Local Parametric Estimation in High Frequency Data

Yoann POTIRON
Faculty of Business and Commerce, Keio University, 2-15-45 Mita, Minato-ku, Tokyo 108-8345, Japan (*potiron@fbc.keio.ac.jp*)

Per MYKLAND
Department of Statistics, The University of Chicago, 5734 S. University Avenue, Chicago, IL (*mykland@pascal.uchicago.edu*)

We give a general time-varying parameter model, where the multidimensional parameter possibly includes jumps. The quantity of interest is defined as the integrated value over time of the parameter process $\Theta = T^{-1}\int_0^T \theta_t^* dt$. We provide a local parametric estimator (LPE) of Θ and conditions under which we can show the central limit theorem. Roughly speaking those conditions correspond to some uniform limit theory in the parametric version of the problem. The framework is restricted to the specific convergence rate $n^{1/2}$. Several examples of LPE are studied: estimation of volatility, powers of volatility, volatility when incorporating trading information and time-varying MA(1).

KEY WORDS: Integrated volatility; Market microstructure noise; Powers of volatility; Quasi-maximum likelihood estimator

1. INTRODUCTION

Modeling dynamics is essential in various fields, including finance, economics, physics, environmental engineering, geology, and sociology. Time-varying parametric models can deal with a specific problem in dynamics, namely, the temporal evolution of systems. The extensive literature on time-varying parameter models and local parametric methods include and are not limited to Fan and Gijbels (1996), Hastie and Tibshirani (1993), or Fan and Zhang (1999) when regression and generalized regression models are involved, locally stationary processes following the work of Dahlhaus (1997, 2000), Dahlhaus and Rao (2006), or any other time-varying parameter models, for example, Stock and Watson (1998) and Kim and Nelson (2006).

In this paper, we propose to specify local parametric methods in the particular context of high-frequency statistics for a broad class of problems. Local methods have been used extensively in the high-frequency data literature, see, for example, Mykland and Zhang (2009, 2011), Kristensen (2010), Reiß (2011), or Jacod and Rosenbaum (2013), among many others. If we define T as the horizon time, the (random) target quantity in this monograph is defined as the integrated parameter

$$\Theta := \frac{1}{T}\int_0^T \theta_s^* ds, \tag{1}$$

which can be equal to the volatility, the covariation between several assets, the variance of the microstructure noise, the friction parameter of the model with uncertainty zones (see Example 4.4 for more details), the time-varying parameters of the MA(1) model, etc. To estimate the integrated parameter, we estimate the *local parameter* on each block by using the parametric estimator on the observations within the block and take a weighted sum of the local parameter estimates, where each weight is equal to the corresponding block length. We call the obtained estimator the *local parametric estimator* (LPE).

In Section 3, we investigate conditions under which we can establish the related central limit theorem with convergence rate $n^{1/2}$, where n is the (possibly expected) number of observations. The framework is such that the local block length vanishes asymptotically. Basically, we aim to provide the statistician with a transparent and as simple as possible device to tackle the time-varying parameter problem based on central limit theory in the parametric version of the problem. The original key probabilistic step of the proof, which formally allows for switching from random to deterministic parameter, is the use of regular conditional distribution theory (see, e.g., Breiman 1992). The price to pay is some kind of uniformity in the parametric limit theory results, and to show that some deviation between the parametric and the time-varying parameter model vanishes asymptotically.

In Section 4, the technology is used on five distinct examples to derive the related central limit theorems. As far as the authors know, all those results are new. Depending on the considered example, the LPE is useful for one or several of the following reasons:

- *Robustness*: The LPE is robust to time-varying parameters (such as the noise variance, η from the model with uncertainty zones, the parameters of the MA(1) process) which

Journal of Business & Economic Statistics
July 2020, Vol. 38, No. 3
DOI: 10.1080/07350015.2019.1566731
Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/r/jbes.

are usually assumed constant. This is the case of all our examples, except for Example 4.3.

- *Efficiency*: The LPE turns out to be more efficient than the global estimator or existing concurrent approaches. This is the case of Example 4.3. In addition, the LPE is conjectured to be efficient in all our examples except for Example 4.4.
- *Definition of new estimators*: It can be the case that the estimator does not work globally but that the LPE provides a good candidate as in Examples 4.2 and 4.3.

We describe the five examples in what follows. To estimate integrated volatility under noisy observations, Xiu (2010) studied the quasi-maximum likelihood-estimator (QMLE) originally examined in Aït-Sahalia, Mykland, and Zhang (2005), showed the corresponding asymptotic theory when the variance of the noise is fixed and obtained a convergence rate $n^{1/4}$, which is optimal (see Gloter and Jacod 2001). More recently, Aït-Sahalia and Xiu (2019) establish that it is robust to shrinking noise satisfying $O_p(1/n^{3/4})$ and Da and Xiu (2017) obtain central limit theorem with rate ranging from $n^{1/4}$ to $n^{1/2}$ depending on the magnitude of the noise. When assuming that it is $O_p(1/\sqrt{n})$, we show that the LPE of the QMLE is optimal (with rate $n^{1/2}$) and furthermore robust to time-varying noise variance.

Another important problem, which goes back to Barndorff-Nielsen and Shephard (2002), is the estimation of higher powers of volatility. To do that, we define a LPE where the local estimates are powers of the QMLE of volatility. Under the assumption on small noise $O_p(1/\sqrt{n})$, we show that this estimator is optimal and robust to time-varying noise variance. This is an example where the global approach does not work as the QMLE is only consistent when estimating volatility.

A more recent problem is the estimation of volatility when incorporating trading information. To do that, Li, Xie, and Zheng (2016) assume that the noise is a parametric function of trading information with a remaining noise component of order $O_p(1/\sqrt{n})$. Their strategy consists in first estimating the parametric part of the noise, and then take the sum of square pre-estimated efficient returns. They also advocate for the use of the QMLE after price pre-estimation although they do not provide the associated limit theory. We show that the latter approach, when considering the LPE of QMLE, is optimal and provides a better asymptotic variance (AVAR) than the former technique. In addition, a modification of the local estimator as in Example 2 allows us to estimate higher powers of volatility.

A concurrent ultrahigh frequency approach to model the observed price was given in Robert and Rosenbaum (2011, 2012), who introduced the semiparametric model with uncertainty zones where η is the one-dimensional friction parameter, observation times are endogenous and observed prices lie on a tick grid. As most likely correlated with the volatility, it is natural to consider η_t as a time-varying parameter. We provide a formal model extension and establish the according limit theory of the LPE of the estimator considered in their work. In addition, our empirical illustration available in the online supplement seems to indicate that η_t is indeed time-varying.

In the last example, we consider an application in time series and introduce a time-varying MA(1) model with null mean. The time series is observed in high frequency on $[0, T]$ and θ_t^* corresponds to the two-dimensional parameter of the MA(1) process. We show that the LPE of the MLE is optimal and document that it outperforms the global MLE and other concurrent approaches in finite sample.

The remaining of this article is organized as follows. The LPM is introduced in the following section. Conditions for the central limit theory are stated in Section 3. We give the examples in Section 4. We investigate the finite sample performance of the local QMLE of volatility and compares it to the global approach in Section 5. We conclude in Section 6. Consistency in a simple model, proofs, additional numerical simulations on MA(1) model, and an empirical illustration on the model with uncertainty zones are gathered in an online appendix.

2. THE LOCALLY PARAMETRIC MODEL (LPM)

2.1. Data-Generating Mechanism

We assume that we observe the d-dimensional vectors $Z_{0,n}, \ldots, Z_{N_n,n}$, where N_n can be random, the observation times satisfy $\tau_{0,n} := 0 < \tau_{1,n} < \cdots < \tau_{N_n,n} \leq T$. The observations and the observation times are both related to the latent parameter θ_t^*.

As an example, the observations can satisfy $Z_{\tau_{i,n},n} = X_{\tau_{i,n}} + \epsilon_{i,n}$, where $X_t = \sigma_t dW_t$ stands for the efficient price, W_t is a standard Brownian motion, $\epsilon_{i,n}$ corresponds to the market microstructure noise (which will be restricted to be of order $\epsilon_{i,n} = O_p(1/\sqrt{n})$ due to the limitation of the technology developed in Section 3), is iid and independent from X_t, and the latent parameter is equal to the volatility, that is, $\theta_t^* = \sigma_t^2$.

We assume that the parameter process θ_t^* takes values in K, a (not necessarily compact) subset of $\mathbb{R}^p$. We do not assume any independence between θ_t^* and the other quantities driving the observations, such as the Brownian motion of the efficient price process. In particular, there can be leverage effect (see, e.g., Wang and Mykland 2014; Aït-Sahalia et al. 2017). Also, the arrival times $\tau_{i,n}$ and the parameter θ_t^* can be correlated, that is, there is (some kind of) endogeneity in sampling times.

2.2. Asymptotics

There are commonly two choices of asymptotics in the literature: the *high-frequency* asymptotics, which makes the number of observations explode on $[0, T]$, and the *low-frequency* asymptotics, which takes T to infinity. We choose the former one. Investigating the low-frequency implementation case is beyond the scope of this article.[1]

2.3. Estimation

The approach taken here is frequent in high-frequency data. We define the block size (i.e., the number of observations in a block) as h_n, and the number of blocks as $B_n := \ulcorner N_n h_n^{-1} \urcorner$. For

[1]If we set down the asymptotic theory in the same way as in Dahlhaus (1997, p. 3), we conjecture that the results of this article would stay true.

$i = 1, \ldots, B_n$ we define the parameter average on the ith block as

$$\Theta_{i,n} := \frac{\int_{\mathrm{T}_{i-1,n}}^{\mathrm{T}_{i,n}} \theta_s^* ds}{\Delta \mathrm{T}_{i,n}}, \tag{2}$$

where $\mathrm{T}_{i,n} := \min(\tau_{ih_n}, T)$ and its corresponding parametric estimator as $\widehat{\Theta}_{i,n}$. Then, we take the weighted sum of $\widehat{\Theta}_{i,n}$ and obtain an estimator of the integrated spot process

$$\widehat{\Theta}_n := \frac{1}{T} \sum_{i=1}^{B_n} \widehat{\Theta}_{i,n} \Delta \mathrm{T}_{i,n}, \tag{3}$$

where $\Delta \mathrm{T}_{i,n} = \mathrm{T}_{i,n} - \mathrm{T}_{i-1,n}$. We call (3) the *LPE*. We assume that

$$h_n/n \to 0 \tag{4}$$

so that when observations are regular the block size $\Delta \mathrm{T}_{i,n} := Th_n/n$ vanishes asymptotically. In view of Condition (T) and Remark 4, we have similarly that $\mathbb{E}[\Delta \mathrm{T}_{i,n}] = O(h_n/n)$ also goes to 0 when observations are not regular.

3. THE CENTRAL LIMIT THEOREM

We present in this section the general technology of our article.[2] It is mainly based on Theorem 2-2 in Jacod (1997), or similarly Theorem IX.7.3 and Theorem IX.7.28 in Jacod and Shiryaev (2003) or Theorem 2.2.15 in Jacod and Protter (2011), along with regular conditional distribution techniques (see, e.g., Breiman 1992, sec. 4.3, pp.77–80). More specifically, we provide sufficient conditions to the aforementioned theorem in the particular context of this article. Those conditions are based on the limit theory in the parametric version of the problem, which we assume pre-obtained by the statistician.

The following methods are specified[3] to the rate of convergence $n^{\frac{1}{2}}$. Formally, we aim to find the limit distribution of

$$n^{\frac{1}{2}} T^{-1} \sum_{i=1}^{B_n} \left(\widehat{\Theta}_{i,n} - \Theta_{i,n}\right) \Delta \mathrm{T}_{i,n}. \tag{5}$$

Specifically, we want to show that (5) converges *stably*[4] to a limit distribution. We first give the definition of stable convergence.

Definition (Stable convergence). A sequence of random variables Z_n is said to converge $\mathcal{J}_T$-stably to Z, which is defined on an extension $(\Omega', \mathcal{F}', P')$ of $(\Omega, \mathcal{F}, P)$, if for any $E \in \mathcal{J}_T$ and for any continuous bounded function f we have

$$\mathbb{E}\left[f(Z_n)\mathbf{1}_E\right] \to \mathbb{E}'\left[f(Z)\mathbf{1}_E\right].$$

3.1. Regular Observation Case

We consider first the simple case when observations are regular, that is, $\tau_{i,n} = iT/n$ and $N_n = n$. We assume that $\mathcal{J}_t$ is a (continuous-time) filtration on $(\Omega, \mathcal{F}, P)$ such that θ_t^* is adapted to it. In the following of this paper, when using the conditional expectation $\mathbb{E}_\tau[Z]$,[5] we will refer to the conditional expectation of Z knowing $\mathcal{J}_\tau$. We define the discrete-time version of the filtration as $\mathcal{I}_{i,n} = \mathcal{J}_{\tau_{i,n}}$. Finally, if we denote the returns of the observations as

$$R_{i,n} = Z_{\tau_{i,n},n} - Z_{\tau_{i-1,n},n}, \tag{6}$$

we assume that the returns can be expressed as

$$R_{i,n} = F_n\left(\{P_{s,n}\}_{0 \le s \le \tau_{i-1,n}}, U_{i,n}, \{\theta_s^*\}_{\tau_{i-1,n} \le s \le \tau_{i,n}}\right), \tag{7}$$

where $F_n(x, y, z)$ is a $\mathbb{R}^d$-dimensional nonrandom function,[6] the random innovation $U_{i,n}$ are iid (although with distribution which can depend on n) adapted to $\mathcal{I}_{i,n}$ and independent of the past information $\mathcal{I}_{i-1,n}$, $P_{t,n}$ is a (possibly multidimensional) process adapted to $\mathcal{J}_t$ which stands for the past that matters in the model. We further assume that $P_{t,n}$ is independent from θ_t^*.

The key example stands as follows. We assume that the observations are following the additive model $Z_{\tau_{i,n},n} = X_{\tau_{i,n}} + \epsilon_{i,n}$, where $X_t = \sigma_t dW_t$ is the efficient price and $\epsilon_{i,n}$ the (shrinking) iid noise independent from X_t, and that the parameter is $\theta_t^* = \sigma_t^2$. In that case $U_{i,n} = (\{W_s\}_{\tau_{i-1,n} \le s \le \tau_{i,n}} - W_{\tau_{i-1,n}}, \epsilon_{i,n})$, and $P_{s,n} = \epsilon_{i,n}$ if $\tau_{i,n} \le s < \tau_{i+1,n}$. The function[7] F_n takes on the form

$$F_n = \int_{\tau_{i-1,n}}^{\tau_{i,n}} \sigma_s dW_s + \epsilon_{i,n} - \epsilon_{i-1,n}. \tag{8}$$

Crucial to the expression (8) is that the dependence in the past is only through the past noise $\epsilon_{i-1,n}$, that is, we do not need to know the whole past of $P_{t,n}$, but rather only the current value. This will be very useful in what follows.

We provide now the outline of the method. Our goal is to investigate the limit distribution of (5) using prior limit result on the parametric version of the problem. A common approach in high frequency statistics proofs consists in decomposing $\left(\widehat{\Theta}_{i,n} - \Theta_{i,n}\right)$ into

$$\left(\widehat{\Theta}_{i,n} - \hat{\hat{\Theta}}_{i,n}\right) + \left(\hat{\hat{\Theta}}_{i,n} - \theta^*_{\mathrm{T}_{i-1,n}}\right) + \left(\theta^*_{\mathrm{T}_{i-1,n}} - \Theta_{i,n}\right), \tag{9}$$

where $\hat{\hat{\Theta}}_{i,n}$ stands for the estimator when we hold the parameter constant on each block. Then, one can usually deal with the first term and the third term (most likely using Burkholder–Davis–Gundy and Markov type of inequalities) and eventually show that they vanish asymptotically. The main work lies in establishing the central limit theory of the second term in (9). A typical proof consists in using locally parametric results along with some Riemann sum argument. But this can be cumbersome as the parameter on each block, although

[2]Note that the local approach in this article is related to the large-T-based-approach and problem of Giraitis, Kapetanios, and Yates (2014).

[3]It is possible to specify the problem with a general rate of convergence, but all the considered examples from this article are with convergence rate $n^{1/2}$.

[4]One can look at definitions of stable convergence in Rényi (1963), Aldous and Eagleson (1978), Chapter 3 (p. 56) of Hall and Heyde (1980), Rootzén (1980), Section 2 (pp. 169–170) of Jacod and Protter (1998), Definition VIII.5.28 in Jacod and Shiryaev (2003) or Definition 1 in Podolskij and Vetter (2010).

[5]The related assumption is that τ is a $\mathcal{J}_t$-stopping time.

[6]We assume that $F_n(x, y, z)$ is jointly measurable, and that $P_{t,n}$ is taking values on a Borel space. Additionally, we assume that for any $(P_{s,n}, U_{i,n}, \theta_s^*)$, we have $\mathbb{E} \mid F_n(P_{s,n}, U_n, \theta_s^*) \mid < \infty$

[7]The advised reader will have noticed that F_n is not a function in the ordinary sense. We still abusively refer to it as a "function."

constant, is random. Instead, we propose to look at the further decomposition of $\big(\hat{\tilde{\Theta}}_{i,n} - \theta^*_{\tau_{i-1,n}}\big)$ into

$$\big(\hat{\tilde{\Theta}}_{i,n} - \hat{\tilde{\Theta}}^{\mathbf{P}}_{i,n}\big) + \big(\hat{\tilde{\Theta}}^{\mathbf{P}}_{i,n} - \theta^*_{\tau_{i-1,n}}\big), \text{ where} \tag{10}$$

$$\hat{\tilde{\Theta}}^{\mathbf{P}}_{i,n} := \hat{\tilde{\Theta}}_{i,n} \mid \{P_{s,n}\}_{0\leq s\leq \tau_{i-1,n}} = \mathbf{P}, \tag{11}$$

and $\mathbf{P}$ is a fixed nonrandom past. In the case of (8), we can choose $\mathbf{P} = 0$. From this new decomposition, it is expected as relatively accessible to show that the first term goes to 0, so that the central limit theory will be investigated on the second term of the decomposition. By conditioning by one particular past in (11), we got rid of some randomness, although the parameter is still random. Using conditional regular distribution results in our proofs, we actually show that we can also take the parameter nonrandom. The price to pay for such method is to show some kind of uniformity in the parameter value when showing the limit results, and that the first term in (10) vanishes asymptotically.

We introduce some definition. For $i = 1, \dots, B_n$ we define the returns on the ith block $R^j_{i,n} := R_{(i-1)h_n+j,n}$ for $j = 1, \dots, h_n$, and similarly $U^j_{i,n}$, $\tau^j_{i,n}$, $W^j_{i,n}$ and $\epsilon^j_{i,n}$. We assume that

$$\widehat{\Theta}_{i,n} := \hat{\theta}_{h_n,n}(R^1_{i,n}; \cdots; R^{h_n}_{i,n}), \tag{12}$$

where $\hat{\theta}_{h_n,n}$ is a function on $\mathbb{R}^{dh_n}$. The approximated returns and the approximated estimates are defined as

$$\widetilde{R}^j_{i,n} := F_n\big(\{P_{s,n}\}_{0\leq s\leq \tau^{j-1}_{i,n}}, U^j_{i,n}, \theta^*_{\mathrm{T}_{i-1,n}}\big), \tag{13}$$

$$\hat{\tilde{\Theta}}_{i,n} := \hat{\theta}_{h_n,n}(\widetilde{R}^1_{i,n}; \cdots; \widetilde{R}^{h_n}_{i,n}). \tag{14}$$

Basically, those two expressions can be seen as the pendant of, respectively, (7) and (12) when we hold the parameter constant equal to its block initial value $\theta^*_{\mathrm{T}_{i-1,n}}$. In the case of the key example (8), we obtain that the approximated returns are of the form

$$\widetilde{R}^j_{i,n} = \sigma_{\mathrm{T}_{i-1,n}}(W^j_{i,n} - W^{j-1}_{i,n}) + (\epsilon^j_{i,n} - \epsilon^{j-1}_{i,n}). \tag{15}$$

We also introduce the conditional parametric version as

$$\widetilde{R}^{j,\mathbf{P}}_{i,n} := \mathbb{E}\big[\widetilde{R}^j_{i,n} \mid \{U^k_{i,n}\}_{k\leq j}, \{P_{s,n}\}_{0\leq s\leq \mathrm{T}_{i-1,n}} = \mathbf{P}\big], \tag{16}$$

$$\hat{\tilde{\Theta}}^{\mathbf{P}}_{i,n} := \hat{\theta}_{h_n,n}(\widetilde{R}^{1,\mathbf{P}}_{i,n}; \cdots; \widetilde{R}^{h_n,\mathbf{P}}_{i,n}). \tag{17}$$

Here, we fix the past equal to $\mathbf{P}$ in (16), which removes some randomness compared with (13). In the key example, we can (arbitrarily) choose $\mathbf{P} = 0$, and this past will only "affect" the first conditional parametric version of the return on the block equal to

$$\widetilde{R}^{1,\mathbf{P}}_{i,n} = \sigma_{\mathrm{T}_{i-1,n}}(W^1_{i,n} - W^0_{i,n}) + \epsilon^1_{i,n}, \tag{18}$$

whereas for $j = 2, \dots, h_n$, we have $\widetilde{R}^{j,\mathbf{P}}_{i,n} = \widetilde{R}^j_{i,n}$. This key example is an instance where the model is 1-Markovian in the sense that the past only affects the value of the first return on the block. This is quite mild assumption, and we will see that more sophisticated models, such as the model with uncertainty zones, naturally exhibit longer past time-dependence. Moreover, we introduce a parametric version of the returns and the estimators when the parameter is equal to θ and the past fixed to $\mathbf{P}$. Accordingly, the randomness is further reduced in the following expressions. This will be useful in Condition (E).

$$R^{j,\mathbf{P},\theta}_{i,n} := \mathbb{E}\big[\widetilde{R}^j_{i,n} \mid \{U^k_{i,n}\}_{k\leq j}, \theta^*_{\mathrm{T}_{i-1,n}} = \theta, \{P_{s,n}\}_{0\leq s\leq \mathrm{T}_{i-1,n}} = \mathbf{P}\big], \tag{19}$$

$$\hat{\tilde{\Theta}}^{\mathbf{P},\theta}_{i,n} := \hat{\theta}_{h_n,n}(R^{1,\mathbf{P},\theta}_{i,n}; \cdots; R^{h_n,\mathbf{P},\theta}_{i,n}). \tag{20}$$

We provide now the assumptions on θ^*_t. The first assumption considers the continuous Itô-semimartingale case.

Condition (P1). The parameter θ^*_t is of the form

$$d\theta^*_t := a^\theta_t dt + \sigma^\theta_t dW^\theta_t, \tag{21}$$

where a^θ_t is adapted locally bounded (of dimension p) and σ^θ_t is a nonnegative continuous Itô-process adapted locally bounded (of dimension $p \times p$), and W^θ_t is a standard p-dimensional Brownian motion.

We introduce a norm for

$$u \in \mathbb{R}^p \text{ as } \mid u \mid= \sqrt{(u^{(1)})^2 + \cdots + (u^{(p)})^2}.$$

The following assumption allows for a more general process than semi-martingales. Nonetheless, this assumption is quite restrictive, in particular since h_n does not show up on the right hand-side of (22). In practice this is useful when considering a smooth parameter which cannot be expressed as a "pure drift."

Condition (P2). θ^*_t satisfies uniformly in $i = 1, \dots, B_n$ that

$$\mathbb{E}_{\mathrm{T}_{i-1,n}}\Big[\sup_{\mathrm{T}_{i-1,n}\leq s\leq \mathrm{T}_{i,n}} \mid \theta^*_s - \theta^*_{\mathrm{T}_{i-1,n}} \mid^2\Big] = o_p(n^{-1}). \tag{22}$$

As the uniformity of limit results on the whole space K might be impossible to obtain, we allow to work on the compact subspace K_M, which grows to K as M increases. Accordingly, we assume that θ^*_t is locally bounded on a compact set K_M in the sense that there exists $\tau_m \xrightarrow{\mathbb{P}} T$ such that for any m, there exists $M_m > 0$ which satisfies $\theta^*_t \in K_{M_m}$ for any $t \in [0, \tau_m]$.

We provide in what follows sufficient conditions to the bias condition (3.10), the increment condition (3.11) and the Lindeberg condition (3.13) in Theorem 3-2 from Jacod (1997). (Almost) equivalently, Theorem IX.7.3 and Theorem IX.7.28 in Jacod and Shiryaev (2003) or Theorem 2.2.15 in Jacod and Protter (2011) could have been used. Those conditions are based on the parametric version of the problem.

Condition (E). For any (nonrandom) parameter $\theta \in K$, we assume that there exists a (nonrandom) covariance matrix V_θ positive definite such that for any $M > 0$, we have V_θ is bounded for any $\theta \in K_M$ and uniformly in $\theta \in K_M$ and in $i = 1, \dots, B_n$ we have

$$\mathbb{E}\big[\big(\hat{\tilde{\Theta}}^{\mathbf{P},\theta}_{i,n} - \theta\big)\big] = o(n^{-\frac{1}{2}}) \tag{23}$$

$$\mathrm{var}\,\big[h_n^{\frac{1}{2}}\big(\hat{\tilde{\Theta}}^{\mathbf{P},\theta}_{i,n} - \theta\big)\big] = V_\theta T + o(1) \tag{24}$$

$$\mathbb{E}\big[h_n \mid \hat{\tilde{\Theta}}^{\mathbf{P},\theta}_{i,n} - \theta \mid^2 \mathbf{1}\{h_n n^{-\frac{1}{2}} \mid \hat{\tilde{\Theta}}^{\mathbf{P},\theta}_{i,n} - \theta \mid > \epsilon\}\big] = o(1)\ , \forall \epsilon > 0. \tag{25}$$

We let $B_n(t)$ be the number of blocks before t, and $\mathcal{M}_b$ the set of all bounded martingales. We now provide the central limit theorem.

Theorem 1 (Central limit theorem with regular observation times). We assume Condition (E). Moreover, we assume Condition (P1) and that the block size h_n is such that

$$n^{-\frac{1}{2}} h_n = o(1), \tag{26}$$

or Condition (P2). Let M_t be a p-dimensional square-integrable continuous martingale. Furthermore, we assume that for all $t \in [0, T]$ we have

$$n^{-\frac{1}{2}} h_n \sum_{i=1}^{B_n(t)} \mathbb{E}_{\mathrm{T}_{i-1,n}}\Big[\big(\hat{\widehat{\Theta}}^{\mathbf{P}}_{i,n} - \theta^*_{\tau_{i-1,n}}\big)\big(M_{\mathrm{T}_{i,n}} - M_{\mathrm{T}_{i-1,n}}\big)^T\Big] \xrightarrow{\mathbb{P}} 0, \tag{27}$$

$$n^{-\frac{1}{2}} h_n \sum_{i=1}^{B_n(t)} \mathbb{E}_{\mathrm{T}_{i-1,n}}\Big[\big(\hat{\widehat{\Theta}}^{\mathbf{P}}_{i,n} - \theta^*_{\tau_{i-1,n}}\big)\big(N_{\mathrm{T}_{i,n}} - N_{\mathrm{T}_{i-1,n}}\big)\Big] \xrightarrow{\mathbb{P}} 0, \tag{28}$$

for all $N \in \mathcal{M}_b(M^{\perp})$, where $\mathcal{M}_b(M^{\perp})$ is the class of all elements of $\mathcal{M}_b$ which are orthogonal to M (i.e., to all components of M). Finally, we assume that

$$n^{-\frac{1}{2}} h_n \sum_{i=1}^{B_n} \big(\widehat{\Theta}_{i,n} - \hat{\widehat{\Theta}}^{\mathbf{P}}_{i,n}\big) \xrightarrow{\mathbb{P}} 0. \tag{29}$$

Then, stably in law as $n \to \infty$, we have

$$n^{\frac{1}{2}}\big(\widehat{\Theta}_n - \Theta\big) \to \widetilde{Z}, \tag{30}$$

where $\langle \widetilde{Z}, \widetilde{Z}\rangle_t = T^{-1}\int_0^t V_{\theta^*_s} ds$, and $\langle \widetilde{Z}, M\rangle_t = 0$. In particular, we have

$$n^{\frac{1}{2}}\big(\widehat{\Theta}_n - \Theta\big) \to \Big(T^{-1}\int_0^T V_{\theta^*_s} ds\Big)^{\frac{1}{2}} \mathcal{N}(0,1). \tag{31}$$

Remark 1 (Parametric model). Note that in the case where the time-varying parameter model is equal to the parametric model with parameter equal to θ^*, the AVAR of $\widehat{\Theta}_n$ is equal to the variance of the parametric model, that is,

$$n^{\frac{1}{2}}\big(\widehat{\Theta}_n - \Theta\big) \to V_{\theta^*}^{\frac{1}{2}} \mathcal{N}(0,1).$$

Remark 2 (Estimating the AVAR). If the statistician does not have a (parametric) variance estimator at hand and that her parametric estimator can be written as in Mykland and Zhang (2017), one can use the techniques of the cited paper to obtain a variance estimate. Investigating if such techniques would work in our setting is beyond the scope of this paper. If she has a variance estimator $\hat{v}_{h_n,n}$, then for any $i = 1, \dots, B_n$ she can estimate the ith block variance $\widehat{V}_{i,n}$ as $\widehat{V}_{i,n} := \hat{v}_{h_n,n}(R^1_{i,n}; \cdots; R^{h_n}_{i,n})$, and the AVAR as the weighted sum

$$\widehat{V}_n = T^{-1}\sum_{i=1}^{B_n} \widehat{V}_{i,n}\Delta \mathrm{T}_{i,n}. \tag{32}$$

This estimator will be consistent under mild uniformity assumptions.

Remark 3 (Nonzero asymptotic bias). If we further assume that in place of condition (27) there is a nonzero continuous process G_t such that

$$n^{-\frac{1}{2}} h_n \sum_{i=1}^{B_n(t)} \mathbb{E}_{\mathrm{T}_{i-1,n}}\Big[\big(\hat{\widehat{\Theta}}^{\mathbf{P}}_{i,n} - \theta^*_{\tau_{i-1,n}}\big)\big(M_{\mathrm{T}_{i,n}} - M_{\mathrm{T}_{i-1,n}}\big)^T\Big] \xrightarrow{\mathbb{P}} G_t, \tag{33}$$

then (30) still holds, where $\langle \widetilde{Z}, \widetilde{Z}\rangle_t = T^{-1}\int_0^t V_{\theta^*_s} ds$ and $\langle \widetilde{Z}, M\rangle_t = G_t$, but (31) no longer holds.

3.2. Nonregular Observation Case

We consider now the case when observations can be random (even endogenous). We define the increment of time as $\Delta\tau_{i,n} := \tau_{i,n} - \tau_{i-1,n}$ and make the first natural assumption.

Condition (T). The observation times are such that

$$\mathbb{E}\big[N_n\big] = O(n), \tag{34}$$

$$\sup_{1 \le i \le N_n} \mathbb{E}_{\tau_{i-1,n}}\big[(\Delta\tau_{i,n})^3\big] = O_p(n^{-3}). \tag{35}$$

Remark 4 (Block length). As an obvious consequence of (35), we have that the block length satisfies $\mathbb{E}\big[\Delta \mathrm{T}_{i,n}\big] = O(h_n n^{-1})$.

The observation times are related to θ^*_t, as are the returns. We assume that $(R_{i,n}, \Delta\tau_{i,n})$ satisfies (7), and that all the definitions (12)–(20) follow. Finally, we define $\Delta\widetilde{\mathrm{T}}^{\mathbf{P}}_{i,n} = \widetilde{\tau}^{\mathbf{P}}_{h_n i,n} - \widetilde{\tau}^{\mathbf{P}}_{(h_n-1)i,n}$ and $\Delta \mathrm{T}^{\mathbf{P},\theta}_{i,n} = \tau^{\mathbf{P},\theta}_{h_n i,n} - \tau^{\mathbf{P},\theta}_{(h_n-1)i,n}$. We adapt Condition (E) in this case.

Condition (E).* For any (nonrandom) parameter $\theta \in K$, we assume that there exists a (nonrandom) covariance matrix $V_\theta > 0$ such that for any $M > 0$, we have V_θ is bounded for any $\theta \in K_M$ and uniformly in $\theta \in K_M$ and in $i = 1, \dots, B_n$ we have

$$\mathbb{E}\big[\big(\hat{\widehat{\Theta}}^{\mathbf{P},\theta}_{i,n} - \theta\big)\Delta \mathrm{T}^{\mathbf{P},\theta}_{i,n}\big] = o(h_n n^{-\frac{3}{2}}), \tag{36}$$

$$\mathrm{var}\big[h_n^{\frac{1}{2}}\big(\hat{\widehat{\Theta}}^{\mathbf{P},\theta}_{i,n} - \theta\big)\Delta \mathrm{T}^{\mathbf{P},\theta}_{i,n}\big] = V_\theta \mathbb{E}\big[\Delta \mathrm{T}^{\mathbf{P},\theta}_{i,n}\big] T h_n n^{-1} + o(h_n^2 n^{-2}), \tag{37}$$

$$\mathbb{E}\big[n^2 h_n^{-1} (A^{\mathbf{P},\theta}_{i,n})^2 \mathbf{1}_{\{n^{\frac{1}{2}} A^{\mathbf{P},\theta}_{i,n} > \epsilon\}}\big] = o(1)\ , \forall \epsilon > 0, \tag{38}$$

where $A^{\mathbf{P},\theta}_{i,n} = \mid \hat{\widehat{\Theta}}^{\mathbf{P},\theta}_{i,n} - \theta \mid \Delta \mathrm{T}^{\mathbf{P},\theta}_{i,n}$.

We also adapt the central limit theorem.

Theorem 2 (Central limit theorem with nonregular observation times). We assume Condition (T) and Condition (E*). Moreover, we assume Condition (P1) and (26), or Condition (P2). Let M_t be a p-dimensional square-integrable continuous martingale. Furthermore, we assume that for all $t \in [0,T]$ we have

$$\frac{n^{\frac{1}{2}}}{T}\sum_{i=1}^{B_n(t)}\mathbb{E}_{\mathrm{T}_{i-1,n}}\Big[\big(\hat{\widehat{\Theta}}^{\mathbf{P}}_{i,n}-\theta^*_{\tau_{i-1,n}}\big)\Delta\widetilde{\mathrm{T}}^{\mathbf{P}}_{i,n}\big(M_{\mathrm{T}_{i,n}}-M_{\mathrm{T}_{i-1,n}}\big)^T\Big]\xrightarrow{\mathbb{P}}0, \tag{39}$$

$$n^{\frac{1}{2}}\sum_{i=1}^{B_n(t)}\mathbb{E}_{\mathrm{T}_{i-1,n}}\Big[\big(\hat{\widehat{\Theta}}^{\mathbf{P}}_{i,n}-\theta^*_{\tau_{i-1,n}}\big)\Delta\widetilde{\mathrm{T}}^{\mathbf{P}}_{i,n}\big(N_{\mathrm{T}_{i,n}}-N_{\mathrm{T}_{i-1,n}}\big)\Big]\xrightarrow{\mathbb{P}}0, \tag{40}$$

for all $N \in \mathcal{M}_b(M^\perp)$. Finally, we assume that

$$n^{\frac{1}{2}}\sum_{i=1}^{B_n}\big(\widehat{\Theta}_{i,n}\Delta\mathrm{T}_{i,n}-\hat{\widehat{\Theta}}^{\mathbf{P}}_{i,n}\Delta\widetilde{\mathrm{T}}^{\mathbf{P}}_{i,n}\big)\xrightarrow{\mathbb{P}}0, \tag{41}$$

$$n^{\frac{1}{2}}\sum_{i=1}^{B_n}\mathbb{E}_{\mathrm{T}_{i-1,n}}\Big[\big|\Delta\widetilde{\mathrm{T}}^{\mathbf{P}}_{i,n}-\Delta\mathrm{T}_{i,n}\big|\Big]\xrightarrow{\mathbb{P}}0, \tag{42}$$

uniformly in $i = 1,\ldots,B_n$. Then, stably in law as $n\to\infty$, we have

$$n^{\frac{1}{2}}\big(\widehat{\Theta}_n-\Theta\big)\to\widetilde{Z}, \tag{43}$$

where $\langle\widetilde{Z},\widetilde{Z}\rangle_t = T^{-1}\int_0^t V_{\theta^*_s}ds$, and $\langle\widetilde{Z},M\rangle_t = 0$. In particular, we have

$$n^{\frac{1}{2}}\big(\widehat{\Theta}_n-\Theta\big)\to\left(T^{-1}\int_0^T V_{\theta^*_s}ds\right)^{\frac{1}{2}}\mathcal{N}(0,1). \tag{44}$$

Remark 5 (Nonzero asymptotic bias). More generally, if we assume that there is a nonzero continuous process G_t such that for all $t \in [0,T]$ we have

$$\frac{n^{\frac{1}{2}}}{T}\sum_{i=1}^{B_n(t)}\mathbb{E}_{\mathrm{T}_{i-1,n}}\Big[\big(\hat{\widehat{\Theta}}^{\mathbf{P}}_{i,n}-\theta^*_{\tau_{i-1,n}}\big)\Delta\widetilde{\mathrm{T}}^{\mathbf{P}}_{i,n}\big(M_{\mathrm{T}_{i,n}}-M_{\mathrm{T}_{i-1,n}}\big)^T\Big]\xrightarrow{\mathbb{P}}G_t, \tag{45}$$

instead of (39), then (43) still holds, where $\langle\widetilde{Z},\widetilde{Z}\rangle_t = T^{-1}\int_0^t V_{\theta^*_s}ds$, and $\langle\widetilde{Z},M\rangle_t = G_t$, but (44) no longer holds.

3.3. Bias Correction

As the parametric estimator must satisfy the bias condition (36), it is useful to consider in some instances a bias-corrected (BC) version of it which provides the estimate on the ith block $\widehat{\Theta}^{(\mathrm{BC})}_{i,n}$. The BC LPE is then constructed as

$$\widehat{\Theta}^{(\mathrm{BC})}_n = \frac{1}{T}\sum_{i=1}^{B_n}\widehat{\Theta}^{(\mathrm{BC})}_{i,n}\Delta\mathrm{T}_{i,n}.$$

4. EXAMPLES

This section provides some applications of the theory introduced in Section 3. The central limit theorems provided in this section are all new. We choose four examples with regular observations in which it is sufficient to show the conditions of Theorem 1. We further consider the model with uncertainty zones where there is endogeneity in observation times implying that we have to verify the more general conditions of Theorem 2.

4.1. Estimation of Volatility With the QMLE

4.1.1. Central Limit Theorem. We assume that the noise has the form

$$\epsilon_{i,n} := n^{-\alpha}v^{\frac{1}{2}}_{\tau_{i,n}}\gamma_{\tau_{i,n}},$$

where $\alpha \geq 1/2$, the noise variance v_t is time-varying, and γ_t are iid with null-mean and unity variance. In other words we have $\epsilon_{i,n} = O_p(1/\sqrt{n})$. The parameter process is defined as the two-dimensional volatility and noise variance process $\theta^*_t = (\sigma^2_t, v_t)$ and thus $\Theta = \big(T^{-1}\int_0^T\sigma^2_t dt, T^{-1}\int_0^T v_t dt\big)$. Correspondingly we work locally with the QMLE considered in Xiu (2010, p. 236) and we introduce the notation for the corresponding LPE $\widehat{\Theta}_n = (\widehat{\sigma}^2_n, \widehat{v}_n)$.

We also consider the bias-corrected version of the QMLE $\widehat{\Theta}^{(BC)}_n$, where the procedure to construct the unbiased estimator is given in Section 4.1.2. In numerical simulations under a realistic framework, this bias is not observed even with small values of n (see Section 6 in Xiu (2010) and Section 5 in Clinet and Potiron (2018b)), and thus it is safe to use $\widehat{\Theta}_n = (\widehat{\sigma}^2_n, \widehat{v}_n)$ in practice.

The assumption of $\alpha \geq 1/2$ is quite restrictive in view of the related literature on the QMLE. Unfortunately in the case $\alpha < 1/2$, the techniques of this article do not apply. Xiu (2010) showed the CLT of the QMLE when v_t is non time-varying and $\alpha = 0$. In the same setting, Clinet and Potiron (2018b) showed that the AVAR can be smaller when using the LPE with $B_n = B$ fixed and documented that in finite sample the LPE was advantageous over the global QMLE. Aït-Sahalia and Xiu (2019) actually establish that the MLE is robust to noise of the form $O_p(1/n^{3/4})$. Da and Xiu (2017) show the central limit theory with rate of convergence ranging from $n^{1/2}$ to $n^{1/4}$ depending on the magnitude of the noise.

However, the techniques allow us to investigate how the LPE behaves in a different asymptotics, that is, when the noise variance is $O_p(1/\sqrt{n})$ and B_n tends to $+\infty$. Moreover, we allow for heteroscedasticity in noise variance. Finally, in the case where the noise variance goes to 0 at the same speed as the variance of the returns, that is, $\alpha = 1/2$, we can also retrieve the integrated variance noise. In accordance with the setting of this paper, the convergence rate of both the volatility and the noise is $n^{1/2}$.

To verify the conditions for the CLT, we use heavily the asymptotic results of the QMLE (see Theorem 6 in Xiu (2010)) and the MLE in the low-frequency asymptotics (see Proposition 1 in p. 369 of Aït-Sahalia, Mykland, and Zhang (2005)). The result is formally embedded in the following theorem.

Theorem 3 (QMLE). We define $\mathcal{F}^X_t$ the filtration generated by X_t.

(i) We assume that $\alpha > \frac{1}{2}$. Then, $\mathcal{F}^X_T$-stably in law as $n\to\infty$,

$$n^{\frac{1}{2}}\left(\widehat{\sigma}^2_n - T^{-1}\int_0^T\sigma^2_s ds\right)\to\left(6T^{-1}\int_0^T\sigma^4_s ds\right)^{\frac{1}{2}}\mathcal{N}(0,1). \tag{46}$$

(ii) When $\alpha = \frac{1}{2}$, we have $\mathcal{F}^X_T$-stable convergence in law of $n^{\frac{1}{2}}\big(\widehat{\Theta}^{(BC)}_n - \Theta\big)$ to a mixed normal random variable with zero mean and AVAR given by

$$T^{-1}\begin{pmatrix} A & -\int_0^T \left(\sigma_s^4+2\sigma_s^2 v_s+4\sigma_s^3\sqrt{4v_s+\sigma_s^2}\right)ds \\ \bullet & \frac{1}{2}\int_0^T \left(2v_s+\sigma_s^2\right)\left(\sigma_s^2+2v_s+\sigma_s\sqrt{4v_s+\sigma_s^2}\right)ds \end{pmatrix}, \quad (47)$$

where $A=\int_0^T \left(2\sigma_s^4+4\sigma_s^3\sqrt{4v_s+\sigma_s^2}\right)ds$.

Remark 6 (Estimation of high-frequency covariance with the QMLE). To estimate integrated covariance under noisy observations and asynchronous observations, Aït-Sahalia, Fan, and Xiu (2010) introduced a QMLE based on a synchronization of observation times. It is clear that their generalized synchronization method can be expressed as a LPM. In view of the close connection between their proposed estimator (2) on p. 1506 and the QMLE studied in Section 4.1, the conditions of our work can be verified and thus Theorem 2 (p. 1506) of the authors can be adapted with the LPE in a framework similar to Section 4.1, that is, when the noise variance is $O(n^{-1/2})$ and time-varying.

4.1.2. Algorithm to Construct the Unbiased Estimator. We describe here the algorithm to obtain $\widehat{\Theta}_n^{(\mathrm{BC})}$. Note that the bias-correction is only required when $\alpha=1/2$.

1. We compute the local QMLEs.
2. From Theorem 6 (p. 241) in Xiu (2010), we compute the corresponding W_1 and W_2.
3. We change some entries of the matrices to ensure unbiased estimates when using formulas (21) and (22) in the aforementioned theorem.
4. We compute the unbiased local QMLE using the formulas (21) and (22) with the corrected matrices.
5. The bias-corrected LPE $\widehat{\Theta}_n^{(\mathrm{BC})}$ is taken as the mean of local bias-corrected estimates.

4.2. Estimation of Powers of Volatility

Here the parameter is $\theta_t^*=g(\sigma_t^2)$ with g not being the identity function. We are concerned with the estimation of powers of volatility $\Theta=T^{-1}\int_0^T g(\sigma_t^2)dt$ under microstructure noise with variance $O(1/\sqrt{n})$ in the same setting as in Section 4.1.

The problem was introduced in Barndorff-Nielsen and Shephard (2002). They showed that the case $g(x)=x^2$ is related to the asymptotic variance of the realized volatility. One can also consult Barndorff-Nielsen et al. (2006), Mykland and Zhang (2017, Proposition 2.17, p. 138) and Renault, Sarisoy, and Werker (2017) for related developments. All those studies assume no microstructure noise.

When there is microstructure noise, Jacod, Podolskij, and Vetter (2010) used the pre-averaging method. In the special case of quarticity, one can also look at Mancino and Sanfelici (2012) and Andersen, Dobrislav, and Schaumburg (2014). In the case of tricicity, see Altmeyer and Bibinger (2015).

Under no microstructure noise, block estimation (Mykland and Zhang 2009, sec. 4.1, pp. 1421–1426) has the ability to make the mentioned estimators approximately or fully efficient. The path followed to do that is to first estimate locally the volatility $\widehat{\sigma}_{i,n}^2$ and then take a Riemann sum of $g(\widehat{\sigma}_{i,n}^2)$. See also Jacod and Rosenbaum (2013) for an extended version of the method in some ways.

In the same spirit when allowing for microstructure noise, we propose to use locally the estimation $g(\widehat{\sigma}_{i,n}^2)$, where $\widehat{\sigma}_{i,n}^2$ is the QMLE estimate of the volatility on the ith block. As pointed out in Jacod and Rosenbaum (2013), even if we use locally the bias-corrected estimator $\left(\widehat{\sigma}_{i,n}^{(\mathrm{BC})}\right)^2$, we will pay a price for the fact that we use the function g in front. In particular, an asymptotic bias quite challenging to correct for will appear in the asymptotic limit theory, as seen in Theorem 3.1 in the cited paper. To get rid of most parts of this bias, we follow the idea at the beginning of Section 3.2 of the cited work and choose h_n such that

$$n^{-1/2}h_n^{3/2}\to\infty. \quad (48)$$

Note that this is not incompatible with the other condition (26), that is, $n^{-1/2}h_n\to 0$, that will be assumed in what follows. With (48), the part of the bias that does not vanish grows to the extent that it explodes asymptotically. This leads us to consider the following two bias-corrected estimators

$$\widehat{\Theta}_n^{(\mathrm{BC},1)}=B_n^{-1}\sum_{i=1}^{B_n}\left(g(\widehat{\sigma}_{i,n}^2)-\frac{3}{h_n}\widehat{\sigma}_{i,n}^4 g''(\widehat{\sigma}_{i,n}^2)\right). \quad (49)$$

$$\begin{aligned}\widehat{\Theta}_n^{(\mathrm{BC},2)}&=B_n^{-1}\sum_{i=1}^{B_n}\Bigg(g\left(\left(\widehat{\sigma}_{i,n}^{(\mathrm{BC})}\right)^2\right)\\&\quad-\frac{(\widehat{\sigma}_{i,n}^{(\mathrm{BC})})^4+2(\widehat{\sigma}_{i,n}^{(\mathrm{BC})})^3\sqrt{4\widehat{v}_{i,n}^{(\mathrm{BC})}+(\widehat{\sigma}_{i,n}^{(\mathrm{BC})})^2}}{h_n}g''\\&\quad\times\left(\left(\widehat{\sigma}_{i,n}^{(\mathrm{BC})}\right)^2\right)\Bigg). \quad (50)\end{aligned}$$

The theorem is given in what follows. The proof uses a local delta method and then follows the proof of Theorem 3.

Theorem 4 (Powers of volatility). Let g a nonnegative function such that

$$|\, g^{(j)}(x)\,|\leq K(1+|\,x\,|^{p-j}),\ j=0,1,2,3, \quad (51)$$

for some constants $K>0, p\geq 3$.

(i) We assume that $\alpha>\frac{1}{2}$. Then, $\mathcal{F}_T^X$-stably in law as $n\to\infty$,

$$n^{\frac{1}{2}}\left(\widehat{\Theta}_n^{(\mathrm{BC},1)}-\Theta\right)\to\left(6T^{-1}\int_0^T\left(g'(\sigma_s^2)\right)^2\sigma_s^4 ds\right)^{\frac{1}{2}}\times\mathcal{N}(0,1). \quad (52)$$

(ii) When $\alpha=\frac{1}{2}$, we have $\mathcal{F}_T^X$-stably in law that

$$n^{\frac{1}{2}}\left(\widehat{\Theta}_n^{(\mathrm{BC},2)}-\Theta\right)\to\left(T^{-1}\int_0^T\left(g'(\sigma_s^2)\right)^2\times\left(2\sigma_s^4+4\sigma_s^3\sqrt{4v_s+\sigma_s^2}\right)ds\right)^{\frac{1}{2}}\mathcal{N}(0,1).$$

To reflect on the powerfulness of the local approach, the reader can note that the global QMLE is estimating the wrong quantity when g is different from the identity function, except when the volatility is constant. To see why this is the case, we consider the estimation of quarticity (i.e., with $g(x) = x^2$) and we note that a global QMLE would estimate $g(\int_0^T \sigma_t^2 dt)$, which is except when volatility is constant different from $\int_0^T \sigma_t^4 dt$. The extensive empirical work in Andersen, Dobrislav, and Schaumburg (2014) also indicates that the two quantities are very different in practice.

4.3. Estimation of Volatility and Higher Powers of Volatility Incorporating Trading Information

To incorporate all the information available in high frequency data (e.g., in addition to transaction prices, we also observe the trading volume, the type of trade, that is, buyer or seller initiated, more generally bid/ask information from the limit order book), Li, Xie, and Zheng (2016) considered the model where the noise is partially observed through a parametric function

$$Z_{\tau_{i,n},n} = X_{\tau_{i,n}} + \epsilon_{i,n} = X_{\tau_{i,n}} + h(I_{i,n}, \nu) + \tilde{\epsilon}_{i,n},$$

where $I_{i,n}$ is the vector of information at time $\tau_{i,n}$ and $\tilde{\epsilon}_{i,n}$ is the noisy part of the original noise $\epsilon_{i,n}$. See also the related papers Chaker (2017) and Clinet and Potiron (2017, 2018c, 2018d). Here again the observation times are assumed to be regular, that is, $\tau_{i,n} = iT/n$.

The authors assumed that $\tilde{\epsilon}_{i,n}$ is with mean 0, finite SD and that $n\,\mathrm{var}[\tilde{\epsilon}_{i,n}] \to \nu$, which in turn implies that $\tilde{\epsilon}_{i,n} = O_p(1/\sqrt{n})$. To embed this assumption in our LPM framework, there is no harm assuming that

$$\tilde{\epsilon}_{i,n} = n^{-\alpha} \nu^{\frac{1}{2}} \gamma_{\tau_{i,n}},$$

where $\alpha \geq 1/2$ and γ_t are iid with null-mean and unity variance. They estimated ν and the underlying price as

$$\begin{aligned} \widehat{\nu} = \arg\min_{\nu} \frac{1}{2} \sum_{i=1}^{N_n} &((Z_{\tau_{i,n},n} - Z_{\tau_{i-1,n},n}) \\ &-(h(I_{i,n}, \nu) - h(I_{i-1,n}, \nu)))^2, \\ \widehat{X}_{\tau_{i,n}} = Z_{\tau_{i,n},n} &- h(I_{i,n}, \hat{\nu}). \end{aligned}$$

The authors then estimated the integrated volatility with

$$\mathrm{ERV}_{\mathrm{ext}} = \sum_{i=1}^{N_n} (\Delta \widehat{X}_{\tau_{i,n}})^2 + 2 \sum_{i=2}^{N_n} \Delta \widehat{X}_{\tau_{i,n}} \Delta \widehat{X}_{\tau_{i-1,n}},$$

where $\Delta \widehat{X}_{\tau_{i,n}} = \widehat{X}_{\tau_{i,n}} - \widehat{X}_{\tau_{i-1,n}}$, and show the according central limit theory. Under suitable assumptions, they obtain the optimal convergence rate $n^{1/2}$ and the AVAR when $T = 1$

$$\mathrm{AVAR}^{(\mathrm{ERV})} = 6 \int_0^1 \sigma_t^4 dt + 8\nu \int_0^1 \sigma_t^2 dt + 8\nu^2.$$

They also considered another estimator (which they call E-QMLE) which consists in using the QMLE from Xiu (2010), which we considered as a local estimator in Example 4.1, on the estimated observations $\widehat{X}_{\tau_{i,n}}$. They indicated that the E-QMLE might yield a smaller AVAR (see their discussion on p. 38), and they report in their numerical study that its finite sample performance is comparable to $\mathrm{ERV}_{\mathrm{ext}}$ (see Table 2 in p. 41). They did not investigate the corresponding central limit theory.

With the theory provided in our article, we cannot investigate the E-QMLE, but rather the E-(LPE of QMLE), that is, we apply Example 4.1 on $\widehat{X}_{\tau_{i,n}}$. To keep notation of our paper, we denote $\widehat{\Theta}_n$ the E-(LPE of QMLE) estimator of volatility and $\widehat{\Theta}_n^{(\mathrm{BC})}$ its bias-corrected version (i.e., E-(BC LPE of QMLE)). The AVARs obtained in Theorem 5 are the same as in Theorem 3. This is due to the fact that the estimation of ν is very accurate featuring n as a rate of convergence and thus the pre-estimation does not impact the AVAR. This was already the case for the $\mathrm{ERV}_{\mathrm{ext}}$ (see the proof of Theorem 3 in pp. 46–47 of Li, Xie, and Zheng (2016)).

Recalling that the LPE of QMLE is conjectured to be more efficient than the QMLE, in particular this implies that E-(LPE of QMLE) is also conjectured to be more efficient than E-QMLE. In Figure 1, we can see that E-(LPE of QMLE) highly improves the AVAR compared to the $\mathrm{ERV}_{\mathrm{ext}}$. The improvement gets bigger as the noise of $\tilde{\epsilon}_{i,n}$ increases. When setting the volatility and the noise variance as in the setting of the numerical study in Li, Xie, and Zheng (2016), the ratio of AVARS is equal to 0.7. When we further assume no jumps in volatility, this ratio goes to 0.2. When choosing a bigger noise variance 1.44e−07 which remains reasonable, this ratio is lower than 0.01. The overall picture is clearly in favor of the E-(LPE of QMLE). We provide the theorem of this estimator in what follows.

Theorem 5 (E-(LPE of QMLE)). Under Assumption A in Li, Xie, and Zheng (2016, p. 7):

(i) We assume that $\alpha > \frac{1}{2}$. Then, stably in law[8] as $n \to \infty$,

$$n^{\frac{1}{2}} \left(\widehat{\Theta}_n - T^{-1} \int_0^T \sigma_s^2 ds \right) \to \left(6T^{-1} \int_0^T \sigma_s^4 ds \right)^{\frac{1}{2}} \mathcal{N}(0,1). \quad (53)$$

(ii) When $\alpha = \frac{1}{2}$, we have stable convergence in law of

$$n^{\frac{1}{2}} \left(\widehat{\Theta}_n^{(\mathrm{BC})} - \Theta \right) \to \left(T^{-1} \int_0^T \left(2\sigma_s^4 + 4\sigma_s^3 \sqrt{4\nu + \sigma_s^2} \right) ds \right)^{\frac{1}{2}} \mathcal{N}(0,1).$$

We discuss now briefly how to estimate the higher powers of volatility, that is, when $\theta_t^* = g(\sigma_t^2)$ with g not being the identity function. We consider the estimators from Example 4.2. The difference with Example 4.2 is that this estimator is used on the estimated price $\widehat{X}_{\tau_{i,n}}$ based on the information rather than on the raw price. The related theorem is given in what follows.

Theorem 6 (Powers of volatility). Under Assumption A in Li, Xie, and Zheng (2016, p.7):

(i) We assume that $\alpha > \frac{1}{2}$. Then, stably in law as $n \to \infty$,

$$n^{\frac{1}{2}} \left(\widehat{\Theta}_n^{(\mathrm{BC},1)} - \Theta \right) \to \left(6T^{-1} \int_0^T \left(g'(\sigma_s^2) \right)^2 \sigma_s^4 ds \right)^{\frac{1}{2}} \mathcal{N}(0,1). \quad (54)$$

[8] Here and in the following statements, the stable convergence in law is with respect to the filtration considered in Li, Xie, and Zheng (2016).

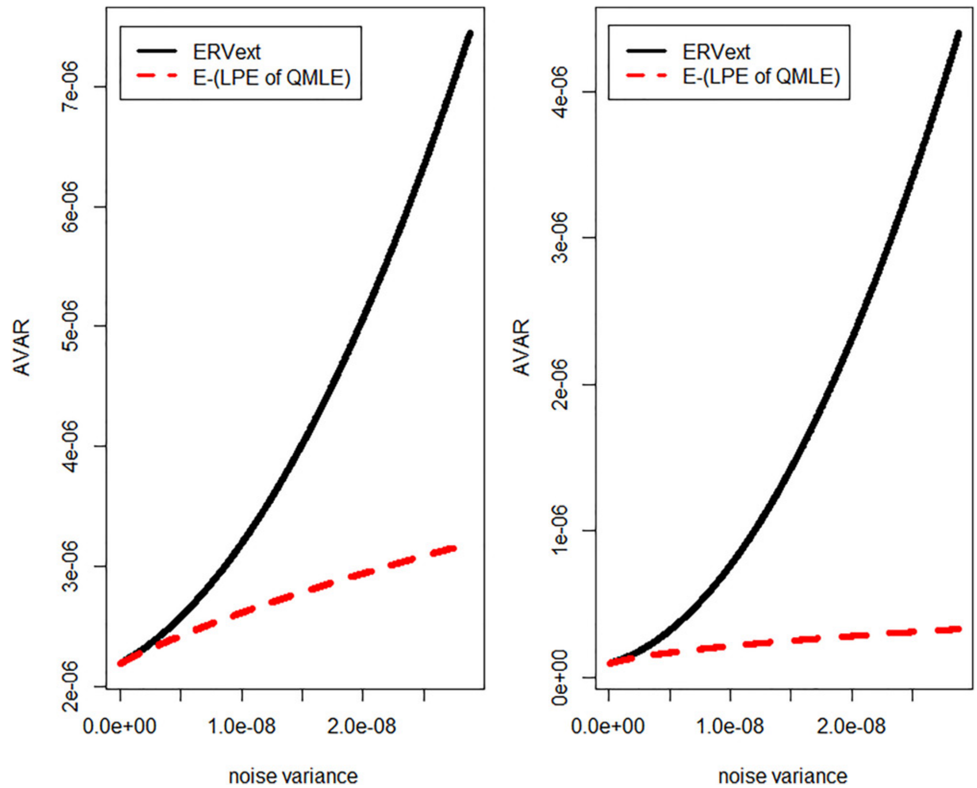

Figure 1. AVAR of $\mathrm{ERV}_{\mathrm{ext}}$ and E-(LPE of QMLE) as a function of the noise variance, that is, the variance of $\tilde{\epsilon}_{i,n}$. The horizon time is set to $T = 1$ (which corresponds to 6.5 hr of intraday trading). On the left hand-side, we follow exactly the setting of the numerical study in Li, Xie, and Zheng (2016), where $\sigma_t^2 = 0.000125$ if $0.05 \leq t < 0.95$ and $\sigma_t^2 = 15 * 0.000125$ otherwise. There is on average one observation a second, which corresponds to $n = 23{,}400$. On the right-hand side, the setting is the same except that we remove the jumps in volatility and consider $\sigma_t^2 = 0.000125$ for $0 \leq t \leq 1$.

(ii) When $\alpha = \frac{1}{2}$, we have

$$n^{\frac{1}{2}}\left(\widehat{\Theta}_n^{(\mathrm{BC},2)} - \Theta\right) \to \left(T^{-1}\int_0^T (g'(\sigma_s^2))^2 \times\left(2\sigma_s^4 + 4\sigma_s^3\sqrt{4v+\sigma_s^2}\right)ds\right)^{\frac{1}{2}} \mathcal{N}(0,1).$$

4.4. Estimation of Volatility Using the Model With Uncertainty Zones

We introduce a time-varying friction parameter extension to the model with uncertainty zones introduced in Robert and Rosenbaum (2011). To incorporate microstructure noise in the model, we define α_n as the tick size, and the related asymptotics is such that $\alpha_n \to 0$. Correspondingly we assume that the observed price $Z_{\tau_{i,n},n}$ takes values on the tick grid (i.e., modulo of size α_n).

We discuss first a simple version of the model with uncertainty zones, which features endogeneity in arrival times. In a frictionless market, we can assume that all the returns (which we recall to be defined as $R_{i,n} = Z_{\tau_{i,n},n} - Z_{\tau_{i-1,n},n}$) have a magnitude of exactly one tick, and that the next transaction will occur when the latent price process crosses the mid-tick value $X_{\tau_{i,n}} + \frac{\alpha_n}{2}$ in case of the price goes up (or $X_{\tau_{i,n}} - \frac{\alpha_n}{2}$ when the price goes down). We extend this toy model in what follows.

The authors introduced the discrete variables $L_{i,n}$ that stands for the absolute size, in tick number, of the next return. In other words, the next observed price has the form $Z_{\tau_{i+1,n},n} = Z_{\tau_{i,n},n} \pm \alpha_n L_{i,n}$. They also introduced a continuous (possibly multidimensional) time-varying parameter χ_t, and assume that conditional on the past, $L_{i,n}$ takes values on $\{1,\ldots,m\}$ with

$$\mathbb{P}_{\tau_{i,n}}(L_{i,n} = k) = p_k(\chi_{\tau_{i,n}})$$

for some unknown positive differentiable with bounded derivative functions p_k such that $\sum_{k=1}^m p_k = 1$.

Also, the frictions induce that the transactions will not occur exactly when the efficient process crosses the mid-tick values. For this purpose, in the notation of Robert and Rosenbaum (2012), let $0 < \eta < 1$ be the parameter that quantifies the aversion to price change. The frictionless scenario corresponds to $\eta = 0$. Conversely, the agents are very averse to trade when η is closer to 1. If we define $X_t^{(\alpha)}$ as the value of X_t rounded to the nearest multiple of α, the sampling times are defined recursively as $\tau_{0,n} := 0$ and for any positive integer i as

$$\tau_{i,n} := \inf\left\{t > \tau_{i-1,n} : X_t = X_{\tau_{i-1,n}}^{(\alpha_n)} - \alpha_n\left(L_{i-1,n} - \frac{1}{2} + \eta\right) \text{ or } X_t = X_{\tau_{i-1,n}}^{(\alpha_n)} + \alpha_n\left(L_{i-1,n} - \frac{1}{2} + \eta\right)\right\}. \quad (55)$$

Correspondingly, the observed price is assumed to be equal to the rounded efficient price $Z_{\tau_{i,n},n} := X_{\tau_{i,n}}^{(\alpha_n)}$.

In the extension of (55) when η_t is time-varying, we assume that the sampling times are defined recursively as $\tau_{i,n} := 0$ and for any positive integer i as

$$\tau_{i,n} := \inf\left\{t > \tau_{i-1,n} : X_t = X^{(\alpha_n)}_{\tau_{i-1,n}} - \alpha_n\left(L_{i,n} - \frac{1}{2} + \eta_{\tau_{i-1,n}}\right)\right. \\ \left.\text{or } X_t = X^{(\alpha_n)}_{\tau_{i-1,n}} + \alpha_n\left(L_{i,n} - \frac{1}{2} + \eta_{\tau_{i-1,n}}\right)\right\}. \tag{56}$$

The idea behind the time-varying friction model with uncertainty zones is that we hold the parameter η_t constant between two observations.

To express the model with uncertainty zones as a LPM, we consider that $\theta^*_t := (\sigma^2_t, \eta_t, \chi_t)$. Following the definition (p. 11) in Robert and Rosenbaum (2012), we further introduce a Brownian motion W'_t independent of all the other quantities, and let Φ denote the cumulative distribution function of a standard Gaussian random variable. We specify the definition of $L_{i,n}$ related to W'_t as

$$g_{t,n} = \sup\{\tau_{j,n} : \tau_{j,n} < t\},$$

$$L'_t = \sum_{k=1}^{m} k\mathbf{1}\left\{\Phi\left(\frac{W'_t - W_{g_{t,n}}}{\sqrt{t - g_{t,n}}}\right) \in \left[\sum_{j=1}^{k-1} p_j(\chi_t), \sum_{j=1}^{k} p_j(\chi_t)\right]\right\},$$

and $L_{i,n} = L'_{\tau_{i,n}}$. If we set the random innovation as the two-dimensional process $U_{i,n} := ((W_t - W_{\tau_{i-1,n}})_{t\geq\tau_{i-1,n}}, ((W'_t - W'_{\tau_{i-1,n}})_{t\geq\tau_{i-1,n}})$, and the past as $P_{\tau_{i,n}} = (L_{i,n}, \text{sign}(R_{i,n}))$, we can deduce the form of F_n in the model.[9]

We provide in what follows the definition of the estimators. We are not interested in estimating directly χ_t and thus we consider the subparameter $\Theta := (\int_0^T \sigma^2_t dt, \int_0^T \eta_t dt)$ to be estimated. For $k = 1, \ldots, m$, we define

$$N^{(a)}_{t,k,n} = \sum_{i=1}^{N_n(t)} \mathbf{1}_{\{R_{i,n}R_{i-1,n}<0\,,\,|R_{i,n}|=k\alpha_n\}},$$

$$N^{(c)}_{t,k,n} = \sum_{i=1}^{N_n(t)} \mathbf{1}_{\{R_{i,n}R_{i-1,n}>0\,,\,|R_{i,n}|=k\alpha_n\}}$$

as, respectively, the number of alternations and continuations of k ticks. By alternation (continuation) of k ticks, we mean that the return magnitude is of k ticks with a direction opposite to (with the same direction as) the previous return. We define the estimator of η as[10]

$$\widehat{\eta}_{t,n} := \sum_{k=1}^{m} \lambda_{t,k,n} u_{t,k,n}, \tag{57}$$

with

$$\lambda_{t,k,n} := \frac{N^{(a)}_{t,k,n} + N^{(c)}_{t,k,n}}{\sum_{j=1}^{m}\left(N^{(a)}_{t,j,n} + N^{(c)}_{\alpha,t,j}\right)},$$

$$u_{t,k,n} := \max\left\{0, \min\left\{1, \frac{1}{2}\left(k\left(\frac{N^{(c)}_{t,k,n}}{N^{(a)}_{t,k,n}} - 1\right) + 1\right)\right\}\right\},$$

where $N_n(t)$ is defined as the integer satisfying $Z_{\tau_{N_n(t),n}} < t < Z_{\tau_{N_n(t)+1,n}}$, we assume that $C/0 := \infty$, and in particular $u_{\alpha,t,k} = 1$ when $N^{(a)}_{\alpha,t,k} = 0$. The key idea is that $u_{\alpha,t,k}$ are consistent estimators of η. Based on the friction parameter estimate, we can construct a consistent latent price estimator as

$$\widehat{X}_{\tau_{i,n}} = Z_{\tau_{i,n},n} - \alpha_n(1/2 - \widehat{\eta}_{t,n})\text{sign}(R_{i,n}).$$

The estimator of integrated volatility is obtained using the usual realized volatility estimator on the estimated price defined as

$$\widehat{RV}_{t,n} = \sum_{i=1}^{N_n(t)} (\widehat{X}_{\tau_{i,n}} - \widehat{X}_{\tau_{i-1,n}})^2.$$

The related local estimators $\widehat{\Theta}_{i,n} := (\widehat{\sigma}^2_{i,n}, \widehat{\eta}_{i,n})$ are constructed from local versions of $\left(\widehat{RV}_{t,n}, \widehat{\eta}_{t,n}\right)$.

Theorem 7 (Time-varying friction parameter model with uncertainty zones). Let $\mathcal{G}_t$ be the filtration generated by X_t, χ_t, and η_t. $\mathcal{G}_T$-stably in law as $n \to \infty$,

$$\alpha_n^{-1}\left(\widehat{\Theta}_n - \Theta\right) \to \left(T^{-1}\int_0^T V_{\theta^*_s} ds\right)^{\frac{1}{2}} \\ \times\mathcal{N}(0, 1), \tag{58}$$

where V_θ can be straightforwardly inferred from the definition of Lemma 4.19 in p. 26 of Robert and Rosenbaum (2012).

Remark 7 (Convergence rate). Note that, equivalently, the convergence rate in (58) is $n^{\frac{1}{2}}$ when n corresponds to the expected number of observations. One can consult Remark 4 in Potiron and Mykland (2017) for more details about this.

4.5. Application in Time Series: The Time-Varying MA(1)

We first specify the LPM for a general one-dimensional time series. In that case, we assume that the observation times are regular. We further assume that the returns $R_{i,n}$ stand for time series observations. Finally, we assume that the time-varying time series can be expressed as the interpolation of θ^*_t via

$$R_{i,n} = F_n\left(\{P_{s,n}\}_{0\leq s\leq\tau_{i-1,n}}, U_{i,n}, \theta^*_{\tau_{i-1,n}}\right), \tag{59}$$

where θ^*_t is assumed to be independent of all the innovations. When θ^*_t is constant, numerous time series[11] are of the form (59).

[9]The advised reader will have noticed that a priori, $\text{sign}(R_{i,n})$ and η_t are not independent, so that the assumptions of the LPM do not hold entirely. This problem can be circumvented as the former is actually conditionally independent from the latter.

[10]Actually, the estimator considered here slightly differs from the original definition (p. 8) in Robert and Rosenbaum (2012) as it provides smaller theoretical finite sample bias. Asymptotically, both estimators are equivalent and thus all the theory provided in Robert and Rosenbaum (2012) can be used to prove Theorem 7.

[11]We can actually show that any time series in state space form can be expressed with a corresponding F_n function.

We now discuss the specific MA(1) representation. Several time-varying extensions are possible and we choose to work with the time-varying parameter model

$$R_{i,n} = \mu_{\tau_{i-1,n}} + \sqrt{\kappa_{\tau_{i-1,n}}}\lambda_{i,n} + \beta_{\tau_{i-1,n}}\sqrt{\kappa_{\tau_{i-1,n}}}\lambda_{i-1,n},$$

where $\lambda_{i,n}$ are standard normally-distributed white noise error terms, and κ_t is the time-varying variance. The three-dimensional parameter is defined as $\theta_t^* := (\mu_t, \beta_t, \kappa_t) \in \mathbb{R}^2 \times \mathbb{R}_*^+$. We fix both the innovation and the past equal to the white noise $U_{i,n} = \lambda_{i,n}$ and $P_{\tau_{i,n},n} = \lambda_{i,n}$. We have thus expressed the MA(1) as a LPM.

We discuss how to estimate the parameters in what follows. For simplicity, we assume that $\mu_t = 0$. The target quantity is thus equal to $\Theta = (\int_0^T \beta_t dt, \int_0^T \kappa_t dt)$. The local estimator is the MLE (see Hamilton 1994, sec. 5.4). On each block (of size h_n), the MLE bias is of order h_n^{-1} (Tanaka 1984) and thus the bias condition (23) is not satisfied. Nonetheless, we can correct for the bias up to the order $O(h_n^{-2})$ as follows. We define the bias-corrected estimator as

$$\widehat{\Theta}_{i,n}^{(\mathrm{BC})} = \widehat{\Theta}_{i,n} - b(\widehat{\Theta}_{i,n}, h_n),$$

where the bias function $b(\theta, h)$ can be derived following the techniques in Tanaka (1984). In particular this implies that the bias-corrected estimator satisfies the bias condition if h_n is chosen such that $n^{1/4} = o(h_n)$. In practice this bias can be obtained by Monte Carlo simulations (see our simulation study).

In the parametric case and in a low frequency asymptotics where $T \to \infty$ and observations times are $0, \Delta, \dots, T = n\Delta$ with $\Delta > 0$, known results (see, e.g., the proof of Proposition I in pp. 391–393 of Aït-Sahalia, Mykland, and Zhang (2005)) show that the AVAR of the MLE is such that

$$n^{1/2}\big((\widehat{\beta},\widehat{\kappa}) - (\beta,\kappa)\big) \to \begin{pmatrix} 1-\beta^2 & 0 \\ 0 & 2\kappa^2 \end{pmatrix}^{1/2} \mathcal{N}(0,1).$$

The following theorem provides the time-varying version of the asymptotic theory when T is fixed.

Theorem 8 (Time-varying MA(1)). Let $\mathcal{F}_t^\theta$ the filtration generated by θ_t^*. We assume that $n^{1/4} = o(h_n)$ and Condition (P2). Then, $\mathcal{F}_T^\theta$-stably in law as $n \to \infty$,

$$n^{\frac{1}{2}}\big(\widehat{\Theta}_n^{(\mathrm{BC})} - \Theta\big) \to \left(T^{-1}\begin{pmatrix} \int_0^T (1-\beta_s^2)ds & 0 \\ 0 & \int_0^T 2\kappa_s^2 ds \end{pmatrix}\right)^{\frac{1}{2}} \times \mathcal{N}(0,1).$$

4.6. Further Examples

Two further examples include our own recent work. Potiron and Mykland (2017) introduced a bias-corrected Hayashi–Yoshida estimator (Hayashi and Yoshida 2005) of the high-frequency covariance and showed the corresponding CLT under endogenous and asynchronous observations. To model duration data, Clinet and Potiron (2018a) built a time-varying Hawkes self-exciting process, derived the bias-corrected MLE and showed the CLT of the corresponding LPE.

4.7. Discussion

We provide in what follows a discussion on the efficiency and robustness of the specific examples considered in this section. The subsequent techniques may also be useful to tackle other examples from the literature.

4.7.1. Efficiency. There are many problems where $n^{1/2}$ is rate-optimal from Gloter and Jacod (2001), such as all the examples considered in this section. In addition, the local feature of the technology should yield efficiency in case the parametric estimator is efficient itself. This is the case of (47) in Example 4.1, Theorem 4(ii) in Example 4.2, Theorem 5(ii) and Theorem 6(ii) in Example 4.3, Theorem 8 in Example 4.5, where the parametric estimator achieves the Cramér–Rao bound of efficiency locally.

In the case of (46) in Example 4.1, that is, when estimating volatility assuming that the noise is very small $\epsilon_{i,n} = o_p(1/\sqrt{n})$, the AVAR is equal to $6T^{-1}\int_0^T \sigma_s^4 ds$, whereas the efficient bound $2T^{-1}\int_0^T \sigma_s^4 ds$ is attained by the RV. This increases the variance by a factor of 3, which is also observed on the MLE (when assuming the volatility is constant) when misspecified on a model which does not incorporate microstructure noise (see, e.g., Barndorff-Nielsen et al. 2008, sec. 2.4, pp. 1486–1487).

4.7.2. Robustness to Drift and Jumps in the Latent Price Process. We focus on the specific case where the observations are related to a latent continuous-Itô price model $dX_t = \int_0^t \sigma_u dW_u$, as in Examples 4.1–4.4 (Example 4.5 considers a time series without any underlying price process). We discuss how we can add a drift and jumps in X_t in those examples.

We first show how to add a drift component. By Girsanov theorem, in conjunction with local arguments (see, e.g., Mykland and Zhang 2012, pp. 158–161), we can weaken the price and volatility local-martingale assumption by allowing them to follow an Itô-process (of dimension 2 in case of volatility or powers of volatility estimation), with a volatility matrix locally bounded and locally bounded away from 0, and drift which is also locally bounded.

It is also easy to see that we can allow for finite activity jumps in X_t. To do that, we assume that $\widehat{\Theta}_{i,n}$ is taking values on a compact set.[12]. Consider $J_n \subset \{1, \dots, B_n\}$ the set of blocks where there is at least one jump in X_t. As the number of blocks $B_n \to \infty$, the cardinality of J_n is at most finite, and thus we have that

$$\frac{1}{T}\sum_{i=1}^{N_n} \widetilde{\Theta}_{i,n}\Delta \mathrm{T}_{i,n} \approx \frac{1}{T}\sum_{i \notin J_n} \widetilde{\Theta}_{i,n}\Delta \mathrm{T}_{i,n}.$$

It is then immediate to adapt the proof of the CLT. On the other hand, if infinitely many jumps are possible in both the price process and the parameter, the theoretical development is beyond the scope of this paper.

4.7.3. Robustness to Jumps in θ_t^*. By a similar reasoning as for when adding jumps in X_t, the techniques of this article are robust to jumps (of finite activity) in θ_t^* in all our examples.

[12]The MLE is always performed on a compact set, so the assumption is trivially satisfied in that case, which corresponds to Examples 4.1–4.3. Moreover, the estimator of η in Example 4.4 is bounded by definition, but one would need to bound the volatility estimator to apply the technique.

4.7.4. Nonregular Observation Times. We also assume here that there is a latent price process and reason about the type of observation times which falls into the LPM. We consider first the time deformation of Barndorff-Nielsen et al. (2008, sec. 5.3, pp. 1505–1507). To express their setting as a LPM, we assume that the observation times are of the form

$$\tau_{i,n} = \Gamma_{i/(nT)}, \tag{60}$$

where Γ_t is a stochastic process satisfying $\Gamma_t = \int_0^t \tilde{\Gamma}_u^2 du$, with $\tilde{\Gamma}_t$ a strictly positive parameter of the LPM. We can then construct a (change of time) process $\tilde{X}_t = X_{\Gamma_t}$ so that for $\tilde{X}_t$ the observations are regular. In view of Dambis Dubins-Schwarz theorem (see, e.g., Revuz and Yor 1999, Theorem 1.6, p. 181) we have that $[X]_T = [\tilde{X}]_{\Gamma_T}$. In addition, it is immediate to see that Condition (T) and (42) hold in that case.

Alternatively one can assume that the quadratic variation of time (see, e.g., Mykland and Zhang 2006, Assumption A, p. 1939) exists and that observation times are independent of the price process. Under suitable assumptions, we can also show that Condition (T) and (42) hold.

Our setting can actually allow for (some kind of) endogenous stopping times as in the case of the model with uncertainty zones detailed in Example 4.4. The type of endogeneity is such that there is no asymptotic bias in the related central limit theorem.

Finally, the model allows for endogenous observation times in the general multidimensional HBT model introduced in Potiron and Mykland (2017). In that case, the central limit theorem features an asymptotic bias.[13]

4.7.5. Estimating Time-Varying Functions of θ_t^*. Another nice corollary about the introduced theory is that we can obtain the central limit theorem of the powers of the integrated parameter $g(t, \theta_t^*)$ for g smooth enough when using the local estimates $g(\mathrm{T}_{i-1,n}, \widehat{\Theta}_{i,n})$. Essentially, the proof uses on each block a Taylor expansion as in the delta method. We apply the technique on the local QMLE in Example 4.2 and on an adapted estimator from Li, Xie, and Zheng (2016) in Example 4.3 to estimate the higher powers of volatility.

5. NUMERICAL STUDY: ESTIMATION OF VOLATILITY WITH THE QMLE

5.1. Goal of the Study

To investigate the finite sample performance of the LPE, we consider the local QMLE introduced in Section 4.1. The goal of the study is 2-fold. First, we want to investigate how the LPE performs compared to the global QMLE. Second, we want to discuss about the choice of the number of blocks B_n in practice. Complementary simulation results can be found in Clinet and Potiron (2018b).

[13]Details about the model can be found in a previous version of the manuscript circulated under the name "Estimating the Integrated Parameter of the Locally Parametric Model in High-Frequency Data."

5.2. Model Design

We perform Monte Carlo simulations of $M = 1000$ days of high-frequency observations where the related horizon time is set to $T = 1/252$ (i.e., annualized). One working day stands for 6.5 hr of trading activity, which can also be expressed as 23,400 sec. We consider three high-frequency sampling frequency scenarios: every second, every other second, and every 3 sec.

We perform local QMLE with number of blocks ranging from $B_n = 1$ (i.e., the global QMLE case) to $B_n = 20$. The corresponding number of observations per block ranges from $h_n = 1170$ to $h_n = 23{,}400$ in the case of 1-sec sampling frequency, from $h_n = 585$ to $h_n = 11{,}700$ if we sample ever other second, and from $h_n = 390$ to $h_n = 7800$ when subsampling every 3 sec. Note that the minimal number of observations per block remains reasonable in view of the finite sample performance of the global QMLE (see the numerical study in Xiu (2010)).

We bring forward the Heston model with U-shape intraday seasonality component and jumps in volatility as

$$dX_t = b\,dt + \sigma_t dW_t,$$
$$\sigma_t = \sigma_{t-,U}\sigma_{t,SV},$$

where

$$\sigma_{t,U} = C + Ae^{-at/T} + De^{-c(1-t/T)} - \beta\sigma_{\tau-,U}\mathbf{1}_{\{t\geq\tau\}},$$
$$d\sigma_{t,SV}^2 = \alpha(\bar{\sigma}^2 - \sigma_{t,SV}^2)dt + \delta\sigma_{t,SV}d\bar{W}_t,$$

where the parameters are set to $b = 0.03$, $C = 0.75$, $A = 0.25$, $D = 0.89$, $a = 10$, $c = 10$, the volatility jump size parameter $\beta = 0.5$, the volatility jump time τ follows a uniform distribution on $[0, T]$, $\alpha = 5$, $\bar{\sigma}^2 = 0.1$, $\delta = 0.4$, $\bar{W}_t$ is a standard Brownian motion such that $d\langle W, \bar{W}\rangle_t = \bar{\phi}dt$, $\bar{\phi} = -0.75$, $\sigma_{0,SV}^2$ is sampled from a Gamma distribution of parameters $(2\alpha\bar{\sigma}^2/\delta^2, \delta^2/2\alpha)$, which corresponds to the stationary distribution of the CIR process. For further reference, see Clinet and Potiron (2018b). The model is almost the same as that of Andersen, Dobrev, and Schaumburg (2012). Finally, the noise is assumed normally distributed with zero-mean and constant variance v set so that the noise to signal ratio defined as

$$\xi^2 = \frac{a_0^2}{\sqrt{T\int_0^T \sigma_u^4 du}} \tag{61}$$

is equal to $\xi^2 = 0.0001$.

5.3. Results

Table 1 reports the sample bias, SD, and the RMSE of the local quasi maximum likelihood volatility estimator. The number of blocks ranges from $B_n = 1$, which corresponds to the global QMLE, to $B_n = 20$. Regardless of the sampling frequency, the numerical experiment results are quite similar. There is a very small sample bias (the bias to SD ratio magnitude is around 0.03), which increases with the number of blocks while staying very small, all of which hinting that the it is not necessary to use a bias correction of the local QMLE in practice. The SD decreases and then stays (roughly)

Table 1. In this table, we report the sample bias ($\times 10^7$), the SD ($\times 10^6$) and the RMSE ($\times 10^6$) for local QMLE with number of blocks ranging from $B_n = 1$ (i.e., the global QMLE case) to $B_n = 20$

Samp. freq.	1 sec.	1 sec.	1 sec.	2 sec.	2 sec.	2 sec.	3 sec.	3 sec.	3 sec.
No. blocks	Bias	SD	RMSE	Bias	SD	RMSE	Bias	SD	RMSE
1	−2.398	8.158	8.162	−2.503	10.813	10.814	−0.492	11.798	11.798
2	−2.614	7.938	7.943	−3.604	10.634	10.640	−0.700	11.642	11.642
3	−2.882	7.820	7.825	−4.041	10.537	10.544	−0.600	11.615	11.615
4	−2.864	7.717	7.722	−4.295	10.500	10.508	−1.210	11.596	11.597
5	−3.181	7.720	7.727	−4.757	10.528	10.539	−1.882	11.587	11.589
6	−3.396	7.695	7.702	−4.918	10.502	10.514	−2.213	11.610	11.612
7	−3.662	7.665	7.674	−5.373	10.523	10.537	−2.919	11.567	11.571
8	−3.561	7.636	7.645	−5.561	10.474	10.489	−3.388	11.601	11.606
9	−4.225	7.636	7.648	−6.344	10.557	10.576	−3.372	11.571	11.576
10	−4.029	7.657	7.668	−6.646	10.536	10.557	−4.400	11.613	11.621
11	−4.503	7.593	7.607	−6.876	10.526	10.548	−5.072	11.638	11.649
12	−4.558	7.634	7.648	−7.495	10.522	10.549	−5.580	11.629	11.642
13	−4.769	7.644	7.659	−8.045	10.548	10.578	−6.485	11.618	11.636
14	−5.058	7.643	7.660	−8.340	10.495	10.529	−7.282	11.533	11.555
15	−5.416	7.591	7.610	−8.394	10.498	10.531	−7.589	11.680	11.704
16	−5.288	7.610	7.629	−8.752	10.491	10.527	−8.452	11.607	11.638
17	−5.638	7.608	7.629	−8.856	10.457	10.494	−8.963	11.619	11.653
18	−5.843	7.604	7.626	−10.093	10.517	10.564	−9.239	11.625	11.661
19	−6.283	7.568	7.594	−10.270	10.499	10.549	−10.611	11.658	11.706
20	−6.109	7.644	7.668	−10.488	10.568	10.620	−10.644	11.603	11.652

NOTE: The number of seconds for one working day is 23,400. The number of Monte Carlo simulations is 1000. Three sampling frequencies are considered: every second, every other second, and every 3 sec.

stable. The picture for the RMSE is the same, all of this very much in line with the fact that almost all the theoretical gain is already obtained in the case of $B = 8$ blocks (see Clinet and Potiron 2018b). Finally, the smallest RMSE is obtained with $B_n = 19$ blocks when sampling at 1-sec frequency, $B_n = 8$ in case of 2-sec frequency and $B_n = 14$ with 3-sec subsampling observations indicating that the finer the sampling frequency the larger the number of blocks should be used. The gains in terms of RMSE goes almost up to 10% when sampling at the finest frequency, whereas less than 5% in the other scenarios.

6. CONCLUSION

In this arcticle, we have introduced a general framework to provide theoretical tools to build central limit theorems of convergence rate $n^{1/2}$ in a time-varying parameter model. We have applied successfully the method to investigate estimation of volatility (possibly under trading information), higher powers of volatility, the time-varying parameters of the model with uncertainty zones and the MA(1). This allowed us to obtain estimators robust to time-varying quantities, more efficient and/or new estimators of quantities (such as in the case of higher powers of volatility).

Subsequently, we believe that many other examples can be solved using the framework of our article, which is simple and natural. This was successfully done in our related papers Potiron and Mykland (2017) and Clinet and Potiron (2018a). In those instances, the regular conditional distribution trick significantly simplified the work of the proofs.

SUPPLEMENTARY MATERIALS

The supplementary materials consist of four distinct sections. First, we investigate consistency in a simple model. Second, the proofs are provided. Third, an additional numerical study, i.e. time-varying MA(1), is explored. Finally, one can find an empirical illustration in the model with uncertainty zones.

ACKNOWLEDGMENTS

We are indebted to the editor, Todd Clark, two anonymous referees, and an anonymous associate editor, Simon Clinet, Takaki Hayashi, Dacheng Xiu, participants of the seminars in Berlin and Tokyo and conferences in Osaka, Toyama, the SoFie annual meeting in Hong Kong, the PIMS meeting in Edmonton for valuable comments, which helped in improving the quality of the paper.

FUNDING

Financial support from the National Science Foundation under grant DMS 14-07812 and Japanese Society for the Promotion of Science Grant-in-Aid for Young Scientists (B) No. 60781119 are greatly acknowledged.

[Received May 2018. Revised August 2018.]

REFERENCES

Aït-Sahalia, Y., Fan, J., Laeven, R., Wang, C. D., and Yang, X. (2017), "Estimation of the Continuous and Discontinuous Leverage Effects," *Journal of the American Statistical Association*, 112, 1744–1758. [680]

Aït-Sahalia, Y., Fan, J., and Xiu, D. (2010), "High-Frequency Covariance Estimates With Noisy and Asynchronous Financial Data," *Journal of the American Statistical Association*, 105, 1504–1517. [685]

Aït-Sahalia, Y., Mykland, P. A., and Zhang, L. (2005), "How Often to Sample a Continuous-Time Process in the Presence of Market Microstructure Noise," *Review of Financial Studies*, 18, 351–416. [680,684,689]

Aït-Sahalia, Y., and Xiu, D. (2019), "A Hausman Test for the Presence of Market Microstructure Noise in High Frequency Data," *Journal of Econometrics* (forthcoming). [680,684]

Aldous, D. J., and Eagleson, G. K., (1978), "On Mixing and sStability of Limit Theorems," *Annals of Probability*, 6, 325–331. [681]

Altmeyer, R., and Bibinger, M. (2015), "Functional Stable Limit Theorems for Quasi-efficient Spectral Covolatility Estimators," *Stochastic Processes and their Applications*, 125, 4556–4600. [685]

Andersen, T. G., Dobrev, D., and Schaumburg, E. (2012), "Jump-Robust Volatility Estimation Using Nearest Neighbor Truncation," *Journal of Econometrics*, 169, 75–93. [690]

Andersen, T. G., Dobrislav, D., and Schaumburg, E. (2014), "A Robust Neighborhood Truncation Approach to Estimation of Integrated Quarticity," *Econometric Theory*, 30, 3–59. [685,686]

Barndorff-Nielsen, O. E., Graversen, S. E., Jacod, J., Podolskij, M., and Shephard, N. (2006), "A Central Limit Theorem for Realised Power and Bipower Variations of Continuous Semimartingales," in *From Stochastic Calculus to Mathematical Finance*, Berlin, Heidelberg: Springer, pp. 33–68. [685]

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, 1481–1536. [689,690]

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, 253–280. [680,685]

Breiman, L. (1992), *Probability*, Classics in Applied Mathematics (Vol. 7), Philadelphia, PA: Society for Industrial and Applied Mathematics (SIAM). [679,681]

Chaker, S. (2017), "On High Frequency Estimation of the Frictionless Price: The Use of Observed Liquidity Variables," *Journal of Econometrics*, 201, 127–143. [686]

Clinet, S., and Potiron, Y. (2017), "Estimation for High-Frequency Data Under Parametric Market Microstructure Noise," arXiv no. 1712.01479. [686]

——— (2018a), "Statistical Inference for the Doubly Stochastic Self-exciting Process," *Bernoulli*, 24, 3469–3493. [689,691]

——— (2018b), "Efficient Asymptotic Variance Reduction When Estimating Volatility in High Frequency Data," *Journal of Econometrics*, 206, 103–142. [684,690,691]

——— (2018c), "Testing If the Market Microstructure Noise Is Fully Explained by the Informational Content of Some Variables From the Limit Order Book," arXiv no. 1709.02502. 686

——— (2018d), "A Relation Between the Efficient, Transaction and Mid Prices: Disentangling Sources of High Frequency Market Microstructure Noise," Working Paper available at SSRN 3167014. [686]

Da, R., and Xiu, D. (2017), "When Moving-Average Models Meet High-Frequency Data: Uniform Inference on Volatility," Working Paper available on Dacheng Xiu's website. [680,684]

Dahlhaus, R. (1997), "Fitting Time Series Models to Nonstationary Processes," *Annals of Statistics*, 25, 1–37. [679,680]

——— (2000), "A Likelihood Approximation for Locally Stationary Processes," *Annals of Statistics*, 28, 1762–1794. [679]

Dahlhaus, R., and Rao, S. S. (2006), "Statistical Inference for Time-Varying ARCH Processes," *Annals of Statistics*, 34, 1075–1114. [679]

Fan, J., and Gijbels, I. (1996), *Local Polynomial Modelling and Its Applications: Monographs on Statistics and Applied Probability* (Vol. 66), Boca Raton, FL: CRC Press. [679]

Fan, J., and Zhang, W. (1999), "Statistical Estimation in Varying Coefficient Models," *Annals of Statistics*, 27, 1491–1518. [679]

Gloter, A., and Jacod, J. (2001), "Diffusions With Measurement Errors. I. Local Asymptotic Normality," *ESAIM: Probability and Statistics*, 5, 225–242. [680,689]

Giraitis, L., Kapetanios, G., and Yates, T. (2014), "Inference on Stochastic Time-Varying Coefficient Models," *Journal of Econometrics*, 179, 46–65. [681]

Hamilton, J. D. (1994), *Time Series Analysis* (Vol. 2), Princeton, NJ: Princeton University Press. [689]

Hall, P., and Heyde, C. C. (1980), *Martingale Limit Theory and Its Application*, Boston, MA: Academic Press. [681]

Hastie T., and Tibshirani, R. (1993), "Varying-Coefficient Models," *Journal of the Royal Statistical Society*, Series B, 55, 757–796. [679]

Hayashi, T., and Yoshida, N. (2005), "On Covariance Estimation of Non-synchronously Observed Diffusion Processes," *Bernoulli*, 11, 359–379. [689]

Jacod, J. (1997), "On Continuous Conditional Gaussian Martingales and Stable Convergence in Law," in *Séminaire de Probabilités XXXI*, pp. 232–246. [681,682]

Jacod, J., Podolskij, M., and Vetter, M. (2010), "Limit Theorems for Moving Averages of Discretized Processes Plus Noise," *Annals of Statistics*, 38, 1478–1545. [685]

Jacod, J., and Protter, P. (1998), "Asymptotic Error Distributions for the Euler Method for Stochastic Differential Equations," *Annals of Probability*, 26, 267–307. [681]

——— (2011), *Discretization of Processes*, Berlin, Heidelberg: Springer. [681,682]

Jacod, J., and Rosenbaum, M. (2013), "Quarticity and Other Functionals of Volatility: Efficient Estimation," *Annals of Statistics*, 41, 1462–1484. [679,685]

Jacod, J., and Shiryaev, A. (2003), *Limit Theorems for Stochastic Processes* (2nd ed.), Berlin: Springer-Verlag. [681,682]

Kim, C. J., and Nelson, C. R. (2006), "Estimation of a Forward-Looking Monetary Policy Rule: A Time-Varying Parameter Model Using Ex Post Data," *Journal of Monetary Economics*, 53, 1949–1966. [679]

Kristensen, D. (2010), "Nonparametric Filtering of the Realized Spot Volatility: A Kernel-Based Approach," *Econometric Theory*, 26, 60–93. [679]

Li, Y., Xie, S., and Zheng, X. (2016), "Efficient Estimation of Integrated Volatility Incorporating Trading Information," *Journal of Econometrics*, 195, 33–50. [680,686,687,690]

Mancino, M. E., and Sanfelici, S. (2012), "Estimation of Quarticity With High-Frequency Data," *Quantitative Finance*, 12, 607–622. [685]

Mykland, P. A., and Zhang, L. (2006), "ANOVA for Diffusions and Itô Processes," *Annals of Statistics*, 34, 1931–1963. [690]

——— (2009), "Inference for Continuous Semimartingales Observed at High Frequency," *Econometrica*, 77, 1403–1445. [679,685]

——— (2011), "The Double Gaussian Approximation for High Frequency Data," *Scandinavian Journal Statistics*, 38, 215–236. [679]

——— (2012), "The Econometrics of High Frequency Data," in *Statistical Methods for Stochastic Differential Equations*, eds. M. Kessler, A. Lindner, and M. Sørensen, Boca Raton, FL: Chapman nad Hall/CRC Press, pp. 109–190. [689]

——— (2017), "Assessment of Uncertainty in High Frequency Data: The Observed Asymptotic Variance," *Econometrica*, 85, 197–231. [683,685]

Podolskij, M., and Vetter, M. (2010), "Understanding Limit Theorems for Semimartingales: A Short Survey," *Statistica Neerlandica*, 64, 329–351. [681]

Potiron, Y., and Mykland, P. A. (2017), "Estimation of Integrated Quadratic Covariation With Endogenous Sampling Times," *Journal of Econometrics*, 197, 20–41. [688,689,690,691]

Reiß, M. (2011), "Asymptotic Equivalence for Inference on the Volatility From Noisy Observations," *Annals of Statistics*, 39, 772–802. [679]

Renault, E., Sarisoy, C., and Werker, B. J. M. (2017), "Efficient Estimation of Integrated Volatility and Related Processes," *Econometric Theory*, 33, 439–478. [685]

Rényi, A. (1963), "On Stable Sequences of Events," *Sanky Series A*, 25, 293–302.

Revuz, D., and Yor, M. (1999), *Continuous Martingales and Brownian Motion* (3rd ed.), Germany: Springer. [681]

Robert, C. Y., and Rosenbaum, M. (2011), "A New Approach for the Dynamics of Ultra-High-Frequency Data: The Model With Uncertainty Zones," *Journal of Financial Econometrics*, 9, 344–366. [690]

——— (2012), "Volatility and Covariation Estimation When Microstructure Noise and Trading Times Are Endogenous," *Mathematical Finance*, 22, 133–164. [680,687]

Rootzén, H., (1980). "Limit Distributions for the Error in Approximations of Stochastic Integrals," *Annals of Probability*, 8, 241–251. [680,687,688]

Stock, J. H., and Watson, M. W. (1998), "Median Unbiased Estimation of Coefficient Variance in a Time-Varying Parameter Model," *Journal of the American Statistical Association*, 93, 349–358. [681]

Tanaka K. (1984), "An Asymptotic Expansion Associated With the Maximum Likelihood Estimators in ARMA Models," *Journal of the Royal Statistical Society*, Series B, 46, 58–67. [679]

Wang, C. D., and Mykland, P. A. (2014), "The Estimation of Leverage Effect With High Frequency Data," *Journal of the American Statistical Association*, 109, 197–215. [689]

Xiu, D. (2010), "Quasi-Maximum Likelihood Estimation of Volatility With High Frequency Data," *Journal of Econometrics*, 159, 235–250. [680] [680,684,685,686,690]