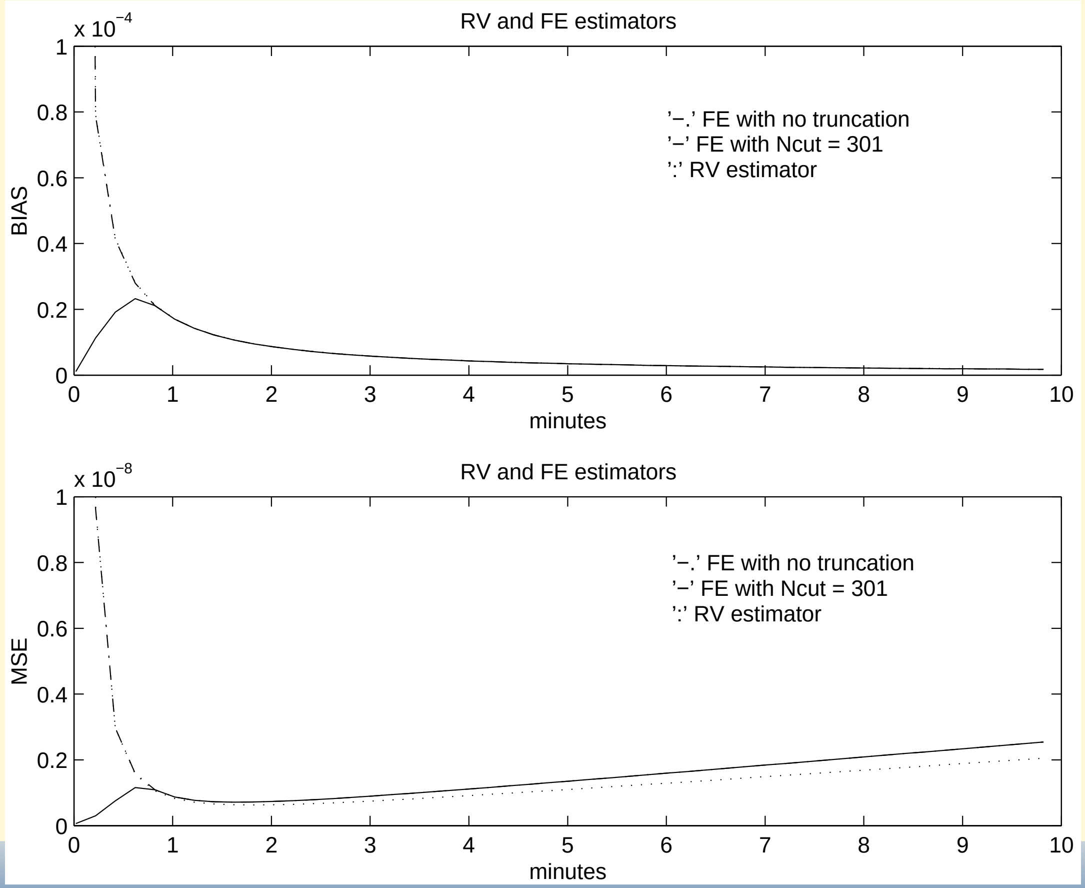
Figure 4: Min true MSE = 2.88e-4 for $k = 264$.

Empirical analysis

- We analyze quote-to-quote logarithmic prices of the Italian stock index futures, named FIB30, for the period January 11, 2000 to January 31, 2001, for a total of 269 trading days.

Empirical analysis

- We analyze quote-to-quote logarithmic prices of the Italian stock index futures, named FIB30, for the period January 11, 2000 to January 31, 2001, for a total of 269 trading days.



Optimal RV estimate:
0.12139e-3

Optimal FE estimate:

0.12155e-3

Conclusions

- We have studied the asymptotic and finite sample properties under the presence of market microstructure noise for the Fourier volatility estimator.

Conclusions

- We have studied the asymptotic and finite sample properties under the presence of market microstructure noise for the Fourier volatility estimator.
- We theoretically and empirically show that the volatility estimates are asymptotically swamped by noise as the number of data observations increases over a fixed time period.

Conclusions

- We have studied the asymptotic and finite sample properties under the presence of market microstructure noise for the Fourier volatility estimator.
- We theoretically and empirically show that the volatility estimates are asymptotically swamped by noise as the number of data observations increases over a fixed time period.
- We derive an easily implementable bias and MSE estimate as a function of the sampling frequency.

Conclusions

- We have studied the asymptotic and finite sample properties under the presence of market microstructure noise for the Fourier volatility estimator.
- We theoretically and empirically show that the volatility estimates are asymptotically swamped by noise as the number of data observations increases over a fixed time period.
- We derive an easily implementable bias and MSE estimate as a function of the sampling frequency.
- Optimal sampling at higher frequency. Lower bias and MSE than for the realized volatility estimator, if the number N of the Fourier coefficients is chosen conveniently.

Conclusions

- We have studied the asymptotic and finite sample properties under the presence of market microstructure noise for the Fourier volatility estimator.
- We theoretically and empirically show that the volatility estimates are asymptotically swamped by noise as the number of data observations increases over a fixed time period.
- We derive an easily implementable bias and MSE estimate as a function of the sampling frequency.
- Optimal sampling at higher frequency. Lower bias and MSE than for the realized volatility estimator, if the number N of the Fourier coefficients is chosen conveniently.
- Future work: analysis of the economic benefits of the optimal sampling in an asset allocation framework.