Asymptotic Theory for Renewal Based High-Frequency Volatility Estimation

Yifan Li[1,2] Ingmar Nolte[1] Sandra Nolte[1]

[1]Lancaster University [2]University of Manchester

4th Konstanz - Lancaster Workshop on Finance and Econometrics

Introduction

- Volatility is an important topic in the area of finance and financial econometrics. (Modern asset pricing, risk management, etc.)
- Conventional (low frequency) measures: daily squared return, daily range, GARCH, etc.
- High-frequency data $\rightarrow$ more precise volatility measures
  - The Realized Volatility (RV) (Andersen et al., 1998)
  - The duration-based volatility estimator (Engle and Russell, 1998; Andersen et al., 2008; Nolte et al., 2017)
  - The intensity-based volatility estimator (Gerhard and Hautsch, 2002)
  - The Realized Range (Christensen and Podolskij, 2007)

  ...

Introduction

- The duration-based volatility estimators have been shown to perform better than the RV-type estimators:
  - The parametric duration-based (PD) volatility estimator (Engle and Russell, 1998; Tse and Yang, 2012; Nolte et al., 2017)
  - The non-parametric duration-based (NPD) volatility estimator (Andersen et al., 2008; Nolte et al., 2017)
- However, the asymptotic behaviours of these estimators are largely unknown, as the findings from these papers are mainly based on simulations and empirical investigations.

Contributions

- We propose a novel approach to estimate high frequency volatility based on a renewal process under business time, and develop the asymptotic theory for the proposed estimator.
- We demonstrate that parametric volatility estimator based on point process can lead to a substantial efficiency gain compared to its non-parametric version.
- We propose a smoothed duration-based volatility estimator that can outperform realized kernel and pre-averaged RV under general MMS noise and jump.

The *NPD* Estimator

Definition 1

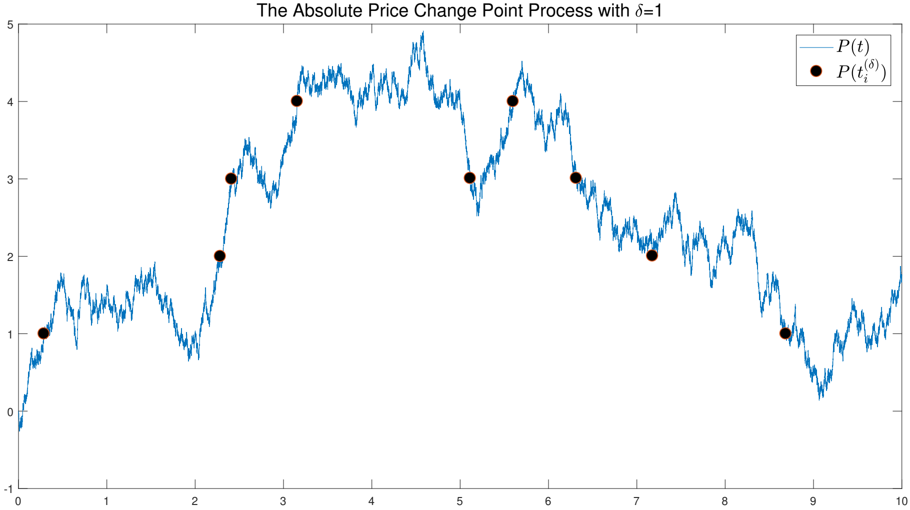
The Absolute Price Change Point Process: *The absolute price change point process* $\{t_i^{(\delta)}\}_{i=0,1,\cdots}$ *for an observed log-price process* $P(t)$ *and a given price change threshold* δ *is constructed as follows:*

1. *Set* $t_0^{(\delta)} = 0$ *and choose a threshold* δ.
2. *For* $i = 1, 2, \cdots$, *compute the first exit time,* $t_i^{(\delta)}$, *of* $P(t_{i-1}^{(\delta)})$ *through the double barrier* $[P(t_{i-1}^{(\delta)}) - \delta, P(t_{i-1}^{(\delta)}) + \delta]$ *as:*

$$t_i^{(\delta)} = \inf_{t > t_{i-1}^{(\delta)}} \{|P(t) - P(t_{i-1}^{(\delta)})| \geq \delta\}.$$

Iterate until the sample is depleted.

The *NPD* Estimator



The *NPD* Estimator

Point Process-Based Volatility Estimation

- Basic idea from Engle and Russell (1998), Gerhard and Hautsch (2002) and Nolte et al. (2017):
  - Each arrival of the price event $t_i^{(\delta)}$ contributes approximately δ^2 to the integrated variance process.
  - Therefore, we can use the number of events multiplied by δ^2 as a measure of volatility within an interval.
  - Similarly, the instantaneous arrival rate (intensity) of the point process multiplied by δ^2 can be used as a measure of the instantaneous volatility.
  - The superscript (δ) denotes that the process is associated with a δ-absolute price change point process.

The *NPD* Estimator

Non-Parametric Duration-Based Volatility Estimator

- Let $X^{(\delta)}(t) = \sum_{i=1}^{\infty} \mathbb{1}_{\{t_i^{(\delta)} \leq t\}}$ denote the counting function of the point process.
- For an interval $(0, t)$, we can formulate a simple volatility estimator:

$$NPD(0,t) = X^{(\delta)}(t)\delta^2 \tag{1}$$

- Our task is to derive the asymptotic distribution of the *NPD* estimator. Obviously this cannot be done without choosing a model for the price process.

The *NPD* Estimator

Asymptotic Distribution for *NPD*

- Suppose that:

$$P(t) = P(0) + \int_0^t \sigma(t) dW(t) \tag{2}$$

  where $\sigma(t)$ is a càdlàg process with $\lim_{t\to\infty} \int_0^t \sigma(t) = \infty$.
- We are usually interested in the integrated variance:

$$IV(0,t) = \int_0^t \sigma^2(t) dt \tag{3}$$

- We show that the *NPD* estimator has the following asymptotic distribution:

$$\lim_{t\to\infty} \frac{X^{(\delta)}(t)\delta^2 - IV(0,t)}{\sqrt{\frac{2}{3} X^{(\delta)}(t)\delta^4}} \xrightarrow{d} \mathcal{N}(0,1) \tag{4}$$

The *NPD* Estimator

Some Discussions

- In (4), it is obvious that *NPD* is consistent. Interestingly, we do not need a separate estimation of the asymptotic variance.
- We also have that $\frac{IV(0,t)}{X^{(\delta)}(t)} \xrightarrow{a.s.} \delta^2$. So if we plug this in the asymptotic variance...

$$V[X^{(\delta)}(t)\delta^2] \to \frac{2IV(0,t)^2}{3X^{(\delta)}(t)}. \quad (5)$$

- Think of $X^{(\delta)}(t)$ as the sampling frequency of the *NPD* estimator. We list asymptotic variances for some RV estimators:
    - Calendar time RV: $\frac{2IQ(0,T)}{N}$.
    - Business time RV: $\frac{2IV(0,T)^2}{N}$.
    - Tick time RV: $\frac{2IQ(0,T)}{3N}$.
- Under the same sampling frequency, the *NPD* estimator outperforms all the RV estimators above in terms of efficiency!

Theoretical Results

Main Idea of the Proof to (4)

- For a price process $P(t)$ and its integrated variance process $IV(0,t)$, we first define a time change $\tau(t) = IV(0,t) \equiv \int_0^t \sigma^2(s)ds$. The changed time $\tau(t)$ is called business time.
- Assume that the time changed process $\tilde{P}(\tau(t)) = P(t)$ follows a Lèvy process **in business time** $\tau(t)$.
- Construct a renewal process based on the Lèvy process in business time.
- Use renewal theory to estimate the time elapse on business clock.

Theoretical Results

Assumptions of Price Process

- We only need to assume that under the business time, $\tilde{P}(\tau(t))$ is a Lèvy process.
- Is this assumption reasonable?

Theorem 1

(Dambis-Dubin Schwarz): *Let $(M_t)_{t\geq 0}$ be a continuous $\mathcal{F}_t$-local martingale such that its quadratic variation $\langle M\rangle_\infty = +\infty$. There exists a Brownian motion $(B_t)_{t\geq 0}$, such that for every $t \geq 0$, $M_t = B_{\langle M\rangle_t}$.*

- It at least holds for ANY continuous local martingale that satisfies the above theorem.
- It also holds for an inhomogeneous compounded Poisson process as in Oomen (2005).

Theoretical Results

Renewal Based Volatility Estimator

- Sample the price process at $\{t_i\}_{i=1,2,\dots}$ so that the business time process $\{\tau(t_i)\}$ is renewal, i.e., $\tilde{D}_i \equiv \tau(t_i) - \tau(t_{i-1})$ is i.i.d.
- The intuition is that we sample the price process so that the IVs between two points are i.i.d. random variables.
- Let $\mu = \mathrm{E}[\tau(t_i) - \tau(t_{i-1})]$ and $\sigma^2 = \mathrm{V}[\tau(t_i) - \tau(t_{i-1})]$ be finite and non-zero. The class of Renewal Based Volatility (RBV) estimators is defined as:

$$RBV(0,t) = X(t)\mu \tag{6}$$

- We show that:

$$\lim_{t\to\infty} \frac{X(t)\mu - IV(0,t)}{\sigma\sqrt{X(t)}} \xrightarrow{d} \mathcal{N}(0,1). \tag{7}$$

Theoretical Results

- The *NPD* estimator belongs to the class of *RBV* estimators. If the price process follows a local martingale, then under business time, $\tilde{P}(\tau(t))$ is a standard Brownian motion. The point process under business time is just the exit time through the double barrier $[-\delta, \delta]$, and we have:

$$\mu = \delta^2, \quad \sigma^2 = \frac{2}{3}\delta^4 \tag{8}$$

- We can also consider a range threshold r, and construct *RBV* based on the exit time when the price range reaches r. Then:

$$\mu = \frac{1}{2}r^2, \quad \sigma^2 = \frac{1}{3}r^4 \tag{9}$$

- One can easily show that under the same sampling frequency, the range duration-based estimator is twice as efficient as the *NPD* estimator.

The *PD* Estimator

A Parametric Design

- Let $\mathcal{F}_t$ be the natural filtration of the point process, the $(\mathcal{F}_t)$-conditional intensity process $\lambda^{(\delta)}(t|\mathcal{F}_t)$ of $X^{(\delta)}(t)$ is defined as:

$$\lambda^{(\delta)}(t|\mathcal{F}_t) \equiv \lim_{\Delta \downarrow 0} \frac{1}{\Delta} \mathrm{E}[X^{(\delta)}(t+\Delta) - X^{(\delta)}(t)|\mathcal{F}_t]. \tag{10}$$

- An instantaneous volatility estimator can be formulated as $\delta^2 \lambda^{(\delta)}(t|\mathcal{F}_t)$. Usually we use a parametric model to estimate the conditional intensity in practice.
- We are more interested in properties of the parametric volatility estimator of the following form:

$$PD(0,t) = \delta^2 \int_0^t \lambda^{(\delta)}(s|\mathcal{F}_s)ds. \tag{11}$$

The *PD* Estimator

Asymptotic Distribution of *PD*

- Assume that $\lambda^{(\delta)}(t|\mathcal{F}_t)$ is known, we have

$$\lim_{t\to\infty} \frac{\delta^2 \int_0^t \lambda^{(\delta)}(s|\mathcal{F}_s)ds - IV(0,t)}{\sqrt{C \cdot X^{(\delta)}(t)\delta^4}} \xrightarrow{d} \mathcal{N}(0,1) \tag{12}$$

- The constant C can be approximated numerically to an arbitrary precision. We find that $C \approx 0.034$ if the price process is a pure diffusion.
- We show that: $\lambda^{(\delta)}(t|\mathcal{F}_t) = \tilde{\lambda}^{(\delta)}(\tau(t)|\mathcal{F}_t)\sigma^2(t)$, where $\tilde{\lambda}^{(\delta)}(\tau(t)|\mathcal{F}_t)$ is the conditional intensity of the renewal process in business time.

The *PD* Estimator

More discussions

- Results in (8) shows that, the *PD* estimator can be much more precise than its non-parametric counterpart if we have a good model for the conditional intensity.
- We can use data beyond the window of volatility estimation, and provide intraday volatility estimation. E.g. use a month's data to estimate volatility for an hour.
- We can add MMS covariates in the parametric model to further improve the performance. To do this we need to augment the information set. This is still under development.

Further Results

Comparison of Non-Parametric Volatility Estimators

- We proceed to compare the performance of different non-parametric volatility estimators by a simulation study.
- We construct three estimators based on the price durations $\{t_i^{(\delta)}\}$:
  1. The *NPD* estimator.
  2. The renewal RV estimator.
  3. The exponentially smoothed NPD^z estimator: constructing *NPD* estimator based on exponentially smoothed price process:

$$S_j = \gamma S_{j-1} + (1-\gamma)P_j, \quad \gamma \in [0,1] \tag{13}$$

- We compare the performance of these estimators against popular calendar time RV estimators.

Further Results

Competing Estimators

Table: List of all volatility estimators considered in the simulation study

Acronym	Description	MMS	Jump
NPD		N	Y
$RV^{(\delta)}$	Renewal RV	N	N
NPD^z		N	Y
RV	Realized Variance	N	N
RBip	Realized Bipower Variation	N	Y
RK	Realized Kernel	Y	N
PRV	Pre-averaged Realized Variance	Y	N
PBip	Pre-averaged Bipwer Variation	Y	Y

Further Results

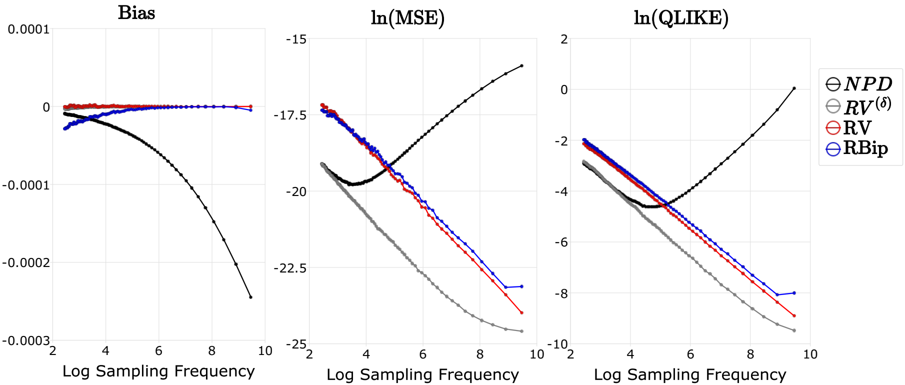
Simulation

- We use a one-factor stochastic volatility (1FSV) model to simulate daily transaction processes.
- MMS noise is a tick-time negatively correlated noise with three different levels.
- We add diurnal patterns of transactions and volatility in the 1FSV model.
- We consider price discretization and flat trades.
- We consider case both with (large infrequent) jumps and without jumps.

Further Results

No Noise Case

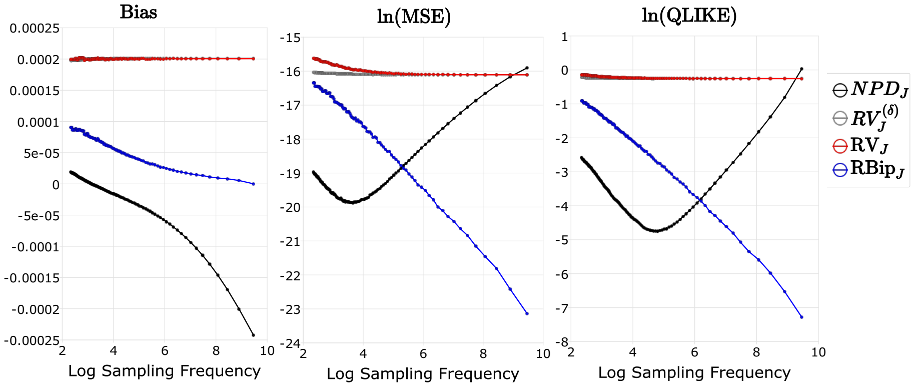
Figure: 1FSV model without jump



Further Results

No Noise Case

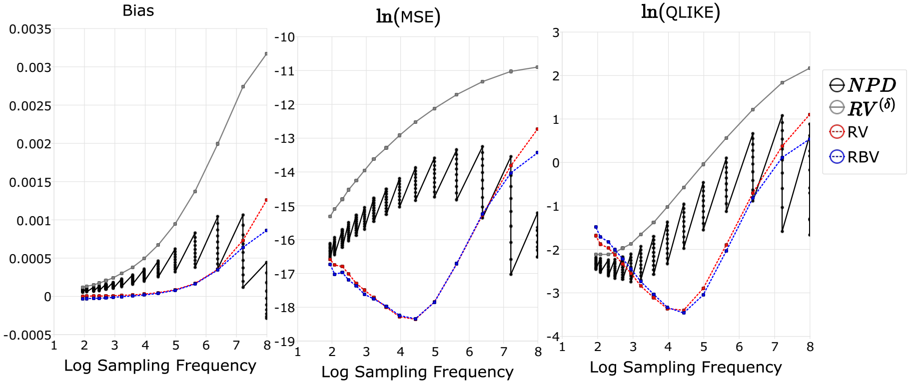
Figure: 1FSV model with jump



Further Results

Medium Noise Case

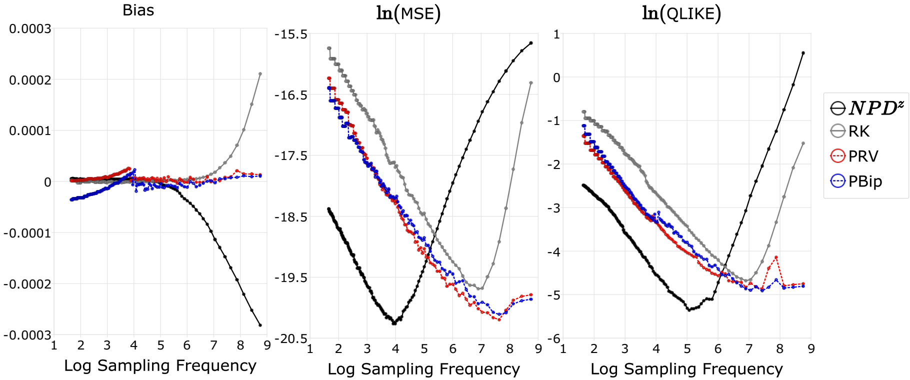
Figure: 1FSV model without jump



Further Results

Medium Noise Case

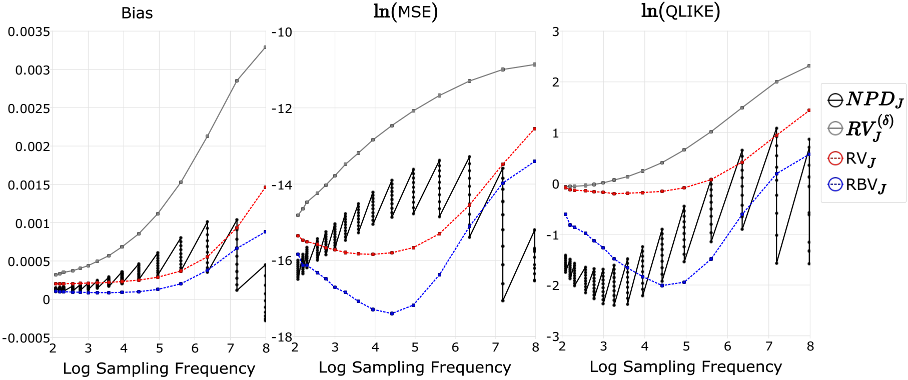
Figure: 1FSV model without jump



Further Results

Medium Noise Case

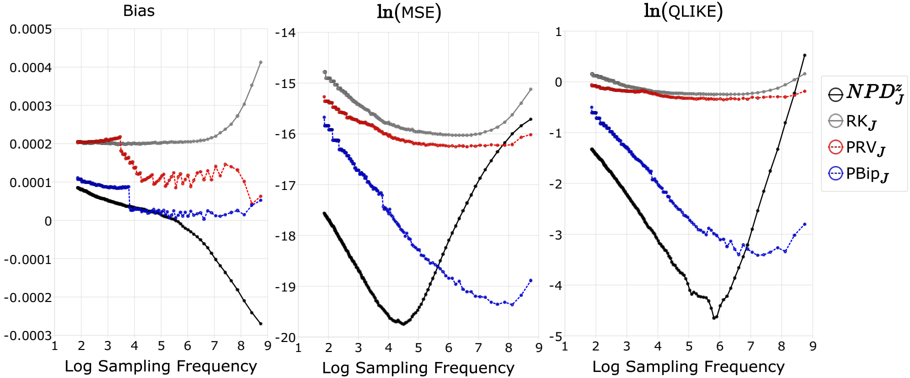
Figure: 1FSV model with jump



Further Results

Medium Noise Case

Figure: 1FSV model with jump



Further Results

Summary of the Simulation Study

- The *NPD* estimator performs better than calendar time RV estimators in theory and is very robust to jumps, but its performance is limited by the MMS noise and the time discretization.
- A large δ is required for the *NPD* estimator to outperform the calendar time RV methods.
- The optimized NPD^z estimator is the overall winner for all noise and jump cases. It outperforms the optimized RK, PRV and PBip for small to moderate sampling frequencies.

Conclusion

- We propose a novel class of volatility estimators and provides the framework to prove its asymptotic properties.
- We show that a parametric RBV estimator can lead to further improvements on the efficiency of volatility estimation.
- We validate the use of the NPD and PD estimator in the existing literature from a theoretical perspective by showing that they are more efficient than the RV-type estimators.
- We propose the NPD^z estimator which has better performance than the calendar time methods in terms of MSE and QLIKE.

Thank you!