Efficient Realized Variance Estimation in Time-Changed Diffusion Processes

Timo Dimitriadis[1]
Roxana Halbleib[2]
Jeannine Polivka[3]
Sina Streicher[4]

Preliminary Version
March 24, 2022

Abstract

In this paper we analyze the statistical properties of realized variance estimators under the assumption that financial logarithmic prices follow a time-changed diffusion process. The time-change takes the form of a counting process implying that the logarithmic price is a pure jump process with stochastic and time-varying tick volatility. This framework is more appropriate to capture the dynamics of observed logarithmic price processes than the standard diffusion model, and it is also more general than the compound Poisson process with constant tick volatility. We show that our approach is particularly suited to model the logarithmic transaction prices of stocks, as they exhibit time-varying tick volatility. Our analysis deals with three types of sampling schemes, namely clock-time sampling, business time sampling and transaction time sampling. We theoretically show that, under no market microstructure noise, realized variance is an unbiased estimator of integrated variance and that business time sampling is optimal in terms of mean squared error. To deal with market microstructure noise, we theoretically and empirically consider various bias-corrected realized variance estimators. Our simulation results show that transaction time sampling outperforms business time sampling for high sampling frequencies and large levels of market microstructure noise.

Keywords: Time-changed Brownian Motion, Diffusion Model, Pure Jump Process, High-Frequency Data, Realized Variance, Market Microstructure Noise, Optimal Sampling

[1]Department of Economics, University of Konstanz, Germany; email: timo.dimitriadis@uni-konstanz.de

[2]Corresponding author. Department of Economics, University of Konstanz, P.O. Box 124, Universitaetstrasse 10, 78464, Konstanz, Germany; email: roxana.halbleib@uni-konstanz.de; telephone: 0049 (0) 7531 88 5373; fax: 0049 (0) 7531 88 4450.

[3]University of St. Gallen, Switzerland; email: jeannine.polivka@unisg.ch

[4]KOF Swiss Economic Institute, ETH Zürich, Switzerland; email: streicher@kof.ethz.ch

1 Introduction

The estimation and forecasting of variance in financial time series plays a major role in risk management, portfolio management and asset pricing. But since the variance is generally unobservable, its estimation is not straightforward. A popular approach is given by the GARCH model introduced by Engle (1982) and Bollerslev (1986), which makes the variance observable conditionally on past information. Other commonly used approaches include the class of stochastic volatility models. However, these methods use (restrictive) parametric frameworks to estimate daily variances based on data sampled once a day. More accurate estimates of daily variation are obtained by using intraday information as in the realized variance estimator introduced by Andersen et al. (2000*a*,*b*) and Andersen et al. (2001). They show that, under the assumption that the logarithmic price process follows a standard continuous-time diffusion model, realized variance is an unbiased estimator of integrated variance as the latter coincides with quadratic variation when no jumps occur. Barndorff-Nielsen and Shephard (2002*a*) and Bandi and Russell (2008) derive the asymptotic normality of the realized variance estimator.

In this paper, differently from the standard diffusion approach, we assume that the logarithmic price process follows a time-changed diffusion model as in Dahlhaus and Tunyavetchakit (2016). The time-change or intrinsic time – a stochastic time transformation of clock-time – takes the form of a counting process. The foundation for this model goes back to Clark (1973), who is the first to use a diffusion approach subordinated by a stochastic process to model prices in finance. More precisely, he uses trading volume as a subordinator of a Brownian motion. On this basis, Jones, Kaul and Lipson (1994) consider the trading volume and the number of trades as subordinators and conclude that it is in fact the latter that generates volatility. Ané and Geman (2000) propose a general time-transformation framework and find that the time-change that leads to the recovery of normal returns is the one computed from the number of transactions, which is in line with the findings of Jones, Kaul and Lipson (1994).

In our framework, the spot volatility splits up into the product of tick volatility and the intensity of the counting process.[1] While Dahlhaus and Tunyavetchakit (2016) use the volatility decomposition to obtain an improved estimate of spot-volatility, we are interested in the analysis and performance of realized variance as an estimator of integrated variance. In particularly, we use a doubly stochastic Poisson process to describe the intrinsic time. For instance, the Poisson process may represent the accumulated number of transactions, quotes, or a combination of the two during a trading day. We call our underlying price model the tick-time stochastic volatility (TTSV) model.[2] The TTSV model is right-continuous, has finite variation and a stochastic jump at each tick.

A concurrent approach to our model is the compound Poisson process as introduced by Press (1967, 1968) combined with a noise term allowing for a moving average

[1] Tick volatility corresponds to the instantaneous volatility in the standard diffusion model.

[2] Here, the tick-times are not to be confused with the tick-time as e.g. in Griffin and Oomen (2008) who define a tick as a price change event. In fact, the tick-times may represent any subset thereof.

structure as assumed by Oomen (2004, 2005, 2006). This type of process is a special case of ours as it exhibits no stochastic tick volatility, but a constant one. In both, the TTSV model and the model of Oomen (2005, 2006), the realized variance is an inconsistent estimator since the number of ticks within a finite interval is bounded by the definition of the Poisson process (Brémaud, 1981). Similarly to Oomen (2005, 2006), we show that realized variance is an unbiased estimator of integrated variance and we derive a closed-form solution of its MSE under various sampling schemes and no market microstructure noise. The sampling schemes we consider are based on clock-time, business time (driven by the intraday pattern of the spot variance) and transaction time (driven by the number of trades). By employing these sampling schemes to IBM and EUR USD exchange rate transaction data, we demonstrate that our time-varying stochastic volatility is more realistic for stocks, whereas a constant volatility seems plausible for foreign exchange markets. When theoretically comparing the MSE's of the three sampling schemes, we find that business time sampling is the optimal choice if its specific approximation error for the unobserved spot variance is ignored.

In practice, asset prices are contaminated with market microstructure noise. Sources of the noise include price discreteness (Harris, 1990, 1991), bid-ask bounces and infrequent trading.[3] Hansen and Lunde (2006) provide empirical evidence on the structure of the noise. They find that market microstructure noise contains time-dependencies, is correlated with the uncontaminated log-price process, and that its properties vary over time. Bandi and Russell (2006, 2008) find significant negative first order serial correlation in IBM quote data and in S&P500 stocks. Among others, Zhang, Mykland and Aït-Sahalia (2005) and Bandi and Russell (2008) consider independent noise and, even though this assumption is not in line with recent empirical evidence, it appears to be reasonable for low sampling frequencies (Hansen and Lunde, 2006).

In order to deal with market microstructure noise, we add independent noise as well as moving average noise to the observed price process and consider bias-corrected versions of the realized variance estimator as proposed by Hansen and Lunde (2006). For transaction time sampling, we derive the maximum number of sampled intra-day returns for which the adjusted realized variance is unbiased.

In a Monte-Carlo experiment, we simulate the logarithmic price process according to our assumption and investigate the performance of realized variance under the three different sampling schemes and various forms of market microstructure noise. We find that, in the absence of market microstructure noise, business time sampling outperforms all other sampling schemes in terms of bias and MSE for reasonable sampling frequencies. Under independent noise, a simple bias correction suffices to ensure that the realized variance estimate under transaction time sampling remains unbiased, which is not the case for the other sampling schemes. This pattern extends to noises that take the form of moving average processes, as long as the number of sampled returns does not exceed a certain maximum value. In general, we observe that clock-time sampling performs worst and we strongly recommend the use of business or transaction time sampling instead. We conduct robustness checks regarding the parameter and data selection for the estimation of the intensity

[3]For extensive surveys on this topic see Madhavan (2000) and O'hara (1995).

process. We find that our simulation is sufficiently insensitive to different variations, strengthening the validity of our results.
The contribution of this paper is twofold. First, on the theoretical side, we investigate the statistical properties of realized variance both with and without market microstructure noise in the general framework of a time-changed diffusion model, which includes the model of Oomen (2004, 2005, 2006) as a special case. Second, through our simulation study, we provide important insights into the optimal usage of the sampling frequency and sampling scheme to obtain unbiased and efficient realized variance estimators.
The remainder of the paper is structured as follows. In Section 2, we detail the theoretical framework for our approach and provide results regarding the statistical properties of realized variance estimators. In Section 3, we conduct a comprehensive simulation study that analyzes the performance of the realized variance estimator under different sampling schemes and for different types of market microstructure noise. We conclude in Section 4. All proofs can be found in Appendix C.

2 The Theory

In this section, we provide the theoretical framework for the time-changed diffusion model which we assume for the logarithmic price process of financial assets. **To this end,** we first introduce some basic theoretical concepts followed by the presentation of the price model we assume in our paper as well as the alternative models that are currently used in the realized variance literature. Before presenting theoretical results on the properties of the realized variance estimator, we present several sampling schemes available under our theoretical setting as well as some standard ways on how we deal with the concept of market microstructure noise.

2.1 Theoretical Concepts

In this section, we introduce some notation and give a model-free definition of spot variance, integrated variance as well as realized variance.
Throughout this paper let $(\Omega, \mathcal{F}_t, \mathbb{P})$ be a filtered probability space and $(\mathcal{F}_t)_{t\geq 0}$ a filtration. For $t \geq 0$ let $P(t)$ denote the right-continuous logarithmic price process of some asset. For $0 \leq s < t$, the logarithmic return process over the interval $[s, t]$ is defined by

$$r(s,t) := P(t) - P(s).$$

For simplicity, we consider just one trading day throughout this paper.[4] On the considered trading day, the interval $[0, T]$ represents the trading hours on that day, e.g., from 9:30am to 4:00pm on the New York Stock Exchange (NYSE). The daily return is given by $r_{\text{daily}} := r(0,T) = P(T) - P(0)$.
Since $P(t)$ is a right-continuous process, there are potentially arbitrarily many returns within a trading day. In order to facilitate the use of intra-daily returns, we

[4]Often, a multi-day framework is considered, see e.g. Barndorff-Nielsen and Shephard (2002*b*). However, all concepts are easily extendable to a framework that includes multiple trading days.

first introduce a very general grid of $[0,T]$. We define the general sampling (GS) scheme or general grid

$$\tau := (\tau_0, \ldots, \tau_M)$$

with $0 = \tau_0 < \tau_1 < \ldots < \tau_M = T$ and where $M \in \mathbb{N}$ denotes the number of subintervals $(\tau_{j-1}, \tau_j]$ for $j = 1, \ldots, M$. Note that the grid is not necessarily equidistantly spaced in time, i.e., the length of the sub-intervals may vary.
Throughout this paper, we refer to M as the number of sampled observations.[5] Then, given $M+1$ price observations on the grid τ, the corresponding intra-daily returns over a trading day are given by

$$r_j := r(\tau_{j-1}, \tau_j) = P(\tau_j) - P(\tau_{j-1}), \quad j = 1, \ldots, M. \tag{1}$$

In what follows we make a clear distinction between volatility and variance: i.e., the volatility is the **square** root of variance.
The *spot variance* of the logarithmic price process $\{P(t)\}_{t\geq 0}$ is given by:

$$\sigma^2_{spot}(t) := \lim_{\delta \searrow 0} \frac{\mathbb{E}\left[(P(t+\delta) - P(t))^2 \middle| \mathcal{F}_t\right]}{\boldsymbol{\delta}}, \tag{2}$$

where $\delta > 0$.
The spot variance can be thought of as instantaneous variance at time t of a price process $P(t)$ (e.g. Engle and Russell, 1998; Dahlhaus and Tunyavetchakit, 2016). Since in our paper, we are interested in a measure of variance within a specific period $[0,T]$, e.g., daily variance, a natural choice is the integral of spot variance within that period, usually referred to as integrated variance in the financial econometrics literature (see e.g. Barndorff-Nielsen and Shephard, 2002*a*, Andersen et al., 2006). Thus, the *integrated variance* (IV) associated with the logarithmic price process $P(t)$ over the interval $[0,T]$ is given by:

$$\mathrm{IV}(0,T) := \int_0^T \sigma^2_{spot}(r)\, dr. \tag{3}$$

The last variance concept introduced here is the *realized variance* (RV), which is defined based on $M \in \mathbb{N}$ returns in the interval $[0,T], T \geq 0$:

$$\mathrm{RV}_M(0,T) := \sum_{j=1}^{M} r_j^2, \tag{4}$$

where the returns r_j are as defined in (1).
As the considered interval is one trading day, realized variance is an ex-post measurement of return variation over that specific trading day. In this paper, we focus on estimating RV by assuming that the price process follows a time-changed diffusion model, as described in the next subsection.

[5] Sometimes, the step size T/M is called sampling frequency (e.g. Hansen and Lunde, 2006) which is not to be mistaken for M.

2.2 Tick-Time Stochastic Volatility Model

In this section, we introduce the model we assume for the logarithmic price process $\{P(t)\}_{t\geq 0}$, namely the tick-time stochastic volatility (TTSV) model which is a time-changed Brownian motion combined with a stochastic volatility process. First, we briefly elaborate on the properties of time-changed Brownian motions and thereafter, we define the TTSV model.

To create a time-changed diffusion subordinated to a Brownian motion via an intrinsic time scheme, we consider a stochastic transformation of clock time. Let $\{N(t)\}_{t\geq 0}$ be a point process:

$$N(t) := \begin{cases} i & \text{if } t \in [t_i, t_{i+1}),\ i \geq 0 \\ +\infty & \text{if } t \geq t_\infty, \end{cases} \tag{5}$$

where **the random arrival times** $t_i \in [0,\infty)$ satisfy $t_0 = 0, t_\infty = \lim_{i\to\infty} t_i = \infty$ and $t_i < t_{i+1}$. The point process $\{N(t)\}_{t\geq 0}$ is a right-continuous step function with jumps of magnitude one at each arrival time $t_i, i \geq 0$. In the context of financial markets, the arrival times can e.g. be thought of as of quote or transaction times of a certain financial asset.
Let $\{B(t)\}_{t\geq 0}$ denote a standard Brownian motion. The time-changed Brownian motion $\{B(N(t))\}_{t\geq 0}$ is a right-continuous process with left limits. Since the time transformation $t \mapsto N(t)$ is merely right-continuous and not strictly increasing, $B(N(t))$ is not a Brownian motion with respect to physical time t as continuity of the process and independence of the increments is not fulfilled. However, we can deduce further properties of the time-changed Brownian motion based on the following assumption.

Assumption 1. $B(t)$ and $N(t)$ are independent.

Under Assumption 1, for the time-changed Brownian motion $\{B(N(t))\}_{t\geq 0}$, it holds that:

$$\begin{aligned} \mathbb{E}\left[B(N(t)) - B(N(s))\right] &= 0, \\ \mathbb{E}\left[\left(B(N(t)) - B(N(s))\right)^2\right] &= \mathbb{E}\left[N(t) - N(s)\right] \end{aligned}$$

for all $0 \leq s < t$. In particular, $B(N(t))\,|N(t) \sim \mathcal{N}(0, N(t))$ which is immediate from the characteristic function, conditioning and using the independence assumption. Defining the difference sequence

$$U_i := U(t_{i-1}, t_i) := B(N(t_i)) - B(N(t_{i-1})),$$

by definition of the arrival times t_i, $\mathbb{E}[U_i] = 0$ and $\mathbb{E}[U_i^2] = \mathbb{E}\left[N(t_i) - N(t_{i-1})\right] = 1$ $\mathbb{P}$-a.s. which implies that

$$U_i \sim \mathcal{N}(0,1).$$

Thus, the time-changed Brownian motion exhibits jumps which are standard normally distributed at the random arrival times t_i. The natural filtration that $B(N(t_i))$ and U_i are adapted to is $\mathcal{F}_{t_i}^{\boldsymbol{BN}}$ where

$$\mathcal{F}_t^{\boldsymbol{BN}} := \sigma\left(\{N(s)\}_{0\leq s\leq t}, \{B(N(s))\}_{0\leq s\leq t}\right).$$

The above definitions imply that $\{\boldsymbol{U}_i\}_{\boldsymbol{i}\geq 0}$ is a $\mathcal{F}_{t_i}^{\boldsymbol{BN}}$-martingale difference sequence and that

$$\mathbb{E}\left[U_{\boldsymbol{i}}|\mathcal{F}_{t_i-}\right]=0 \quad \text{and} \quad \mathbb{E}\left[U_{\boldsymbol{i}}^2\middle|\mathcal{F}_{t_i-}\right]=1 \tag{6}$$

for all $i=1,\ldots,N(T)$, where $\mathcal{F}_{t_i-}^{\boldsymbol{BN}}$ is the information set including everything until t_i, but excluding t_i.
In order to illustrate that $\{B(N(t))\}_{t\geq 0}$ is in fact a pure jump process, we consider the formal differential $dB(N(t))$. By definition of the stochastic integral, it follows that

$$\int_s^t dB\left(N\left(\boldsymbol{r}\right)\right)=\sum_{s<t_i\leq t} U_i. \tag{7}$$

Assuming the counting process $\boldsymbol{\{N(t)\}_{t\geq 0}}$ e.g. to be a Poisson process, $\boldsymbol{B(N(t))}$ is a compound Poisson process with i.i.d. standard normal jumps, and therefore it belongs to the class of semi-martingales with regard to the filtration defined by $\boldsymbol{\mathcal{F}_t^{BN}}$.
To specify the logarithmic price process $\{P(t)\}_{t\geq 0}$, we assume the existence of an underlying filtered probability space given by $\left(\Omega,\mathcal{F},(\mathcal{F})_{t\geq 0},\mathbb{P}\right)$ where $\mathcal{F}_0$ contains all null sets and the filtration is right-continuous.
We define the logarithmic price process $\{P(t)\}_{t\geq 0}$ by

$$dP(t)=\varsigma(t)\,dB(N(t)), \quad t\in[0,T], \tag{8}$$

where the stochastic process $\varsigma(t)$ is called tick volatility. We henceforth refer to the log-price model in (8) as the tick-time stochastic volatility (TTSV) model. The TTSV model is also considered by Dahlhaus and Tunyavetchakit (2016), however, with the purpose of finding an improved estimate of spot volatility. In addition to Assumption 1, we make the following three assumptions:

Assumption 2. The point process $N(t)$ is a doubly stochastic Poisson process, i.e. N_t is adapted to $\mathcal{F}_t$ and has a non-negative, $\mathcal{F}_t$-measurable, continuous stochastic intensity $\lambda(t)$ satisfying for all $t\geq 0$ that $\int_0^t \lambda_s ds<\infty$ $\mathbb{P}$-a.s.[6]

Allowing the point process to follow a Poisson process with a random time-dependent intensity is a natural extension to account for different intensities of market activity throughout a trading period. The definition ensures that the point process is adapted. Furthermore, the intensity is set to be non-negative and $\mathcal{F}_t$-measurable. The boundedness of $\lambda(t)$ in L^1 prevents explosive behaviour in market activity, such that e.g. the accumulated number of transactions in a finite time interval cannot diverge to infinity.

Assumption 3. The tick volatility $\varsigma(t)$ is a non-negative, continuous and $\mathcal{F}_t$-predictable process.

This implies in particular that $\varsigma(t)$ is $\mathcal{F}_{t-}$-measurable. Note that by assuming ς to be continuous adaptedness and thus predictability with regard to $\mathcal{F}_t$ follow immediately.

[6]See Brémaud (1981) for a formal definition of doubly stochastic Poisson processes.

Assumption 4. It holds that $\mathbb{E}\left[\int_0^t \varsigma^2(r)\,\lambda(r)\,dr\right] < \infty$ for all $0 \le t \le T$.

The latter integrability condition ensures the quadratic variation to be $\mathbb{P}$-a.s. finite and allows for the stochastic integral with respect to the time-changed Brownian motion to exist.

Assumption 5. It holds that $\forall t \ge 0, \exists \varepsilon > 0:\ \mathbb{E}\left[\sup_{\delta\in[0,\varepsilon]} \int_t^{t+\delta} \varsigma^2(r)\,\lambda(r)\,dr|\mathcal{F}_t\right] < \infty$ for all $\mathbb{P}$-a.s.

This uniform integrability assumption guarantees that the spot variance converges to the product of intensity and squared tick volatility at time t such that the instantaneous volatility cannot explode. The filtration that naturally fulfills the assumptions above is given by

$$\mathcal{F}_t := \sigma\left(\{N(s)\}_{0\le s\le t}, \{\lambda(s)\}_{0\le s\le t}, \{\varsigma(s)\}_{0\le s\le t}, \{B(N(s))\}_{0\le s\le t}\right). \tag{9}$$

Note that we do not use the entire path of $B(\cdot)$, but only $B(i)$ for $i \in \mathbb{N}_0$. In summary, the logarithmic price in the TTSV model is a càdlàg process with jumps at the random arrival times $t_i, i \ge 0$ of the point process $N(t)$ and of random size $\varsigma(t_i)\,U_i$. Since the distribution of the tick volatility $\varsigma(t)$ remains unspecified, the distribution of the return process is unknown. As we have seen that $B(N(t))$ is a semi-martingale, by Assumption 4, $\int_0^t \varsigma(r)\,dB(N(r))$ is also a semi-martingale or a local martingale only?, i.e., the log-price process $P(t)$ in Equation (8) belongs to the class of semi-martingales.
Then, the TTSV process has the following representation:

$$P(t) = P(s) + \sum_{s<t_i\le t} \varsigma(t_i)\,U_i, \tag{10}$$

i.e. it is a purely discontinuous semi-martingale. Figure 1 plots the log-prices of IBM traded on the NYSE on the 2nd of July, 2015 between 10:00am and 10:10am. The figure shows that the observed price is composed of jumps of random sizes at each transaction time and no price changes in between. This graph supports the TTSV model assumption we impose for the logarithmic prices in Equation (8).

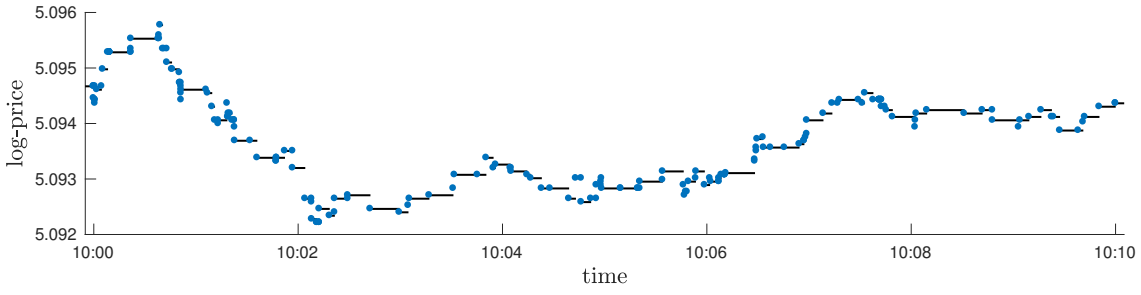

Figure 1: IBM transaction log-price data on the 2nd of July, 2015 between 10:00am and 10:10am.

The price process defined in Equation (8) has finite variation, i.e., it has almost surely bounded variation over every finite interval. Moreover, if $\{\varsigma(t)\}_{t\ge0}$ and $\{\lambda(t)\}_{t\ge0}$ are continuous processes, it holds that

$$\sigma^2_{spot}(t) = \varsigma^2(t)\,\lambda(t), \tag{11}$$

where $\left\{\sigma^2_{spot}(t)\right\}_{t\geq 0}$ is the spot variance process related to $P(t)$.[7] The variance decomposition from Equation (11) allows us to specify the integrated variance of the log-price following a TTSV model:

$$\mathrm{IV}(0,T) = \int_0^T \sigma^2_{spot}(r)\,dr = \int_0^T \varsigma^2(r)\,\lambda(r)\,dr. \tag{12}$$

Unfortunately, we are unable to specify a characteristic function for the TTSV model, even if we condition on the latent processes $\lambda(t)$ and $\varsigma(t)$. Here, even if $\varsigma(t)$ is assumed to be known, the arrival times t_i remain unknown, which is why $\varsigma(t_i)$ for some $t_i < t$ is not $\mathcal{F}^{\varsigma}_t$-measurable where $\mathcal{F}^{\varsigma}_t = \sigma(\{\varsigma(s)\}_{0\leq s\leq t})$. This difficulty can largely be overcome by exploiting the martingale property of the process $M(t) = N(t) - \int_0^t \lambda(r)\,dr$.[8]
For the theoretical results presented in Section 2.5, we define the information set generated by the latent processes, i.e., the intensity process $\{\lambda(s)\}_{0\leq s\leq t}$ and the tick volatility process $\{\varsigma(s)\}_{0\leq s\leq t}$ as follows:

$$\mathcal{F}^{\lambda,\varsigma}_t := \sigma\left(\{\lambda(s)\}_{0\leq s\leq t}, \{\varsigma(s)\}_{0\leq s\leq t}\right) \subset \mathcal{F}_t. \tag{13}$$

The variance decomposition in Equation (11) does not require a specification of the dependence structure of the tick-volatility $\varsigma(t)$ and the processes $N(t)$ and $B(t)$.[9] However, when conditioning on the latent processes, the following assumption is necessary.

Assumption 6. The process $\varsigma(t)$ is independent of $N(t)$ and $B(t)$.

Representation (10) and Assumptions 1-6 imply that for the daily return,

$$\begin{aligned}\mathbb{E}\left[r^2_{\text{daily}}\middle|\mathcal{F}^{\lambda,\varsigma}_T\right] &= \mathbb{E}\left[\left(\sum_{0<t_i\leq T}\varsigma(t_i)\,U_i\right)^2\middle|\mathcal{F}^{\lambda,\varsigma}_T\right] = \mathbb{E}\left[\int_0^T \varsigma^2(r)\,\lambda(r)\,dr\middle|\mathcal{F}^{\lambda,\varsigma}_T\right] \\ &= \mathrm{IV}(0,T).\end{aligned}$$

by 2.1. Hence, under the TTSV model, the variance of the conditional daily return equals the integrated variance.[10]

2.3 Alternative Models

In this section, we briefly describe two alternative models for the logarithmic price process, which are currently used in the realized variance literature: the standard diffusion model and the compound Poisson model.

[7] See Dahlhaus and Tunyavetchakit (2016) for a formal proof.

[8] See Appendix A for details.

[9] In fact, the tower property and the fact that $\varsigma(t)$ is $\mathcal{F}_t$-predictable suffice to proof the variance decomposition. However, when we condition on $\mathcal{F}^{\lambda,\varsigma}_t$, the tower property cannot be applied anymore.

[10] In the realized variance literature, it is common to condition on the latent variable, see e.g. Barndorff-Nielsen and Shephard (2002*a*); Oomen (2006); Andersen, Davis, Kreiß and Mikosch (2006); Hansen and Lunde (2006) among many others.

2.3.1 Standard Diffusion Model

It is commonly assumed that the logarithmic price $P(t)$ follows a continuous time diffusion model (e.g. Barndorff-Nielsen and Shephard, 2002*a*), i.e., $P(t)$ solves the stochastic differential equation:

$$dP(t) = \mu(t)\,dt + \sigma(t)\,dB(t), \quad t \in [0,T], \tag{14}$$

where $\{B(t)\}_{t\geq 0}$ is the standard Brownian motion, $\mu(t)$ is called drift term, $\sigma(t)$ denotes the instantaneous volatility and $\sigma(t)$ and $B(t)$ are independent. It is further assumed that $\mu(t)$ and $\sigma(t)$ are $\mathcal{F}_t$-predictable processes, $\mu(t)$ is of finite variation and $\sigma(t)$ satisfies $\mathbb{E}\big[\int_0^t \sigma^2(r)\,dr\big] < \infty$. For simplicity, we impose $\mu(t) = 0$, which is often done in the literature since its effect on the realized variance is of third order (Barndorff-Nielsen and Shephard, 2002*a*). In the diffusion setting, it holds that

$$\sigma^2_{spot}(t) = \sigma^2(t),$$

which implies that the integrated variance over a trading day is given by

$$\mathrm{IV}(0,T) = \int_0^T \sigma^2(r)\,dr.$$

Let now $\mathcal{F}_t^\sigma = \sigma(\{\sigma(s) : 0 \leq s \leq t\})$ denote the information set containing the sample path of the instantaneous volatility. It is readily established that:

$$r_{\text{daily}}\,|\mathcal{F}_t^\sigma \sim \mathcal{N}(0, \mathrm{IV}(0,T))$$

and thus, in the diffusion model, the conditional daily return variance corresponds to the integrated variance.[11]

A convenient feature of the standard diffusion framework is that the quadratic variation and the integrated variance of the price process $P(t)$ coincide. As realized variance converges to quadratic variation in probability (see e.g. Protter, 1990, Chapter II), it is a consistent estimator of integrated variance. Barndorff-Nielsen and Shephard (2002*a*) and Bandi and Russell (2008) show that, asymptotically, realized variance is normally distributed, which allows for statistical inference.[12]

2.3.2 Compound Poisson Process (CPP)

Alternatives to the diffusion-based price modeling are jump processes. See e.g. Cont and Tankov (2004) for an overview of mixed diffusion and jump processes as well as pure jump and subordinated processes. Press (1967, 1968) uses a compound Poisson process mixed with a Wiener process to study security price fluctuations. Oomen (2004, 2005, 2006) applies a compound Poisson process to model the logarithmic price of an asset in the realized variance context, i.e., he assumes that:

$$P(t) = P(0) + \sum_{i=1}^{N(t)} \varepsilon_i, \tag{15}$$

[11] For a formal proof and an extension to a multivariate framework, see Andersen, Bollerslev, Diebold and Labys (2003).

[12] Among others, Barndorff-Nielsen and Shephard (2005) study the finite sample behaviour of the asymptotic normality and find that it is poorly sized. Therefore, they find that the performance can be improved drastically by a log-transformation.

where $\varepsilon_i \sim$ i.i.d. $\mathcal{N}(0, \sigma_\varepsilon^2)$ and $N(t)$ is a doubly stochastic Poisson process with instantaneous stochastic intensity $\lambda(t)$, independent of the innovation process ε_i. Clearly, the logarithmic price process is modelled by a right-continuous jump process with sample paths of finite variation and therefore, it is a semi-martingale. Here, the spot variance is specified by

$$\sigma_{spot}^2(t) = \sigma_\varepsilon^2 \lambda(t), \tag{16}$$

where $\left\{\sigma_{spot}^2(t)\right\}_{t\geq 0}$ is the spot variance process related to $P(t)$ in Equation (15). It follows that the integrated variance for the CPP log-price process is given by:

$$\mathrm{IV}(0,T) = \int_0^T \sigma_{spot}^2(r)\, dr = \sigma_\varepsilon^2 \int_0^T \lambda(r)\, dr. \tag{17}$$

In a pure jump setting with a non-explosive point process, the quadratic variation differs from the integrated variance. Moreover, it becomes clear that realized variance is an inconsistent estimator of integrated variance, since increasing the sampling frequency beyond a certain point does not generate any additional information. However, Oomen (2006) shows that, conditionally on the intensity process, RV under the CPP is still an unbiased estimator of integrated variance. By means of the conditional characteristic function of the log-price increments, it can be easily shown that

$$\mathbb{E}\left[r_{\text{daily}}^2 \middle| \mathcal{F}_T^\lambda\right] = \sigma_\varepsilon^2 \int_0^T \lambda(r)\, dr = \mathrm{IV}(0,T), \tag{18}$$

where $\mathcal{F}_t^\lambda$ denotes the information set that contains the sample path of the latent process until point t, which is the intensity $\lambda(t)$ in the CPP model. Hence, as in the continuous diffusion model and the TTSV model, the conditional variance of daily returns coincides with integrated variance.

2.4 Sampling Schemes

In practice, prices are recorded whenever an event such as a transaction or a quote takes place. Hence, the price process is observed at discrete points in time, which are not necessarily equidistantly spaced. In order to obtain clock-time prices, the observed price process has to be transformed. Possible transformation procedures are the linear interpolation method, e.g. used in Andersen and Bollerslev (1997), and the previous tick sampling method proposed by Wasserfallen and Zimmermann (1985). It has been shown that for a fixed interval, realized variance converges to zero in probability under the linear interpolation method (Hansen and Lunde, 2006). The practical implication of this result is that using high-frequency data, increasing the sampling frequency leads to a downward drift in RV, as e.g. observed by Hansen and Lunde (2004). We therefore use the previous tick method for the construction of the clock-time price process.
Let $\{t_i\}_{i=1\ldots,N(T)}$ again denote the tick times in the interval $[0,T]$. Then, according to the TTSV model, the price at time $t \in [t_i, t_{i+1})$ is given by

$$P(t) = P(t_i).$$

For a sampling interval of e.g. one second, the latest tick price that occurred before the particular second defines its price. In particular, if T is the number of seconds within one trading day, the price process per second is obtained by

$$P(j) = P(\bar{t}) \quad \text{where} \quad \bar{t} = \sup\{t_i : t_i < j\}.$$

for $j = 1, \ldots, T$ and $i = 1, \ldots, N(T)$.
On the interval $[0, T]$, let again $\tau = (\tau_0, \ldots, \tau_M)$ be the general sampling (GS) grid. As shown by Oomen (2005), the type of sampling scheme can have a large impact on the quality of the realized variance estimator. Therefore, here we follow his lead and use several specific sampling schemes:

(a) Clock-time sampling (CTS), for which the sampling intervals are equidistant in calendar time, i.e., $\tau_j^{\text{CTS}} - \tau_{j-1}^{\text{CTS}}$ is constant for all $j = 1, \ldots, M$. The logarithmic price process is sampled at the times

$$\tau_j^{\text{CTS}} = j\Delta_{\text{CTS}}, \quad j = 1, \ldots, M,$$

where $\Delta_{\text{CTS}} := {}^{T}/_{M}$ is assumed to be integer-valued.

Its simple implementation and intuition make it the most widespread sampling scheme in the empirical finance literature so far. However, as it relies on an artificial clock-time price process, it neglects information on intraday trading as well as volatility patterns. Nevertheless, due to its wide application, we use it as a benchmark sampling scheme.

(b) Business time sampling (BTS), where the points at time τ_j^{BTS} are chosen such that the integrated variance is equal across all M subintervals, i.e., τ_j^{BTS} such that

$$\text{IV}\left(\tau_{j-1}^{\text{BTS}}, \tau_j^{\text{BTS}}\right) = \frac{1}{M}\,\text{IV}(0, T).$$

This means that the sampling intervals are "equidistant" in terms of integrated variance. Note that this definition of the sample grid is infeasible in practice since the integrated variance is unobservable. However, it is commonly approximated by the sum of all available intra-daily squared returns, i.e., for $j = 1, \ldots, M$, the τ_j^{BTS} are chosen such that

$$\text{RV}_{\text{all}}\left(\tau_{j-1}^{\text{BTS}}, \tau_j^{\text{BTS}}\right) \approx \frac{1}{M}\,\text{RV}_{\text{all}}(0, T) := \Delta_{\text{BTS}},$$

where RV_{all} denotes the realized variance using all available intra-daily returns in the respective interval. We set $\tau_0^{\text{BTS}} = 0$ and

$$\tau_j^{\text{BTS}} = \sup\left\{t_i \in \left(\tau_{j-1}^{\text{BTS}}, T\right] : \text{RV}_{\text{all}}(0, t_i) \leq j\Delta_{\text{BTS}}\right\}, \quad j = 1, \ldots, M,$$

where t_i are again the tick-times.

(c) Transaction time sampling (TTS), for which the data is sampled "equidistantly" in terms of number of transactions. Let $N(T) \in \mathbb{N}$ denote the cumulated number of transactions in the interval $[0, T]$. Then, under TTS, the sampled points in time τ_j^{TTS} for $j = 0, \ldots, M$ are chosen such that

$$N\left(\tau_j^{\text{TTS}}\right) - N\left(\tau_{j-1}^{\text{TTS}}\right) = \frac{N(T)}{M} =: \Delta_t$$

which for simplicity is assumed to be integer-valued. In terms of the transaction times $t_i, i = 0, \ldots, N(T)$, we have that

$$\tau_j^{\text{TTS}} = t_{j\Delta_{\text{TTS}}}, \quad j = 1, \ldots, M.$$

This sampling scheme is a natural one when the price process is observed at each transaction and, differently from CTS, it takes the information about the intraday trading activity into account.[13]

(d) Tick-time sampling (TiTS), for which the number of intervals is set to the number of ticks, i.e., $M = N(T)$. Since t_i denotes the time at which tick i occurs, one gets that

$$\tau_j^{\text{TiTS}} = t_j, \quad j = 1, \ldots, N(T)$$

and, therefore, TiTS is a special case of TTS.

There exist numerous other sampling schemes that can be applied in a similar way. For instance, Engle and Russell (1998) introduce a sampling based on the absolute price change, Clark (1973) introduces sampling based on trading volume and Wu (2012) based on the average number of transactions over past days.

Figure 2 shows a snippet of the grid of TTS, BTS, and CTS. The graph shows a simulated log-price process plotted against the number of transactions. The underlying transaction-time process is taken from the IBM stock on the 2$^{\text{nd}}$ of July, 2015 during the trading hours from 9:30am to 4:00pm. The plot depicts roughly the first half of the trading day, since on that day, we have $N(T) = 3872$ transactions. Naturally, the TTS sampled points are equidistant in terms of the number of transactions. For CTS, the sampled points are more widespread at the beginning and more frequent for the second half of the plot. Since the trading activity is typically high at the beginning of a trading day (see Figure 3), this implies that CTS misses important intra-day trading activity. Similarly, BTS samples less often at the beginning and more often towards the end of the snippet. Since the scale is equidistant in the number of transactions, it is, loosely speaking, equidistant in trading intensity. Hence, changes in spot volatility explain the frequency of sampled points for BTS, which implies that in this snippet, the spot volatility is lower for the first 1000 transactions than afterwards.

Figure 4 shows the monthly average sampling frequency per minute for the EUR USD exchange rate (upper panel) and IBM stock price (lower panel). In the case of the EUR USD foreign exchange rate, trading takes place 24 hours on different markets around the world. We observe that at the beginning of the trading day, TTS and BTS sample less often than CTS while when the European and American markets are open, TTS and BTS sample more often than CTS. Contrary, IBM is traded on the NYSE which is only open from 9:30am to 04:00pm (GMT-5). In the beginning of the trading day, BTS peaks and samples by far more than TTS and CTS while at the end of the trading day, TTS shows a similar peak.

From Equation (11) we know that in the TTSV framework, the spot variance splits up into the product of innovation variance $\varsigma^2(t)$ and the intensity $\lambda(t)$ of the point

[13] Instead of transactions, one can also use quotes.

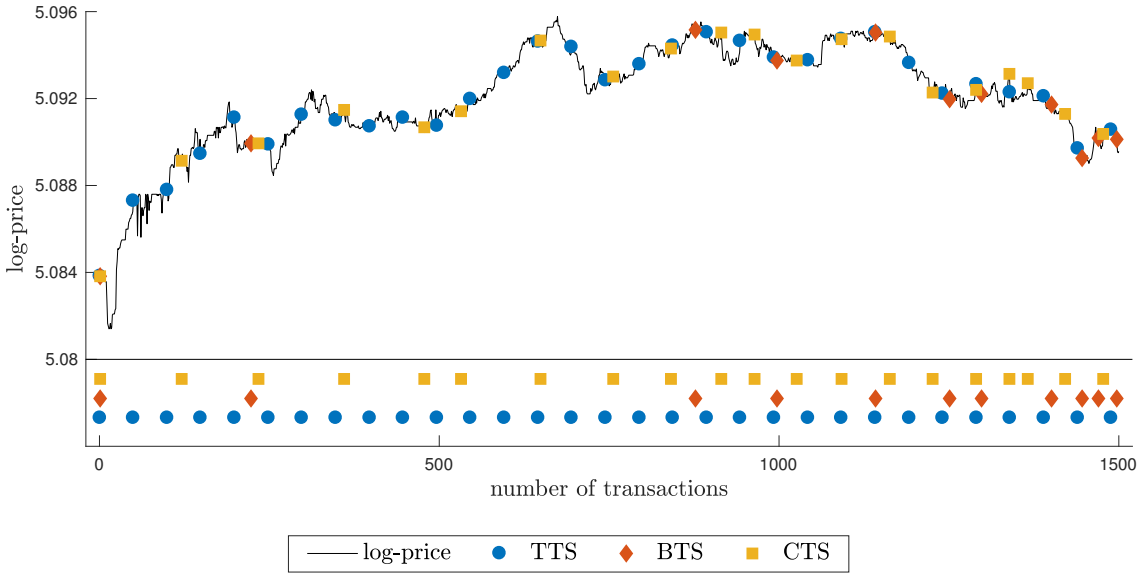

Figure 2: Sampling grid comparison. We plot a snippet of the TTS, BTS, and CTS grid for a simulated log-price process with transaction times taken from the IBM stock on the 2nd of July, 2015 during the trading hours from 9:30am to 4:00pm. The data was simulated according to Simulation 1 (see Tables 1 and 2).

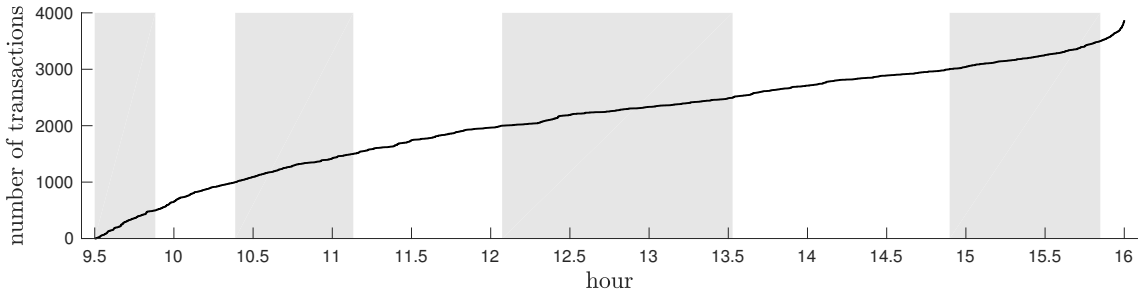

Figure 3: Number of transactions as a function of time for IBM. The data is taken from the IBM stock on the 2nd of July, 2015 during the trading hours from 9:30am to 4:00pm. The shaded and non-shaded regions correspond to 500 transactions.

process. Since BTS samples based on $\sigma^2_{spot}(t) = \varsigma^2(t)\,\lambda(t)$ and TTS based on the frequency $\lambda(t)$, Figure 4 indicates that a constant variance σ^2_ε as in the CCP framework is not a reasonable assumption for the IBM stock. In fact, the variance of the return innovations seems to be time-varying, explaining the differences between BTS and TTS. For the EUR USD exchange rate on the other hand, a constant innovation variance seems plausible since TTS and BTS appear to sample similarly. Oomen (2005) shows that in the CPP setting and in the absence of market microstructure noise, the MSE is smaller for TTS and BTS than for CTS. Moreover, he finds that the efficiency gaps are positively related to the innovation variance σ^2_ε and the trading intensity. More precisely, the larger the trading intensity or the innovation variance, the larger is the efficiency gain. In the next two sections, we undergo a similar analysis on the properties of the RV estimator under the three sampling schemes and the TTSV assumption.

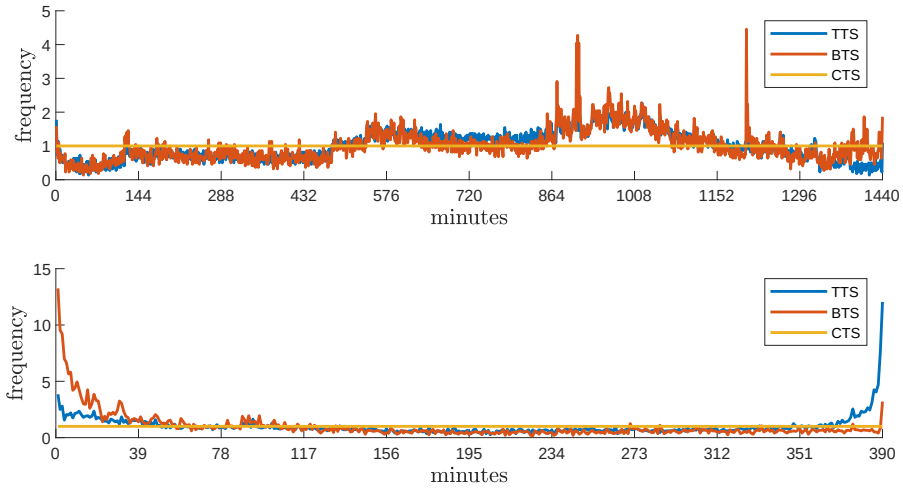

Figure 4: Monthly average sampling frequency per minute.. The underlying data are prices from the EUR USD exchange rate in June, 2016 (upper panel) and the IBM stock in January, 2015 (lower panel) during the respective trading hours (24 hours for EUR USD and 9:30am to 4:00pm for IBM).

2.5 Theoretical Results

I this section, we derive statistical properties of RV under the TTSV framework specified in Equation (8).

Proposition 2.1. Assume that Assumptions 1-6 hold true. Then,

$$\mathbb{E}\left[\mathrm{RV}_M(0,T) - \mathrm{IV}\,(0,T)\middle|\mathcal{F}_T^{\lambda,\varsigma}\right] = 0,$$

i.e., RV conditional on the latent processes is an unbiased estimator of integrated variance.

Similar to Oomen (2006), we derive below the MSE of RV as a function of integrated variance and integrated quarticity. For simplicity, we make a slight abuse of notation and define $\lambda\,(s,t) := \int_s^t \lambda\,(r)\,dr$ for $0 \leq s < t$. Moreover, regarding higher moments we make the following two assumptions:

Assumption 7. The 4-th moment of the tick volatility $\varsigma\,(t)$ exists, i.e., for all $0 \leq t \leq T$, $\mathbb{E}\left[\varsigma^4\,(t)\right] < \infty$.

Assumption 8. It holds that $\mathbb{E}\left[\int_0^t \varsigma^4\,(r)\,\lambda\,(r)\,dr\right] < \infty$ for all $0 \leq t \leq T$.

Proposition 2.2. Let Assumptions 1-8 hold true. Then,

$$\begin{aligned}\mathrm{MSE}_M\,(\mathrm{GS}) &= \mathbb{E}\left[\left(\mathrm{RV}_M(0,T) - \mathrm{IV}(0,T)\right)^2\middle|\mathcal{F}_T^{\lambda,\varsigma}\right] \\ &= 2\sum_{j=1}^{M}\mathrm{IV}\,(\tau_{j-1},\tau_j)^2 + 3\,\mathrm{IQ}\,(0,T)\,,\end{aligned}$$

where $\mathrm{IQ}\,(s,t) := \int_s^t \varsigma^4\,(r)\,\lambda\,(r)\,dr$ defines the integrated quarticity.

From Proposition 2.2, we obtain a first intuition on what drives the magnitude of the MSE. In particular, we see that the MSE is bounded from below by the constant factor $3\,\mathrm{IQ}\,(0,T)$. Moreover, for different sampling schemes, the first and second summand of the MSE may vary, i.e., the MSE across sampling schemes differs.
The results of this section reveal many of the similarities between the TTSV model and the CPP log-price model. Both log-price processes are càdlàg, have finite variation and stochastic jumps at each tick. Additionally, in both models, realized variance can be seen as a inconsistent estimate of integrated variance since the number of ticks within a finite interval is bounded by construction. The main difference is that under the TTSV model, the tick volatility $\varsigma\,(t)$ is stochastic and time-varying whereas it is constant under the CPP log-price model. In fact, the CPP model is a special case of TTSV model. Therefore, under the CPP model, the result of Proposition 2.2 is in fact the same.[14]
Similarly to Oomen (2005) we start by investigating the MSE for specific sampling schemes. To be more precise, we compare the MSE of TTS, BTS, and CTS theoretically and empirically for both real and simulated data.
The expression given in Proposition 2.2 can be further simplified under the BTS scheme. Following the definition of BTS, $\mathrm{IV}\left(\tau_{j-1}^{\mathrm{BTS}},\tau_{j}^{\mathrm{BTS}}\right)=\mathrm{IV}(0,T)/M$ and therefore

$$\mathrm{MSE}_M\,(\mathrm{BTS})=\frac{2}{M}\,\mathrm{IV}\,(0,T)^2+3\,\mathrm{IQ}\,(0,T)\,.$$

By the Cauchy-Schwarz inequality, it follows that

$$\mathrm{MSE}_M\,(\mathrm{GS})-\mathrm{MSE}_M\,(\mathrm{BTS})=2\sum_{j=1}^{M}\mathrm{IV}\,(\tau_{j-1},\tau_j)^2-\frac{2}{M}\,\mathrm{IV}\,(0,T)^2\geq 0,$$

implying that the MSE for BTS is lower than for any other sampling scheme. For $M\to\infty$, the MSE under BTS is bounded from below by $3\,\mathrm{IQ}\,(0,T)$. Note that after M has been increased to a certain point, BTS will coincide with TiTS. Therefore, for $M\to\infty$, we have $\mathrm{MSE}_M\,(\mathrm{BTS})\to\mathrm{MSE}\,(\mathrm{TiTS})$. However, due to the latency of integrated variance and due to the discrete nature of the tick-times, RV under BTS contains some approximation error. Therefore, the MSE for the specific sampling schemes implemented here, including the BTS for which the IV measure is estimated, is investigated empirically in Section 3.
Naturally, for TTS, the sampling frequency M cannot be increased beyond the total number of ticks $N\,(T)$ and thus

$$\begin{aligned}\mathrm{MSE}_M\,(\mathrm{TTS})&=3\,\mathrm{IQ}\,(0,T)+2\sum_{j=1}^{M}\mathrm{IV}\left(t_{(j-1)\Delta_t},t_{j\Delta_t}\right)^2\\&\geq 3\,\mathrm{IQ}\,(0,T)+2\sum_{i=1}^{N(T)}\mathrm{IV}\,(t_{i-1},t_i)^2\\&=\mathrm{MSE}\,(\mathrm{TiTS})\end{aligned}$$

[14]However, note that under the CPP model, $\mathrm{IV}\,(s,t):=\sigma_\varepsilon^2\int_s^t\lambda\,(r)\,dr$ and $\mathrm{IQ}\,(s,t):=\sigma_\varepsilon^4\int_s^t\lambda\,(r)\,dr$.

with equality for $M = N(T)$ and where t_i denote the tick-times with $t_0 = 0$ by convention. From a theoretical perspective and in the absence of market microstructure noise, the optimal sampling frequency under TTS is $M = N(T)$.
The comparison of the MSE under TTS and BTS is different from that by Oomen (2004, 2006) due to the time varying tick volatility $\varsigma(t)$. In his setting, BTS is based on the expected number of ticks while TTS is based on the realized number of ticks. This is not the case for the TTSV model, where BTS also accounts for intra-daily fluctuations of the tick volatility.

2.6 Market Microstructure

So far we have ignored possible measurement errors that occur when observing the price process. In practice however, we have to account for market microstructure noise (MMN). MMN may arise for several reasons such as price discreteness (Harris, 1990, 1991), bid-ask bounces, or infrequent trading (e.g. Madhavan, 2000), among others.
A popular way of visualizing the bias resulting from MMN is by means of volatility signature plots (Andersen et al., 2000*b*). Let $\mathrm{RV}_M^i(0,T)$ be the RV estimate on the i-th trading day using M observations. The volatility signature plot displays the sample average of the RV_M^i estimates over a sample of $n \in \mathbb{N}$ trading days as a function of the sampled number of observations M, i.e.,

$$\overline{RV}(M) = \frac{1}{n}\sum_{i=1}^{n} \mathrm{RV}_M^i(0,T).$$

Figure 5 depicts the volatility signature plot and the variation of RV_M^i using IBM transaction data during the daily trading hours from 9:30am to 4:00pm from 2015 until 2017 under the three different sampling schemes. Without noise, we expect the RV estimate to stabilize as the number of observations M increases. However, we observe a rather sharp increase under CTS for $M > 250$ which approximately corresponds to a clock-time sampling frequency of 90 seconds. The average RV estimate under TTS and BTS increases only slightly, indicating that the two sampling schemes are much less sensitive to market microstructure noise. Regarding the sample variance of RV, under TTS and BTS, the variance appears to be the lowest for low sampling frequencies and the highest for a frequency of $M \approx 120$ observations. Under CTS, the variance increases with an increasing number of observations M. Ideally, we want to find the number of sampled observations such that the RV estimate remains stable and such that the variance of the RV estimate is the lowest.
In the following, we theoretically investigate the bias under market microstructure noise assuming that the TTSV model is the underlying data generating process. We now suppose that we observe the logarithmic price process $\tilde{P}(t)$ which follows the decomposition

$$\tilde{P}(t) = P(t) + v_i \text{ for } t \in [t_i, t_{i+1})$$

where $P(t)$ follows the TTSV model in (8), v_i is a noise component yet to be specified and t_i denotes the time of the i-th tick. This implies a moving average (MA) structure for the error term of the log-return process, i.e.,

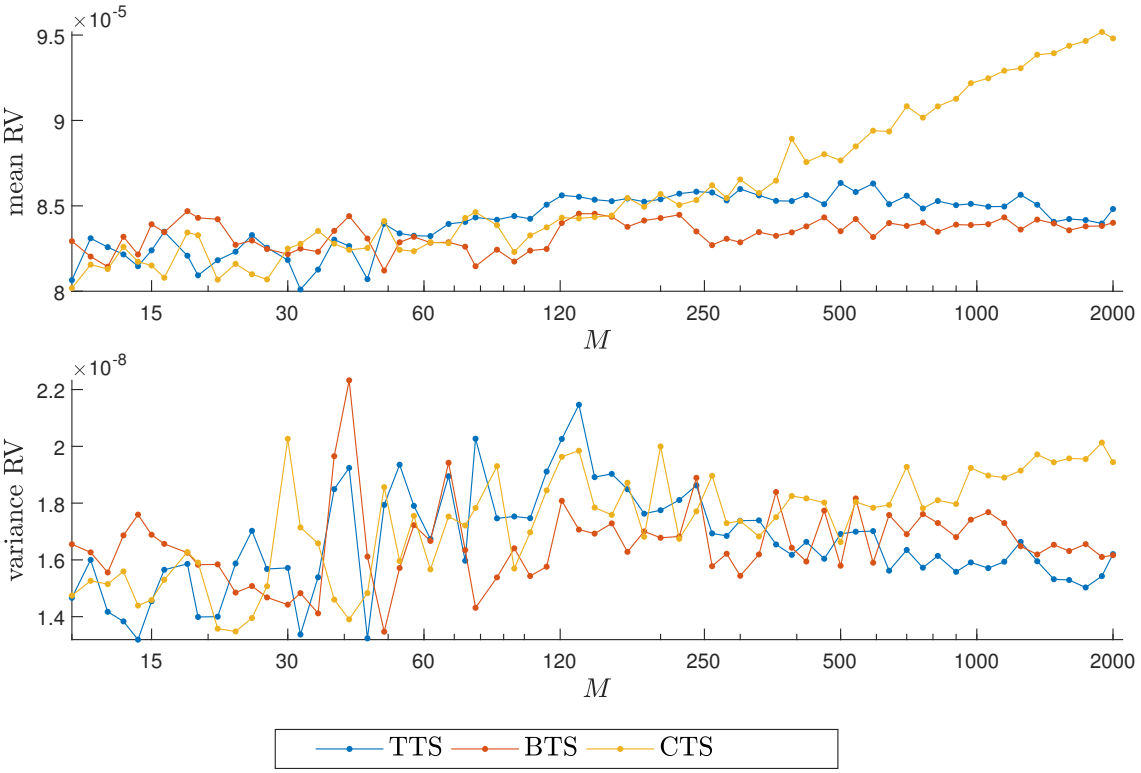

Figure 5: Volatility signature plot and variance of RV for IBM. We plot the mean RV and the variance of RV as a function of the sampling frequency M using IBM transaction time data from 2015 to 2017. The number of observations M is plotted on a log-scale.

$$\tilde{r}_i := \tilde{P}(t_i) - \tilde{P}(t_{i-1}) = r_i + v_i - v_{i-1} = r_i + u_i,$$

where $u_i = v_i - v_{i-1}$. Henceforth, with v_i and u_i we associate the log-price noise and log-return noise corresponding to the i-th tick-time t_i.

For illustration purposes, lets consider a grid τ other than the TTS grid: e.g., τ be the clock-time grid where τ_j now corresponds to the j-th minute within a trading day. Using this grid, 390 log-returns are sampled and used to compute the realized variance. If there were, say $m - n \in \mathbb{N}$ trades in the interval $(\tau_{j-1}, \tau_j]$ with corresponding tick-times $t_{n+1}, \ldots, t_m$, the noisy log-price process is given by

$$\tilde{P}(\tau_j) = \tilde{P}(t_m) = P(t_m) + v_m = P(\tau_j) + v_m$$

and for the log-return process we have that

$$\begin{aligned}\tilde{r}_j &= \tilde{P}(\tau_j) - \tilde{P}(\tau_{j-1}) = \tilde{P}(t_m) - \tilde{P}(t_n) \\ &= P(t_m) - P(t_n) + v_m - v_n = P(\tau_j) - P(\tau_{j-1}) + v_m - v_n \\ &= r_j + v_m - v_n. \end{aligned} \tag{19}$$

Hence, the bias of the realized variance will depend on the covariance structure of the noise process. Generally speaking, if there is just one trade within a minute, the covariance term of the noise is larger in absolute terms than if there were 100 trades within that minute.

We now move from this simple example to a more formal approach. For the general sampling grid $\tau = (\tau_0, \ldots, \tau_T)$, the structure of the noise depends on how many

trades happen within each interval. Let $N(\tau_j)$ now denote the number of trades that occur until τ_j. Then, the noisy log-price process is given by

$$\tilde{P}(\tau_j) = \tilde{P}\left(t_{N(\tau_j)}\right) = P\left(t_{N(\tau_j)}\right) + v_{N(\tau_j)}$$

and the corresponding noisy log-return by

$$\begin{aligned}
\tilde{r}_j &= \tilde{P}(\tau_j) - \tilde{P}(\tau_{j-1}) \\
&= \tilde{P}\left(t_{N(\tau_j)}\right) - \tilde{P}\left(t_{N(\tau_{j-1})}\right) \\
&= P\left(t_{N(\tau_j)}\right) - P\left(t_{N(\tau_{j-1})}\right) + v_{N(\tau_j)} - v_{N(\tau_{j-1})} \\
&= P(\tau_j) - P(\tau_{j-1}) + v_{N(\tau_j)} - v_{N(\tau_{j-1})} \\
&= r_j + v_{N(\tau_j)} - v_{N(\tau_{j-1})}.
\end{aligned}$$

If no trade has occured in the interval $(\tau_{j-1}, \tau_j]$, then $\tilde{P}(\tau_j) = \tilde{P}(\tau_{j-1})$ and therefore $v_{N(\tau_j)} = v_{N(\tau_{j-1})}$. It follows that

$$\begin{aligned}
\widetilde{\mathrm{RV}}_M(0,T) &= \sum_{j=1}^{M} \tilde{r}_j^2 = \sum_{j=1}^{M} \left(r_j + v_{N(\tau_j)} - v_{N(\tau_{j-1})}\right)^2 \\
&= \mathrm{RV}_M(0,T) + 2\sum_{j=1}^{M} r_j \left(v_{N(\tau_j)} - v_{N(\tau_{j-1})}\right) + \sum_{j=1}^{M} \left(v_{N(\tau_j)} - v_{N(\tau_{j-1})}\right)^2.
\end{aligned}$$

From this expression it is clear that even in the simplest case when v_i is a white noise process, the realized variance is now a biased estimator of IV. If instead we were in a diffusive price setting, the third summand would actually diverge. More precisely, Bandi and Russell (2008) show that the RV estimate under additive log-price market microstructure noise diverges $\mathbb{P}$-a.s. as the sampling frequency M goes to infinity. A distinct difference in the TTSV setting is that when increasing M, the accumulation of noise is bounded since the overall number of ticks $N(T)$ is bounded.

Assumption 9. For the noise process $\{v_j\}_{j\in\mathbb{N}}$ we make the following assumptions:

(N-1) $\mathbb{E}[v_j] = 0$ for all $j \in \mathbb{N}$.
(N-2) $\mathrm{cov}(v_j, v_{j+l}) = \gamma_v(l) < \infty$ for all $j, l \in \mathbb{N}$.
(N-3) v_j and $P(t)$ and $N(t)$ are independent for all $j \in \mathbb{N}$ and $t \geq 0$.

Properties (N-1) and (N-2) of Assumption 9 imply that v_j is a covariance stationary process. Examples of covariance stationary processes are white noise processes, i.e., independent noise with $\gamma_v(j) = 0$ for all $j \neq 0$, and autoregressive moving average processes with suitable parameter restrictions.

Lemma 2.3. Assume that $P(t)$ satisfies Assumptions 1-8 and that the noise process $\{v_j\}_{j\in\mathbb{N}}$ satisfies Assumption 9. Then,

$$\mathbb{E}\left[\widetilde{\mathrm{RV}}_M(0,T) - \mathrm{IV}(0,T)\middle|\mathcal{F}_T^{\lambda,\varsigma}\right] = 2M\gamma_v(0) - 2\sum_{j=1}^{M} \mathbb{E}\left[v_{N(\tau_j)} v_{N(\tau_{j-1})}\middle|\mathcal{F}_T^{\lambda,\varsigma}\right],$$

i.e., realized variance is a biased estimator of IV.

While the first bias term $2M\gamma_v(0)$ is the same regardless of the sampling scheme, the second may vary drastically. For TTS, $N(\tau_j) - N(\tau_{j-1}) = j\Delta_t - (j-1)\Delta_t = \Delta_t$ is constant and therefore the second bias term simplifies to a particular covariance function of v_i, i.e., $-2M\mathbb{E}[v_{\Delta_t} v_0 | \mathcal{F}_T^{\lambda,\varsigma}]$ which cannot be further specified since $N(T)$ and thus also Δ_t remain random. For other sampling schemes, the bias term may include covariance terms of v_i of high and low lag order. In a sense, this can be seen as a disadvantage of TTS, because e.g. for a decreasing autocovariance function, covariance terms of low order are larger and, thus, can offset the common bias term $2M\gamma_v(0)$.

Lemma 2.4. Assume that $P(t)$ satisfies Assumptions 1-8 and that the noise process $\{v_j\}_{j\in\mathbb{N}}$ satisfies Assumption 9. Then, under GS:

$$\begin{aligned}\widetilde{\mathrm{MSE}}_M(\mathrm{GS}) &= \mathbb{E}\left[\left(\widetilde{\mathrm{RV}}_M(0,T) - \mathrm{IV}(0,T)\right)^2 \middle| \mathcal{F}_T^{\lambda,\varsigma}\right] \\ &= \mathrm{MSE}_M(\mathrm{GS}) + \mathbb{E}\left[\left(\sum_{j=1}^{M}\left(v_{N(\tau_j)} - v_{N(\tau_{j-1})}\right)^2\right)^2 \middle| \mathcal{F}_T^{\lambda,\varsigma}\right] \\ &\quad + 8\left(\mathrm{IV}(0,T)\,\gamma_v(0) - \sum_{j=1}^{M}\mathbb{E}\left[r_j^2 \middle| \mathcal{F}_T^{\lambda,\varsigma}\right]\mathbb{E}\left[v_{N(\tau_j)} v_{N(\tau_{j-1})} \middle| \mathcal{F}_T^{\lambda,\varsigma}\right]\right),\end{aligned}$$

where $\mathrm{MSE}_M(\mathrm{GS})$ is given by Proposition 2.2.

Lemma 2.4 implies that the MSE under noise increases for every sampling scheme. If $v_{N(\tau_j)}$ and $v_{N(\tau_{j-1})}$ are independent, then BTS remains the sampling scheme with the lowest MSE. Generally, we cannot rank the considered sampling schemes due to the unkown error terms. We empirically investigate the bias and its corresponding MSE under noise and for several sampling schemes in Section 3. In the following two sections, we investigate the properties of two RV estimators that are biased-corrected based on information on the autocorrelation (AC) of the noise.

2.6.1 AC(1) Bias-Correction

We next consider a bias-corrected realized variance estimator as introduced by Zhou (1996), i.e.,

$$\widetilde{\mathrm{RV}}_M^{\mathrm{AC}(1)}(0,T) = \sum_{j=1}^{M}\tilde{r}_j^2 + \sum_{j=1}^{M}\tilde{r}_j\tilde{r}_{j-1} + \sum_{j=1}^{M}\tilde{r}_j\tilde{r}_{j+1}. \tag{20}$$

Note that the bias-corrected version involves two returns that do not lie within the interval $[0,T]$. Empirically, this problem can be circumvented by instead considering the estimator $\sum_{j=1}^{M}\tilde{r}_j^2 + \sum_{j=2}^{M}\tilde{r}_j\tilde{r}_{j-1} + \sum_{j=1}^{M-1}\tilde{r}_j\tilde{r}_{j+1}$ (Hansen and Lunde, 2006) since the additional bias decreases as the sampling frequency increases. For the following results, we stick to the version in (20) for simplicity. The idea of the bias-correction is to offset the constant bias term $2M\gamma_v(0)$ and add covariance terms of lag order greater than zero, which are potentially of smaller magnitude.

Lemma 2.5. Assume that $P(t)$ satisfies Assumptions 1-6 and that the noise process $\{v_j\}_{j\in\mathbb{N}}$ satisfies Assumption 9. Then,

$$\mathbb{E}\left[\widetilde{\mathrm{RV}}_M^{\mathrm{AC}(1)}(0,T) - \mathrm{IV}(0,T)\middle|\mathcal{F}_T^{\lambda,\varsigma}\right] = \sum_{j=1}^{M}\mathbb{E}\left[-v_{N(\tau_j)}v_{N(\tau_{j-2})} + v_{N(\tau_{j-1})}v_{N(\tau_{j-2})} + v_{N(\tau_j)}v_{N(\tau_{j+1})} - v_{N(\tau_{j-1})}v_{N(\tau_{j+1})}\middle|\mathcal{F}_T^{\lambda,\varsigma}\right],$$

where τ_{-1} is the last sampling point of the previous day and τ_{M+1} is the first sampling point of the following day.[15]

Again, we may simplify the bias of the bias-corrected realized variance estimator under TTS to

$$\mathrm{bias}_M(\mathrm{TTS}) = 2M\mathbb{E}\left[v_{\Delta_t}v_0 - v_{2\Delta_t}v_0\middle|\mathcal{F}_T^{\lambda,\varsigma}\right].$$

It is easy to see that its magnitude depends on the number of sampled points M as well as on the dependence structure of v_i. If, e.g., v_i has an MA structure, the autocovariance function decreases with increasing lag size. This implies that the bias under TTS is positive and that if M is increased, the effect of the noise and, thus, the bias is increased as well.
Figure 6 depicts the volatility signature plot of the same IBM transaction data as above using the RV with the AC(1) bias-correction. We observe that the increase of the average RV estimate under CTS is not as sharp as without bias-correction. Still, for all sampling schemes, the mean RV increased with increased sampling frequency.

Example 2.6. We impose that v_j is a white noise process, i.e., additionally to Assumption 9, we assume $\gamma_v(l) = 0$ for all $l \geq 1$. In fact, in this case the noise of the log-return process $\tilde{r}_j$ is a MA(1) process

$$u_j = v_j + \theta_1 v_{j-1}, \quad j \in \mathbb{N}$$

with $\theta_1 = -1$. Given that in each sampling interval $(\tau_j, \tau_{j+1}]$ there exists at least one tick, $\widetilde{\mathrm{RV}}_M^{\mathrm{AC}(1)}(0,T)$ is an unbiased estimator of integrated variance for every sampling scheme. The necessary condition is always fulfilled under TTS since naturally, one chooses $M \leq N(T)$.

Empirically, it has been observed that the noise process possesses negative first-order autocorrelation. For instance, Bandi and Russell (2008) investigate IBM quotes and find significant negative first order autocorrelation. Bandi and Russell (2006) examine S&P500 stocks and find highly significant first order serial correlation and some significant serial correlation up to lag four. We therefore consider a MA(1) noise process in our next example.

Example 2.7. We assume v_j follows a MA(1) process, i.e.,

$$v_j = \nu_j + \theta_1 \nu_{j-1} \quad j \in \mathbb{N}$$

[15]Note that we do not rely on data outside the interval $[0,T]$ since $P(\tau_{-1}) = P(0)$ and $P(\tau_{M+1}) = P(T)$.

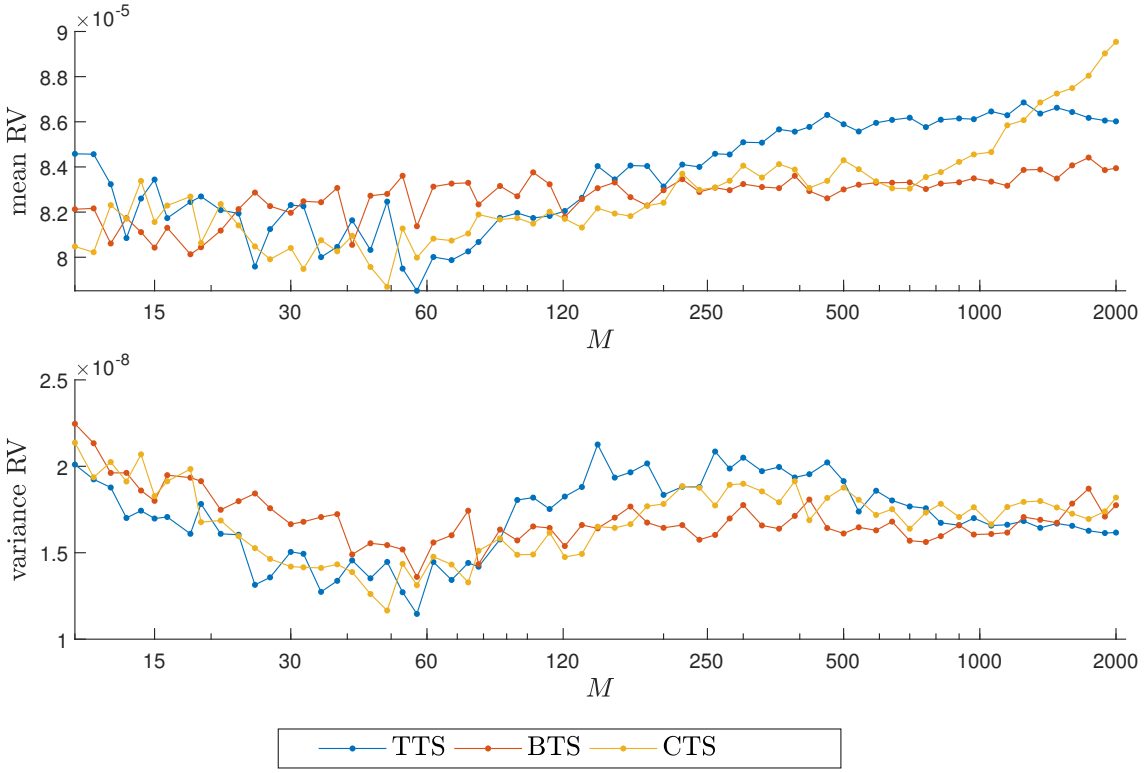

Figure 6: Volatility signature plot and variance of RV for IBM using bias correction. We plot the mean RV and variance of RV as a function of the number of observations M using IBM transaction time data from 2015 to 2017. The RV is AC(1) bias-corrected. The number of observations M is plotted on a log-scale.

where ν_j is a white noise process with variance σ_ν^2. Clearly, Assumption 9 is satisfied and

$$\gamma_v(l) = \begin{cases} (1+\theta_1^2)\,\sigma_\nu^2 & \text{if } l = 0, \\ \theta_1 \sigma_\nu^2 & \text{if } l = 1, \\ 0 & \text{if } l > 1. \end{cases}$$

For the AC(1) bias-corrected RV estimate under TTS, this implies that

$$\text{bias}_M(\text{TTS}) = 0$$

as long as $\Delta_t = {}^{N(T)}/_{M} \geq 2$. Hence, the number of sampled points M should not exceed $1/2$ of the total number of trades during a day. Figure 7 shows the bias and MSE for the AC(1)-corrected RV estimate for TTS, BTS, and CTS. The number of observations M is plotted on a log scale. The data is simulated using the procedure described in Section 3.1 using the parameters from Table 1. The transaction times are taken from the IBM stock on the 1st of May, 2015 during the trading hours from 9:30am to 4:00pm. On that day, the number of transactions amount to $N(T) = 2093$. Hence, for sampling frequencies $M > \lfloor {}^{2093}/_{2} \rfloor = 1046$, the bias under TTS is different from zero and negative. In Figure 7, the last seven points represent the sampling frequencies $M = \{1000, 1138, 1309, 1500, 1731, 1991, 2093\}$. The downward slope of the bias and upward slope of the MSE start after $M = 1000$, which is expected.

The auto-covariance function of v_i also implies that TiTS induces a bias since $M = N(T)$. Unfortunately, for the other sampling schemes, we cannot make similar statements. If, e.g., the price process is sampled in clock-time and there is an interval $(\tau_j, \tau_{j+1}]$ in which only one trade happens, then,

$$\mathbb{E}\left[v_{N(\tau_j)} v_{N(\tau_{j+1})} \middle| \mathcal{F}_T^{\lambda,\varsigma}\right] > 0$$

and the AC(1)-corrected RV estimate remains biased.

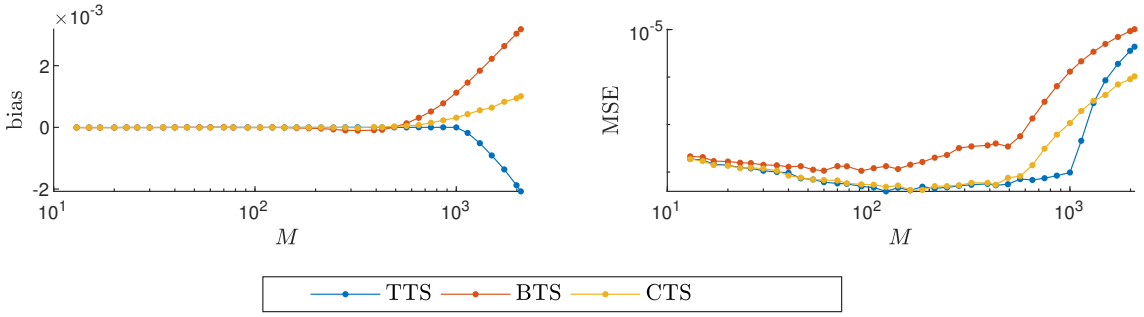

Figure 7: Example of Bias and MSE under MA(1) noise. We plot the bias and MSE of the AC(1)-corrected RV estimate for different sampling schemes under market microstructure noise. The noise follows an MA(1) process $v_i = \nu_i + \theta_1 \nu_{i-1}$ with $\theta_1 = -0.9$ and $\nu_i \sim$ i.i.d. $\mathcal{N}\left(0, \sigma_\nu^2\right)$ with $\sigma_\nu^2 =$ 1e−6. The number of observations M is plotted on a log scale. The data was simulated with the parameters in Table 1. The transaction times were taken from the IBM stock on the 1st of May, 2015 during the trading hours from 9:30am to 4:00pm. The number of transactions on that day amounts to $N(T) = 2093$.

2.6.2 AC(p) Bias-Correction

Hansen and Lunde (2006) examine signature plots for Alcoa (AA) and Microsoft (MSFT) stock data from the years 2000 and 2004 using both CTS and TTS. They find that the AC(1)-corrected RV estimator in (20) is still significantly biased under CTS when the sampling frequency is large.[16] This is due to the fact that the same price is sampled multiple times which introduces artificial autocorrelation. Moreover, the same phenomenon is observed for TTS for step sizes Δ_t smaller than 4 ticks for AA and smaller than 14 ticks for MSFT with some variation across the considered years. This indicates that the market microstructure noise has significant autocorrelation beyond the first lag. We therefore generalize the bias-corrected RV estimator in what follows.

Inspired by Newey and West (1987), Zhou (1996) introduces a bias-corrected realized variance estimator based on the first p autocovariance terms of the form

$$\widetilde{\mathrm{RV}}_M^{\mathrm{AC}(p)}(0,T) = \sum_{j=1}^{M} \tilde{r}_j^2 + \sum_{k=1}^{p} \sum_{j=k}^{M} \tilde{r}_j \tilde{r}_{j-k} + \sum_{k=1}^{p} \sum_{j=1}^{M} \tilde{r}_j \tilde{r}_{j+k}. \tag{21}$$

This bias-corrected version is also investigated by Oomen (2005) under CPP and by Hansen and Lunde (2006) in a diffusion setting.

[16] Hansen and Lunde (2006) use confidence bounds constructed from the average of RV estimates based on transaction prices, bid prices, and ask prices.

As outlined above, we additionally include noise with higher order serial correlation. Therefore, we now consider noise that follows an MA(q) process, i.e.,

$$v_i = \nu_i + \theta_1 \nu_{i-1} + \ldots + \theta_q \nu_{i-q}, \tag{22}$$

where ν_i is a white noise process with variance $\sigma_\nu^2 := \mathbb{E}\left[\nu_i^2\right] < \infty$.

Lemma 2.8. Assume that Assumptions 1-8 hold true and that $\{v_j\}_{j\in\mathbb{N}}$ is an MA(q) process as specified in (22). Then, for the AC (p)-corrected RV from Equation (21) it holds that

$$\mathbb{E}\left[\widetilde{\mathrm{RV}}_M^{\mathrm{AC}(p)}(0,T) - \mathrm{IV}\,(0,T)\middle|\mathcal{F}_T^{\lambda,\varsigma}\right] = \sum_{j=1}^{M} \mathbb{E}\left[v_{N(\tau_{j-1})}v_{N(\tau_{j-1-p})} + v_{N(\tau_j)}v_{N(\tau_{j+p})} + v_{N(\tau_j)}v_{N(\tau_{j-1-p})} + v_{N(\tau_{j-1})}v_{N(\tau_{j+p})}\middle|\mathcal{F}_T^{\lambda,\varsigma}\right].$$

The bias expression in Lemma 2.8 involves sampled points that potentially lie outside the interval $[0,T]$, e.g., $\tau_{-p} < \tau_0 = 0$ or $\tau_{M+p} > \tau_M = T$. Since the total number of ticks on a trading day $N\,(T)$ varies across days, the TTS step size $\Delta_t = {}^{N(T)}/_{M}$ varies as well. Therefore, when considering TTS, the simplification used above is no longer valid in general, namely,

$$\mathbb{E}\left[v_{N(\tau_j)}v_{N(\tau_{j-1})}\middle|\mathcal{F}_T^{\lambda,\varsigma}\right] \neq \mathbb{E}\left[v_{\Delta_t}v_0\middle|\mathcal{F}_T^{\lambda,\varsigma}\right].$$

However, we can choose a sampling strategy where we keep the step size Δ_t fixed across trading days by adjusting the sampling frequency M. Alternatively, we may adjust the bias-corrected estimator such that no returns outside the interval $[0,T]$ are taken, i.e.,

$$\widetilde{\mathrm{RV}}_M^{\mathrm{AC}(p)}(0,T) = \sum_{j=1}^{M}\tilde{r}_j^2 + \frac{M}{M-k+1}\left\{\sum_{k=1}^{p}\sum_{j=k+1}^{M}\tilde{r}_j\tilde{r}_{j-k} + \sum_{k=1}^{p}\sum_{j=1}^{M-k+1}\tilde{r}_j\tilde{r}_{j+k}\right\}. \tag{23}$$

where we scale by ${}^{M}/_{M-k+1}$ to compensate for the missing terms.

Proposition 2.9. Assume that Assumptions 1-6 hold true and that $\{v_j\}_{j\in\mathbb{N}}$ follows the MA(q) process in (22) for which Assumption 9 is satisfied. Then, the bias under TTS of the AC(p)-corrected and interval-adjusted RV from (23) simplifies to

$$\mathrm{bias}_M\,(\mathrm{TTS}) = 2M\mathbb{E}\left[v_{p\Delta_t}v_0 + v_{(p+1)\Delta_t}v_0\middle|\mathcal{F}_T^{\lambda,\varsigma}\right]$$

and if

$$M \leq \left\lfloor \frac{p}{q+1}N\,(T)\right\rfloor$$

it holds that the bias is zero.

In summary, by adding more autocovariance terms to the bias-correction by increasing p, we may increase the sampling frequency and remain bias-free. Furthermore, the more ticks occur during a day, the more often we may sample without adding bias. Lastly, the optimal magnitude of M is negatively related to the order q of the MA noise process.

3 Simulation Study

In order to evaluate the performance of the realized volatility estimator in the setting of the TTSV model, we conduct several simulation studies. We investigate the bias and MSE under different sampling schemes and for different levels of market microstructure noise.

3.1 Simulation Setup

To simulate TTSV model and in order to resemble realistic tick times t_i and thus a realistic counting process, we take $\{t_i\}_{i=1,\dots,N(T)}$ and $\{N(t)\}_{0\leq t\leq T}$ from real transaction based data. In particular, we choose IBM transaction arrival times on a specific trading day and select multiple other days for robustness checks (see Appendix B.2). We simulate the continuous stochastic volatility according to

$$\begin{aligned} \varsigma(t) &= s_c(t)\exp\left(\beta_0+\beta_1\varsigma^*(t)\right), \\ d\varsigma^*(t) &= \alpha\varsigma^*(t)\,dt + dB_2(t), \end{aligned} \tag{24}$$

where $\alpha, \beta_0 < 0$ and $\beta_1 > 0$ are constants and $B_2(t)$ is a Brownian motion independent of the Brownian motion $B(t)$ from Equation (8).[17] The deterministic function $s_c(t)$ produces a concave L-shaped form. The process $\varsigma^*(t)$ belongs to the class of Ornstein-Uhlenbeck processes, where $-\alpha$ characterizes the speed of mean reversion towards zero. Due to stationarity, it holds that $\varsigma^*(t)$ is normally distributed with mean zero and variance $^{-1}/_{2\alpha}$ which we use to initialize the process.
We set $T = 23400$ such that the interval $[0, T]$ represents one trading day with 23400 seconds, i.e., 6.5 hours specific to the trading period on the NYSE. Let τ now be the equidistant clock-time grid $\tau = (\tau_0, \dots, \tau_T) \in \mathbb{R}^{T+1}$ with unit intervals $\Delta_c = 1$ representing one second, i.e., $\tau_j = j$ and $\tau_{j+1} = \tau_j + \Delta_c$ for $j = 0, \dots, T$. We simulate the TTSV model according to the following procedure:

(i) Tick-time grid: The transaction times t_i are taken from real IBM stock data on one trading day and are normalized such that $t_i \in [0, T]$. The tick-time grid is then given by $(t_1, \dots, t_{\overline{N}})$, where $\overline{N} = N(T)$ denotes the total number of transactions on that day. It differs substantially from the clock-time grid in the sense that transaction times are not equidistant in time.

(ii) Counting process $N(t)$: Given the tick-time grid $(t_1, \dots, t_{\overline{N}})$, the counting process is obtained through setting

$$N(t) = \begin{cases} 0 & \text{if } t \in [0, t_1), \\ i & \text{if } t \in [t_i, t_{i+1}),\ i = 1, \dots, \overline{N}-1, \\ N(T) & \text{if } t \in [t_{\overline{N}}, T]. \end{cases}$$

(iii) Intensity process $\lambda(t)$: Since the transaction times are taken from real data, we rely on the estimation of the intensity process $\{\lambda(t)\}_{t\in[0,T]}$. We follow

[17] This simulation approach has for instance been used by Huang and Tauchen (2005), Barndorff-Nielsen et al. (2008) and Li, Nolte and Nolte (2018).

Oomen (2006) and use a non-parametric kernel estimate proposed by Diggle and Marron (1988) to estimate $\lambda(t)$. The estimator includes an adjustment for a possible bias at the edges of the interval $[0, T]$ and is given by

$$\hat{\lambda}_h(\tau_j) = \begin{cases} \frac{1}{h}\sum_{i=1}^{N(T)} \left\{ \kappa\left(\frac{t_i - \tau_j}{h}\right) + \kappa\left(\frac{t_i + \tau_j}{h}\right) \right\} & \text{if } \tau_j \in [0, h), \\ \frac{1}{h}\sum_{i=1}^{N(T)} \kappa\left(\frac{t_i - \tau_j}{h}\right) & \text{if } \tau_j \in [h, T-h], \\ \frac{1}{h}\sum_{i=1}^{N(T)} \left\{ \kappa\left(\frac{t_i - \tau_j}{h}\right) + \kappa\left(\frac{t_i + \tau_j - 2T}{h}\right) \right\} & \text{if } \tau_j \in (T-h, T], \end{cases}$$

for $j = 0, \ldots, T$ where h is the bandwidth and the kernel κ satisfies $\int_{\mathbb{R}} \kappa(x)dx = 1$ and $\kappa(x) = 0$ for $|x| \leq 1$. Following Dahlhaus and Tunyavetchakit (2016), we choose an Epanechnikov kernel function, i.e., $\kappa(x) = {}^3\!/_4 (1 - x^2) \mathbb{1}_{\{|x| \leq 1\}}$. The choice of bandwidth is discussed in Appendix B.1.

(iv) Stochastic volatility $\varsigma(t)$: The volatility process is simulated according to the model in (24), i.e., $\varsigma(\tau_j) = s_c(t) \exp(\beta_0 + \beta_1 \varsigma^*(\tau_j)), j = 1, \ldots, T$ and the process $\varsigma^*(\tau_j)$ is obtained using the Euler discretization scheme

$$\begin{aligned} \varsigma^*(\tau_{j+1}) &= \varsigma^*(\tau_j) + \alpha \Delta \varsigma^*(\tau_j) + \sqrt{\Delta}\varepsilon_{2,j} \\ &= \varsigma^*(\tau_j)(1 + \alpha) + \varepsilon_{2,j}, \end{aligned} \qquad j = 0, \ldots, T-1,$$

where $\varepsilon_{2,j}$ are i.i.d. standard normal random variables. We use $\varsigma^*(t) \sim \mathcal{N}(0, {}^{-1}\!/_{2\alpha})$ to initialize the process. The concavity is added via the function

$$s_c(t) := \frac{1}{a_1 t + a_2} - \frac{\log({}^{a_1}\!/_{a_2} + 1)}{a_1} + 1$$

with $a_1, a_2 > 0$ and $\int_0^T s_c(r)dr = T$ (see e.g. Li et al., 2018).

(v) Integrated variance $\mathrm{IV}(0, T)$: We numerically approximate the integral via the trapezoidal method, i.e.,

$$\begin{aligned} \mathrm{IV}(0, T) &= \int_0^T \varsigma(t)^2 \lambda(t) dt \\ &\approx \frac{\Delta}{2} \sum_{j=0}^{T-1} \left(\varsigma^2(\tau_j) \hat{\lambda}(\tau_j) + \varsigma^2(\tau_{j+1}) \hat{\lambda}(\tau_{j+1}) \right) = \widehat{\mathrm{IV}}(0, T) \end{aligned}$$

where we replace the true intensity process $\lambda(t)$ by its estimate $\hat{\lambda}_h(t)$ from step (iii).

(vi) Logarithmic price process $P(t)$: Since the log-price process is equidistant in terms of the number of transactions, the natural discretization grid is in fact the transaction-time grid. The stochastic volatility process is simulated on the

clock-time grid. This poses a problem, since we need the stochastic volatility at the transaction times t_i (see representation (10)). However, the tick-time grid is not a subset of the clock-time grid because transaction times are recorded with precision of up to a micro-second. Therefore, we use simple linear interpolation to construct $\tilde{\varsigma}(t_{i+1})$ and obtain

$$\begin{aligned} P(t_{i+1}) &= P(t_i) + \tilde{\varsigma}(t_{i+1})\,\varepsilon_{1,i}, \\ \tilde{\varsigma}(t_i) &= (t_i - \lfloor t_i \rfloor)\,\varsigma(\lceil t_i \rceil) + (\lceil t_i \rceil - t_i)\,\varsigma(\lfloor t_i \rfloor), \end{aligned} \qquad i = 0, \ldots, N(T),$$

where $\varepsilon_{1,i}$ are i.i.d. standard normal random variables.

(vii) Market microstructure noise v_i: We add different kinds of noise to the log-price process, i.e.,

$$\tilde{P}(t_i) = P(t_i) + v_i, \quad i = 0, \ldots, N(T)$$

with v_i specified below.

(viii) Price sampling: We use three sampling schemes, namely TTS, BTS and CTS.

Let $K \in \mathbb{N}$ be the number of repetitions of the simulation. For each repetition $k = 1, \ldots, K$ we use bias and mean squared error (MSE) as performance measures. For the true integrated variance $\mathrm{IV}_k(0,T)$ of the k-th simulation, we define

$$\begin{aligned} \mathrm{bias}(k, M, h) &:= \mathrm{RV}_{M,k}(0,T) - \widehat{\mathrm{IV}}_k(0,T), \\ \mathrm{MSE}(k, M, h) &:= \left(\mathrm{RV}_{M,k}(0,T) - \widehat{\mathrm{IV}}_k(0,T)\right)^2, \end{aligned}$$

where the subscript k in $\mathrm{RV}_{M,k}(0,t)$ and $\widehat{\mathrm{IV}}_k(0,t)$ relates to the k-th repetition. Ultimately, we are interested in

$$\mathrm{bias}(M,h) := \frac{1}{K}\sum_{k=1}^{K} \mathrm{bias}(k,M,h), \quad \mathrm{MSE}(M,h) := \frac{1}{K}\sum_{k=1}^{K} \mathrm{MSE}(k,M,h). \quad (25)$$

For simplicity, if h is fixed we write $\mathrm{bias}(M) = \mathrm{bias}(M,h)$ and $\mathrm{MSE}(M) = \mathrm{MSE}(M,h)$ and vice versa. We write bias and MSE without subscripts if the meaning is clear from the context.

3.2 Parameter Setup

The simulation procedure described above requires several parameter choices. Regarding the estimation of the intensity function, we rely on the specification of the bandwidth h. Ideally, h is chosen such that the variance of the estimator is small and the accuracy of the estimator is large, which, as common for kernel estimators, constitutes a trade-off. On the one hand, decreasing the bandwidth h reduces the bias and on the other hand, increasing h extends the smoothing horizon and thus the variance decreases (e.g. Oomen, 2006). More intuitively, a large bandwidth will lead to an over-smoothed intensity that misses important intra-day trading information whereas the estimator becomes more efficient. We choose a bandwidth of $h = 500$ as outlined in Appendix B.1.

The estimated intensity function alongside two additional bandwidths can be found in Figure 8. The underlying counting process was taken from the IBM stock on the 2nd of January, 2015 during the trading hours from 9:30am to 4:00pm. The number of transactions on that day amounts to $N(T) = 3872$. The shape suggests that the intensity varies during the day and that, common for financial markets, it slightly peaks at the beginning and gets more pronounced at the end of the trading day, which is often referred to as U-shaped form.

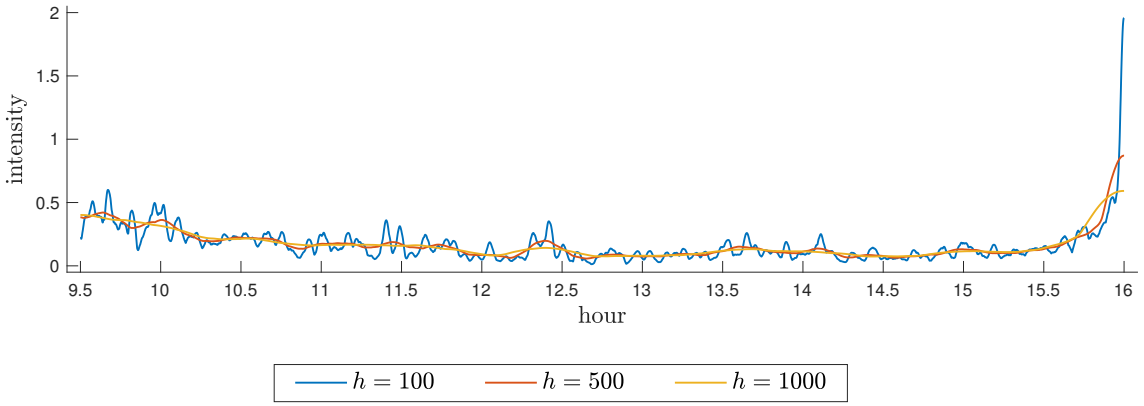

Figure 8: Estimated intensity function for IBM. We plot a bias-adjusted kernel estimate of the intensity function $\lambda(t)$ for different bandwidths. The transaction times were taken from the IBM stock on the 2nd of January, 2015 during the trading hours from 9:30am to 4:00pm.

For each of the following simulations, we choose the parameters reported in Table 1. The parameters related to the volatility process $\varsigma(t)$ correspond to an annual unconditional mean of the spot volatility process $\sigma_{spot}(t)$ of approximately 26.8%.[18]

Parameter setup	
a_1	10
a_2	0.5
β_0	-8.2
β_1	1.0e$-$03
α	-2.5e$-$05
h	500
K	10000
M	$\{13, \ldots, N(T)\}$
$P(0)$	5

Table 1: Experimental design: parameter setup. We report the chosen parameters for the stochastic volatility simulation, $a_1, a_2, \beta_0, \beta_1, \alpha$, the bandwidth h for the estimation of the intensity function, the number of repetitions K, the number of observations M, and the initial log-price $P(0)$.

An example of a simulated price path and an artificial price path in clock-time is given in Figure 9. For an easy comparison, the shaded and non-shaded regions correspond to 500 transactions. Again, we note that the difference between the

[18]Note that we made use of the volatility decomposition $\sigma^2_{spot}(t) = \varsigma^2(t)\lambda(t)$ to approximate σ^2_{spot} (see Proposition C.2).

two price processes is clearly visible. The trading frequency is the lowest at around 12:45pm and the largest in the morning when the market opens and in the afternoon before the market closes. The corresponding simulated volatility process is presented in Figure 10.

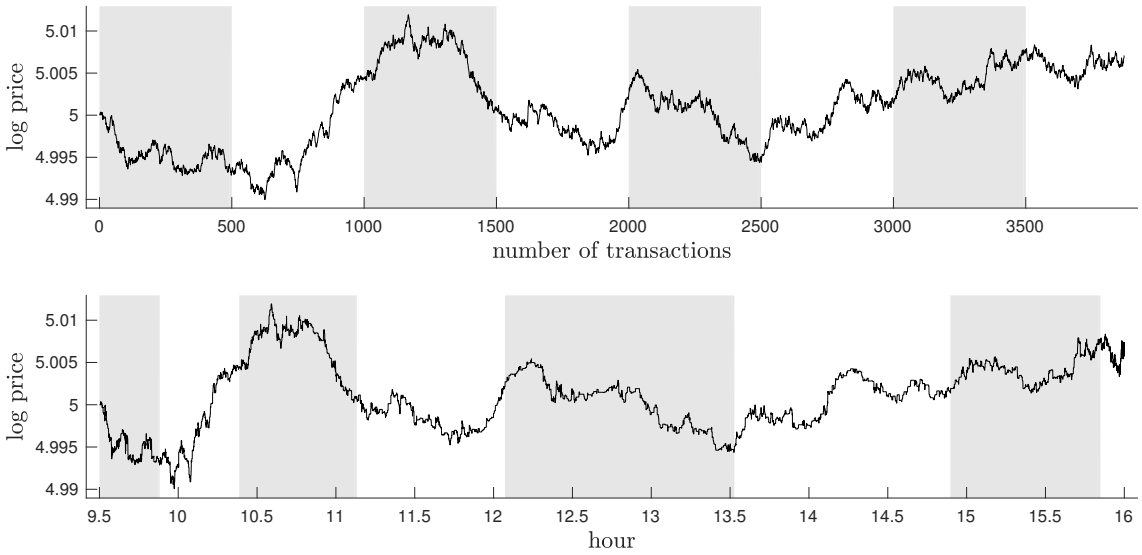

Figure 9: Simulated price path. We plot a simulated price path as a function of the number of transactions (top) and as a function of time (bottom). The underlying parameters can be found in Table 1. The transaction times were taken from the IBM stock on the 2nd of January, 2015 during the trading hours from 9:30am to 4:00pm. On that day there were $N(T) = 3872$ transactions.

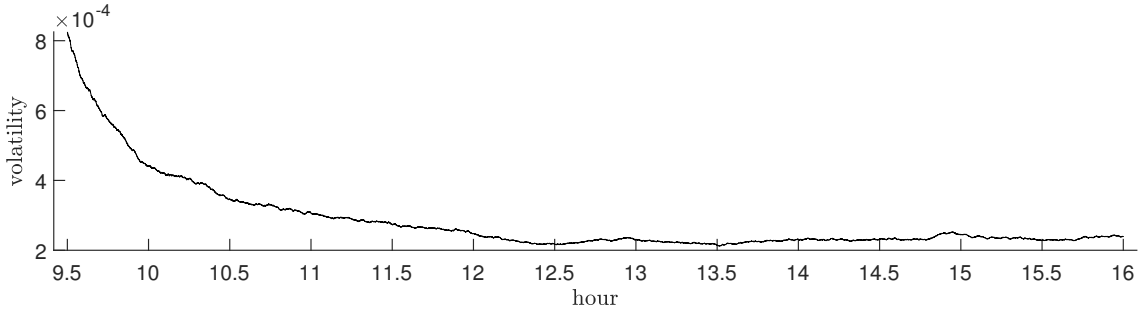

Figure 10: Simulated volatility path. We plot a simulated volatility as a function of time. The underlying parameters can be found in Table 1. The transaction times were taken from the IBM stock on the 2nd of January, 2015 during the trading hours from 9:30am to 4:00pm.

We run four Monte-Carlo simulation studies, for which the market microstructure noise processes and potential bias-corrections are specified in Table 2. The noise specifications include independent, normal noise and an MA(1) process, both with different levels of variance.

3.3 TTSV Model without Market Microstructure Noise

In this section, we evaluate the performance of the RV estimator for different sampling schemes when there is no market microstructure noise (Simulation 1). A plot

Noise setup

	Bias correction	Noise
Simulation 1	–	$v_i = 0$
Simulation 2 Simulation 3	– AC(1)	$v_i \sim$ i.i.d. $\mathcal{N}(0, \sigma_v^2)$ $\sigma_v^2 \in \{1e{-}8, 1e{-}7, 1e{-}6\}$
Simulation 4	AC(1)	$v_i = \nu_i + \theta_1 \nu_{i-1}$ $\nu_i \sim$ i.i.d. $\mathcal{N}(0, \sigma_\nu^2)$ $\sigma_\nu^2 \in 5.525 \cdot \{1e{-}9, 1e{-}8, 1e{-}7\}$

Table 2: Experimental design: noise setup. We report the chosen parameters for the simulation of the market microstructure noise.

of the bias (M) and MSE (M) defined in (25) under the considered sampling schemes is presented in Figure DA.6. The number of observations M is plotted on a log scale. The numerical results can be found in Table EA.1.
For all sampling schemes, the bias is fairly similar and the fluctuations appear to diminish with an increased number of observations M. The MSE under all sampling schemes appears to converge towards zero with seemingly equal speed. However, under CTS, the MSE is clearly the largest regardless of M. Corresponding to the results of the previous sections, BTS has a smaller MSE than TTS for all small sampling frequencies except for very large M. This is due to the fact that the larger M is, the more noise gets added to BTS by resampling the same point. For all sampling frequencies, BTS and TTS outperform CTS in terms of MSE. As expected, the optimal sampling frequency M under no microstructure noise is as large as possible since the bias as well as MSE decrease with increasing M.

3.4 TTSV Model with Market Microstructure Noise

As detailed in Section 2.6, the absence of market microstructure noise is not a reasonable assumption. Thus, in this section we want to investigate the performance of realized volatility under different kinds of noise for the above mentioned sampling schemes.
In a second simulation exercise (Simulation 2), we consider noise v_j that is independent and simulated according to $\mathcal{N}(0, \sigma_\nu^2)$ with different levels of variance σ_ν^2. Figure DA.8 depicts the bias (M) and MSE (M) of the realized variance estimates plotted against the number of transactions M on a log-scale. The level of noise is given by $\sigma_\nu^2 = 1e{-}8$ (upper panel), $\sigma_\nu^2 = 1e{-}7$ (middle panel), and $\sigma_\nu^2 = 1e{-}6$ (lower panel). The numerical results are presented in Table EA.2 in Appendix E.
We observe that bias and MSE increase with increased sampling frequency M and with increased noise variability σ_ν^2. This is in line with what is observed in real data, namely, that the effect of market microstructure noise is more pronounced at high sampling frequencies. Regarding the bias, all sampling schemes perform similar for small frequencies, while CTS outperforms TTS and BTS for large frequencies. We observe that for small sampling frequencies, the MSE under BTS is still the lowest

while for large frequencies, CTS has the lowest MSE. These effects intensify with increasing noise level. The poor performance of TTS is not unexpected considering the results of Lemma 2.3: the negative bias term counteracting the constant bias is the lowest under TTS. A possible additional explanation for the comparably good performance of CTS is that the conversion of the price process from transaction-time to clock-time also induces a bias which is possibly counteractive.
In a third simulation (Simulation 3), we induce the same noise as in the second simulation but use the AC(1)-corrected realized variance estimator $\mathrm{RV}_M^{\mathrm{AC(1)}}(0,T)$ from (20). The results are depicted in Figure DA.9 as well as in Tablthe e EA.3 in Appendix E.
For low levels of noise (upper two panels) and comparably low sampling frequencies, BTS remains the preferred sampling scheme in terms of MSE. For a high level of noise, the realized variance estimate under TTS performs best in terms of bias and MSE, followed by CTS and BTS. According to Example 2.6, $\mathrm{RV}_M^{\mathrm{AC(1)}}(0,T)$ is unbiased under TTS as long as $M \leq N(T)$, i.e., as long as no price is sampled twice. In our simulation we have $N(T) = 3872$ and thus, for $M \leq 3872$, the necessary condition is satisfied. Hence, in this setting the optimal sampling scheme is TTS.
In the last simulation (Simulation 4), we consider a noise process that follows an MA(1) process with $\theta_1 = -0.9$ because there is strong empirical evidence that the noise is serially correlated (e.g. Hansen and Lunde, 2006; Bandi and Russell, 2006, 2008).[19] The noise of the innovations ν_i is chosen such that the variance of v_i is approximately the same as under the independence and normality assumptions of Simulations 2 and 3. For example, $(1+\theta_1^2)\cdot 5.525\mathrm{e}{-9} \approx 1\mathrm{e}{-8}$, where the right side is the lowest noise level from Simulations 2 and 3 (see Table 2 for the exact noise setup). Since the RV estimate will be biased under MA(1) noise, we again consider the AC(1)-corrected estimate $\mathrm{RV}_M^{\mathrm{AC(1)}}(0,T)$.
The results are presented in Figure DA.10 and Table EA.4 in Appendices D and E. Similarly to the previous simulation, RV under TTS outperforms all other sampling schemes in terms of bias and MSE for higher levels of noise. Interestingly, for $M = 2000$ onward, the bias under TTS exhibits a drift. From Example 2.7 and Proposition 2.9, we know that $\mathrm{RV}_M^{\mathrm{AC(1)}}(0,T)$ remains unbiased if $M \leq \lfloor N(T)/2 \rfloor$. However, this bias is not surprising since $N(T) = 3872 < 2 \cdot 2000 = 2M$. Moreover, for higher sampling frequencies, the bias under BTS and CTS shows an upward drift. Again, TTS remains the preferred sampling scheme for high levels of noise.
Robustness checks presented in Appendix B.2 provide evidence that the results are sufficiently insensitive to different choices of bandwidth h as well as transaction times $\{t_i\}_{i=1,\ldots,N(T)}$, in particular, for M large enough.

4 Conclusions

In this paper, we analyze the realized variance estimator under the assumption that the logarithmic price process follows a time-changed diffusion model. The time

[19] This parameter choice induces negative first-order autocorrelation as observed by e.g. Hansen and Lunde (2006) for real equity data.

change is driven by the number of ticks, e.g. transaction times or quote times. We assume that the accumulated number of ticks on a trading day follows a doubly stochastic Poisson process. In this setting, we make use of the volatility decomposition of Dahlhaus and Tunyavetchakit (2016) to define the integrated variance. In the absence of market microstructure noise, we show that the realized variance estimator is unbiased. Additionally, we derive a closed-form solution of the MSE and deduce that business time sampling performs best in terms of MSE if its approximation error is ignored.
Since the jump volatility in our model is stochastic and time-varying, our logarithmic price assumption is more general than that of Oomen (2005, 2006). Using IBM and EUR USD transaction data, we demonstrate that our assumption is more realistic for stocks whereas a constant volatility seems plausible for foreign exchange markets. Similarly to Oomen (2005, 2006), we investigate the impact of several sampling schemes on the quality of realized variance, namely: transaction time sampling, business time sampling and clock-time sampling. On a theoretical side, we derive explicit expressions of the MSE and show that business time sampling has the smallest MSE compared to all other sampling schemes. In a Monte-Carlo experiment, we show that, in the absence of market microstructure noise, transaction time sampling outperforms business time sampling for large sampling frequencies due to the approximation error resulting from the definition of business time sampling.
In our theoretical as well as empirical analysis we also include market microstructure noise. For the case of independent noise, adding first-order autocovariance terms to the realized variance estimator suffices to remain bias-free under transaction time sampling. In simulations, this result is confirmed whereas we observe an increase in bias and MSE for both other sampling schemes. When the noise has a moving average structure of order one, the number of sampled prices under transaction time sampling must remain below one half of the number of transactions on the particular trading day. Similarly to the independent noise, simulations confirm our theoretical findings, and they additionally reveal that transaction time sampling outperforms business time sampling for high sampling frequencies and large levels of market microstructure noise.

Acknowledgements

We would like to thank Winfried Pohlmeier and members of the chairs of Econometrics and Statistics at the University of Konstanz for helpful comments. All remaining errors are ours. Timo Dimitriadis gratefully acknowledges financial support from the University of Konstanz through the Graduate School of Decision Sciences and from the Heidelberger Akademie der Wissenschaften through the WIN-Kolleg. Roxana Halbleib gratefully acknowledges financial support from the European Social Fund, Ministry of Science, Research and Arts of Baden–Württemberg, Germany, from the Deutsche Forschungsgemeinschaft through the PO375/11-1 titled "Robust Risk Measures in Real Time Settings" and from the Zukunftskolleg at the University of Konstanz.

References

Andersen, T. G. and Bollerslev, T. (1997), 'Intraday periodicity and volatility persistence in financial markets', *Journal of Empirical Finance* **4**(2-3), 115–158.

Andersen, T. G., Bollerslev, T., Diebold, F. X. and Ebens, H. (2001), 'The distribution of realized stock return volatility', *Journal of Financial Economics* **61**(1), 43–76.

Andersen, T. G., Bollerslev, T., Diebold, F. X. and Labys, P. (2000*a*), 'Exchange rate returns standardized by realized volatility are (nearly) gaussian', *Multinational Finance Journal* **4**, 159–179.

Andersen, T. G., Bollerslev, T., Diebold, F. X. and Labys, P. (2000*b*), 'Great realizations', *Risk* **13**, 105–108.

Andersen, T. G., Bollerslev, T., Diebold, F. X. and Labys, P. (2003), 'Modeling and forecasting realized volatility', *Econometrica* **71**(2), 579–625.

Andersen, T. G., Davis, R. A., Kreiß, J.-P. and Mikosch, T. V. (2006), *Handbook of Financial Time Series*, Springer Science & Business Media.

Ané, T. and Geman, H. (2000), 'Order flow, transaction clock, and normality of asset returns', *Journal of Finance* **55**(5), 2259–2284.

Bandi, F. M. and Russell, J. R. (2006), 'Separating microstructure noise from volatility', *Journal of Financial Economics* **79**(3), 655–692.

Bandi, F. M. and Russell, J. R. (2008), 'Microstructure noise, realized variance, and optimal sampling', *Review of Economic Studies* **75**(2), 339–369.

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A. and Shephard, N. (2008), 'Designing realized kernels to measure the ex post variation of equity prices in the presence of noise', *Econometrica* **76**(6), 1481–1536.

Barndorff-Nielsen, O. E. and Shephard, N. (2002*a*), 'Econometric analysis of realized volatility and its use in estimating stochastic volatility models', *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **64**(2), 253–280.

Barndorff-Nielsen, O. E. and Shephard, N. (2002*b*), 'Estimating quadratic variation using realized variance', *Journal of Applied Econometrics* **17**(5), 457–477.

Barndorff-Nielsen, O. E. and Shephard, N. (2005), 'How accurate is the asymptotic approximation to the distribution of realized variance', *Identification and inference for econometric models. A Festschrift in honour of TJ Rothenberg* pp. 306–311.

Bollerslev, T. (1986), 'Generalized autoregressive conditional heteroskedasticity', *Journal of Econometrics* **31**(3), 307–327.

Brémaud, P. (1981), *Point Processes and Queues: Martingale Dynamics*, Springer Series in Statistics.

Clark, P. K. (1973), 'A subordinated stochastic process model with finite variance for speculative prices', *Econometrica* pp. 135–155.

Cont, R. and Tankov, P. (2004), *Financial Modelling with Jump Processes* **2**, CRC press.

Dahlhaus, R. and Tunyavetchakit, S. (2016), 'Volatility decomposition and estimation in time-changed price models'. Working paper, available at arXiv:1605.02205.

Diggle, P. and Marron, J. S. (1988), 'Equivalence of smoothing parameter selectors in density and intensity estimation', *Journal of the American Statistical Association* **83**(403), 793–800.

Engle, R. F. (1982), 'Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation', *Econometrica* **50**(4), 987–1007.

Engle, R. F. and Russell, J. R. (1998), 'Autoregressive conditional duration: A new model for irregularly spaced transaction data', *Econometrica* **66**(5), 1127–1162.

Griffin, J. E. and Oomen, R. C. (2008), 'Sampling returns for realized variance calculations: Tick time or transaction time?', *Econometric Reviews* **27**(1-3), 230–253.

Hansen, P. R. and Lunde, A. (2004), 'An unbiased measure of realized variance'. Manustricpt, Stanford University and University of Aarhus.

Hansen, P. R. and Lunde, A. (2006), 'Realized variance and market microstructure noise', *Journal of Business & Economic Statistics* **24**(2), 127–161.

Harris, L. (1990), 'Estimation of stock price variances and serial covariances from discrete observations', *Journal of Financial and Quantitative Analysis* **25**(3), 291–306.

Harris, L. (1991), 'Stock price clustering and discreteness', *Review of Financial Studies* **4**(3), 389–415.

Huang, X. and Tauchen, G. (2005), 'The relative contribution of jumps to total price variance', *Journal of Financial Econometrics* **3**(4), 456–499.

Jones, C. M., Kaul, G. and Lipson, M. L. (1994), 'Transactions, volume, and volatility', *Review of Financial Studies* **7**(4), 631–651.

Li, Y., Nolte, I. and Nolte, S. (2018), 'Asymptotic theory for renewal based high-frequency volatility estimation'. Working paper.

Madhavan, A. (2000), 'Market microstructure: A survey', *Journal of Financial Markets* **3**(3), 205–258.

Newey, W. K. and West, K. D. (1987), 'A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix', *Econometrica* **55**(3), 703–708.

O'hara, M. (1995), *Market Microstructure Theory* **108**, Blackwell Publishers Cambridge.

Oomen, R. (2004), 'Properties of realized variance for a pure jump process: Calendar time sampling versus business time sampling'. Manuscript, University of Warwick.

Oomen, R. C. (2005), 'Properties of bias-corrected realized variance under alternative sampling schemes', *Journal of Financial Econometrics* **3**(4), 555–577.

Oomen, R. C. A. (2006), 'Properties of realized variance under alternative sampling schemes', *Journal of Business & Economic Statistics* **24**(2), 219–237.

Press, S. J. (1967), 'A compound events model for security prices', *Journal of Business* **40**(2), 317–335.

Press, S. J. (1968), 'A modified compound poisson process with normal compounding', *Journal of the American Statistical Association* **63**(322), 607–613.

Protter, P. E. (1990), *Stochastic Integration and Differential Equations: A New Approach*, Springer, Berlin.

Wasserfallen, W. and Zimmermann, H. (1985), 'The behavior of intra-daily exchange rates', *Journal of Banking & Finance* **9**(1), 55–72.

Wu, Z. (2012), 'On the intraday periodicity duration adjustment of high-frequency data', *Journal of Empirical Finance* **19**(2), 282–291.

Zhang, L., Mykland, P. A. and Aït-Sahalia, Y. (2005), 'A tale of two time scales: Determining integrated volatility with noisy high-frequency data', *Journal of the American Statistical Association* **100**(472), 1394–1411.

Zhou, B. (1996), 'High-frequency data and volatility in foreign-exchange rates', *Journal of Business & Economic Statistics* **14**(1), 45–52.

A Mathematical Prerequisites

A.1 Point Processes

The following two lemmata are results from Brémaud (1981) that allow us to link the point process $N(t)$ as a stochastic integrand to its intensity function.

Lemma A.1. Let $\{N(t)\}_{t\geq 0}$ be a doubly stochastic Poisson process adapted to some filtration $\mathcal{F}_t$ with stochastic $\boldsymbol{\mathcal{F}_t}$-intensity $\lambda(t)$ such that $\mathbb{E}\left[\int_0^t \lambda(r)\,dr\right] < \infty$ for all $t \geq 0$. Then,

$$M(t) := N(t) - \int_0^t \lambda(r)\,dr$$

is a $\mathcal{F}_t$-martingale.

Lemma A.2. Let $\{N(t)\}_{t\geq 0}$ be a doubly stochastic Poisson process with stochastic intensity $\lambda(t)$ such that $\mathbb{E}\left[\int_0^t \lambda(r)\,dr\right] < \infty$ for all $t \geq 0$. If $H(t)$ is a $\mathcal{F}_t$-predictable process such that

$$\mathbb{E}\left[\int_0^t |H(r)|\,\lambda(r)\,dr\right] < \infty$$

for $t \geq 0$, then $\int_0^t H(r)\,dM(r)$ is a $\mathcal{F}_t$-martingale.

By general martingale theory, Lemma A.2 yields that $\mathbb{E}\left[\int_0^t H(r)\,dM(r)\right] = 0$ for all $t \geq 0$ which implies that

$$\mathbb{E}\left[\int_s^t H(r)\,dN(r)\right] = \mathbb{E}\left[\int_s^t H(r)\,\lambda(r)\,dr\right] \tag{26}$$

for all non-negative $\mathcal{F}_t$-predictable processes $H(t)$.
An implication of (26) is that $N(t)$ is $\mathbb{P}$-nonexplosive, i.e., $N(t) < \infty$ $\mathbb{P}$-a.s. for all $t \geq 0$ (Brémaud, 1981, Theorem T8 in Chapter II). We may therefore assume that there exists a $\tilde{N} \in \mathbb{N}$ such that for the interval $[0, T]$ the point process $N(t)$ is $\mathbb{P}$-a.s. bounded by $\tilde{N}$. When the t_i denote tick-times, this makes intuitive sense, since there will only be a finite number of trades or quotes within a fixed interval.

B Simulation

B.1 Bandwidth Selection

In order to select a suitable bandwidth, we consider IBM transaction data from 2015. For each of the $K_d = 250$ trading days during that year, we compute the intensity function for different bandwidths $h = 1, \ldots, H$. Then, for each bandwidth, we approximate the bias by

$$\text{bias}_\lambda(h) := \frac{1}{T+1}\sum_{j=0}^{T}\frac{1}{K_d}\sum_{k=1}^{K_d}\left\{\bar{\hat{\lambda}}_{h,k}(0,\tau_j) - N_k(\tau_j)\right\}$$

where $N_k(\tau_j)$ is the number of trades until the j-th second on the k-th day. Similar to before,

$$\bar{\hat{\lambda}}_{h,k}(0,\tau_j) = \sum_{k=0}^{j} \hat{\lambda}_{h,k}(\tau_k) \approx \int_0^{\tau_j} \hat{\lambda}_{h,k}(t)\, dt,$$

where $\hat{\lambda}_{h,k}(t)$ is the estimated intensity process on the k-th day. Moreover, we compute

$$\mathrm{var}_\lambda(h) := \frac{1}{T+1} \sum_{j=0}^{T} \frac{1}{K_d} \sum_{k=1}^{K_d} \left(\hat{\lambda}_{h,k}(\tau_j) - \sum_{k=1}^{K_d} \hat{\lambda}_{h,k}(\tau_j) \right)^2,$$

where $\tau_j = j$ for $j = 1, \ldots, T$ again represents the j-th second. Hence, for each second τ_j we compute the sample variance of $\hat{\lambda}_{h,i}(\tau_j)$ across days and average across all seconds during a trading day. The left and right panel in Figure BA.1 depict $\mathrm{bias}_\lambda(h)$ and $\mathrm{var}_\lambda(h)$ for different sample sizes, respectively. As expected, with increasing bandwidth, the approximate bias increases and the variance decreases. We therefore choose a bandwidth of $h = 500$ as indicated by the grey vertical lines, which appears as an acceptable trade-off.

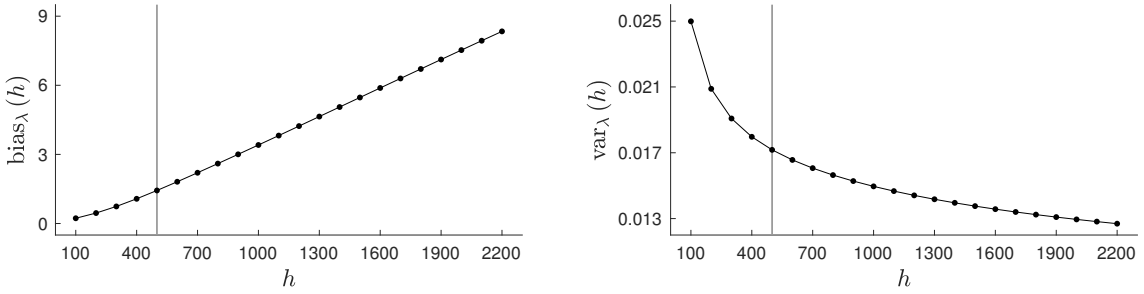

Figure BA.1: Approximate bias and sample variance of the estimated intensity. We plot the approximate bias $\mathrm{bias}_\lambda(h)$ (left panel) of the estimated intensity process $\lambda(t)$ and its variance $\mathrm{var}_\lambda(h)$ (right panel) as a function of the bandwidth h. The grey vertical lines represent a bandwidth of $h = 500$. The transaction times were taken from the IBM stock on the 2$^{\text{nd}}$ of January, 2015 during the trading hours from 9:30am to 4:00pm.

B.2 Robustness

In the following, we conduct robustness checks regarding the chosen transaction time process $\{t_i\}_{i=1,\ldots,N(T)}$ and the bandwidth h for the estimation of the intensity function.

In a first experiment we keep $\{t_i\}_{i=1,\ldots,N(T)}$ fixed and conduct Simulation 1 from Section 3.3 for various bandwidths $h \in \{100, 120, \ldots, 980, 1000\} \in \mathbb{N}^H$ with $H = 37$ (see Table 1 for the parameter setup of Simulation 1). The transaction times are taken from the IBM stock on the 2$^{\text{nd}}$ of January, 2015. We then consider each sampling scheme individually and investigate the bias of the RV estimate as a function of the bandwidth. The bandwidth h does not have an impact on the RV estimate itself, but is only reflected in the integrated variance $\widehat{\mathrm{IV}}(0,T) = \int_0^T \varsigma(t)\, \hat{\lambda}_h(t)\, dt$.

Hence, checking the robustness with regard to the bandwidth is only necessary when the data is simulated, as otherwise $\widehat{\mathrm{IV}}(0,T)$ is not observable. Still, as each sampling scheme produces a different RV estimate, we want to investigate the impact of a change in the bandwidth by considering the bias instead of just the integrated variance. We compute $\mathrm{bias}(h)$ defined in (25) as a function of the bandwidth h while keeping the number of observations M fixed. As the difference in bias across bandwidths cannot be observed visually, we compute the standardized bias across bandwidths, i.e.,

$$\text{std. bias}(h) = \frac{\mathrm{bias}(h) - \overline{\mathrm{bias}}(h)}{\sqrt{\mathrm{var}(\mathrm{bias}(h))}}$$

for each number of observations M. A plot of the standardized bias and the volatility of the bias can be found in Figure BA.2. For each sampling scheme, the dotted black line depicts $\text{std. bias}(500)$ which represents the bandwidth used throughout all simulation studies. The shaded grey area in the left panel depicts a 95% confidence interval.[20] We observe that there is no significant difference in bias for almost all considered bandwidths. Moreover, $\text{std. bias}(h)$ increases with increasing bandwidth h regardless of the sampling scheme and regardless of the number of observations M.

We next want to investigate the impact of the choice of the chosen trading day on our results. Throughout the simulation experiments, we have taken the transaction times $\{t_i\}_{i=1,\dots,N(T)}$ from the IBM stock on the 2nd of January, 2015. They have a direct impact on the log-price process $P(t)$ and the intensity process $\lambda(t)$ and thereby on the integrated variance. We visually compare the estimated intensity function on different days and present a snippet of the comparison in Figure BA.3. The top left panel depicts the estimated intensity on the 2nd of January, 2015 while the three other panels depict the intensity on randomly chosen days from the years 2015 to 2017. We note that the general shape is similar across days. When the market opens and closes, the intensity is larger leading to a U-shaped form. Furthermore, all inspected days mark a drastic peak at the end of the trading day. In general, the intensity function across days shares important and comparable characteristics.

In order to examine the impact of the transaction times beyond the intensity function, we repeat Simulation 1 from Section 3.3 for randomly chosen dates within the period from January 2015 until December 2017. In particular, we select 100 unique trading days and conduct Simulation 1 with $K = 1000$ repetitions. For each of the 100 trading days we obtain $\mathrm{bias}(M)$ and $\mathrm{MSE}(M)$ which we compare across days. The results are presented in Figure BA.4. In the left panel, we plot the maximum, mean, and minimum bias among all trading days and for each sampling scheme. The right panel depicts the maximum, mean, and minimum MSE. We observe that there is a visual difference in performance for all sampling schemes which decreases with increasing sampling frequency M. Similarly, the MSE shows some variation across days which appears to diminish when M is increased. In order to evaluate whether the performance between the RV estimate across days is significantly different, similar to the robustness check on the bandwidth, we compute the standardized

[20] We assume that the price process as well as the volatility process are independent across repetitions and that the intensity can be treated as deterministic as it is taken from real data. Hence, the central limit theorem can be applied to construct confidence intervals.

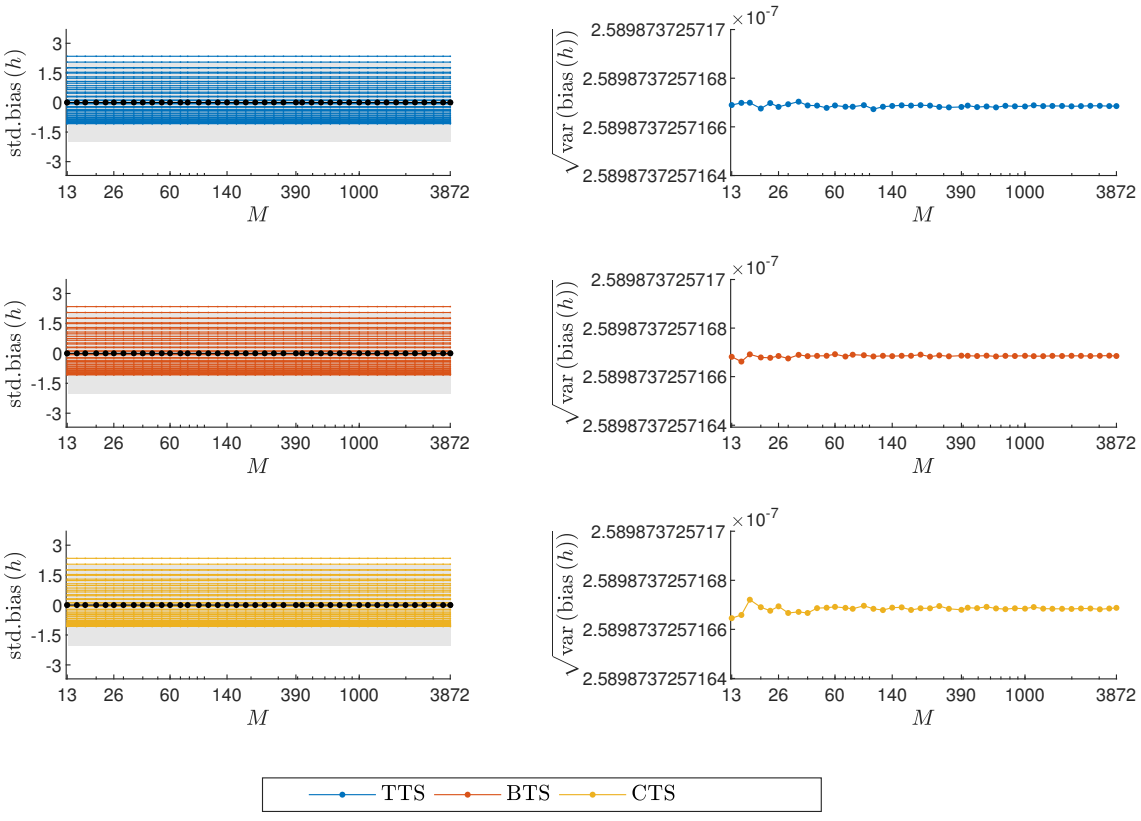

Figure BA.2: Robustness among bandwidths. For each sampling scheme, the left panel depicts std. bias (h) for different bandwidths $h \in \{100, 120, \ldots, 980, 1000\}$ and for different sampling frequencies M. The right panel displays the volatility of bias (h) across bandwidths and for different sampling frequencies M. The sampling frequency M is plotted on a log scale. The shaded area in the left panel depicts a 95% confidence interval. The transaction times were taken from the IBM stock during the trading hours from 9:30am to 4:00pm on the 2$^{\text{nd}}$ of January, 2015.

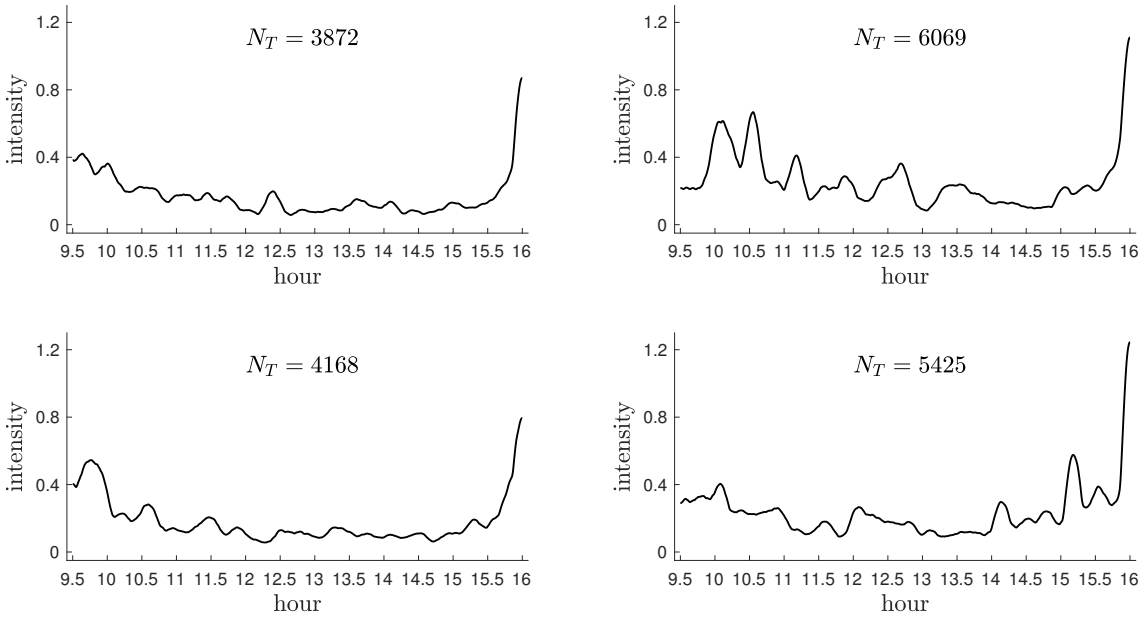

Figure BA.3: Comparison of estimated intensity functions of the IBM stock. We plot adjusted kernel estimates with a bandwidth of $h = 500$ of the intensity function $\lambda(t)$ on different trading days. The transaction times were taken from the IBM stock during the trading hours from 9:30am to 4:00pm on the 2$^{\text{nd}}$ of January, 2015 (top left), and on three randomly chosen days, namely, the 23$^{\text{rd}}$ of April, 2015 (top right), the 13$^{\text{th}}$ of August, 2015 (bottom left), and the 3$^{\text{rd}}$ of November, 2015 (bottom right).

bias across trading days while keeping M fixed. In Figure BA.5, we plot the percentage of days on which the standardized bias is significantly different from zero on a 5% significance level which we denote by $p \in [0, 1]$. Starting from $M > 120$, there seems to be no significant difference in the performance of RV among each sampling scheme. As previously established, a large M may increase the effect of market microstructure noise and choosing a sampling frequency that is too low causes bias due to the lack of information between sampled points. Figure BA.5 is in line with these findings and can be helpful in terms of selecting the sampling frequency M. In general, the results appear sufficiently insensitive to different choices of bandwidth h and transaction times $\{t_i\}_{i=1,\ldots,N(T)}$, in particular, for M large enough.

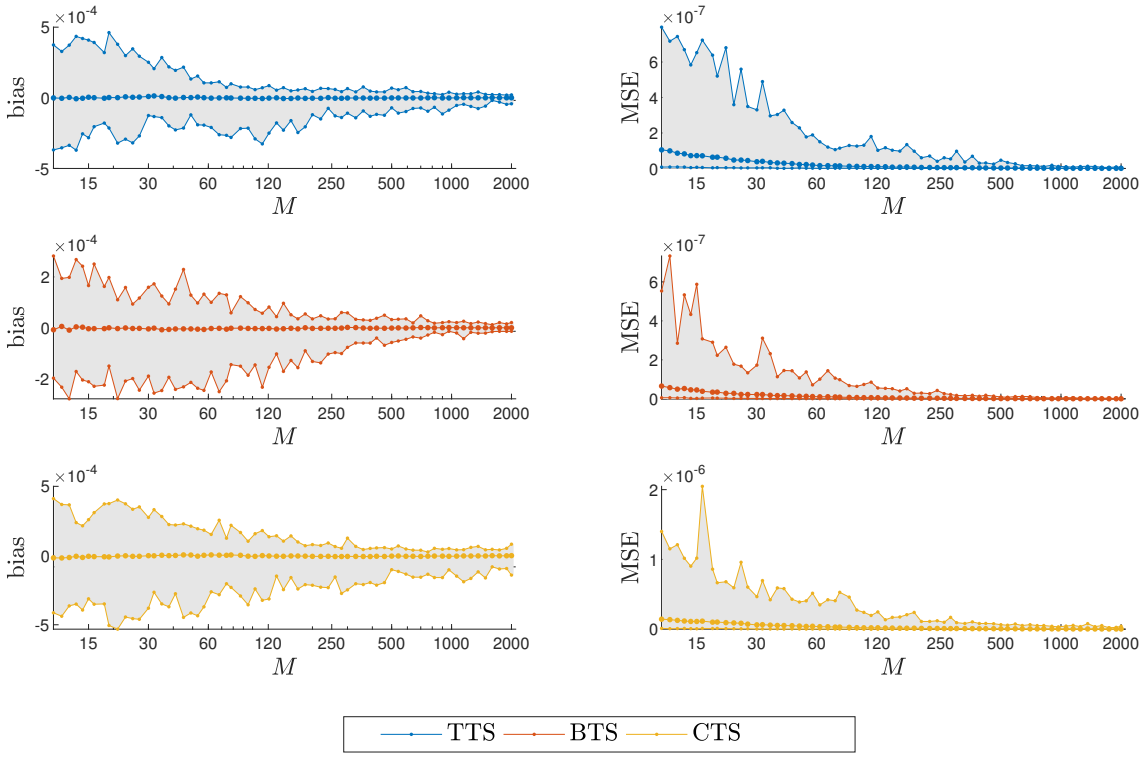

Figure BA.4: Robustness among transaction times $\{t_i\}_{i=1,\ldots,N(T)}$. The left panels depict the maximum, mean, and minimum bias (M) for TTS, BTS, and CTS (top to bottom). The right panels display the corresponding maximum, mean, and minimum MSE (M). The number of observations M is plotted on a log-scale. The underlying transaction times were taken from the IBM stock on 100 unique and randomly chosen trading days from 2015 to 2017 during the trading hours from 9:30am to 4:00pm.

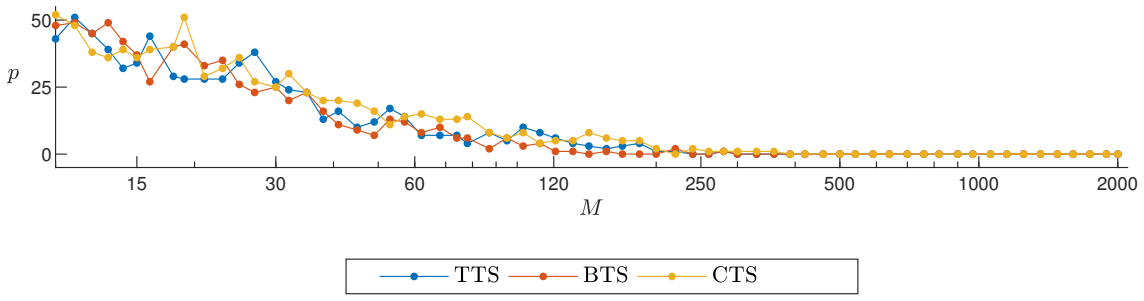

Figure BA.5: Difference in performances across trading days. We plot the percentage of days p for which bias (h) is significantly different from the mean bias across days for fixed M and fixed sampling scheme. The number of observations M is plotted on a log-scale. The underlying transaction times were taken from the IBM stock on 100 unique and randomly chosen trading days from 2015 to 2017 during the trading hours from 9:30am to 4:00pm.

C Proofs

In the diffusion setting, the price process allows for an application of the Ito-Isometry. For the TTSV model, this is not the case. However, there exists an adjusted isometry presented in the following lemma.

Lemma C.1 (TTSV Ito Isometry)**.** Let $\{P(t)\}_{t\geq 0}$ follow the TTSV model as defined in (8). Further, let Assumptions 1-6 hold true. Then, it holds that

$$\mathbb{E}\left[\left(\int_s^t \varsigma(r)\, dB(N(r))\right)^2\right] = \mathbb{E}\left[\int_s^t \varsigma^2(r)\,\lambda(r)\, dr\right],$$
$$\mathbb{E}\left[\left(\int_s^t \varsigma(r)\, dB(N(r))\right)^2 \middle| \mathcal{F}_s\right] = \mathbb{E}\left[\int_s^t \varsigma^2(r)\,\lambda(r)\, dr \middle| \mathcal{F}_s\right],$$

which we refer to as TTSV Ito isometry.

The TTSV Ito-Isometry will be useful in the derivation of the spot variance and thus integrated variance and the bias and MSE of the realized variance estimator.

Proof of Lemma C.1. We split the proof into three parts. 1. Without loss of generality, let $\{t_i\}_{i=n}^m$ with $t_n < \ldots < t_m, n, m \in \mathbb{N}$ and $n \leq m$ denote the sequence of arrival times in the interval $(s,t]$. Using representation (10), it holds that

$$\begin{aligned}
\mathbb{E}\left[\left(\int_s^t \varsigma(r)\, dB(N(r))\right)^2\right] &= \mathbb{E}\left[\left(\sum_{s<t_i\leq t} \varsigma(t_i)\, U_i\right)^2\right] \\
&= \mathbb{E}\left[\left(\sum_{t_n\leq t_i\leq t_m} \varsigma(t_i)\, U_i\right)^2\right] \\
&= \mathbb{E}\left[\left(\sum_{t_n\leq t_i\leq t_{m-1}} \varsigma(t_i)\, U_i + \varsigma(t_m)\, U_m\right)^2\right] \\
&= \mathbb{E}\left[\left(\sum_{t_n\leq t_i\leq t_{m-1}} \varsigma(t_i)\, U_i\right)^2 + (\varsigma(t_m)\, U_m)^2 \right. \\
&\qquad \left. +2\left(\sum_{t_n\leq t_i\leq t_{m-1}} \varsigma(t_i)\, U_i\right)\varsigma(t_m)\, U_m\right].
\end{aligned}$$

By (6), the tower property and the fact that $\varsigma(t)$ is $\mathcal{F}_t$-predictable and therefore $\mathcal{F}_{t-}$-measurable by Assumption 3, it follows that

$$\mathbb{E}[\varsigma(t_m)\, U_m] = \mathbb{E}[\varsigma(t_m)\,\mathbb{E}[U_m|\mathcal{F}_{t_m-}]] = 0$$

and thus the third summand is zero. Similarly,

$$\mathbb{E}\left[(\varsigma(t_m)\, U_m)^2\right] = \mathbb{E}\left[\varsigma^2(t_m)\,\mathbb{E}\left[U_m^2\middle|\mathcal{F}_{t_m-}\right]\right] = \mathbb{E}\left[\varsigma^2(t_m)\right].$$

2. Repeatedly splitting up the squared sum yields

$$\mathbb{E}\left[\left(\int_s^t \varsigma(r)\, dB(N(r))\right)^2\right] = \mathbb{E}\left[\sum_{t_n \le t_i \le t_m} \varsigma^2(t_i)\right] = \mathbb{E}\left[\sum_{s < t_i \le t} \varsigma^2(t_i)\right]$$
$$= \mathbb{E}\left[\int_s^t \varsigma^2(r)\, dN(r)\right].$$

Since $\varsigma(t)$ is a $\mathcal{F}_t$-predictable and positive process, we obtain that

$$\mathbb{E}\left[\int_s^t \varsigma^2(r)\, dN(r)\right] = \mathbb{E}\left[\int_s^t \varsigma^2(r)\, \lambda(r)\, dr\right]$$

by (26) in Appendix A.1 as implied by Lemma A.2.
3. Conditioning on $\mathcal{F}_s$ does not change the first part of the proof since the tower property still holds for $\mathcal{F}_s \subset \mathcal{F}_{t_i-}, i = n \ldots, m$. Hence, we have that

$$\mathbb{E}\left[\left(\int_s^t \varsigma(r)\, dB(N(r))\right)^2 \middle| \mathcal{F}_s\right] = \mathbb{E}\left[\sum_{s < t_i \le t} \varsigma^2(t_i) \middle| \mathcal{F}_s\right] = \mathbb{E}\left[\int_s^t \varsigma^2(r)\, dN(r) \middle| \mathcal{F}_s\right].$$

By Lemma A.2, we get that the process $Y(t) = \int_0^t \varsigma^2(r)\, dM(r)$ is a $\mathcal{F}_t$-martingale and thus,

$$\mathbb{E}\left[\int_s^t \varsigma^2(r)\, dN(r) \middle| \mathcal{F}_s\right] = \mathbb{E}\left[\int_s^t \varsigma^2(r)\, dM(r) \middle| \mathcal{F}_s\right] + \mathbb{E}\left[\int_s^t \varsigma^2(r)\, \lambda(r)\, dr \middle| \mathcal{F}_s\right]$$
$$= \mathbb{E}\left[\int_s^t \varsigma^2(r)\, \lambda(r)\, dr \middle| \mathcal{F}_s\right]$$

which concludes the proof. □

The following lemma is from Dahlhaus and Tunyavetchakit (2016).

Lemma C.2 (Spot variance decomposition)**.** Let $\{P(t)\}_{t\ge 0}$ follow the TTSV model in (8) and let Assumptions 1-4 hold true. Then, if $\{\varsigma(t)\}_{t\ge 0}$ and $\{\lambda(t)\}_{t\ge 0}$ are continuous processes, it holds that

$$\sigma^2_{spot}(t) = \varsigma^2(t)\, \lambda(t),$$

where $\left\{\sigma^2_{spot}(t)\right\}_{t\ge 0}$ is the spot variance process related to $P(t)$.

Proof of Lemma C.2. Throughout this proof, we follow Dahlhaus and Tunyavetchakit (2016), though we have already shown parts of it in Lemma C.1. It holds that

$$\mathbb{E}\left[\int_t^{t+\delta} \varsigma^2(r)\, dN(r) \middle| \mathcal{F}_t\right] = \mathbb{E}\left[\int_t^{t+\delta} \varsigma^2(r)\, \lambda(r)\, dr \middle| \mathcal{F}_t\right]$$

by Lemma C.1, Due to the continuity of $\varsigma^2(t)$ and $\lambda(t)$, it holds that

$$\lim_{\delta \searrow 0} \int_t^{t+\delta} \varsigma^2(r)\, \lambda(r)\, dr = \varsigma^2(t)\, \lambda(t)$$

and $\left|\int_t^{t+\delta} \varsigma^2(r)\,\lambda(r)\,dr\right|$ is $\mathbb{P}$-a.s. bounded by some integrable function by Assumption 4. Then, using dominated convergence, it follows that

$$\begin{aligned}\sigma^2_{spot}(t) &= \lim_{\delta \searrow 0} \mathbb{E}\left[(P(t+\delta) - P(t))^2 \middle| \mathcal{F}_t\right] = \lim_{\delta \searrow 0} \mathbb{E}\left[\int_t^{t+\delta} \varsigma^2(r)\,\lambda(r)\,dr \middle| \mathcal{F}_t\right] \\ &= \mathbb{E}\left[\lim_{\delta \searrow 0} \int_t^{t+\delta} \varsigma^2(r)\,\lambda(r)\,dr \middle| \mathcal{F}_t\right] = \varsigma^2(t)\,\lambda(t). \qquad \square\end{aligned}$$

Proof of Proposition 2.1. Using the definition of the TTSV price process in (10), for $j = 1, \ldots, M$, it follows that

$$\begin{aligned}\mathbb{E}\left[r_j^2 \middle| \mathcal{F}_T^{\lambda,\varsigma}\right] &= \mathbb{E}\left[\left(\int_{\tau_{j-1}}^{\tau_j} \varsigma(r)\,dB(N(r))\right)^2 \middle| \mathcal{F}_T^{\lambda,\varsigma}\right] \\ &= \mathbb{E}\left[\left(\sum_{\tau_{j-1} < t_i \leq \tau_j} \varsigma(t_i)\,U_i\right)^2 \middle| \mathcal{F}_T^{\lambda,\varsigma}\right] \\ &= \mathbb{E}\left[\sum_{\tau_{j-1} < t_i \leq \tau_j} \varsigma(t_i)^2 \middle| \mathcal{F}_T^{\lambda,\varsigma}\right] \\ &= \mathbb{E}\left[\int_{\tau_{j-1}}^{\tau_j} \varsigma^2(r)\,dN(r) \middle| \mathcal{F}_T^{\lambda,\varsigma}\right]\end{aligned}$$

using similar arguments as in the proof of the adjusted Ito isometry (see Lemma C.1) and additionally using the independence of $\varsigma(t)$ and U_i by Assumption 6. Relationship (26) implied by Lemma A.2 yields that the process $Y(t) = \int_0^t \varsigma^2(r)\,dM(r)$ is a martingale and thus $\mathbb{E}\left[\int_0^t \varsigma^2(r)\,dM(r)\right] = 0$. Conditioning on $\mathcal{F}_T^{\lambda,\varsigma}$ does not alter the martingale property of $Y(t)$ since we coincide with a setting in which the intensity and volatility process are given. Hence, it follows that

$$\mathbb{E}\left[\int_{\tau_{j-1}}^{\tau_j} \varsigma^2(r)\,dN(r) \middle| \mathcal{F}_T^{\lambda,\varsigma}\right] = \mathbb{E}\left[\int_{\tau_{j-1}}^{\tau_j} \varsigma^2(r)\,\lambda(r)\,dr \middle| \mathcal{F}_T^{\lambda,\varsigma}\right] = \int_{\tau_{j-1}}^{\tau_j} \varsigma^2(r)\,\lambda(r)\,dr.$$

Summing up, we get that

$$\mathbb{E}\left[\mathrm{RV}_M(0,T) \middle| \mathcal{F}_T^{\lambda,\varsigma}\right] = \mathbb{E}\left[\sum_{j=1}^{M} r_j^2 \middle| \mathcal{F}_T^{\lambda,\varsigma}\right] = \sum_{k=1}^{M} \int_{\tau_{j-1}}^{\tau_j} \varsigma^2(r)\,\lambda(r)\,dr = \mathrm{IV}(0,T)$$

which finishes the proof. $\square$

Proof of Proposition 2.2. We split the proof into several parts. 1. From Lemma 2.1 it follows that

$$\begin{aligned}&\mathbb{E}\left[(\mathrm{RV}_M(0,T) - \mathrm{IV}(0,T))^2 \middle| \mathcal{F}_T^{\lambda,\varsigma}\right] \\ &\quad = \mathbb{E}\left[(\mathrm{RV}_M(0,T))^2 - 2\,\mathrm{RV}_M(0,T)\,\mathrm{IV}(0,T) + \mathrm{IV}(0,T)^2 \middle| \mathcal{F}_T^{\lambda,\varsigma}\right]\end{aligned}$$

$$= \mathbb{E}\left[\left(\mathrm{RV}_M(0,T)\right)^2 \middle| \mathcal{F}_T^{\lambda,\varsigma}\right] - \mathrm{IV}\left(0,T\right)^2.$$

Moreover, we have that

$$\begin{aligned}
\mathbb{E}\left[\left(\mathrm{RV}_M(0,T)\right)^2 \middle| \mathcal{F}_T^{\lambda,\varsigma}\right] &= \mathbb{E}\left[\left(\sum_{j=1}^{M} r_j^2\right)^2 \middle| \mathcal{F}_T^{\lambda,\varsigma}\right] \\
&= \mathbb{E}\left[\left(\sum_{j=2}^{M} r_j^2\right)^2 + 2\sum_{k=2}^{M} r_k^2 r_1^2 + r_1^4 \middle| \mathcal{F}_T^{\lambda,\varsigma}\right] \\
&= \mathbb{E}\left[2\sum_{j=1}^{M-1}\sum_{k=j+1}^{M} r_j^2 r_k^2 + \sum_{j=1}^{M} r_j^4 \middle| \mathcal{F}_T^{\lambda,\varsigma}\right]. \qquad (27)
\end{aligned}$$

2. We simplify the left summand in (27). For the non-overlapping intervals $(\tau_{k-1},\tau_k], k = j+1,\ldots,M$ and $(\tau_{j-1},\tau_j]$ it holds that

$$\begin{aligned}
\mathbb{E}\left[r_j^2 r_k^2 \middle| \mathcal{F}_T^{\lambda,\varsigma}\right] &= \mathbb{E}\left[\left(\sum_{\tau_{j-1}<t_i\leq\tau_j} \varsigma\left(t_i\right) U_i\right)^2 \left(\sum_{\tau_{k-1}<t_i\leq\tau_k} \varsigma\left(t_i\right) U_i\right)^2 \middle| \mathcal{F}_T^{\lambda,\varsigma}\right] \\
&= \mathbb{E}\left[\left(\sum_{\tau_{j-1}<t_i\leq\tau_j} \varsigma\left(t_i\right) U_i\right)^2 \left(\sum_{\tau_{k-1}<t_i\leq\tau_k} \varsigma^2\left(t_i\right)\right) \middle| \mathcal{F}_T^{\lambda,\varsigma}\right] \\
&= \mathbb{E}\left[\left(\sum_{\tau_{j-1}<t_i\leq\tau_j} \varsigma^2\left(t_i\right)\right) \left(\sum_{\tau_{k-1}<t_i\leq\tau_k} \varsigma^2\left(t_i\right)\right) \middle| \mathcal{F}_T^{\lambda,\varsigma}\right] \\
&= \mathbb{E}\left[\left(\int_{\tau_{j-1}}^{\tau_j} \varsigma^2\left(r\right) dN\left(r\right)\right) \left(\int_{\tau_{k-1}}^{\tau_k} \varsigma^2\left(r\right) dN\left(r\right)\right) \middle| \mathcal{F}_T^{\lambda,\varsigma}\right]
\end{aligned}$$

due to the independence of $\varsigma\left(t_i\right)$ and U_i. Since $\varsigma\left(t\right)$ is $\mathcal{F}_T^{\lambda,\varsigma}$-measurable for all $t \leq T$ and the arrival times t_i of the point process are mutually independent, it follows that

$$\begin{aligned}
\mathbb{E}\left[r_j^2 r_k^2 \middle| \mathcal{F}_T^{\lambda,\varsigma}\right] &= \mathbb{E}\left[\int_{\tau_{j-1}}^{\tau_j} \varsigma^2\left(r\right) dN\left(r\right) \middle| \mathcal{F}_T^{\lambda,\varsigma}\right] \mathbb{E}\left[\int_{\tau_{k-1}}^{\tau_k} \varsigma^2\left(r\right) dN\left(r\right) \middle| \mathcal{F}_T^{\lambda,\varsigma}\right] \\
&= \mathbb{E}\left[\int_{\tau_{j-1}}^{\tau_j} \varsigma^2\left(r\right) \lambda\left(r\right) dr \middle| \mathcal{F}_T^{\lambda,\varsigma}\right] \mathbb{E}\left[\int_{\tau_{k-1}}^{\tau_k} \varsigma^2\left(r\right) \lambda\left(r\right) dr \middle| \mathcal{F}_T^{\lambda,\varsigma}\right] \\
&= \mathrm{IV}\left(\tau_{j-1},\tau_j\right) \mathrm{IV}\left(\tau_{k-1},\tau_k\right).
\end{aligned}$$

3. Next, we take on the right summand in (27). Without loss of generality, let $\{t_i\}_{i=n}^m$ with $t_n < \ldots < t_m, n, m \in \mathbb{N}$ and $n \leq m$ denote the series of arrival times in the interval $(\tau_{j-1},\tau_j]$. For $j = 1,\ldots,M$, we have that

$$\mathbb{E}\left[r_j^4 \middle| \mathcal{F}_T^{\lambda,\varsigma}\right] = \mathbb{E}\left[\left(\sum_{\tau_{j-1}<t_i\leq\tau_j} \varsigma\left(t_i\right) U_i\right)^4 \middle| \mathcal{F}_T^{\lambda,\varsigma}\right]$$

$$
\begin{aligned}
&= \mathbb{E}\left[\left(\sum_{t_{n+1}\leq t_i\leq t_m}\varsigma\left(t_i\right)U_i\right)^4+\varsigma^4\left(t_n\right)U_n^4\right.\\
&\qquad +4\left(\sum_{t_{n+1}\leq t_i\leq t_m}\varsigma\left(t_i\right)U_i\right)^3\varsigma\left(t_n\right)U_n\\
&\qquad +6\left(\sum_{t_{n+1}\leq t_i\leq t_m}\varsigma\left(t_i\right)U_i\right)^2\varsigma^2\left(t_n\right)U_n^2\\
&\qquad \left.+4\left(\sum_{t_{n+1}\leq t_i\leq t_m}\varsigma\left(t_i\right)U_i\right)\varsigma^3\left(t_n\right)U_n^3\middle|\mathcal{F}_T^{\lambda,\varsigma}\right]\\
&= \mathbb{E}\left[3\sum_{\tau_{j-1}<t_i\leq\tau_j}\varsigma^4\left(t_i\right)+6\sum_{\tau_{j-1}<t_i<\tau_j}\sum_{t_{i+1}\leq t_h\leq\tau_j}\varsigma^2\left(t_h\right)\varsigma^2\left(t_i\right)\middle|\mathcal{F}_T^{\lambda,\varsigma}\right]\\
&= \mathbb{E}\left[3\left(\sum_{\tau_{j-1}<t_i\leq\tau_j}\varsigma^2\left(t_i\right)\right)^2\middle|\mathcal{F}_T^{\lambda,\varsigma}\right],
\end{aligned}
$$

where we used Assumption 6, and the fact that $U_i\sim\mathcal{N}\left(0,1\right)$. Moreover, it holds that

$$
\begin{aligned}
&\mathbb{E}\left[\left(\sum_{\tau_{j-1}<t_i\leq\tau_j}\varsigma^2\left(t_i\right)\right)^2\middle|\mathcal{F}_T^{\lambda,\varsigma}\right]\\
&\quad= \mathbb{E}\left[\left(\int_{\tau_{j-1}}^{\tau_j}\varsigma^2\left(r\right)dN\left(r\right)\right)^2\middle|\mathcal{F}_T^{\lambda,\varsigma}\right]\\
&\quad= \mathbb{E}\left[\left(\int_{\tau_{j-1}}^{\tau_j}\varsigma^2\left(r\right)dM\left(r\right)+\int_{\tau_{j-1}}^{\tau_j}\varsigma^2\left(r\right)\lambda\left(r\right)dr\right)^2\middle|\mathcal{F}_T^{\lambda,\varsigma}\right]\\
&\quad= \mathbb{E}\left[\left(\int_{\tau_{j-1}}^{\tau_j}\varsigma^2\left(r\right)dM\left(r\right)\right)^2+2\int_{\tau_{j-1}}^{\tau_j}\varsigma^2\left(r\right)dM\left(r\right)\int_{\tau_{j-1}}^{\tau_j}\varsigma^2\left(r\right)\lambda\left(r\right)dr\right.\\
&\qquad\left.+\left(\int_{\tau_{j-1}}^{\tau_j}\varsigma^2\left(r\right)\lambda\left(r\right)dr\right)^2\middle|\mathcal{F}_T^{\lambda,\varsigma}\right]\\
&\quad= \mathbb{E}\left[\left(\int_{\tau_{j-1}}^{\tau_j}\varsigma^2\left(r\right)dM\left(r\right)\right)^2\middle|\mathcal{F}_T^{\lambda,\varsigma}\right]+2\mathbb{E}\left[\int_{\tau_{j-1}}^{\tau_j}\varsigma^2\left(r\right)dM\left(r\right)\middle|\mathcal{F}_T^{\lambda,\varsigma}\right]\mathrm{IV}\left(\tau_{j-1},\tau_j\right)\\
&\qquad+\mathrm{IV}\left(\tau_{j-1},\tau_j\right)^2.
\end{aligned}
$$

The second summand is zero due to the martingale property of $Y\left(t\right)=\int_0^t\varsigma^4\left(r\right)dM\left(r\right)$ from Lemma A.2. To further simplify the first summand, we need the quadratic

variation $[M]_t$ since by Ito's isometry for martingales it holds that

$$\mathbb{E}\left[\left(\int_{\tau_{j-1}}^{\tau_j}\varsigma^2(r)\,dM(r)\right)^2\middle|\mathcal{F}_T^{\lambda,\varsigma}\right]=\mathbb{E}\left[\int_{\tau_{j-1}}^{\tau_j}\varsigma^4(r)\,d[M]_r\middle|\mathcal{F}_T^{\lambda,\varsigma}\right].$$

Let $0=s_0<s_1<\ldots<s_n=t, \max_{1\leq k\leq n}|s_k-s_{k-1}|\to 0$ as $n\to\infty$. Then, using that $N(t)$ is a pure jump process and that $t\mapsto\int_0^t\lambda(r)\,dr$ is continuous, we have that

$$\begin{aligned}
[M]_t &= \text{plim}_{n\to\infty}\sum_{k=1}^{n}\left(M(s_k)-M(s_{k-1})\right)^2\\
&= \text{plim}_{n\to\infty}\sum_{k=1}^{n}\left(N(s_k)-N(s_{k-1})+\int_{s_{k-1}}^{s_k}\lambda(r)\,dr\right)^2\\
&= \text{plim}_{n\to\infty}\sum_{k=1}^{n}\left\{\left(N(s_k)-N(s_{k-1})\right)^2+\left(\int_{s_{k-1}}^{s_k}\lambda(r)\,dr\right)^2\right\}\\
&= [N]_t+\left[\int_0^{\cdot}\lambda(r)\,dr\right]_t\\
&= \sum_{0<s\leq t}\left(N(s)-N(s-)\right)^2\\
&= \sum_{0<s\leq t}\left(N(s)-N(s-)\right)\\
&= N(t).
\end{aligned}$$

Hence, it follows that

$$\begin{aligned}
\mathbb{E}\left[\left(\int_{\tau_{j-1}}^{\tau_j}\varsigma^2(r)\,dM(r)\right)^2\middle|\mathcal{F}_T^{\lambda,\varsigma}\right] &= \mathbb{E}\left[\int_{\tau_{j-1}}^{\tau_j}\varsigma^4(r)\,dN(r)\middle|\mathcal{F}_T^{\lambda,\varsigma}\right]\\
&= \mathbb{E}\left[\int_{\tau_{j-1}}^{\tau_j}\varsigma^4(r)\,dM(r)+\int_{\tau_{j-1}}^{\tau_j}\varsigma^4(r)\lambda(r)\,dr\middle|\mathcal{F}_T^{\lambda,\varsigma}\right]\\
&= \mathbb{E}\left[\int_{\tau_{j-1}}^{\tau_j}\varsigma^4(r)\lambda(r)\,dr\middle|\mathcal{F}_T^{\lambda,\varsigma}\right]\\
&= \text{IQ}(\tau_{j-1},\tau_j),
\end{aligned}$$

where we again used the martingale property of $Y(t)=\int_0^t\varsigma^4(r)\,dM(r)$.
4. Summing up,

$$\begin{aligned}
&\mathbb{E}\left[\left(\text{RV}_M(0,T)-\text{IV}(0,T)\right)^2\middle|\mathcal{F}_T^{\lambda,\varsigma}\right]\\
&\quad=2\sum_{j=1}^{M-1}\sum_{k=j+1}^{M}\text{IV}(\tau_{j-1},\tau_j)\,\text{IV}(\tau_{k-1},\tau_k)-\text{IV}(0,T)^2
\end{aligned}$$

$$
\begin{aligned}
&\quad + 3\sum_{j=1}^{M}\left\{\mathrm{IQ}\left(\tau_{j-1},\tau_j\right) + \mathrm{IV}\left(\tau_{j-1},\tau_j\right)^2\right\} \\
&= 2\sum_{j=1}^{M-1}\sum_{k=j+1}^{M}\mathrm{IV}\left(\tau_{j-1},\tau_j\right)\mathrm{IV}\left(\tau_{k-1},\tau_k\right) \\
&\quad + 3\sum_{j=1}^{M}\mathrm{IV}\left(\tau_{j-1},\tau_j\right)^2 + 3\,\mathrm{IQ}\left(0,T\right) - \mathrm{IV}\left(0,T\right)^2 \\
&= 3\,\mathrm{IQ}\left(0,T\right) + 2\sum_{j=1}^{M}\mathrm{IV}\left(\tau_{j-1},\tau_j\right)^2 . \qquad \square
\end{aligned}
$$

Proof of Lemma 2.3. On the general grid $\tau = (\tau_0, \ldots, \tau_M)$, equation (19) and the independence of r_j and v_i for all $j = 1, \ldots, M$ and $i = 1, \ldots, N(T)$ yield

$$
\begin{aligned}
\mathbb{E}\left[\widetilde{\mathrm{RV}}_M(0,T)\middle|\mathcal{F}_T^{\lambda,\varsigma}\right] &= \mathbb{E}\left[\sum_{j=1}^{M} r_j^2\middle|\mathcal{F}_T^{\lambda,\varsigma}\right] + \sum_{j=1}^{M}\mathbb{E}\left[\left(v_{N(\tau_j)} - v_{N(\tau_{j-1})}\right)^2\middle|\mathcal{F}_T^{\lambda,\varsigma}\right] \\
&= \mathbb{E}\left[\mathrm{RV}_M(0,T)\middle|\mathcal{F}_T^{\lambda,\varsigma}\right] \\
&\quad + \sum_{j=1}^{M}\mathbb{E}\left[v_{N(\tau_j)}^2 - 2v_{N(\tau_j)}v_{N(\tau_{j-1})} + v_{N(\tau_{j-1})}^2\middle|\mathcal{F}_T^{\lambda,\varsigma}\right] \\
&= \mathrm{IV}\left(0,T\right) + 2M\gamma_v\left(0\right) - 2\sum_{j=1}^{M}\mathbb{E}\left[v_{N(\tau_j)}v_{N(\tau_{j-1})}\middle|\mathcal{F}_T^{\lambda,\varsigma}\right].
\end{aligned}
$$

The last equality holds due to the independence of $N(t)$ and v_i for all $i = 1, \ldots, N(T)$. $\square$

Proof of Lemma 2.4. As for the bias we consider the general grid $\tau = (\tau_0, \ldots, \tau_M)$. For simplicity we define $\xi_j = v_{N(\tau_j)} - v_{N(\tau_{j-1})}$ for which $\mathbb{E}\left[\xi^2\middle|\mathcal{F}_T^{\lambda,\varsigma}\right] = 2\gamma_v(0) - 2\mathbb{E}\left[v_{N(\tau_j)}v_{N(\tau_{j-1})}\middle|\mathcal{F}_T^{\lambda,\varsigma}\right]$. Equation (19) and the independence of r_j and v_i for all $j = 1, \ldots, M$ and $i = 1, \ldots, N(T)$ yield

$$
\begin{aligned}
\widetilde{\mathrm{MSE}}_M\left(\mathrm{GS}\right) &= \mathbb{E}\left[\left(\widetilde{\mathrm{RV}}_M(0,T) - \mathrm{IV}\left(0,T\right)\right)^2\middle|\mathcal{F}_T^{\lambda,\varsigma}\right] \\
&= \mathbb{E}\left[\left(\mathrm{RV}_M(0,T) - \mathrm{IV}\left(0,T\right)\right)^2 + \left(2\sum_{j=1}^{M} r_j\xi_j + \sum_{j=1}^{M}\xi_j^2\right)^2\right. \\
&\qquad \left. + \left(\mathrm{RV}_M(0,T) - \mathrm{IV}\left(0,T\right)\right)\left(2\sum_{j=1}^{M} r_j\xi_j + \sum_{j=1}^{M}\xi_j^2\right)\middle|\mathcal{F}_T^{\lambda,\varsigma}\right] \\
&= \mathrm{MSE}_M\left(\mathrm{GS}\right) + \mathbb{E}\left[4\left(\sum_{j=1}^{M} r_j\xi_j\right)^2 + 2\sum_{j=1}^{M} r_j\xi_j\sum_{j=1}^{M}\xi_j^2 + \left(\sum_{j=1}^{M}\xi_j^2\right)^2\middle|\mathcal{F}_T^{\lambda,\varsigma}\right].
\end{aligned}
$$

The second term in the expectation is zero and moreover,

$$\mathbb{E}\left[4\left(\sum_{j=1}^{M} r_j \xi_j\right)^2 \middle| \mathcal{F}_T^{\lambda,\varsigma}\right] = 4\sum_{j=1}^{M} \mathbb{E}\left[r_j^2 \middle| \mathcal{F}_T^{\lambda,\varsigma}\right] \mathbb{E}\left[\xi_j^2 \middle| \mathcal{F}_T^{\lambda,\varsigma}\right]$$
$$= 4\sum_{j=1}^{M} \mathbb{E}\left[r_j^2 \middle| \mathcal{F}_T^{\lambda,\varsigma}\right] \mathbb{E}\left[2\gamma_v(0) - 2v_{N(\tau_j)} v_{N(\tau_{j-1})} \middle| \mathcal{F}_T^{\lambda,\varsigma}\right].$$

Finally

$$\widetilde{\mathrm{MSE}}_M(\mathrm{GS}) = \mathrm{MSE}_M(\mathrm{GS}) + 8\left(\mathrm{IV}(0,T)\gamma_v(0) - \sum_{j=1}^{M} \mathbb{E}\left[r_j^2 \middle| \mathcal{F}_T^{\lambda,\varsigma}\right] \mathbb{E}\left[v_{N(\tau_j)} v_{N(\tau_{j-1})} \middle| \mathcal{F}_T^{\lambda,\varsigma}\right]\right)$$
$$+ \mathbb{E}\left[\left(\sum_{j=1}^{M} \left(v_{N(\tau_j)} - v_{N(\tau_{j-1})}\right)^2\right)^2 \middle| \mathcal{F}_T^{\lambda,\varsigma}\right]. \qquad \square$$

Proof of Lemma 2.5. On the general grid $\tau = (\tau_0, \ldots, \tau_M)$, where $0 = \tau_0 < \ldots < \tau_M = T$ the intervals $(\tau_{j-1}, \tau_j], j = 1, \ldots, M$ are disjoint. Recall that

$$r_j = r(\tau_{j-1}, \tau_j) = \sum_{\tau_{j-1} < t_i \leq \tau_j} \varsigma(t_i) U(t_i)$$

where the t_i are the tick-times. Thus, due to the structure of the process $U(t_i)$ (see Assumptions 1-6), it holds that $\mathbb{E}\left[r_j \middle| \mathcal{F}_T^{\lambda,\varsigma}\right] = 0$ and $\mathbb{E}\left[r_j r_k \middle| \mathcal{F}_T^{\lambda,\varsigma}\right] = 0$ for $j \neq k$. Furthermore, $\mathbb{E}[v_i] = 0$ and the processes r_j and v_i are independent for all $j = 1, \ldots, M$ and $i = 1, \ldots, N(T)$. From equation (19), the definition of the bias-corrected realized variance and Lemma 2.3, it follows that

$$\mathbb{E}\left[\widetilde{\mathrm{RV}}_M^{\mathrm{AC}(1)}(0,T) \middle| \mathcal{F}_T^{\lambda,\varsigma}\right] = \mathbb{E}\left[\sum_{j=1}^{M} \tilde{r}_j^2 + \sum_{j=1}^{M} \tilde{r}_j \tilde{r}_{j-1} + \sum_{j=1}^{M} \tilde{r}_j \tilde{r}_{j+1} \middle| \mathcal{F}_T^{\lambda,\varsigma}\right]$$
$$= \mathbb{E}\left[\widetilde{\mathrm{RV}}_M(0,T) \middle| \mathcal{F}_T^{\lambda,\varsigma}\right]$$
$$+ \sum_{j=1}^{M} \mathbb{E}\left[\left(v_{N(\tau_j)} - v_{N(\tau_{j-1})}\right)\left(v_{N(\tau_{j-1})} - v_{N(\tau_{j-2})}\right) \middle| \mathcal{F}_T^{\lambda,\varsigma}\right]$$
$$+ \sum_{j=1}^{M} \mathbb{E}\left[\left(v_{N(\tau_j)} - v_{N(\tau_{j-1})}\right)\left(v_{N(\tau_{j+1})} - v_{N(\tau_j)}\right) \middle| \mathcal{F}_T^{\lambda,\varsigma}\right]$$
$$= \mathrm{IV}(0,T) + 2M\gamma_v(0) - 2\sum_{j=1}^{M} \mathbb{E}\left[v_{N(\tau_j)} v_{N(\tau_{j-1})} \middle| \mathcal{F}_T^{\lambda,\varsigma}\right]$$
$$+ \sum_{j=1}^{M} \mathbb{E}\left[2v_{N(\tau_j)} v_{N(\tau_{j-1})} - v_{N(\tau_j)} v_{N(\tau_{j-2})} + v_{N(\tau_{j-1})} v_{N(\tau_{j-2})}\right.$$
$$\left. + v_{N(\tau_j)} v_{N(\tau_{j+1})} - v_{N(\tau_{j-1})} v_{N(\tau_{j+1})} \middle| \mathcal{F}_T^{\lambda,\varsigma}\right]$$

$$
\begin{aligned}
&\quad - 2M\gamma_v(0) \\
&= \mathrm{IV}(0,T) + \sum_{j=1}^{M} \mathbb{E}\left[-v_{N(\tau_j)} v_{N(\tau_{j-2})} + v_{N(\tau_{j-1})} v_{N(\tau_{j-2})} \right. \\
&\qquad \left. + v_{N(\tau_j)} v_{N(\tau_{j+1})} - v_{N(\tau_{j-1})} v_{N(\tau_{j+1})} \middle| \mathcal{F}_T^{\lambda,\varsigma} \right]
\end{aligned}
$$

which shows the claim. □

Proof of Lemma 2.8. Using similar arguments as in the proof of Lemma 2.5, straightforward computations yield

$$
\begin{aligned}
&\mathbb{E}\left[\sum_{k=1}^{p} \sum_{j=1}^{M} \tilde{r}_j \tilde{r}_{j-k} \middle| \mathcal{F}_T^{\lambda,\varsigma} \right] \\
&\quad = \sum_{k=1}^{p} \sum_{j=1}^{M} \mathbb{E}\left[\left(v_{N(\tau_j)} - v_{N(\tau_{j-1})} \right) \left(v_{N\left(\tau_{j-k}\right)} - v_{N\left(\tau_{j-k-1}\right)} \right) \middle| \mathcal{F}_T^{\lambda,\varsigma} \right] \\
&\quad = \sum_{k=1}^{p} \sum_{j=1}^{M} \mathbb{E}\left[v_{N(\tau_j)} v_{N\left(\tau_{j-k}\right)} - v_{N(\tau_j)} v_{N\left(\tau_{j-k-1}\right)} \right. \\
&\qquad\qquad \left. - v_{N(\tau_{j-1})} v_{N\left(\tau_{j-k}\right)} + v_{N(\tau_{j-1})} v_{N\left(\tau_{j-k-1}\right)} \middle| \mathcal{F}_T^{\lambda,\varsigma} \right] \\
&\quad = \sum_{j=1}^{M} \mathbb{E}\left[v_{N(\tau_j)} v_{N(\tau_{j-1})} + v_{N(\tau_{j-1})} v_{N(\tau_{j-1-p})} - v_{N(\tau_j)} v_{N(\tau_{j-1-p})} - v_{N(\tau_{j-1})}^2 \middle| \mathcal{F}_T^{\lambda,\varsigma} \right]
\end{aligned}
$$

and similarly

$$
\begin{aligned}
&\mathbb{E}\left[\sum_{k=1}^{p} \sum_{j=1}^{M} \tilde{r}_j \tilde{r}_{j+k} \middle| \mathcal{F}_T^{\lambda,\varsigma} \right] \\
&\quad = \sum_{k=1}^{p} \sum_{j=1}^{M} \mathbb{E}\left[\left(v_{N(\tau_j)} - v_{N(\tau_{j-1})} \right) \left(v_{N\left(\tau_{j+k}\right)} - v_{N\left(\tau_{j+k-1}\right)} \right) \middle| \mathcal{F}_T^{\lambda,\varsigma} \right] \\
&\quad = \sum_{k=1}^{p} \sum_{j=1}^{M} \mathbb{E}\left[v_{N(\tau_j)} v_{N\left(\tau_{j+k}\right)} - v_{N(\tau_j)} v_{N\left(\tau_{j+k-1}\right)} \right. \\
&\qquad\qquad \left. - v_{N(\tau_{j-1})} v_{N\left(\tau_{j+k}\right)} + v_{N(\tau_{j-1})} v_{N\left(\tau_{j+k-1}\right)} \middle| \mathcal{F}_T^{\lambda,\varsigma} \right] \\
&\quad = \sum_{j=1}^{M} \mathbb{E}\left[v_{N(\tau_j)} v_{N(\tau_{j+p})} + v_{N(\tau_{j-1})} v_{N(\tau_j)} - v_{N(\tau_{j-1})} v_{N(\tau_{j+p})} - v_{N(\tau_j)}^2 \middle| \mathcal{F}_T^{\lambda,\varsigma} \right].
\end{aligned}
$$

From the proof of Lemma 2.3, we know that

$$
\mathbb{E}\left[\sum_{j=1}^{M} \tilde{r}_j^2 \middle| \mathcal{F}_T^{\lambda,\varsigma} \right] = \mathrm{IV}(0,T) + 2M\gamma_v(0) - 2M\mathbb{E}\left[v_{N(\tau_j)} v_{N(\tau_{j-1})} \middle| \mathcal{F}_T^{\lambda,\varsigma} \right].
$$

Combining the three summands, we obtain that

$$
\mathbb{E}\left[\widetilde{\mathrm{RV}}_M^{\mathrm{AC}(p)}(0,T) - \mathrm{IV}(0,T) \middle| \mathcal{F}_T^{\lambda,\varsigma} \right] = \sum_{j=1}^{M} \mathbb{E}\left[v_{N(\tau_{j-1})} v_{N(\tau_{j-1-p})} + v_{N(\tau_j)} v_{N(\tau_{j+p})} \right.
$$

$$-v_{N(\tau_{j-1})}v_{N(\tau_{j+p})} - v_{N(\tau_j)}v_{N(\tau_{j-1-p})}\Big|\mathcal{F}_T^{\lambda,\varsigma}\Big]$$

which concludes the proof. □

Proof of Proposition 2.9 . Recall that $N(\tau_j) - N(\tau_{j-1}) = {}^{jN(T)}/_{M} = j\Delta_t$ for all $j = 1, \ldots, M$. Then, using (23) we obtain that

$$\begin{aligned}
&\mathbb{E}\left[\sum_{k=1}^{p}\frac{M}{M-k+1}\sum_{j=k+1}^{M}\tilde{r}_j\tilde{r}_{j-k}\Big|\mathcal{F}_T^{\lambda,\varsigma}\right] \\
&\quad=\sum_{k=1}^{p}\frac{M}{M-k+1}\sum_{j=k+1}^{M}\mathbb{E}\left[\left(v_{N(\tau_j)}-v_{N(\tau_{j-1})}\right)\left(v_{N\left(\tau_{j-k}\right)}-v_{N\left(\tau_{j-k-1}\right)}\right)\Big|\mathcal{F}_T^{\lambda,\varsigma}\right] \\
&\quad=\sum_{k=1}^{p}\frac{M}{M-k+1}\sum_{j=k+1}^{M}\mathbb{E}\Big[v_{N(\tau_j)}v_{N\left(\tau_{j-k}\right)}-v_{N(\tau_j)}v_{N\left(\tau_{j-k-1}\right)} \\
&\qquad\qquad -v_{N(\tau_{j-1})}v_{N\left(\tau_{j-k}\right)}+v_{N(\tau_{j-1})}v_{N\left(\tau_{j-k-1}\right)}\Big|\mathcal{F}_T^{\lambda,\varsigma}\Big] \\
&\quad=\sum_{k=1}^{p}\frac{M}{M-k+1}\sum_{j=k+1}^{M}\mathbb{E}\left[v_{k\Delta_t}v_0-v_{k\Delta_t}v_{\Delta_t}-v_{(k+1)\Delta_t}v_0+v_{k\Delta_t}v_0\Big|\mathcal{F}_T^{\lambda,\varsigma}\right] \\
&\quad=M\sum_{k=1}^{p}\mathbb{E}\left[2v_{k\Delta_t}v_0-v_{k\Delta_t}v_{\Delta_t}-v_{(k+1)\Delta_t}v_0\Big|\mathcal{F}_T^{\lambda,\varsigma}\right] \\
&\quad=M\sum_{k=1}^{p}\mathbb{E}\left[2v_{k\Delta_t}v_0\Big|\mathcal{F}_T^{\lambda,\varsigma}\right]-M\mathbb{E}\left[v_{\Delta_t}^2\Big|\mathcal{F}_T^{\lambda,\varsigma}\right] \\
&\qquad-M\sum_{k=1}^{p-1}\mathbb{E}\left[v_{k\Delta_t}v_0\Big|\mathcal{F}_T^{\lambda,\varsigma}\right]-M\sum_{k=2}^{p+1}E\left[v_{k\Delta_t}v_0\Big|\mathcal{F}_T^{\lambda,\varsigma}\right] \\
&\quad=-M\gamma_v(0)+M\mathbb{E}\left[v_{p\Delta_t}v_0\Big|\mathcal{F}_T^{\lambda,\varsigma}\right]+M\mathbb{E}\left[v_{\Delta_t}v_0\Big|\mathcal{F}_T^{\lambda,\varsigma}\right]-M\mathbb{E}\left[v_{(p+1)\Delta_t}v_0\Big|\mathcal{F}_T^{\lambda,\varsigma}\right]
\end{aligned}$$

Similarly, we have that

$$\begin{aligned}
&\mathbb{E}\left[\sum_{k=1}^{p}\frac{M}{M-k+1}\sum_{j=k}^{M-k+1}\tilde{r}_j\tilde{r}_{j+k}\Big|\mathcal{F}_T^{\lambda,\varsigma}\right] \\
&\quad=\sum_{k=1}^{p}\frac{M}{M-k+1}\sum_{j=1}^{M-k+1}\mathbb{E}\Big[v_{N(\tau_j)}v_{N\left(\tau_{j+k}\right)}-v_{N(\tau_{j-1})}v_{N\left(\tau_{j+k}\right)} \\
&\qquad\qquad -v_{N(\tau_j)}v_{N\left(\tau_{j+k-1}\right)}+v_{N(\tau_{j-1})}v_{N\left(\tau_{j+k-1}\right)}\Big|\mathcal{F}_T^{\lambda,\varsigma}\Big] \\
&\quad=\sum_{k=1}^{p}\frac{M}{M-k+1}\sum_{j=1}^{M-k+1}\mathbb{E}\left[v_{k\Delta_t}v_0-v_{(k+1)\Delta_t}v_0-v_{k\Delta_t}v_{\Delta_t}+v_{k\Delta_t}v_0\Big|\mathcal{F}_T^{\lambda,\varsigma}\right] \\
&\quad=-M\gamma_v(0)+M\mathbb{E}\left[v_{p\Delta_t}v_0\Big|\mathcal{F}_T^{\lambda,\varsigma}\right]+M\mathbb{E}\left[v_{\Delta_t}v_0\Big|\mathcal{F}_T^{\lambda,\varsigma}\right]-M\mathbb{E}\left[v_{(p+1)\Delta_t}v_0\Big|\mathcal{F}_T^{\lambda,\varsigma}\right]
\end{aligned}$$

Using Lemma 2.5, we know that

$$\mathbb{E}\left[\sum_{j=1}^{M}\tilde{r}_j^2\Big|\mathcal{F}_T^{\lambda,\varsigma}\right]=\mathrm{IV}(0,T)+2M\gamma_v(0)-2M\mathbb{E}\left[v_{\Delta_t}v_0\Big|\mathcal{F}_T^{\lambda,\varsigma}\right]$$

and thus

$$\text{bias}\left(\text{TTS}\right) = \mathbb{E}\left[\widetilde{\text{RV}}_M^{\text{AC}(p)}(0,T) - \text{IV}\left(0,T\right)\middle|\mathcal{F}_T^{\lambda,\varsigma}\right] = 2M\mathbb{E}\left[v_{p\Delta_t}v_0 - v_{(p+1)\Delta_t}v_0\middle|\mathcal{F}_T^{\lambda,\varsigma}\right].$$

Since v_i is an MA (q) process, the largest bias term $\mathbb{E}\left[v_{p\Delta_t}v_0\middle|\mathcal{F}_T^{\lambda,\varsigma}\right]$ is zero if $p\Delta_t \geq q+1$ or equivalently if

$$M \leq \frac{p}{q+1}N\left(T\right)$$

is satisfied. Further, if $\mathbb{E}\left[v_{p\Delta_t}v_0\middle|\mathcal{F}_T^{\lambda,\varsigma}\right] = 0$, then $\mathbb{E}\left[v_{(p+1)\Delta_t}v_0\middle|\mathcal{F}_T^{\lambda,\varsigma}\right] = 0$. □

D Figures

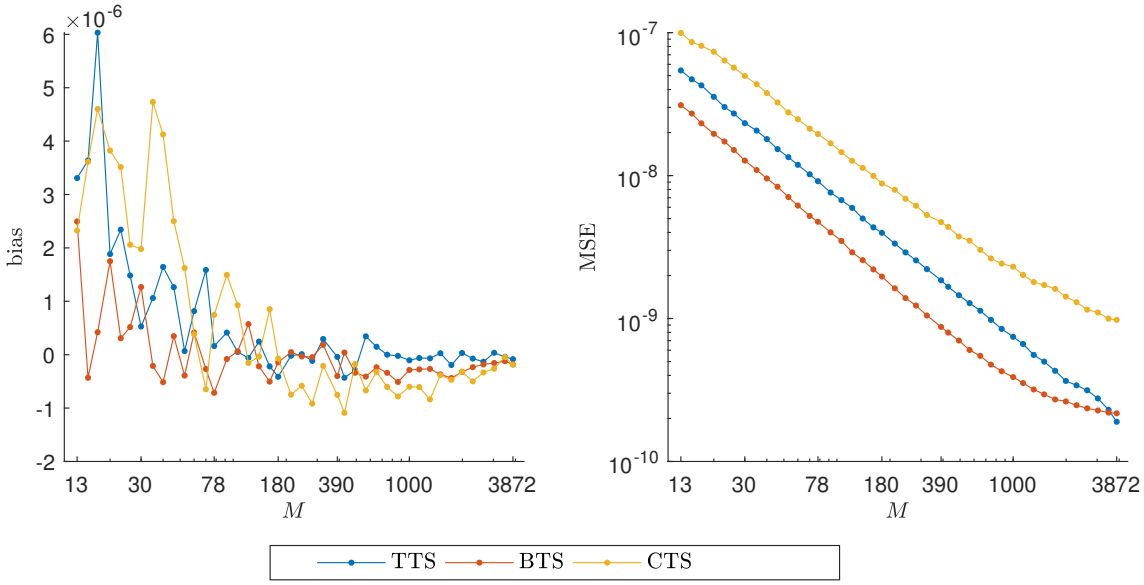

Figure DA.6: Bias and MSE for Simulation 1. We plot the bias and MSE for different sampling schemes under no market microstructure noise. The number of observations M and the MSE are plotted on a log scale. The data was simulated with the parameters in Table 1. The transaction times were taken from the IBM stock on the 2$^{\text{nd}}$ of January, 2015 during the trading hours from 9:30am to 4:00pm.

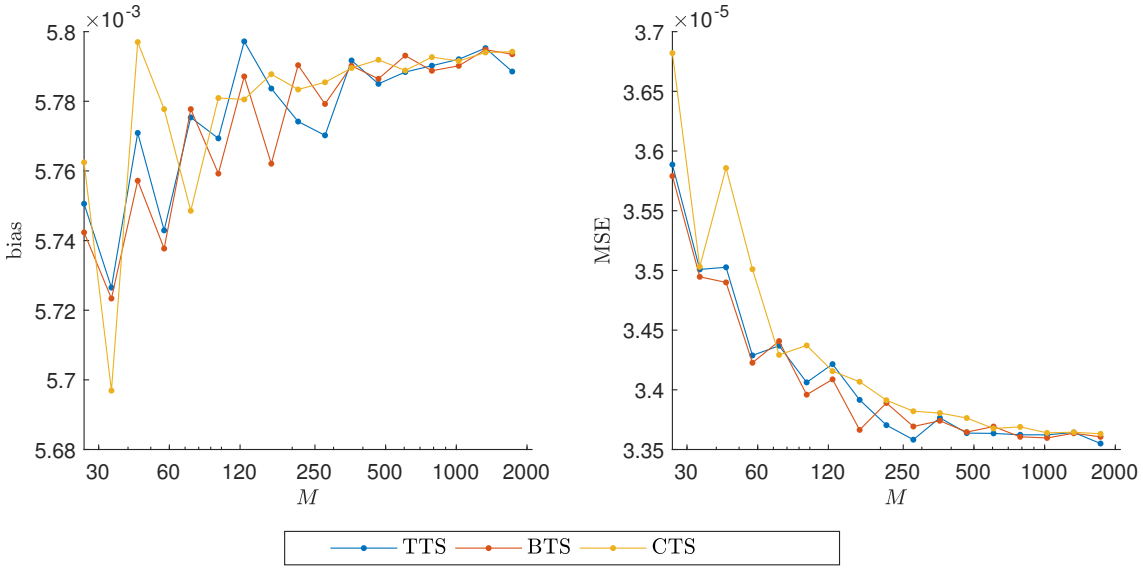

Figure DA.7: Bias and MSE for Simulation 1 for the EUR USD rate. We plot the bias and MSE for different sampling schemes under no market microstructure noise. The number of observations M and the MSE are plotted on a log scale. The data was simulated with the parameters in Table 1 and $\varsigma(t) = 6.8605\text{e}{-5}$. The transaction times were taken from the EUR USD exchange rate on the 1$^{\text{st}}$ of June, 2016.

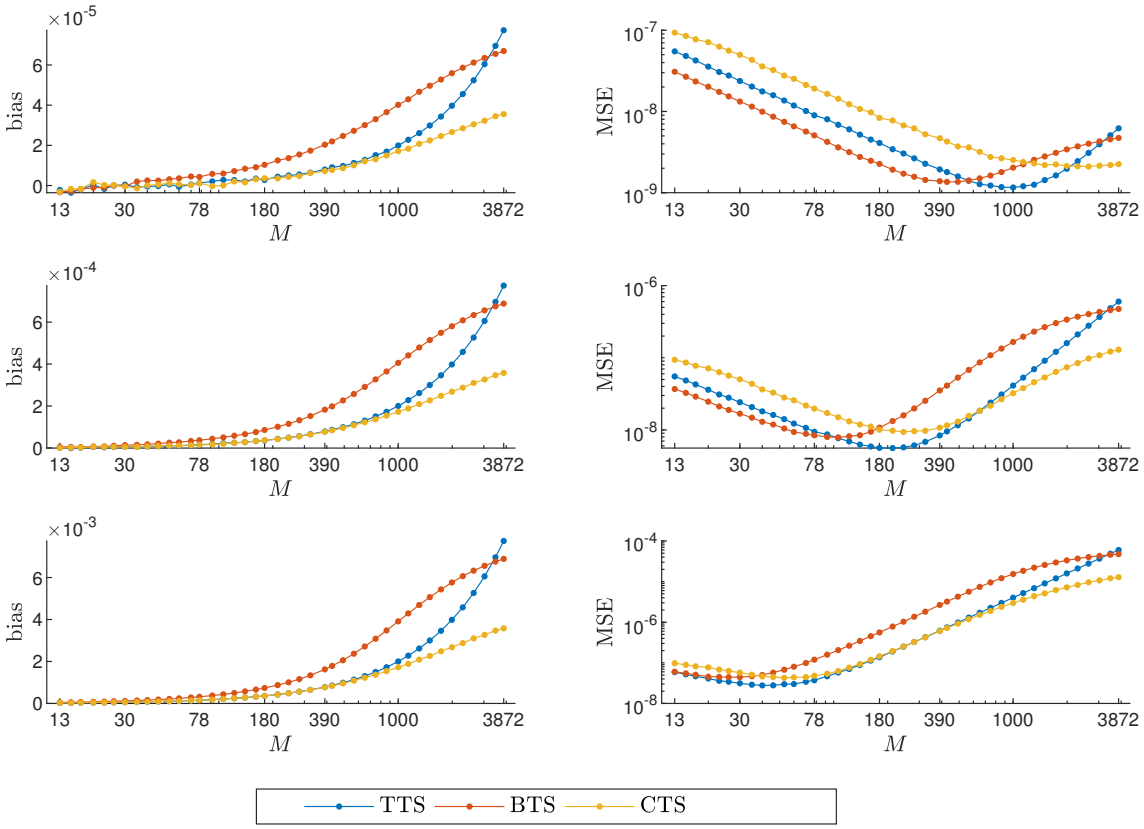

Figure DA.8: Bias and MSE for Simulation 2. We plot the bias and MSE for different sampling schemes under i.i.d. noise distributed according to $\mathcal{N}\left(0, \sigma_v^2\right)$ with $\sigma_v^2 = 1\text{e}{-8}$ (upper panel), $\sigma_v^2 = 1\text{e}{-7}$ (middle panel) and $\sigma_v^2 = 1\text{e}{-6}$ (lower panel). The number of observations M and the MSE are plotted on a log scale. The data was simulated with the parameters in Table 1. The transaction times were taken from the IBM stock on the 2$^{\text{nd}}$ of January, 2015 during the trading hours from 9:30am to 4:00pm.

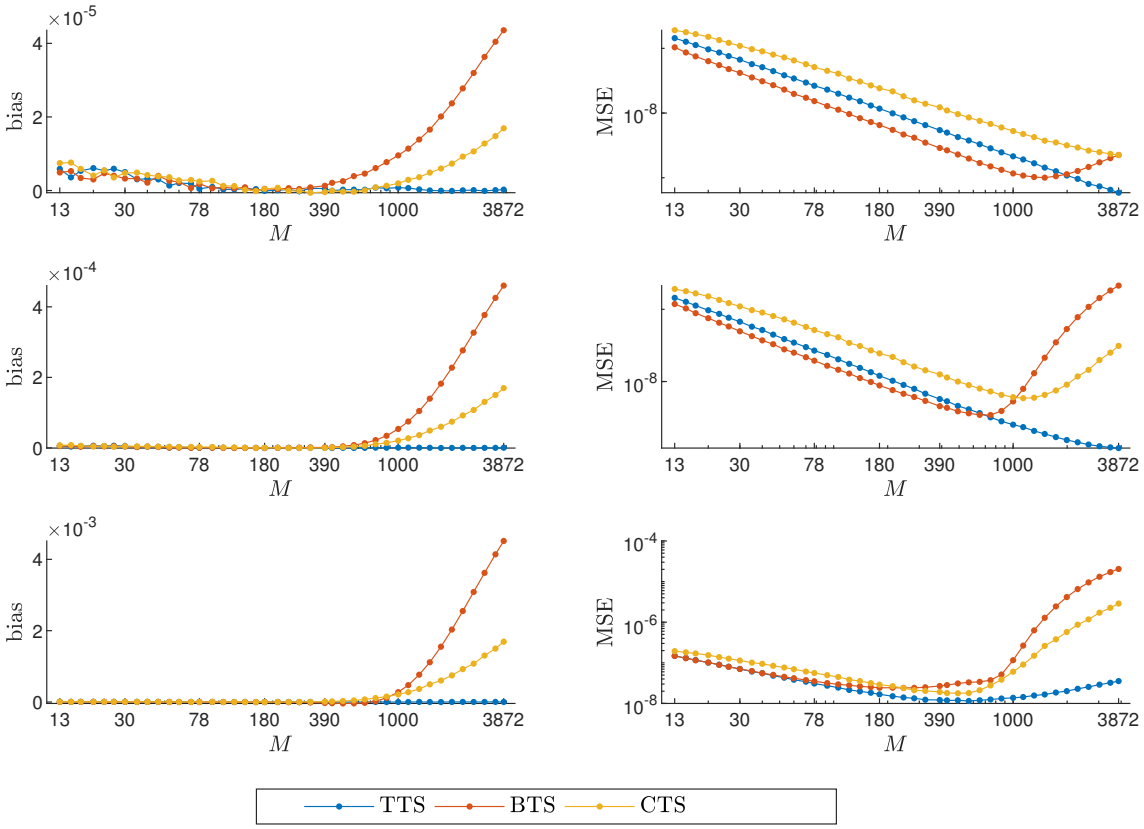

Figure DA.9: Bias and MSE for Simulation 3. We plot the AC(1)-corrected bias and MSE for different sampling schemes under i.i.d. noise distributed according to $\mathcal{N}\left(0, \sigma_v^2\right)$ with $\sigma_v^2 = 1\text{e}{-8}$ (upper panel), $\sigma_v^2 = 1\text{e}{-7}$ (middle panel) and $\sigma_v^2 = 1\text{e}{-6}$ (lower panel). The number of observations M and the MSE are plotted on a log scale. The data was simulated with the parameters in Table 1. The transaction times were taken from the IBM stock on the 2$^{\text{nd}}$ of January, 2015 during the trading hours from 9:30am to 4:00pm.

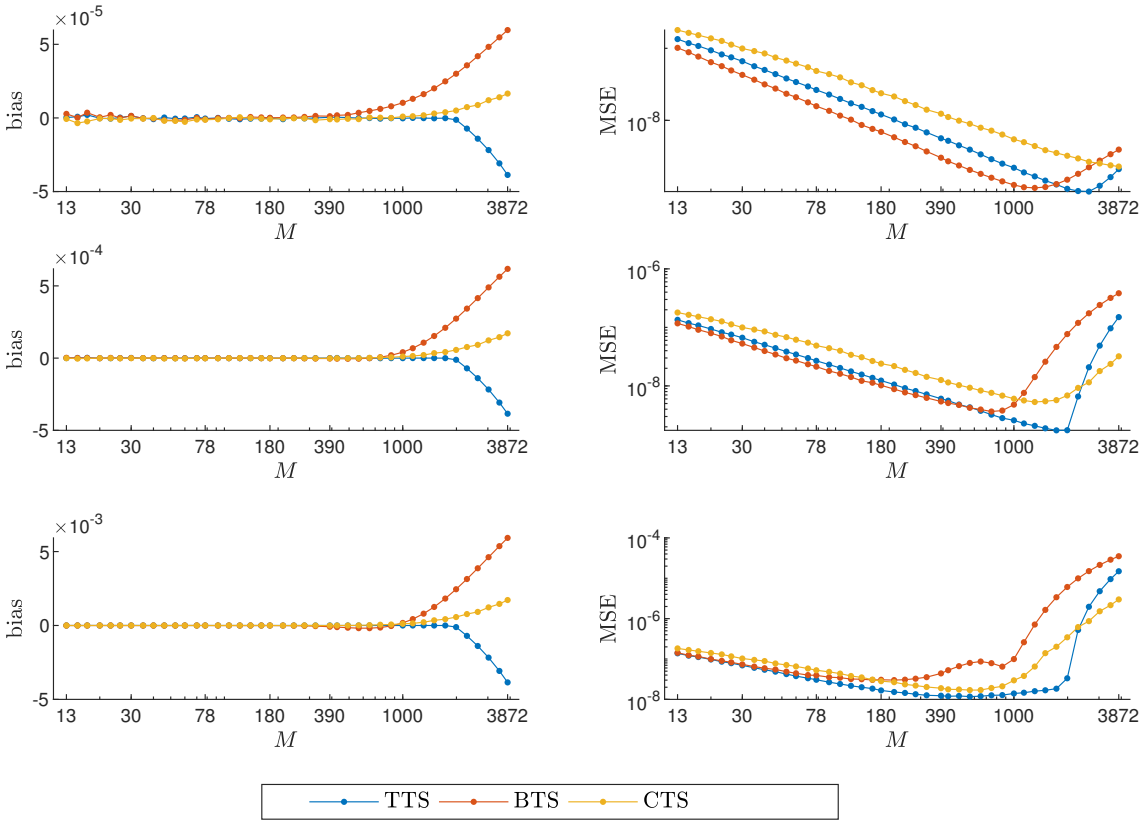

Figure DA.10: Bias and MSE for Simulation 4. We plot the AC(1)-corrected bias and MSE for different sampling schemes under MA(1) noise with $\theta_1 = -0.9$ and with $\sigma_v^2 \approx$ 1e−8 (upper panel), $\sigma_v^2 \approx$ 1e−7 (middle panel) and $\sigma_v^2 \approx$ 1e−6 (lower panel). The number of observations M and the MSE are plotted on a log scale. The data was simulated with the parameters in Table 1. The transaction times were taken from the IBM stock on the 2nd of January, 2015 during the trading hours from 9:30am to 4:00pm.

E Tables

	TTS	BTS	CTS	TTS	BTS	CTS
M	bias $\cdot 10^6$			MSE $\cdot 10^8$		
13	3.4747	2.2392	3.1296	4.6987	2.6693	8.5418
26	0.6216	0.6774	2.8455	2.3378	1.3058	4.9474
78	−0.2010	0.4210	0.5080	0.7777	0.4072	1.6588
390	0.1261	−0.1345	−0.7545	0.1596	0.0763	0.4043
1000	−0.1424	−0.1253	−0.4866	0.0630	0.0333	0.1977
1500	−0.0072	−0.1677	−0.2774	0.0427	0.0255	0.1452
3872	−0.0772	−0.0851	−0.1444	0.0161	0.0184	0.0830

Table EA.1: Bias and MSE for Simulation 1. We report the bias and MSE for different sampling schemes under no market microstructure noise. The data was simulated with the parameters in Table 1. The transaction times were taken from the IBM stock on the 2nd of January, 2015 during the trading hours from 9:30am to 4:00pm.

		TTS	BTS	CTS	TTS	BTS	CTS
σ_ν^2	M	bias $\cdot 10^5$			MSE $\cdot 10^8$		
1.0e−08	13	0.4210	0.2306	0.4489	4.6934	2.6527	8.7405
	26	0.1043	0.2272	0.2985	2.3369	1.3422	4.9822
	78	0.1454	0.4520	0.1268	0.7778	0.4379	1.7113
	390	0.8279	2.0111	0.7657	0.1678	0.1250	0.4128
	1000	1.9599	4.0044	1.7374	0.1037	0.1980	0.2320
	1500	3.0155	4.9669	2.2537	0.1365	0.2753	0.1982
	3872	7.7376	6.7064	3.5573	0.6184	0.4711	0.2116
σ_ν^2	M	bias $\cdot 10^4$			MSE $\cdot 10^7$		
1.0e−07	13	0.0666	0.0764	0.0669	0.4740	0.3336	0.8761
	26	0.0585	0.1307	0.0764	0.2368	0.1743	0.5012
	78	0.1540	0.3679	0.1494	0.0834	0.0775	0.1764
	390	0.7863	1.8185	0.7704	0.0812	0.3491	0.1035
	1000	1.9945	4.0553	1.7255	0.4084	1.6560	0.3218
	1500	2.9964	5.1473	2.2645	0.9071	2.6594	0.5322
	3872	7.7425	6.8866	3.5786	6.0044	4.7516	1.2942
σ_ν^2	M	bias $\cdot 10^3$			MSE $\cdot 10^5$		
1.0e−06	13	0.0302	0.0555	0.0307	0.0051	0.0053	0.0092
	26	0.0514	0.1069	0.0544	0.0029	0.0041	0.0057
	78	0.1551	0.3174	0.1557	0.0036	0.0118	0.0046
	390	0.7802	1.6222	0.7705	0.0618	0.2654	0.0606
	1000	2.0001	3.9090	1.7229	0.4016	1.5313	0.2984
	1500	3.0014	5.0638	2.2620	0.9029	2.5678	0.5135
	3872	7.7426	6.8886	3.5794	5.9999	4.7496	1.2837

Table EA.2: Bias and MSE for Simulation 2. We report the bias and MSE for different sampling schemes under i.i.d. noise distributed according to $\mathcal{N}\left(0, \sigma_v^2\right)$ with $\sigma_v^2 = 1\text{e}{-8}$ (upper panel), $\sigma_v^2 = 1\text{e}{-7}$ (middle panel) and $\sigma_v^2 = 1\text{e}{-6}$ (lower panel). The number of observations M is plotted on a log scale. The data was simulated with the parameters in Table 1. The transaction times were taken from the IBM stock on the 2$^{\text{nd}}$ of January, 2015 during the trading hours from 9:30am to 4:00pm.

		TTS	BTS	CTS	TTS	BTS	CTS
σ_ν^2	M		bias $\cdot 10^5$			MSE $\cdot 10^7$	
1.0e−08	13	−0.3864	−0.3538	−0.5252	1.1692	0.8465	1.5294
	26	0.0879	−0.3208	−0.2171	0.6398	0.4102	0.9993
	78	−0.0906	−0.1895	−0.0672	0.2271	0.1319	0.4371
	390	−0.0047	0.2017	−0.0265	0.0465	0.0250	0.0986
	1000	0.0390	0.9556	0.1381	0.0184	0.0105	0.0443
	1500	0.0030	1.6719	0.4539	0.0123	0.0092	0.0318
	3872	−0.0072	4.4061	1.6592	0.0050	0.0225	0.0193
σ_ν^2	M		bias $\cdot 10^4$			MSE $\cdot 10^7$	
1.0e−07	13	−0.0361	−0.0160	−0.0532	1.1723	0.9962	1.5297
	26	0.0094	−0.0285	−0.0198	0.6417	0.4936	1.0026
	78	−0.0106	−0.0051	−0.0048	0.2292	0.1699	0.4413
	390	−0.0015	0.0136	0.0063	0.0506	0.0424	0.1019
	1000	0.0024	0.5364	0.2067	0.0221	0.0516	0.0515
	1500	0.0034	1.3979	0.4921	0.0163	0.2129	0.0599
	3872	0.0003	4.6151	1.6923	0.0113	2.1397	0.3075
σ_ν^2	M		bias $\cdot 10^3$			MSE $\cdot 10^6$	
1.0e−06	13	−0.0049	−0.0046	−0.0051	0.1207	0.1218	0.1565
	26	0.0010	−0.0018	−0.0013	0.0673	0.0707	0.1035
	78	−0.0009	−0.0016	−0.0002	0.0270	0.0311	0.0479
	390	−0.0006	−0.0245	0.0108	0.0111	0.0257	0.0164
	1000	0.0001	0.2709	0.2125	0.0135	0.1134	0.0595
	1500	−0.0000	1.1109	0.4956	0.0161	1.2736	0.2608
	3872	0.0005	4.5110	1.6901	0.0349	20.3830	2.8754

Table EA.3: Bias and MSE for Simulation 3. We report the AC(1)-corrected bias and MSE for different sampling schemes under i.i.d. noise distributed according to $\mathcal{N}\left(0, \sigma_v^2\right)$ with $\sigma_v^2 = 1\text{e}{-8}$ (upper panel), $\sigma_v^2 = 1\text{e}{-7}$ (middle panel) and $\sigma_v^2 = 1\text{e}{-6}$ (lower panel). The number of observations M is plotted on a log scale. The data was simulated with the parameters in Table 1. The transaction times were taken from the IBM stock on the 2$^{\text{nd}}$ of January, 2015 during the trading hours from 9:30am to 4:00pm.

		TTS	BTS	CTS	TTS	BTS	CTS
σ_ν^2	M		bias $\cdot 10^5$			MSE $\cdot 10^7$	
1.0e−08	13	−0.4215	−0.2844	−0.5880	1.1701	0.8632	1.5267
	26	0.0370	−0.2568	−0.2739	0.6403	0.4184	1.0000
	78	−0.1005	−0.1518	−0.1025	0.2272	0.1341	0.4374
	390	−0.0076	0.1717	−0.0270	0.0465	0.0263	0.0989
	1000	0.0410	1.1141	0.0353	0.0184	0.0113	0.0443
	1500	0.0161	2.0533	0.3031	0.0122	0.0110	0.0318
	3872	−3.8492	6.0449	1.6938	0.0199	0.0398	0.0194
σ_ν^2	M		bias $\cdot 10^4$			MSE $\cdot 10^7$	
1.0e−07	13	−0.0412	−0.0427	−0.0610	1.1722	1.0505	1.5285
	26	0.0050	−0.0203	−0.0266	0.6433	0.5247	1.0039
	78	−0.0115	−0.0113	−0.0091	0.2304	0.1825	0.4414
	390	−0.0037	−0.0389	−0.0085	0.0501	0.0484	0.1016
	1000	0.0015	0.4070	0.1024	0.0220	0.0458	0.0487
	1500	0.0004	1.5356	0.3443	0.0165	0.2597	0.0476
	3872	−3.8495	6.1873	1.7237	1.4982	3.8432	0.3191
σ_ν^2	M		bias $\cdot 10^3$			MSE $\cdot 10^5$	
1.0e−06	13	−0.0050	−0.0079	−0.0062	0.0121	0.0125	0.0156
	26	0.0006	−0.0044	−0.0026	0.0067	0.0072	0.0103
	78	−0.0016	−0.0050	0.0000	0.0027	0.0034	0.0048
	390	0.0001	−0.0856	−0.0047	0.0011	0.0043	0.0016
	1000	−0.0014	0.1672	0.1121	0.0013	0.0098	0.0028
	1500	−0.0018	1.2501	0.3495	0.0017	0.1637	0.0140
	3872	−3.8512	5.9215	1.7204	1.4917	3.5131	0.2987

Table EA.4: Bias and MSE for Simulation 4. We report the AC(1)-corrected bias and MSE for different sampling schemes under MA(1) noise with $\theta_1 = -0.9$ and with $\sigma_v^2 \approx$ 1e−8 (upper panel), $\sigma_v^2 \approx$ 1e−7 (middle panel) and $\sigma_v^2 \approx$ 1e−6 (lower panel). The number of observations M is plotted on a log scale. The data was simulated with the parameters in Table 1. The transaction times were taken from the IBM stock on the 2$^{\text{nd}}$ of January, 2015 during the trading hours from 9:30am to 4:00pm.