A ReMeDI for Microstructure Noise *

Z. Merrick Li† Oliver Linton‡

September 8, 2020

Abstract

We introduce the Realized moMents of Disjoint Increments (ReMeDI) paradigm to measure microstructure noise (the deviation of the observed asset prices from the fundamental values caused by market imperfections). We propose consistent estimators of arbitrary moments of the microstructure noise process based on high-frequency data, where the noise process could be serially dependent, endogenous, and nonstationary. We characterize the limit distributions of the proposed estimators and construct confidence intervals under infill asymptotics. Our simulation and empirical studies show that the ReMeDI approach is very effective to measure the scale and the serial dependence of microstructure noise. Moreover, the estimators are quite robust to model specifications, sample sizes and data frequencies.

KEYWORDS: Microstructure noise, semimartingale, serial dependence, stable convergence, mixing sequence, infill asymptotics, finite sample bias

*Part of the paper was circulated under various titles, including, "A simple measure of microstructure noise"and "Robust measures of microstructure noise". We benefit from discussions with Yacine Aït-Sahalia, Torben Andersen, Peter Boswijk, Rob Engle, Christian Gouriéroux, Peter Reinhard Hansen, Jean Jacod, Ilze Kalnina, Frank Kleibergen, Eben Lazarus, Yingying Li, Albert Menkveld, Per Mykland, George Tauchen, Michel Vellekoop, Bas Werker, Dacheng Xiu, Xiye Yang, Xinghua Zheng, and seminar participants at the 2017 North American Summer Meeting (St. Louis, June 2017), 2017 European Meeting (Lisbon, August 2017) and 2018 China Meeting (Shanghai, June 2018) of the Econometric Society, International Conference on Quantitative Finance and Financial Econometrics (Marseille, June 2018), the 12th Annual SoFiE Conference (Shanghai, June 2019), the Econometric Society and Bocconi University Virtual World Congress (August, 2020), Hong Kong University of Science and Technology, NYU Shanghai, University of Amsterdam and University of Cambridge. The project is partially sponsored by the Keynes Fund (JHUL).

†Faculty of Economics, University of Cambridge, Austin Robinson Building, Sidgwick Avenue, Cambridge, CB3 9DD, United Kingdom. Email: z.merrick.li@gmail.com

‡Faculty of Economics, University of Cambridge, Austin Robinson Building, Sidgwick Avenue, Cambridge, CB3 9DD, United Kingdom. Email: obl20@cam.ac.uk.

1 Introduction

Economic time series are often modelled as the sum of a latent process obtained from an underlying economic model and another term that reflects a variety of adjustments to or departures from the frictionless theoretical model, thus

$$\underbrace{Y}_{\text{observed series}} = \underbrace{X}_{\text{underlying process}} + \underbrace{\varepsilon}_{\text{deviation}} . \tag{1}$$

The two processes X and ε are generated by different mechanisms, and can have quite distinct statistical properties and economic interpretations. Both quantities may be of interest as they give interpretation of some underlying economic theory and its relevance for the observed data. However, since only the sum process Y is observable, this makes estimation and inference challenging.

We are concerned with applications of this framework in financial markets where the observed asset price[1] (Y) subsumes both *market microstructure noise* (ε) and the *efficient price* (or *fundamental value)* (X). The fundamental theorem of asset pricing says that X should be a semimartingale process (Delbaen and Schachermayer (1994)). In practice however, many market frictions, such as: transaction costs, price discreteness, inventory holdings, information asymmetry, measurement errors, may cause observed prices to deviate from this ideal price. One may also want to allow for temporary mispricing (French and Roll (1986)) or fad effects (Lehmann (1990)); see also O'Hara (1995), Hasbrouck (2007) and Foucault et al. (2013) for insightful reviews. A lot of early work proceeded on the basis that the microstructure noise process was i.i.d., but recently this assumption has been shown to be too strong; both theoretically and empirically the microstructure noise may exhibit rich dynamics depending on its origin. If the microstructure effects are negligible, the observed price should be close to the efficient price and be unpredictable. Therefore, the dispersion and persistence of the microstructure noise serve as natural measures of market quality. Market quality is of concern to regulators and practitioners as well as academics; proxies for market quality are widely used in empirical analysis, see Linton and Mahmoodzadeh (2018). In the working paper Li and Linton (2019), we proposed two market liquidity measures, called IBAS and ABAS, that were defined in terms of the autocovariance function of the noise process. Such liquidity measures are robust to the pattern of order flows, and have an intuitive economic interpretation.

We introduce a general econometric approach to measuring microstructure noise in a nonparametric setting. Specifically, we propose a new estimator of the moments

[1] By *price* it always means the *logarithmic price* in this paper unless stated otherwise.

of a general dependent noise process based on the observed noisy high-frequency transaction prices; we call our estimator the *Realized moMents of Disjoint Increments* (ReMeDI). The estimation method is based on the differencing paradigm, which is widely used in microeconometrics to eliminate nuisance parameters, see, e.g., Athey and Imbens (2006). We build on the general setup introduced in the seminal work of Jacod et al. (2017). Specifically, we assume that the underlying efficient price follows a semimartingale, which may accommodate stochastic volatility, jumps, etc. We allow the microstructure noise to be weakly dependent and to have a serial correlation of an unknown form that may decay at an algebraic rate; this may capture, for instance, the effects of clustered (or hidden) order flows or herding (Park and Sabourian (2011)). The microstructure noise is allowed to have time-varying and stochastic heteroskedasticity, which allows for intraday variation in the scale of the noise. The general setting we consider allows for random and endogenous observation schemes. We develop estimators of arbitrary moments of the microstructure noise; this includes the autocovariance function of powers of the noise process as well as other quantities of interest. We derive the stable convergence in law of the estimated quantities as the sample size increases on a given domain. We provide a consistent estimator of the asymptotic variance that allows us to quantify the accuracy of our estimator.

We present some simulation studies comparing the ReMeDI approach with the method of Jacod et al. (2017). We find that the ReMeDI approach is relatively robust to: the data frequency, the sample size, the tuning parameter, and to model specification. We provide an empirical study on an individual stock price, which reveals that the microstructure noise has non-trivial serial dependence, but that the dependence structure falls short of being long memory. This is consistent with leading microstructure models,[2] and differs from the findings in Jacod et al. (2017).

The robustness of the ReMeDI approach as demonstrated in our simulation and empirical studies has an intuitive explanation. The differencing method works because the increments of X over disjoint intervals (the efficient returns) are small and/or uncorrelated, and what remains is attributed to ε. This property distinguishes the ReMeDI approach from alternative high-frequency estimators that rely structurally on the *infill asymptotics*.

1.1 Related literature

There are a number of methods for estimation of the moments of noise and the parameters of the efficient price. Specifically: the two-scale/multi-scale realized volatility

[2]For example, Hasbrouck and Ho (1987), Choi et al. (1988) and Huang and Stoll (1997) model the probability of order reversal, and microstructure noise becomes an AR(1) process.

by Zhang et al. (2005), Zhang (2006), Aït-Sahalia et al. (2011); the optimal-sampling realized variance by Bandi and Russell (2006, 2008); the maximum likelihood estimators by Aït-Sahalia et al. (2005), Xiu (2010); the pre-averaging method developed in Podolskij and Vetter (2009), Jacod et al. (2009); and the realized kernel by Hansen and Lunde (2006), Barndorff-Nielsen et al. (2008). Most of this literature only considers i.i.d. microstructure noise.

Several recent papers explore richer microstructure models by allowing for autocorrelated noise. The estimators of the second moments of noise in Da and Xiu (2019) and Li et al. (2020) are by-products of the integrated volatility estimators in the presence of autocorrelated noise. In a recent seminal paper, Jacod et al. (2017) introduced the first feasible procedure, called the *local averaging* (LA) method, to estimate arbitrary moments of microstructure noise using high-frequency data. They also introduced a general framework allowing for a stochastic observation scheme and a microstructure noise with a semimartingale "size process". We follow their general setup and derive asymptotic properties of our estimators under this general framework. We differentiate our paper from Jacod et al. (2017) as follows. First, the ReMeDI method is based on *differencing*; while the LA method is based on deviations from *local averages*, both ideas are widely used in other contexts such as panel data and semiparametric estimation to eliminate nuisance parameters, see Yatchew (1997) and Athey and Imbens (2006). Second, the ReMeDI approach works beyond the infill framework. Specifically, in the working paper version, Li and Linton (2019), we proved that the ReMeDI estimator is consistent and has an associated CLT in a long-span, non-infill setting. In this case, the method works provided the efficient price is a martingale in which case its increments are uncorrelated at any horizon. The LA method, however, is inconsistent when applied to low-frequency data. Next, the finite sample performance of the LA estimators heavily depends on the sample size and the noise-to-signal ratio (the ratio of noise variance to the integrated volatility of the efficient price), see an analysis in Jacod et al. (2017). This may cause many issues in the implementations with real data.[3] The bias of the ReMeDI estimators by contrast only depends on the slope of the microstructure autocovariance function, and in short memory contexts this bias can be very small. Last, the ReMeDI approach has another two advantages in real implementations: it is computationally very efficient,[4] and it is very robust to a wide range of tuning parameters.

[3]One can easily verify the following scenarios by simulation: (1) the LA estimator may report positive autocovariances when the true noise process is uncorrelated or even negatively autocorrelated; (2) the LA estimator has larger bias and variance if there are bursts of volatility in the efficient price process, e.g., when the volatility process jumps; (3) the LA estimator gives very different estimates over two samples where the noise processes are identical but the efficient prices have different variances.

[4]For example, the LA (ReMeDI) takes 99.77% (0.23%) of the CPU time to estimate the variance of the noise using noisy price data from a random walk plus AR(1) noise model, based on 1,000 simulated

Our paper introduces an econometric approach to richer microstructure models. It aims to integrate the market microstructure and financial econometrics literature. It is, however, not the first attempt to push towards the integration of the two fields. Diebold and Strasser (2013) focus on the correlation of efficient price and noise in several leading microstructure models, and study the implications for integrated volatility estimation. Li et al. (2016), Chaker (2017) and Clinet and Potiron (2017) model the microstructure noise as a parametric function of the observable trading information and develop efficient volatility estimators. Bandi et al. (2017) develop a novel measure of the staleness of stock returns under the infill asymptotic framework. Bollerslev et al. (2018) study the relationship between trading volume and return volatility around important public news announcements using intraday high-frequency data. The study relies critically on high-frequency econometric techniques to identify jumps. Da and Xiu (2019) advocate the quasi-maximum likelihood approach to estimate both the volatility and the autocovariances of moving-average microstructure noise.

2 Continuous-Time Framework and Assumptions

We follow the general framework of Jacod et al. (2017) to specify the continuous-time efficient price process, the observation scheme, and the microstructure noise. We have almost the same regularity conditions as Jacod et al. (2017).[5] Hence, for brevity of exposition, we put some details of the specifications in Appendix A.

2.1 Efficient price process

We assume that the efficient price process X is an Itô semimartingale defined on a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$ with the Grigelionis representation

$$X_t = X_0 + \int_0^t b_s \mathrm{d}s + \int_0^t \sigma_s \mathrm{d}W_s + \left(\vartheta \mathbf{1}_{\{|\vartheta|\leq 1\}}\right) \star (\mathfrak{p} - \mathfrak{q})_t + \left(\vartheta \mathbf{1}_{\{|\vartheta|>1\}}\right) \star \mathfrak{p}_t, \qquad (2)$$

where $W, \mathfrak{p}$ are a Wiener process and a Poisson random measure on $\mathbb{R}_+$and E respectively. Here, $(E, \mathcal{E})$ is a measurable Polish space on $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$ and the predictable compensator of $\mathfrak{p}$ is $\mathfrak{q}(\mathrm{d}s, \mathrm{d}z) = \mathrm{d}s \otimes \lambda(\mathrm{d}z)$ for some given σ-finite measure on $(E, \mathcal{E})$, see Jacod and Shiryaev (2003) for detailed introduction of the last two integrals. Further regularity conditions on X are discussed in Assumption (H) in Appendix A. Note that the setting of the efficient price is very general, and it allows for

samples of size 23,400.

[5] The only difference is that we have more restriction on the parameter that controls the degree of serial dependence of the stationary noise, see Remark 2.1.

stochastic volatility and jumps in both the price and volatility processes.

2.2 Observation scheme

We assume a triangular array structure. For each n, let $\{t_i^n : i \in \mathbb{N}_+\}$ be a sequence of random finite observed times (usually when a transaction occurs) with $0 = t_0^n < t_1^n < \ldots$, where $\mathbb{N}_+$ is the set of nonnegative integers. We denote

$$n_t := \sum_{i \geq 0} \mathbf{1}_{\{t_i^n \leq t\}}, \quad \delta(n,i) := t_i^n - t_{i-1}^n,\ i \geq 1. \tag{3}$$

Here, n_t is the stochastic number of observations recorded on the interval $[0,t]$ for $t \in \mathbb{R}_+$, while $\delta(n,i)$ is the i^{th} spacing of the observation times. For any process V, we denote $V_i^n := V_{t_i^n}$.

Let $\{\delta_n\}_n$ be a positive sequence of real numbers satisfying $\delta_n \to 0$ as $n \to \infty$. We may think of δ_n as the average magnitude of the spacings between successive observation times. If the observation times were equally spaced (the regular observation scheme), then δ_n would be proportional to that spacing. The difference between the regular observation scheme and the general scheme is characterized by two semimartingales $\alpha, \overline{\alpha}$, which are approximately the conditional mean and variance of the time differences, and characterizing the density of the observations. Specifically, conditional upon an appropriate σ-algebra, the expectations of $\delta(n,i)/\delta_n$ and $(\delta(n,i)\alpha_{i-1}^n - \delta_n)^2/\delta_n^2$ are approximately equal to $1/\alpha_{i-1}^n$ and $\overline{\alpha}_{i-1}^n$, respectively. For brevity, we move the details of the specifications to Assumption (O) in Appendix A. A useful consequence of our setting is the following convergence in probability:

$$\delta_n n_t \xrightarrow{\mathbb{P}} A_t := \int_0^t \alpha_s \mathrm{d}s. \tag{4}$$

The observation times framework is very general, and includes, e.g., *regular sampling scheme, time-changed regular sampling scheme, modulated Poisson sampling scheme, and predictably-modulated random walk sampling scheme,* see the discussion in Jacod et al. (2017).

2.3 Microstructure noise

The microstructure noise has a multiplicative form that allows for serial dependence, stochastic scale and dependence of the scale on the efficient price process. At time t_i^n, the microstructure noise is given by

$$\varepsilon_i^n := \gamma_i^n \cdot \chi_i.$$

Here, γ is a nonnegative Itô semimartingale on $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$.[6] The process $\{\chi_i\}_{i\in\mathbb{Z}}$ is stationary and ρ-mixing, and its degree of serial dependence is controlled by a parameter v. Specifically, the autocovariance function of $\{\chi_i\}_i$ decays at a polynomial rate, i.e.,

$$|\mathbf{Cov}(\chi_i, \chi_{i+k})| \leq \frac{K}{k^v}, \tag{5}$$

where $K > 0$ is some positive constant. The reader is referred to Assumption (N) for the detailed specifications of γ and χ.

Remark 2.1. *To obtain limit results, we shall suppose that $v > 1$ for consistency and that $v > 2$ to derive the limit distribution, which allows for quite strong dependence close to the long memory boundary. Jacod et al. (2017) require $v > 0$ for consistency and $v > 1$ to establish the limit distribution.*

2.4 The observed noisy price

Finally, the observed noisy price Y_i^n is given by (for $i = 1, \ldots, n_t$)

$$Y_i^n = X_i^n + \varepsilon_i^n. \tag{6}$$

Note that both X and ε are latent, only Y is observable, and we aim to estimate the moments of ε using Y only.

3 The Design and the Intuition of the ReMeDI Estimators

3.1 The estimator of the autocovariance function

The intuition of the ReMeDI design can be best seen in a simpler setting. Let $\{\varepsilon_i\}_{i\in\mathbb{Z}}$ be a stationary mixing sequence with mean zero; we would like to estimate its covariance

[6] The semimartingale assumption for γ is adopted for comparison with Jacod et al. (2017), but other assumptions can be accommodated and might be more reasonable.

$r_\ell := \mathbf{Cov}(\varepsilon_i, \varepsilon_{i+\ell})$. The natural estimator is the sample analogue

$$\widehat{r}_\ell := \frac{1}{n} \sum_{i=0}^{n-\ell} \varepsilon_i \varepsilon_{i+\ell}, \tag{7}$$

which is consistent and asymptotically normal under very mild conditions.

We consider instead an estimator that replaces the "observations" $\varepsilon_i, \varepsilon_{i+\ell}$ by the "differences", i.e.,

$$\widetilde{r}_\ell^n := \frac{1}{n} \sum_{i=k_n'}^{n-\ell-k_n} \left(\varepsilon_i - \varepsilon_{i-k_n'}\right) \left(\varepsilon_{i+\ell} - \varepsilon_{i+\ell+k_n}\right),$$

where k_n, k_n' are integers that grow at certain rates as the sample size increases. The estimator $\widetilde{r}_\ell^n$ follows the ReMeDI design and it provides another consistent estimator of r_ℓ, provided $k_n \wedge k_n' \to \infty$, and $\frac{k_n \vee k_n'}{n} \to 0$. The intuition of the consistency becomes immediate if one rewrites $\widetilde{r}_\ell^n$ as

$$\widetilde{r}_\ell^n = \frac{1}{n} \sum_{i=k_n'}^{n-\ell-k_n} \varepsilon_i \varepsilon_{i+\ell} - \frac{1}{n} \sum_{i=k_n'}^{n-\ell-k_n} \varepsilon_i \varepsilon_{i+\ell+k_n} - \frac{1}{n} \sum_{i=k_n'}^{n-\ell-k_n} \varepsilon_{i-k_n'} \varepsilon_{i+\ell} + \frac{1}{n} \sum_{i=k_n'}^{n-\ell-k_n} \varepsilon_{i-k_n'} \varepsilon_{i+\ell+k_n}. \tag{8}$$

The first average is (asymptotically) equivalent to the sample analogue, thus it converges in probability to r_ℓ; the remaining three averages are centred at $r_{\ell+k_n}$, $r_{\ell+k_n'}$, and $r_{\ell+k_n+k_n'}$, which themselves converge to zero as $n \to \infty$ at a rate depending on (5).

Taking differences seems redundant if the time series $\{\varepsilon_i\}_i$ is observable. However, in our framework, ε is masked by the efficient price X, and we only observe $Y = X + \varepsilon$. Taking time differences removes the effect of the efficient price. The intuition of such removal under infill asymptotics is that the differences of the efficient prices, say, $X_i^n - X_{i-k_n'}^n$, are much smaller than the differences of the noise as n increases.

3.2 The general ReMeDI design

We next formally define our ReMeDI (Realized moMents of Disjoint Increments) estimator of a general class of parameters. First, we provide some notations that we will use below. Let $\mathfrak{J}$ be the set of all finite sequences of integers satisfying

$$\mathfrak{J} := \left\{ \mathbf{j} = (j_1, j_2, \ldots, j_q) : j_l \in \mathbb{Z}, l = 1, 2, \ldots, q; q \geq 2 \right\}.$$

In the sequel, we will assume without loss of generality that $j_1 = \max\{j_l : j_l \in \mathbf{j}\}$ for any $\mathbf{j} \in \mathfrak{J}$. The $\mathbf{j}$-moments of χ, the stationary component of microstructure noise, are given by

$$\boldsymbol{r}(\mathbf{j}) := \mathbb{E}\left(\prod_{l=1}^{q} \chi_{j_l}\right). \tag{9}$$

This is our parameter of interest (after proper scaling by the γ process); it includes the autocovariance function of the noise process and many other examples as special cases.

Let $\boldsymbol{k} = (k_1, \ldots, k_q)$ be a q-tuple of integers. For any $\mathbf{j} \in \mathfrak{J}$ and any process V, let $\mathbb{I}(\boldsymbol{k}, \mathbf{j})^n_t$ be the set of observation indices on $[0, t]$ for which the following *multi-difference operator* $\Delta^{\boldsymbol{k}}_{\mathbf{j}}(\cdot)^n_i$ is well defined[7]:

$$\Delta^{\boldsymbol{k}}_{\mathbf{j}}(V)^n_i := \prod_{l=1}^{q} \left(V^n_{i+j_l} - V^n_{i+j_l-k_l}\right). \tag{10}$$

Then the ReMeDI estimator corresponding to $\boldsymbol{r}(\mathbf{j})$ based on data $\{Y^n_i\}_{i=1}^{n_t}$ and tuning parameters $\boldsymbol{k}$ is defined by

$$\text{ReMeDI}(Y; \mathbf{j}, \boldsymbol{k})^n_t := \sum_{i \in \mathbb{I}(\boldsymbol{k}, \mathbf{j})^n_t} \Delta^{\boldsymbol{k}}_{\mathbf{j}}(Y)^n_i. \tag{11}$$

Remark 3.1. *Using the above notations, we rewrite the estimator $\widetilde{r}^n_\ell$ as follows*

$$\widetilde{r}^n_\ell = \frac{1}{n} \sum_{i=k'_n}^{n-\ell-k_n} \Delta^{-k_n, k'_n}_{0,\ell}(\varepsilon)^n_i.$$

The general ReMeDI *approach inherits two salient features of this estimator determined by the choices of $\boldsymbol{k}$: 1) the first entry of $\boldsymbol{k}$ will be negative whereas the remaining ones are positive, i.e., the first difference is a* forward difference *and the remaining ones are* backward differences*; 2) $\forall 1 \leq l \leq q$, $|k_l| \to \infty$ as $n \to \infty$, and we will often write $\boldsymbol{k}_n = (k_{1,n}, \ldots, k_{q,n})$ in the sequel to reflect such dependence.*

We discuss a little more why the general ReMeDI procedure works under infill asymptotics. For this purpose, suppose that the noise size process γ is constant and we are estimating $\mathbb{E}\left(\prod_{l=1}^{q} \varepsilon^n_{i+j_l}\right)$. Suppose that $\boldsymbol{k}_n$ satisfies the two properties in Remark 3.1. Next, we explain how to connect $\mathbb{E}\left(\prod_{l=1}^{q} \varepsilon^n_{i+j_l}\right)$ and $\Delta^{\boldsymbol{k}_n}_{\mathbf{j}}(Y)^n_i$ with $\Delta^{\boldsymbol{k}_n}_{\mathbf{j}}(\varepsilon)^n_i$. To see this, we first note that $\{i + j_l - k_{l,n}\}_l$ are the "distant" indices of the intervals on which the backward and forward differences are taken. Figure 1 illustrates a simple example with $\mathbf{j} = (j_1, j_2, j_3)$, $\boldsymbol{k}_n = (-k_n, 2k_n, 4k_n)$ for some $k_n \in \mathbb{N}_+$. The forward difference starts at the $(i + j_1)$-th observation and ends at the $(i + j_1 + k_n)$-th observation; for

[7] By convention we set $\Delta^{\boldsymbol{k}}_{\mathbf{j}}(V)^n_i = 1$ if $\mathbf{j} = \emptyset$ and $\Delta^{\boldsymbol{k}}_{\mathbf{j}}(V)^n_i = 0$ if $\mathbf{j}$ is a singleton.

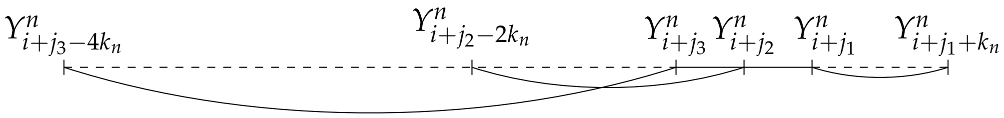

Figure 1: Illustration the ReMeDI estimator of $\mathbf{j}$-moments with $\mathbf{j} = (j_1, j_2, j_3)$ and $\boldsymbol{k}_n = (-k_n, 2k_n, 4k_n)$.

the remaining indices in $\mathbf{j}$, the associated differences start from $i + j_2, i + j_3$ and end at $i + j_2 - 2k_n, i + j_3 - 4k_n$, respectively. The intuition of the ReMeDI approach is that the "distant" noise terms are approximately independent of each other, and are also independent of the "clustered" noise $\{\varepsilon^n_{i+j_l}\}_l$ (recall a special case outlined in (8)), therefore

$$\mathbb{E}\left(\Delta^{\boldsymbol{k}_n}_{\mathbf{j}}(\varepsilon)^n_i\right) \approx \mathbb{E}\left(\prod_{l=1}^q \varepsilon^n_{i+j_l}\right).$$

If $k_{l,n}$ is still relatively small such that $\sup_l \delta_n |k_{l,n}| \to 0$, the differences/increments of the efficient price over the intervals are asymptotically negligible. That is, $\Delta^{\boldsymbol{k}_n}_{\mathbf{j}}(Y)^n_i \approx \Delta^{\boldsymbol{k}_n}_{\mathbf{j}}(\varepsilon)^n_i$. Thus the averages of $\Delta^{\boldsymbol{k}_n}_{\mathbf{j}}(Y)^n_i$ will converge in probability to $\mathbb{E}\left(\prod_{l=1}^q \varepsilon^n_{i+j_l}\right)$ by the law of large numbers. This is the intuition of the identification.

4 The Asymptotic Properties of the ReMeDI Estimators

4.1 Consistency

We next give the large sample properties of the ReMeDI estimator (for a given choice of $\boldsymbol{k}_n$) in our general setting. For a general γ process that satisfies Assumption (N), the "average size" of the noise moments $\mathbb{E}\left(\prod_{l=1}^q \varepsilon^n_{i+j_l}\right)$ is $\int_0^t \gamma_s^q \mathrm{d}A_s / A_t$, and this scaling appears in the probability limit of the ReMeDI estimators. Also recall (5) that v is the parameter that controls the degree of serial dependence in the noise.

Theorem 4.1. *Let Assumptions (H, O, N) hold, assume $v > 1$ and $\boldsymbol{k}_n$ satisfies*

$$\begin{cases} -k_{1,n} \to \infty, k_{l,n} \to \infty, \ \forall\, l \geq 2, \\ \sup_l \left|\delta_n^{\eta} k_{l,n}\right| \to 0, \ \eta \in (0, 1/2), \ \forall\, l \geq 1, \\ k_{l+1,n} - k_{l,n} \to \infty, \forall\, l \geq 2, \end{cases} \tag{12}$$

as $n \to \infty$. For $\mathbf{j} \in \mathfrak{J}$, we have the following convergence in probability:

$$\frac{\mathrm{ReMeDI}(Y; \mathbf{j}, \boldsymbol{k}_n)^n_t}{n_t} \xrightarrow{\mathbb{P}} \mathbf{R}(\mathbf{j})_t := \frac{\int_0^t \gamma_s^q \mathrm{d}A_s}{A_t} \boldsymbol{r}(\mathbf{j}), \tag{13}$$

where $\boldsymbol{r}(\mathbf{j})$ is defined in (9) and A_t in (4).

This says that our estimator consistently estimates $\boldsymbol{r}(\mathbf{j})$ up to a time t-varying scaling factor that depends on the average scale of the noise and on the stochastic process governing observation times.

Let $\{k_n\}_n$ be a sequence of integers satisfying $k_n \to \infty$, $k_n\delta_n \to 0$. Let $\boldsymbol{k}_n$ be specified as follows: $k_{l,n} = -k_n$ if $l = 1$, and $k_{l,n} = (l-1)k_n$ if $l \geq 2$. Then, $\boldsymbol{k}_n$ satisfies the conditions in (12).

4.2 Limit distribution

We first restrict further the values of $\boldsymbol{k}_n$ in order to facilitate the limit theory.[8] Among many possibilities, we propose the following specification of $\boldsymbol{k}_n$, which is solely determined by a single integer k_n:

$$k_{l,n} = \begin{cases} -k_n & \text{if } l = 1, \\ 2^{l-1}k_n & \text{if } l \geq 2, \end{cases} \tag{14}$$

where k_n is related to v as follows:

$$v > 2, \quad k_n\delta_n^{\eta} \to 0, \quad \eta \in \left(\frac{1}{2v}, \frac{1}{3}\right).$$

Remark 4.1. *Note that* (14) *implies* (12). *In the sequel, we will omit* $\boldsymbol{k}_n$ *and simply write* $\Delta_{\mathbf{j}}(Y)^n_i$ *and* $\mathrm{ReMeDI}(Y;\mathbf{j})^n_t$ *instead of* $\Delta^{\boldsymbol{k}_n}_{\mathbf{j}}(Y)^n_i$ *and* $\mathrm{ReMeDI}(Y;\mathbf{j},\boldsymbol{k}_n)^n_t$ *when* $\boldsymbol{k}_n$ *satisfies* (14).

We establish the CLT for both the following centered stochastic processes:

$$Z(\mathbf{j})^n_t := \frac{1}{\sqrt{\delta_n}}\left(\delta_n \mathrm{ReMeDI}(Y;\mathbf{j})^n_t - \boldsymbol{r}(\mathbf{j})\int_0^t \gamma^q_s \mathrm{d}A_s\right); \; \overline{Z}(\mathbf{j})^n_t := \sqrt{n_t}\left(\frac{\mathrm{ReMeDI}(Y;\mathbf{j})^n_t}{n_t} - \mathbf{R}(\mathbf{j})_t\right).$$

The first process involves unknown but deterministic norming, whereas the second process is normed by the observed stochastic sample size. Thus the second one is "feasible" in practice. Now let $\mathcal{F}_\infty := \bigvee_{t>0}\mathcal{F}_t$.

Theorem 4.2. *Let Assumptions (H), (O) and (N) hold, and* k_n, v *satisfy* (14). *For any* $t > 0, \mathbf{j}, \mathbf{j}' \in \mathfrak{J}$, *we have the following* $\mathcal{F}_\infty$*-stable convergence in law*

(a) $(Z(\mathbf{j})^n_t, Z(\mathbf{j}')^n_t) \xrightarrow{\mathcal{L}_s - \mathcal{F}_\infty} (Z(\mathbf{j})_t, Z(\mathbf{j}')_t)$ *, where the limit is defined on an extension* $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathbb{P}})$ *of* $(\Omega, \mathcal{F}, \mathbb{P})$. *Conditionally on* $\mathcal{F}$, $(Z(\mathbf{j})_t, Z(\mathbf{j}')_t)$ *are centred Gaussian with (co)variances* $\sigma(\mathbf{j},\mathbf{j}')_t$ *that is given by*

$$\sigma(\mathbf{j},\mathbf{j}')_t := \boldsymbol{s}(\mathbf{j},\mathbf{j}')\int_0^t \gamma^{q+q'}_s \mathrm{d}A_s + \boldsymbol{r}(\mathbf{j})\boldsymbol{r}(\mathbf{j}')\int_0^t \gamma^{q+q'}_s \overline{\alpha}_s \mathrm{d}A_s, \tag{15}$$

[8] In the supplementary material Li and Linton (2020), we discuss how to select k_n in practice.

where $\mathbf{s}(\mathbf{j},\mathbf{j}')$ *is given by* (B.6)*.*

(b) $\left(\overline{Z}(\mathbf{j})^n_t, \overline{Z}(\mathbf{j}')^n_t\right) \xrightarrow{\mathcal{L}_s-\mathcal{F}_\infty} \left(\overline{Z}(\mathbf{j})_t, \overline{Z}(\mathbf{j}')_t\right)$ *, where the limit is defined on an extension* $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathbb{P}})$ *of* $(\Omega, \mathcal{F}, \mathbb{P})$*. Conditionally on* $\mathcal{F}$*,* $\left(\overline{Z}(\mathbf{j})_t, \overline{Z}(\mathbf{j}')_t\right)$ *are centred Gaussian with (co)variances* $\overline{\sigma}(\mathbf{j},\mathbf{j}')_t$ *that is given by*

$$\begin{aligned}\overline{\sigma}(\mathbf{j},\mathbf{j}')_t := &\frac{\mathbf{s}(\mathbf{j},\mathbf{j}')}{A_t}\int_0^t \gamma_s^{q+q'}\mathrm{d}A_s + \frac{\boldsymbol{r}(\mathbf{j})\boldsymbol{r}(\mathbf{j}')}{A_t}\int_0^t \gamma_s^{q+q'}\overline{\alpha}_s\mathrm{d}A_s + \frac{\mathbf{R}(\mathbf{j})_t\mathbf{R}(\mathbf{j}')_t}{A_t}\int_0^t \overline{\alpha}_s\mathrm{d}A_s \\ &- \frac{\mathbf{R}(\mathbf{j})_t\boldsymbol{r}(\mathbf{j}')}{A_t}\int_0^t \gamma_s^{q'}\overline{\alpha}_s\mathrm{d}A_s - \frac{\boldsymbol{r}(\mathbf{j})\mathbf{R}(\mathbf{j}')_t}{A_t}\int_0^t \gamma_s^{q}\overline{\alpha}_s\mathrm{d}A_s.\end{aligned} \tag{16}$$

Remark 4.2. $\mathbf{s}(\mathbf{j},\mathbf{j}')$ *is the asymptotic variance of the ReMeDI estimators contributed by the stationary part of the noise. The explicit form is given by* (B.6) *in Appendix B.1. When the observation scheme is simpler, e.g., when* $\overline{\alpha}_t \equiv 0$ *or* 1*, the asymptotic variance (as well as its consistent estimator) are much simplified, see the discussion in Appendix B.2.*

Remark 4.3 (Asymptotic variances of ReMeDI and LA)**.** *Note that the ReMeDI and LA estimators have very similar asymptotic (co)variances. The only difference lies in the* $\mathbf{s}(\mathbf{j},\mathbf{j}')$ *part, which represents the asymptotic variance contributed by the stationary part of noise. The* $\mathbf{s}(\mathbf{j},\mathbf{j}')$ *of the ReMeDI estimators includes the asymptotic (co)variances of the "distant" noise terms (recall the discussion in Section 3.2). It is therefore larger than the counterpart of the LA estimators. Hence, the LA estimators are asymptotically more efficient (although one can improve the efficiency of ReMeDI by taking averages of estimators computed using different* $\boldsymbol{k}_n$*). However, simulation studies show that the ReMeDI class works better in finite samples with realistic sample sizes (or equivalently, data frequency) — it has smaller finite sample variance and is almost unbiased under various model specifications. Moreover, the ReMeDI approach has greater computational efficiency, which pays off when one is working with massive high-frequency datasets (recall Footnote 4).*

Theorem 4.3. *Suppose that all the conditions of Theorem 4.2 hold. Moreover,* $\{i_n\}_n, \{\phi_n\}_n$ *are two sequences of integers satisfying:*

$$\frac{i_n}{k_n^v} \to 0, \quad i_n\delta_n^\eta \to 0, \quad \frac{\phi_n}{k_n\delta_n} \to 0, \quad \frac{k_n^{3/4}\delta_n}{\phi_n} \to 0. \tag{17}$$

For any $\mathbf{j} \in \mathfrak{J}$*, we have the following* $\mathcal{F}_\infty$*-stable convergence in law*

$$\frac{\sqrt{n_t}}{\sqrt{\widehat{\sigma}(\mathbf{j},\mathbf{j}')^n_t}}\left(\frac{\mathrm{ReMeDI}(Y;\mathbf{j})^n_t}{n_t} - \mathbf{R}(\mathbf{j})_t\right) \xrightarrow{\mathcal{L}_s-\mathcal{F}_\infty} \Phi, \tag{18}$$

where Φ is a standard normal random variable that is defined on an extension of the space and is independent of $\mathcal{F}$, and $\widehat{\sigma}(\mathbf{j},\mathbf{j}')_t^n$ is a consistent estimator of the asymptotic variance constructed in (B.7).

4.3 Estimating the autocovariances of microstructure noise

In this section we consider the special case concerning the estimation of the autocovariance function of the microstructure noise. Let $\mathbf{j}_\ell = (0,\ell), \ell \in \mathbb{N}_+$, and let

$$\widehat{R}_{t,\ell}^n := \frac{1}{n_t}\mathrm{ReMeDI}(Y;\mathbf{j}_\ell)_t^n = \frac{1}{n_t}\sum_{i=2k_n}^{n_t-k_n-\ell}\left(Y_{i+\ell}^n - Y_{i+\ell+k_n}^n\right)\left(Y_i^n - Y_{i-2k_n}^n\right). \tag{19}$$

The following corollary provides the limit distribution.

Corollary 4.1 (ReMeDI estimators of autocovariances)**.** *Under the conditions of Theorem 4.2, we have*

$$\sqrt{n_t}\left(\widehat{R}_{t,\ell}^n - R_{t,\ell}\right) \stackrel{\mathcal{L}_s-\mathcal{F}_\infty}{\longrightarrow} \mathcal{N}\left(0,\overline{\sigma}(\mathbf{j}_\ell,\mathbf{j}_\ell)_t\right),$$

where

$$\overline{\sigma}(\mathbf{j}_\ell,\mathbf{j}_\ell)_t := \frac{1}{A_t}\left(\boldsymbol{s}_\ell\int_0^t\gamma_s^4\mathrm{d}A_s + r_\ell^2\int_0^t\gamma_s^4\overline{\alpha}_s\mathrm{d}A_s + R_{t,\ell}^2\int_0^t\overline{\alpha}_s\mathrm{d}A_s - 2R_{t,\ell}r_\ell\int_0^t\gamma_s^2\overline{\alpha}_s\mathrm{d}A_s\right); \tag{20}$$

$$R_{t,\ell} := r_\ell\frac{\int_0^t\gamma_s^2\mathrm{d}A_s}{A_t}, \quad \boldsymbol{s}_\ell := \sum_{k=-\infty}^{\infty}\left(\mathbb{E}((\chi_0\chi_\ell - r_\ell)(\chi_k\chi_{k+\ell} - r_\ell)) + 3r_k^2\right).$$

Moreover, under the assumptions of Theorem 4.3, we have

$$\frac{\sqrt{n_t}}{\sqrt{\widehat{\sigma}(\mathbf{j}_\ell,\mathbf{j}_\ell)_t^n}}\left(\widehat{R}_{t,\ell}^n - R_{t,\ell}\right) \stackrel{\mathcal{L}_s-\mathcal{F}_\infty}{\longrightarrow} \Phi, \tag{21}$$

where Φ is a standard normal random variables as in Theorem 4.3 and $\widehat{\sigma}(\mathbf{j}_\ell,\mathbf{j}_\ell)_t^n$ is provided in (B.7).

Remark 4.4. *$\boldsymbol{s}_\ell$ represents the variance of the ReMeDI estimators contributed by the stationary part of noise. It has two components. The first part $\sum_{k=-\infty}^{\infty}\mathbb{E}((\chi_0\chi_\ell - r_\ell)(\chi_k\chi_{k+\ell} - r_\ell))$ is in fact the asymptotic variance of the sample analogue, recall* (7)*. The second part $3\sum_{k=-\infty}^{\infty}r_k^2$ is the asymptotic variance of the three additional terms appear in* (8) *that arise in differencing.*

Remark 4.5. *The last three terms of $\overline{\sigma}(\mathbf{j}_\ell,\mathbf{j}_\ell)_t$ that appear in* (20) *arise because of the stochastic sampling scheme; it is positive and is zero whenever $r_\ell = 0$, $\overline{\alpha}_s \equiv 0$ or $\gamma_s \equiv K$, where K is a constant.*

We note that while the multiplicative structure of microstructure noise (recall (A.3)) allows for a time-varying and stochastic size of the noise, the serial correlation of the noise is not affected by the size process. This structure allows us to estimate the autocorrelations of noise directly once we have an estimator of the autocovariances. Define the ReMeDI estimator of the noise autocorrelation, $\widehat{r}(\ell)_t^n := \widehat{R}_{t,\ell}^n / \widehat{R}_{t,0}^n$, and its asymptotic variance estimator

$$\widehat{\sigma}(\ell)_t^n := \frac{\widehat{\sigma}(\mathbf{j}_\ell, \mathbf{j}_\ell)_t^n n_t^2}{(\mathrm{ReMeDI}(Y, \mathbf{j}_0)_t^n)^2} - \frac{2n_t^2 \widehat{\sigma}(\mathbf{j}_0, \mathbf{j}_\ell)_t^n \mathrm{ReMeDI}(Y, \mathbf{j}_\ell)_t^n}{(\mathrm{ReMeDI}(Y, \mathbf{j}_0)_t^n)^3} + \frac{n_t^2 (\mathrm{ReMeDI}(Y, \mathbf{j}_\ell)_t^n)^2 \widehat{\sigma}(\mathbf{j}_0, \mathbf{j}_0)_t^n}{(\mathrm{ReMeDI}(Y, \mathbf{j}_0)_t^n)^4}.$$

The following corollary spells out the limit distribution of the proposed estimators.

Corollary 4.2 (ReMeDI estimators of autocorrelations). *Under the conditions of Theorem 4.3, we have the following $\mathcal{F}_\infty$-stable convergence in law*

$$\sqrt{\frac{n_t}{\widehat{\sigma}(\ell)_t^n}} \left(\widehat{r}(\ell)_t^n - r(\ell)\right) \xrightarrow{\mathcal{L}_s - \mathcal{F}_\infty} \Phi,$$

where Φ is a standard normal random variable as in Theorem 4.3.

5 Simulation Study

5.1 Model settings

We suppose that the efficient price process has stochastic volatility and jumps that appear in both the price and volatility processes:

$$\begin{aligned} &\mathrm{d}X_t = \kappa_1(\mu_1 - X_t)\mathrm{d}t + \sigma_t \mathrm{d}W_{1,t} + \xi_{1,t}\mathrm{d}N_t; \quad \mathrm{d}\sigma_t^2 = \kappa_2(\mu_2 - \sigma_t^2)\mathrm{d}t + \eta\sigma_t \mathrm{d}W_{2,t} + \xi_{2,t}\mathrm{d}N_t; \\ &\mathbf{Corr}(W_1, W_2) = v; \quad \xi_{1,t} \sim \mathcal{N}(0, \mu_2/10); \quad N_t \sim \mathrm{Poi}(\lambda); \quad \xi_{2,t} \sim \mathrm{Exp}(\delta). \end{aligned} \tag{22}$$

We set

$$\kappa_1 = 0.5;\ \mu_1 = 3.6;\ \kappa_2 = 5/252;\ \mu_2 = 0.04/252;\ \eta = 0.05/252;\ v = -0.5;\ \lambda = 1;\ \delta = \eta.$$

This setting is motivated by some empirical facts that jumps in price levels and volatility tend to occur together, see Todorov and Tauchen (2011).

We further suppose that the stationary component of the microstructure noise follows an AR(1) process with Gaussian innovations

$$\chi_{i+1} = \rho\chi_i + e_i, \quad e_i \overset{\text{i.i.d.}}{\sim} \mathcal{N}\left(0, 1 - \rho^2\right), \quad |\rho| < 1.$$

Note that χ has unit variance. We set $\rho = 0.7$, motivated by the empirical studies in Aït-Sahalia et al. (2011) and Li et al. (2020).

5.2 LA versus ReMeDI

We estimate the autocovariances of microstructure noise using the ReMeDI estimator and the local averaging (LA) estimators (Jacod et al., 2017). We assume that the noise is stationary so that we can compare the estimates to the true parameters. We also assume that the observation scheme is regular so that we know explicitly the data frequency, which is a key factor that affects the finite sample performance of many high-frequency estimators.

The top and middle panels of Figure 2 present the estimation of the first 20 autocovariances of the noise by ReMeDI and LA.[9] The solid lines are the mean estimates over 1,000 replications; the shaded region represents the 95% simulated confidence intervals. We simulate 23,400 observations for each sample path, corresponding to the number of seconds in a business day (6.5 trading hours). The ReMeDI estimators perform well: the estimates are approximately unbiased with narrow confidence bands. Surprisingly, there is a significant average deviation of the LA estimates from the true parameters, and the confidence bands are much larger as well.

The deviation of the LA estimates is elicited by a *finite sample bias*, which is known to be a fraction of the *prior unknown* quadratic variation (QV) of the efficient price, see the discussion in Jacod et al. (2017). Thus to correct the bias, we need an estimate of the QV. But the estimation of QV in the presence of dependent noise is not trivial, see a discussion in Li et al. (2020). In a simulation context, we can obtain the QV and thus can give the LA estimators the privilege to make the bias correction, which is, of course not feasible in practice. The bottom panel of Figure 2 displays the bias corrected estimation of LA. Even with accurate bias correction, however, the ReMeDI estimators still outperform the LA estimators with almost no bias but greater accuracy.

It is interesting to compare ReMeDI and LA when the data frequencies vary. However, increasing the data frequency in a fixed time span has two effects: both the number of observations and the noise-to-signal ratio of tick returns will increase. We design a simulation study to separate the two effects and examine how sensitive ReMeDI and LA are to these changes.

The left panel of Figure 3 presents the mean-squared-error (MSE) of the ReMeDI and LA estimators for the first 20 autocovariances of the noise. The sample size varies from 23,400 (1 trading day) to 117,000 (1 trading week), and 468,000 (1 trading

[9] We select the same tuning parameter for the LA estimator as in Jacod et al. (2017); we also check other alternatives, and we find $k_n = 6$ leads to smaller bias.

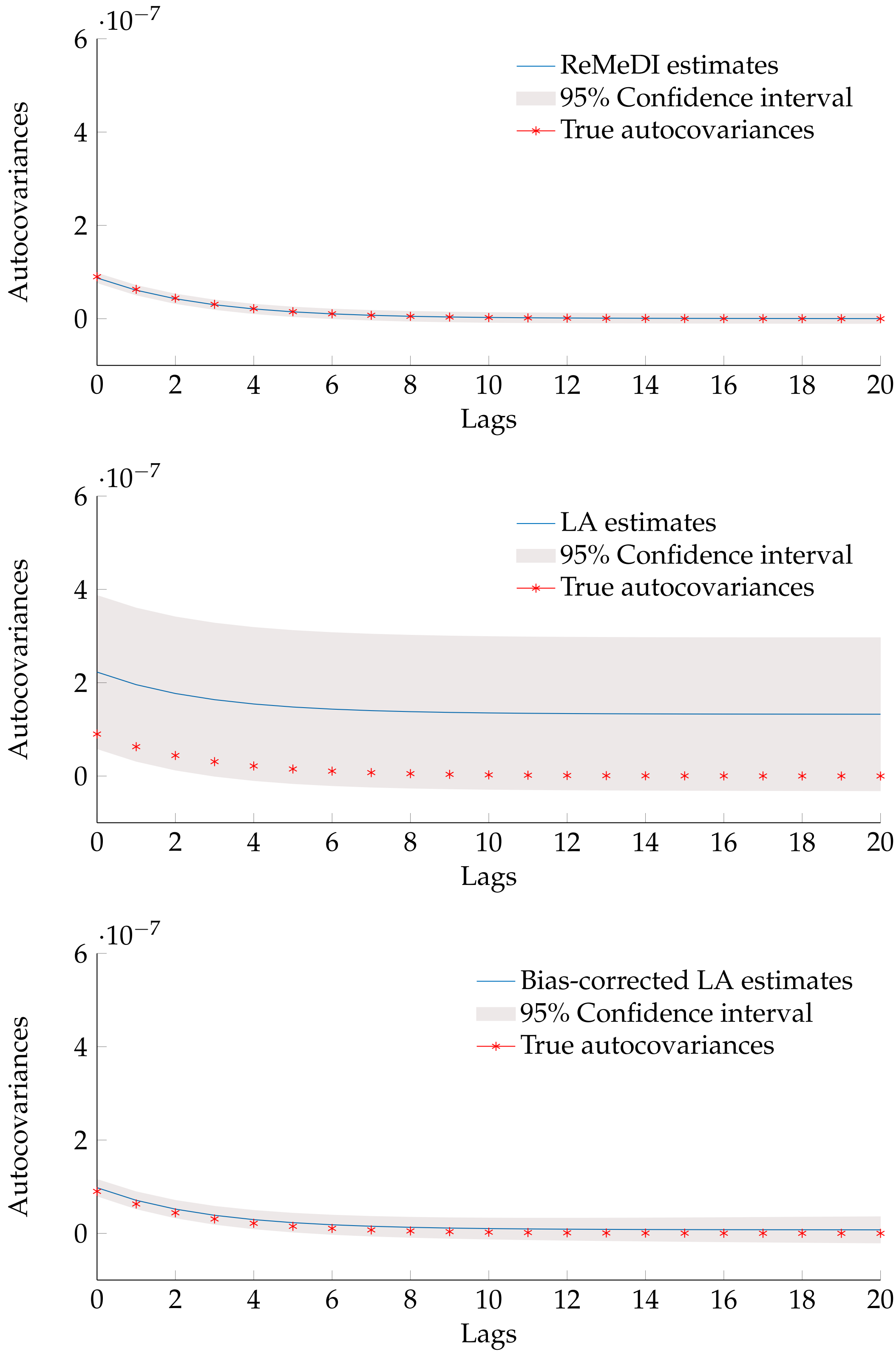

Figure 2: Estimation of the autocovariances of noise by the ReMeDI method (top panel), the local averaging method (middle panel) and the bias corrected local averaging method (bottom panel). The blue solid line is the mean estimates of 1,000 simulations by the three estimators. The tuning parameters of the ReMeDI and LA estimators are 10 and 6, respectively. The noise scale is fixed at $\gamma \equiv 5 \times 10^{-4}$.

month). The MSE of the ReMeDI estimators remains low and slightly drops when the sample sizes increases. The LA estimators, however, has larger MSE in a larger sample! This is statistically counterintuitive. However, it does make sense if we recall that the integrated volatility contributes to the finite sample bias of the LA estimators. Hence longer time span induces larger integrated volatility (relatively to the number of observations), which in turn leads to a larger finite sample bias. This is especially so if the sample covers a period of volatility burst, and the likelihood of an such event increases if the sampling period becomes large, see our empirical studies with real transaction prices.

The right panel of Figure 3 compares ReMeDI and LA when noise variance varies from 10^{-8} (small noise) to 10^{-6} (large noise). We note that the advantage of ReMeDI over LA is more prominent when the noise is smaller. Indeed, the size of noise in practice is closer to the small noise scenario, see an extensive empirical study by Christensen et al. (2014). Thus in an extreme case when the noise has *identical* statistical properties in two samples, LA may give very different estimates due to the differences in sample sizes or noise-to-signal ratios. The ReMeDI approach remains robust and accurate.

5.3 Random noise size and observation times

As the last robustness check, we now allow for stochastic observation times and random scales of noise. Following Jacod et al. (2017), we let $\{t_i^n\}$ follow an inhomogeneous Poisson process with rate $n\alpha_t$ where $\alpha_t = (1 + \cos(2\pi t))/2$ and the process γ satisfies

$$\gamma_t = C_\gamma \gamma'_t, \quad \mathrm{d}\gamma'_t = -\rho_\gamma(\gamma'_t - \mu_t)\mathrm{d}t + \sigma_\gamma \mathrm{d}W_t.$$

We set $\rho_\gamma = 10$, $\mu_t = 1 + 0.1\cos(2\pi t)$, $\sigma_\gamma = 0.1$, $C_\gamma = 5 \times 10^{-4}$. Figure 4 reports the estimation of the autocorrelation functions by the two estimators. We observe similar patterns presented in Figure 2: compared to the ReMeDI estimators, the LA estimators have large biases with wide confidence band.

The supplementary material Li and Linton (2020) provides additional simulation studies to examine the CLT, the effect of rounding error due to the discreteness of price and sensitivity to the choice of tuning parameters.

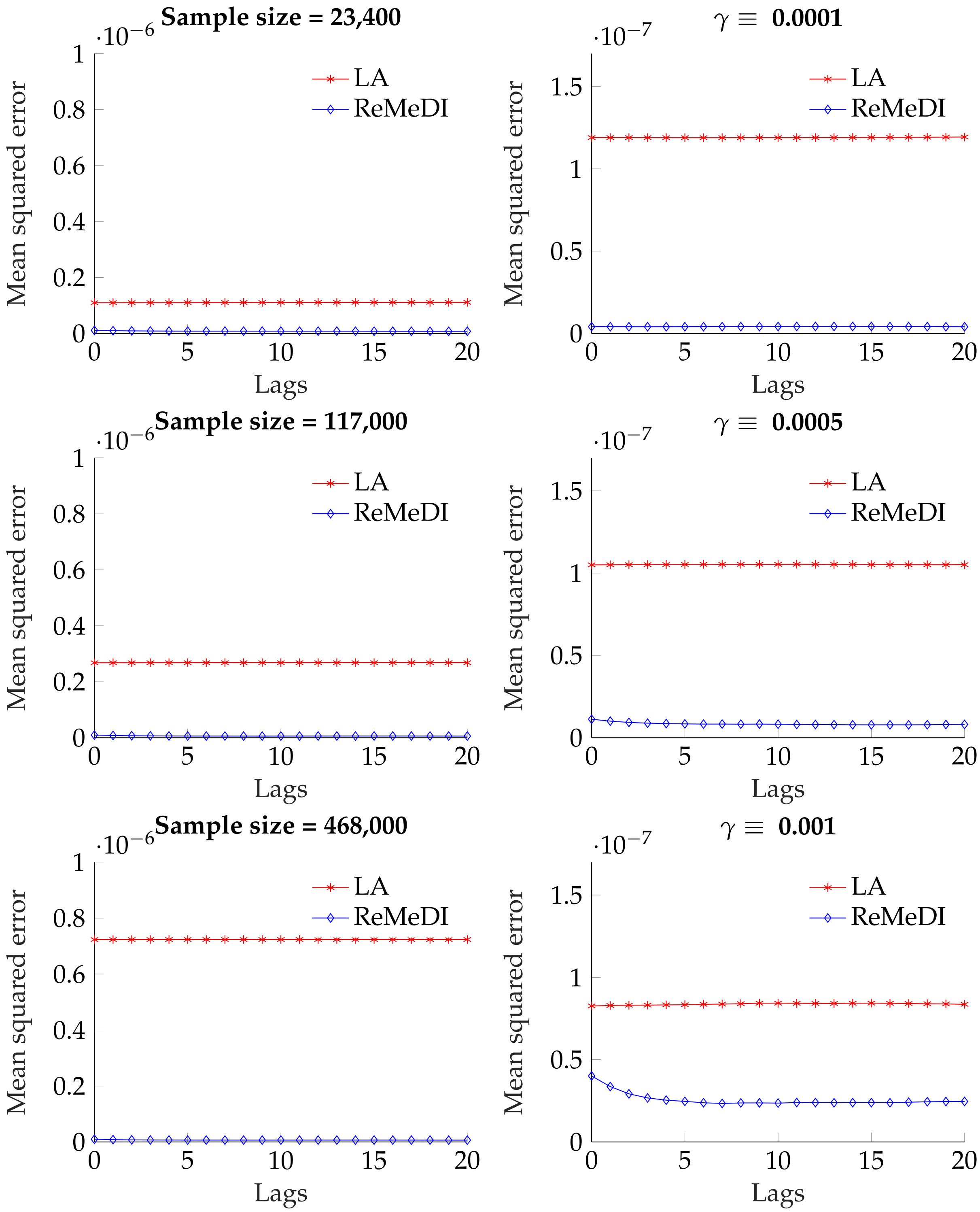

Figure 3: Mean squared error (MSE) of the ReMeDI and LA estimators for the first 20 autocovariances of noise based on 1,000 simulations. In the left panel, the noise scale is fixed at $\gamma = 5 \times 10^{-4}$ and the sample size varies; in the right panel, the size sample is 23,400 while the noise scale parameter varies. The tuning parameters of the ReMeDI and LA estimators are 10 and 6, respectively.

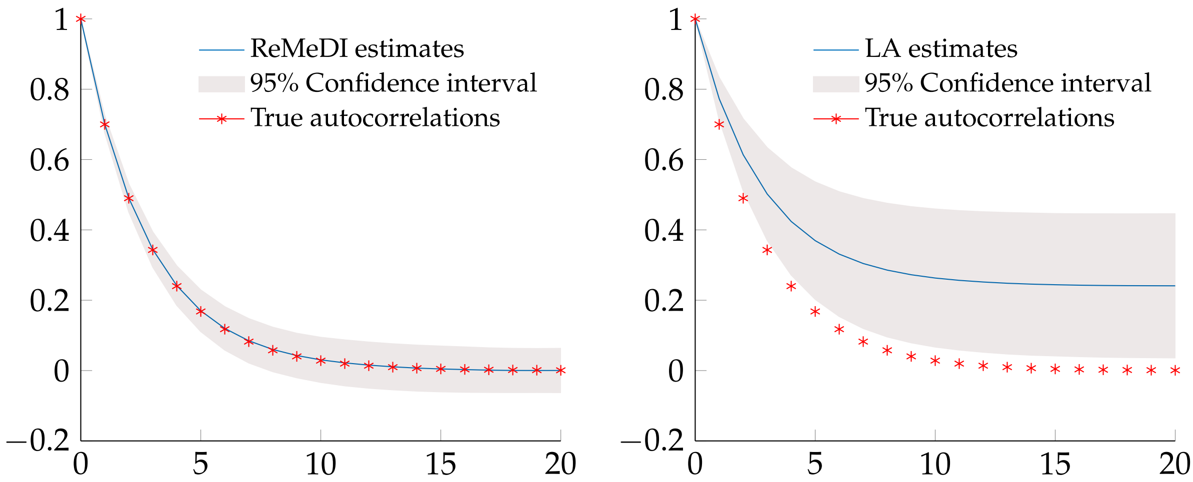

Figure 4: Estimation of the autocorrelations of noise by the ReMeDI method (left panel) and the local averaging method (right panel). The blue solid line is the mean estimates of 1,000 simulations by the two estimators. The tuning parameters of the ReMeDI and LA estimators are 10 and 6, respectively. The noise has stochastic scales and the observation times are random, see the specifications in Section 5.3.

6 Empirical Study

We obtain the transaction prices of Coca-Cola (trading symbol KO)[10] from the TAQ database for January 2018 (21 trading days). We remove prices before 9:30 and after 16:00. We collect approximately 50,000 observations per day, i.e., 2.1 transactions per second on average. The average price is 46.84$, with a standard deviation of 0.85.

Figure 5 plots the estimated autocovariances of noise by the ReMeDI estimators (the blue plots) based on samples of different sizes. The autocorrelation pattern is non-trivial: noise exhibits positive autocorrelations up to 4 lags and shortly thereafter, the sign switches to negative for a few lags, and then reverts to positive autocorrelations before decaying to zero around 20 lags. The pointwise confidence interval[11] includes zero or excludes positive values after lag 5, which is incompatible with simple long memory.

The ReMeDI estimates of microstructure noise presented in Figure 5 are economically intuitive. The positive autocovariances at the first several lags may be a consequence of the order splitting strategies by high-frequency traders (Biais et al. (1995)), or the successive transactions executed by limit orders (Parlour (1998)).[12] The negative

[10] In the supplementary material Li and Linton (2020), we use the transaction prices of General Electric (GE) and Citigroup (Citi), and we obtain similar results.

[11] Recall Section B.2 that the duration of successive observed prices is part of the asymptotic variance estimator. We do not plot the confidence intervals when we use transaction prices on different trading days since the prices will cover overnight non-trading hours.

[12] Hasbrouck and Ho (1987) and Choi et al. (1988) model the continuation of order flows by an AR(1)

autocovariances at the intermediate lags are consistent with the prediction of inventory models (Ho and Stoll (1981), Hendershott and Menkveld (2014)), in which the market makers induce negatively autocorrelated order flows to balance his inventories. However, the LA method gives very different estimates: it says that the noise is strongly autocorrelated without any sign of decay after 20 lags. This is economically counterintuitive — such a pattern, if it exists, would be exploited by high-frequency traders and we would expect it to disappear rapidly. Moreover, the serial dependence, according to the LA estimates, is even stronger when estimation is performed in a larger sample. Since we only estimate autocovariances of noise up to 20 ticks/lags, or a few seconds, it is statistically counterintuitive to obtain stronger autocovariance estimates using the prices of a week than using the prices in a single trading day. This is in line with our simulation study that the LA estimates are subject to a finite sample bias that depends on the noise-to-signal ratio and sample size. The ReMeDI approach retains its accuracy and robustness.

7 Concluding Remarks

We introduced a nonparametric method to separate the microstructure noise from the underlying semimartingale efficient prices in a general setting. We demonstrate the robustness of the proposed method compared to the main existing approach. We have concentrated on the infill setting primarily and the univariate case. The method naturally extends to the multivariate case, although in that case, several issues arise. First, the nonsynchronous trading issue has to be faced. Second, even when the assets trade on a common clock, there are some remaining theoretical results that need to be established for the infill case. We have not discussed efficiency in a great deal, but one can improve efficiency in two ways: first, by combining the estimators associated with different choices of $\boldsymbol{k}_n$ by a minimum distance, and second by doing a kind of GLS procedure using a local estimator of γ_u. We leave these problems for future research.

Appendix A Assumptions and Regularity Conditions

Assumption (H). *The process b is locally bounded, the process σ is càdlàg , there is a localizing sequence $\{\tau_n\}_n$ of stopping times and for each n a deterministic nonnegative function Γ_n on E satisfying $\int \Gamma_n^2(z)\lambda(\mathrm{d}z) < \infty$ such that $|\vartheta(\omega, t, z)| \wedge 1 \leq \Gamma_n(z)$ for all (ω, t, z) satisfying $t \leq \tau_n(\omega)$.*

process.

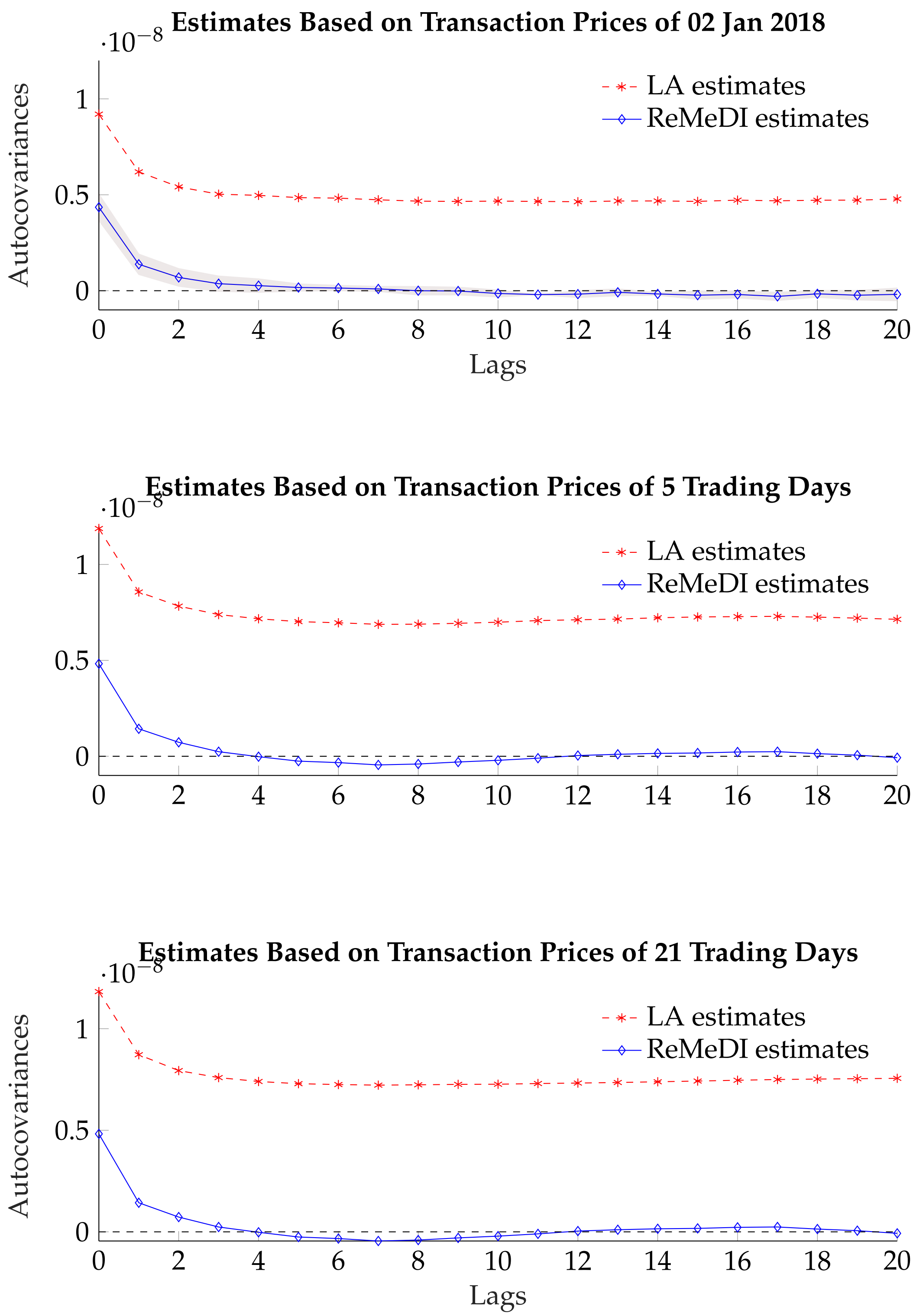

Figure 5: Estimation of autocovariances of noise for Coca-Cola (KO) in January 2018. In the top panel, we use the transaction prices of KO on 2 January 2018; in the middle panel, we use the transaction prices of KO in the second trading week (8 January 2018 to 12 January 2018); we employ the entire transaction prices of KO in January 2018 in the bottom panel. The tuning parameters for ReMeDI and LA are 10 and 6, respectively. The shaded area in the top panel represents the 95% confidence interval, and we set $i_n = 5, \phi_n = k_n^{3/5}/n$ to compute the asymptotic variances of the ReMeDI estimators, where n is the number of observations.

Definition A.1. *Let $\{\chi_i\}_{i\in\mathbb{Z}}$ be a sequence of stationary random variables. For any $k \in \mathbb{N}_+$, we define the following* mixing coefficients *for $k \in \mathbb{N}_+$:*

$$\rho_k := \sup\left\{|\mathbb{E}(V_h V_{k+h})| : \mathbb{E}(V_k) = \mathbb{E}(V_{k+h}) = 0, \|V_h\|_2 \le 1, \|V_{k+h}\|_2 \le 1, V_h \in \mathcal{G}_h, V_{k+h} \in \mathcal{G}^{k+h}\right\}, \tag{A.1}$$

where $\mathcal{G}_p := \sigma\{\chi_k : p \ge k\}$, $\mathcal{G}^q := \sigma\{\chi_k : q \le k\}$. The sequence $\{\chi_i\}_{i\in\mathbb{Z}}$ is ρ mixing *if $\rho_k \to 0$ as $k \to \infty$.*

Assumption (O). *$\alpha, \overline{\alpha}$ are two Itô semimartingales defined on $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\ge0}, \mathbb{P})$ satisfying Assumption (H). We further assume there is a localizing sequence $\{\tau_m\}_m$ of stopping times and positive constants $\kappa_{m,p}$ and κ such that:*

(i) For $t < \tau_m$, we have $\frac{1}{\kappa_{m,1}} \le \alpha_{t-} \le \kappa_{m,1}$ and $\overline{\alpha}_{t-} \le \kappa_{m,1}$.

(ii) Let $(\mathcal{F}^n_t)_{t\ge0}$ be the smallest filtration satisfying

(a) $\mathcal{F}_t \subset \mathcal{F}^n_t$,

(b) t^n_i is a $\{\mathcal{F}^n_t\}_{t\ge0}$ stopping time for $i = 0, 1, 2, \dots$,

(c) $\delta(n,i)$, conditional $\mathcal{F}^n_{i-1} := \mathcal{F}^n_{t^n_{i-1}}$, is independent of $\mathcal{F}_\infty := \bigvee_{t\ge0} \mathcal{F}_t$ for $i = 0, 1, 2, \dots$

(iii) With the restriction $\{t^n_{i-1} < \tau_m\}$, and for all $p > 0$,

$$\begin{aligned}
&\left|\mathbb{E}\left(\delta(n,i)\,\middle|\,\mathcal{F}^n_{i-1}\right) - \frac{\delta_n}{\alpha^n_{i-1}}\right| \le \kappa_{m,1}\delta_n^{\frac{3}{2}+\kappa},\\
&\left|\mathbb{E}\left(\left(\delta(n,i)\alpha^n_{i-1} - \delta_n\right)^2 \,\middle|\, \mathcal{F}^n_{i-1}\right) - \delta_n^2\overline{\alpha}^n_{i-1}\right| \le \kappa_{m,2}\delta_n^{2+\kappa},\\
&\mathbb{E}\left(\delta(n,i)^p\,\middle|\,\mathcal{F}^n_{i-1}\right) \le \kappa_{m,p}\delta_n^p.
\end{aligned} \tag{A.2}$$

Assumption (N). *Let $\{\chi_i\}_{i\in\mathbb{Z}}$ be a stationary ρ-mixing random sequence with mixing coefficients $\{\rho_k\}_{k\in\mathbb{N}_+}$. At stage n, the noise at time t^n_i is given by*

$$\varepsilon^n_i = \gamma^n_i \cdot \chi_i, \tag{A.3}$$

where γ is a nonnegative Itô semimartingale on $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\ge0}, \mathbb{P})$ satisfying Assumption (H) and is not identically zero on any interval. We further assume that $\{\chi_i\}_{i\in\mathbb{Z}}$ is centred at 0 with variance 1 and finite moments of all orders, and is independent of $\mathcal{F}_\infty$. Moreover, there is some $K > 0, v > 0$ such that

$$\rho_k \le \frac{K}{k^v} \quad \forall\, k \in \mathbb{N}_+. \tag{A.4}$$

Appendix B Asymptotic (Co)Variance and its Estimation

B.1 The asymptotic (co)variance

This section introduces $\boldsymbol{s}(\mathbf{j},\mathbf{j}')$ that appears in the asymptotic variance in Theorem 4.2. In the sequel whenever we have two vectors $\mathbf{j}=(j_1,\ldots,j_q),\mathbf{j}'=(j'_1,\ldots,j'_{q'})\in\mathfrak{J}$, we suppose without loss of generality that $q\le q'$. We denote

$$
\begin{aligned}
&\mathbf{j}\oplus\mathbf{j}'=(j_1,j_2,\ldots,j_q,j'_1,j'_2,\ldots,j'_{q'}),\quad \mathbf{j}_{-l}=\mathbf{j}\backslash\{j_l\},\\
&\mathbf{j}(+k)=(j_1+k,j_2+k,\ldots,j_q+k),\ \text{for } k\in\mathbb{Z},\\
&\mathbf{j}_{Q_q}=(j_l:l\in Q_q)\ \text{for } Q_q\subset\{1,2,\ldots,q\},\\
&\mathcal{Q}_q:=\{Q_q:Q_q\subsetneq\{1,2,\ldots,q\}\}.
\end{aligned}
$$

For each $Q_q\subset\{1,2,\ldots,q\}$, there is an associated (unique) pair of subsets:

$$
Q_q^c:=\{1,2,\ldots,q\}\backslash Q_q,\quad Q_{q'}:=Q_q\cup\{q+1,\ldots,q'\}. \tag{B.5}
$$

We denote for each $k\in\mathbb{Z}$ the following moments[13]

$$
\begin{aligned}
\boldsymbol{s}_0(\mathbf{j},\mathbf{j}';k)&:=\boldsymbol{r}\left(\mathbf{j}\oplus\left(\mathbf{j}'(+k)\right)\right)-\boldsymbol{r}\left(\mathbf{j}\right)\boldsymbol{r}\left(\mathbf{j}'\right);\\
\boldsymbol{s}_1(\mathbf{j},\mathbf{j}';k)&:=\sum_{Q_q\in\mathcal{Q}_q}\boldsymbol{r}\left(\mathbf{j}_{Q_q}\oplus\left(\mathbf{j}'_{Q_{q'}}(+k)\right)\right)\prod_{l\in Q_q^c}\boldsymbol{r}(j_l,j'_l+k);\\
\boldsymbol{s}_2(\mathbf{j},\mathbf{j}';k)&:=\sum_{\substack{j_l\in\mathbf{j},j'_{l'}\in\mathbf{j}'\\ l\neq l'}}\boldsymbol{r}(j_l,j'_{l'}+k)\boldsymbol{r}(\mathbf{j}_{-l})\boldsymbol{r}(\mathbf{j}'_{-l'})-\sum_{j_l\in\mathbf{j}}\boldsymbol{r}(\{j_l\}\oplus\mathbf{j}'(+k))\boldsymbol{r}(\mathbf{j}_{-l})\\
&\quad-\sum_{j'_{l'}\in\mathbf{j}'}\boldsymbol{r}(\{j'_{l'}+k\}\oplus\mathbf{j})\boldsymbol{r}(\mathbf{j}'_{-l'});
\end{aligned}
$$

Then $\boldsymbol{s}(\mathbf{j},\mathbf{j}')$ is given by

$$
\boldsymbol{s}(\mathbf{j},\mathbf{j}'):=\sum_{k\in\mathbb{Z}}\boldsymbol{s}_0(\mathbf{j},\mathbf{j}';k)+\boldsymbol{s}_1(\mathbf{j},\mathbf{j}';k)+\boldsymbol{s}_2(\mathbf{j},\mathbf{j}';k). \tag{B.6}
$$

B.2 The estimation of the asymptotic (co)variance

First, we introduce a sequence of notations

$$
\widehat{\delta}_i^n:=\left(\frac{k_n\delta(n,i+1+k_n)-t^n_{i+2+2k_n}+t^n_{i+2+k_n}}{(t^n_{i+k_n}-t^n_i)\vee\phi_n}\right)^2,\quad U(1)^n_t:=\sum_{i=0}^{n_t-w(1)_n}\widehat{\delta}_i^n,
$$

[13] By convention we let $\boldsymbol{r}(\emptyset)=1$.

$$U(2;\mathbf{j})_t^n := \sum_{i=0}^{n_t - w(2)_n} \widehat{\delta}_i^n \Delta_{\mathbf{j}}(Y)^n_{i+w(2)_2^n};$$

$$U(3,\mathbf{j},\mathbf{j}')_t^n := \sum_{i=0}^{n_t - w(3)_n} \widehat{\delta}_i^n \Delta_{\mathbf{j}}(Y)^n_{i+w(3)_2^n} \Delta_{\mathbf{j}'}(Y)^n_{i+w(3)_3^n},$$

$$U(4;\mathbf{j},\mathbf{j}')_t^n := -\sum_{i=2^{q-1}k_n}^{n_t - w(4)_n} \Delta_{\mathbf{j}}(Y)^n_i \Delta_{\mathbf{j}'}(Y)^n_{i+w(4)_2^n},$$

$$U(5,k;\mathbf{j},\mathbf{j}')_t^n := \sum_{Q_q \in \mathcal{Q}_q} \sum_{i=2^{e(Q_q)}k_n}^{n_t - w(5)_n} \Delta_{\mathbf{j}_{Q_q} \oplus (\mathbf{j}'_{Q_{q'}}(+k))}(Y)^n_i \prod_{\ell : l_\ell \in Q_q^c} \Delta_{(j_{l_\ell}, j'_{l_\ell}+k)}(Y)^n_{i+w(5)^n_{\ell+1}},$$

$$U(6,k;\mathbf{j},\mathbf{j}')_t^n := \sum_{\substack{j_l \in \mathbf{j}, j'_{l'} \in \mathbf{j}' \\ l \neq l'}} \sum_{i=2k_n}^{n_t - w(6)_n} \Delta_{(j_l, j'_{l'}+k)}(Y)^n_i \Delta_{\mathbf{j}_{-l}}(Y)^n_{i+w(6)_2^n} \Delta_{\mathbf{j}'_{-l'}}(Y)^n_{i+w(6)_3^n}$$

$$- \sum_{j_l \in \mathbf{j}} \sum_{i=2^{q'}k_n}^{n_t - w'(6)_n} \Delta_{\{j_l\} \oplus \mathbf{j}'(+k)}(Y)^n_i \Delta_{\mathbf{j}_{-l}}(Y)^n_{i+w'(6)_2^n}$$

$$- \sum_{j'_{l'} \in \mathbf{j}'} \sum_{i=2^q k_n}^{n_t - w''(6)_n} \Delta_{\{j'_{l'}+k\} \oplus \mathbf{j}}(Y)^n_i \Delta_{\mathbf{j}'_{-l'}}(Y)^n_{i+w''(6)_2^n},$$

$$U(7,k;\mathbf{j},\mathbf{j}')_t^n := \mathrm{ReMeDI}(\mathbf{j} \oplus \mathbf{j}'(+k))_t^n; \quad U(k;\mathbf{j},\mathbf{j}')_t^n := \sum_{\ell=5}^{7} U(\ell,k;\mathbf{j},\mathbf{j}')_t^n,$$

where the indices appear above are given by

$$w(1)_n := 2 + 2k_n,\ w(2)_2^n := 2 + (3 + 2^{q-1})k_n,\ w(2)_n := w(2)_2^n + j_1 + k_n;$$
$$w(3)_2^n := 2 + (3 + 2^{q-1})k_n,\ w(3)_3^n := 2 + (5 + 2^{q-1} + 2^{q'-1})k_n + j_1;$$
$$w(3)_n := w(3)_3^n + j'_1 + k_n;\ w(4)_2^n := 2k_n + q'_n + j_1,\ w(4)_n := w(4)_2^n + j'_1 + k_n;$$
$$e(Q_q) := (2|Q_q| + q' - q - 1) \vee 1, \quad w(5)^n_{\ell+1} := 4\ell k_n + \sum_{\ell'=1}^{\ell} j_{l_{\ell'}} \vee (j'_{l_{\ell'}} + k) \ \text{ for } \ell \geq 1,$$
$$w(5)_n := w(5)^n_{|Q_q^c|+1} + j_{l_{|Q_q^c|}} \vee (j'_{l_{|Q_q^c|}} + k) + k_n;$$
$$w(6)_2^n := (2^{q-2} + 2)k_n + j_\ell \vee (j'_{\ell'} + k),\ w(6)_3^n := (2^{q-2} + 2^{q'-2} + 2)k_n + j_1 + j_\ell \vee (j'_{\ell'} + k),$$
$$w'(6)_2^n := (2^{q-2} + 2)k_n + j_\ell \vee (j'_1 + k),\ w''(6)_2^n := (2^{q'-2} + 1)k_n + (j'_{\ell'} + k) \vee j_1,$$
$$w(6)_n := w(6)_3^n + j'_1 + k_n,\ w'(6)_n := w'(6)_2^n + j_1 + k_n,\ w''(6)_n := w''(6)_2^n + j'_1 + k_n.$$

The asymptotic variance estimator is given by

$$\widehat{\sigma}(\mathbf{j},\mathbf{j}')_t^n := \frac{1}{n_t} \sum_{\ell=1}^{3} \widehat{\sigma}_\ell(\mathbf{j},\mathbf{j}')_t^n, \tag{B.7}$$

where

$$
\begin{aligned}
\widehat{\sigma}_1(\mathbf{j},\mathbf{j}')^n_t &:= U(0;\mathbf{j},\mathbf{j}')^n_t + \sum_{k=1}^{i_n}\left(U(k;\mathbf{j},\mathbf{j}')^n_t + U(k;\mathbf{j}',\mathbf{j})^n_t\right) + (2i_n+1)U(4;\mathbf{j},\mathbf{j})^n_t;\\
\widehat{\sigma}_2(\mathbf{j},\mathbf{j}')^n_t &:= U(3,\mathbf{j},\mathbf{j}');\\
\widehat{\sigma}_3(\mathbf{j},\mathbf{j}')^n_t &:= \frac{1}{n_t^2}\mathrm{ReMeDI}(Y;\mathbf{j})^n_t\mathrm{ReMeDI}(Y;\mathbf{j}')^n_t U(1)^n_t\\
&\quad - \frac{1}{n_t}\left(\mathrm{ReMeDI}(Y;\mathbf{j})^n_t U(2,\mathbf{j}')^n_t + \mathrm{ReMeDI}(Y;\mathbf{j}')^n_t U(2,\mathbf{j})^n_t\right).
\end{aligned}
$$

The estimators seem quite complicated. However, the intuition will be clear in light of the following convergences, which are proved in the supplementary material Li and Linton (2020):

$$
\begin{aligned}
&\frac{1}{n_t}\widehat{\sigma}_1(\mathbf{j},\mathbf{j}')^n_t \xrightarrow{\mathbb{P}} \frac{\boldsymbol{s}(\mathbf{j},\mathbf{j}')}{A_t}\int_0^t \gamma_s^{q+q'}\mathrm{d}A_s, \quad \frac{1}{n_t}\widehat{\sigma}_2(\mathbf{j},\mathbf{j}')^n_t \xrightarrow{\mathbb{P}} \frac{\boldsymbol{r}(\mathbf{j})\boldsymbol{r}(\mathbf{j}')}{A_t}\int_0^t \gamma_s^{q+q'}\overline{\alpha}_s\mathrm{d}A_s;\\
&\frac{1}{n_t}\widehat{\sigma}_3(\mathbf{j},\mathbf{j}')^n_t \xrightarrow{\mathbb{P}} \frac{\mathbf{R}(\mathbf{j})\mathbf{R}(\mathbf{j}')}{A_t}\int_0^t \overline{\alpha}_s\mathrm{d}A_s - \frac{\mathbf{R}(\mathbf{j})\boldsymbol{r}(\mathbf{j}')}{A_t}\int_0^t \gamma_s^{q'}\overline{\alpha}_s\mathrm{d}A_s - \frac{\boldsymbol{r}(\mathbf{j})\mathbf{R}(\mathbf{j}')}{A_t}\int_0^t \gamma_s^{q}\overline{\alpha}_s\mathrm{d}A_s.
\end{aligned}
$$

Now we consider some special cases where the asymptotic (co)variances are simpler. As a consequence, the asymptotic variance estimators are also much simplified.

First, we consider the scenario $\overline{\alpha}_t \equiv 0$. The observations schemes that satisfy this condition include the *regular sampling scheme*, the *time-changed regular sampling scheme*; next, let $\overline{\alpha}_t \equiv 1$, one can verify that the *modulated Poisson sampling scheme* satisfies this condition, see the discussion in Jacod et al. (2017). The asymptotic (co)variance becomes

$$
\overline{\sigma}(\mathbf{j},\mathbf{j}')_t = \begin{cases} \frac{\boldsymbol{s}(\mathbf{j},\mathbf{j}')}{A_t}\int_0^t \gamma_s^{q+q'}\mathrm{d}A_s, & \text{if } \overline{\alpha}_t \equiv 0;\\ \frac{\boldsymbol{s}(\mathbf{j},\mathbf{j}')+\boldsymbol{r}(\mathbf{j})\boldsymbol{r}(\mathbf{j}')}{A_t}\int_0^t \gamma_s^{q+q'}\mathrm{d}A_s - \mathbf{R}(\mathbf{j})_t\mathbf{R}(\mathbf{j}')_t, & \text{if } \overline{\alpha}_t \equiv 1.\end{cases}
$$

and a consistent estimator is given by

$$
\widehat{\sigma}(\mathbf{j},\mathbf{j}')_t = \begin{cases} \frac{1}{n_t}\widehat{\sigma}_1(\mathbf{j},\mathbf{j}')^n_t, & \text{if } \overline{\alpha}_t \equiv 0;\\ \frac{1}{n_t}\left(\widehat{\sigma}_1(\mathbf{j},\mathbf{j}')^n_t + \widehat{\sigma}'_2(\mathbf{j},\mathbf{j}')^n_t + \widehat{\sigma}'_3(\mathbf{j},\mathbf{j}')^n_t\right), & \text{if } \overline{\alpha}_t \equiv 1;\end{cases}
$$

where

$$
\begin{aligned}
\widehat{\sigma}'_2(\mathbf{j},\mathbf{j}')^n_t &= \sum_{i=0}^{n_t - w(3)_n} \Delta_{\mathbf{j}}(Y)^n_{i+w(3)^n_2}\Delta_{\mathbf{j}'}(Y)^n_{i+w(3)^n_3},\\
\widehat{\sigma}'_3(\mathbf{j},\mathbf{j}')^n_t &= -\frac{1}{n_t}\mathrm{ReMeDI}(Y;\mathbf{j})^n_t\mathrm{ReMeDI}(Y;\mathbf{j}')^n_t.
\end{aligned}
$$

References

AÏT-SAHALIA, Y., P. A. MYKLAND, AND L. ZHANG (2005): "How often to sample a continuous-time process in the presence of market microstructure noise," *Review of Financial Studies*, 18, 351–416.

——— (2011): "Ultra high frequency volatility estimation with dependent microstructure noise," *Journal of Econometrics*, 160, 160–175.

ATHEY, S. AND G. W. IMBENS (2006): "Identification and inference in nonlinear difference-in-differences models," *Econometrica*, 74, 431–497.

BANDI, F. M., D. PIRINO, AND R. RENO (2017): "Excess idle time," *Econometrica*, 85, 1793–1846.

BANDI, F. M. AND J. R. RUSSELL (2006): "Separating microstructure noise from volatility," *Journal of Financial Economics*, 79, 655–692.

——— (2008): "Microstructure noise, realized variance, and optimal sampling," *Review of Economic Studies*, 75, 339–369.

BARNDORFF-NIELSEN, O. E., P. R. HANSEN, A. LUNDE, AND N. SHEPHARD (2008): "Designing realized kernels to measure the ex post variation of equity prices in the presence of noise," *Econometrica*, 76, 1481–1536.

BIAIS, B., P. HILLION, AND C. SPATT (1995): "An empirical analysis of the limit order book and the order flow in the Paris Bourse," *Journal of Finance*, 50, 1655–1689.

BOLLERSLEV, T., J. LI, AND Y. XUE (2018): "Volume, Volatility and Public News Announcements," *Review of Economic Studies*, 85, 2005—2041.

CHAKER, S. (2017): "On high frequency estimation of the frictionless price: The use of observed liquidity variables," *Journal of Econometrics*, 201, 127–143.

CHOI, J. Y., D. SALANDRO, AND K. SHASTRI (1988): "On the estimation of bid-ask spreads: Theory and evidence," *Journal of Financial and Quantitative Analysis*, 23, 219–230.

CHRISTENSEN, K., R. C. OOMEN, AND M. PODOLSKIJ (2014): "Fact or friction: Jumps at ultra high frequency," *Journal of Financial Economics*, 114, 576–599.

CLINET, S. AND Y. POTIRON (2017): "Estimation for high-frequency data under parametric market microstructure noise," Tech. rep.

DA, R. AND D. XIU (2019): "When moving-average models meet high-frequency data: Uniform inference on volatility," Tech. rep.

DELBAEN, F. AND W. SCHACHERMAYER (1994): "A general version of the fundamental theorem of asset pricing," *Mathematische Annalen*, 300, 463–520.

DIEBOLD, F. X. AND G. STRASSER (2013): "On the correlation structure of microstructure noise: A financial economic approach," *Review of Economic Studies*, 80, 1304–1337.

FOUCAULT, T., M. PAGANO, A. ROELL, AND A. RÖELL (2013): *Market Liquidity: Theory, Evidence, and Policy*, Oxford University Press.

FRENCH, K. R. AND R. ROLL (1986): "Stock return variances: The arrival of information and the reaction of traders," *Journal of Financial Economics*, 17, 5–26.

HANSEN, P. R. AND A. LUNDE (2006): "Realized variance and market microstructure noise," *Journal of Business & Economic Statistics*, 24, 127–161.

HASBROUCK, J. (2007): *Empirical market microstructure: The institutions, economics, and econometrics of securities trading*, Oxford University Press.

HASBROUCK, J. AND T. S. HO (1987): "Order Arrival, Quote Behavior, and the Return-Generating Process," *Journal of Finance*, 42, 1035–1048.

HENDERSHOTT, T. AND A. J. MENKVELD (2014): "Price pressures," *Journal of Financial Economics*, 114, 405–423.

HO, T. AND H. R. STOLL (1981): "Optimal dealer pricing under transactions and return uncertainty," *Journal of Financial Economics*, 9, 47–73.

HUANG, R. D. AND H. R. STOLL (1997): "The components of the bid-ask spread: A general approach," *Review of Financial Studies*, 10, 995–1034.

JACOD, J., Y. LI, P. A. MYKLAND, M. PODOLSKIJ, AND M. VETTER (2009): "Microstructure noise in the continuous case: the pre-averaging approach," *Stochastic Processes and their Applications*, 119, 2249–2276.

JACOD, J., Y. LI, AND X. ZHENG (2017): "Statistical properties of microstructure noise," *Econometrica*, 85, 1133–1174.

JACOD, J. AND A. N. SHIRYAEV (2003): *Limit Theorems for Stochastic Processes*, vol. 288, Springer-Verlag Berlin.

LEHMANN, B. N. (1990): "Fads, martingales, and market efficiency," *Quarterly Journal of Economics*, 105, 1–28.

LI, Y., S. XIE, AND X. ZHENG (2016): "Efficient estimation of integrated volatility incorporating trading information," *Journal of Econometrics*, 195, 33–50.

LI, Z. M., R. J. LAEVEN, AND M. H. VELLEKOOP (2020): "Dependent microstructure noise and integrated volatility estimation from high-frequency data," *Journal of Econometrics*, 215, 536–558.

LI, Z. M. AND O. LINTON (2019): "A ReMeDI for microstructure noise," https://ssrn.com/abstract=3423607.

——— (2020): "Supplementary material for 'A ReMeDI for microstructure noise'," https://papers.ssrn.com/abstract=3688787.

LINTON, O. AND S. MAHMOODZADEH (2018): "Implications of high-frequency trading for security markets," *Annual Review of Economics*, 10, 237–259.

O'HARA, M. (1995): *Market Microstructure Theory*, vol. 108, Blackwell Cambridge, MA.

PARK, A. AND H. SABOURIAN (2011): "Herding and contrarian behavior in financial markets," *Econometrica*, 79, 973–1026.

PARLOUR, C. A. (1998): "Price dynamics in limit order markets," *Review of Financial Studies*, 11, 789–816.

PODOLSKIJ, M. AND M. VETTER (2009): "Estimation of volatility functionals in the simultaneous presence of microstructure noise and jumps," *Bernoulli*, 15, 634–658.

TODOROV, V. AND G. TAUCHEN (2011): "Volatility jumps," *Journal of Business & Economic Statistics*, 29, 356–371.

XIU, D. (2010): "Quasi-maximum likelihood estimation of volatility with high frequency data," *Journal of Econometrics*, 159, 235–250.

YATCHEW, A. (1997): "An elementary estimator of the partial linear model," *Economics Letters*, 57, 135–143.

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

ZHANG, L., P. A. MYKLAND, AND Y. AÏT-SAHALIA (2005): "A tale of two time scales: determining integrated volatitility with noisy high-frequency data," *Journal of the American Statistical Association*, 100, 1394–1411.