The Maximal Range-Return Divergence Statistic*

Yifan Li† Ingmar Nolte‡ Sandra Nolte§

University of Manchester Lancaster University Lancaster University

This Version: October 5, 2022

Abstract

We propose the Maximal rAnge-rEturn Divergence (MAED) statistic, defined as the maximal distance between the price range and the absolute return on a fixed time interval. The statistic can be easily constructed when high-frequency transaction data is available. The MAED statistic summarizes the inward movement of price paths, which contains substantially different information as the candlestick data (i.e., high, low, open and close price in an interval) that mainly captures the outward movement of prices. We propose a spot volatility estimator based on the MAED-augmented candlestick data and establish its asymptotic properties in the fixed-k asymptotic setting with discrete price observations. Our analytical and simulation results show that our MAED-augmented estimator can reduce the asymptotic variance of the optimal candlestick-based spot volatility estimator by as much as 40%. In the presence of extreme price movements such as a jump or a drift burst, the MAED statistic has very different behaviour from the candlestick-based statistics. This allows us to monitor and detect explosive directional price movements in real-time.

JEL classification: C02, C13, C14, C51

Keywords: High-Frequency Data, Nonparametric Inference, Spot Volatility Estimation, Jumps, Drift Bursts.

*We thank Torben Andersen, Kim Christensen, Dobrislav Dobrev, Seok Young Hong, Aleksey Kolokolov, Jia Li, Roberto Renò, Shuping Shi, Viktor Todorov, Jun Yu, as well as participants at the Singapore Management University Workshop, the 5th International Conference on Econometrics and Statistics 2022, and the Volatility, Jumps and Bursts Workshop at Lancaster University for their constructive comments. The authors would like to acknowledge the assistance given by Research IT and the use of the Computational Shared Facility at The University of Manchester. The usual disclaimer applies.

†Corresponding author: Alliance Manchester Business School, Booth Street W, Manchester, M15 6PB, UK. e-mail: yifan.li@manchester.ac.uk.

‡Lancaster University Management School Bailrigg, Lancaster LA1 4YX, UK.

§Lancaster University Management School, Bailrigg, Lancaster, LA1 4YX, UK.

1 Introduction

The recent availability of high-quality tick-by-tick asset transaction data has led to many important developments in financial econometrics. The abundant intraday observations from asset price trajectories allow us to construct various intraday statistics, which can be used to make inference about the volatility of the asset, either in the form of the integrated variance over some fixed interval (e.g., 1 trading day), or the spot volatility at a fixed point in time. Most of the existing methods are built from two fundamental statistics: return — the difference of open and close log-prices of an interval, and range — the difference between the high and low log-prices of an interval. For example, in terms of integrated variance estimation, the widely applied realized volatility (RV) estimator popularized by Andersen et al. (1999, 2003); Barndorff-Nielsen and Shephard (2002) is constructed by summing the intraday squared returns sampled over a trading day. The realized range estimator of Christensen and Podolskij (2007) has the same structure as RV, but sums the squared intraday ranges instead. Andersen et al. (2008) consider duration-based estimators sampled using return and range. As to spot volatility estimation, return-based spot volatility estimators include Foster and Nelson (1996), Kristensen (2010), Zu and Peter Boswijk (2014), and Bollerslev et al. (2021), among others. Recently, Li et al. (2022) and Bollerslev et al. (2022) develop spot volatility estimators as optimal combinations of intraday return and ranges, which are shown to have much better accuracy than a return-based spot volatility estimator.

The statistical advantage of range over return for volatility estimation is well-studied in the literature. In the simple setting of a Brownian motion with a constant variance[1], Parkinson (1980) and Garman and Klass (1980) point out the statistical superiority of range-based over return-based variance estimators and demonstrate that an optimal combination of high, low, open and close (HLOC) prices can further improve the estimation accuracy. Intuitively, while the return summarizes the price variation from the endpoints of an interval and discards the entire price path within the interval, the range describes price extrema in the interval which extracts additional information from the price path and improves the precision of the variance measurements. These results lay the foundation for the high-frequency range-based estimators in Christensen and Podolskij (2007); Li et al. (2022), as high-frequency prices are locally Gaussian in short intervals under the standard assumption that asset prices are semi-martingales.

Despite these theoretical advances on return- and range-based volatility estimators, there are two major problems with using only return and range. First, both return and range summarize 'outward' price movements, which expand the high and low prices in an interval. The 'inward' price movements,

[1]See also Beckers (1983); Rogers and Satchell (1991); Yang and Zhang (2000); Meilijson (2011) for further developments of the HLOC estimators under more general conditions.

which drive the price trajectory towards the open price from an extremum, are largely ignored[2] by return and range. This leads to a substantial information loss, as the inward price path contains as much information about volatility as the outward price path by a reflection principle argument. Second, return and range are heavily influenced by the presence of explosive outward price movements such as jumps and drift bursts[3], which are salient features of empirical asset prices that can introduce substantial bias to the return- and range-based volatility estimates.

To retrieve the information embedded in the inward price movements, we propose the Maximal rAnge-rEturn Divergence (MAED) statistic, defined as the maximal difference between the running range and running absolute return of the price path in an interval. The precise definition and additional ideas are discussed in Section 2. The MAED statistic is simple and fast to construct whenever intraday price observations are available, and it conveys information about price volatility from the inward price paths which is not measured by return or range. Moreover, it is by construction highly robust to jumps and drift bursts as the inward price movements are unaffected by the magnitude of price extrema.

Exploiting these interesting features of the novel MAED statistic discussed above, the paper contributes to the literature by developing two significant MAED-based applications. First, in the fixed-k asymptotic framework of Bollerslev et al. (2021), we proposed the Optimal MAED-candlesticK (OMK) spot volatility estimator, which is in essence a variance-optimal and unbiased linear combination of MAED, range and absolute return. We establish its asymptotic properties using the coupling technique of Bollerslev et al. (2021) and show that it nests the Optimal candlesticK (OK) estimator of Li et al. (2022), a state-of-the-art spot volatility estimator. Importantly, the OMK estimator can in theory reduce the asymptotic variance of the OK estimator by more than 40% with an about 25% tighter confidence interval. This result clearly reflects the importance of MAED in spot volatility estimation, which adds substantial information to the linear span of the return and range data and greatly improves the precision of the OK estimator.

Second, we show that a spot volatility estimator based solely on MAED, referred to as the MAED estimator, is only slightly worse than the OK estimator in terms of asymptotic variance, but it is robust to the presence of a jump with unknown size or location, and behaves very differently to the OK estimator in the presence of drift bursts. Therefore, the MAED estimator provides a simple method for jump-robust spot volatility estimation. More importantly, this feature allows us to construct a simple pivotal test for the detection of jumps or drift bursts in real time by directly comparing the OK estimator with the MAED estimator, which adds a unique method to the literature of local jump and

[2]For example, the range statistic is invariant to different inward price trajectories after reaching the (global) high and low prices of an interval. The return statistic becomes smaller when the close price approaches the open price of an interval, regardless of the size of the inward price movements.

[3]Jumps in asset prices are discussed in e.g., Barndorff-Nielsen and Shephard (2004); Huang and Tauchen (2005), and Andersen et al. (2021); Christensen et al. (2022) document drift bursts and the implications to volatility estimation.

drift burst detection (Lee and Mykland, 2007, 2012; Laurent and Shi, 2020; Christensen et al., 2022).

The paper also makes a non-trivial theoretical contribution to the fixed-k asymptotic framework by analysing the impact of discrete price observations and measurement errors, which are imperfections in empirical prices that are not fully addressed in Bollerslev et al. (2021); Li et al. (2022); Bollerslev et al. (2022). Specifically, we quantify the bias of the OMK estimator due to discretely observed prices and propose a simple bias correction based on a discretely observed Brownian motion. We also show that the impact of the measurement errors can be mitigated by adopting the pre-averaging method of Jacod et al. (2009). Our simulation and empirical analyses demonstrate that these modifications are indeed necessary for valid statistical inference when we observe prices discretely with measurement errors, instead of a continuous price path.

The remainder of the paper is organized as follows. Section 2 gives the definition of the MAED statistic and discusses its basic properties. Main theoretical results of the paper are presented in Section 3. Simulation studies and empirical illustrations of the theoretical results can be found in Section 4 and Section 5, respectively. Section 6 concludes.

2 The MAED Statistic

We start with the definition of the MAED statistic, which is the main innovation of this paper:

Definition 2.1 (The MAED Statistic)**.** *Let $P = (P_t)_{t\geq 0}$ denote some stochastic process understood as the log-price of a financial asset, and consider an arbitrary interval $[s,t]$. Let r_h and w_h denote the running return and range of P from time s till time $h \in [s,t]$:*

$$r_h := P_h - P_s, \quad w_h := \sup_{h_1,h_2\in[s,h]} |P_{h_1} - P_{h_2}|. \tag{2.1}$$

The Maximal rAnge-rEturn Divergence (MAED) m of P on $[s,t]$ is defined as follows:

$$m := \sup_{h\in[s,t]} m_h, \quad m_h := w_h - |r_h|. \tag{2.2}$$

For notational convenience, we shall also use the following notation to denote the full return and range of P associated with the interval $[s,t]$, which is widely used in the estimation of price variation:

$$r := r_t = P_t - P_s, \quad w := w_t = \sup_{h_1,h_2\in[s,t]} |P_{h_1} - P_{h_2}|. \tag{2.3}$$

By construction, m is the maximal distance between the running range w_h and the running absolute return $|r_h|$ of P on $[s,t]$, hence the name. The following properties of m can be easily proved, which holds for any P and $[s,t]$:

Properties of MAED

1. $0 \leq m \leq w$.

2. $m = 0$ iff P is either non-increasing or non-decreasing on $[s, t]$.
3. $m = w$ iff $\exists h^* \in (t^*, t]$ such that $r_{h^*} = 0$, where $t^* = \text{argsup}_{h\in[s,t]} |P_h - P_0|$.
4. Define $P' = \mu + \sigma P$ for some fixed $\mu \in \mathbb{R}$ and $\sigma \in \mathbb{R} \setminus \{0\}$, and let m' denote the MAED of P' on $[s, t]$. Then it holds that $m' = |\sigma| m$.

By Property 1, we learn that m is dominated by w, so its moments always exist as long as all moments of w exist. Property 2 is very special, as it suggests that m does not summarize information from unidirectional price movements, which is otherwise reflected in w and $|r|$. Property 3 reveals when $m = w$ holds, that is, when $r_{h^*} = 0$ at some point h^* after P_h reaches its extremum on $[s, t]$. Property 4 suggests that m is translation invariant and scales linearly with P, which follows from the linearity and symmetry of the absolute return and the range functionals. This property is ideal for the purpose of estimating the scale parameter σ, which is exploited throughout this paper.

To provide more intuitions on what information is captured by the MAED, we present a graphical illustration of the construction of m from a path of P in Fig. 2.1.

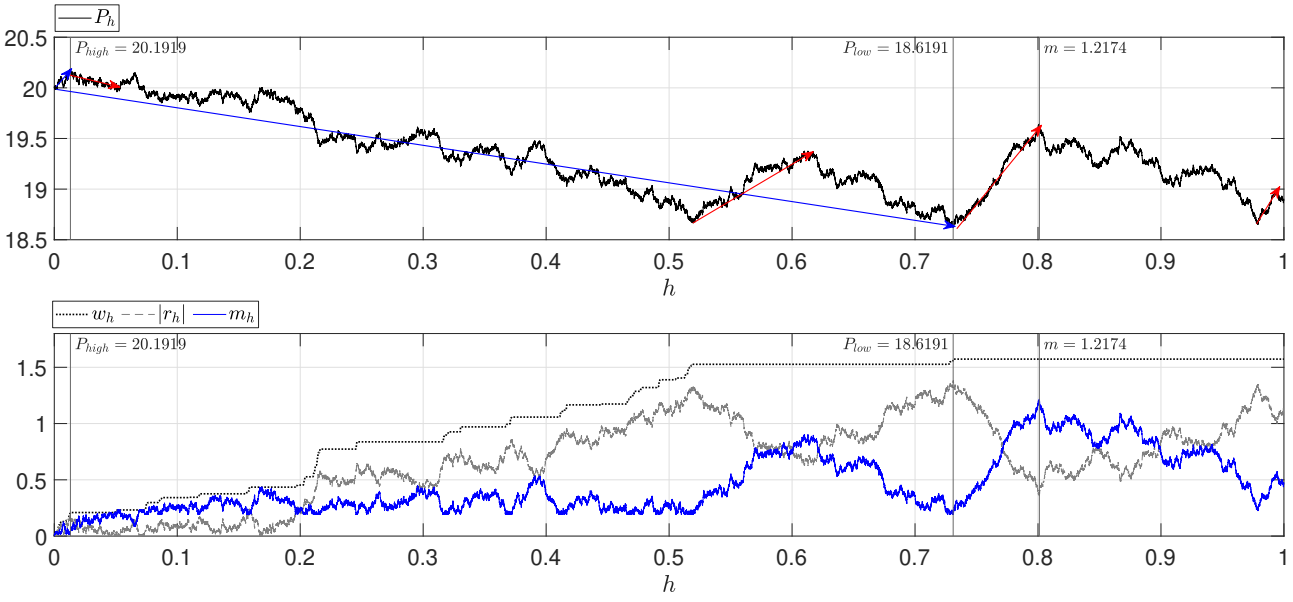


Figure 2.1: Plot of a path of the process $P_h = 20 + W_h$ in the upper panel and the associated processes w_h, $|r_h|$ and m_h in the lower panel for $h \in [0, 1]$, where W is a standard Brownian motion. In the upper panel, the blue and red arrows point out outward and inward price movements, respectively. The price paths are simulated with 10^5 Euler steps. P_{high} and P_{low} refer to the high and low price on $[0, 1]$, respectively.

In the upper panel of Fig. 2.1, we present an example price path on the interval $[0, 1]$ and indicate the times when we obtain the maximum and minimum prices as well as the MAED. Observe that, as the price moves 'outward' relative to P_0 which is indicated by the two blue arrows, the running range of P, plotted by w_h in the lower panel of the figure, increases gradually towards the value of w and remains constant afterwards. Intuitively, the range statistic w only summarizes how much P expands outwards relative to the starting point P_0 on $[0, 1]$, while any price movements between P_{high} and P_{low}

afterwards are completely discarded. Similarly, r only summarizes a single outward price movement from time 0 to time 1 and discards all the information in between.

Crucially, both w and $|r|$ do not capture the 'inward' price movements, which are price movements from the local extrema towards P_0. We point out some large inward price movements by red arrows in Fig. 2.1. These inward price movements do not alter the extrema of P on $[0,1]$ and are ignored by w and $|r|$. The lower panel of Fig. 2.1 shows that m intuitively captures the largest inward price movement. Intuitively, whenever P_h extends outwards by refreshing its extrema, the increment in w_h cancels exactly with the increment in $|r_h|$, so m_h does not increase as P_h moves outwards. On the contrary, when P_h moves inwards from a local extrema, w_h stays constant while $|r_h|$ decreases as P_h moves towards P_0, driving m_h upwards. As a result, m effectively[4] captures the largest inward price movements on the interval $[0,1]$.

The joint distribution of the MAED, the range, and the absolute return of a standard Brownian motion $\tilde{W}$ on $[0,1]$ plays a central role in the theoretical results of this paper. For notational convenience, we shall denote:

$$\tilde{\eta} := \sup_{t\in[0,1]}\{\sup_{h_1,h_2\in[0,t]}|\tilde{W}_{h_1}-\tilde{W}_{h_2}|-|\tilde{W}_t|\},\quad \tilde{\xi}_1 := \sup_{t_1,t_2\in[0,1]}|\tilde{W}_{t_1}-\tilde{W}_{t_2}|,\quad \tilde{\xi}_2 := |\tilde{W}_1|.$$

The densities and moments of $\tilde{\xi}_1$ and $\tilde{\xi}_2$ are known analytically (see e.g., Feller (1951)), and we can easily simulate the density of $\tilde{\eta}$ and the associated moments. We plot the density functions of the three quantities in Fig. 2.2. The figure clearly shows that the mode of the distribution of $\tilde{\eta}$ is approximately 1 with a smaller dispersion when compared with those of $\tilde{\xi}_1$ and $\tilde{\xi}_2$.

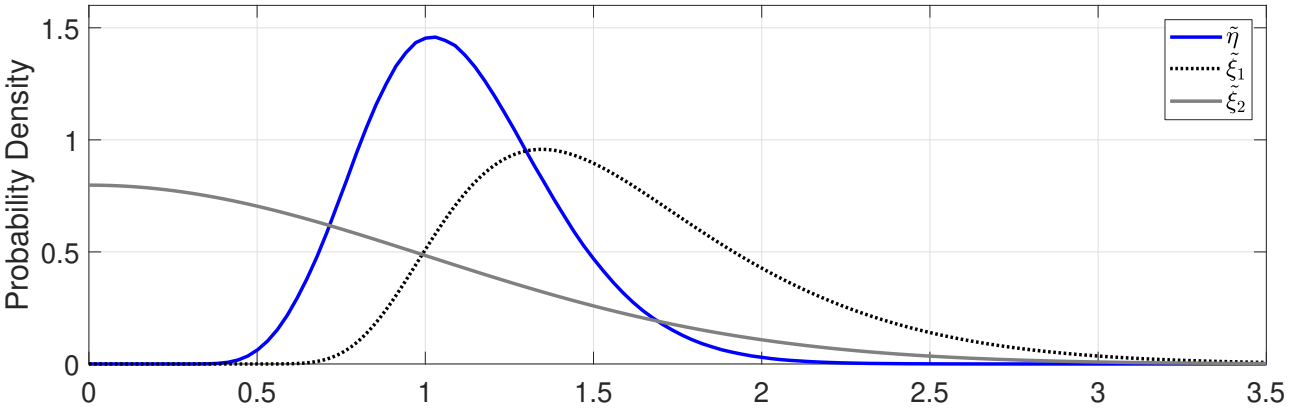


Figure 2.2: Density functions of the MAED ($\tilde{\eta}$), the range ($\tilde{\xi}_1$) and the absolute return ($\tilde{\xi}_2$) of a standard Brownian motion $\tilde{W}$ on $[0,1]$. The density function of $\tilde{\eta}$ is simulated based on $\tilde{W}$ with 10^8 increments. The density functions of $\tilde{\xi}_1$ and $\tilde{\xi}_2$ are $f_1(x) = 8\sum_{k=1}^{\infty}(-1)^{k-1}k^2\phi(kx)$ and $f_2(x) = 2\phi(x)$, respectively, where $x \in \mathbb{R}^+$ and $\phi(x)$ is the density function of a standard normal distribution.

To further understand the properties of the three distributions, we present the first two moments of $\tilde{\eta}$, $\tilde{\xi}_1$ and $\tilde{\xi}_2$. The moments related to $\tilde{\eta}$ are computed via the simulation in Section C, while other

[4]To be precise, the MAED of P on $[0,1]$ equals $|P_{high}-P_0| \wedge |P_{low}-P_0|$ plus the largest inward price movement.

moments have closed forms that can be found in Garman and Klass (1980); Meilijson (2011). Define $\mu_k := \mathrm{E}[\tilde{\eta}^k]$, $\nu_k := \mathrm{E}[\tilde{\xi}_1^k]$, $\psi_k := \mathrm{E}[\tilde{\xi}_2^k]$, $\gamma_0 := \mathrm{E}[\tilde{\xi}_1\tilde{\xi}_2]$, $\gamma_1 := \mathrm{E}[\tilde{\eta}\tilde{\xi}_1]$, $\gamma_2 := \mathrm{E}[\tilde{\eta}\tilde{\xi}_2]$, we have:

$$\begin{aligned} &\mu_1 \approx 1.106, \quad \nu_1 = \sqrt{8/\pi} \approx 1.596, \quad \psi_1 = \sqrt{2/\pi} \approx 0.798, \\ &\mu_2 \approx 1.303, \quad \nu_2 = 4\ln 2 \approx 2.773, \quad \psi_2 = 1, \\ &\gamma_0 = 3/2, \quad \gamma_1 \approx 0.0091, \quad \gamma_2 \approx -0.0755. \end{aligned} \tag{2.4}$$

The above moments imply the following variances and correlations:

$$\begin{gathered} \mathrm{Var}[\tilde{\eta}] \approx 0.0798, \quad \mathrm{Var}[\tilde{\xi}_1] \approx 0.226, \quad \mathrm{Var}[\tilde{\xi}_2] \approx 0.363, \\ \mathrm{Corr}(\tilde{\eta}, \tilde{\xi}_1) \approx 0.0677, \quad \mathrm{Corr}(\tilde{\eta}, \tilde{\xi}_2) \approx -0.443, \quad \mathrm{Corr}(\tilde{\xi}_1, \tilde{\xi}_2) \approx 0.791. \end{gathered}$$

We therefore find that: (1) $\tilde{\eta}$ has a significantly smaller variance than $\tilde{\xi}_1$ and $\tilde{\xi}_2$; (2) $\tilde{\eta}$ is very weakly positively correlated with $\tilde{\xi}_1$, and is negatively correlated with $\tilde{\xi}_2$. These results suggest that $\tilde{\eta}$ measures the variation of W with higher precision in terms of a more concentrated distribution and smaller variance when compared to $\tilde{\xi}_1$ and $\tilde{\xi}_2$. The correlations among the variables further confirm that the inward price movements measured by $\tilde{\eta}$ is substantially different from the outward price movements measured by $\tilde{\xi}_1$ and $\tilde{\xi}_2$.

Concluding from above, we see that the MAED measures the inward price movements of P on an interval, which differs distinctively from the return and the range statistics that only reflect outward price movements. In this paper, we discuss two important applications of MAED: (1) the inward price movements can be more informative than the outward price movements on the variation of price. Therefore, we expect that the precision of volatility estimators based on an optimal combination of w and r, such as Garman and Klass (1980); Li et al. (2022), can be significantly improved when we add m to the combination; (2) by construction, w and r capture the full magnitudes of extreme outward price movements such as jumps or drift bursts (Christensen et al., 2022). On the contrary, m is naturally much more robust as it measures inward price movements from local extrema, which are less affected by jumps or drift bursts that generate local extrema. This feature allows us to detect jumps or drift bursts by comparing the relative magnitudes of m with w or r. Finally, we note that the MAED is very simple to construct in practice. For example, in MATLAB, with the observations of P on $[s, t]$ stored in the vector `P`, the associated MAED `m` can be computed efficiently by the following single line of code:

```
m = max(cummax(P)-cummin(P)-abs(P-P(1)));
```

3 Theoretical Results

3.1 Econometric Setting

In this section, we formulate the econometric setting of the theoretical results in this paper. We consider the problem of spot volatility estimation in a fixed-k framework, following Bollerslev et al. (2021) and Li et al. (2022). On a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$, we assume that the log-price of an asset, denoted by $P = (P_t)_{t\geq 0}$, follows an Itô-semimartingale of the following form:

$$P_t = P_0 + \int_0^t b_s ds + \int_0^t \sigma_s dW_s + J_t, \tag{3.1}$$

where b_t and σ_t are optional processes known as the drift and the spot volatility of P_t, and W_t is the standard Brownian motion. The J process is a pure jump process driven by a random Poisson measure. We are mainly interested in estimating the spot volatility σ_t at an arbitrary time $t \in [0, T]$ for some fixed $T > 0$ representing the time span of the complete dataset.

We require the following mild regularity condition for the fixed-k inference theory:

Assumption 3.1. *For P defined in Eq. (3.1), we assume that there exists an increasing and diverging sequence of stopping times $(T_m)_{m\geq 1}$ and a sequence of constants $(K_m)_{m\geq 1}$ such that the following conditions hold for each $m \geq 1$: (1) for some $r \in [0, 2)$, $|b_t| + |\sigma_t| + |\sigma_t|^{-1} + \int (|x|^r \wedge 1) F_t(dx) \leq K_m$ for all $t \in [0, T_m]$, where F_t is the spot Lèvy measure of J; (2) for some constant $\kappa > 0$, $\mathrm{E}[|\sigma_{t\wedge T_m} - \sigma_{s\wedge T_m}|^2] \leq K_m |t - s|^{2\kappa}$ for all $t, s \in [0, T]$.*

As discussed in Bollerslev et al. (2021) and Li et al. (2022), Assumption 3.1 is highly flexible and allows for many well-established asset price features, such as the leverage effect, intraday diurnal patterns and long-memory in volatility, and jumps in prices or volatility. Specifically, condition (1) in Assumption 3.1 imposes local boundedness to component processes of P, and condition (2) requires σ_t to be locally κ-Hölder continuous under the L_2-norm.

Given a realization of P, an estimator of σ_t is typically constructed over asymptotically shrinking blocks around t, see for example Foster and Nelson (1996), Kristensen (2010), and Chapter 13 in Jacod and Protter (2012). For a generic index $n \to \infty$, one can choose a sequence of numbers of blocks k_n, a vanishing sequence of sampling interval length $\Delta_n \to 0$, and consider the block $t \in [s, s + k_n\Delta_n]$. An important insight from Bollerslev et al. (2021) is that one can make valid inference for estimators of σ_t by using a fixed k instead of a diverging k_n. In the fixed-k setting, the estimator is inconsistent with a non-standard limiting distribution (which is usually known or can be easily simulated), allowing us to construct confidence bounds easily. As the asymptotic analysis is analogous for any fixed k, without much loss of generality we shall fix $k = 1$ in this paper to simplify notations and exposition.

Under the setting above, to construct an estimator of σ_t, we consider an interval $I_n = [s, s + \Delta_n]$ such that $t \in I_n$. We impose the following assumption about what is observed from I_n:

Assumption 3.2. *For all intervals of the form $I_n = [s, s + \Delta_n]$, we observe the MAED m, the return r, and the range w of P on I_n.*

Assumption 3.2 effectively assumes that P is observed continuously on I_n so that m and w associated with P can be computed, which is also adopted by Bollerslev et al. (2021); Li et al. (2022). In practice, for a fixed interval I_n, the assumption holds approximately true when high quality tick-by-tick data for P is available on I_n, but is less appropriate when P is observed sparsely with large measurement errors. We shall firstly establish theoretical results based on the ideal Assumption 3.2 and discuss the implications of these results when Assumption 3.2 is violated.

3.2 The Optimal MAED-Candlestick Estimator of Spot Volatility

In this section, we show that the MAED can be used to obtain more precise volatility measurements when compared to volatility estimators based on the return and the range statistics. Our benchmark estimator is the Optimal-candlesticK (OK) estimator of Li et al. (2022), which is the best linear unbiased estimator (BLUE) of σ_t based on the optimal linear combination of $|r|$ and w of P observed on I_n. By adding m to the combination, we propose the Optimal MAED-candlesticK (OMK) estimator of σ_t, which is a natural generalization of the OK estimator when m, w and r are all available. To this end, we first define the MAED-candlestick vector $\boldsymbol{c} := (m, w, |r|)'$ and denote the 3-by-3 diagonal matrix $\boldsymbol{\Theta} := \text{diag}(\mu_1^{-1}, \nu_1^{-1}, \psi_1^{-1})$. Pick any weight vector $\boldsymbol{\lambda} := (\lambda_m, \lambda_w, \lambda_r)'$ such that $\boldsymbol{\lambda}'\boldsymbol{\iota} = 1$ where $\boldsymbol{\iota} := (1, 1, 1)'$, the OMK estimator is defined as:

$$\hat{\sigma}_t(\boldsymbol{\lambda}) := \Delta_n^{-1/2} \boldsymbol{\lambda}' \boldsymbol{\Theta} \boldsymbol{c}. \tag{3.2}$$

From the construction of $\hat{\sigma}_t(\boldsymbol{\lambda})$, it is not immediate that the estimator is even positive for a particular choice of $\boldsymbol{\lambda}$, which is an essential property for a spot volatility estimator. The following result shows that non-negativeness of the OMK estimator can be guaranteed by choosing an appropriate $\boldsymbol{\lambda}$:

Proposition 3.1. *Under Assumptions 3.1 and 3.2, construct $\hat{\sigma}_t(\boldsymbol{\lambda})$ on an arbitrary interval I_n, then* $\text{Prob}(\hat{\sigma}_t(\boldsymbol{\lambda}) \geq 0) = 1$ *if $\boldsymbol{\lambda} \in \boldsymbol{\Lambda}$, where:*

$$\boldsymbol{\Lambda} := \{\boldsymbol{\lambda} \in \mathbb{R}^3 : \boldsymbol{\lambda}'\boldsymbol{\iota} = 1, \boldsymbol{\lambda}'\boldsymbol{\Theta}\boldsymbol{\iota} \geq \max\{0, \boldsymbol{\lambda}'\boldsymbol{\Theta}\boldsymbol{e}_1, \boldsymbol{\lambda}'\boldsymbol{\Theta}\boldsymbol{e}_3\}\}, \tag{3.3}$$

in which $\{\boldsymbol{e}_k\}_{k\in\{1,2,3\}}$ is the set of (3-by-1) standard basis vectors of the $\mathbb{R}^3$ space.

As negative volatility estimates are undesirable, in this paper we shall restrict ourselves to the choices of weights $\boldsymbol{\lambda} \in \boldsymbol{\Lambda}$ to guarantee non-negative spot volatility estimates. We derive the asymptotic properties of the OMK estimator in Theorem 3.1:

Theorem 3.1. *Suppose that Assumptions 3.1 and 3.2 hold true. For any $t \in I_n = [s, s + \Delta_n]$ and weight vector $\boldsymbol{\lambda} \in \boldsymbol{\Lambda}$, it holds as $\Delta_n \to 0$ that:*

$$\frac{\hat{\sigma}_t(\boldsymbol{\lambda})}{\sigma_t} = \boldsymbol{\lambda}'\boldsymbol{\Theta}\boldsymbol{z} + o_p(1), \tag{3.4}$$

where $\boldsymbol{z} = (\eta, \xi_1, \xi_2)'$ in which η, ξ_1, ξ_2 are the MAED, the range, and the absolute return of the scaled Brownian motion $(\Delta_n^{-1/2} W_t)_{t \in I_n}$, respectively.

Remark 3.1. Eq. (3.4) is a coupling result in the spirit of Bollerslev et al. (2021); Li et al. (2022), which exploits the fact that P behaves locally like the scaled Brownian motion $\sigma_t W$ on I_n as $\Delta_n \to 0$. By the scaling law of Brownian motion, we find $(\eta, \xi_1, \xi_2) \overset{d}{=} (\tilde{\eta}, \tilde{\xi}_1, \tilde{\xi}_2)$, which implies that $\mathrm{E}[\boldsymbol{\Theta}\boldsymbol{z}] = \boldsymbol{\iota}$ and hence $\mathrm{E}[\boldsymbol{\lambda}'\boldsymbol{\Theta}\boldsymbol{z}] = 1$. It should be clear now that the matrix $\boldsymbol{\Theta}$ ensures the asymptotic unbiasedness of $\hat{\sigma}_t(\boldsymbol{\lambda})$. Similar to the estimators in Bollerslev et al. (2021) and Li et al. (2022), $\hat{\sigma}_t(\boldsymbol{\lambda})$ is not consistent as the limiting variable $\boldsymbol{\lambda}'\boldsymbol{\Theta}\boldsymbol{z} = O_p(1)$. However, Theorem 3.1 provides a simple construction of the confidence bounds which allows for valid statistical inference. For a significance level $\alpha\%$, define the $(1-\alpha)\%$ Highest Density Interval (HDI) of a continuous random variable X as the shortest interval $\ell_\alpha(X) := [B_{\alpha^-}(X), B_{\alpha^+}(X)]$ that satisfies:

$$\mathrm{Prob}(X^{-1} \in \ell_\alpha(X)) = 1 - \alpha. \tag{3.5}$$

Based on Theorem 3.1 and given the estimate $\hat{\sigma}_t(\boldsymbol{\lambda})$, the $(1-\alpha)\%$ confidence interval of σ_t can be constructed as $[B_{\alpha^-}(\boldsymbol{\lambda}'\boldsymbol{\Theta}\boldsymbol{z})\hat{\sigma}_t(\boldsymbol{\lambda}), B_{\alpha^+}(\boldsymbol{\lambda}'\boldsymbol{\Theta}\boldsymbol{z})\hat{\sigma}_t(\boldsymbol{\lambda})]$, whose validity can be seen from a standard continuous mapping argument:

$$\begin{aligned} &\lim_{\Delta_n \to 0} \mathrm{Prob}(\sigma_t \in [B_{\alpha^-}(\boldsymbol{\lambda}'\boldsymbol{\Theta}\boldsymbol{z})\hat{\sigma}_t(\boldsymbol{\lambda}), B_{\alpha^+}(\boldsymbol{\lambda}'\boldsymbol{\Theta}\boldsymbol{z})\hat{\sigma}_t(\boldsymbol{\lambda})]) \\ &\quad = \mathrm{Prob}((\boldsymbol{\lambda}'\boldsymbol{\Theta}\boldsymbol{z})^{-1} \in \ell_\alpha(\boldsymbol{\lambda}'\boldsymbol{\Theta}\boldsymbol{z})) = 1 - \alpha, \end{aligned} \tag{3.6}$$

in which the critical values $B_{\alpha^\pm}(\boldsymbol{\lambda}'\boldsymbol{\Theta}\boldsymbol{z})$ can be computed based on the simulated distribution of $\boldsymbol{z}$. The width of the $(1-\alpha)\%$ HDI of $\boldsymbol{\lambda}'\boldsymbol{\Theta}\boldsymbol{z}$, namely $B_{\alpha^+}(\boldsymbol{\lambda}) - B_{\alpha^-}(\boldsymbol{\lambda})$, is a pivotal quantity that determines the width of the confidence interval of $\hat{\sigma}_t(\boldsymbol{\lambda})$. This provides a direct measure for the precision of $\hat{\sigma}_t(\boldsymbol{\lambda})$ with different choices of $\boldsymbol{\lambda}$.

Although Theorem 3.1 holds for arbitrary choices of $\boldsymbol{\lambda} \in \boldsymbol{\Lambda}$, we are only interested in a few important special cases. Following the discussion in Li et al. (2022), an optimal choice $\boldsymbol{\lambda}^*$ can be constructed by minimizing $\mathrm{Var}[\boldsymbol{\lambda}'\boldsymbol{\Theta}\boldsymbol{z}]$ subject to the asymptotic unbiasedness constraint $\boldsymbol{\lambda}'\boldsymbol{\iota} = 1$, which yields the BLUE estimator $\hat{\sigma}_t(\boldsymbol{\lambda}^*)$ among the class of estimators $\{\hat{\sigma}_t(\boldsymbol{\lambda}) : \boldsymbol{\lambda}'\boldsymbol{\iota} = 1\}$. Define the variance-covariance matrix of the limiting variable $\boldsymbol{\Theta}\boldsymbol{z}$ as:

$$\boldsymbol{\Sigma} := \mathrm{Var}[\boldsymbol{\Theta}\boldsymbol{z}] = \begin{pmatrix} \frac{\mu_2}{\mu_1^2} & \frac{\gamma_1}{\mu_1\nu_1} & \frac{\gamma_2}{\mu_1\psi_1} \\ \bullet & \frac{\nu_2}{\nu_1^2} & \frac{\gamma_0}{\nu_1\psi_1} \\ \bullet & \bullet & \frac{\psi_2}{\psi_1^2} \end{pmatrix} - \boldsymbol{\iota}\boldsymbol{\iota}' \approx \begin{pmatrix} 0.0650 & 0.0049 & -0.0857 \\ \bullet & 0.0888 & 0.1781 \\ \bullet & \bullet & 0.5708 \end{pmatrix}, \tag{3.7}$$

which follows from Eq. (2.4), the optimal choice is the solution to the following standard global minimum variance portfolio (GMVP) problem in the finance literature:

$$\boldsymbol{\lambda}^* := \underset{\{\boldsymbol{\lambda}:\boldsymbol{\lambda}'\boldsymbol{\iota}=1\}}{\operatorname{argmin}} \boldsymbol{\lambda}'\boldsymbol{\Sigma}\boldsymbol{\lambda} = \frac{\boldsymbol{\Sigma}^{-1}\boldsymbol{\iota}}{\boldsymbol{\iota}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\iota}} \approx (0.832, -0.030, 0.198)'. \tag{3.8}$$

It is interesting to see that the optimal weight assigned to w is almost zero, so that w contributes little to the optimal linear combination.[5] The minimized variance factor of the OMK estimator is thus $\operatorname{Var}[\boldsymbol{\lambda}^{*\prime}\boldsymbol{\Theta}\boldsymbol{z}] = (\boldsymbol{\iota}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\iota})^{-1} \approx 0.0368$. This should be compared to the variance factor for the optimal OK estimator in Li et al. (2022) which is $\operatorname{Var}[\boldsymbol{\lambda}^{o\prime}\boldsymbol{\Theta}\boldsymbol{z}] \approx 0.0625$, where $\boldsymbol{\lambda}^o$ solves the following restricted GMVP problem[6]:

$$\boldsymbol{\lambda}^o := \underset{\{\boldsymbol{\lambda}:\boldsymbol{\lambda}'\boldsymbol{\iota}=1,\boldsymbol{\lambda}'\boldsymbol{e}_1=0\}}{\operatorname{argmin}} \boldsymbol{\lambda}'\boldsymbol{\Sigma}\boldsymbol{\lambda} \approx (0, 1.294, -0.294)'. \tag{3.9}$$

One should verify that $\hat{\sigma}_t(\boldsymbol{\lambda}^o)$ is identical to the optimal OK estimator in Li et al. (2022). Importantly, we see that $\operatorname{Var}[\boldsymbol{\lambda}^{*\prime}\boldsymbol{\Theta}\boldsymbol{z}]/\operatorname{Var}[\boldsymbol{\lambda}^{o\prime}\boldsymbol{\Theta}\boldsymbol{z}] \approx 0.592$, which implies that the optimal OMK estimator shrinks the asymptotic variance of the optimal OK estimator by an impressive 40.8%! In fact, even if we construct the spot volatility estimator based solely on m by choosing $\boldsymbol{\lambda} = \boldsymbol{e}_1$, the corresponding variance factor is $\operatorname{Var}[\boldsymbol{e}_1'\boldsymbol{\Theta}\boldsymbol{z}] \approx 0.065$, which is only slightly larger than that of the optimal OK estimator but is considerably smaller than a spot volatility estimator constructed solely from w or $|r|$.

In this paper, we shall focus on the three choices of the weight vectors above, which corresponds to three spot volatility estimators. To simplify the notation, we write:

$$\hat{\sigma}_{t,OMK} := \hat{\sigma}_t(\boldsymbol{\lambda}^*), \quad \hat{\sigma}_{t,OK} := \hat{\sigma}_t(\boldsymbol{\lambda}^o), \quad \hat{\sigma}_{t,MAED} := \hat{\sigma}_t(\boldsymbol{e}_1),$$

which are referred to as the (optimal) OMK estimator, the (optimal) OK estimator, and the MAED estimator, respectively. One should also verify that all three estimators above are almost surely non-negative by Proposition 3.1. To further compare the precision of the three estimators, in Fig. 3.1 we plot their limiting distributions and the widths of the corresponding 90% HDIs. The figure clearly shows that: (1) the distributions of $(\boldsymbol{\lambda}^{o\prime}\boldsymbol{\Theta}\boldsymbol{z})^{-1}$ and $(\boldsymbol{e}_1'\boldsymbol{\Theta}\boldsymbol{z})^{-1}$ are almost identical, while the width of the 90% HDI is slightly wider for $(\boldsymbol{e}_1'\boldsymbol{\Theta}\boldsymbol{z})^{-1}$. This is in line with the finding that $\hat{\sigma}_{t,MAED}$ has a slightly higher variance than $\hat{\sigma}_{t,OK}$; (2) the distribution of $(\boldsymbol{\lambda}^{*\prime}\boldsymbol{\Theta}\boldsymbol{z})^{-1}$ is more concentrated than the other two statistics with a substantially tighter 90% HDI.

To quantify the precision gain of the OMK estimator in terms of the HDI width, we present numerical values and the critical values of the HDIs for the three densities in Table 3.1 with various choices of α below. The table shows that, for all choices of α, the HDI widths for $\boldsymbol{\lambda}^{o\prime}\boldsymbol{\Theta}\boldsymbol{z}$ and $\boldsymbol{e}_1'\boldsymbol{\Theta}\boldsymbol{z}$ are

[5] The optimal weight vector with the constraint $\lambda_w = 0$ is approximately $\boldsymbol{\lambda}^\star \approx (0.813, 0, 0.137)'$, with the asymptotic variance factor $\operatorname{Var}[\boldsymbol{\lambda}^{\star\prime}\boldsymbol{\Theta}\boldsymbol{z}] \approx 0.0368$ which is virtually the same as the case without the constraint.

[6] It is worth noting that the solution to this problem has a closed from $\boldsymbol{\lambda}^o = ((4\ln 2 - 2)^{-1}, 1 - (4\ln 2 - 2)^{-1})'$ with the minimized variance $\operatorname{Var}[\boldsymbol{\lambda}^{o\prime}\boldsymbol{\Theta}\boldsymbol{z}] = \frac{\pi}{2} + \frac{\pi}{16(1-2\ln 2)} - 1$, which can be derived from the analytical moments in Eq. (2.4).

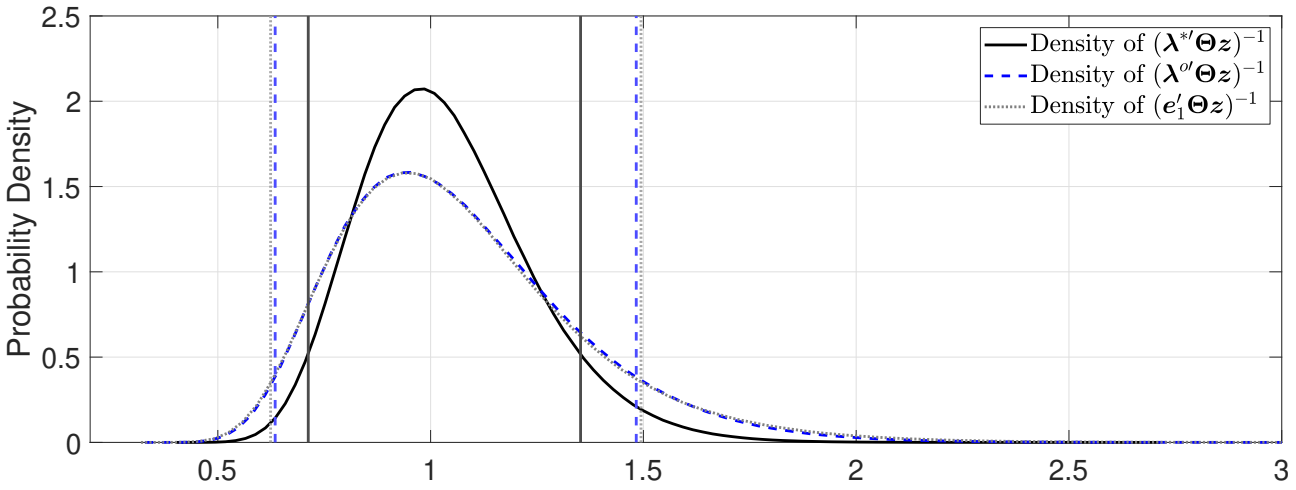


Figure 3.1: Plot of the simulated densities of $(\boldsymbol{\lambda}^{*\prime}\boldsymbol{\Theta}\boldsymbol{z})^{-1}$, $(\boldsymbol{\lambda}^{o\prime}\boldsymbol{\Theta}\boldsymbol{z})^{-1}$, and $(\boldsymbol{e}_1'\boldsymbol{\Theta}\boldsymbol{z})^{-1}$, and the corresponding critical values of the 90% HDIs. The densities are simulated based on 10^7 simulated Brownian paths with an Euler step size of 10^{-8}. For each density, the critical values for the 90% HDI, namely $B_{0.1\pm}(\boldsymbol{\lambda}'\boldsymbol{\Theta}\boldsymbol{z})$, are plotted as vertical lines that are of the same style as the density line.

very close. For $\boldsymbol{\lambda}^{*\prime}\boldsymbol{\Theta}\boldsymbol{z}$, we observe an about 25% reduction of the HDI widths at every α compared to that of $\boldsymbol{\lambda}^{o\prime}\boldsymbol{\Theta}\boldsymbol{z}$. Note that although our simulation setting is different to that in Li et al. (2022), the results of $\boldsymbol{\lambda}^{o\prime}\boldsymbol{\Theta}\boldsymbol{z}$ in Table 3.1 are consistent with those in Table 1 of Li et al. (2022) up to 2 digits after the decimal point, which proves the credibility of our simulation results.

X	$\boldsymbol{\lambda}^{o\prime}\boldsymbol{\Theta}\boldsymbol{z}$			$\boldsymbol{\lambda}^{*\prime}\boldsymbol{\Theta}\boldsymbol{z}$			$\boldsymbol{e}_1'\boldsymbol{\Theta}\boldsymbol{z}$		
α	Lower	Upper	Width	Lower	Upper	Width	Lower	Upper	Width
0.5	0.791	1.132	0.341	0.856	1.115	0.260	0.793	1.135	0.342
0.4	0.765	1.192	0.427	0.829	1.153	0.324	0.762	1.191	0.429
0.3	0.729	1.256	0.527	0.799	1.199	0.400	0.722	1.254	0.532
0.2	0.687	1.342	0.655	0.762	1.258	0.496	0.679	1.343	0.664
0.1	0.635	1.483	0.848	0.713	1.352	0.640	0.626	1.496	0.870

Table 3.1: Simulated critical values and widths for the highest density intervals of $\boldsymbol{\lambda}^{o\prime}\boldsymbol{\Theta}\boldsymbol{z}$, $\boldsymbol{\lambda}^{*\prime}\boldsymbol{\Theta}\boldsymbol{z}$ and $\boldsymbol{e}_1'\boldsymbol{\Theta}\boldsymbol{z}$ with different choices of α. The densities are simulated based on 10^7 simulated Brownian paths with an Euler step size of 10^{-8}. The Lower and Upper columns correspond to $B_{\alpha-}(X)$ and $B_{\alpha+}(X)$ as defined in Eq. (3.5) with X defined in the first row of the table, and the Width columns display the value of $B_{\alpha+}(X) - B_{\alpha-}(X)$.

To sum up, the comparisons of variances and HDI widths among $\hat{\sigma}_{t,OMK}$, $\hat{\sigma}_{t,OK}$ and $\hat{\sigma}_{t,MAED}$ provide strong evidence supporting the superiority of $\hat{\sigma}_{t,OMK}$ over $\hat{\sigma}_{t,OK}$, which shows the rich information about σ_t embedded in m. Indeed, $\hat{\sigma}_{t,MAED}$ performs almost as good as $\hat{\sigma}_{t,OK}$, which suggests that m alone is almost as informative about σ_t as w and $|r|$ combined. Importantly, as m mainly measures the inward price movements which are largely ignored by w and $|r|$, the addition of m to the candlestick data leads to a significant precision gain in the estimation of σ_t. Moreover, in Appendix B, we show that the MAED statistic can improve the asymptotic variance of an optimal candlestick-based spot variance estimator[7] by about 41%. These results clearly demonstrate the significance of MAED

[7]Note that candlestick-based spot variance estimator has a slightly smaller (about 3.4%) variance than the classic Garman

in estimating spot volatility and variance.

3.3 Jump-Robustness of the MAED Estimator and A Spot Test For Jumps

The asymptotic results in Theorem 3.1 hold under Assumption 3.1 which allows for Poisson-type jumps in P, implying that the OMK estimator is robust to jumps. However, as discussed in footnote 8 of Li et al. (2022), such jump robustness is purely probabilistic, as it exploits the fact that the probability of observing one jump in I_n is of the order $O_p(\Delta_n)$ which vanishes in the limit. Bollerslev et al. (2021) and Li et al. (2022) suggest both a truncation technique in the spirit of Mancini (2009) or a bipower-type extension similar to Barndorff-Nielsen and Shephard (2004) to guard against the possible occurrence of a jump in I_n. Both approaches require non-trivial modification to the original estimator, and the finite sample jump-robustness of the two approaches are not reported.

Instead of pursuing these jump-robust modifications to the OMK estimator, we show that the MAED estimator, $\hat{\sigma}_{t,MAED}$, is jump-robust in its original form:

Theorem 3.2. *Suppose Assumptions 3.1 and 3.2 hold true. Consider the jump-augmented process* $P_t^\circ = P_t + J\mathbb{1}_{\{t\geq\tau\}}$ *where* $\tau \in (s, s+\Delta_n)$ *is an arbitrary (n-dependent) jump time and* J *is any random variable strictly bounded away from zero. Define* $\hat{\sigma}^\circ_{t,OK}$ *and* $\hat{\sigma}^\circ_{t,MAED}$ *as the counterpart of* $\hat{\sigma}_{t,OK}$ *and* $\hat{\sigma}_{t,MAED}$ *constructed from* P° *instead of* P *on* I_n*, then it holds that:*

$$\frac{|\hat{\sigma}^\circ_{t,OK} - \hat{\sigma}_{t,OK}|}{\sigma_t} = O_p(\Delta_n^{-1/2}), \quad \frac{|\hat{\sigma}^\circ_{t,MAED} - \hat{\sigma}_{t,MAED}|}{\sigma_t} \leq \frac{3\xi_1}{\mu_1} + o_p(1). \tag{3.10}$$

Remark 3.2. Theorem 3.2 shows that the OK estimator explodes in the presence of a jump in the limit, which formally justifies the necessity of jump-robust modifications for the OK estimator as discussed in Li et al. (2022). On the contrary, the presence of a jump only introduces a stochastic bias to the $\hat{\sigma}_{t,MAED}$ estimator whose size is bounded above by $3\xi_1/\mu_1$, which is of order $O_p(1)$. This implies that $\hat{\sigma}_{t,MAED}$ does not explode in the limit, but it has an asymptotic relative absolute bias bounded above by $3\nu_1/\mu_1 \approx 4.328$. We stress that this result holds path-wise and independent of the size, direction or location of J, so no knowledge about the jump is needed for the result to be applicable. Consequently, the bias upper bound of $\hat{\sigma}_{t,MAED}$ is far from sharp. We analyse the actual finite sample bias of $\hat{\sigma}_{t,MAED}$ via simulation in Fig. 4.3, which suggests that the relative bias of $\hat{\sigma}_{t,MAED}$ is much smaller than the upper bound suggests.

Remark 3.3. We caution that $\hat{\sigma}_{t,MAED}$ is not robust to multiple jumps with opposite directions on I_n. Intuitively, two jumps that (partially) offset each other are inward price movements which are captured by the MAED and cause $\hat{\sigma}_{t,MAED}$ to explode in the limit. This could be caused by erroneous price entries that are quickly corrected in the dataset, which should be eliminated by appropriate data filtering rules such as Barndorff-Nielsen et al. (2009) before applying the estimator. As jumps are

and Klass (1980) estimator. Detailed analysis is presented in Appendix B.

found to be infrequent in equity prices (Huang and Tauchen, 2005; Christensen et al., 2014), robustness to a single jump in a short interval (e.g., up to 10 minutes) should be sufficient for the purpose of jump-robust spot volatility estimation.

The jump-robustness of $\hat{\sigma}_{t,MAED}$ leads naturally to a spot jump test. Among many possible specifications, we consider the following log-ratio statistic:

Corollary 3.1. *For some price process P observed on I_n, construct the following S-test statistic as the log-ratio of the OK and the MAED estimator:*

$$S_t := \ln \frac{\hat{\sigma}_{t,OK}}{\hat{\sigma}_{t,MAED}} = \ln \frac{\boldsymbol{\lambda}^{o\prime}\boldsymbol{\Theta}\boldsymbol{c}}{\boldsymbol{e}_1'\boldsymbol{\Theta}\boldsymbol{c}}. \tag{3.11}$$

Under the null hypothesis that P follows Assumptions 3.1 and 3.2, it holds that:

$$S_t = \ln \frac{\boldsymbol{\lambda}^{o\prime}\boldsymbol{\Theta}\boldsymbol{z}}{\boldsymbol{e}_1'\boldsymbol{\Theta}\boldsymbol{z}} + o_p(1). \tag{3.12}$$

Under the alternative hypothesis that P° in Theorem 3.2 is the underlying price process, we have $S_t \xrightarrow{p} \infty$.

Notice that the S-test statistic is pivotal as the limiting distribution is independent of the unknown spot volatility σ_t. This is much easier to construct than the local jump tests considered in e.g., Lee and Mykland (2007, 2012). Despite the fixed-k setting, the S-test has correct asymptotic size under the null and is consistent under the alternative. For the purpose of jump detection, one can compare S_t to the critical values of a right-tailed test, which can be retrieved from the simulated density of $\boldsymbol{z}$. The simulated 10%, 5% and 1% critical values are 0.375, 0.535 and 0.847, respectively (see Appendix C for details of this simulation).

3.4 MAED and Drift Bursts

Recently, Christensen et al. (2022) propose a drift-burst model for P which allows the drift and spot volatility processes b_t and σ_t to explode locally, violating Assumption 3.1. This model provides an econometric framework to analyse 'gradual jumps' in the observed stock prices, i.e., a gradual directional movement of stock prices occurring in a very short window. Andersen et al. (2021) show that the presence of drift bursts can lead to non-trivial finite sample bias when measuring the integrated variance of P which need to be correct. As the MAED is robust to extreme outward price movements, it is interesting to study to behaviour of the MAED in the presence of a drift burst, which is the purpose of this section.

To this end, we introduce a drift-burst alternative price process in the spirit of Christensen et al. (2022):

Assumption 3.3. *On the interval $I_n = [s, s+\Delta_n] \subset [0,T]$, the drift-burst alternative price process P^{db} is given by:*

$$P_t^{db} = P_s + b_s \int_s^t (u-s)^{-\alpha} du + \sigma_s \int_s^t (u-s)^{-\beta} dW_u, \quad t \in [s, s+\Delta_n],$$

where b_s and σ_s satisfy Assumption 3.1, and the constants $\alpha \in (1/2, 1)$ and $\beta \in (0, 1/2)$.

The above assumption states that the starting point of the interval I_n is the drift burst time, τ_{db}, in the terminology of Christensen et al. (2022). Effectively, I_n is a short time window right after the drift burst at time τ_{db} when the price reverts back to the pre-burst level. This is an important design as it ensures that P^{db} explodes in a unidirectional manner, which we exploit in this section. Alternatively, one can also consider the interval $[s-\Delta_n, s]$ where the drift burst occurs at the end of the interval. If $\tau_{db} \notin I_n$, then there is no drift burst in the limit and Theorem 3.1 continuous to hold.

Assumption 3.3 implies that as $\Delta_n \to 0$, both the drift and the spot volatility of $P^{db}_{s+\Delta_n}$ may explode. The parameters α and β control for the explosion rates of the drift and the volatility on I_n, respectively. Christensen et al. (2022) show that when $0 < \alpha - \beta < 1/2$, there is no local arbitrage opportunity, while a local arbitrage opportunity exists when $\alpha - \beta > 1/2$. We shall refer to the case $\alpha > 1/2$ and $\beta = 0$ as the pure drift burst case. Similarly, the case $\alpha = 0$ and $\beta > 0$ is called the pure volatility burst case, and a drift-volatility burst corresponds to the case $\alpha > 1/2$ and $\beta > 0$.

The OMK estimator has the following asymptotic properties when constructed from the drift-burst augmented price process in Assumption 3.3:

Proposition 3.2. *Suppose 3.2 holds true and construct the OMK estimators constructed from P^{db} on I_n as specified in Assumption 3.3. For any $\boldsymbol{\lambda} \in \boldsymbol{\Lambda}$ as $\Delta_n \to 0$, if $0 < \alpha - \beta < 1/2$, then:*

$$\Delta_n^{\beta} \frac{\hat{\sigma}_t(\boldsymbol{\lambda})}{\sigma_s} \xrightarrow{d} \frac{\boldsymbol{\lambda}' \boldsymbol{\Theta} \check{\boldsymbol{z}}}{\sqrt{1-2\beta}}, \tag{3.13}$$

where $\check{\boldsymbol{z}}$ is defined analogously as $\boldsymbol{z}$ in Theorem 3.1, but is constructed from a Brownian motion $(B_t)_{t \in I_n}$ independent of W. If $\alpha - \beta > 1/2$, then it holds that:

$$\Delta_n^{\beta} \hat{\sigma}_{t,OK} = O_p(\Delta_n^{1/2-\alpha+\beta}), \quad \Delta_n^{\beta} \hat{\sigma}_{t,MAED} \xrightarrow{p} 0. \tag{3.14}$$

Proposition 3.2 shows that the OMK estimator behaves very differently depending on the presence of a local arbitrage opportunity. When $0 < \alpha - \beta < 1/2$, $\hat{\sigma}_t(\boldsymbol{\lambda})$ still converges to a limiting distribution identical to that in Theorem 3.1 (since $\check{\boldsymbol{z}} \overset{d}{=} \boldsymbol{z}$ by construction), but with additional scaling factors Δ_n^{β} and $\sqrt{1-2\beta}$ due to the volatility burst. Intuitively, in this case the volatility burst dominates the drift burst, thus the drift-burst augmented price is equivalent in distribution to a price process without a drift burst when rescaled by the exploding volatility in the limit, which gives the same limiting distribution as in Theorem 3.1. However, when $\alpha - \beta > 1/2$, Eq. (3.14) suggests that $\hat{\sigma}_{t,MAED}$

is biased towards zero while $\hat{\sigma}_{t,OK}$ explodes in the limit, after adjusting for the volatility burst. In this case, the drift burst dominates the volatility burst, and the rescaled price process converges to a monotonic explosive drift process that drives $\hat{\sigma}_{t,MAED}$ towards zero.

Unfortunately, as in practice we do not know the value of β, none of the results in Proposition 3.2 are informative in practice. Nevertheless, Proposition 3.2 directly implies the following result for the S-test statistic as the unknown scaling factors cancel in the log-ratio:

$$\begin{aligned} S_t &\xrightarrow{d} \ln \frac{\boldsymbol{\lambda}^{o\prime}\boldsymbol{\Theta}\breve{\boldsymbol{z}}}{\boldsymbol{e}_1'\boldsymbol{\Theta}\breve{\boldsymbol{z}}}, \quad 0 < \alpha - \beta < 1/2, \\ S_t &\xrightarrow{p} \infty, \qquad\qquad \alpha - \beta > 1/2. \end{aligned} \tag{3.15}$$

Notice that in the no arbitrage case ($0 < \alpha - \beta < 1/2$), the asymptotic distribution of S_t is equal in distribution to the setting without a drift burst in Corollary 3.1. This shows that S_t has no power in the no arbitrage case, but it can consistently detect a drift burst with local arbitrage opportunity ($\alpha - \beta > 1/2$), which is in line with the testing procedure proposed in Christensen et al. (2022). However, as pointed out by Christensen et al. (2022), drift bursts tend to generate local directional outward price movements regardless of whether a local arbitrage opportunity exists, which has a larger impact on the OK estimator than the MAED estimator. Consequently, our simulation results in Fig. 4.5 show that the S-test still has non-trivial power against the presence of a drift burst in the non-arbitrage case.

3.5 Discretely Observed Prices

Assumption 3.2 plays a central role in determining the limiting distribution of the OMK estimator and the S-test statistic. However, empirically we only observe a finite number of price records on the interval I_n, so Assumption 3.2 is violated. Suppose that the price process is not observed with measurement error, the discretely observed price records lead to underestimations of the range w and the MAED m on I_n. This not only biases the candlestick and MAED-based estimators downwards, but also distorts the asymptotic distributions and the confidence intervals for these estimators. In this section, we quantify the discretization bias and propose an explicit and easy-to-implement correction method.

To examine the impact of discrete observations formally, we need to make an assumption about the price observations in the interval $[s, s + \Delta_n]$ as $\Delta_n \to 0$. In particular, we are interested in a scenario where the number of observations are fixed for any subinterval $[s, s+\Delta_n] \subset [0, T]$ in the limit, which naturally requires an increasing number of observations globally on $[0, T]$, a setting similar to the infill asymptotics in the classic realized volatility literature (see e.g. Jacod (1997); Andersen et al. (2001); Barndorff-Nielsen and Shephard (2002)).

Borrowing ideas from the infill asymptotics literature, we make the following assumption about

the observation times of P on $[0,T]$:

Assumption 3.4. *On $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$, let $\alpha = (\alpha_t)$ denote a semimartingale which satisfies the same conditions as P in Assumption 3.1. For each n, P is observed at a sequence of strictly increasing stopping times $0 \leq \tau_{n,0} < \tau_{n,1} < \ldots < \tau_{n,N_n} \leq T$, where N_n is the total number of observations on $[0,T]$. We assume that:*

(1) For every $t < T_m$, we have $1/\bar{\alpha} \leq \alpha_{t-} \leq \bar{\alpha}$ for some strictly positive constant $\bar{\alpha}$.

(2) If $(\mathcal{F}_t^n)$ is the smallest filtration containing $(\mathcal{F}_t)$ and w.r.t. which all $\{\tau_{n,i}\}_{i=1,2,\ldots}$ are stopping times, then for each i, the variable $\Delta\tau_{n,i} := \tau_{n,i} - \tau_{n,i-1}$ is, conditionally on $\mathcal{F}^n_{\tau_{n,i-1}}$, independent of $\mathcal{F}_\infty := \bigvee_{t>0} \mathcal{F}_t$.

(3) In restriction to the set $\{\tau_{n,i} < T_m\}$, there exists a strictly positive and decreasing sequence $\delta_n \to 0$ and some constant $\kappa > 0$ such that for each i, $|\Delta\tau_{n,i} - \delta_n/\alpha_{\tau_{n,i-1}}| = O_p(\delta_n^{1+\kappa})$.

Remark 3.4. Assumption 3.4 is a special case of the more general setting in Assumption (O) of Jacod et al. (2017). In detail, the process α_t controls for the 'spot' observation arrival rate at time t, which is assumed to be bounded above and away from zero by condition (1). Condition (2) is a conditional exogeneity assumption of the observation times which ensures that P is still a semimartingale relative to the filtration $(\mathcal{F}_t^n)$ with the same dynamics. Condition (3) allows the sampling times to be time-varying and random, but requires the observation times to be locally equidistant in a vanishing window of order δ_n. In the context of the examples given in Jacod et al. (2017), this condition is trivially satisfied by the regular or the time-changed regular sampling scheme, but is not satisfied by the modulated Poisson or the predictably-modulated random walk sampling scheme which has $|\Delta\tau_{n,i} - \delta_n/\alpha_{\tau_{n,i-1}}| = O_p(\delta_n)$. In this case, the asymptotic distribution of the OMK estimator depends on the distribution of the sampling times which can in theory be an arbitrary positive distribution. Therefore, to derive a feasible asymptotic theory for the OMK estimator in the discrete case, we shall stick with condition (3).

Under Assumption 3.4, it suffices to consider the time-deformed process $(P(i))_{i=0:N_n}$ where $P(i) := P_{\tau_{n,i}}$. Also, we assume that the interval of interest $I_n = [s, s+\Delta_n]$ contains the grid of discrete observations $I_n^{(q)} := \{i_s, i_s+1, \ldots i_s+q\}$ for some fixed q and $i_s \leq N_n - q$ such that $I_n = [\tau_{n,i_s}, \tau_{n,i_s+q}]$. This design ensures that we observe exactly $q+1$ price observations in the limit[8]. In this case, the length of the interval $\Delta_n = \tau_{n,i_s+q} - \tau_{n,i_s}$ is implicitly determined by the choice of q and the (unobserved) sequence δ_n which represents the 'average mesh size'. One can show (see Eq. (A.44)) that by Assumption 3.4(3), we have $\Delta_n = O_p(\delta_n)$ for any fixed q, so I_n indeed shrinks as $n \to \infty$,

[8]Clearly, adding leading and trailing intervals without observations to I_n does not change the statistics associated with I_n, thus we can restrict ourselves to consider I_n which both starts and ends with an observation.

which is similar to the continuous case in the previous section, but with a possibly random interval length.

For the interval I_n and the observed prices $(P(i))_{i\in I_n^{(q)}}$, the associated MAED, range and return statistics reduce to:

$$m = \max_{1\le j\le q}\{w(j) - |r(j)|\}, \quad w = w(q), \quad r = r(q), \tag{3.16}$$

where $w(j) := \max_{1\le i,i'\le j}|P(i_s+i) - P(i_s+i')|$ and $r(j) := P(i_s+j) - P(i_s)$ for $j \in \{1,\dots,q\}$ are the discrete running range and return of P on the grid of observations, respectively. As two special cases, when $q = 1$, $\Pr(m = 0) = \Pr(w = |r|) = 1$ as only the return information is available and there is no 'inward' price movement. When $q = 2$, it is trivial to show that $m + |r| = w$, so the MAED does not add any information to the candlestick data.

To construct the OMK estimator in the discrete case (which we refer to as the *discrete OMK estimator*), we introduce some further notations related to a discretely observed Brownian motion. For some fixed $q > 1$, consider the standard Brownian motion $\tilde{W}$ on $[0,1]$ observed equidistantly with q intervals, i.e., $(\tilde{W}^{(q)}(i))_{i=0:q}$, where $\tilde{W}^{(q)}(i) := \tilde{W}_{i/q}$ denotes the ith equidistant observation from $\tilde{W}$. Denote the associated discrete MAED and range statistics of $(\tilde{W}^{(q)}(i))_{i=0:q}$ as $\tilde{\eta}^{(q)}$ and $\tilde{\xi}_1^{(q)}$, and note that the absolute return statistic $\tilde{\xi}_2$ is not affected by the discrete observations. The corresponding moments of these discrete statistics are defined similarly in Eq. (2.4) with the additional superscript (q), e.g., $\mu_1^{(q)} := \mathrm{E}[\tilde{\eta}^{(q)}]$. For any $q > 1$, these moments can be easily simulated from discretely observed Brownian motions, and the moments presented in Eq. (2.4) can be considered as the limits of their discrete counterparts as $q \to \infty$. In Table C.1 of Appendix C, we present the simulated values of the discrete moments and a polynomial approximation inspired by the asymptotic expansion in Proposition 3 of Asmussen et al. (1995). This allows us to compute the discrete moments for any choice of q without additional simulation. As will be shown, these discrete moments play the same role as their continuous counterparts in the construction and the asymptotic properties of the OMK estimator in the continuous case.

We are now ready to construct the discrete OMK estimator. Start with the MAED-candlestick vector $\boldsymbol{c}$ constructed from $(P(i))_{i\in I_n^{(q)}}$, the discrete OMK estimator is defined as:

$$\hat{\sigma}_t^{(q)}(\boldsymbol{\lambda}) := \Delta_n^{-1/2}\boldsymbol{\lambda}'\boldsymbol{\Theta}^{(q)}\boldsymbol{c}, \tag{3.17}$$

where $\boldsymbol{\Theta}^{(q)} := \mathrm{diag}((\mu_1^{(q)})^{-1}, (\nu_1^{(q)})^{-1}, \psi_1^{-1})$ is the discrete counterpart of $\boldsymbol{\Theta}$ that ensures the asymptotic unbiasedness of the discrete OMK estimator. To guarantee the non-negativity of $\hat{\sigma}_t^{(q)}(\boldsymbol{\lambda})$, we require $\boldsymbol{\lambda} \in \boldsymbol{\Lambda}^{(q)}$, where:

$$\boldsymbol{\Lambda}^{(q)} := \{\boldsymbol{\lambda} \in \mathbb{R}^3 : \boldsymbol{\lambda}'\boldsymbol{\iota} = 1, \boldsymbol{\lambda}'\boldsymbol{\Theta}^{(q)}\boldsymbol{\iota} \ge \max\{0, \boldsymbol{\lambda}'\boldsymbol{\Theta}^{(q)}\boldsymbol{e}_1, \boldsymbol{\lambda}'\boldsymbol{\Theta}^{(q)}\boldsymbol{e}_3\}\}, \tag{3.18}$$

which can be proved in the same manner as Proposition 3.1 in the continuous case. The following result characterizes the asymptotic distribution of $\hat{\sigma}_t^{(q)}(\boldsymbol{\lambda})$:

Proposition 3.3. *Suppose Assumption 3.1 and Assumption 3.4 hold true. For any fixed integer $q > 1$, $\boldsymbol{\lambda} \in \boldsymbol{\Lambda}^{(q)}$ and $t \in I_n$, it holds as $\delta_n \to 0$ that:*

$$\frac{\hat{\sigma}_t^{(q)}(\boldsymbol{\lambda})}{\sigma_t} = \boldsymbol{\lambda}' \boldsymbol{\Theta}^{(q)} \boldsymbol{z}^{(q)} + o_p(1), \tag{3.19}$$

where $\boldsymbol{z}^{(q)} := (\eta^{(q)}, \xi_1^{(q)}, \xi_2)'$, in which $\eta^{(q)}$ and $\xi_1^{(q)}$ are the MAED and the range of the scaled equidistantly observed Brownian motion $(\Delta_n^{-1/2} W^{(q)}(i))_{i=0:q}$ such that $W^{(q)}(i) := W_{\tau_{n,i_s} + i\Delta_n/q}$ is the ith equidistant observation of W on I_n.

Remark 3.5. Analogous to the discussion in Remark 3.1, it holds by the Brownian scaling law that $(\eta^{(q)}, \xi_1^{(q)}, \xi_2) \overset{d}{=} (\tilde{\eta}^{(q)}, \tilde{\xi}_1^{(q)}, \tilde{\xi}_2)$, which implies that $\hat{\sigma}_t^{(q)}(\boldsymbol{\lambda})$ is asymptotically unbiased for all $\boldsymbol{\lambda} \in \boldsymbol{\Lambda}^{(q)}$. It is worth noting that the distribution of the limiting variable $\boldsymbol{z}^{(q)}$ comes from the equidistantly observed Brownian motion $(\Delta_n^{-1/2} W^{(q)}(i))_{i=0:q}$ which does not depend on the (possibly random) observation times of P. This coupling result is due to Assumption 3.4 which ensures that the observation times on I_n converges to a deterministic equidistant grid as Δ_n. As the distribution of $\boldsymbol{z}^{(q)}$ can be simulated for any fixed q to an arbitrary precision, Proposition 3.3 allows us to construct valid confidence intervals for $\hat{\sigma}_t^{(q)}(\boldsymbol{\lambda})$ in the same manner as described in Remark 3.1.

Analogous to Eqs. (3.8) and (3.9), we can construct the optimal discrete OMK and OK estimators by choosing the variance-minimizing weight vectors $\boldsymbol{\lambda}^*$ and $\boldsymbol{\lambda}^o$ for any fixed $q > 1$ based on the discrete variance-covariance matrix $\boldsymbol{\Sigma}^{(q)}$, whose inputs can be computed from Table C.1 for any choice of q. The notation $\hat{\sigma}_{t,OMK}^{(q)}$, $\hat{\sigma}_{t,OK}^{(q)}$ and $\hat{\sigma}_{t,MAED}^{(q)}$ are understood as the discrete counterparts of $\hat{\sigma}_{t,OMK}$, $\hat{\sigma}_{t,OK}$ and $\hat{\sigma}_{t,MAED}$, respectively. It should also be clear that $\hat{\sigma}_{t,OMK}^{(q)}$ is BLUE among the class of estimators $\{\hat{\sigma}_t^{(q)}(\boldsymbol{\lambda}) : \boldsymbol{\lambda}'\boldsymbol{\iota} = 1\}$. We present the optimal OK and OMK weights as a function of q in Fig. 3.2. The figure shows that both $\boldsymbol{\lambda}^*$ and $\boldsymbol{\lambda}^o$ converge to their corresponding limiting values when $q \to \infty$ as expected, and the speed of convergence for $\boldsymbol{\lambda}^*$ appears to be slower than that of $\boldsymbol{\lambda}^o$. For small q (e.g., $q \leq 50$), the optimal weight vectors are very different from their limiting versions based on the continuous assumption, which reveals the impact of sparsely observed prices.

We proceed to examine the precision of $\hat{\sigma}_{t,OMK}^{(q)}$ and $\hat{\sigma}_{t,MAED}^{(q)}$ relative to $\hat{\sigma}_{t,OK}^{(q)}$ across various choices of q. To this end, we compute the simulated asymptotic variances and 90% HDIs of the limiting variables of the three estimators and benchmark them to the values of $\hat{\sigma}_{t,OK}^{(q)}$ for selective choices of qs, which are presented in Table 3.2.

Table 3.2 shows that, as $q \to \infty$, both the asymptotic variances and the 90% HDI widths for all three estimators decrease monotonically and converge to the corresponding value in the continuous case. Comparing the performance of the OK to the OMK estimator, we find that: (1) both the

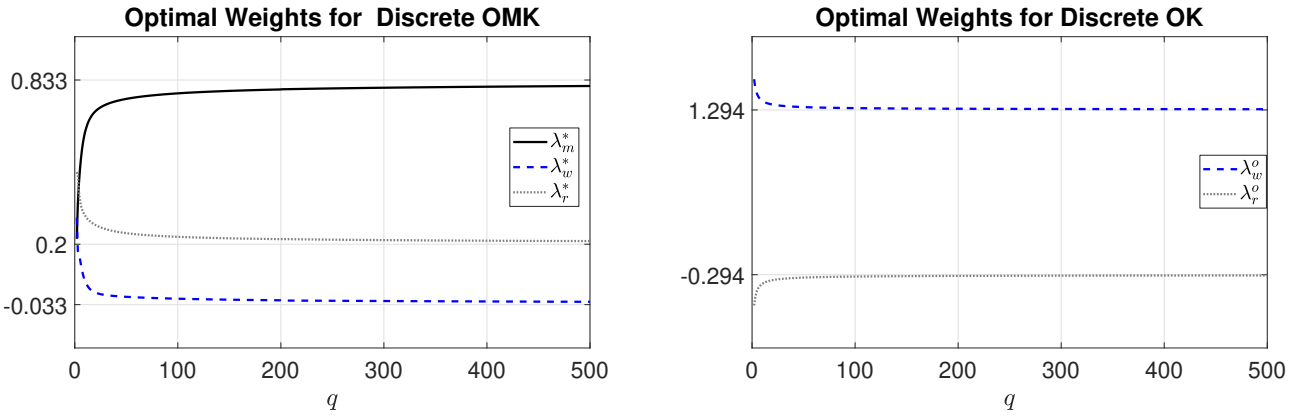


Figure 3.2: Plot of the optimal weight vectors $\boldsymbol{\lambda}^* = (\lambda_m^*, \lambda_w^*, \lambda_r^*)'$ for the discrete OMK estimator and $\boldsymbol{\lambda}^o = (0, \lambda_w^o, \lambda_r^o)'$ for the discrete OK estimator with $q \leq 500$. The optimal weight vectors are computed based on the simulated/approximated moments in Table C.1. The y-axis tick labels show the limiting values of the weights as $q \to \infty$.

	Asymptotic Variance					90% HDI Width				
q	OMK	OK	MAED	$\frac{\text{OMK}}{\text{OK}}$	$\frac{\text{MAED}}{\text{OK}}$	OMK	OK	MAED	$\frac{\text{OMK}}{\text{OK}}$	$\frac{\text{MAED}}{\text{OK}}$
1	0.5	0.5	–	1	–	6.148	6.148	–	1	–
2	0.284	0.284	2.327	1	8.203	2.450	2.450	∞	1	∞
3	0.191	0.206	1.025	0.928	4.977	1.744	1.827	∞	0.955	∞
4	0.148	0.171	0.635	0.863	3.710	1.448	1.581	∞	0.915	∞
5	0.123	0.152	0.458	0.814	3.016	1.284	1.450	7.317	0.886	5.048
6	0.108	0.139	0.360	0.777	2.589	1.180	1.367	3.681	0.863	2.694
7	0.0977	0.130	0.301	0.749	2.306	1.108	1.309	2.748	0.847	2.099
8	0.0902	0.124	0.261	0.729	2.109	1.056	1.267	2.315	0.834	1.828
9	0.0845	0.119	0.233	0.712	1.967	1.016	1.233	2.063	0.824	1.673
10	0.0801	0.115	0.213	0.699	1.858	0.985	1.206	1.899	0.816	1.574
20	0.0614	0.0953	0.139	0.644	1.461	0.846	1.079	1.385	0.784	1.284
30	0.0553	0.0881	0.119	0.628	1.351	0.798	1.030	1.253	0.775	1.217
40	0.0521	0.0840	0.109	0.620	1.296	0.772	1.002	1.186	0.770	1.183
50	0.0500	0.0814	0.103	0.615	1.261	0.755	0.984	1.144	0.768	1.162
60	0.0486	0.0796	0.0985	0.611	1.237	0.744	0.971	1.114	0.766	1.147
70	0.0475	0.0782	0.0953	0.608	1.219	0.735	0.961	1.092	0.764	1.136
80	0.0467	0.0770	0.0929	0.606	1.205	0.728	0.954	1.075	0.763	1.127
90	0.0460	0.0761	0.0909	0.605	1.194	0.722	0.947	1.061	0.762	1.120
100	0.0455	0.0754	0.0893	0.604	1.185	0.717	0.942	1.049	0.762	1.114
500	0.0403	0.0680	0.0747	0.593	1.098	0.672	0.889	0.944	0.756	1.062
1000	0.0392	0.0663	0.0717	0.591	1.080	0.662	0.877	0.921	0.755	1.051
10000	0.0375	0.0637	0.0670	0.589	1.052	0.647	0.857	0.886	0.754	1.033
100000	0.0370	0.0629	0.0656	0.589	1.043	0.642	0.851	0.875	0.754	1.028
∞	0.0368	0.0625	0.0650	0.589	1.039	0.640	0.848	0.870	0.754	1.026

Table 3.2: Simulated asymptotic variances and the 90% HDI widths of the limiting variables of $\hat{\sigma}_{t,OMK}^{(q)}$, $\hat{\sigma}_{t,OK}^{(q)}$ and $\hat{\sigma}_{t,MAED}^{(q)}$, namely $\boldsymbol{\lambda}^{*\prime}\boldsymbol{\Theta}^{(q)}\boldsymbol{z}^{(q)}$, $\boldsymbol{\lambda}^{o\prime}\boldsymbol{\Theta}^{(q)}\boldsymbol{z}^{(q)}$ and $\boldsymbol{e}_1'\boldsymbol{\Theta}^{(q)}\boldsymbol{z}^{(q)}$, with various choices of q. The simulation setting is described in Appendix C. The columns with heading $\frac{\text{OMK}}{\text{OK}}$ and $\frac{\text{MAED}}{\text{OK}}$ present the ratio of the corresponding statistics to those of the OK estimator.

asymptotic variance and the 90% HDI width of the OMK estimator are superior to those of the OK estimator, starting from $q = 3$; (2) with $q \geq 20$, the discrete OMK estimator starts to outperform the continuous OK estimator in terms of both the asymptotic variance and the 90% HDI width; (3) for any $q \geq 100$, the asymptotic variance (resp. 90% HDI width) reduction of the OMK estimator over the OK estimator is about 40% (resp. 24%) already, which is close to the precision gain in the continuous case.

For the performance of the MAED estimator, its asymptotic variance and the 90% HDI width converge slower to the limiting values as $q \to \infty$. We note that the MAED estimator is undefined for $q = 1$ since $\eta^{(1)} \equiv 0$, and its 90% HDI width for $q \leq 4$ is infinite due to the non-trivial probability[9] of the event $\eta^{(q)} \equiv 0$. In general, the MAED estimator for small q performs rather poorly when compared to the OK estimator. With $q \geq 100$, the performance of the MAED estimator becomes somewhat more reliable.

Concluding from above, Table 3.2 clearly demonstrates that the OMK estimator dominates the OK estimator for any choice of q, and the precision gain increases with q. Therefore, when ultra high-frequency data is available, the OMK estimator is always preferred over the OK estimator, regardless of the choice of q. With about 100 observations in the interval I_n, the OMK estimator can reduce the variance of the OK estimator by about 40% and the 90% HDI width by about 24%. The HDI widths with different confidence levels are also reduced by a similar amount, which can be seen from Table C.2 in Appendix C. As to the MAED estimator, we do not recommend to use it as a standalone spot volatility estimator due to its inferior precision relative to the OMK estimator.

However, $\sigma^{(q)}_{t,MAED}$ preserves the jump-robustness as a discrete version of Theorem 3.2 holds under Assumption 3.4. To see this, we note that for any $q > 1$, m constructed from $(P(i))_{i \in I_n^{(q)}}$ is bounded above by its counterpart constructed from $(P_t)_{t \in I_n}$, which also holds when a jump is present. As a special case, with $q = 2$ one can show that $m \leq \min(|r_1|, |r_2|)$, where r_1 and r_2 are the returns over the two subintervals on I_n which satisfy $r_1 + r_2 = r$. Hence, the jump-robustness of m in this case can also be seen from the nearest neighbourhood truncation method of Andersen et al. (2012). Consequently, $\sigma^{(q)}_{t,MAED}$ provides jump-robust spot volatility estimates in the discrete case at the cost of a mild precision loss.

The jump-robustness of $\sigma^{(q)}_{t,MAED}$ leads naturally to a discrete S-test with the statistic $S^{(q)}_t := \ln(\hat{\sigma}^{(q)}_{t,OK}/\hat{\sigma}^{(q)}_{t,MAED})$. From a discrete version of Corollary 3.1, the asymptotic distribution of $S^{(q)}_t$ under the null is the discrete counterpart of Eq. (3.12) while the test statistic diverges under the alternative. We provide critical values of the discrete version of the test for any $q \in \mathbb{N}$ in Table C.2 of Appendix C. We caution that such critical values are not always available when $\text{Prob}(\eta^{(q)} = 0)$ is non-trivial,

[9] It is easy to show that $\text{Prob}(\eta^{(q)} = 0) = 2^{1-q}$, which is the probability that a q step Gaussian random walk is monotonic.

which can be avoided by choosing some moderately large q (e.g., $q \geq 10$). Also, a discrete version[10] of Proposition 3.2 holds in the presence of a drift burst since discrete observations do not change the asymptotic order of the drift burst. As a result, the test statistic $S_t^{(q)}$ is still valid for drift burst detection in the discrete case.

As to the performance of the continuous OMK estimator in the discrete case, we have $\hat{\sigma}_t(\boldsymbol{\lambda})/\sigma_t = \boldsymbol{\lambda}'\boldsymbol{\Theta}\boldsymbol{z}^{(q)} + o_p(1)$ due to Proposition 3.3, which suggests that the continuous OMK estimator is in general not asymptotically unbiased, since $\mathrm{E}[\boldsymbol{\lambda}'\boldsymbol{\Theta}\boldsymbol{z}^{(q)}] = \lambda_m\mu^{(q)}/\mu_1 + \lambda_w\nu_1^{(q)}/\nu_1 + \lambda_r < 1$. To quantify the bias of the continuous OMK estimator and compare its performance against the discrete OMK estimator, we compute the asymptotic relative bias (ARBias), the asymptotic mean squared relative error (AMSRE) and the asymptotic 90% coverage rate (ACR90%) of $\hat{\sigma}_t(\boldsymbol{\lambda})$ for a wide range of q, which are defined as follows:

$$\begin{aligned} \mathrm{ARBias}^{(q)} &:= \mathrm{E}[\boldsymbol{\lambda}'\boldsymbol{\Theta}\boldsymbol{z}^{(q)}], \quad \mathrm{AMSRE}^{(q)} := \frac{\mathrm{E}[(\boldsymbol{\lambda}'\boldsymbol{\Theta}\boldsymbol{z}^{(q)} - 1)^2]}{\mathrm{Var}[\boldsymbol{\lambda}'\boldsymbol{\Theta}^{(q)}\boldsymbol{z}^{(q)}]}, \\ \mathrm{ACR90\%}^{(q)} &:= \mathrm{Prob}((\boldsymbol{\lambda}'\boldsymbol{\Theta}\boldsymbol{z}^{(q)})^{-1} \in \ell_{0.9}(\boldsymbol{\lambda}'\boldsymbol{\Theta}\boldsymbol{z})). \end{aligned} \tag{3.20}$$

Note that $\mathrm{AMSRE}^{(q)}$ compares the asymptotic mean squared relative error of the continuous OMK estimator against the asymptotic variance factor of the discrete OMK estimator (since the latter is unbiased), which measures the relative performance of the continuous OMK estimator to its discrete version in the mean squared sense. We present the above statistics for the continuous optimal OMK, OK and the MAED estimators in Table 3.3.

Table 3.3 clearly presents the bias induced by discrete observations for the three continuous estimators, which can be quite substantial for small q (e.g., more than 20% in magnitude for $q \leq 10$), but in general diminishes as $q \to \infty$. The MAED estimator is most sensitive to this bias, followed by the OMK and the OK estimators in descending order. This bias in general inflates the AMSRE of the continuous estimators relative to their discrete versions when q is small (except for the MAED estimator with $q \leq 5$ where the variance of the discrete MAED estimator explodes), but small AMSRE reductions are expected for moderate to large q due to a shrinkage effect[11]. However, regardless of the choice of q, the coverage rates of the continuous estimators are all distorted from the nominal level of 90%, and the distortion is sizeable even with $q = 100$. These results demonstrate the importance of the correction for discrete observations when constructing the OMK estimators, especially when q is relatively small. For $q \geq 1000$ and suppose the measurement errors are negligible, one can safely ignore the effect of discrete observations and use the simple continuous OMK estimators.

[10] Note that one needs to alter the asymptotic distribution in Eq. (3.13) to account for discrete observations. However, this asymptotic distribution is infeasible in practice as it depends on the unknown parameter β. We therefore omit this result in the paper.

[11] This is also known as the Stein effect (Stein, 1956). Intuitively, the continuous estimators are biased towards zero, which also simultaneously reduce their variances and improve the mean squared error of the unbiased discrete estimators.

	ARBias$^{(q)}$			AMSRE$^{(q)}$			ACR90%$^{(q)}$		
q	OMK	OK	MAED	OMK	OK	MAED	OMK	OK	MAED
2	0.305	0.487	0.149	1.816	1.165	0.333	0.030	0.261	0.053
3	0.393	0.559	0.257	2.117	1.257	0.605	0.063	0.346	0.101
4	0.458	0.607	0.337	2.233	1.272	0.806	0.101	0.412	0.148
5	0.508	0.641	0.398	2.253	1.261	0.950	0.142	0.463	0.194
6	0.547	0.667	0.446	2.225	1.241	1.050	0.182	0.505	0.236
7	0.579	0.688	0.485	2.177	1.219	1.118	0.222	0.538	0.275
8	0.605	0.706	0.517	2.120	1.197	1.162	0.260	0.567	0.311
9	0.627	0.721	0.543	2.062	1.178	1.189	0.294	0.590	0.343
10	0.645	0.733	0.566	2.006	1.159	1.205	0.326	0.611	0.373
20	0.747	0.805	0.690	1.611	1.048	1.166	0.532	0.722	0.558
30	0.792	0.838	0.746	1.415	1.001	1.100	0.630	0.768	0.645
40	0.820	0.859	0.779	1.302	0.976	1.055	0.686	0.795	0.695
50	0.839	0.873	0.802	1.229	0.961	1.024	0.723	0.812	0.728
60	0.852	0.883	0.819	1.179	0.952	1.003	0.748	0.823	0.751
70	0.863	0.892	0.833	1.142	0.946	0.987	0.767	0.832	0.768
80	0.872	0.898	0.843	1.114	0.942	0.976	0.782	0.839	0.781
90	0.879	0.904	0.852	1.093	0.939	0.967	0.793	0.845	0.792
100	0.886	0.909	0.860	1.075	0.937	0.959	0.803	0.849	0.801
500	0.949	0.958	0.937	0.966	0.944	0.931	0.877	0.889	0.875
1000	0.964	0.970	0.956	0.963	0.955	0.941	0.889	0.896	0.887
∞	1	1	1	1	1	1	0.9	0.9	0.9

Table 3.3: Performances of the continuous optimal OMK ($\hat{\sigma}_{t,OMK}$), optimal OK ($\hat{\sigma}_{t,OK}$) and MAED ($\hat{\sigma}_{t,MAED}$) estimators in the discrete observation setting. The evaluation metrics ARBias, AMSRE and CR90% are defined in Eq. (3.20) with various choices of q in the first column of the table, where $q+1$ represents the number of observations on the interval I_n used to construct the estimators.

3.6 Measurement Errors

In practice, we do not observe the efficient prices $(P(i))_{i=0:N_n}$, but rather prices contaminated by measurement errors due to the presence of bid-ask spread, rounding, and other trading activities, commonly known as the market microstructure (MMS) noise. The impact of measurement errors on the realized volatility (RV) estimator or the spot volatility estimator in the infinite-k setting is well-studied (see e.g., Aït-Sahalia and Jacod (2014); Bandi and Russell (2008); Hansen and Lunde (2006); Zhang et al. (2005); Zu and Peter Boswijk (2014); Li et al. (2020); Li and Linton (2022) among many others). The impact of measurement errors on spot volatility estimators in the fixed-k framework, however, is not yet studied in the literature, as Bollerslev et al. (2021) and Li et al. (2022) implicitly assume Assumption 3.2 in their analyses, which ignores both discrete observations and measurement errors. As the MAED exploits the full price path on I_n which can be more sensitive to the range and the return, in this section we analyse the impact of measurement errors on the OMK estimator and discuss potential remedies.

We start with the following assumption about the measurement error on some interval I_n with q intervals:

Assumption 3.5. *Suppose Assumptions 3.1 and 3.4 hold true. We assume that the observed price*

process on $[0,T]$*, denoted by* $(P^{\epsilon}(i))_{i=0:N_n}$*, takes the following form:*

$$P^{\epsilon}(i) := P(i) + \epsilon(i), \forall i, \tag{3.21}$$

where $(\epsilon(i))_{i\in\mathbb{Z}}$ *is a zero-mean time series such that for all* $i \in \mathbb{Z}$*, we have* $\epsilon(i) = O_p(1)$ *and:*

$$k_n^{-1}\sum_{j=1}^{k_n}|\epsilon(i+j)| = O_p(k_n^{d-1/2}), \quad d \in [0, 1/2), \tag{3.22}$$

as $k_n \to \infty$*.*

Remark 3.6. The $O_p(1)$ noise assumption is referred to as the 'fixed noise' setting in Aït-Sahalia and Jacod (2014), which is commonly discussed in the context of integrated variance estimation. The above assumption only requires the noise process to be zero mean and asymptotically negligible when we average over k_n consecutive terms as $k_n \to \infty$. Apart from these conditions, we do not restrict the form of heterescedasticity, endogeneity, or dependence of the noise, as these are irrelevant in the asymptotic analysis below. This is in stark contrast to the study of noise-robust integrated variance estimation, which normally require intricate assumptions on the structure of the MMS noise process for identification (e.g., Hansen and Lunde (2006); Bandi and Russell (2008); Barndorff-Nielsen et al. (2008); Aït-Sahalia et al. (2011); Varneskov (2017); Jacod et al. (2019)). The parameter d controls for the level of long memory in $\epsilon(i)$, similar to the degree of fractional integration as in Granger and Joyeux (1980); Baillie (1996), and $d < 1/2$ is required to ensure that the variance of the sample mean in Eq. (3.22) vanishes in the limit. This assumption is satisfied by almost all existing models for the MMS noise. For example, any short memory noise specification as in Zhang (2006) and Aït-Sahalia et al. (2011) trivially satisfies the above assumption with $d = 0$. It can be shown that the more complicated noise specifications in Varneskov (2017); Jacod et al. (2019) also satisfy this assumption for $d < 1/2$ according to their assumption on the polynomial α- or ρ-mixing coefficient.

We first show that the coupling result in Proposition 3.3 no longer holds and the discrete OMK estimator diverges in the presence of MMS noise:

Proposition 3.4. *Under Assumptions 3.1 and 3.5, construct* $\hat{\sigma}_t(\boldsymbol{\lambda})$ *from* $(P^{\epsilon}(i))_{i\in I_n^{(q)}}$ *on the interval* I_n *for some fixed* $q > 1$ *and* $\boldsymbol{\lambda} \in \boldsymbol{\Lambda}^{(q)}$*. As* $\delta_n \to 0$*, it holds that:*

$$\hat{\sigma}_t(\boldsymbol{\lambda}) = \Delta_n^{-1/2}\boldsymbol{\lambda}'\boldsymbol{\Theta}^{(q)}\boldsymbol{\epsilon}^{(q)} + O_p(1). \tag{3.23}$$

where $\boldsymbol{\epsilon}^{(q)} := (\eta^{\epsilon}, \xi_1^{\epsilon}, \xi_2^{\epsilon})'$*, in which* η^{ϵ}*,* ξ_1^{ϵ} *and* ξ_2^{ϵ} *are the MAED, the range and the absolute return of the process* $(\epsilon(i))_{i\in I_n^{(q)}}$*, respectively.*

Proposition 3.4 suggests that $\hat{\sigma}_t(\boldsymbol{\lambda}) = O_p(\delta_n^{-1/2})$, so the OMK estimator explodes towards infinity in the limit when the noise is present. Intuitively, the observed price process is dominated by the

noise process in the limit, and the MAED-candlestick vector of the observed price process $\boldsymbol{c}$ converges to that of the noise process, $\boldsymbol{\epsilon}^{(q)}$, which diverges when scaled by $\Delta_n^{-1/2}$. This result implies that one cannot choose Δ_n to be as small as possible in practice, as the noise-induced bias completely swamps the spot volatility estimates when Δ_n is small. Therefore, the OMK estimator is not robust to the presence of MMS noise, and empirically one needs to choose a large Δ_n to 'sparse sample' the price process and dampen the impact of noise. However, for the impact of noise to drop to a tolerable level, one might require an oversized Δ_n which heavily distorts the finite sample distribution of the OMK estimator and defeats the purpose of spot volatility estimation.

To mitigate the above problem, we develop a noise-robust OMK estimator using the pre-averaging method of Jacod et al. (2009, 2010) in the fixed-k framework. We consider an interval $I_n = [s, s+\Delta_n]$ with q_n+1 observations $(P^\epsilon(i))_{i\in I_n^{(q_n)}}$ such that $I_n^{(q_n)} := \{i_s, i_s+1, \ldots, i_s+q_n\}$. Similar to the discussion in the previous section we shall assume that the first and the last observations are at the edges of I_n. We set $q_n = O(\delta_n^{-\theta})$ for some $\theta \in (0,1)$. Under Assumption 3.4 and from our discussion in the previous section, we have $\Delta_n = O_p(\delta_n^{1-\theta}) = o_p(1)$, so I_n is a shrinking interval in the limit as desired, but it contains an increasing (instead of fixed) number of observations in the limit.

On I_n, we pick a pre-averaging window $k_n = O(\delta_n^{\varepsilon-\theta})$ strictly bounded above by q_n for some $\varepsilon \in (0,\theta)$ so that $k_n \to \infty$ while $k_n/q_n \to 0$. Construct the following pre-averaged price process:

$$\bar{P}(i) = \frac{1}{k_n}\sum_{j=0}^{k_n-1} P^\epsilon(i-j), \quad i \in I_n^{(q_n)}, \tag{3.24}$$

where we use at most k_n observations prior to time s to compute the first k_n averaged prices. Using $(\bar{P}(i))_{i\in I_n^{(q_n)}}$ as the price observations, we compute the pre-averaged MAED-candlestick vector $\bar{\boldsymbol{c}} := (\bar{m}, \bar{w}, |\bar{r}|)'$, where the bar notation highlights the pre-averaging procedure. The pre-averaged OMK estimator is defined as:

$$\hat{\bar{\sigma}}_t(\boldsymbol{\lambda}) := \Delta_n^{-1/2}\boldsymbol{\lambda}'\boldsymbol{\Theta}\bar{\boldsymbol{c}}, \tag{3.25}$$

and note that $\boldsymbol{\Theta}$ is the same as in Theorem 3.1. We deduce the following asymptotic result:

Proposition 3.5. *Under Assumption 3.5, choose $\theta \in (0,1)$ and $\varepsilon \in (0,\theta)$ that satisfy $(\theta-\varepsilon)/(1-\varepsilon) > 1/(2-2d)$. As $\delta_n \to 0$, it holds for all $t \in I_n$ and $\boldsymbol{\lambda} \in \boldsymbol{\Lambda}$ that:*

$$\frac{\hat{\bar{\sigma}}_t(\boldsymbol{\lambda})}{\sigma_t} = \boldsymbol{\lambda}'\boldsymbol{\Theta}\boldsymbol{z} + o_p(1), \tag{3.26}$$

where $\boldsymbol{z}$ is defined as in Theorem 3.1.

Remark 3.7. Note that the asymptotic distribution of $\hat{\bar{\sigma}}_t(\boldsymbol{\lambda})$ is the same as that of Theorem 3.1 in the continuous case, which may seem surprising at first. Intuitively, the impact of noise diminishes as the pre-averaging window $k_n \to \infty$. However, the condition $k_n = o(q_n)$ ensures that the pre-averaged price process is asymptotically equivalent to the efficient price, as the pre-averaging window

is asymptotically negligible compared to the length of the interval. In the absence of MMS noise, one can set $k_n = 1$, and $\hat{\bar{\sigma}}_t(\boldsymbol{\lambda})$ simply reduces to $\hat{\sigma}_t^{(q)}(\boldsymbol{\lambda})$ with $q = \infty$. This asymptotic setting can thus be considered as an 'in-fill' asymptotic limit on the fixed-k interval I_n. As pre-averaging does not alter the asymptotic orders of jumps or drift bursts, one can show that $\hat{\bar{\sigma}}_{t,MAED}$ is jump-robust in view of Theorem 3.2, and the pre-averaged S-test, denoted as $\bar{S}_t := \ln \hat{\bar{\sigma}}_{t,OK} - \ln \hat{\bar{\sigma}}_{t,MAED}$, has the same asymptotic distribution as S_t in the presence of jumps or drift-bursts as discussed in Corollary 3.1 and Proposition 3.2.

Remark 3.8. It is worth noting that we can consider θ, ε and d as nuisance parameters as they do not play a role in the asymptotic distribution of $\hat{\bar{\sigma}}_t(\boldsymbol{\lambda})$. The condition $(\theta - \varepsilon)/(1 - \varepsilon) > 1/(2 - 2d)$ ensures that the impact of the MMS noise is asymptotically negligible. For example, with short memory noise so that $d = 0$, we need $(\theta - \varepsilon)/(1 - \varepsilon) > 1/2$, or $2\theta - \varepsilon > 1$. There are clearly infinitely many possible choices of θ and ε, such as $\theta = 3/4$ and $\varepsilon = 1/4$. In general, we need larger θ and ε to account for a larger d, but choices of θ and ε for any $d < 1/2$ are always available, which allows Eq. (3.26) to hold in the presence of long memory noises. For $d \geq 1/2$, no choices of θ and ε could satisfy this condition, and Eq. (3.26) no longer holds true due to the excess persistence in the noise process.

Proposition 3.5 shows that $\hat{\bar{\sigma}}_t(\boldsymbol{\lambda})$ is asymptotically robust to a very flexible specification of the MMS noise. In practice, $\hat{\bar{\sigma}}_t(\boldsymbol{\lambda})$ is very simple to construct as one only needs to choose k_n and q_n and construct the continuous OMK estimator on the pre-averaged prices. Although in theory k_n and q_n should be chosen based on the unknown parameters δ_n, θ, ε and d, they only control for the rate of explosion for k_n and q_n which is required for the asymptotic result, but they are irrelevant in finite sample as we only pick finite k_n and q_n to construct $\hat{\bar{\sigma}}_t(\boldsymbol{\lambda})$. As a result, $\hat{\bar{\sigma}}_t(\boldsymbol{\lambda})$ is still biased in finite sample due to MMS noise and discretely observed prices.

To provide some guidance on the choices of k_n and q_n, we note that one should always choose k_n large enough to fully dampen the impact of MMS noise. Therefore, for the purpose of noise reduction it is tempting to choose k_n as large as possible. However, we stress that one needs to pick q_n much larger relative to k_n (in theory $q_n/k_n \to \infty$) for Eq. (3.26) to hold approximately true. A finite q_n/k_n introduces a negative finite sample bias (and hence a distortion of the finite sample distribution) to $\hat{\bar{\sigma}}_t(\boldsymbol{\lambda})$ due to the pre-averaging of price increments. Intuitively, in finite sample, pre-averaging the price process also smooths the efficient price path, which effectively reduces the corresponding MAED and range statistics and generates a negative bias. As spot volatility estimators are typically constructed in intervals of pre-determined lengths (e.g., a 10-minute window used in Bollerslev et al. (2021); Li et al. (2022)), we recommend to choose the smallest k_n which shrinks the impact of the MMS noise to a negligible level. Therefore, if one believes that the noise is already negligible without pre-averaging (which may be the case when one chooses a large Δ_n), then one can simply pick $k_n = 1$.

We conclude this section by proposing a plausible finite sample correction to the bias due to pre-averaging. Via simulation, we discover the following approximated asymptotic distribution of $\overline{\boldsymbol{c}}$:

$$\Delta_n^{-1/2}\overline{\boldsymbol{c}} \overset{d}{\approx} \sigma_t \boldsymbol{z}^{(q)}, \tag{3.27}$$

when we fix $q_n/k_n = q$ for any $q \geq 10$ and let $q_n \to \infty$. Based on this observation, we propose the following 'discrete' pre-averaged OMK estimator in practice:

$$\hat{\bar{\sigma}}_t^{(\tilde{q})}(\boldsymbol{\lambda}) := \Delta_n^{-1/2}\boldsymbol{\lambda}'\boldsymbol{\Theta}^{(\tilde{q})}\overline{\boldsymbol{c}}, \tag{3.28}$$

where $\tilde{q} := q_n/k_n$, and $\boldsymbol{\Theta}^{(\tilde{q})}$ is defined as in Eq. (3.17) but with non-integer numbers of observations, whose values can be obtained using our polynomial approximations[12] in Table C.1. For the estimators $\hat{\bar{\sigma}}_{t,OMK}^{(\tilde{q})}$ and $\hat{\bar{\sigma}}_{t,OK}^{(\tilde{q})}$, the corresponding optimal weight vectors are computed based on Eqs. (3.8) and (3.9) with the variance-covariance matrix $\boldsymbol{\Sigma}^{(\tilde{q})}$ whose elements can also be obtained from Table C.1. For statistical inference of the estimators and the modified S-test denoted as $\bar{S}_t^{(\tilde{q})} := \ln \hat{\bar{\sigma}}_{t,OK}^{(\tilde{q})} - \ln \hat{\bar{\sigma}}_{t,MAED}^{(\tilde{q})}$, we obtain the related critical values for the confidence intervals and the S-test based on the polynomial approximations in Table C.2. Note that $\tilde{q}$ plays a similar role as q in the discrete OMK estimator, so a larger $\tilde{q}$ is preferred as it reduces the asymptotic variance factor of $\hat{\bar{\sigma}}_t^{(\tilde{q})}(\boldsymbol{\lambda})$, which indicates that the choice of k_n is a classic bias-variance trade-off for a fixed q_n.

Eq. (3.27) is a surprising observation, as it suggests that the impact of pre-averaging on the MAED and range statistics is similar to the impact of discrete observations (which is discussed in Table 3.3). The polynomial approximations are a natural interpolation method to obtain the moments and critical values for non-integer $\tilde{q}$. We are unable to justify these results theoretically, as it requires a careful study of the probabilistic nature of Brownian moving averages, which is beyond the scope of this paper. Nevertheless, our simulation results in Fig. 4.2 show convincingly that $\hat{\bar{\sigma}}_t^{(\tilde{q})}(\boldsymbol{\lambda})$ has much better finite sampler performance than $\hat{\bar{\sigma}}_t(\boldsymbol{\lambda})$ with less bias and better coverage rates, which provides a credible finite sample correction to the pre-averaged OMK estimator and is thus recommended in practice.

4 Simulation Study

We conduct a comprehensive simulation study to examine the properties of the MAED-related statistics proposed in this paper. The log-price process P is generated according to a two-factor stochastic volatility model, following Li et al. (2022):

$$\begin{aligned}
dP_t &= \sigma_t dW_t + J_t, \quad \sigma_t^2 = V_{1,t} + V_{2,t} \\
dV_{1,t} &= 0.0128(0.4068 - V_{1,t})dt + 0.0954\sqrt{V_{1,t}}(\rho dW_t + \sqrt{1-\rho^2}dB_{1,t}) \\
dV_{2,t} &= 0.6930(0.4068 - V_{2,t})dt + 0.7023\sqrt{V_{1,t}}(\rho dW_t + \sqrt{1-\rho^2}dB_{2,t}),
\end{aligned} \tag{4.1}$$

[12] Note that if $\tilde{q} \in [1, 10]$, we compute the associate moments using cubic interpolation from the simulated moments with integer q in Table C.1.

where W, B_1 and B_2 are independent standard Brownian motions, and we set $\rho = -0.7$ to capture the leverage effect between financial price and volatility shocks. We simulate 10000 paths of P based on Eq. (4.1) on the unit interval $[0, 1]$ normalized as a trading day with 390 minutes or 23400 seconds. The Euler discretization step is set to be 2.34×10^{-5}, or 10 steps per second. We initialize $P_0 = \ln 20$ so that the asset is traded at roughly \$20 on the trading day, and we fix $V_{1,0} = V_{2,0} = 0.5$. To account for discretely observed prices with different number of observations, we generate equidistant observation times $\tau_{n,i} = i/N_n$ with $N_n = 46800$, which corresponds to 2 observations per second. The efficient price observed at a discrete grid is therefore $\{P(i)\}_{i=0:N_n}$.

For the specification of the measurement error, we follow Christensen et al. (2022) and consider the following model for the observed price:

$$P^{\epsilon}(i) = P(i) + \epsilon(i), \quad \epsilon(i) \sim \mathcal{N}(0, \omega_i^2), \tag{4.2}$$

where $\omega_i = \varphi \sigma_{\tau_{n,i}} / \sqrt{N_n}$. The parameter φ captures the overall magnitude of the measurement error. Following Christensen et al. (2022), we set $\varphi = 2$, which represents a moderate contamination level (Christensen et al., 2014).

We first examine the performance of the OMK, OK and MAED estimators for the end-of-day volatility σ_1 in the absence of jumps and drift bursts. We compute these estimators based on sub-intervals of the form $I_n = [1 - \Delta_n, 1]$ with $390 \cdot \Delta_n \in \{1, 5, 10, 15, \cdots, 120\}$, so the interval length ranges from one minute to two hours with a step of five minutes. To demonstrate the importance of correcting for discrete price observations, we consider both the discrete and the continuous versions of the three estimators. Taking the OMK estimator as an example, we consider both $\hat{\sigma}^{(q)}_{1,OMK}$ with the correct choice of q and the corresponding optimal $\boldsymbol{\lambda}^*$ as shown in Fig. 3.2, as well as its continuous counterpart $\hat{\sigma}_{1,OMK}$ by treating $q = \infty$ with the optimal weight vector given in Eq. (3.8). The estimators $\hat{\sigma}^{(q)}_{1,OK}$, $\hat{\sigma}_{1,OK}$, $\hat{\sigma}^{(q)}_{1,MAED}$ and $\hat{\sigma}_{1,MAED}$ are defined in a similar fashion.

To evaluate the performance of the estimators, we compute the following evaluation measures for a general estimator $\hat{\sigma}_1 \in \{\hat{\sigma}^{(q)}_{1,OMK}, \hat{\sigma}_{1,OMK}, \hat{\sigma}^{(q)}_{1,OK}, \hat{\sigma}_{1,OK}, \hat{\sigma}^{(q)}_{1,MAED}, \hat{\sigma}_{1,MAED}\}$:

$$\text{RBias} := \text{E}[\hat{\sigma}_1/\sigma_1 - 1], \quad \text{MSRE} := \text{E}[(\hat{\sigma}_1/\sigma_1 - 1)^2], \quad \text{CR90\%} := \text{Prob}[\sigma_1 \in CI_{90\%}(\hat{\sigma}_1)], \tag{4.3}$$

where RBias, MSRE and CR90% stand for the relative bias, the mean squared relative error and the 90% coverage rate of $CI_{90\%}(\hat{\sigma}_1)$, the 90% confidence interval[13] of $\hat{\sigma}_1$, respectively.

The simulation results for the estimators of σ_1 are presented in Fig. 4.1. Several interesting findings can be concluded from this figure. First, the results constructed from the efficient price are largely in line with our results in Table 3.3 by comparing RBias, MSRE and CR090% with their asymptotic versions defined in Eq. (3.20). In general, as $\Delta_n \to 0$, the discrete estimators appear unbiased with

[13] See Remark 3.1 and Table C.2 for the construction of the 90% confidence intervals of the related estimators.

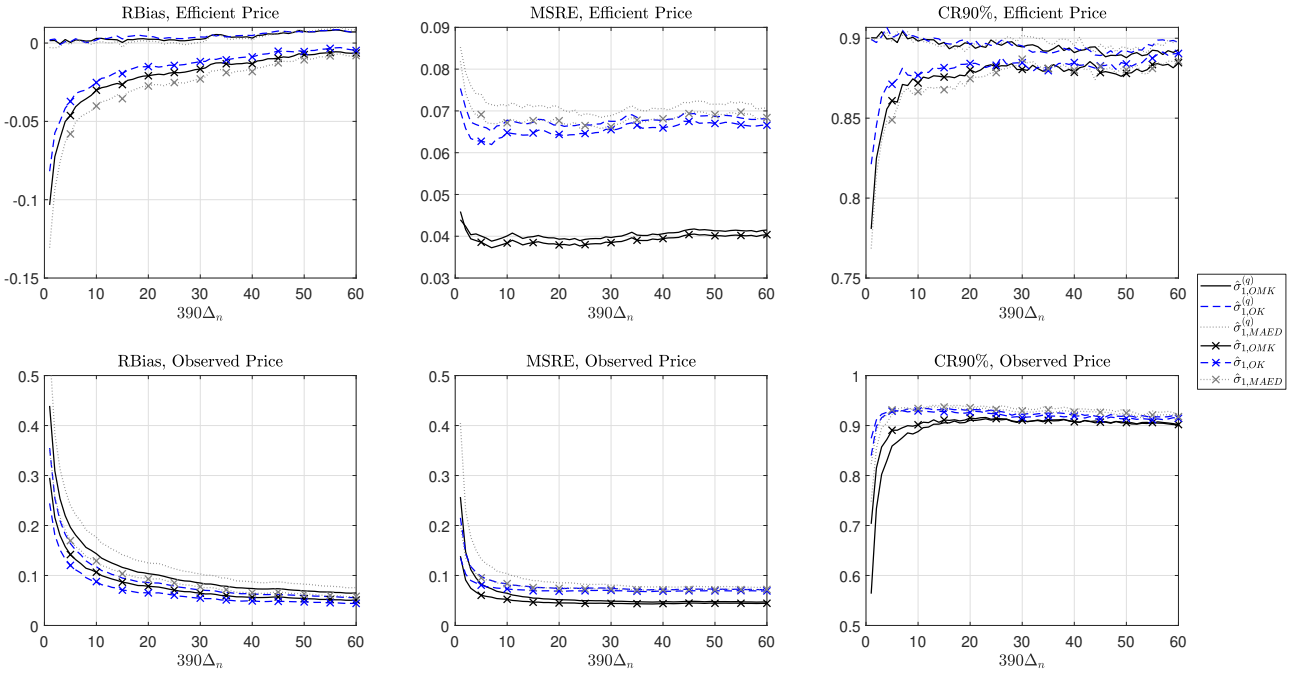


Figure 4.1: Simulated RBias, MSRE and CR90% of $\hat{\sigma}^{(q)}_{1,OMK}$, $\hat{\sigma}^{(q)}_{1,OK}$ and $\hat{\sigma}^{(q)}_{1,MAED}$ and their continuous versions. All estimators are constructed on the interval $[1-\Delta_n, 1]$ with the choice of Δ_n indicated by the x-axis. The results are computed based on 10000 simulated paths of the model in Eqs. (4.1) and (4.2) in the absence of jumps and drift bursts. The Efficient (resp. Observed) Price in the figure headings indicate that the results are computed from $\{P_i\}$ (resp. $\{P^\epsilon_i\}$).

correct coverage rates, while the continuous estimators are biased downwards with distorted coverage rates. The MSREs of the estimators show that the OMK estimator clearly dominates the OK and the MAED estimators which have similar performances. It also shows that the continuous versions of the estimators can indeed have smaller MSREs than their discrete counterparts, which is consistent with Table 3.3.

In the presence of measurement errors, all six estimators are heavily biased upwards, and the size of the bias explodes as $\Delta_n \to 0$. This leads to inflated MSREs and distorted coverage rates, as is shown in the second row of Fig. 4.1, which is consistent with Proposition 3.4. It is worth pointing out that the OK estimator is the least biased estimator among the three estimators, followed by the OMK and the MAED estimators. This is due to the fact that MAED is more sensitive to the presence of measurement errors, relative to range or return. Also, the continuous versions of the estimators appear less biased than the discrete versions as the positive noise-induced bias is partially offset by the negative bias due to discretely observed price process. The MSRE advantage of the OMK estimator is less pronounced in the presence of measurement errors, and the CR90% of the estimators are visibly distorted even for a relatively large Δ_n.

The results in Fig. 4.1 clearly demonstrate the impact of discrete price observations and measurement errors on the estimators considered. In the absence of measurement errors, the bias introduced by discrete observations can be corrected by using the discrete estimators, which is particularly rele-

vant when q is small. However, in the presence of measurement errors, all the estimators are biased upwards, which can have a non-trivial effect on the distribution of the estimators even for a relatively large choice of Δ_n. For example, with a window length of 10 minutes which is used in Bollerslev et al. (2021); Li et al. (2022), all six estimators are biased upwards by about 10% with various degrees of coverage rate distortions. These results suggest that the aforementioned estimators may be unreliable in the presence of measurement errors, especially when Δ_n is small.

We proceed to show that the impact of measurement errors can be greatly alleviated by the pre-averaging procedure, and the discrete pre-averaged OMK estimator can effectively eliminate the finite sample bias due to pre-averaging. To simplify the exposition, we consider the 10-minute interval $I_n = [1 - 10/390, 1]$ with $q_n = 1200$. On I_n, we construct the pre-averaged estimators $\hat{\bar{\sigma}}_{1,OMK}$, $\hat{\bar{\sigma}}_{1,OK}$ and $\hat{\bar{\sigma}}_{1,MAED}$, as well as their bias-corrected versions $\hat{\bar{\sigma}}^{(\tilde{q})}_{1,OMK}$, $\hat{\bar{\sigma}}^{(\tilde{q})}_{1,OK}$ and $\hat{\bar{\sigma}}^{(\tilde{q})}_{1,MAED}$ using $1 \le k_n \le 60$, so that $\tilde{q} \ge 20$. We present the results in Fig. 4.2.

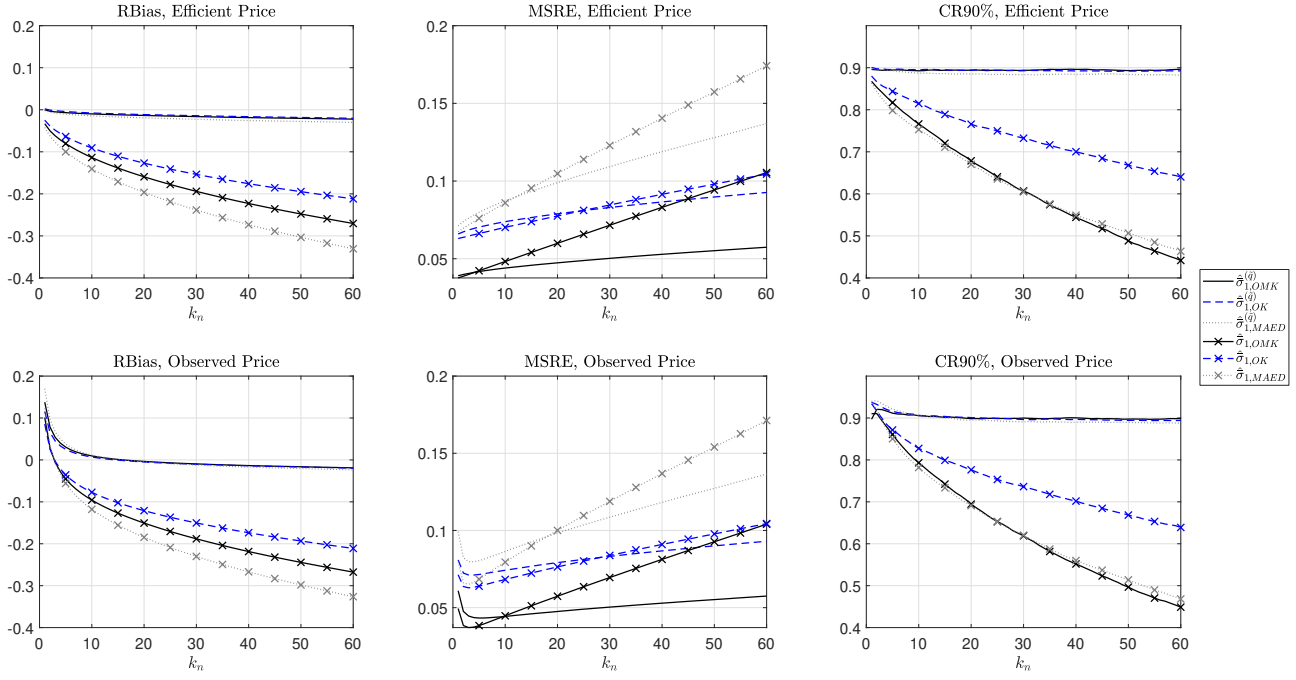


Figure 4.2: Simulated RBias, MSRE and CR90% of the pre-averaged spot volatility estimators. All estimators are constructed on $[1 - 10/390, 1]$ with $q_n = 1200$ and $1 \le k_n \le 60$. The results are computed based on 10000 simulated paths of the model in Eqs. (4.1) and (4.2) in the absence of jumps and drift bursts. The Efficient (resp. Observed) Price in the figure headings indicate that the results are computed from $\{P_i\}$ (resp. $\{P^\epsilon_i\}$).

The first row of Fig. 4.2 presents the performance of the pre-averaged estimators in the absence of measurement errors. We clearly see that $\hat{\bar{\sigma}}_{1,OMK}$, $\hat{\bar{\sigma}}_{1,OK}$ and $\hat{\bar{\sigma}}_{1,MAED}$ have a non-trivial negative bias[14] which enlarges as k_n grows, resulting in significantly inflated MSRE and highly distorted CR90%. On the contrary, the bias of the bias-corrected estimators are much less pronounced (less than 3% in

[14]The size of this bias is consistent with those reported in Table 3.3. For example, with $k_n = 60$ so that $\tilde{q} = 20$, the bias of the continuous OK estimator is about 20%, which is consistent with Fig. 4.2.

absolute value) for all values of k_n with much lower MSER and reliable coverage rates. The MSREs of the bias-corrected estimators increases almost linearly as a function of k_n, which is consistent with the numerical values in Table 3.2. For example, with $k_n = 30$ and $\tilde{q} = 40$, the simulated MSRE of $\hat{\bar{\sigma}}^{(\tilde{q})}_{1,OMK}$ is around 0.05, which coincides with the asymptotic variance factor of $\hat{\sigma}^{(q)}_{1,OMK}$ with $q = 40$ documented in Table 3.2. If the measurement error is absent, then it is optimal to choose $k_n = 1$ which minimizes the MSRE, and the bias-correction is also less important since q_n is large.

When measurement error is present, the second row of Fig. 4.2 shows that the bias induced by the measurement error diminishes as one increases k_n. With $k_n \geq 20$, all estimators have almost identical performance when compared to the case without measurement error. This is strong evidence supporting the effectiveness of the pre-averaging procedure in eliminating the impact of measurement error. More importantly, as one needs to choose an adequate k_n to eliminate the impact of measurement error, the bias-corrected pre-averaged estimators are clearly superior to their original counterparts for any $k_n \geq 20$.

We now examine the behaviour of the MAED estimator and the S-test statistic in the presence of jumps or drift bursts. We consider a practical setting where the estimators and the S-test are constructed in a rolling window fashion. Starting with the presence of a jump, we consider the following simple specification of J_t:

$$J_t = \mathbb{1}_{\{t \geq \tau_J\}} \Xi, \tag{4.4}$$

where a fixed jump occurs at time τ_J with a fixed size Ξ. We consider $\Xi \in \{0, 0.1, 0.5, 1\}$ to represent the no/small/medium/large jump case, and we fix $\tau_J = 330/390$ so the jump occurs at 15:30 on the simulated trading day. The jump process is then added to P_t to generated the jump-augmented prices. To conserve space, we only report the simulation results in the presence of measurement errors.

We construct $\hat{\bar{\sigma}}^{(\tilde{q})}_{t,OK}$, $\hat{\bar{\sigma}}^{(\tilde{q})}_{t,MAED}$, and $\bar{S}^{(\tilde{q})}_t$ on intervals of the form $[t-\Delta_n, t]$ for $390 \cdot t \in \{300, \ldots, 360\}$, which corresponds to the minute-by-minute grid points from 3:00-4:00 in a trading day. To conserve space, we fix $\Delta_n \in \{5/390, 10/390\}$ and $k_n = 20$ so that $\tilde{q} = 6 \cdot 390\Delta_n$ and the impact of measurement error is negligible (as shown in Fig. 4.2). For each t, we report the RBias of $\hat{\bar{\sigma}}^{(\tilde{q})}_{t,OK}$ and $\hat{\bar{\sigma}}^{(\tilde{q})}_{t,MAED}$, as well as the rejection rates of $\bar{S}^{(\tilde{q})}_t$ under conventional significance levels. The simulation results are presented in Fig. 4.3.

Focusing on the case $\Xi = 0$ in Fig. 4.3, we first see that both $\hat{\bar{\sigma}}^{(\tilde{q})}_{t,OK}$ and $\hat{\bar{\sigma}}^{(\tilde{q})}_{t,MAED}$ are almost unbiased, and the rejection rates of $\bar{S}^{(\tilde{q})}_t$ are close to the nominal level for all t and both choices of Δ_n. When $\Xi \neq 0$, both $\hat{\bar{\sigma}}^{(\tilde{q})}_{t,OK}$ and $\hat{\bar{\sigma}}^{(\tilde{q})}_{t,MAED}$ are biased whenever a jump is present in I_n, and the bias in general increases with Ξ. However, we clearly see that $\hat{\bar{\sigma}}^{(\tilde{q})}_{t,OK}$ is always more biased than $\hat{\bar{\sigma}}^{(\tilde{q})}_{t,MAED}$. More importantly, the relative bias of $\hat{\bar{\sigma}}^{(\tilde{q})}_{t,MAED}$ is bounded above (by about 0.4) regardless of the size of the jump, while the relative bias of $\hat{\bar{\sigma}}^{(\tilde{q})}_{t,OK}$ explodes as Ξ increases. This result is consistent with

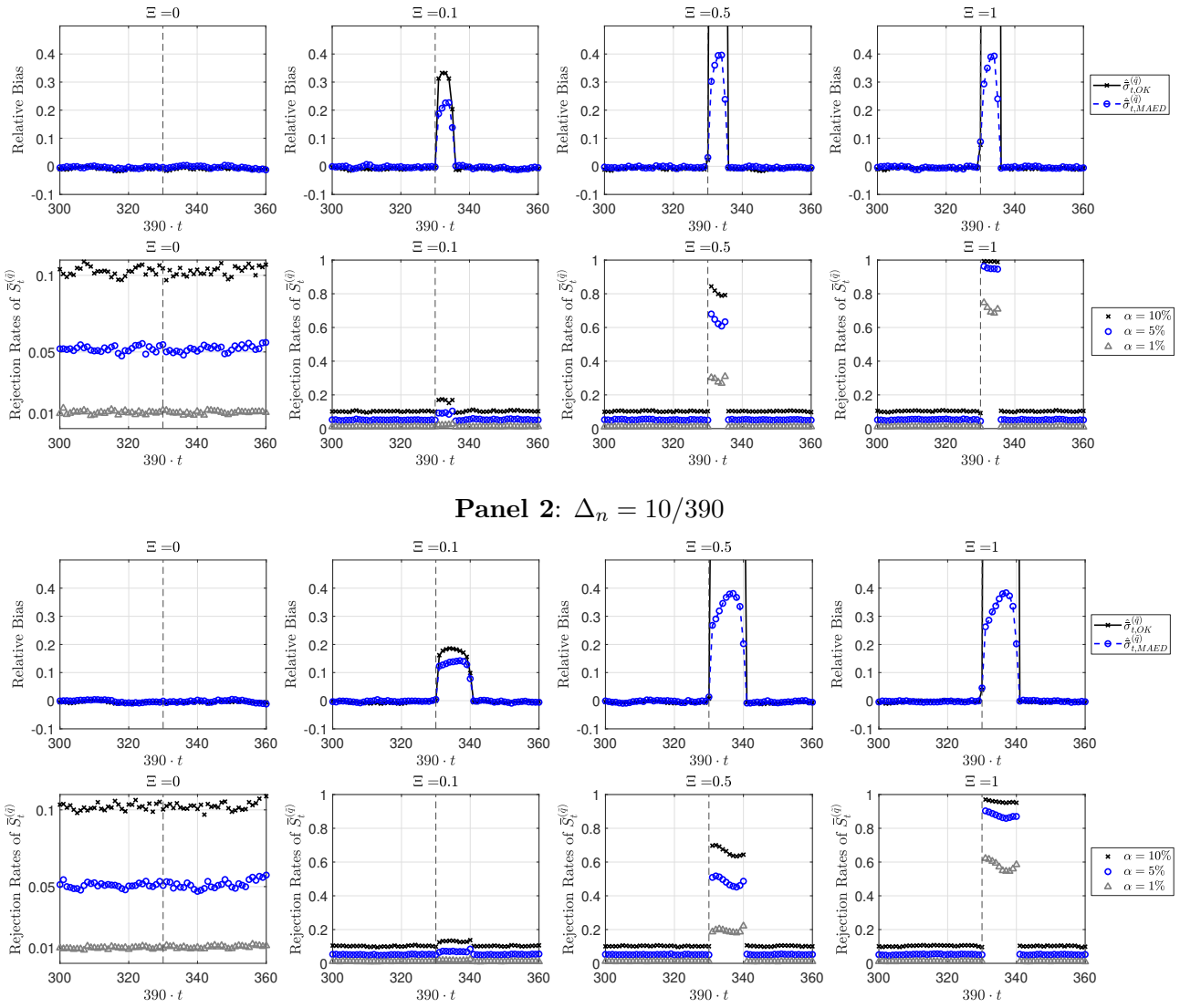


Figure 4.3: Simulated RBias of $\hat{\bar{\sigma}}_{t,OK}^{(\tilde{q})}$ and $\hat{\bar{\sigma}}_{t,MAED}^{(\tilde{q})}$ and the rejection rates of $\bar{S}_t^{(\tilde{q}))}$ in the presence of a jump. For each $t \in \{300/390, \ldots, 360/390\}$, we compute $\hat{\bar{\sigma}}_{t,OK}^{(\tilde{q})}$, $\hat{\bar{\sigma}}_{t,MAED}^{(\tilde{q})}$, and $\bar{S}_t^{(\tilde{q})}$ on the interval $[t - \Delta_n, t]$. The results are computed based on 10000 random draws of the price process according to Eq. (4.1) with $N_n = 23400$ and the jump specification in Eq. (4.4). The vertical dashed black line indicates the location of the jump, and Ξ shows the size of the jump.

Theorem 3.2 which shows that the MAED estimator has a bounded bias in the presence of a jump regardless of the size and location of the jump, while the bias of the OK estimator explodes as jump size increases. The discrepancy between $\hat{\tilde{\sigma}}^{(\tilde{q})}_{t,OK}$ and $\hat{\tilde{\sigma}}^{(\tilde{q})}_{t,MAED}$ generates power for the S-test, which is clearly higher as the size of the jump becomes larger. The performance of the S-test is largely similar for both choices of Δ_n, but the power of the test is in general better for smaller choices of Δ_n, as the (fixed) jump becomes larger relative to the (shrinking) MAED and candlestick statistics of the continuous price path.

Lastly, we examine the impact of a drift burst on the MAED and the OK estimators and the performance of the S-test. We follow the simulation design of Christensen et al. (2022) and consider the following alternative drift burst-augmented price process P^{db}:

$$
\begin{aligned}
dP^{db}_t &= -\frac{3\mathrm{sgn}(\tau_{db}-t)}{|\tau_{db}-t|^{\alpha}}\mathbb{1}_{\{t\in[\tau_-,\tau_+]\}}dt + \sigma^{db}_t dW_t, \\
\sigma^{db}_t &= \begin{cases} \sqrt{V_{1,t}+V_{2,t}} & t\in[0,\tau_-)\cup(\tau_+,1] \\ \sqrt{V_{1,t}+V_{2,t}}\frac{|\tau_{db}-\tau_-|^{\beta}}{|\tau_{db}-t|^{\beta}}dW_t, & t\in[\tau_-,\tau_{db}] \\ \sqrt{V_{1,t}+V_{2,t}}\frac{|\tau_{db}-\tau_+|^{\beta}}{|\tau_{db}-t|^{\beta}}dW_t, & t\in(\tau_{db},\tau_+] \end{cases},
\end{aligned}
\tag{4.5}
$$

where $\mathrm{sgn}(x)$ returns the sign of x, and jumps are assumed absent in the above model. The model assumes that a drift burst occurs in the interval $[\tau_-,\tau_+]$ with the drift burst time $\tau_{db} := (\tau_- + \tau_+)/2$. We set $[\tau_-,\tau_+] = [320/390, 340/390]$, so that the drift burst interval is 20 minutes, which occurs at around 15:30 in a trading day. The specification of σ^{db}_t ensures that it is continuous on $[0,1]$ with a possible singularity at τ_{db} if $\beta > 0$.

Following the suggestions in Christensen et al. (2022), we consider the following four cases of the drift burst parameters: (1) pure volatility burst, $(\alpha,\beta) = (0, 0.2)$; (2) pure drift burst, $(\alpha,\beta) = (0.55, 0)$; (3) drift-volatility burst without arbitrage, $(\alpha,\beta) = (0.55, 0.4)$; (4) drift-volatility burst with arbitrage, $(\alpha,\beta) = (0.75, 0.2)$. An example of the simulated price and spot volatility paths for each case is presented in Fig. 4.4. The figure illustrates the magnitude of the drift burst and volatility burst with various choices of parameters. In detail, larger α and β induce more pronounced drifts and volatility bursts. In particular, it is very difficult to distinguish drift bursts with different β, which is consistent with the observation in Christensen et al. (2022).

We construct $\hat{\tilde{\sigma}}^{(\tilde{q})}_{t,OK}$, $\hat{\tilde{\sigma}}^{(\tilde{q})}_{t,MAED}$, and $\bar{S}^{(\tilde{q})}_t$ using the same simulation setting as Fig. 4.3. As the spot volatility estimates explode during a drift burst or a volatility burst, we present the average volatility estimates alongside σ^{db}_t instead of the relative bias of the volatility estimates for a more transparent visualization. We present the simulation results in Fig. 4.5.

Several interesting findings can be concluded from Fig. 4.5. In the pure volatility burst case, both $\hat{\tilde{\sigma}}^{(\tilde{q})}_{t,OK}$ and $\hat{\tilde{\sigma}}^{(\tilde{q})}_{t,MAED}$ explode as $t \to \tau_{db}$. However, the rejection rates of $\bar{S}^{(\tilde{q})}_t$ are only mildly

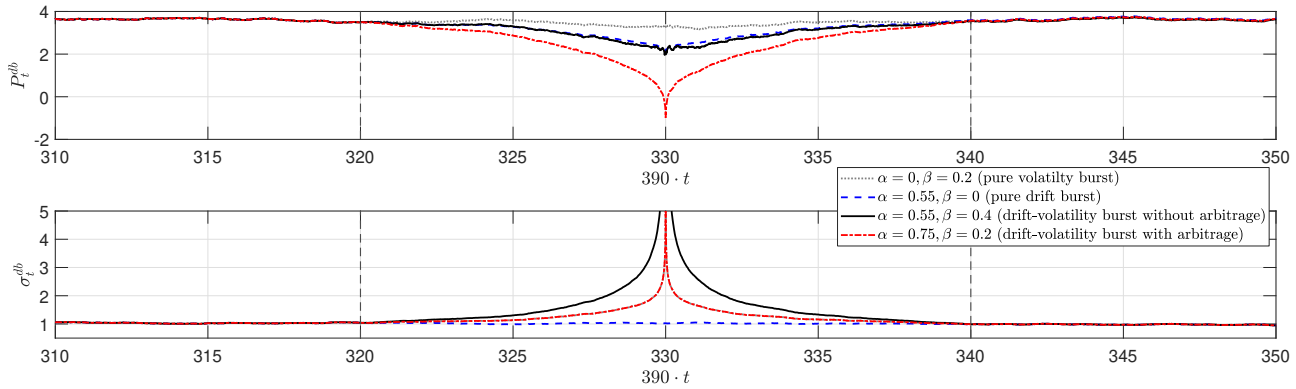


Figure 4.4: Simulated drift burst-augmented efficient price and spot volatility paths. For different choices of α and β, the paths are simulated according to Eq. (4.5) with the choices of α and β given in the legend. The two vertical lines shows the interval $[\tau_-, \tau_+]$.

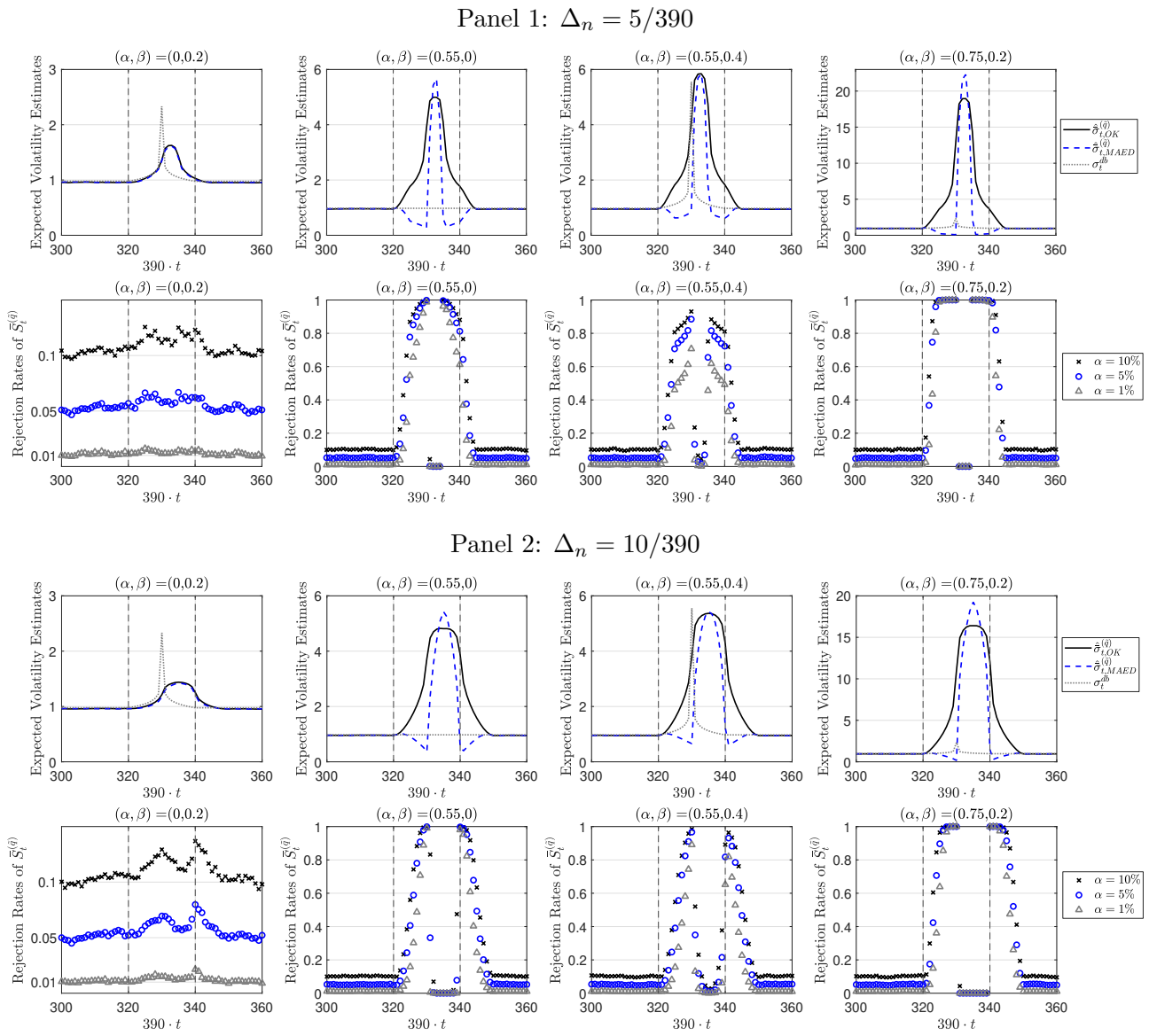


Figure 4.5: Simulated expectation of $\hat{\bar{\sigma}}_{t,OK}^{(\tilde{q})}$, $\hat{\bar{\sigma}}_{t,MAED}^{(\tilde{q})}$ and the rejection rates of $\bar{S}_t^{(\tilde{q})}$ in the presence of drift/volatility burst. For each $t \in \{300/390, \ldots, 360/390\}$, we compute $\hat{\bar{\sigma}}_{t,OK}^{(\tilde{q})}$, $\hat{\bar{\sigma}}_{t,MAED}^{(\tilde{q})}$, and $\bar{S}_t^{(\tilde{q})}$ on the interval $[t - \Delta_n, t]$. The results are computed based on 10000 random draws of the price process according to Eq. (4.5) with $N_n = 23400$ and the drift burst parameters specified in the titles of the figures. The two vertical lines shows the interval $[\tau_-, \tau_+]$.

distorted from the nominal rates as volatility explodes, suggesting that pure volatility bursts are unlikely to be detected by the S-test. When the drift burst is present, $\hat{\bar{\sigma}}^{(\tilde{q})}_{t,OK}$ always explodes as $t \to \tau_{db}$, while $\hat{\bar{\sigma}}^{(\tilde{q})}_{t,MAED}$ first decreases as we enter $[\tau_-, \tau_+]$, but may explode towards infinity together with $\hat{\bar{\sigma}}^{(\tilde{q})}_{t,OK}$ if $\tau_{db} \in I_n$. This is due to the drift burst reversal after the local price infimum at τ_{db}, which by construction induces an explosive inward price movement and drives $\hat{\sigma}^{(q)}_{t,MAED}$ upwards. Also, it is clear that all the spot volatility estimates are strikingly different from the true spot volatility, especially when a drift burst is present.

In the presence of drift bursts, the rejection rates of $\bar{S}^{(\tilde{q})}_t$ clearly demonstrate its usefulness in monitoring the birth and turning point of drift bursts for different choices of Δ_n, regardless of whether arbitrage opportunities are present. For the interval $[t-\Delta_n, t]$ and as t gradually increases pass τ_-, the unidirectional price movements drive $\hat{\bar{\sigma}}^{(\tilde{q})}_{t,OK}$ and $\hat{\bar{\sigma}}^{(\tilde{q})}_{t,MAED}$ toward opposite directions, which inflates $\bar{S}^{(\tilde{q})}_t$ and leads to rejections of the S-test. This result is consistent with our theoretical results in Proposition 3.2. Importantly, as the MAED estimator has the opposite sign of the bias in the presence of a jump or a drift burst (see Fig. 4.3), the behaviour of $\hat{\bar{\sigma}}^{(\tilde{q})}_{t,MAED}$ also allows us to identify whether a rejection of the S-test is likely due to a single jump or a drift burst.

5 Empirical Illustrations

We conduct a small empirical study to demonstrate the usefulness of the OMK estimator and the S-test in practice. We obtain the tick-by-tick transaction prices for Apple (ticker: AAPL) from the TAQ dataset in WRDS during the period 2016-2021. The transaction prices are timestamped in microseconds. We apply the standard data filters[15] in Barndorff-Nielsen et al. (2009); Holden and Jacobsen (2014) to remove possible outliers. We also ignore trading days with early market closure, namely the day before Independence Day and Christmas Day, and the day after Thanksgiving Day. The final sample comprises of 1499 trading days.

We split each trading day into 5-minute sampling windows rolling forward every minute, which gives 386 intervals in total. On each interval, we construct $\hat{\bar{\sigma}}^{(\tilde{q})}_{t,OK}$, $\hat{\bar{\sigma}}^{(\tilde{q})}_{t,OMK}$ and $\hat{\bar{\sigma}}^{(\tilde{q})}_{t,MAED}$, as well as the $\bar{S}^{(\tilde{q})}_t$ test statistic. To determine the choice of k_n required to dampen the impact of measurement error, we construct a volatility signature plot in the spirit of Andersen et al. (2000) by plotting the mean and median of the (pre-averaged) OMK estimators evaluated over all intervals against the choices of k_n, which is presented in Fig. 5.1. The volatility signature plots clearly suggest the presence of measurement error, as the average volatility estimates are higher for $k_n = 1$ and decay as k_n becomes larger. This pattern is in general consistent with Fig. 4.2 and suggests that a plain OMK estimator

[15] We use the SAS code from Holden and Jacobsen (2014) to extract transaction data with the prevailing NBBO quotes from the TAQ dataset. We then apply the filter rules P1-2 and T1-4 in Barndorff-Nielsen et al. (2009) to the transaction prices to remove potential outliers.

can be biased upwards due to the presence of measurement errors. As a conservative choice, we pick $k_n = 30$ to construct the pre-averaged estimators throughout the empirical analysis, which yields an average $\tilde{q}$ of approximately 130 across all intervals, so the expected variance reduction of the OMK estimator to the OK estimator is close to the theoretical limit of 40%.

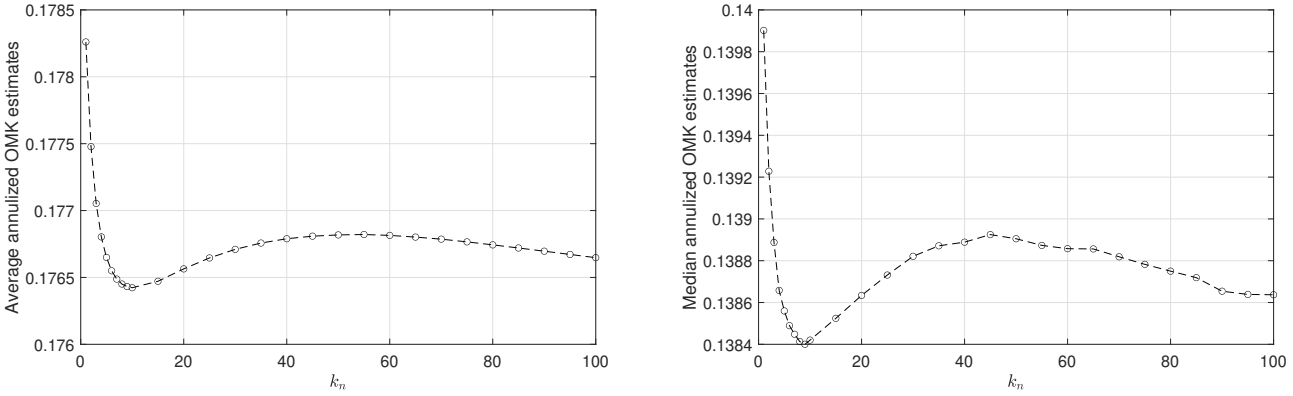


Figure 5.1: Volatility signature plots of OMK estimates against choices of k_n. The left (resp. right) figure plots the average (resp. median) of $\hat{\sigma}^{(\tilde{q})}_{t,OMK}$ evaluated over all 5-minute intervals in the sample against different choices of k_n.

We first show that the OMK estimator can be used to improve the precision of the OK estimator for spot volatility estimation. To this end, we split the trading days into two subsets: days with pre-scheduled FOMC announcements and those without announcements. We present the average spot volatility estimates based on the pre-averaged OMK and the OK estimators alongside with the averaged 90% HDIs in Fig. 5.2.

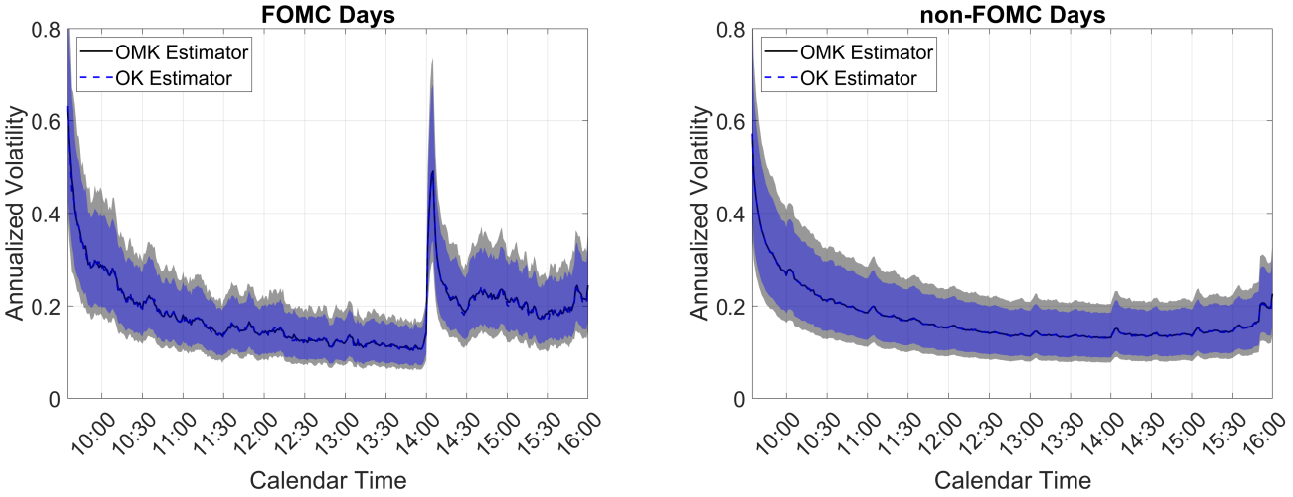


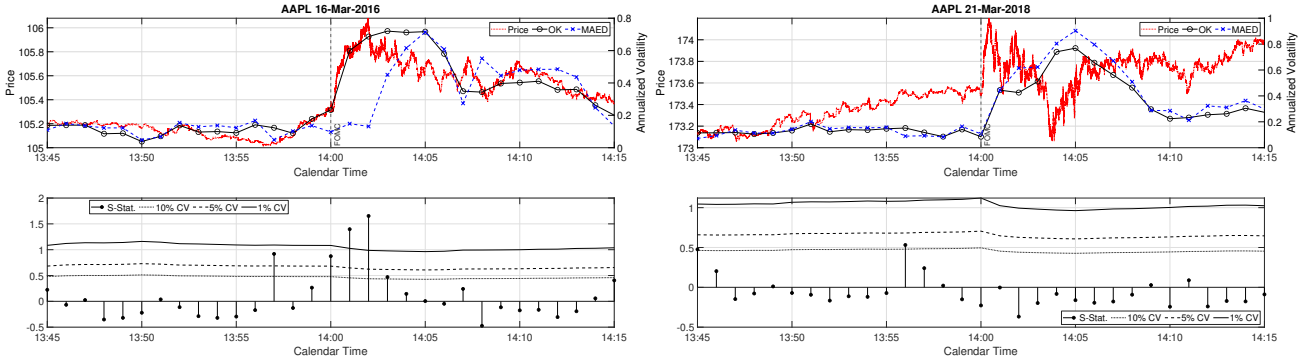
Figure 5.2: Averaged spot volatility estimates for AAPL based on the pre-averaged OMK and the OK estimator on FOMC and non-FOMC days during 2016-2021. The x-axis represents the right endpoints of a 5-minute interval, starting from 9:00-9:05, rolling forward every one minute. The blue and grey shaded areas are the average 90% HDIs of the OMK and OK estimators, respectively.

From Fig. 5.2, we clearly see a large volatility spike right after 14:00 on FOMC days, which is absent on non-FOMC days. This shows that FOMC meetings are associated with big volatility spikes, which is consistent with the findings in Bollerslev et al. (2021). More importantly, the OMK and

the OK estimators provide almost identical average spot volatility estimates in both subsets, but the average 90% HDIs of the OMK estimator are visibly narrower than those of the OK estimator. This clearly demonstrates the superiority of the OMK estimator to the OK estimator in generating more precise spot volatility measurements.

To illustrate the usefulness of the S-test, we select some representative cases and present them in Fig. 5.3. Panel 1 of the figure presents the 5-minute volatility estimates and the S-tests in a 30-minute window around two FOMC meetings at 14:00. The price paths of both sub-figures in panel 1 react immediately after the FOMC announcement time with a sharp increase in the OK volatility estimates. However, the S-test results suggest that the two FOMC meetings have fundamentally different implications to the price paths. In detail, the FOMC meeting on 16th March 2016 leads to a price increase with nearly monotonic price paths around 14:00. This generates large discrepancy between the OK and the MAED spot volatility estimates and leads to strong rejections of the S-test. On the contrary, the FOMC meeting on 21st March 2018 generates oscillatory price paths, which can be interpreted as a volatility burst that drives up both the OK and the MAED estimators. Consequently, the S-test is not rejected in this interval.

Panel 1: FOMC announcements with different reactions



Panel 2: Local price trends and gradual jumps

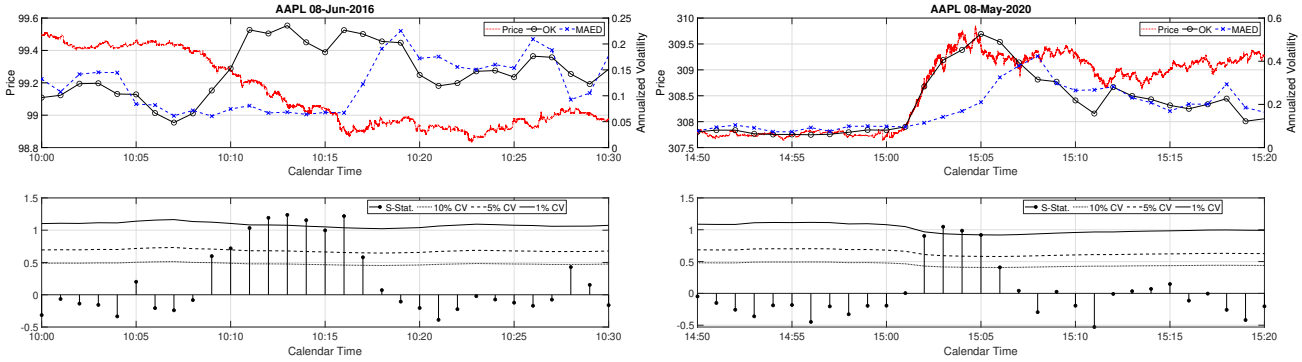


Figure 5.3: Price paths, volatility estimates, and the associated S-tests. In the top panel of the four figures, the observed price path is plotted against the left y-axis, while the 5-minute pre-averaged OK and MAED estimators are plotted against the right y-axis. The lower panel of the four figures presents the associated 5-minute S-test statistics and the critical values.

Panel 2 of Fig. 5.3 presents two different scenarios where the S-test is strongly rejected in the absence of macroeconomic news announcements. The left sub-figure of Panel 2 depicts a gradual and roughly monotonic price decrease from 10:00-10:15, whereas the right sub-figure shows a rapid price increase, or a gradual jump, during 15:00-15:05. In both sub-figures, the local trends are immediately reflected in the OK volatility estimates but ignored by the MAED volatility estimates, generating a strong sequence of rejections of the S-test.

Fig. 5.3 provides strong empirical evidence supporting the effectiveness of the S-test in detecting local price trends. It is worth noting that we are unable to find individual jumps in the tick-by-tick prices of AAPL which leads to a rejection of the S-test. Instead, the rejections are typically caused by 'gradual jumps' similar to what we have discussed in Fig. 5.3. This finding is consistent with the finding in Christensen et al. (2014) that price jumps are in fact very rare events, and the empirically detected price jumps are likely caused by the gradual jumps in the tick-by-tick data.

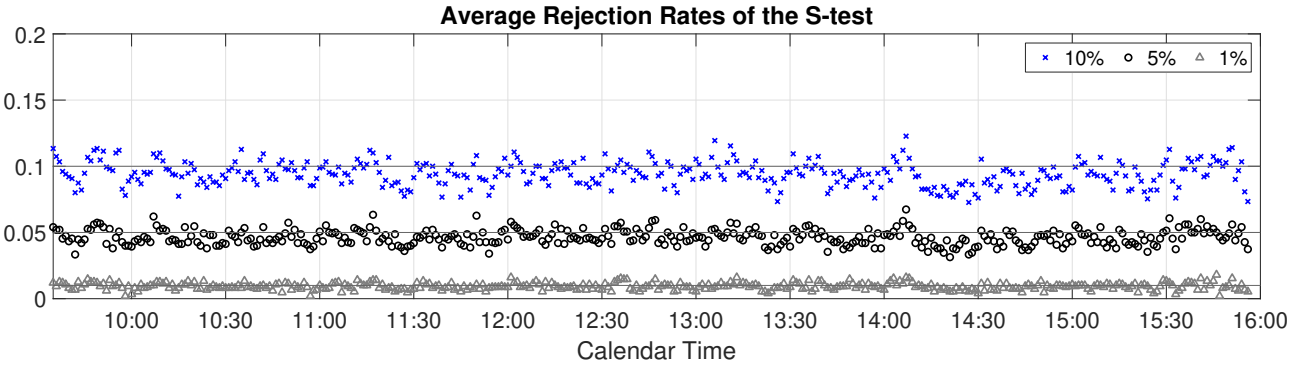


Figure 5.4: Average rejection rates of the S-test. The x-axis represents the right endpoints of a 5-minute interval, starting from 9:00-9:05, rolling forward every one minute. Each symbol represents the percentage of rejections in the same time of day averaged over the sampling period. The significance level of the rejection rates are given in the legend of the figure.

As the S-test can reliably detect drift bursts or flash crashes, it is meaningful to know how frequent the S-test is rejected in practice, which implies the frequency of drift bursts in practice. We plot the average rejection rates of the S-test against calendar time in Fig. 5.4. The figure shows that the rejection rates are fairly close to the nominal level through the trading day, with only minor excess rejections after 14:00, which is possibly due to the FOMC meetings. This result suggests that observed prices on average behave like a continuous martingale in a 5-minute interval, and genuine drift bursts or flash crashes are rare events in the tick-by-tick prices of AAPL that do not substantially increase the average rejection rates of the S-test.

6 Concluding Remarks

In this paper, we propose the MAED statistic, which is a novel and easy-to-construct statistic based on high-frequency asset prices. We show that MAED summarizes the inward movements from a price path, which contains unique information about the price variation that complements the candlestick

statistics. By combining the MAED statistic with the candlestick statistics, we propose the OMK estimator for spot volatility estimation which can in theory improve the asymptotic variance of the OK estimator of Li et al. (2022) by about 40% percent in the fixed-k framework of Bollerslev et al. (2021). We further show that the MAED estimator can be used to provide jump-robust spot volatility estimates. This result is used to construct the S-test, a pivotal test for the spot detection of explosive directional price movements. Pre-averaged versions of the new estimators and the S-test are constructed to mitigate the impact of measurement errors. Our simulation and empirical illustrations verify the superiority of the OMK estimator to the OK estimator and demonstrate the effectiveness of the S-test. Overall, these results highlight the empirical relevance of the MAED statistic. As the candlestick data for regular trading intervals is widely available from online trading platforms, we advocate data providers to also provide the MAED statistic, which allows regular investors to obtain more precise volatility measurements and detect local explosive price movements at a lower cost than purchasing tick-by-tick data.

We conclude this paper by discussing some unsolved questions for future research. First, although moments and critical values of the MAED statistic of a Brownian motion can be easily simulated, its analytical distribution and moments are still unknown and deserve individual investigation. Second, MAED is not the only functional that summarizes the inward price movements, which can also be (partially) summarized by, e.g., the maximal drawdown (Magdon-Ismail et al., 2004). It is therefore important to study whether the inward price movements can be more efficiently summarized by an alternative statistic. Third, as pointed out by Li et al. (2022); Bollerslev et al. (2022), one can also construct optimal estimators for σ_t^p for a general p based on the candlestick data. It is likely that the performance of these candlestick-based estimators can be substantially improved by also incorporating the MAED statistic. However, how to optimally combine the MAED statistic with the candlestick data is a non-trival task, which provides ample room for future research.

References

Aït-Sahalia, Y. and Jacod, J. (2014). *High-frequency financial econometrics*. Princeton University Press.

Aït-Sahalia, Y., Mykland, P. A., and Zhang, L. (2011). Ultra high frequency volatility estimation with dependent microstructure noise. *Journal of Econometrics*, 160(1):160–175.

Andersen, T. G., Bollerslev, T., Diebold, F. X., and Labys, P. (2000). Great Realizations. *Risk*, 13(3):105–108.

Andersen, T. G., Bollerslev, T., Diebold, F. X., and Labys, P. (2001). The distribution of realized exchange rate volatility. *Journal of the American Statistical Association*, 96(453):42–55.

Andersen, T. G., Bollerslev, T., Diebold, F. X., and Labys, P. (2003). Modeling and Forecasting Realized Volatility. *Econometrica*, 71:579–625.

Andersen, T. G., Bollerslev, T., and Lange, S. (1999). Forecasting financial market volatility: Sample frequency vis-à-vis forecast horizon. *Journal of Empirical Finance*, 6(5):457–477.

Andersen, T. G., Dobrev, D., and Schaumburg, E. (2008). Duration-Based Volatility Estimation. *Working Paper, Northwestern University*.

Andersen, T. G., Dobrev, D., and Schaumburg, E. (2012). Jump-robust volatility estimation using nearest neighbor truncation. *Journal of Econometrics*, 169(1):75–93.

Andersen, T. G., Li, Y., Todorov, V., and Zhou, B. (2021). Volatility measurement with pockets of extreme return persistence. *Journal of Econometrics*, In Press.

Asmussen, S., Glynn, P., and Pitman, J. (1995). Discretization error in simulation of one-dimensional reflecting Brownian motion. *The Annals of Applied Probability*, 5(4):875–896.

Baillie, R. T. (1996). Long memory processes and fractional integration in econometrics. *Journal of Econometrics*, 73(1):5–59.

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., Hansen, P. R., Lunde, A., and Shephard, N. (2009). Realized kernels in practice: Trades and quotes. *Econometrics Journal*, 12(3):C1—C32.

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

Barndorff-Nielsen, O. E. and Shephard, N. (2004). Power and Bipower Variation with Stochastic Volatility and Jumps. *Journal of Financial Econometrics*, 2:1–48.

Barndorff-Nielsen, O. E. and Shiryaev, A. (2015). *Change of time and change of measure*, volume 21 of *Advanced Series on Statistical Science & Applied Probability*. World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, second edition.

Beckers, S. (1983). Variances of Security Price Returns Based on High, Low, and Closing Prices. *The Journal of Business*, 56(1):97–112.

Bollerslev, T., Li, J., and Li, Q. (2022). Optimal Nonparametric Range-Based Volatility Estimation. *Working Paper*.

Bollerslev, T., Li, J., and Liao, Z. (2021). Fixed-k inference for volatility. *Quantitative Economics*, 12(4):1053–1084.

Christensen, K., Oomen, R., and Renò, R. (2022). The drift burst hypothesis. *Journal of Econometrics*, 227(2):461–497.

Christensen, K., Oomen, R. C. A., and Podolskij, M. (2014). Fact or friction: Jumps at ultra high frequency. *Journal of Financial Economics*, 114(3):576–599.

Christensen, K. and Podolskij, M. (2007). Realized range-based estimation of integrated variance. *Journal of Econometrics*, 141(2):323–349.

Feller, W. (1951). The asymptotic distribution of the range of sums of independent random variables. *Annals of Mathematical Statistics*, 22:427–432.

Foster, D. P. and Nelson, D. B. (1996). Continuous Record Asymptotics for Rolling Sample Variance Estimators. *Econometrica*, 64(1):139–174.

Garman, M. B. and Klass, M. J. (1980). On the Estimation of Security Price Volatilities from Historical Data. *The Journal of Business*, 53(1):67–78.

Granger, C. W. J. and Joyeux, R. (1980). An Introduction to Long-Memory Time Series Models and Fractional Differencing. *Journal of Time Series Analysis*, 1(1):15–29.

Hansen, P. R. and Lunde, A. (2006). Realized variance and market microstructure noise. *Journal of Business & Economic Statistics*, 24(2):127–218.

Holden, C. W. and Jacobsen, S. (2014). Liquidity measurement problems in fast, competitive markets: Expensive and cheap solutions. *Journal of Finance*, 69(4):1747–1785.

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

Jacod, J. (1997). On continuous conditional gaussian martingales and stable convergence in law. In Azéma, J., Yor, M., and Emery, M., editors, *Séminaire de Probabilités XXXI*, pages 232–246, Berlin, Heidelberg. Springer.

Jacod, J., Li, Y., Mykland, P. A., Podolskij, M., and Vetter, M. (2009). Microstructure noise in the continuous case: the pre-averaging approach. *Stochastic Processes and their Applications*, 119(7):2249–2276.

Jacod, J., Li, Y., and Zheng, X. (2017). Statical Property of Market Microstructure Noise. *Econometrica*, 8(4):1133–1174.

Jacod, J., Li, Y., and Zheng, X. (2019). Estimating the integrated volatility with tick observations. *Journal of Econometrics*, 208(1):80–100.

Jacod, J., Podolskij, M., and Vetter, M. (2010). Limit Theorems for moving averages of discretized processes plus noise. *Annals of Statistics*, 38(3):1478–1545.

Jacod, J. and Protter, P. E. (2012). *Discretization of Processes*. Springer-Verlag, New York, first ed. edition.

Kristensen, D. (2010). Nonparametric Filtering of the Realized Spot Volatility: A Kernel-Based Approach. *Econometric Theory*, 26(1):60–93.

Laurent, S. and Shi, S. (2020). Volatility estimation and jump detection for driftâ€“diffusion processes. *Journal of Econometrics*, 217(2):259–290.

Lee, S. S. and Mykland, P. A. (2007). Jumps in Financial Markets: A New Nonparametric Test and Jump Dynamics. *The Review of Financial Studies*, 21(6):2535–2563.

Lee, S. S. and Mykland, P. A. (2012). Jumps in equilibrium prices and market microstructure noise. *Journal of Econometrics*, 168(2):396–406.

Li, J., Wang, D., and Zhang, Q. (2022). Reading the Candlesticks: An OK Estimator for Volatility. *Review of Economics and Statistics*, forthcoming.

Li, Z. M., Laeven, R. J. A., and Vellekoop, M. H. (2020). Dependent microstructure noise and integrated volatility estimation from high-frequency data. *Journal of Econometrics*, 215(2):536–558.

Li, Z. M. and Linton, O. (2022). Robust Estimation of Integrated and Spot Volatility. *Working Paper*.

Magdon-Ismail, M., Atiya, A. F., Pratap, A., and Abu-Mostafa, Y. S. (2004). On the Maximum Drawdown of a Brownian Motion. *Journal of Applied Probability*, 41(1):147–161.

Mancini, C. (2009). Non-parametric threshold estimation for models with stochastic diffusion coefficient and jumps. *Scandinavian Journal of Statistics*, 36(2):270–296.

Meilijson, I. (2011). The Garman-Klass volatility estimator revisited. *Revstat Statistical Journal*, 9(3):199–212.

Parkinson, M. (1980). The Extreme Value Method for Estimating the Variance of the Rate of Return. *The Journal of Business*, 53:61–65.

Rogers, L. C. G. and Satchell, S. E. (1991). Estimating Variance From High, Low and Closing Prices. *Annals of Applied Probability*, 1(4):504–512.

Stein, C. (1956). Inadmissibility of the Usual Estimator for the Mean of a Multivariate Normal Distribution. *Berkeley Symposium on Mathematical Statistics and Probability*, I:197–206.

Varneskov, R. T. (2017). Estimating the Quadratic Variation Spectrum of Noisy Asset Prices Using Generalized Flat-Top Realized Kernels. *Econometric Theory*, 33(6):1457–1501.

Yang, D. and Zhang, Q. (2000). Drift-Independent Volatility Estimation Based on High, Low, Open, and Close Prices. *The Journal of Business*, 73(3):477–492.

Zhang, L. (2006). Efficient estimation of stochastic volatility using noisy observations: a multi-scale approach. *Bernoulli*, 12(6):1019–1043.

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.

Zu, Y. and Peter Boswijk, H. (2014). Estimating spot volatility with high-frequency financial data. *Journal of Econometrics*, 181(2):117–135.

Appendices

A Proofs

Proof of Proposition 3.1. We begin with the following set of inequalities for m, w and $|r|$ constructed from $(P_t)_{t\in I_n}$, which holds with probability 1 for the three statistics constructed from arbitrary processes by construction:

$$\begin{aligned} m \geq 0, \quad w \geq 0, \quad |r| \geq 0, \\ w - m \geq 0, \quad w - |r| \geq 0, \quad m + |r| - w \geq 0. \end{aligned} \tag{A.1}$$

The first five inequalities should be immediate. For the last one, note that:

$$m = \sup_{h\in I_n} \{w_h - |r_h|\} \geq w_{s+\Delta} - |r_{s+\Delta}| = w - |r|, \tag{A.2}$$

which gives the desired result. Therefore, for arbitrary non-negative real numbers $\{\alpha_k\}_{k=1:6}$, the following sum is guaranteed to be positive:

$$\begin{aligned} &\alpha_1(m + |r| - w) + \alpha_2(w - m) + \alpha_3(w - |r|) + \alpha_4 m + \alpha_5 w + \alpha_6|r| \\ =& \underbrace{(\alpha_1 - \alpha_2 + \alpha_4)}_{\alpha_m} m + \underbrace{(-\alpha_1 + \alpha_2 + \alpha_3 + \alpha_5)}_{\alpha_w} w + \underbrace{(\alpha_1 - \alpha_3 + \alpha_6)}_{\alpha_r} |r| \geq 0. \end{aligned} \tag{A.3}$$

The non-negativity condition of α_k is equivalent to the following set of inequalities, which can be derived by eliminating the negative terms in α_m, α_w and α_r:

$$\alpha_m + \alpha_w \geq 0, \quad \alpha_w + \alpha_r \geq 0, \quad \alpha_m + \alpha_w + \alpha_r \geq 0. \tag{A.4}$$

Therefore, we can conclude that, for any vector $\boldsymbol{\alpha} := (\alpha_m, \alpha_w, \alpha_r)'$, $\boldsymbol{\alpha}'\boldsymbol{c} \geq 0$ with probability 1 if the above set of inequality holds. In vector form, the above three inequalities can be rewritten as:

$$\begin{aligned} \alpha_m + \alpha_w = \boldsymbol{\alpha}'(\boldsymbol{e}_1 + \boldsymbol{e}_2) = \boldsymbol{\alpha}'(\boldsymbol{\iota} - \boldsymbol{e}_3) \geq 0 \Leftrightarrow \boldsymbol{\alpha}'\boldsymbol{\iota} \geq \boldsymbol{\alpha}'\boldsymbol{e}_3, \\ \alpha_w + \alpha_r = \boldsymbol{\alpha}'(\boldsymbol{e}_2 + \boldsymbol{e}_3) = \boldsymbol{\alpha}'(\boldsymbol{\iota} - \boldsymbol{e}_1) \geq 0 \Leftrightarrow \boldsymbol{\alpha}'\boldsymbol{\iota} \geq \boldsymbol{\alpha}'\boldsymbol{e}_1, \\ \alpha_m + \alpha_w + \alpha_r = \boldsymbol{\alpha}'\boldsymbol{\iota} \geq 0. \end{aligned} \tag{A.5}$$

We can thus summarize the above three inequalities succinctly by $\boldsymbol{\alpha}'\boldsymbol{\iota} \geq \max\{0, \boldsymbol{\alpha}'\boldsymbol{e}_1, \boldsymbol{\alpha}'\boldsymbol{e}_3\}$. Finally, the proposition is recovered by setting $\boldsymbol{\alpha} = \boldsymbol{\Theta}\boldsymbol{\lambda}$, and clearly adding the linear constraint $\boldsymbol{\lambda}'\boldsymbol{\iota} = 1$ does not violate the set of inequalities above. One can also easily show that there are infinitely many choices of $\boldsymbol{\lambda} \in \boldsymbol{\Lambda}$. This completes the proof. □

Proof of Theorem 3.1. The proof is in the spirit of the proofs of Theorem 1 in Bollerslev et al. (2021) and Li et al. (2022). Following their proofs, we shall assume that jumps are absent in the interval such that P is continuous on I_n, as the probability of observing a jump converges to zero as $\Delta_n \to 0$.

We also use the strengthened version of Assumption 3.1 by assuming that the conditions hold with $T_1 = \infty$ using a localization procedure (see the discussion in Li et al. (2022)).

We now turn to the proof of the theorem. We start with a fundamental inequality which is used repeatedly in the proofs:

Lemma A.1. *For arbitrary processes $X = (X_t)$ and $Y = (Y_t)$ defined on some interval $I = [s, h]$ satisfying $X_s = Y_s = 0$, it holds that:*

$$|\sup_{t\in[s,h]} X_t - \sup_{t\in[s,h]} Y_t| \leq \sup_{t\in[s,h]} |X_t - Y_t|, \quad |\inf_{t\in[s,h]} X_t - \inf_{t\in[s,h]} Y_t| \leq \sup_{t\in[s,h]} |X_t - Y_t|. \tag{A.6}$$

Proof. Since $X_s = Y_s = 0$, we must have $\sup_{t\in[s,h]} X_t \wedge \sup_{t\in[s,h]} Y_t \geq 0$. To prove the first inequality of the lemma, we first suppose that $\sup_{t\in[s,h]} X_t \geq \sup_{t\in[s,h]} Y_t$. Then:

$$|\sup_{t\in[s,h]} X_t - \sup_{t\in[s,h]} Y_t| = \sup_{t\in[s,h]} (X_t - \sup_{\tau\in[s,h]} Y_\tau) \leq \sup_{t\in[s,h]} (X_t - Y_t) \leq \sup_{t\in[s,h]} |X_t - Y_t|, \tag{A.7}$$

And the first inequality of the lemma follows by a symmetry argument. For the second inequality in the lemma, it suffices to notice that:

$$|\inf_{t\in[s,h]} X_t - \inf_{t\in[s,h]} Y_t| = |\sup_{t\in I_n}(-X_t) - \sup_{t\in[s,h]} (-Y_t)| = |\sup_{t\in[s,h]} X_t - \sup_{t\in[s,h]} Y_t| \leq \sup_{t\in[s,h]} |X_t - Y_t|, \tag{A.8}$$

where the last estimate follows from the first inequality of the lemma. This completes the proof. □

Let us define the processes $\xi_{1,h}$ and $\xi_{2,h}$ as the range and absolute return of the scaled Brownian motion $(\Delta^{-1/2} W_t)_{t\in[s,h]}$. We shall prove the following estimates:

$$\sup_{h\in I_n} |w_h - \sqrt{\Delta_n}\sigma_s \xi_{1,h}| = o_p(\Delta_n^{1/2}), \quad \sup_{h\in I_n} ||r_h| - \sqrt{\Delta_n}\sigma_s \xi_{2,h}| = o_p(\Delta_n^{1/2}). \tag{A.9}$$

We first look at the second supremum above:

$$\begin{aligned}
\sup_{h\in I_n} \left||r_h| - \sqrt{\Delta_n}\sigma_s\xi_{2,h}\right| &= \sup_{h\in I_n} \left||\textstyle\int_s^h b_t dt + \int_s^h \sigma_t dW_t| - \sigma_s|W_h - W_s|\right| \\
&\leq \sup_{h\in I_n} \left|\left|\textstyle\int_s^h b_t dt\right| + \left|\textstyle\int_s^h \sigma_t dW_t\right| - \sigma_s|W_h - W_s|\right| \\
&\leq \sup_{h\in I_n} |\textstyle\int_s^h b_t dt| + \sup_{h\in I_n} \left||\textstyle\int_s^h \sigma_t dW_t| - \sigma_s|W_h - W_s|\right| \\
&\leq \textstyle\int_s^{s+\Delta_n} |b_t| dt + \sup_{h\in I_n} \left|\textstyle\int_s^h (\sigma_t - \sigma_s) dW_t\right| \\
&= O_p(\Delta_n) + O_p(\Delta_n^{1/2+\kappa}) = o_p(\Delta_n^{1/2}),
\end{aligned} \tag{A.10}$$

which follows by repeatedly applying the (reverse) triangle inequality, and the last estimate follows

from Eqs. (A.1) and (A.2) of Li et al. (2022). For the first supremum in Eq. (A.9), we note that:

$$\begin{aligned}
&|w_h - \sqrt{\Delta_n}\xi_{1,h}| \\
=&\Big| \sup_{t\in[s,h]}\{P_t - P_s\} - \inf_{t\in[s,h]}\{P_t - P_s\} - \sigma_s \sup_{t\in[s,h]}\{W_t - W_s\} + \sigma_s \inf_{t\in[s,h]}\{W_t - W_s\}\Big| \\
\leq&\Big| \sup_{t\in[s,h]}\{P_t - P_s\} - \sigma_s \sup_{t\in I_n(h)}\{W_t - W_s\}\Big| + \Big| \inf_{t\in[s,h]}\{P_t - P_s\} - \sigma_s \inf_{t\in[s,h]}\{W_t - W_s\}\Big| \\
\leq&2 \sup_{t\in[s,h]}\Big|(P_t - P_s) - \sigma_s(W_t - W_s)\Big|,
\end{aligned} \tag{A.11}$$

where the first inequality follows from the triangle inequality and the second inequality follows from Lemma A.1. Consequently:

$$\begin{aligned}
\sup_{h\in I_n}|w_h - \sqrt{\Delta_n}\xi_{1,h}| \leq&2 \sup_{h\in I_n}\sup_{t\in[s,h]}\Big|(P_t - P_s) - \sigma_s(W_t - W_s)\Big| \\
\leq&2 \sup_{h\in I_n}\Big|(P_h - P_s) - \sigma_s(W_h - W_s)\Big| \\
\leq&2 \sup_{h\in I_n}\Big|\textstyle\int_s^h b_\tau d\tau + \int_s^h(\sigma_\tau - \sigma_s)dW_\tau\Big| \\
\leq&2\textstyle\int_s^{s+\Delta_n}|b_t|dt + \sup_{h\in I_n}\Big|\int_s^h(\sigma_\tau - \sigma_s)dW_\tau\Big| = o_p(\Delta_n^{1/2}),
\end{aligned} \tag{A.12}$$

where the last estimate is identical to that in Eq. (A.10). As a direct consequence of Eq. (A.9) with the help of Lemma A.1, we obtain the following estimates for the MAED, range and return of $(P_t)_{t\in I_n}$:

$$\begin{aligned}
|m - \sqrt{\Delta_n}\sigma_s\eta| =& |\sup_{h\in I_n}\{w_h - |r_h|\} - \sqrt{\Delta_n}\sigma_s \sup_{h\in I_n}\{\xi_{1,h} - \xi_{2,h}\}| \\
\leq& \sup_{h\in I_n}|w_h - \sqrt{\Delta_n}\sigma_s\xi_{1,h} - |r_h| + \sqrt{\Delta_n}\sigma_s\xi_{2,h}| \\
\leq& \sup_{h\in I_n}|w_h - \sqrt{\Delta_n}\xi_{1,h}| + \sup_{h\in I_n}||r_h| - \sqrt{\Delta_n}\xi_{2,h}| = o_p(\Delta_n^{1/2}), \\
|w - \sqrt{\Delta_n}\sigma_s\xi_1| \leq& \sup_{h\in I_n}|w_h - \sqrt{\Delta_n}\xi_{1,h}| = o_p(\Delta_n^{1/2}), \\
||r| - \sqrt{\Delta_n}\sigma_s\xi_2| \leq& \sup_{h\in I_n}||r_h| - \sqrt{\Delta_n}\xi_{2,h}| = o_p(\Delta_n^{1/2}).
\end{aligned} \tag{A.13}$$

An alternative proof of the above result related to w and $|r|$ can be found in Li et al. (2022). Therefore, by the Slutsky theorem, we conclude the following coupling result for the MAED-candlestick vector:

$$\boldsymbol{c} = (m, w, |r|)' = \sqrt{\Delta_n}\sigma_s(\eta, \xi_1, \xi_2)' + o_p(\Delta_n^{1/2}) = \sqrt{\Delta_n}\sigma_s\boldsymbol{z} + o_p(\Delta_n^{1/2}). \tag{A.14}$$

where the $o_p(\Delta_n^{1/2})$ term is understood as a compatible vector with $o_p(1)$ elements. By a standard continuous mapping argument, we arrive at:

$$\hat{\sigma}_t(\boldsymbol{\lambda}) = \Delta_n^{-1/2}\boldsymbol{\lambda}'\boldsymbol{\Theta}\boldsymbol{c} = \sigma_s\boldsymbol{\lambda}'\boldsymbol{\Theta}\boldsymbol{z} + o_p(1). \tag{A.15}$$

Finally, it suffices to notice that, for any $t \in I_n$, $\sigma_s/\sigma_t \xrightarrow{p} 1$ by the right continuity and boundedness of σ_t. We can therefore divide both sides of the above convergence by σ_t and obtain:

$$\frac{\hat{\sigma}_t(\boldsymbol{\lambda})}{\sigma_t} = \boldsymbol{\lambda}'\boldsymbol{\Theta}\boldsymbol{z} + o_p(1), \tag{A.16}$$

which is the desired coupling result. This completes the proof. □

Proof of Theorem 3.2. We begin by introducing some notations. For a process A constructed from P, we use the notation A° to denote the corresponding process constructed from P° instead. We shall also set $h < \tau$ and $h^\circ \geq \tau$ to denote the time before and after the jump. It should be clear that $w_h - |r_h| \equiv w_h^\circ - |r_h^\circ|$.

To prove the result related to the OK estimator, one simply needs to realize that as $\Delta_n \to 0$, it holds that:

$$|w^\circ - w| = J + O_p(\Delta_n^{1/2}), \quad ||r^\circ| - |r|| = J + O_p(\Delta_n^{1/2}), \tag{A.17}$$

due to the scaling of the Brownian motion. Since $J = O_p(1)$, we conclude that:

$$\begin{aligned} |\hat{\sigma}_{t,OK}^\circ - \hat{\sigma}_{t,OK}| &= \Delta_n^{-1/2} |\lambda_w^o \nu_1^{-1}(w - w^\circ) + \lambda_r^o \psi_1^{-1}(|r| - |r^\circ|)| \\ &\leq \Delta_n^{-1/2}(|\lambda_w^o \nu_1^{-1}||w - w^\circ| + |\lambda_r^o \psi_1^{-1}|||r| - |r^\circ||) \\ &= \Delta_n^{-1/2}(|\lambda_w^o \nu_1^{-1}| + |\lambda_r^o \psi_1^{-1}|)J + O_p(1) = O_p(\Delta_n^{-1/2}). \end{aligned} \tag{A.18}$$

This implies the result related to the OK estimator in Eq. (3.10).

We proceed to prove the result related to the MAED estimator in Eq. (3.10). Consider two scenarios, $J > 0$ and $J < 0$. We use the superscripts $+$ and $-$ to distinguish the signs of the jump, e.g. $P_{h^\circ}^{\circ,+}$ represents the jump-augmented price process at time h° conditioning on that the jump is positive. We also define the supremum and infimum processes of P on $I_n(h)$ as:

$$u_h := \sup_{t \in I_n(h)} \{P_t - P_s\}, \quad l_h := \inf_{t \in I_n(h)} \{P_t - P_s\}, \tag{A.19}$$

and the jump-augmented versions $u_h^{\circ,\pm}$ and $l_h^{\circ,\pm}$ are well understood. Let us first assume $J > 0$. Since J is strictly bounded away from zero, in the limit with probability 1, we have $P_{h^\circ}^{\circ,+} = P_{h^\circ} + J > P_h$ as the jump dominates the price increments before and after the jump which are of order $O_p(\Delta_n^{1/2})$. Consequently, we see that in the limit with probability 1:

$$u_{h^\circ}^{\circ,+} = P_\tau - P_s + J + \sup_{t \in [\tau, h^\circ]} (P_t - P_\tau), \quad l_{h^\circ}^{\circ,+} = l_\tau, \quad |r_{h^\circ}^{\circ,+}| = P_{h^\circ} - P_s + J, \tag{A.20}$$

which implies that:

$$w_{h^\circ}^{\circ,+} - |r_{h^\circ}^{\circ,+}| = \sup_{t \in [\tau, h^\circ]} (P_t - P_\tau) - (P_{h^\circ} - P_\tau) - l_\tau. \tag{A.21}$$

Similarly, when $J < 0$, we have in the limit with probability 1:

$$u_{h^\circ}^{\circ,-} = u_\tau, \quad l_{h^\circ}^{\circ,-} = P_\tau - P_s + J + \inf_{t \in [\tau, h^\circ]} (P_t - P_\tau), \quad |r_{h^\circ}^{\circ,-}| = -(P_{h^\circ} - P_s + J), \tag{A.22}$$

which leads to:

$$w^{\circ,-}(h^\circ) - |r^{\circ,-}(h^\circ)| = u(\tau) - \inf_{t \in [\tau, h^\circ]} (P_t - P_\tau) + (P_{h^\circ} - P_\tau). \tag{A.23}$$

As we do not know the direction of the jump, we can bound $|m - m^\circ|$ by the maximum absolute difference in both cases:

$$\begin{aligned} &|m - m^\circ| \leq |m - m^{\circ,+}| \vee |m - m^{\circ,-}| \\ &\leq \sup_{h^\circ \geq \tau} \{|w_{h^\circ} - |r_{h^\circ}| - w_{h^\circ}^{\circ,+} + |r_{h^\circ}^{\circ,+}|| \vee |w_{h^\circ} - |r_{h^\circ}| - w_{h^\circ}^{\circ,-} + |r_{h^\circ}^{\circ,-}||\}. \end{aligned} \tag{A.24}$$

Note that by the triangle inequality:

$$\begin{aligned} &\sup_{h^\circ \geq \tau} \{|w_{h^\circ} - |r_{h^\circ}| - w_{h^\circ}^{\circ,+} + |r_{h^\circ}^{\circ,+}||\} \\ =&\sup_{h^\circ \geq \tau} \Big|u_{h^\circ} - \sup_{t \in [\tau, h^\circ]} (P_t - P_\tau) - l_{h^\circ} + l_\tau - |r_{h^\circ}| + (P_{h^\circ} - P_\tau)\Big| \\ \leq &\sup_{h^\circ \geq \tau} |u_{h^\circ} - |r_{h^\circ}|| + \sup_{h^\circ \geq \tau} |l_{h^\circ} - l_\tau| + \sup_{h^\circ \geq \tau} |\sup_{t \in [\tau, h^\circ]} (P_t - P_\tau) - (P_{h^\circ} - P_\tau)| \\ \leq &u_{\Delta_n} + \sup_{t \in [\tau, s+\Delta_n]} (P_t - P_\tau) - 2 \inf_{t \in [\tau, s+\Delta_n]} (P_t - P_\tau). \end{aligned} \tag{A.25}$$

Similarly:

$$\begin{aligned} &\sup_{h^\circ \geq \tau} \{|w_{h^\circ} - |r_{h^\circ}| - w_{h^\circ}^{\circ,-} + |r_{h^\circ}^{\circ,-}||\} \\ =&\sup_{h^\circ \geq \tau} \Big|u_{h^\circ} - u_\tau - l_{h^\circ} + \inf_{t \in [\tau, h^\circ]} (P_t - P_\tau) - |r_{h^\circ}| - (P_{h^\circ} - P_\tau)\Big| \\ \leq &\sup_{h^\circ \geq \tau} |u_{h^\circ} - u_\tau| + \sup_{h^\circ \geq \tau} |-|r_{h^\circ}| - l_{h^\circ}| + \sup_{h^\circ \geq \tau} |\inf_{t \in [\tau, h^\circ]} (P_t - P_\tau) - (P_{h^\circ} - P_\tau)| \\ \leq &2 \sup_{t \in [\tau, s+\Delta_n]} (P_t - P_\tau) - l_{\Delta_n} - \inf_{t \in [\tau, s+\Delta_n]} (P_t - P_\tau). \end{aligned} \tag{A.26}$$

Combining the results from two cases, we arrive at:

$$\begin{aligned} &\sup_{h^\circ \geq \tau} \{|w_{h^\circ} - |r_{h^\circ}| - w_{h^\circ}^{\circ,+} + |r_{h^\circ}^{\circ,+}|| \vee |w_{h^\circ} - |r_{h^\circ}| - w_{h^\circ}^{\circ,-} + |r_{h^\circ}^{\circ,-}||\} \\ \leq &\Big(u_{\Delta_n} + \sup_{t \in [\tau, s+\Delta_n]} (P_t - P_\tau) - 2 \inf_{t \in [\tau, s+\Delta_n]} (P_t - P_\tau)\Big) \vee \\ &\Big(2 \sup_{t \in [\tau, s+\Delta_n]} (P_t - P_\tau) - l_{\Delta_n} - \inf_{t \in [\tau, s+\Delta_n]} (P_t - P_\tau)\Big) \\ \leq &w + 2 \sup_{t \in [\tau, s+\Delta_n]} (P_t - P_\tau) - 2 \inf_{t \in [\tau, s+\Delta_n]} (P_t - P_\tau). \end{aligned} \tag{A.27}$$

Finally, accounting for the unknown location of the jump, we take supremum over τ to arrive at:

$$\sup_{\tau \in (s, s+\Delta_n)} |m - m^\circ| \leq w + 2 \sup_{t \in [s, s+\Delta_n]} (P_t - P_s) - 2 \inf_{t \in [s, s+\Delta_n]} (P_t - P_s) = 3w. \tag{A.28}$$

Since we have already shown in Theorem 3.1 that $\Delta_n^{-1/2} w = \sigma_s \xi_1 + o_p(1)$, the required result follows by a standard continuous mapping theorem. This completes the proof. □

Proof of Corollary 3.1. The corollary directly follows from Theorems 3.1 and 3.2 by the continuous mapping theorem. This completes the proof. □

Proof of Proposition 3.2. We start with the normalized price process $\breve{P}_t^{db} := P_t^{db}\Delta_n^{\beta-1/2}$, $\forall t \in I_n$, and without loss of generality, we shall take $b_s > 0$ as the negative case follows by a symmetry argument. For the case $0 < \alpha - \beta < 1/2$, we first prove the following estimate:

$$\sup_{t\in I_n} |\breve{P}_t^{db} - \breve{P}_s^{db} - \Delta_n^{-1/2}\sigma_s \textstyle\int_s^t f(u)dW_u| = o_p(1), \tag{A.29}$$

where $f(t) = (\frac{t-s}{\Delta_n})^{-\beta}$ for $t \in I_n$, and it is understood that $\breve{P}_s^{db} := P_s\Delta_n^{\beta-1/2}$. This is easily seen as follows:

$$\begin{aligned}&\sup_{t\in I_n} |\breve{P}_t^{db} - \breve{P}_s^{db} - \Delta_n^{-1/2}\sigma_s \textstyle\int_s^t f(u)dW_u| = b_s \sup_{t\in I_n} \Delta_n^{\beta-1/2} \textstyle\int_s^t (u-s)^{-\alpha}du \\ &= b_s\Delta_n^{\beta-1/2} \textstyle\int_s^{s+\Delta_n} (u-s)^{-\alpha}du = b_s O(\Delta_n^{1/2-\alpha+\beta}) = o_p(1).\end{aligned} \tag{A.30}$$

Eq. (A.30) directly implies the following results, which can be derived in a similar manner as the proof of Eq. (A.13):

$$\Delta_n^{-1/2}|\Delta_n^\beta m - \sigma_s\breve{m}| = o_p(1), \quad \Delta_n^{-1/2}|\Delta_n^\beta w - \sigma_s\breve{w}| = o_p(1), \quad \Delta_n^{-1/2}|\Delta_n^\beta|r| - \sigma_s|\breve{r}|| = o_p(1), \tag{A.31}$$

where $\breve{m}$, $\breve{w}$ and $\breve{r}$ are the MAED, range, and return of $(\int_s^t f(u)dW_u)_{t\in I_n}$, respectively. Note that $\Delta_n^{\beta-1/2}m$ is the MAED of $(\breve{P}_t^{db})_{t\in I_n}$ due to the scaling property of MAED, and likewise for $\Delta_n^{\beta-1/2}|r|$ and $\Delta_n^{\beta-1/2}w$. It now suffices to notice that:

$$(\textstyle\int_s^t f(u)dW_u)_{t\in I_n} = (\breve{W}_{\tau_t})_{t\in I_n} \overset{d}{=} \frac{1}{\sqrt{1-2\beta}}(B_t)_{t\in I_n}, \tag{A.32}$$

where $\tau_t = \int_s^t f(u)^2du$ and $\breve{W}$ is another standard Brownian motion independent of W and B. The first equality above is an application of the Dambis-Dubins-Schwarz theorem (see e.g., Barndorff-Nielsen and Shiryaev (2015)), which states that every continuous martingale is a time-changed Brownian motion, with the time change $t \mapsto \tau_t$ given by the integrated variance of the continuous martingale. The second equality follows by the scaling law of a Brownian motion and the fact that:

$$\textstyle\int_s^{s+\Delta_n} f(u)^2du = \frac{\Delta_n}{1-2\beta}, \tag{A.33}$$

so $1/\sqrt{1-2\beta}$ is the correct scaling factor for $(B_t)_{t\in I_n}$. Eq. (A.32) directly implies the following joint equality in distribution by the continuous mapping theorem:

$$(\breve{m}, \breve{w}, |\breve{r}|)' \overset{d}{=} \sqrt{\frac{\Delta_n}{1-2\beta}}\breve{\boldsymbol{z}}. \tag{A.34}$$

Finally, by the Slutsky theorem, we can combine the three estimates in Eq. (A.31) to arrive at:

$$\Delta_n^\beta \boldsymbol{c} = \sigma_s(\breve{m}, \breve{w}, |\breve{r}|)' + o_p(1) \overset{d}{=} \sigma_s\sqrt{\frac{\Delta_n}{1-2\beta}}\breve{\boldsymbol{z}}, \tag{A.35}$$

which implies Eq. (3.13) by a continuous mapping argument.

We now turn to the case $\alpha - \beta > 1/2$. In this case, we can use the following convenient result of Christensen et al. (2022):

$$|\breve{P}_t^{db} - \breve{P}_s^{db}| = O_p(\Delta_n^{1/2-\alpha+\beta}) + O_p(1), \tag{A.36}$$

where $O_p(\Delta_n^{1/2-\alpha+\beta})$ is due to a drift that explodes in the limit, and $O_p(1)$ is due to a diffusion which is the same as the previous case. This directly implies that:

$$\Delta_n^{\beta-1/2} w = \Delta_n^{\beta-1/2}|r| = O_p(\Delta_n^{1/2-\alpha+\beta}), \tag{A.37}$$

as w and $|r|$ are linear in P_t^{db}, which implies the result related to the OK estimator in Eq. (3.14). Now to prove the result related to m, we shall establish the following result, which states that the probability for $\tilde{P}_t^{db}$ (and hence P_t^{db}) to be monotonically non-decreasing on I_n approaches 1 in the limit:

$$\lim_{\Delta_n \to 0} \Pr(\breve{P}_{t_1}^{db} - \breve{P}_{t_2}^{db} \leq 0) = 1, \forall t_1 \leq t_2 \in I_n. \tag{A.38}$$

To prove this, we examine the event $\breve{P}_{t_1}^{db} - \breve{P}_{t_2}^{db} \leq 0$ in detail:

$$\begin{aligned} \breve{P}_{t_1}^{db} - \breve{P}_{t_2}^{db} &= -b_s \Delta_n^{\beta-1/2} \textstyle\int_{t_1}^{t_2} (u-s)^{-\alpha} du - \Delta_n^{\beta-1/2} \sigma_s \int_{t_1}^{t_2} f(u) dW_u \\ &\sim \mathcal{N}\Big(b_s \Delta_n^{\beta-1/2} \frac{(t_1-s)^{1-\alpha} - (t_2-s)^{1-\alpha}}{1-\alpha}, \Delta_n^{2\beta-1} \sigma_s^2 \frac{(t_2-s)^{1-2\beta} - (t_1-s)^{1-2\beta}}{1-2\beta} \Big), \end{aligned} \tag{A.39}$$

which implies that:

$$\Pr(\breve{P}_{t_1}^{db} - \breve{P}_{t_2}^{db} \leq 0) = \Phi\left(\frac{b_s \sqrt{1-2\beta}}{\sigma_s (1-\alpha)} \frac{(t_2-s)^{1-\alpha} - (t_1-s)^{1-\alpha}}{\sqrt{(t_2-s)^{1-2\beta} - (t_1-s)^{1-2\beta}}} \right), \tag{A.40}$$

where $\Phi(x)$ is the cumulative density function of a standard normal distribution. Now consider the change of variable $\tilde{t}_1 = (t_1 - s)/\Delta_n$, $\tilde{t}_2 = (t_2 - s)/\Delta_n$ such that $\tilde{t}_1 < \tilde{t}_2 \in [0, 1]$. The above probability becomes:

$$\Pr(\breve{P}_{t_1}^{db} - \breve{P}_{t_2}^{db} \leq 0) = \Phi\left(\frac{b_s \sqrt{1-2\beta}}{\sigma_s (1-\alpha)} \cdot \Delta_n^{1/2-\alpha+\beta} \frac{(\tilde{t}_2^{1-\alpha} - \tilde{t}_1^{1-\alpha})}{\sqrt{\tilde{t}_2^{1-2\beta} - \tilde{t}_1^{1-2\beta}}} \right). \tag{A.41}$$

The above result directly implies that, for any choice of $\tilde{t}_1 < \tilde{t}_2 \in [0, 1]$,

$$\lim_{\Delta_n \to 0} \Pr(\breve{P}_{t_1}^{db} - \breve{P}_{t_2}^{db} \leq 0) \to \Phi(K \cdot \Delta_n^{1/2-\alpha+\beta}) = 1, \tag{A.42}$$

for some constant $K > 0$ by the assumption that $b_s > 0$, while the case $\tilde{t}_1 = \tilde{t}_2$ is trivial since $\Pr(\breve{P}_{t_1}^{db} - \breve{P}_{t_1}^{db} \leq 0) = 1$ for all n. This proves Eq. (A.38). Also, we note that $b_s < 0$ would imply an asymptotically non-increasing P_t^{db} in this case. Finally, it suffices to notice that the event $m = 0$ is identical to the event that P_t^{db} is monotonically non-increasing or non-decreasing on I_n. Therefore, Eq. (A.42) imply that, for every fixed $\epsilon > 0$:

$$\begin{aligned} &\Pr(\Delta_n^{\beta-1/2} m > \epsilon) = \Pr(\Delta_n^{\beta-1/2} \sup_{h \in I_n} \{w_h - |r_h|\} > \epsilon) \\ =& \Pr(\breve{P}_{t_1}^{db} - \breve{P}_{t_2}^{db} > \epsilon, \exists t_1 < t_2 | P_t^{db} \geq P_s^{db}, \forall t \in I_n) \Pr(P_t^{db} \geq P_s^{db}, \forall t \in I_n) \\ &+ \Pr(\Delta_n^{\beta-1/2} m > \epsilon | P_t^{db} < P_s^{db}, \exists t \in I_n) \Pr(P_t^{db} < P_s^{db}, \exists t \in I_n) \\ \to& 0 \cdot 1 + 0 = 0, \end{aligned} \tag{A.43}$$

Eq. (A.43) shows that, by definition, $\Delta_n^{\beta-1/2} m \xrightarrow{p} 0$, which leads to the result related to the MAED estimator in Eq. (3.14) by a standard continuous mapping argument. This completes the proof. □

Proof of Proposition 3.3. We shall continue on the premise of the discussion in the beginning of the proof of Theorem 3.1. We start by showing that $\Delta_n = O_p(\delta_n)$ for any fixed q under Assumption 3.4. Indeed, we have:

$$\Big|\Delta_n - \sum_{j=0}^{q-1} \frac{\delta_n}{\alpha_{\tau_{n,i_s+j}}}\Big| = \Big|\sum_{j=1}^{q}(\Delta\tau_{n,i_s+j} - \frac{\delta_n}{\alpha_{\tau_{n,i_s+j-1}}})\Big| \leq \sum_{j=1}^{q}\Big|\Delta\tau_{n,i_s+j} - \frac{\delta_n}{\alpha_{\tau_{n,i_s+j-1}}}\Big| = O_p(\delta_n^{1+\kappa}), \quad \text{(A.44)}$$

by Assumption 3.4(3) and the triangle inequality. This implies that:

$$\Delta_n = \delta_n \cdot O_p(1) + O_p(\delta_n^{1+\kappa}) = O_p(\delta_n) + o_p(\delta_n), \quad \text{(A.45)}$$

by the boundedness of α_t, as desired.

We now return to the proof. We start by the precise definitions of the limiting variables of the OMK estimator in $\boldsymbol{z}^{(q)}$:

$$\begin{aligned}
\eta^{(q)} &:= \max_{1\leq j\leq q}\{\xi_1^{(q)}(j) - |\xi_2^{(q)}(j)|\}, \quad \xi_1^{(q)} := \xi_1^{(q)}(q), \quad \xi_2^{(q)} := \xi_2^{(q)}(q), \\
\xi_1^{(q)}(j) &:= \Delta_n^{-1/2} \max_{1\leq i,i'\leq j} |W^{(q)}(i_s+i) - W^{(q)}(i_s+i')|, \\
\xi_2^{(q)}(j) &:= \Delta_n^{-1/2} |W^{(q)}(i_s+j) - W^{(q)}(i_s)|,
\end{aligned} \quad \text{(A.46)}$$

and recall that $W^{(q)}(i) := W_{\tau_{n,i}^{(q)}}$ is the ith equidistant observation of W on I_n, where $\tau_{n,i}^{(q)} := \tau_{n,i_s} + i\Delta_n/q$ is the ith equidistant grid on I_n. We shall also define $W(i) := W_{\tau_{n,i_s+i}}$ as the non-equidistant observations of W, and construct η, ξ_1, ξ_2, $\xi_1(j)$, $\xi_2(j)$ as in Eq. (A.46) but based on $(W(i))_{i\in I_n^{(q)}}$ instead of $(W^{(q)}(i))_{i\in I_n^{(q)}}$. The only difference between $(\eta^{(q)}, \xi_1^{(q)}, \xi_2^{(q)})$ and (η, ξ_1, ξ_2) is the observation times of W, and the superscript (q) denotes whether the underlying process is observed equidistantly or not. However, as we shall show later in this proof, these two versions of the limiting variables are asymptotically equivalent due to Assumption 3.4(3).

We start with a lemma which shows that the observation times $\{\tau_{n,i}\}_{i\in I_n^{(q)}}$ and $\{\tau_{n,i}^{(q)}\}_{i\in I_n^{(q)}}$ are asymptotically equivalent:

Lemma A.2. *It holds for all* $1 \leq j \leq q$ *that* $|\tau_{n,i_s+j} - \tau_{n,i_s+j}^{(q)}| = O_p(\delta_n^{1+\kappa})$.

Proof. Since $\Delta_n = O_p(\delta_n)$, we may deduce that α_t contains no jump on I_n with probability 1 in the limit. Therefore, we can assume that α has the following representation without loss of generality:

$$\alpha_t = \alpha_{\tau_{n,i_s}} + \int_{\tau_{n,i_s}}^t b_s' ds + \int_{\tau_{n,i_s}}^t \sigma_s' dW_s, \quad \text{(A.47)}$$

where $t \in I_n$ and the processes b_s' and σ_s' satisfy Assumption 3.1. In view of Eq. (A.10), we can further simplify α_t to a continuous martingale with fixed volatility in the limit on I_n:

$$\alpha_t = \alpha_{\tau_{n,i_s}} + \sigma_{\tau_{n,i_s}}' \int_{\tau_{n,i_s}}^t dW_s + o_p(\delta_n^{1/2}). \quad \text{(A.48)}$$

We therefore have $\alpha_{\tau_{n,i_s+j}} = \alpha_{\tau_{n,i_s}} + O_p(\delta_n^{1/2})$ for all $1 \le j \le q$. By Assumption 3.4(3) and the triangle inequality, we have:

$$\left|\Delta\tau_{n,i_s+j} - \frac{\delta_n}{\alpha_{\tau_{n,i_s}}}\right| \le \left|\Delta\tau_{n,i_s+j} - \frac{\delta_n}{\alpha_{\tau_{n,i_s+j-1}}}\right| + \left|\frac{\delta_n(\alpha_{\tau_{n,i_s+j-1}} - \alpha_{\tau_{n,i_s}})}{\alpha_{\tau_{n,i_s+j-1}}\alpha_{\tau_{n,i_s}}}\right| = O_p(\delta_n^{1+\kappa}) + O_p(\delta_n^{3/2}). \quad \text{(A.49)}$$

For $\kappa < 1/2$ we can ignore the last term and conclude that $|\Delta\tau_{n,i_s+j} - \delta_n/\alpha_{\tau_{n,i_s}}| = O_p(\delta_n^{1+\kappa})$. This first leads to the estimate:

$$\left|\Delta_n - \frac{q\delta_n}{\alpha_{\tau_{n,i_s}}}\right| = \left|\tau_{n,i_s+q} - \tau_{n,i_s} - \frac{q\delta_n}{\alpha_{\tau_{n,i_s}}}\right| \le \sum_{i=1}^{q} |\Delta\tau_{n,i_s+i} - \delta_n/\alpha_{\tau_{n,i_s}}| = O_p(\delta_n^{1+\kappa}). \quad \text{(A.50)}$$

Note that this implies $\Delta_n/\delta_n = q/\alpha_{\tau_{n,i_s}} + o_p(1)$. We thus have, for each j:

$$\begin{aligned} |\tau_{n,i_s+j} - \tau^{(q)}_{n,i_s+j}| &= \left|\tau_{n,i_s+j} - \tau_{n,i_s} - \frac{j\Delta_n}{q}\right| \le \left|\sum_{i=1}^{j}(\Delta\tau_{n,i_s+i} - \delta_n/\alpha_{\tau_{n,i_s}})\right| + \frac{j}{q}\left|\frac{q\delta_n}{\alpha_{\tau_{n,i_s}}} - \Delta_n\right| \\ &= O_p(\delta_n^{1+\kappa}) + O_p(\delta_n^{1+\kappa}) = O_p(\delta_n^{1+\kappa}), \end{aligned} \quad \text{(A.51)}$$

as desired. This completes the proof. □

Lemma A.2 implies the following result for every $1 \le j \le q$:

$$|W(i_s+j) - W(i_s+j)^{(q)}| = O_p(|\tau_{n,i_s+j} - \tau^{(q)}_{n,i_s+j}|^{1/2}) = O_p(\delta_n^{1/2+\kappa/2}), \quad \text{(A.52)}$$

since $|W_t - W_s| = O_p(|t-s|^{1/2})$, thus we can also conclude that:

$$\max_{1\le j\le q} |W(i_s+j) - W^{(q)}(i_s+j)| = o_p(\delta_n^{1/2}), \quad \text{(A.53)}$$

as q is assumed to be finite and $\kappa > 0$. We now prove the following estimates, which is the discrete counterpart of Eq. (A.9):

$$\max_{1\le j\le q} |w(j) - \sqrt{\Delta_n}\sigma_s\xi_1^{(q)}(j)| = o_p(\delta_n^{1/2}), \quad \max_{1\le j\le q} ||r(j)| - \sqrt{\Delta_n}\sigma_s\xi_2^{(q)}(j)| = o_p(\delta_n^{1/2}). \quad \text{(A.54)}$$

Starting with the second claim:

$$\begin{aligned} \max_{1\le j\le q} ||r(j)| - \sqrt{\Delta_n}\sigma_s\xi_2^{(q)}(j)| &= \max_{1\le j\le q} ||r(j)| - \sqrt{\Delta_n}\sigma_s\xi_2(j) + \sqrt{\Delta_n}\sigma_s\xi_2(j) - \sqrt{\Delta_n}\sigma_s\xi_2^{(q)}(j)| \\ &\le \max_{1\le j\le q} ||r(j)| - \sqrt{\Delta_n}\sigma_s\xi_2(j)| + \sqrt{\Delta_n}\sigma_s \max_{1\le j\le q} |\xi_2(j) - \xi_2^{(q)}(j)| \\ &\le \sup_{h\in I_n} ||r_h| - \sqrt{\Delta_n}\sigma_s\xi_{2,h}| + \sigma_s \max_{1\le j\le q} |W(j) - W^{(q)}(j)| \\ &\le o_p(\delta_n^{1/2}) + o_p(\delta_n^{1/2}) = o_p(\delta_n^{1/2}), \end{aligned} \quad \text{(A.55)}$$

where in the second inequality, the first term is due to the fact that the supremum of a continuous process is not smaller than the maximum of the discrete version of the same process. The estimates in the third line of the above equation therefore follow from Eq. (A.10) and Eq. (A.52). The first claim

in Eq. (A.54) can be proven in a similar manner:

$$\begin{aligned}\max_{1\le j\le q}|w(j)-\sqrt{\Delta_n}\sigma_s\xi_1^{(q)}(j)| &= \max_{1\le j\le q}|w(j)-\sqrt{\Delta_n}\sigma_s\xi_1(j)+\sqrt{\Delta_n}\sigma_s\xi_1(j)-\sqrt{\Delta_n}\sigma_s\xi_1^{(q)}(j)|\\ &\le \max_{1\le j\le q}|w(j)-\sqrt{\Delta_n}\sigma_s\xi_1(j)|+\sqrt{\Delta_n}\sigma_s\max_{1\le j\le q}|\xi_1(j)-\xi_1^{(q)}(j)|\\ &\le \sup_{h\in I_n}|w_h-\sqrt{\Delta_n}\sigma_s\xi_{1,h}|+2\sigma_s\max_{j\in I_n^{(q)}}|W(j)-W^{(q)}(j)|\\ &\le o_p(\delta_n^{1/2})+o_p(\delta_n^{1/2})=o_p(\delta_n^{1/2}),\end{aligned} \tag{A.56}$$

where the second inequality follows since:

$$\begin{aligned}&\max_{1\le j\le q}\sqrt{\Delta_n}|\xi_1(j)-\xi_1^{(q)}(j)|\\ =&\max_{1\le j\le q}\Big|\max_{1\le i,i'\le j}|W(i_s+i)-W(i_s+i')|-\max_{1\le i,i'\le j}|W^{(q)}(i_s+i)-W^{(q)}(i_s+i')|\Big|\\ \le&\max_{1\le i,i'\le q}\Big||W(i_s+i)-W(i_s+i')|-|W^{(q)}(i_s+i)-W^{(q)}(i_s+i')|\Big|\\ \le&\max_{1\le i,i'\le q}\Big|W(i_s+i)-W^{(q)}(i_s+i)-W(i_s+i')+W^{(q)}(i_s+i')|\Big|\\ \le&\max_{1\le i\le q}|W(i_s+i)-W^{(q)}(i_s+i)|+\max_{1\le i'\le q}|W(i_s+i')+W^{(q)}(i_s+i')|\\ =&2\max_{j\in I_n^{(q)}}|W(j)-W^{(q)}(j)|=o_p(\delta_n^{1/2}).\end{aligned} \tag{A.57}$$

In the same vein as the derivation of Eq. (A.13), Eq. (A.54) directly leads to the following result:

$$\begin{aligned}|m-\sqrt{\Delta_n}\sigma_s\eta^{(q)}| &\le \max_{1\le j\le q}|w(j)-\sqrt{\Delta_n}\sigma_s\xi_1^{(q)}(j)|+\max_{1\le j\le q}||r(j)|-\sqrt{\Delta_n}\sigma_s\xi_2^{(q)}(j)|\\ &\le o_p(\delta_n^{1/2})+o_p(\delta_n^{1/2})=o_p(\delta_n^{1/2}),\\ |w-\sqrt{\Delta_n}\sigma_s\xi_1^{(q)}| &\le \max_{1\le j\le q}|w(j)-\sqrt{\Delta_n}\sigma_s\xi_1^{(q)}(j)|=o_p(\delta_n^{1/2}),\\ ||r|-\sqrt{\Delta_n}\sigma_s\xi_2| &\le \max_{1\le j\le q}||r(j)|-\sqrt{\Delta_n}\sigma_s\xi_2^{(q)}(j)|=o_p(\delta_n^{1/2}).\end{aligned} \tag{A.58}$$

Following the same procedure as in Eqs. (A.14) and (A.15), the above results leads to the desired coupling result in Eq. (3.19) via the Slutsky theorem and the continuous mapping theorem, which completes the proof. □

Proof of Proposition 3.4. Following the notation in Eq. (3.16), we write m^ϵ, w^ϵ, and $|r^\epsilon|$ (resp. m, w, and r) as the MAED, range, and absolute return of $(P^\epsilon(i))_{i\in I_n^{(q)}}$ (resp. $(P(i))_{i\in I_n^{(q)}}$), respectively. We also define the processes $w^\epsilon(j)$ and $r^\epsilon(j)$ as the running range and return up to the jth observation, where $1\le j\le q$. Similarly, define $\xi_1^\epsilon(j)$ and $\xi_2^\epsilon(j)$ as the running range and absolute return up to the jth observation constructed from $(\epsilon(i))_{i=i_s:i_s+j}$. The key to the proof is the following result:

Lemma A.3. *It holds for all* $1\le j\le q$ *that:* $w^\epsilon(j)=\xi_1^\epsilon(j)+O_p(\delta_n^{-1/2})$, $|r^\epsilon(j)|=\xi_2^\epsilon(j)+O_p(\delta_n^{-1/2})$.

Proof. For the first claim, we have:

$$\begin{aligned}
|w^\epsilon(j) - \xi_1^\epsilon(j)| &= \Big| \max_{1\le h,h'\le j} |P^\epsilon(i_s+h) - P^\epsilon(i_s+h')| - \max_{1\le h,h'\le j} |\epsilon(i_s+h) - \epsilon(i_s+h')| \Big| \\
&\le \max_{1\le h,h'\le j} \Big| |P^\epsilon(i_s+h) - P^\epsilon(i_s+h')| - |\epsilon(i_s+h) - \epsilon(i_s+h')| \Big| \\
&\le \max_{1\le h,h'\le j} |P^\epsilon(i_s+h) - \epsilon(i_s+h) - (P^\epsilon(i_s+h) - \epsilon(i_s+h))| \\
&= \max_{1\le h,h'\le j} |P(i_s+h) - P(i_s+h)| = w(j) = O_p(\delta_n^{-1/2}).
\end{aligned} \tag{A.59}$$

where the first inequality follows from Lemma A.1, the second inequality is the reverse triangle inequality, and the last estimate follows from the proof of Proposition 3.3. Similarly:

$$||r^\epsilon(j)| - \xi_2^\epsilon(j)| \le |r^\epsilon(j) - \epsilon(i_s+j) + \epsilon(i_s))| = |r(j)| = O_p(\delta_n^{-1/2}). \tag{A.60}$$

The two results above directly imply the lemma, and the proof is complete. □

As a direct consequence of Lemma A.3 with an application of Lemma A.1, we see that:

$$\begin{aligned}
|m^\epsilon - \eta^\epsilon| &= \Big| \max_{1\le j\le q} \{w^\epsilon(j) - |r^\epsilon(j)|\} - \max_{1\le j\le q} \{\xi_1^\epsilon(j) - \xi_2^\epsilon(j)\} \Big| \\
&\le \max_{1\le j\le q} \Big| w^\epsilon(j) - \xi_1^\epsilon(j) - (|r^\epsilon(j)| - \xi_2^\epsilon(j)) \Big| \\
&\le \max_{1\le j\le q} |w^\epsilon(j) - \xi_1^\epsilon(j)| + \max_{1\le j\le q} ||r^\epsilon(j)| - \xi_2^\epsilon(j)| = O_p(\delta_n^{-1/2}), \\
|w^\epsilon - \xi_1^\epsilon| &= |w^\epsilon(q) - \xi_1^\epsilon(q)| = O_p(\delta_n^{-1/2}), \\
||r^\epsilon| - \xi_2^\epsilon| &= ||r^\epsilon(q)| - \xi_2^\epsilon(q)| = O_p(\delta_n^{-1/2}).
\end{aligned} \tag{A.61}$$

By the Slutsky theorem, we have the following result in vector form:

$$(m^\epsilon, w^\epsilon, |r^\epsilon|) = (\eta^\epsilon, \xi_1^\epsilon, \xi_2^\epsilon) + O_p(\delta_n^{-1/2}), \tag{A.62}$$

which can be written in the vector form $\boldsymbol{c}^\epsilon = \boldsymbol{\epsilon}^{(q)} + O_p(\delta_n^{-1/2})$, where $\boldsymbol{c}^\epsilon := (m^\epsilon, w^\epsilon, |r^\epsilon|)'$. It now suffices to notice that, by a continuous mapping argument:

$$\hat{\sigma}_t^\epsilon(\boldsymbol{\lambda}) = \frac{\boldsymbol{\lambda}'\boldsymbol{\Theta}^{(q)}\boldsymbol{c}^\epsilon}{\sqrt{\Delta_n}} = \Delta_n^{-1/2}(\boldsymbol{\lambda}'\boldsymbol{\Theta}^{(q)}\boldsymbol{\epsilon}^{(q)} + O_p(\delta_n^{-1/2})) = \Delta_n^{-1/2}\boldsymbol{\lambda}'\boldsymbol{\Theta}^{(q)}\boldsymbol{\epsilon}^{(q)} + O_p(1), \tag{A.63}$$

as desired. This completes the proof. □

Proof of Proposition 3.5. We shall continue on the premise of the discussion in the beginning of the proof of Theorem 3.2. Following the reasoning in Eq. (A.44), we can conclude that $\Delta_n = O_p(\delta_n^{1-\theta}) = o_p(1)$. We first show that the pre-averaging procedure can 'recover' the efficient price from noisy observations in the following sense:

$$\max_{i\in I_n^{(q_n)}} |\bar{P}(i) - P(i)| = O_p(\delta_n^{1/2+\varepsilon-\theta}) + O_p(\delta_n^{(\varepsilon-\theta)(d-1/2)}), \tag{A.64}$$

To see this, note that we have:

$$\begin{aligned}
\max_{i\in I_n^{(q_n)}} |\bar{P}(i) - P(i)| &\le \max_{i\in I_n^{(q_n)}} \Big| \frac{1}{k_n} \sum_{j=0}^{k_n-1} (P(i-j) - P(i)) \Big| + \max_{i\in I_n^{(q_n)}} \Big| \frac{1}{k_n} \sum_{j=0}^{k_n-1} \epsilon(i-j) \Big| \\
&= O_p(\sqrt{k_n\delta_n}) + O_p(k_n^{d-1/2}) = O_p(\delta_n^{(1+\varepsilon-\theta)/2}) + O_p(\delta_n^{(\varepsilon-\theta)(d-1/2)}),
\end{aligned} \tag{A.65}$$

where the second estimate is a direct result of Assumption 3.5. For the first estimate:

$$\begin{aligned}&\max_{i\in I_n^{(q_n)}}\left|\frac{1}{k_n}\sum_{j=0}^{k_n-1}(P(i-j)-P(i))\right| = \max_{i\in I_n^{(q_n)}}\left|\sum_{j=1}^{k_n-1}\frac{k_n-j}{k_n}(P(i-j)-P(i-j+1))\right| \\ &= \max_{i\in I_n^{(q_n)}}\left|\int_{\tau_{n,i-k_n+1}}^{\tau_{n,i}} g(s)b_s ds + \int_{\tau_{n,i-k_n+1}}^{\tau_{n,i}} g(s)\sigma_s dW_s\right| = O_p(\sqrt{k_n\delta_n}),\end{aligned} \tag{A.66}$$

where $g(t) = 1 - j/k_n$ for $t \in [\tau_{n,i-j}, \tau_{n,i-j+1})$ which summarizes the impact of the pre-averaging, and the last estimate is due to the boundedness of b_t and σ_t and the estimate $\tau_{n,i} - \tau_{n,i-k_n+1} = O(k_n\delta_n)$ (in view of Eq. (A.44)). Note that $g(t)$ does not change the asymptotic order of the estimates as it is bounded above by 1. Let $\boldsymbol{c} = (m, w, |r|)$ denote the MAED-candlestick vector constructed from $(P(i))_{i\in I_n^{(q_n)}}$, then Eq. (A.64) directly implies that:

$$|\overline{\boldsymbol{c}} - \boldsymbol{c}| = O_p(\delta_n^{(1+\varepsilon-\theta)/2}) + O_p(\delta_n^{(\varepsilon-\theta)(d-1/2)}). \tag{A.67}$$

To conserve space, we shall only prove the above estimate for the range statistic, and the corresponding results for the MAED and the return follow by a similar argument. We have:

$$\begin{aligned}|\bar{w} - w| &\leq \max_{i,i'\in I_n^{(q_n)}} ||\bar{P}(i) - \bar{P}(i')| - |P(i) - P(i')|| \\ &\leq 2\max_{i\in I_n^{(q_n)}} |\bar{P}(i) - P(i)| = O_p(\delta_n^{(1+\varepsilon-\theta)/2}) + O_p(\delta_n^{(\varepsilon-\theta)(d-1/2)}),\end{aligned} \tag{A.68}$$

which is the desired asymptotic order, where the first inequality is due to Lemma A.1. We thus see that:

$$\Delta_n^{-1/2}|\boldsymbol{\lambda}'\boldsymbol{\Theta}\overline{\boldsymbol{c}} - \boldsymbol{\lambda}'\boldsymbol{\Theta}\boldsymbol{c}| = O_p(\delta_n^{\varepsilon/2}) + O_p(\delta_n^{(\varepsilon-\theta)(d-1/2)-(1-\theta)/2}) = o_p(1), \tag{A.69}$$

where the last estimate holds iff $\varepsilon > 0$ and $(\varepsilon-\theta)(d-1/2)-(1-\theta)/2 > 0$. One should verify that the latter condition is equivalent to $(\theta-\varepsilon)/(1-\varepsilon) > 1/(2-2d)$, which gives the require range of θ and ε.

We proceed to show that $|\Delta_n^{-1/2}\boldsymbol{\lambda}'\boldsymbol{\Theta}\boldsymbol{c} - \sigma_t\boldsymbol{\lambda}'\boldsymbol{\Theta}\boldsymbol{z}| = o_p(1)$ in the limit, which implies Eq. (3.26) in view of Eq. (A.69). Note that this result is very similar to Proposition 3.3 except that the number of observations q_n in I_n diverges instead of being constant. By modifying the asymptotic order of Δ_n in Lemma A.2, it follows that:

$$|\tau_{n,i_s+j} - \tau_{n,i_s+j}^{(q_n)}| = O_p(\delta_n^{1-\theta+\kappa}) + O_p(\delta_n^{3(1-\theta)/2}), \quad 1 \leq j \leq q_n, \tag{A.70}$$

where $\tau_{n,i}^{(q_n)} := \tau_{n,i_s} + i\Delta_n/q_n$ is the ith equidistant grid on I_n as in Lemma A.2. This implies that:

$$\max_{1\leq j\leq q_n} |W(i_s+j) - W^{(q_n)}(i_s+j)| = O_p(\delta_n^{(1-\theta+\kappa)/2}) + O_p(\delta_n^{3(1-\theta)/4}) = o_p(\delta_n^{(1-\theta)/2}), \tag{A.71}$$

where $(W^{(q_n)}(i))_{i\in I_n^{(q_n)}}$ is the (q_n+1)-times equidistant observations of W on I_n. Let us also write $(W_t^{(q_n)})_{t\in I_n}$ as the right continuous version of $(W^{(q_n)}(i))_{i\in I_n^{(q_n)}}$, which is understood as a discretized version of W on I_n. Proposition 1 of Asmussen et al. (1995) gives an estimate of the asymptotic order of the discretization error:

$$\sup_{t\in I_n} |W_t - W_t^{(q_n)}| = O_p(\sqrt{(\Delta_n/q_n)|\ln q_n|}) = O_p(\delta_n^{1/2}\sqrt{|\ln\delta_n|}). \tag{A.72}$$

for any $\alpha > 0$. This result and Eq. (A.71) further imply that $|\boldsymbol{c} - \sigma_s \Delta_n^{1/2} \boldsymbol{z}| = o_p(\delta_n^{(1-\theta)/2})$. We note that this can be proven using similar arguments as for the proof of Eq. (A.58), and for brevity we only show the result for the range statistic. We have:

$$\begin{aligned} |w - \sigma_s \Delta_n^{1/2} \xi_1| &\leq |w - \sigma_s \Delta_n^{1/2} \xi_1^{(q_n)}| + \sigma_s \Delta_n^{1/2} |\xi_1^{(q_n)} - \xi_1| \\ &\leq o_p(\sqrt{\Delta_n}) + 2\sigma_s \Delta_n^{1/2} \sup_{t \in I_n} |W_t - W_t^{(q_n)}| \\ &\leq o_p(\delta_n^{(1-\theta)/2}) + O_p(\delta_n^{1-\theta/2} \sqrt{|\ln \delta_n|}) = o_p(\delta_n^{(1-\theta)/2}), \end{aligned} \tag{A.73}$$

where $\xi_1^{(q_n)}$ is understood as the range statistic of $(W^{(q_n)}(i))_{i \in I_n^{(q_n)}}$. The first estimate in the first inequality is due to Eq. (A.13), and the final estimate follows since $O(\sqrt{\delta_n |\ln \delta_n|}) = o(1)$. Finally, it suffices to notice that:

$$\begin{aligned} |\hat{\tilde{\sigma}}_t(\boldsymbol{\lambda}) - \sigma_s \boldsymbol{\lambda}' \boldsymbol{\Theta} \boldsymbol{z}| &\leq \Delta_n^{-1/2} |\boldsymbol{\lambda}' \boldsymbol{\Theta} \bar{\boldsymbol{c}} - \boldsymbol{\lambda}' \boldsymbol{\Theta} \boldsymbol{c}| + |\Delta_n^{-1/2} \boldsymbol{\lambda}' \boldsymbol{\Theta} \boldsymbol{c} - \sigma_s \boldsymbol{\lambda}' \boldsymbol{\Theta} \boldsymbol{z}| \\ &\leq o_p(1) + o_p(1) = o_p(1), \end{aligned} \tag{A.74}$$

which implies $\hat{\tilde{\sigma}}_t(\boldsymbol{\lambda}) = \sigma_s \boldsymbol{\lambda}' \boldsymbol{\Theta} \boldsymbol{z} + o_p(1)$. We can thus obtain Eq. (3.26) by dividing σ_t on both sides of this relation following the same logic as Eq. (A.16), and the proof is complete. □

B The OMK Estimator for Spot Variance

Follow the setting in Section 3.2, we start by defining an OMK estimator for σ_t^2. Let the vector $\boldsymbol{c}_2 := (m^2, mw, m|r|, w^2, |r|^2, w|r|)'$ collects all second order statistics computed from MAED, range and absolute return, and let $\boldsymbol{\Theta}_2 := \text{diag}(\mu_2^{-1}, \gamma_1^{-1}, \gamma_2^{-1}, \nu_2^{-1}, \psi_2^{-1}, \gamma_0^{-1})'$ denote the corresponding matrix of normalizing factors. The OMK estimator of spot variance is defined as:

$$\hat{\sigma}_t^2(\boldsymbol{\lambda}) := \Delta_n^{-1/2} \boldsymbol{\lambda}' \boldsymbol{\Theta}_2 \boldsymbol{c}_2, \tag{B.1}$$

for some weight vector $\boldsymbol{\lambda} = (\lambda_{mm}, \lambda_{mw}, \lambda_{mr}, \lambda_{ww}, \lambda_{rr}, \lambda_{wr})'$ that adds up to one. Using a similar argument as Theorem 3.1, one can prove the following coupling result:

$$\frac{\hat{\sigma}_t^2(\boldsymbol{\lambda})}{\sigma_t^2} = \boldsymbol{\lambda}' \boldsymbol{\Theta}_2 \boldsymbol{z}_2 + o_p(1), \tag{B.2}$$

where $\boldsymbol{z}_2 := (\eta^2, \eta\xi_1, \eta\xi_2, \xi_1^2, \xi_2^2, \xi_1\xi_2)'$. Following the discussion after Remark 3.1, we can pick optimal weight vectors by minimizing $\text{Var}[\boldsymbol{\lambda}' \boldsymbol{\Theta}_2 \boldsymbol{z}_2]$ subject to $\boldsymbol{\lambda}' \boldsymbol{\iota} = 1$ where the variance-covariance matrix $\text{Var}[\boldsymbol{\Theta}_2 \boldsymbol{z}_2]$ can be easily simulated. We find the optimal OMK weights and the corresponding minimized variance to be approximately:

$$\boldsymbol{\lambda}_2^* \approx (0.523, 0.683, -0.004, -0.816, -0.170, 0.784)', \quad \text{Var}[\boldsymbol{\lambda}_2^{*\prime} \boldsymbol{\Theta}_2 \boldsymbol{z}_2] \approx 0.153. \tag{B.3}$$

Restricting $\lambda_{mm} = \lambda_{mw} = \lambda_{mr} = 0$, the optimal OK weight and the corresponding minimized variance are (which can be derived in closed form):

$$\boldsymbol{\lambda}_2^o \approx (0, 0, 0, 1.7103, 0.0544, -0.7647)', \quad \text{Var}[\boldsymbol{\lambda}_2^{o\prime} \boldsymbol{\Theta}_2 \boldsymbol{z}_2] \approx 0.2594. \tag{B.4}$$

Therefore, the optimal OMK estimator has an asymptotic variance factor that is about 41% smaller than that of the optimal OK estimator for σ_t^2, so that the variance reduction is almost the same as the OMK estimator for σ_t.

It is worth noting that the 'practical version' of the Garman-Klass estimator can be obtained by further restricting $\lambda_{mm} = \lambda_{mw} = \lambda_{mr} = \lambda_{wr} = 0$, which takes the form[16] $\hat{\sigma}^2_{t,GK} \approx \Delta_n^{-1/2}(0.5031w^2 - 0.3949r^2)$. We find that $\text{avar}(\hat{\sigma}^2_{t,GK}/\sigma_t^2) \approx 0.2686$, which is worse than our OK estimator here due to the restriction on the weight vector. In fact, we have $\text{Var}[\boldsymbol{e}_1\boldsymbol{\Theta}_2\boldsymbol{z}_2] \approx 0.270$, so the MAED estimator of σ_t^2 has almost identical performance to the Garman-Klass estimator in the continuous case. Finally, we note that the optimal OK variance factor of 0.2594 can be improved slightly to 0.2587 if we decompose w and $|r|$ further into high, low and close returns based on the method of Meilijson (2011) (see also the discussion in Bollerslev et al. (2022)). As the improvement is empirically insignificant, we do not pursue this approach further in this paper.

C Simulated Moments and Critical Values of the MAED-Based Statistics

This section provides a comprehensive guidance on the choice of various moments and critical values for the OMK estimator and the S-test in the discrete case under Assumption 3.4. We note that the results in the continuous case are obtained by letting $q \to \infty$.

We start by simulating the vector $(\tilde{\eta}^{(q)}, \tilde{\xi}_1^{(q)}, \tilde{\xi}_2)'$ with various choices of q, which can be easily constructed from a standard Brownian motion equidistantly observed with $q+1$ observations on $[0,1]$. As q can take any natural number, we consider the following simulation scheme to make full use of our computational resources:

1. For $q \in \{2, 3, 4, \dots 10\}$, we draw 10^9 replications of $(\tilde{\eta}^{(q)}, \tilde{\xi}_1^{(q)}, \tilde{\xi}_2)'$.
2. For $q \in \{11, 12, \dots, 200\}$, we draw 10^8 replications of $(\tilde{\eta}^{(q)}, \tilde{\xi}_1^{(q)}, \tilde{\xi}_2)'$.
3. For $q \in \{201, 202, \dots, 2000, 2005, 2010, \dots, 5000, 5010, 5020, \dots, 10^4, 10^5, 10^6, 10^7, 10^8\}$, we draw 10^7 replications of $(\tilde{\eta}^{(q)}, \tilde{\xi}_1^{(q)}, \tilde{\xi}_2)'$.

First, in Table C.1 we present the simulation results related to the first two moments of $(\tilde{\eta}^{(q)}, \tilde{\xi}_1^{(q)}, \tilde{\xi}_2)'$, namely $\mu_1^{(q)}$, $\mu_2^{(q)}$, $\nu_1^{(q)}$, $\nu_2^{(q)}$, $\gamma_0^{(q)}$, $\gamma_1^{(q)}$ and $\gamma_2^{(q)}$. We compute these moments by their corresponding Monte Carlo averages for each q. For $q \leq 10$, we directly use the simulated values, which has the highest precision in our simulation schemes. The standard error of the simulated moments are also presented in Panel 2 of that table to assess the precision of the simulated moments. For $q \geq 11$, we

[16] Note that the original definition in Garman and Klass (1980) is $\hat{\sigma}^2_{t,GK} \approx \Delta_n^{-1/2}(0.5w - (\ln 2 - 1)|r|)$, which has a slightly larger variance than the variance-optimal version given here.

propose to use the following polynomial approximation for the aforementioned moments:

$$Y^{(q)} = \sum_{i=0}^{k} \beta_i q^{-i/2} + u^{(q)}, \tag{C.1}$$

where $Y^{(q)} \in \{\mu_1^{(q)}, \mu_2^{(q)}, \nu_1^{(q)}, \nu_1^{(q)}, \gamma_0^{(q)}, \gamma_1^{(q)}, \gamma_2^{(q)}\}$ are the simulated moments to be approximated and $u^{(q)}$ is a nearly zero-mean residual term. The specification of the regression is inspired by Proposition 3 of Asmussen et al. (1995), which derive an asymptotic expansion for moments of reflected Brownian motion up to the $q^{-1/2}$ term. In fact, the analytical coefficient β_1 in Panel 3 of Table C.1 is due to Asmussen et al. (1995), which is also used in Andersen et al. (2008). We find that a cubic approximation works very well across all $q \geq 11$ for the candlestick-based moments, i.e., $\nu_1^{(q)}$, $\nu_2^{(q)}$ and $\gamma_0^{(q)}$, while a quartic approximation is needed for the MAED-based moments with $11 \leq q \leq 200$. The regression standard error and the R^2 are presented in the table to evaluate the precision of the polynomial approximation. It should be clear that one can view β_0s in Panel 4 as the simulated moments in the continuous case by letting $q \to \infty$.

The simulation precision in Table C.1 is worth discussing. As the Monte Carlo means are asymptotically normal by a standard central limit theorem, one can interpret a simulated moment $\pm 2\times$SE as its 95% confidence bounds. Panel 1 thus shows that, for $q \leq 10$, the simulated $\mu_1^{(q)}$, $\mu_2^{(q)}$, $\gamma_1^{(q)}$, $\gamma_2^{(q)}$ and $\nu_2^{(q)}$ are precise up to the 4th digits after the decimal point, as the width of the 95% confidence bounds are narrower than 10^{-4}. $\nu_2^{(q)}$ and $\gamma_0^{(q)}$ are slightly less accurate, but the 95% confidence bounds indicate that the simulated error is likely to be smaller than ± 0.0001. Turning to Panels 2 and 3, all the R^2s of the regressions are virtually 1, which suggests an almost perfect fit of the polynomials to the simulated moments for the range of q considered. The fitted values are more precise for $11 \leq q \leq 200$ due to the larger Monte Carlo size, and one can observe that the relative size of the SEs among different panels roughly correspond to the Monte Carlo size. Importantly, assuming that the regression specification is correct, one can use the fitted value $\pm 2\times$regression SE to gauge the precision of approximated moments for any choice of q. The table thus shows that the largest approximation error of the moments is likely to be within ± 0.001 for $\nu_2^{(q)}$ and $\gamma_0^{(q)}$. In particular, $\mu_1^{(q)}$ and $\nu_1^{(q)}$, which determine the bias of the OMK estimator, can be approximated accurately up to the third digit after the decimal point. This result demonstrates the quality of the simulated moments and the validity of our polynomial approximation.

Based on the simulated values of $(\tilde{\eta}^{(q)}, \tilde{\xi}_1^{(q)}, \tilde{\xi}_2)'$, we also provide approximated critical values for $\hat{\sigma}_{t,OK}^{(q)}$, $\hat{\sigma}_{t,OMK}^{(q)}$, $\hat{\sigma}_{t,MAED}^{(q)}$ and $S_t^{(q)}$ with $\alpha \in \{10\%, 5\%, 1\%\}$. Critical values for non-conventional choices of α are available upon request, but they are omitted here for conciseness. Taking the critical values for the α% HDI of $\hat{\sigma}_{t,OMK}^{(q)}$ as an example. For each q and the associated Monte Carlo size, we first compute $\tilde{\boldsymbol{z}}^{(q)} := (\tilde{\eta}^{(q)}, \tilde{\xi}_1^{(q)}, \tilde{\xi}_2)'$ and $\boldsymbol{\Theta}^{(q)} = \text{diag}((\mu_1^{(q)})^{-1}, (\nu_1^{(q)})^{-1}, \psi_1^{-1})$, where $\mu_1^{(q)}$ and $\nu_1^{(q)}$ can be found in Table C.1. This allows us to compute $\text{Var}[\boldsymbol{\Theta}^{(q)}\tilde{\boldsymbol{z}}^{(q)}]$, and hence $\boldsymbol{\lambda}^*$, from the Monte Carlo draws based

$q =$ \ $Y^{(q)} =$	$\mu_1^{(q)}$	$\mu_2^{(q)}$	$\gamma_1^{(q)}$	$\gamma_2^{(q)}$	$\nu_1^{(q)}$	$\nu_2^{(q)}$	$\gamma_0^{(q)}$
Panel 1: Simulated Moments							
2	0.165	0.0908	0.159	0.0683	0.963	1.228	1.068
3	0.284	0.164	0.294	0.143	1.052	1.382	1.117
4	0.373	0.227	0.406	0.207	1.111	1.496	1.152
5	0.440	0.283	0.500	0.259	1.153	1.584	1.178
6	0.494	0.331	0.578	0.302	1.186	1.655	1.199
7	0.536	0.374	0.643	0.337	1.212	1.714	1.216
8	0.571	0.412	0.699	0.367	1.233	1.764	1.231
9	0.601	0.445	0.747	0.391	1.251	1.807	1.243
10	0.626	0.475	0.789	0.412	1.267	1.845	1.254
Panel 2: Simulated SE of the Simulated Moments$\times 10^4$							
2	0.0797	0.1732	0.0998	0.0423	0.0696	0.4427	0.4442
3	0.0911	0.1658	0.1224	0.0603	0.0865	0.4486	0.4470
4	0.0939	0.1619	0.1345	0.0704	0.0986	0.4547	0.4502
5	0.0942	0.1595	0.1421	0.0771	0.1080	0.4602	0.4532
6	0.0937	0.1579	0.1474	0.0822	0.1154	0.4650	0.4559
7	0.0930	0.1568	0.1515	0.0864	0.1215	0.4691	0.4583
8	0.0923	0.1560	0.1549	0.0902	0.1265	0.4728	0.4604
9	0.0918	0.1553	0.1579	0.0935	0.1308	0.4760	0.4623
10	0.0914	0.1548	0.1606	0.0965	0.1345	0.4788	0.4639
Panel 3: Quartic Approximation for $11 \leq q \leq 200$							
β_0	1.107	1.306	1.776	0.807	$\sqrt{8/\pi}$	$4\ln 2$	$3/2$
β_1	-1.568	-3.483	-3.860	-1.310	$\sqrt{2/\pi}\zeta(1/2)$	$8\zeta(1/2)/\pi$	$2\zeta(1/2)/\pi$
β_2		2.756	2.532	0.116	0.400	2.745	0.501
β_3	1.005			1.002	-0.0128	0.841	-0.0694
β_4	-1.689	-0.479	-1.924	-2.438			
Reg SE$\times 10^4$	0.2838	0.5826	0.6532	0.4039	0.4419	1.5430	1.3476
R^2	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Panel 4: Cubic Approximation for $201 \leq q$							
β_0	1.106	1.303	1.774	0.807	Same as $11 \leq q \leq 200$		
β_1	-1.533	-3.368	-3.791	-1.302	Same as $11 \leq q \leq 200$		
β_2	-0.541	1.067	1.596	0.127	Same as $11 \leq q \leq 200$		
β_3	3.994	9.344	4.727		Same as $11 \leq q \leq 200$		
Reg SE$\times 10^4$	0.8799	2.0336	2.2531	1.6825	1.5086	5.5544	5.0471
R^2	1.0000	1.0000	1.0000	0.9999	0.9999	0.9999	0.9981

Table C.1: Simulated and approximated values for the moments of $(\tilde{\eta}^{(q)}, \tilde{\xi}_1^{(q)}, \tilde{\xi}_2)'$. In Panel 1, we present the simulated moments for each $2 \leq q \leq 10$, and the corresponding standard errors (SE) are presented in Panel 2. Panels 3 and 4 present the estimated coefficients of the polynomial regression $Y^{(q)} = \sum_{i=0}^{4} \beta_i q^{-i/2} + u^{(q)}$ based on different ranges of q, with the regression SE and R^2. Note that $\zeta(x)$ is the Riemann Zeta function.

on Eq. (3.8). Consequently, we obtain the simulated distribution of $\boldsymbol{\lambda}^{*\prime}\boldsymbol{\Theta}^{(q)}\tilde{\boldsymbol{z}}^{(q)}$, which coincides with the asymptotic distribution of the OMK estimator according to Proposition 3.3. With a significance level of $(1-\alpha)\%$, we construct $\ell_\alpha(\boldsymbol{\lambda}^{*\prime}\boldsymbol{\Theta}^{(q)}\tilde{\boldsymbol{z}}^{(q)}) = [B_{\alpha^-}(\boldsymbol{\lambda}^{*\prime}\boldsymbol{\Theta}^{(q)}\tilde{\boldsymbol{z}}^{(q)}), B_{\alpha^+}(\boldsymbol{\lambda}^{*\prime}\boldsymbol{\Theta}^{(q)}\tilde{\boldsymbol{z}}^{(q)})]$ such that the number of observations in $\ell_\alpha(\boldsymbol{\lambda}^{*\prime}\boldsymbol{\Theta}^{(q)}\tilde{\boldsymbol{z}}^{(q)})$ is exactly $(1-\alpha)\%$ times the Monte Carlo size, where the critical values $B_{\alpha^\pm}(\boldsymbol{\lambda}^{*\prime}\boldsymbol{\Theta}^{(q)}\tilde{\boldsymbol{z}}^{(q)})$ are chosen by minimizing $B_{\alpha^+}(\boldsymbol{\lambda}^{*\prime}\boldsymbol{\Theta}^{(q)}\tilde{\boldsymbol{z}}^{(q)}) - B_{\alpha^-}(\boldsymbol{\lambda}^{*\prime}\boldsymbol{\Theta}^{(q)}\tilde{\boldsymbol{z}}^{(q)})$ over the grid of simulated values of $(\boldsymbol{\lambda}^{*\prime}\boldsymbol{\Theta}^{(q)}\tilde{\boldsymbol{z}}^{(q)})^{-1}$. Similar to Table C.1, we suggest to use the exact simulated critical values for $q \leq 10$ and provide polynomial approximations for $q > 10$. We present the values and their associated SEs of the simulated/approximated critical values in Table C.2.

We summarize some key results in Table C.2. First, we focus on the results in Panel 1a. By comparing critical values of OMK and OK estimators for $3 \leq q \leq 10$, we find that the HDI widths of the OMK estimator is always narrower than that of the OK estimator, regardless of the choices of q or α. This is in line with our conclusion in Table 3.2 that the OMK estimator outperforms the OK estimator for any choices of $q > 2$. The critical values of the MAED estimator and the associated test statistic is not available for various choices of q due to the non-negligible probability of $\tilde{\eta}^{(q)} = 0$ when q is small. In Panel 1b, the simulated standard errors are computed from the asymptotic variance of an order statistic, as all the critical values presented here are in essence an order statistic of the simulated distribution. Panel 1b thus shows that the simulated critical values for small q is reasonably precise for the OMK and the OK estimators, with SEs below 0.0001 for most choices of q. The SEs of the simulated critical values for the MAED estimator and the test statistic is noisier, especially for small q and α, but as q increases, the SEs shrinks to a similar scale relative to those of the OK and OMK estimators.

Polynomial approximation coefficients of the critical values for different ranges of q are presented in Panels 2 and 3 of Table C.2. For each regression, the order of the polynomial approximation is chosen such that the coefficient of an additional term is statistically insignificant. By comparing Panel 2 with Panel 3, we see that the polynomial approximation is somewhat more complicated for $11 \leq q \leq 200$ and for the MAED-related critical values, with a maximum of 8 polynomial terms for the MAED estimator with $\alpha = 1\%$. However, all the regressions have $R^2 \approx 1$ and similar levels of regression SEs around 0.001. These results suggest a very accurate fit to the simulated critical values, which are used in our simulation and empirical analyses. It is worth noting that in Panel 3, β_0 can be regarded as the critical values for the limiting case when the continuous price path is observed. MATLAB codes to compute the critical values according to Table C.2 are provided in the supplementary material of the paper.

	$B_{\alpha-}(\boldsymbol{\lambda}^{*\prime}\boldsymbol{\Theta}^{(q)}\tilde{\boldsymbol{z}}^{(q)})$			$B_{\alpha+}(\boldsymbol{\lambda}^{*\prime}\boldsymbol{\Theta}^{(q)}\tilde{\boldsymbol{z}}^{(q)})$			$B_{\alpha-}(\boldsymbol{\lambda}^{o\prime}\boldsymbol{\Theta}^{(q)}\tilde{\boldsymbol{z}}^{(q)})$			$B_{\alpha+}(\boldsymbol{\lambda}^{o\prime}\boldsymbol{\Theta}^{(q)}\tilde{\boldsymbol{z}}^{(q)})$			$B_{\alpha-}(\boldsymbol{e}_1'\boldsymbol{\Theta}^{(q)}\tilde{\boldsymbol{z}}^{(q)})$			$B_{\alpha+}(\boldsymbol{e}_1'\boldsymbol{\Theta}^{(q)}\tilde{\boldsymbol{z}}^{(q)})$			CVs of $S_t^{(q)}$		
$q=$ \ $\alpha=$	1%	5%	10%	10%	5%	1%	1%	5%	10%	10%	5%	1%	1%	5%	10%	10%	5%	1%	10%	5%	1%
Panel 1a: Simulated critical values, $q \leq 10$																					
3	0.327	0.386	0.423	2.167	2.779	4.841	0.313	0.371	0.407	2.234	2.879	5.040	-	-	-	-	-	-	-	-	-
4	0.373	0.435	0.474	1.922	2.342	3.609	0.347	0.407	0.446	2.027	2.493	3.889	-	-	-	-	-	-	-	-	-
5	0.405	0.470	0.510	1.794	2.123	3.049	0.369	0.431	0.470	1.920	2.302	3.369	-	-	0.233	7.550	-	-	1.958	-	-
6	0.429	0.494	0.535	1.715	1.992	2.737	0.386	0.448	0.488	1.855	2.189	3.081	-	0.240	0.287	3.969	11.243	-	1.335	2.456	-
7	0.446	0.513	0.554	1.663	1.906	2.539	0.397	0.461	0.501	1.810	2.115	2.898	-	0.284	0.327	3.075	5.568	-	1.090	1.763	-
8	0.460	0.527	0.569	1.625	1.846	2.405	0.407	0.471	0.511	1.778	2.061	2.773	0.229	0.315	0.356	2.671	4.141	40.239	0.955	1.473	3.900
9	0.471	0.539	0.581	1.597	1.801	2.308	0.415	0.480	0.520	1.753	2.021	2.682	0.268	0.339	0.380	2.443	3.496	12.060	0.870	1.307	2.697
10	0.480	0.549	0.591	1.575	1.766	2.235	0.422	0.486	0.527	1.733	1.988	2.612	0.291	0.358	0.399	2.297	3.131	7.829	0.811	1.200	2.266
Panel 1b: Simulated SEs of the critical values$\times 10^4$, $q \leq 10$																					
3	0.522	0.233	0.173	0.797	1.392	5.226	0.554	0.245	0.183	0.836	1.471	5.493	-	-	-	-	-	-	-	-	-
4	0.422	0.203	0.154	0.578	0.947	3.080	0.460	0.220	0.168	0.635	1.038	3.346	-	-	-	-	-	-	-	-	-
5	0.362	0.180	0.140	0.474	0.738	2.148	0.415	0.207	0.161	0.540	0.850	2.484	-	-	4.706	16.048	-	-	2.341	-	-
6	0.322	0.168	0.133	0.412	0.624	1.699	0.385	0.198	0.155	0.486	0.739	1.985	-	9.226	1.274	3.865	34.903	-	1.119	3.360	-
7	0.300	0.160	0.126	0.371	0.552	1.426	0.365	0.191	0.152	0.448	0.665	1.736	-	2.230	0.704	2.082	7.764	-	0.805	1.580	-
8	0.279	0.152	0.121	0.344	0.502	1.248	0.351	0.187	0.149	0.423	0.618	1.535	119.294	1.188	0.514	1.447	3.944	516.337	0.660	1.120	13.482
9	0.269	0.147	0.118	0.323	0.466	1.125	0.342	0.183	0.147	0.402	0.584	1.416	10.212	0.823	0.413	1.130	2.635	44.224	0.579	0.911	4.063
10	0.260	0.142	0.115	0.307	0.439	1.036	0.333	0.181	0.145	0.388	0.560	1.337	4.534	0.649	0.360	0.957	1.992	18.029	0.526	0.790	2.604
Panel 2: Polynomial approximation coefficients and goodness-of-fit, $11 \leq q \leq 200$																					
β_0	0.603	0.672	0.713	1.356	1.439	1.626	0.526	0.594	0.634	1.482	1.615	1.900	0.518	0.587	0.624	1.474	1.579	2.680	0.374	0.530	0.775
β_1	-0.259	-0.254	-0.279	0.292	0.568	0.792	-0.296	-0.302	-0.287	0.611	0.839	1.369	-0.587	-0.621	-0.460	1.916	4.207	-40.05	0.846	1.338	5.555
β_2	-0.408	-0.413	-0.197	1.876	1.062	3.456	-0.102	-0.121	-0.165	0.094	0.332	1.985	0.591	0.578	-1.192	-12.575	-38.583	1135	-0.834	-3.703	-81.99
β_3			-0.453	-5.686	-1.526	-9.192				1.539	2.362	-2.147	-3.129	-2.863	4.359	113.153	322.345	-16493	15.15	43.57	969.1
β_4				11.967	9.070	30.250						14.71			-9.976	-403.905	-1182.782	146983	-59.18	-175.4	-6351
β_5																611.463	1804.714	-814505	115.4	313.8	24043
β_6																		2743590			-48988
β_7																		-5140428			43483
β_8																		4136140			
R^2	0.999	0.999	0.999	1.000	1.000	1.000	0.999	0.999	0.999	1.000	1.000	1.000	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Reg SE$\times 10^3$	0.562	0.457	0.419	0.414	0.468	0.549	0.587	0.546	0.529	0.531	0.555	0.616	0.718	0.635	0.528	0.547	0.637	0.862	0.089	0.123	0.220
Panel 3: Polynomial approximation coefficients and goodness-of-fit, $q \geq 200$																					
β_0	0.604	0.673	0.713	1.352	1.439	1.623	0.527	0.595	0.635	1.483	1.614	1.898	0.516	0.584	0.626	1.496	1.644	1.974	0.375	0.535	0.847
β_1	-0.262	-0.272	-0.268	0.410	0.566	0.931	-0.314	-0.316	-0.313	0.573	0.833	1.454	-0.523	-0.552	-0.535	1.004	1.559	3.067	0.781	1.085	1.708
β_2	-0.600	-0.360	-0.278	0.678	0.946	1.648				0.444	0.532	1.201			-0.283	2.174	3.452	6.473	0.662	1.105	2.099
R^2	0.924	0.951	0.952	0.980	0.988	0.993	0.907	0.939	0.942	0.983	0.991	0.995	0.959	0.977	0.980	0.995	0.998	0.999	0.999	1.000	0.999
Reg SE$\times 10^3$	1.130	0.876	0.840	0.848	0.892	1.173	1.308	1.042	1.009	1.028	1.070	1.441	1.409	1.108	1.031	1.057	1.154	1.569	0.243	0.316	0.579

Table C.2: Simulated and approximated critical values for the HDIs of the discrete OMK, OK, and MAED estimators and the right-tailed test of $S_t^{(q)}$. The simulated standard errors in Panel 1b are computed based on the asymptotic standard error of order statistics. For the polynomial approximation results in Panels 2 and 3, we present the coefficients of the polynomial, $\sum_{i=0}^{k} \beta_i q^{-i/2}$, estimated based on linear regression from the simulated critical values for the range of q considered. The regression R^2 and the regression standard errors (reg SE) are presented under the coefficients.