Volatility Regressions with Fat Tails

Jihyun Kim* Nour Meddahi†

Toulouse School of Economics

May 24, 2017

Abstract

Nowadays, a common practice to forecast integrated variance is to do simple OLS auto-regressions of the observed realized variance data. However, non-parametric estimates of the tail index of this realized variance process reveal that its second moment is possibly unbounded. In this case, the behavior of the OLS estimators and the corresponding statistics are unclear. We prove that when the second moment of the spot variance is unbounded, the slope of the spot variance's auto-regression converges to a random variable when the sample size diverges. Likewise, the same result holds when one consider either integrated variance's auto-regression or the realized variance one. We also characterize the connection between these slopes whether the second moment of the spot variance is finite or not. Our theory also allows for a nonstationary spot variance process. We derive the results for the case of several lags in the auto-regressions and multifactor volatility process. A simulation study corroborates our theoretical findings.

Key words: volatility; auto-regression; fat tails; random limits.

*Email: jihyun.kim@tse-fr.eu

†Email: nour.meddahi@tse-fr.eu

1. Introduction

In this paper, we are interested in using high frequency based measures to forecast future variance. A common practice is to approximate the latent daily integrated variance by high-frequency based realized measures like realized variance (Andersen et al. (2001)) or robust-to-noise measures (Zhang et al. (2011); Barndorff-Nielsen et al. (2008); Jacod et al. (2009)) and then estimate by OLS a simple auto-regressive regression of this realized measures to get a forecast of the integrated variance. This auto-regressive regression is often misspecified because the dynamics of the integrated variance is more complex. For instance, if the true instantaneous (or spot) variance is a square-root process, then the integrated and realized variances are ARMA (1,1) processes (Barndorff-Nielsen and Shephard (2002), Meddahi (2003)). Still, even if the auto-regression model is misspecified, it provides a very accurate forecast because integrated variance as well as high-frequency realized measures are persistent and therefore few lags are sufficient to predict well future volatility (Andersen et al. (2004)).

On the other hand, the GARCH era (Engle (1982), Bollerslev (1986)) based on parametric models of daily data provides very useful information about the variance process. One of them which is a primary interest in this paper is fat tails. When one estimates a daily GARCH model on stock returns and exchange rates, one often finds that the returns' fourth moment is not bounded or close to be unbounded. If the fourth moment of the returns is unbound, then the second moment of the daily realize variance defined as the sum of intra-daily squared returns is also unbounded. Consequently, the interpretation of the auto-regressive regression and the OLS estimation are questionable. Likewise, the delivered forecast and all the statistical tools used to assess the quality of the forecast could be not valid. Note also that the same concerns are in place when the fourth moment of the returns is close to be unbounded, that is traditional statistical tools are less reliable when for instance the fifth or sixth moment of the returns is unbounded.

The doubt about the finiteness of the fourth moment of the returns is based on a parametric model of the volatility. In contrast, an important contribution of the high-frequency volatility literature is that the availability of a lot of information allows one to get non-parametric measures of the variance and therefore get rid of these volatility parametric models. It is therefore necessary to assess the finiteness of the second moment of realized measures in a non-parametric way. The solution hinges on a non-parametric estimation of the tail index. We use the Hill's (1975) estimator to our data and we get the same result.

Observe that when one considers a continuous time model without jumps and without market microstructure noise, the fourth moment of the intra-day returns is unbounded if and only if the second moment of the instantaneous variance is infinite.

In this paper, we revisit the results about the auto-regressive regression of the variance process like Andersen et al. (2004) when the the second moment of the spot variance is possibly unbounded, implying that the second moment of integrated and realized variances are unbounded. We also allow for nonstationary stochastic volatility models. When the instantaneous variance is either stationary with an unbounded second moment or nonstationary, then the results in Andersen et al. (2004) are no more valid because one can not compute population auto-regression parameters.

We consider empirical regressions instead of population regressions. More precisely, we study the asymptotic behavior of the OLS estimator of the auto-regressions. We study auto-regressions of three variables: the spot variance, the integrated variance and the realized variance. Of course, the first two auto-regressions are not doable in practice because the variables are not observed, but still the two auto-regressions provide good benchmarks. In particular, the third auto-regression will try to mimic the second one.

We study two types of asymptotic approaches. We start by considering the regression of $V_{(i+1)\Delta}$ on a constant and $V_{i\Delta}$, where V_t is the spot variance process. In the first asymptotic approach, we assume that $\Delta \to 0$ while $T = N\Delta$ is fix. Here, T should be interpreted as the time span of the data. In contrast, Δ is the length of sub-periods. In practice, Δ would be the length of one day while N is the number of days. We do this type of asymptotic because we want to characterize the behavior of the OLS estimators without making a parametric model assumption as did Andersen et al. (2004). The same asymptotic is done when we consider the auto-regression of scaled integrated variance, that is the regression of $\frac{1}{\Delta}\int_{i\Delta}^{(i+1)\Delta} V_s ds$ on a constant and its lagged value $\frac{1}{\Delta}\int_{(i-1)\Delta}^{i\Delta} V_s ds$. The third regression involves the realized variance where an interval of length Δ is divided into equal intervals of length δ. For this case, δ will also shrink to zero in order to ensure that realized variance will converge to the integrated variance.

In this first asymptotic analysis, we characterize the behavior of the OLS estimator of the three regressions slopes. In addition, we characterize the connection between them. We then consider the second asymptotic by allowing $T \to \infty$, that is we consider long span asymptotics and derive various results depending on the spot variance process. When the spot variance process is stationary and has a bounded second moment, we prove that the OLS estimators converge to finite quantities, which are the same ones as the population parameters derived in Andersen et al. (2004). In contrast, when the spot variance is stationary with unbounded second moment, we prove that the OLS estimators converge to random variables. The same result holds when the spot variance process is nonstationary. Both the simulations and the comparison with the results in Andersen et al. (2004) when the spot variance has a finite second moment corroborate the good quality of our approach.

The paper is organized as follows. The next section provides the setup, an empirical motivation for fat tails, and various regressions. The next section studies the asymptotic behavior of the OLS estimators when $\Delta \to 0$. Section 4 studies the long-span asymptotics, that is the limits of the OLS estimators when $T \to \infty$. The following section provides simulations to assess the finite sample properties of the estimators, while the last section concludes. All the proofs are provided in Appendix.

2. Model and Preliminary

2.1. Integrated and Realized Variances

Let $(P_t, 0 \leq t \leq T)$ be a price process given by

$$d\log(P_t) = D_t dt + V_t^{1/2} dW_t^P,$$

where W^P is a Brownian motion, D and V are, respectively, drift and variance processes of P. Then the integrated variance x and realized variance y of P, for a given Δ-period, are defined respectively as, for $i = 1, \cdots, N$ with $N\Delta = T$,

$$x_i = \frac{1}{\Delta}\int_{(i-1)\Delta}^{i\Delta} V_t dt \quad \text{and} \quad y_i^n = \frac{1}{\Delta}\sum_{j=1}^{n}\left(r_{(i-1)\Delta+j\delta}^{(\delta)}\right)^2,$$

where the δ-period return is given by $r^{(\delta)}_{(i-1)\Delta+j\delta} = \log(P_{(i-1)\Delta+j\delta}) - \log(P_{(i-1)\Delta+(j-1)\delta})$, for $j = 1, \cdots, n$ with $n\delta = \Delta$. It is well known that

$$(n/2)^{1/2}(y_i^n - x_i) \to_d \eta_i \mathbb{N}(0,1), \tag{2.1}$$

where $\eta_i^2 = \Delta^{-1}\int_{(i-1)\Delta}^{i\Delta} V_t^2 dt$, as $n \to \infty$ for fixed Δ and for each $i = 1, \cdots, N$. See, e.g., Barndorff-Nielsen and Shephard (2004). Moreover, the convergence (2.1) holds jointly for $i = 1, ..., N$ if $T = N\Delta$ is fixed (see, e.g., Jacod and Protter (1998)).

Instead of analyzing the realized variance y_i^n directly, we simply consider a limiting analogue y_i of y_i^n, where y_i is defined in the following assumption.

Assumption 2.1. *For $i = 1, \cdots, N$, we let $y_i = x_i + \eta_i g_i$, where $\eta_i^2 = \Delta^{-1}\int_{(i-1)\Delta}^{i\Delta} V_t^2 dt$ and g_i is defined as*

$$g_i = \left((2\delta)^{1/2}/\Delta\right)\left(G_{i\Delta} - G_{(i-1)\Delta}\right),$$

where (W, G) is a standard bivariate Brownian motion such that $\mathbb{E}[W_t G_t] = 0$ for all t.

Assumption 2.1 implies

$$(n/2)^{1/2}(y_i - x_i) =_d \eta_i \mathbb{N}(0,1)$$

for $i = 1, \cdots, N$, since $\mathbb{N}(0,1) =_d \left(G_{(i+1)\Delta} - G_{i\Delta}\right)/\sqrt{\Delta}$. In particular, y_i becomes a good proxy of y_i^n for fixed T due to the joint convergence of (2.1). In our asymptotics below, we also allow $T \to \infty$. In this long span case, the joint convergence of (2.1) for growing T is required to approximate y_i^n using y_i uniformly in $i = 1, \cdots, N$.

2.2. Fat Tails: An Empirical Assessment

We will now assess the magnitude of tails of empirical data. We will use trade data on the SPDR S&P 500 ETF (SPY), which is an exchange traded fund (ETF) that tracks the S&P 500 index. Our primary sample comprises 10 years of trade data on SPY starting from June 15, 2004 through June 13, 2014 as available in the New York Stock Exchange Trade and Quote (TAQ) database. This tick-by-tick dataset has been cleaned according to the procedure outlined by Barndorff-Nielsen et al. (2008). We also removed short trading days leaving us with 2497 days of trade data.

We will estimate the tail index of the daily returns and daily realized variance based on five minutes intra-day returns. Because we could have jumps that may affect the tail of the realized variance data, we also consider daily bipower variation which is a consistent estimator of integrated variance under the presence of jumps; see Barndorff-Nielsen and Shephard (2006).

We estimate the tail index by using the Hill's (1975) estimator. Let $(X_i)_{i=1}^n$ be a stationary time series with

$$\mathbb{P}[X_t > x] \sim x^{-\alpha}\ell(x), \quad x \to \infty,$$

where ℓ being a slowly varying function.

The Hill's estimator for α^{-1} which arose in the i.i.d. context as a conditional MLE is defined as

$$h = \frac{1}{k_n}\sum_{i=1}^{k_n} \log(X_{(i)}/X_{(k_n)}),$$

where $(X_{(i)})_{i=1}^n$ is the order statistics $X_{(n)} \leq \cdots \leq X_{(k_n)} \leq \cdots \leq X_{(1)}$ for some $k_n \leq n$ such that $k_n \to \infty$ and $k_n/n \to 0$ as $n \to \infty$.

The results by Hsing (1991) and Resnick and Stărică (1995) indicate that the Hill estimator is asymptotically quite robust with respect to deviations from independence; Resnick and Stărică (1998) prove consistency under ARCH-type dependence. See also Hill (2010) for some other processes including ARFIMA, FIGARCH, explosive GARCH, nonlinear ARMA-GARCH and etc.

Let k_n be the truncation levels for extreme observations. Let $\hat{\alpha}_H = h^{-1}$. Then the standard errors of $\hat{\alpha}_H$ is given by $\sqrt{1/k_n}\hat{\alpha}_H$. Following Ibragimov et al. (2015), who recommended to take k_n in $[0.025 \times n, 0.15 \times n]$, we consider three values: $0.025 \times n$, $0.0875 \times n$ and $0.15 \times n$.

k_n/n	Returns	RV	BPV
0.025	2.569	1.590	1.423
	(0.325)	(0.201)	(0.180)
0.0875	2.125	1.322	1.289
	(0.144)	(0.089)	(0.087)
0.15	1.817	1.176	1.213
	(0.094)	(0.061)	(0.063)

Table 1: Estimated Tail Index

Table 1 provides the estimators of the tail index and the corresponding standard errors. They clearly show that the fourth moment of the returns and the second moment of both realized variance and bipower variation are unbounded.

2.3. Notation

To facilitate our exposition, we need to make some notational conventions. Following the markov process literature, we use the same notation for both a measure and its density with respect to Lebesgue measure. As an example, for a given measure or a density m and a function f on $\mathcal{D} \subset \mathbb{R}$, we write $m(\mathcal{D})$ and $m(f)$ interchangeably with $\int_{\mathcal{D}} m(x)dx$ and $\int_{\mathcal{D}} m(x)f(x)dx$ respectively. The identity function on $\mathbb{R}$ is denoted by ι, i.e., $\iota(x) = x$, and we write the κ-th order power function as ι^κ so that $\iota^\kappa(x) = x^\kappa$. Moreover, for a stochastic process $(V_t, 0 \leq t \leq T)$, we let $T(f) = \max_{0 \leq t \leq T} |f(V_t)|$ for a function f defined on the domain of V. Finally, we use "$P_T \sim_p Q_T$" and "$P_T \sim_d Q_T$" to denote $P_T = Q_T\big(1 + o_p(1)\big)$ and $P_T =_d Q_T\big(1 + o_p(1)\big)$, respectively, as $T \to \infty$.

2.4. Spot Variance

Let $(V_t, 0 \leq t \leq T)$ be a diffusion process on $\mathcal{D} = (\underline{v}, \overline{v}) \subset \mathbb{R}$ driven by

$$dV_t = \mu(V_t)dt + \sigma(V_t)dW_t, \tag{2.2}$$

where W is a Brownian motion, and μ and σ are respectively drift and diffusion functions of V. We let s be the scale function defined as

$$s(v) = \int_y^v \exp\left(-\int_y^x \frac{2\mu(z)}{\sigma^2(z)}dz\right)dx, \tag{2.3}$$

where the lower limits of the integrals can be arbitrarily chosen to be any point $y \in \mathcal{D}$. Defined as such, the scale function s is uniquely identified up to any increasing affine transformation, i.e., if s is a scale function, then so is $as + b$ for any constants $a > 0$ and $-\infty < b < \infty$. We also define the speed density

$$m(v) = \frac{1}{(\sigma^2 s')(v)} \tag{2.4}$$

on $\mathcal{D}$, where s' is the derivative of s, often called the scale density, which is assumed to exist. The speed measure is defined to be the measure on $\mathcal{D}$ given by the speed density with respect to the Lebesgue measure.

Throughout this paper, we assume

Assumption 2.2. *We assume that (a) $\sigma^2(v) > 0$ for all $v \in \mathcal{D}$, and (b) $\mu(v)/\sigma^2(v)$ and $1/\sigma^2(v)$ are locally integrable at every $v \in \mathcal{D}$.*

Assumption 2.2 provides a simple sufficient set of conditions to ensure that a weak solution to the stochastic differential equation (2.2) exists uniquely up to an explosion time. See, e.g., Theorem 5.5.15 in Karatzas and Shreve (1991). Note, under Assumption 2.2, that both the scale function and speed density are well defined, and that the scale function is strictly increasing, on $\mathcal{D}$. Moreover, under Assumption 2.2, the diffusion V is recurrent if and only if the scale function s in (2.3) is unbounded at both boundaries $\underline{v}$ and $\overline{v}$, i.e.,

$$s(\underline{v}) = -\infty \quad \text{and} \quad s(\overline{v}) = \infty.$$

Furthermore, the recurrent diffusion V becomes positive recurrent or null recurrent, depending upon

$$m(\mathcal{D}) < \infty \quad \text{or} \quad m(\mathcal{D}) = \infty,$$

where m is the speed measure defined in (2.4). A diffusion which is not recurrent is said to be transient.

Positive recurrent diffusions are stationary. More precisely, they have time invariant distributions, and if they are started from the time invariant distributions they become stationary. The time invariant density of the positive recurrent diffusion V is given by $\pi(v) = m(v)/m(\mathcal{D})$. Null recurrent and transient diffusions are nonstationary. They do not have time invariant distributions, and their marginal distributions change over time. Out of these two different types of nonstationary processes, we mainly consider null recurrent diffusions in the paper. Brownian motion is the prime example of null recurrent diffusions. Like unit root processes in discrete time, null recurrent processes have stochastic trends and the standard law of large numbers and central limit theory in continuous time are not applicable. See, e.g., Kim and Park (2017) for more details on the statistical properties of null recurrent diffusions.

Let $V^s = s(V)$ be the scale transformation of V, which may be defined as $dV_t^s = m_s^{-1/2}(V_t^s)dW_t$ with speed measure m_s given by

$$m_s = \frac{1}{(s'\sigma)^2 \circ s^{-1}}.$$

Both recurrence and stationarity are preserved under scale transformation. First, V is recurrent on $\mathcal{D}$ if and only if V^s is recurrent on $\mathbb{R}$. Trivially, the scale function of V^s is identity, since it is already in natural scale, and therefore, V^s is recurrent if and only if its domain is given by the entire real line $\mathbb{R}$. However, the domain of V^s becomes $\mathbb{R}$ if and only if V is recurrent, i.e., $s(\underline{v}) = -\infty$ and $s(\overline{v}) = \infty$. Second, V is stationary on $\mathcal{D}$ if and only if V^s is stationary on $\mathbb{R}$, since $m_s(\mathbb{R}) = m(\mathcal{D})$.

2.5. Regressions with Spot, Integrated and Realized Variances

In this paper, we consider the auto-regression of $(z_i)_{i=1}^N$ for $z = v, x, y$, where $v_i = V_{i\Delta}$ is a discrete sample of underlying diffusion V, x and y are, respectively, the integrated and realized variance of V defined in the earier section. Specifically, we consider

$$z_{i+1} = \alpha_z + \beta_z z_i + u_i, \tag{2.5}$$

and test the null hypothesis $\beta = 1$ against the alternative hypothesis $\beta \neq 1$ based on the ordinary least square method. The least square estimator and the t-statistic for β_z in (2.5), denoted respectively as $\hat{\beta}_z$ and $t(\hat{\beta}_z)$, are given by

$$\hat{\beta}_z = \frac{\sum_{i=1}^N (z_i - \overline{z}_N) z_{i+1}}{\sum_{i=1}^N (z_i - \overline{z}_N)^2} \quad \text{and} \quad t(\hat{\beta}_z) = \frac{\hat{\beta}_z - 1}{\hat{\tau}_z \left(\sum_{i=1}^N (z_{i-1} - \overline{z}_N)^2 \right)^{-1/2}},$$

where $\overline{z}_N$ is the sample mean of $(z_i, i = 1, ..., N)$, $\triangle$ is the usual difference operator, i.e., $\triangle z_i = z_{i+1} - z_i$, and $\hat{\tau}_z^2$ is the usual estimator for the variance of regression errors (u_i), i.e., $\hat{\tau}_z^2 = \frac{1}{N} \sum_{i=1}^N \left(z_{i+1} - \hat{\alpha}_z - \hat{\beta}_z z_i \right)^2$. We also consider the usual coefficient R_z^2 of determination for the discrete time regression (2.5).

In the followings, the asymptotic theory for $\hat{\beta}_z$, $t(\hat{\beta}_z)$ and R_z^2 is developed when the spot variance V is either nonstationary or stationary with possibly unbounded variance or mean. Our asymptotics for $z = v, x$ involves two parameters, the sampling interval Δ and the time span T, and it is developed under the assumption that $\Delta \to 0$ and $T \to \infty$ simultaneously. On the other hand, the asymptotics for $z = y$ involves three parameters, the sampling interval Δ at low frequency, the sampling interval δ at high frequency, and the time span T. In this case, the asymptotics is developed under the assumption that $\delta/\Delta \to 0$, $\Delta \to 0$ and $T \to \infty$ simultaneously. For the case of stationarity with bounded variance, the fixed Δ asymptotics are considered in Andersen et al. (2004) using the eigenfunction approach.

2.6. Population Regressions with Spot, Integrated and Realized Variances

In this section, we revisit the results of Andersen et al. (2004) where various regressions in population were studied. These authors considered the Eigenfunction Stochastic Volatility (ESV) model of Meddahi (2001) to derive analytical forecast results. Examples of ESV includes the square-root model, the log-normal stochastic volatility model and the GARCH diffusion model. We will focus here on the GARCH diffusion model of Nelson (1990) because this example allows for unbounded moments while the two other ones lead to bounded ones. More precisely, assume that the spot variance process V_t is given by

$$dV_t = \kappa(\mu - V_t)dt + \sigma V_t dW_t,$$

where W_t is a standard Brownian process, possibly correlated with the Brownian process W_t^p that drives the price process. One can easily prove that the second moment of V_t is bounded when $\sigma^2 < 2\kappa$.

Andersen et al. (2004) computed the population values of the auto-covariances and variances of spot, integrated and realized variances. From these quantities, one gets the corresponding auto-regressive coefficients β_v, β_x and β_y. In particular, one has

$$
\begin{aligned}
\beta_v &= \exp(-\kappa\Delta), \\
\beta_x &= \frac{1}{2}\frac{(1-\exp(-\kappa\Delta))^2}{\exp(-\kappa\Delta)+\kappa\Delta-1}, \\
\beta_y &= \frac{a_1^2}{\Delta^2\kappa^2}\frac{(1-\exp(-\kappa\Delta))^2}{Var(y)},
\end{aligned}
$$

with

$$
Var(y) = 2\frac{a_1^2}{\Delta^2\kappa^2}(\exp(-\kappa\Delta)+\kappa\Delta-1) + \frac{4}{\Delta^2}\frac{\Delta}{\delta}\left(\frac{a_0^2\delta^2}{2} + \frac{a_1^2}{\kappa^2}(\exp(-\kappa\delta)+\kappa\delta-1)\right),
$$

and

$$
a_0 = \mathbb{E}(V_t) = \mu, \quad a_1^2 = Var(V_t) = \mu^2\frac{\psi}{(1-\psi)}, \quad \text{with} \quad \psi = \frac{\sigma^2}{2\kappa}.
$$

One should notice that in this example, the spot variance is an AR(1) process while both integrated and realized variances are ARMA(1,1) processes. In addition, the three processes have the same auto-regressive root which equals $\exp(-\kappa\Delta)$.

When Δ is small, one gets

$$
\begin{aligned}
\beta_v &= 1-\kappa\Delta+o(\Delta), \\
\beta_x &= 1-\frac{2}{3}\kappa\Delta+o(\Delta).
\end{aligned}
$$

Likewise, when both Δ and δ/Δ are small, one gets

$$
\beta_y = 1-\frac{2}{3}\kappa\Delta-2\frac{\delta}{\Delta}\frac{\mathbb{E}(V_t^2)}{Var(V_t)}+o(\Delta)+o(\delta/\Delta).
$$

It is interesting to notice that

$$
\begin{aligned}
\beta_x &= \beta_v+\frac{1}{3}\kappa\Delta+o(\Delta) \\
\beta_y &= \beta_x-2\frac{\delta}{\Delta}\frac{\mathbb{E}(V_t^2)}{Var(V_t)}+o(\Delta)+o(\delta/\Delta),
\end{aligned}
$$

that is, integrated variance has a larger first order auto-correlation than the spot and realized variances.

3. Primary Asymptotics

Recall that $T = N\Delta$ and $\Delta = n\delta$. For our asymptotics here we let $\Delta, \delta/\Delta \to 0$, with T being fixed or $T \to \infty$ simultaneously as $\Delta, \delta/\Delta \to 0$. In case we have $\Delta, \delta/\Delta \to 0$ and $T \to \infty$ simultaneously, we assume that $\Delta, \delta/\Delta \to 0$ sufficiently fast relative to $T \to \infty$. It is indeed more relevant in a majority of practical applications, which rely on observations collected at small sampling intervals over only moderately long span. More precisely, we assume

Assumption 3.1. *We let σ^2 be twice continuously differentiable on $\mathcal{D}$, and we also let $f = \mu, \sigma^2, \sigma^{2'}, \sigma^{2''}, \iota$ be all majorized by a locally bounded function ω on $\mathcal{D}$, for which we have (a) $\Delta T(\omega^4)T^2 \log(T/\Delta) \to_p 0$ and (b) $(\delta/\Delta^{3/2})T(\omega^3)T\sqrt{\log(T/\Delta)} \to_p 0$.*

Assumption 3.1 (a) is similar to Assumption 5.1 in Kim and Park (2017), and provides a sufficient condition for our primary asymptotics of $z = v, x$. On the other hand, the asymptotics for realized variance $z = y$ involves three parameters, and requires Assumption 3.1 (b) in addition to Assumption 3.1 (a). In particular, Assumption 3.1 (b) requires $\delta/\Delta^{3/2} \to 0$, which is more restrictive than $\delta/\Delta \to 0$, regardless of $T = \overline{T}$ and $T \to \infty$. The role of Assumption 3.1 (b) is to analyze the asymptotic effect of the noise $(\eta_i g_i)$ of (y_i) in the least square estimates.

3.1. First Order Auto-Regressions

The following lemma is useful in our primary asymptotics.

Lemma 3.1. *Under Assumption 3.1, we have*

$$(a) \quad \sum_{i=1}^{N}(z_{i+1} - z_i)^2 \sim_p \begin{cases} [V]_T, & \text{for } z = v \\ (2/3)[V]_T, & \text{for } z = x \\ (2/3)[V]_T + (4\delta/\Delta^2)\int_0^T V_t^2 dt, & \text{for } z = y, \end{cases}$$

$$(b) \quad z_N^2 - z_1^2 - \overline{z}_N(z_N - z_1) \sim_p \left(V_T^2 - V_0^2 - \overline{V}_T(V_T - V_0)\right) \quad \text{for } z = v, x, y,$$

$$(c) \quad \sum_{i=1}^{N}(z_i - \overline{z}_N)^2\Delta \sim_p \int_0^T (V_t - \overline{V}_T)^2 dt \quad \text{for } z = v, x, y,$$

where $\overline{V}_T = T^{-1}\int_0^T V_t dt$.

Remark 3.1. (a) The primary asymptotics of $\sum_{i=1}^{N}(z_{i+1}-z_i)^2$ in Lemma 3.1 (a) is depending upon z. In particular, Lemma 3.1 (a) implies

$$\sum_{i=1}^{N}(x_{i+1}-x_i)^2 < \sum_{i=1}^{N}(v_{i+1}-v_i)^2 \quad \text{and} \quad \sum_{i=1}^{N}(x_{i+1}-x_i)^2 < \sum_{i=1}^{N}(y_{i+1}-y_i)^2 \qquad (3.1)$$

with probability approaching one as $\Delta, \delta/\Delta \to 0$ under Assumption 3.1. An intuitive explanation for the inequalities in (3.1) are as follow. We can naturally expect that the integrated variance (x_i) has more smoother sample path compare to its corresponding spot variance (v_i). As a result, the sum of squared increments of (x_i) tends to be smaller than that of (v_i), and we have the first inequality in (3.1). On the other hand, the additional noise $(\eta_i g_i)$ in the realized variance (y_i) generates additional variation, and hence, the sample path of (y_i) becomes more rough compare to the integrated variance (x_i). Therefore, we have the second inequality in (3.1).

(b) Unlike Lemma 3.1 (a), each result in Lemma 3.1 (b) and (c) is identical for all $z = v, x, y$ under Assumption 3.1. The results in Lemma 3.1 (b) and (c) are well expected since $|z_i - V_{(i-1)\Delta}| \to_p 0$ for all z as long as δ/Δ and Δ are sufficiently small relative to T.

(c) It follows from Lemma 3.1 and Ito's lemma that

$$\sum_{i=1}^{N}(z_i-\overline{z}_N)\triangle z_i \sim_p \begin{cases} \int_0^T (V_t-\overline{V}_T)dV_t, & \text{for } z=v \\ \int_0^T (V_t-\overline{V}_T)dV_t + (1/6)[V]_T, & \text{for } z=x \\ \int_0^T (V_t-\overline{V}_T)dV_t + (1/6)[V]_T - (2\delta/\Delta^2)\int_0^T V_t^2 dt, & \text{for } z=y \end{cases} \qquad (3.2)$$

as $\Delta, \delta/\Delta \to 0$ under Assumption 3.1, since

$$\sum_{i=1}^{N}(z_i-\overline{z}_N)\triangle z_i = \frac{1}{2}\left(\left(z_N^2 - z_1^2 - \overline{z}_N(z_N-z_1)\right) - \sum_{i=1}^{N}(z_{i+1}-z_i)^2\right).$$

The result (3.2) for $z = v$ is quite natural and expectable by the asymptotic negligibility of discretization errors when $\Delta \to 0$. In a similar argument, one may expect

$$\sum_{i=1}^{N}(z_i - \overline{z}_N)\triangle z_i \sim_p \sum_{i=1}^{N}(v_i - \overline{v}_N)\triangle v_i \sim_p \int_0^T (V_t - \overline{V}_T)dV_t \quad \text{for} \quad z = x, y \tag{3.3}$$

since $\sup_{0\leq i\leq N} |z_i - v_i| \to_p 0$ as $\delta/\Delta, \Delta \to 0$. In this case, however, our result in (3.2) is different from the conjecture (3.3). This is not surprising at all since the convergence of stochastic process does not necessarily imply the convergence of stochastic integral associated with the stochastic process. Here, in particular, $|z_i\triangle z_i - v_i\triangle v_i| \not\to_p 0$ as $\delta/\Delta, \Delta \to 0$. The reader is referred to see Kurtz and Protter (1991) for more detailed discussions about the weak convergence of stochastic integrals.

Now we are to show the primary asymptotics for $\hat{\beta}_z$ and $t(\hat{\beta}_z)$. To effectively explain our asymptotics, it is useful to note that

$$\hat{\beta}_z - 1 = \frac{\sum_{i=1}^{N}(z_i - \overline{z}_N)\triangle z_i}{\sum_{i=1}^{N}(z_i - \overline{z}_N)^2} = \frac{1}{2}\frac{(z_N^2 - z_1^2 - \overline{z}_N(z_N - z_1)) - \sum_{i=1}^{N}(z_{i+1} - z_i)^2}{\sum_{i=1}^{N}(z_i - \overline{z}_N)^2}. \tag{3.4}$$

We then can easily obtain the primary asymptotics for $\hat{\beta}_z$ and $t(\hat{\beta}_z)$ from Lemma 3.1 and (3.4) with Ito's lemma.

Proposition 3.2. *Let Assumption 3.1 hold.*

(a) For $\hat{\beta}_z$, we have

$$\begin{aligned}
\hat{\beta}_v &\sim_p 1 + \Delta\frac{\int_0^T (V_t - \overline{V}_T)dV_t}{\int_0^T (V_t - \overline{V}_T)^2 dt}, \\
\hat{\beta}_x &\sim_p 1 + \Delta\frac{\int_0^T (V_t - \overline{V}_T)dV_t + (1/6)[V]_T}{\int_0^T (V_t - \overline{V}_T)^2 dt}, \\
\hat{\beta}_y &\sim_p 1 + \Delta\frac{\int_0^T (V_t - \overline{V}_T)dV_t + (1/6)[V]_T - (2\delta/\Delta^2)\int_0^T V_t^2 dt}{\int_0^T (V_t - \overline{V}_T)^2 dt}.
\end{aligned}$$

(b) For $z = v, x, y$, we have

$$(T/\Delta)\hat{\tau}_z^2 \sim_p \sum_{i=1}^{N}(z_{i+1} - z_i)^2 \quad \textit{and} \quad R_z^2 \sim_p \hat{\beta}_z^2.$$

(c) For $t(\hat{\beta}_z)$, we have

$$t(\hat{\beta}_v) \sim_p \sqrt{T}\frac{\int_0^T (V_t - \overline{V}_T)dV_t}{[V]_T^{1/2}\left(\int_0^T (V_t - \overline{V}_T)^2 dt\right)^{1/2}},$$

$$t(\hat{\beta}_x) \sim_p \sqrt{T}\frac{\int_0^T (V_t - \overline{V}_T)dV_t + (1/6)[V]_T}{((2/3)[V]_T)^{1/2}\left(\int_0^T (V_t - \overline{V}_T)^2 dt\right)^{1/2}},$$

$$t(\hat{\beta}_y) \sim_p \sqrt{T}\frac{\int_0^T (V_t - \overline{V}_T)dV_t + (1/6)[V]_T - (2\delta/\Delta^2)\int_0^T V_t^2 dt}{\left((2/3)[V]_T + (4\delta/\Delta^2)\int_0^T V_t^2 dt\right)^{1/2}\left(\int_0^T (V_t - \overline{V}_T)^2 dt\right)^{1/2}}.$$

Remark 3.2. (a) If we define

$$\gamma_v = \frac{\int_0^T (V_t - \overline{V}_T)dV_t}{\int_0^T (V_t - \overline{V}_T)^2 dt}, \quad \gamma_x = \frac{(1/6)[V]_T}{\int_0^T (V_t - \overline{V}_T)^2 dt}, \quad \gamma_y = \frac{2\int_0^T V_t^2 dt}{\int_0^T (V_t - \overline{V}_T)^2 dt},$$

then it follows from Proposition 3.2 (a) that

$$\hat{\beta}_v \sim_p 1 + \Delta\gamma_v, \quad \hat{\beta}_x \sim_p 1 + \Delta(\gamma_v + \gamma_x), \quad \hat{\beta}_y \sim_p 1 + \Delta(\gamma_v + \gamma_x) - (\delta/\Delta)\gamma_y.$$

In particular, we have $\hat{\beta}_y < \hat{\beta}_x$ and $\hat{\beta}_v < \hat{\beta}_x$ with probability approaching one as $\Delta, \delta/\Delta \to 0$ under Assumption 3.1.

(b) Note that Assumption 3.1 (b) does not necessarily imply $\delta/\Delta^2 \to 0$. Therefore, the speeds of $\delta \to 0$ and $\Delta \to 0$ are important in the asymptotics of $\hat{\beta}_y$. For instance, if $\delta/\Delta^2 \to 0$ sufficiently quickly, then $\hat{\beta}_y \sim_p \hat{\beta}_x$. Otherwise, $\hat{\beta}_y \not\sim_p \hat{\beta}_x$.

3.2. Extensions

As extensions of the first order auto-regression (2.5), we consider the following two regressions

$$z_{i+1} = \alpha_z + \beta_z^{(k)} z_{i-k} + u_i \quad \text{for some} \quad k \geq 1 \tag{3.5}$$

$$z_{i+1} = \alpha_z + \beta_{1,z} z_i + \beta_{2,z} z_{i-1} + u_i. \tag{3.6}$$

The regression (3.5) is a multi-lag auto-regression, and the regression (3.6) is a second order auto-regression. Below we analyze each regression separately.

We first consider the multi-lag auto-regression (3.5). We define $\hat{\beta}_z^{(k)}$ and R_z^2 as before. We then have

Proposition 3.3. *Let Assumption 3.1 hold.*

$$(a) \quad \sum_{i=k+1}^{N} (z_{i+1} - z_{i-k})^2 \sim_p \begin{cases} (1+k)[V]_T, & \text{for } z = v \\ (2/3+k)[V]_T, & \text{for } z = x \\ (2/3+k)[V]_T + (4\delta/\Delta^2)\int_0^T V_t^2 dt, & \text{for } z = y, \end{cases}$$

$$(b) \quad \hat{\beta}_z^{(k)} \sim_p \hat{\beta}_z + k(\hat{\beta}_v - 1) \quad \text{for } z = v, x, y,$$

$$(c) \quad R_z^2 \sim_p (\hat{\beta}_z^{(k)})^2 \quad \text{for } z = v, x, y.$$

Remark 3.3. (a) For Proposition 3.3 (a), it is shown in the proof that

$$\begin{aligned} \sum_{i=k+1}^{N} (z_{i+1} - z_{i-k})^2 &= \sum_{i=k+1}^{N} \Big((z_{i+1} - v_i) + (v_i - v_{i-k}) + (v_{i-k} - z_{i-k}) \Big)^2 \\ &\sim_p \underbrace{\sum_{i=k+1}^{N} (z_{i+1} - v_i)^2 + \sum_{i=k+1}^{N} (v_{i-k} - z_{i-k})^2}_{\sim_p \sum_{i=k+1}^{N} (z_{i+1}-z_i)^2} + \underbrace{\sum_{i=k+1}^{N} (v_i - v_{i-k})^2}_{\sim_p k\sum_{i=1}^{N}(v_i - v_{i-1})^2}. \end{aligned} \tag{3.7}$$

The stated result in Proposition 3.3 (a) follows immediately from (3.7) with Lemma 3.1 (a).

(b) As in (3.4), we have

$$\begin{aligned} &\sum_{i=k+1}^{N} (z_{i-k} - \overline{z}_N)(z_{i+1} - z_{i-k}) \\ &= \frac{1}{2} \sum_{j=0}^{k} \Big((z_{N+1+j}^2 - z_{1+k-j}^2) - \overline{z}_N (z_{N+1+j} - z_{1+k-j}) \Big) - \frac{1}{2} \sum_{i=k+1}^{N} (z_{i+1} - z_{i-k})^2. \end{aligned}$$

Moreover, we can deduce from (3.7) and Lemma 3.1 (b) that

$$\begin{aligned}
&\frac{1}{2}\sum_{j=0}^{k}\Big((z_{N+1+j}^2 - z_{1+k-j}^2) - \overline{z}_N(z_{N+1+j} - z_{1+k-j})\Big) - \frac{1}{2}\sum_{i=k+1}^{N}(z_{i+1} - z_{i-k})^2, \\
&\sim_p \frac{1}{2}\Big(z_{N+1}^2 - z_{1+k}^2 - \overline{z}_N(z_{N+1} - z_{1+k}) - \sum_{i=k+1}^{N}(z_{i+1} - z_i)^2\Big) \\
&\qquad + \frac{k}{2}\Big(v_N^2 - v_0^2 - \overline{v}_N(v_N - v_0) - \sum_{i=1}^{N}(v_i - v_{i-1})^2\Big).
\end{aligned} \tag{3.8}$$

It then follows from (3.8) and Lemma 3.1 that

$$\begin{aligned}
\hat{\beta}_z^{(k)} - 1 &\sim_p \frac{1}{2}\left(\frac{z_{N+1}^2 - z_{1+k}^2 - \overline{z}_N(z_{N+1} - z_{1+k}) - \sum_{i=k+1}^{N}(z_{i+1} - z_i)^2}{\sum_{i=k+1}^{N}(z_{i+1} - \overline{z}_N)^2}\right) \\
&\quad + \frac{k}{2}\left(\frac{v_N^2 - v_0^2 - \overline{v}_N(v_N - v_0) - \sum_{i=1}^{N}(v_i - v_{i-1})^2}{\sum_{i=k+1}^{N}(z_{i+1} - \overline{z}_N)^2}\right) \\
&\sim_p (\hat{\beta}_z - 1) + k(\hat{\beta}_v - 1).
\end{aligned}$$

Therefore, we can verify that the additional term, $k(\hat{\beta}_v - 1)$, in $\hat{\beta}_z^{(k)}$ is induced by $\sum_{i=k+1}^{N}(v_i - v_{i-k})^2$ in (3.7).

Now we consider the second order auto-regression (3.6), and define the least square estimator $(\hat{\beta}_{1,z}, \hat{\beta}_{2,z})'$ as

$$\begin{pmatrix} \hat{\beta}_{1,z} \\ \hat{\beta}_{2,z} \end{pmatrix} = \left(\sum_{i=2}^{N} w_i w_i'\right)^{-1}\left(\sum_{i=2}^{N} w_i \tilde{z}_{i+1}\right),$$

where $w_i = (\tilde{z}_i, \tilde{z}_{i-1})'$ with the demeaned series $(\tilde{z}_i, \tilde{z}_{i-1})_{i=2}^N$ for $(z_i, z_{i-1})_{i=2}^N$, and we also define R_z^2 correspondingly.

Proposition 3.4. *Let Assumption 3.1 hold. For $z = v, x, y$, we have the followings.*

(a) $\frac{1}{\sum_{i=2}^N \tilde{z}_i^2}\left(\sum_{i=2}^N w_i \tilde{z}_{i+1}\right) \sim_p \begin{pmatrix} \hat{\beta}_z \\ \hat{\beta}_z^{(1)} \end{pmatrix}$ *and* $\frac{1}{\sum_{i=2}^N \tilde{z}_i^2}\left(\sum_{i=2}^N w_i w_i'\right) \sim_p \begin{pmatrix} 1 & \hat{\beta}_z \\ \hat{\beta}_z & 1 \end{pmatrix}$,

(b) $\hat{\beta}_{1,z} \sim_p \frac{\hat{\beta}_z}{\hat{\beta}_z + 1} + \frac{\hat{\beta}_z}{\hat{\beta}_z + 1}\frac{\hat{\beta}_v - 1}{\hat{\beta}_z - 1}$ *and* $\hat{\beta}_{2,z} \sim_p \frac{\hat{\beta}_z}{\hat{\beta}_z + 1} - \frac{1}{\hat{\beta}_z + 1}\frac{\hat{\beta}_v - 1}{\hat{\beta}_z - 1}$,

(c) $\hat{\beta}_{1,z} + \hat{\beta}_{2,z} \sim_p \hat{\beta}_z - \Delta\Gamma_z$,

(d) $R_z^2 \sim_p (\hat{\beta}_{1,z} + \hat{\beta}_{2,z})^2 + 2\hat{\beta}_{1,z}\hat{\beta}_{2,z}(\hat{\beta}_z - 1)$,

where $\Gamma_v = 0$, $\Gamma_x = \gamma_x/2$ *and* $\Gamma_y = \gamma_x/2 - (\delta/\Delta^2)\gamma_y/2$.

Remark 3.4. (a) Proposition 3.4 (c), together with Proposition 3.2 (a), implies that

$$\begin{aligned} \hat{\beta}_{1,v} + \hat{\beta}_{2,v} &\sim_p 1 + \Delta\gamma_v \\ \hat{\beta}_{1,x} + \hat{\beta}_{2,x} &\sim_p 1 + \Delta\gamma_v + \Delta\gamma_x/2 \\ \hat{\beta}_{1,y} + \hat{\beta}_{2,y} &\sim_p 1 + \Delta\gamma_v + \Delta\gamma_x/2 - (\delta/\Delta)\gamma_y/2, \end{aligned}$$

and therefore, $\hat{\beta}_{1,v} + \hat{\beta}_{2,v} \sim_p \hat{\beta}_v$ and

$$\hat{\beta}_{1,x} + \hat{\beta}_{2,x} < \hat{\beta}_x \quad \text{and} \quad \hat{\beta}_{1,y} + \hat{\beta}_{2,y} < \hat{\beta}_y$$

with probability approaching one as $\Delta, \delta/\Delta \to 0$ under Assumption 3.1.

(b) As in Proposition 3.2 (see also Remark 3.2 (b)), the speeds of $\delta \to 0$ and $\Delta \to 0$ are important in the asymptotics of $\hat{\beta}_{1,y}$ and $\hat{\beta}_{2,y}$, and we have $\hat{\beta}_{1,y} \sim_p \hat{\beta}_{1,x}$ and $\hat{\beta}_{2,y} \sim_p \hat{\beta}_{2,x}$ as long as $\delta/\Delta^2 \to 0$ sufficiently quickly. Otherwise, neither $\hat{\beta}_{1,y} \sim_p \hat{\beta}_{1,x}$ nor $\hat{\beta}_{2,y} \sim_p \hat{\beta}_{2,x}$.

(c) In Proposition 3.2, we consider only the first order approximations for $\hat{\beta}_z$. However, the higher order approximation is important in the asymptotics for each of $\hat{\beta}_{1,z}$ and $\hat{\beta}_{2,z}$. To see this, we let γ_z' be the second order approximation term such that

$$\begin{aligned} \hat{\beta}_v &= 1 + \Delta\gamma_v + \Delta^2\gamma_v' + o_p(\Delta^2), \\ \hat{\beta}_x &= 1 + \Delta(\gamma_v + \gamma_x) + \Delta^2\gamma_x' + o_p(\Delta^2), \\ \hat{\beta}_y &= 1 + \Delta(\gamma_v + \gamma_x) - (\delta/\Delta)\gamma_y + \Delta^2\gamma_y' + o_p(\Delta^2). \end{aligned} \tag{3.9}$$

By applying Taylor expansion to Proposition 3.4 (b) with (3.9), we may show that

$$\hat{\beta}_{1,z} = \frac{\hat{\beta}_z + 1}{2} - \Gamma_{1,z} + \Delta\Gamma_{2,z} - \Delta\Gamma_{3,z} + o_p(\Delta)$$
$$\hat{\beta}_{2,z} = \frac{\hat{\beta}_z - 1}{2} + \Gamma_{1,z} - \Delta\Gamma_{2,z} - \Delta\Gamma_{3,z} + o_p(\Delta),$$

where $\Gamma_{1,v} = \Gamma_{2,v} = \Gamma_{3,v} = 0$ and

$$\Gamma_{1,x} = \frac{\gamma_x}{2(\gamma_v + \gamma_x)}, \qquad \Gamma_{1,y} = \frac{\gamma_x - (\delta/\Delta^2)\gamma_y}{2(\gamma_v + \gamma_x - (\delta/\Delta^2)\gamma_y)},$$
$$\Gamma_{2,x} = \frac{\gamma_v'}{2(\gamma_v + \gamma_x)} - \frac{\gamma_v \gamma_x'}{2(\gamma_v + \gamma_x)^2}, \quad \Gamma_{2,y} = \frac{\gamma_v'}{2(\gamma_v + \gamma_x - (\delta/\Delta^2)\gamma_y)} - \frac{\gamma_v \gamma_y'}{2(\gamma_v + \gamma_x - (\delta/\Delta^2)\gamma_y)^2},$$
$$\Gamma_{3,x} = \frac{\gamma_x}{4}, \qquad \Gamma_{3,y} = \frac{\gamma_x}{4} - \frac{\delta}{\Delta^2}\frac{\gamma_y}{4}.$$

Clearly, $\Gamma_{2,x}$ and $\Gamma_{2,y}$ have the first order effects in the approximations of $(\hat{\beta}_{1,x}, \hat{\beta}_{2,x})$ and $(\hat{\beta}_{1,y}, \hat{\beta}_{2,y})$, respectively. Moreover, $\Gamma_{2,x}$ and $\Gamma_{2,y}$ contain the second order approximation term γ_z' of $\hat{\beta}_z$, and hence, we may conclude that the second order approximation term γ_z' has a first order effect in both $\hat{\beta}_{1,z}$ and $\hat{\beta}_{2,z}$. In this paper, we don't consider the higher order approximation terms explicitly, and we leave them for future research.

(d) Unlike $\hat{\beta}_{1,z}$ and $\hat{\beta}_{2,z}$, it can be shown that the higher order approximation terms do not have the first order effects on both $\hat{\beta}_{1,z} + \hat{\beta}_{2,z}$ and R_z^2.

4. Long Span Asymptotics

We now consider the long span property of the least square estimates by letting $T \to \infty$ under Assumption 3.1. In addition to Assumptions 2.1 and 3.1, we assume the followings.

Assumption 4.1. *We assume that (a) s' is regularly varying or rapidly varying with index $\kappa \neq -1$, (b) σ^2 is regularly varying and (c) m is either integrable or regularly varying.*

Assumption 4.1 is identical to Assumption 2.2 in Kim and Park (2016), and is mild enough to include most diffusion processes used in practice. The reader is also referred to see Bingham et al. (1993) for more discussions about the regular and rapid variations.

In the following, we let $f_s = f \circ s^{-1}$ for any function f other than m. We may easily show by a change of variables in integrals that $m_s(f_s) = m(f)$ for any f defined on $\mathcal{D}$. Moreover, for a regularly varying function f on $\mathbb{R}$, we define its limit homogeneous function $\overline{f}$ as $f(\lambda v)/f(\lambda) \to \overline{f}(v)$ as $\lambda \to \infty$ for all $v \neq 0$. Finally, for a locally integrable f on $\mathbb{R}$, we define $[f]$ as $[f](\lambda) = \int_{|v|<\lambda} f(v)dv$. This notation will be used without reference in what follows.

4.1. Basic Asymptotics

The long span asymptotics in this paper is closely related to the mean reversion property in Kim and Park (2016). We follow their approach, and consider the following three conditions.

(ST) m is either integrable or nearly integrable,

(DD) $1/s'$ is either integrable or nearly integrable, and

(SI) ι^2 is either m-integrable or m-nearly integrable.

As in Kim and Park (2017), we say that a null recurrent diffusion V is *strongly nonstationary* if m is strongly nonintegrable, and it is *nearly stationary* if m is nearly integrable. Therefore, the condition ST holds if and only if V is either stationary or nearly stationary. See Kim and Park (2016, 2017) for more detailed discussions about the strong nonstationarity and near stationarity. The condition DD is about the m-integrability of σ^2 since $m\sigma^2 = 1/s'$ by the definition of speed density (2.4). It is shown in Kim and Park (2016) that under the condition DD the drift term dominates the diffusion term in the following sense

$$\int_0^T (V_t - \overline{V}_T)dV_t = \int_0^T (V_t - \overline{V}_T)\mu(V_t)dt + \int_0^T (V_t - \overline{V}_T)\sigma(V_t)dW_t \sim_p \int_0^T (V_t - \overline{V}_T)\mu(V_t)dt$$

as $T \to \infty$. Lastly, the condition SI is about the m-square integrability of the linear function ι, and it implies that V is a stationary diffusion satisfying $\int_0^T V_t^2 dt/(T\ell(T)) = O_p(1)$ for some slowly varying function ℓ. In particular, if ι^2 is m-integrable, then V has a finite second moment and $\int_0^T V_t^2 dt/T \to_p \mathbb{E}(V_t^2)$ as $T \to \infty$. The reader is referred to see Kim and Park (2017) for general asymptotics of diffusion functionals.

We let (λ_T) be the normalizing sequence satisfying

$$T = \lambda_T[m_s](\lambda_T) \quad \text{or} \quad \lambda_T^2 m_s(\lambda_T) \tag{4.1}$$

depending upon whether or not ST holds. In case either ST or DD holds, we subsequently define

$$a_T = \begin{cases} \lambda_T[m_s\sigma_s^2](\lambda_T) \\ \lambda_T^2(m_s\sigma_s^2)(\lambda_T) \end{cases} \text{if} \quad \begin{array}{l} \text{DD holds} \\ \text{DD does not hold and ST holds} \end{array}$$

$$b_T = \begin{cases} \lambda_T[m_s\iota_s^2](\lambda_T) \\ \lambda_T^2(m_s\iota_s^2)(\lambda_T) \end{cases} \text{if} \quad \begin{array}{l} \text{SI holds} \\ \text{SI does not hold} \end{array}$$

from (λ_T), and let

$$P = \begin{cases} L(\tau,0) \\ \int_0^\tau \overline{m_s\sigma_s^2}(B_t)dt \end{cases} \quad \text{if} \quad \begin{cases} \text{DD holds} \\ \text{DD does not hold and ST holds} \end{cases}$$

$$Q_1 = \begin{cases} \left((\pi(\iota))^2/\pi(\iota^2)\right)^{1/2} \\ 0 \\ \int_0^\tau \overline{m_s\iota_s}(B_t)dt \end{cases} \quad \text{if} \quad \begin{cases} \text{SI holds} \\ \text{SI does not hold and ST holds} \\ \text{DD holds and ST does not hold} \end{cases}$$

$$Q_2 = \begin{cases} 1 \\ \int_0^\tau \overline{m_s\iota_s^2}(B_t)dt \end{cases} \quad \text{if} \quad \begin{cases} \text{SI holds} \\ \text{SI does not hold} \end{cases}$$

where $\pi = m/m(\mathcal{D})$ is the time invariant density of V, and τ is a stopping time defined as

$$\tau = \inf\left\{t \,\middle|\, L(t,0) > 1\right\} \quad \text{or} \quad \inf\left\{t \,\middle|\, \int_{\mathbb{R}} L(t,x)\overline{m}_s(dx) > 1\right\}, \tag{4.2}$$

depending upon whether or not ST holds, from the local time L of Brownian motion B. Numerical sequences (a_T, b_T) and random variables (P, Q_1, Q_2) introduced here will be used repeatedly in what follows.

Lemma 4.1. *Let Assumption 4.1 hold. If either* ST *or* DD *holds, we have* $Ta_T/b_T \to$

∞ *and*

$$\frac{1}{a_T}[V]_T \to_d P, \qquad \frac{1}{a_T}\int_0^T (V_t - \overline{V}_T)dV_t \to_d -\frac{P}{2},$$
$$\frac{1}{b_T}\int_0^T V_t^2 dt \to_d Q_2, \qquad \frac{1}{b_T}\int_0^T (V_t - \overline{V}_T)^2 dt \to_d Q_2 - Q_1^2 \equiv Q$$

as $T \to \infty$.

Remark 4.1. (a) If ST holds, we have $L(\tau, 0) = 1$ a.s. Moreover, if V is stationary, then $\lambda_T \sim T/m(\mathcal{D})$, where we mean $P_T \sim Q_T$ by $P_T/Q_T \to 1$ as $T \to \infty$.

(b) If a stationary V has a finite second moment, i.e., $\pi(\iota^2) < \infty$, then $b_T \sim T\pi(\iota^2)$, $Q_1^2 = (\pi(\iota))^2/\pi(\iota^2)$ and $Q_2 = 1$. Therefore, if $\pi(\iota^2) < \infty$, then Lemma 4.1 implies

$$\frac{1}{T}\int_0^T V_t^2 \to_p \pi(\iota^2) = \mathbb{E}(V_t^2) \quad \text{and} \quad \frac{1}{T}\int_0^T (V_t - \overline{V}_T)^2 \to_p \pi(\iota^2) - (\pi(\iota))^2 = Var(V_t)$$

as $T \to \infty$.

(c) If a stationary V satisfies $\pi(\sigma^2) < \infty$, then $a_T \sim T\pi(\sigma^2)$ and $P = 1$. Therefore, Lemma 4.1 implies

$$\frac{1}{T}[V]_T = \frac{1}{T}\int_0^T \sigma^2(V_t) \to_p \pi(\sigma^2) = \mathbb{E}(\sigma^2(V_t))$$

as $T \to \infty$.

If neither ST nor DD holds, we would have a quite different asymptotics. Let $Y = s(V)$ be the scale transformation of V and define Y^T by $Y_t^T = \lambda_T^{-1} Y_{Tt}$ with the normalizing sequence (λ_T) in (4.1). It then follows from Proposition 3.2 of Kim and Park (2017) that $Y^T \to_d Y^\circ$ as $T \to \infty$ in the space $C[0, 1]$ of continuous functions with uniform topology, where using Brownian motion B and its local time L we may represent the limit process Y° as

$$Y_t^\circ = B \circ \overline{A}_t \quad \text{with} \quad \overline{A}_t = \inf\left\{ s \,\middle|\, \int_{\mathbb{R}} L(s, x)\overline{m}_s(dx) > t \right\}. \tag{4.3}$$

For the asymptotics of a general diffusion V, we write it as $V = s^{-1}(Y)$, and define

V^T as $V_t^T = V_{Tt}/s^{-1}(\lambda_T) = s^{-1}(Y_{Tt})/s^{-1}(\lambda_T)$. We then may well expect that

$$V^T = \frac{s^{-1}(\lambda_T Y^T)}{s^{-1}(\lambda_T)} \to_d \overline{s^{-1}}(Y^\circ) = V^\circ,$$

in $C[0,1]$ as $T \to \infty$. In particular, it is shown in Kim and Park (2017) that the limiting process V° is a nontrivial stochastic process on $[0,1]$ when neither ST nor DD holds.

Lemma 4.2. *Let Assumption 4.1 hold. If neither* ST *nor* DD *holds, then*

$$\begin{aligned}
\frac{1}{(s^{-1}(\lambda_T))^2}[V]_T &\to_d [V^\circ]_1 \\
\frac{1}{(s^{-1}(\lambda_T))^2}\int_0^T (V_t - \overline{V}_T)dV_t &\to_d \int_0^1 (V_t^\circ - \overline{V}_1^\circ)dV_t^\circ \\
\frac{1}{T\,(s^{-1}(\lambda_T))^2}\int_0^T (V_t - \overline{V}_T)^2 dt &\to_d \int_0^1 (V_t^\circ - \overline{V}_1^\circ)^2 dt
\end{aligned}$$

with $\overline{V}_1^\circ = \int_0^1 V_t^\circ dt$, *as* $T \to \infty$.

Remark 4.2. Combining Lemmas 4.1 and 4.2, Kim and Park (2016) shows that V has a mean reversion if and only if either ST or DD holds. They also show that the unit root test for $z = v$ becomes a test for no mean reversion. In the following section, we will provide that the unit root test is still a test for no mean reversion if it applied to $z = x$.

4.2. *Main Asymptotics*

The long span asymptotics for the one-factor model follows immediately from Proposition 3.2 with Lemmas 4.1 and 4.2.

Theorem 4.3. *Let Assumptions 3.1 and 4.1 hold. In addition, we let either* ST *or* DD *hold. As* $\Delta, \delta/\Delta \to 0$ *and* $T \to \infty$, *we have the followings.*

(a) For $\hat{\beta}_z$, we have

$$N(\hat{\beta}_v - 1) \sim_d -\frac{Ta_T}{b_T}\frac{P}{2Q},$$
$$N(\hat{\beta}_x - 1) \sim_d -\frac{Ta_T}{b_T}\frac{P}{3Q},$$
$$N(\hat{\beta}_y - 1) \sim_d -\frac{Ta_T}{b_T}\frac{P}{3Q} - \frac{\delta T}{\Delta^2}\frac{2Q_2}{Q}.$$

(b) For $t(\hat{\beta}_z)$, we have

$$t(\hat{\beta}_v) \sim_d -\left(\frac{Ta_T}{b_T}\frac{P}{4Q}\right)^{1/2},$$
$$t(\hat{\beta}_x) \sim_d -\left(\frac{Ta_T}{b_T}\frac{P}{6Q}\right)^{1/2},$$
$$t(\hat{\beta}_y) \sim_d -\left(\frac{Ta_T}{b_T}\right)^{1/2}\frac{P + (6\delta/\Delta^2)(b_T/a_T)Q_2}{\left(6P + (36\delta/\Delta^2)(b_T/a_T)Q_2\right)^{1/2} Q^{1/2}}.$$

Remark 4.3. (a) If V is stationary with $\pi(\iota^2), \pi(\sigma^2) < \infty$, then Theorem 4.3 (a) implies

$$\hat{\beta}_v - 1 \sim_p -\Delta\frac{\mathbb{E}(\sigma^2(V_t))}{2Var(V_t)},$$
$$\hat{\beta}_x - 1 \sim_p -\Delta\frac{\mathbb{E}(\sigma^2(V_t))}{3Var(V_t)},$$
$$\hat{\beta}_y - 1 \sim_p -\Delta\frac{\mathbb{E}(\sigma^2(V_t))}{3Var(V_t)} - \frac{\delta}{\Delta}\frac{2\mathbb{E}(V_t^2)}{Var(V_t)}$$

as $\Delta, \delta/\Delta \to 0$ and $T \to \infty$. It is very interesting to note that these results are the same as those derived by Andersen et al. (2004) and given in Section 2.6.

(b) For a stationary Ornstein-Uhlenbeck process V, given as

$$dV_t = \kappa(\mu - V_t)dt + \sigma dW_t,$$

we have $\mathbb{E}(\sigma^2(V_t)) = \sigma^2$, $Var(V_t) = \sigma^2/(2\kappa)$ and $\mathbb{E}(V_t^2) = \sigma^2/(2\kappa) + \mu^2$, and hence, we have

$$\hat{\beta}_v - 1 \sim_p -\Delta\kappa,$$
$$\hat{\beta}_x - 1 \sim_p -\Delta\frac{2}{3}\kappa,$$
$$\hat{\beta}_y - 1 \sim_p -\Delta\frac{2}{3}\kappa - 2\frac{\delta}{\Delta}\left(1 + \frac{2\kappa\mu^2}{\sigma^2}\right)$$

as $\Delta, \delta/\Delta \to 0$ and $T \to \infty$.

(c) Let V be a stationary Ornstein-Uhlenbeck process with $V_0 = 0$, $\mu = 0$ and $\sigma = 1$. It then follows from Proposition 3.2 (a) that

$$\hat{\beta}_x - 1 \sim_p -\Delta\kappa + \Delta\frac{\int_0^T (V_t - \overline{V}_T)dW_t + T/6}{\int_0^T (V_t - \overline{V}_T)^2 dt} \tag{4.4}$$

as long as Δ is sufficiently small. Moreover, if we let $T \to \infty$, then (4.4) becomes

$$\hat{\beta}_x - 1 \sim_p -\Delta\frac{2}{3}\kappa \not\sim_p \hat{\beta}_v - 1 \sim_p -\Delta\kappa \tag{4.5}$$

as shown in Remark 4.3 (b). For the Ornstein-Uhlenbeck process here, Chambers (2004) shows that (i) (4.4) holds when $\Delta \to 0$ and T is fixed and (ii) $\hat{\beta}_x - 1 \sim_p \hat{\beta}_v - 1$ when $\Delta \to 0$ and $T \to \infty$ such that $\Delta T^{1/2} \to \infty$. Our result (4.5) does not contradict with the result in Chambers (2004) since we require $\Delta T^{1/2} \to 0$ due to Assumption 3.1 (a).

(d) Let V be a stationary GARCH diffusion

$$dV_t = \kappa(\mu - V_t)dt + \sigma V_t dW_t$$

with $\sigma^2 < 2\kappa$ so that $\mathbb{E}(V_t^2) < \infty$. In this case, we have $\mathbb{E}(\sigma^2(V_t)) = \sigma^2\mathbb{E}(V_t^2)$, $Var(V_t) = \mu^2\sigma^2/(2\kappa - \sigma^2)$ and $\mathbb{E}(V_t^2) = 2\kappa\mu^2/(2\kappa - \sigma^2)$, and hence, we have

$$\hat{\beta}_v - 1 \sim_p -\Delta\kappa,$$
$$\hat{\beta}_x - 1 \sim_p -\Delta\frac{2}{3}\kappa,$$
$$\hat{\beta}_y - 1 \sim_p -\Delta\frac{2}{3}\kappa - 4\frac{\delta}{\Delta}\frac{\kappa}{\sigma^2}$$

as $\delta/\Delta, \Delta \to 0$ and $T \to \infty$. Again, it is very interesting to note that these results are the same as those derived by Andersen et al. (2004) and given in Section 2.6.

(e) Let V be a linear drift diffusion

$$dV_t = \kappa(\mu - V_t)dt + \sigma(V_t)dW_t$$

with $\mathbb{E}(V_t) = \mu$ and $\mathbb{E}(\sigma^2(V_t)) < \infty$. In Lemma A6 of Kim and Park (2016), it is shown that $\mathbb{E}(\sigma^2(V_t)) = -2\mathbb{E}(V_t\mu(V_t))$ for a stationary V with $\mathbb{E}(\sigma^2(V_t)) < \infty$. In this case, we have $\mathbb{E}(\sigma^2(V_t)) = 2\kappa Var(V_t)$, and hence,

$$\hat{\beta}_v - 1 \sim_p -\Delta\kappa \quad \text{and} \quad \hat{\beta}_x - 1 \sim_p -\Delta\frac{2}{3}\kappa.$$

(f) Let V be a stationary GARCH diffusion with $2\kappa < \sigma^2$ so that $\mathbb{E}(V_t^2) = \infty$. In this case, neither DD nor SI holds. Moreover, $a_T P = \sigma^2 b_T Q$ and $Q_2 = Q$ since $a_T = \sigma^2 b_T$ and $P = Q$, and therefore, we have

$$\begin{aligned}
\hat{\beta}_v - 1 &\sim_p -\Delta\frac{1}{2}\sigma^2, \\
\hat{\beta}_x - 1 &\sim_p -\Delta\frac{1}{3}\sigma^2, \\
\hat{\beta}_y - 1 &\sim_p -\Delta\frac{1}{3}\sigma^2 - 2\frac{\delta}{\Delta}.
\end{aligned}$$

(g) Our example in Remark 4.3 (f) is comparable to the limit theory for the sample autocorrelations of a GARCH(1,1) process obtained in Mikosch and Starica (2000). Let

$$X_i = \sigma_i Z_i \quad \text{with} \quad \sigma_i^2 = \alpha_0 + \beta_1\sigma_{i-1}^2 + \alpha_1 X_{i-1}^2 \quad \text{for} \quad i = 1, 2, \cdots, N,$$

where (Z_i) is a sequence of iid symmetric random variables with $\mathbb{E}Z_i^2 = 1$. Under some assumptions, which imply that the vector (X_i, σ_i) is regularly varying with index $\kappa > 0$, it is shown that for $\kappa \in (0, 4)$ the variance process (σ_i^2) has unbounded variance and satisfies

$$\left(\frac{\sum_{i=1}^{N-h} X_i^2 X_{i+h}^2}{\sum_{i=1}^{N} X_i^4} - 1, \frac{\sum_{i=1}^{N-h} \sigma_i^2 \sigma_{i+h}^2}{\sum_{i=1}^{N} \sigma_i^4} - 1 \right) \sim_d \left(\frac{\Sigma_{1,X^2} - \Sigma_{0,X^2}}{\Sigma_{0,X^2}}, \frac{\Sigma_{1,\sigma^2} - \Sigma_{0,\sigma^2}}{\Sigma_{0,\sigma^2}} \right),$$

where the limit distribution is nondegenerated since the vector $(\Sigma_{m,X^2}, \Sigma_{m,\sigma^2})_{m=0,1}$ is

$\kappa/2$-stable. This contrasts with our result for a GARCH diffusion with unbounded variance (see Remark 4.3 (f)) since $(\hat{\beta}_z - 1)/\Delta$ has a degenerate limit for $z = v, x$.

Theorem 4.4. *Let Assumptions 3.1 and 4.1 hold. In addition, we let neither* ST *nor* DD *hold. As* $\Delta, \delta/\Delta \to 0$ *and* $T \to \infty$, *we have the followings.*

(a) For $\hat{\beta}_z$, *we have*

$$N(\hat{\beta}_v - 1) \sim_d \frac{\int_0^1 (V_t^\circ - \overline{V}_1^\circ) dV_t^\circ}{\int_0^1 (V_t^\circ - \overline{V}_1^\circ)^2 dt},$$

$$N(\hat{\beta}_x - 1) \sim_d \frac{\int_0^1 (V_t^\circ - \overline{V}_1^\circ) dV_t^\circ + (1/6)[V^\circ]_1}{\int_0^1 (V_t^\circ - \overline{V}_1^\circ)^2 dt},$$

$$N(\hat{\beta}_y - 1) \sim_d \frac{\int_0^1 (V_t^\circ - \overline{V}_1^\circ) dV_t^\circ + (1/6)[V^\circ]_1 - 2(\delta/\Delta^2) \int_0^1 V_t^{\circ 2} dt}{\int_0^1 (V_t^\circ - \overline{V}_1^\circ)^2 dt}.$$

(b) For $t(\hat{\beta}_z)$, *we have*

$$t(\hat{\beta}_v) \sim_d \frac{\int_0^1 (V_t^\circ - \overline{V}_1^\circ) dV_t^\circ}{[V^\circ]_1^{1/2} \left(\int_0^1 (V_t^\circ - \overline{V}_1^\circ)^2 dt \right)^{1/2}},$$

$$t(\hat{\beta}_x) \sim_d \frac{\int_0^1 (V_t^\circ - \overline{V}_1^\circ) dV_t^\circ + (1/6)[V^\circ]_1}{((2/3)[V^\circ]_1)^{1/2} \left(\int_0^1 (V_t^\circ - \overline{V}_1^\circ)^2 dt \right)^{1/2}},$$

$$t(\hat{\beta}_y) \sim_d \frac{\int_0^1 (V_t^\circ - \overline{V}_1^\circ) dV_t^\circ + (1/6)[V^\circ]_1 - 2(\delta/\Delta^2) \int_0^1 V_t^{\circ 2} dt}{\left((2/3)[V^\circ]_1 + 4(\delta/\Delta^2) \int_0^1 V_t^{\circ 2} dt \right)^{1/2} \left(\int_0^1 (V_t^\circ - \overline{V}_1^\circ)^2 dt \right)^{1/2}}.$$

Remark 4.4. (a) Theorems 4.3 and 4.4 imply that the unit root test, applied to $z = v, x$, can effectively discriminate the following null and alternative hypothesis

$$H_0 : \text{neither ST nor DD holds (no mean reversion)}$$
$$H_A : \text{ST or DD holds (mean reversion)},$$

instead of nonstationarity and stationarity for the underlying process V. The test statistics, both $N(\hat{\beta}_z - 1)$ and $t(\hat{\beta}_z)$ for $z = v, x$, have well defined limiting distributions under H_0, whereas it diverges to negative infinity under H_A. However, the test applied to $z = y$ can discriminate H_0 and H_A only when $\delta/\Delta^2 \to 0$. The reader is referred to see Kim and Park (2016) for more discussions about the unit root and mean reversion properties when $z = v$.

(b) If V is a Brownian motion, then

$$N(\hat{\beta}_x - 1) \sim_d \frac{\int_0^1 (W_t - \overline{W}_1) dW_t + (1/6)}{\int_0^1 (W_t - \overline{W}_1)^2 dt}$$

as long as $\Delta \to 0$ due to Proposition 3.2 and Theorem 4.4. This result can be also found in Chambers (2004).

Theorem 4.5. *Let Assumptions 3.1 and 4.1 hold. As $\Delta, \delta/\Delta \to 0$ and $T \to \infty$, we have the followings.*

(a) For $z = v, x, y$, we have

$$\hat{\beta}_{1,z} + \hat{\beta}_{2,z} - 1 \sim_d \hat{\beta}_z - 1 - \Delta\Gamma_z, \quad \hat{\beta}_{1,z} \sim_p 1 - \Gamma_{1,z}, \quad \hat{\beta}_{2,z} \sim_p \Gamma_{1,z}.$$

(b) If either ST *or* DD *holds, we have $\Gamma_v = \Gamma_{1,v} = 0$ and*

$$\Gamma_x \sim_d \frac{1}{12} \frac{a_T P}{b_T Q}, \qquad \Gamma_y \sim_d \frac{1}{12} \frac{a_T P}{b_T Q} - \frac{\delta}{\Delta^2} \frac{Q_2}{Q},$$

$$\Gamma_{1,x} \sim_p -\frac{1}{4}, \qquad \Gamma_{1,y} \sim_d -\frac{a_T P - 12(\delta/\Delta^2) b_T Q_2}{4 a_T P + 24(\delta/\Delta^2) b_T Q_2}.$$

(c) If neither ST *nor* DD *holds, we have $\Gamma_v = \Gamma_{1,v} = 0$ and*

$$\Gamma_x \sim_d \frac{1}{12} \frac{[V^\circ]_1}{\int_0^1 (V_t^\circ - \overline{V}_1^\circ)^2 dt},$$

$$\Gamma_y \sim_d \frac{1}{12} \frac{[V^\circ]_1}{\int_0^1 (V_t^\circ - \overline{V}_1^\circ)^2 dt} - \frac{\delta}{\Delta^2} \frac{\int_0^1 V_t^{\circ 2} dt}{\int_0^1 (V_t^\circ - \overline{V}_1^\circ)^2 dt},$$

$$\Gamma_{1,x} \sim_d \frac{[V^\circ]_1}{12 \int_0^1 (V_t^\circ - \overline{V}_1^\circ) dV_t^\circ + 2[V^\circ]_1},$$

$$\Gamma_{1,y} \sim_d \frac{[V^\circ]_1 - 12(\delta/\Delta^2) T \int_0^1 V_t^{\circ 2} dt}{12 \int_0^1 (V_t^\circ - \overline{V}_1^\circ) dV_t^\circ + 2[V^\circ]_1 - 24(\delta/\Delta^2) T \int_0^1 V_t^{\circ 2} dt}.$$

Remark 4.5. (a) If V is stationary, then we have $\hat{\beta}_{1,x} \to_p 5/4$ and $\hat{\beta}_{2,x} \to_p -1/4$ as

$\Delta \to 0$ and $T \to \infty$. Moreover, if $\pi(\iota^2), \pi(\sigma^2) < \infty$, then

$$\hat{\beta}_{1,y} \sim_p 1 + \frac{\mathbb{E}(\sigma^2(V_t)) - 12(\delta/\Delta^2)\mathbb{E}(V_t^2)}{4\mathbb{E}(\sigma^2(V_t)) + 24(\delta/\Delta^2)\mathbb{E}(V_t^2)} \quad \text{and} \quad \hat{\beta}_{2,y} \sim_p 1 - \hat{\beta}_{1,y}.$$

and

$$\begin{aligned}
\hat{\beta}_{1,v} + \hat{\beta}_{2,v} - 1 &\sim_p -\Delta \frac{\mathbb{E}(\sigma^2(V_t))}{2Var(V_t)}, \\
\hat{\beta}_{1,x} + \hat{\beta}_{2,x} - 1 &\sim_p -\Delta \frac{5\mathbb{E}(\sigma^2(V_t))}{12Var(V_t)}, \\
\hat{\beta}_{1,y} + \hat{\beta}_{2,y} - 1 &\sim_p -\Delta \frac{5\mathbb{E}(\sigma^2(V_t))}{12Var(V_t)} - \frac{\delta}{\Delta} \frac{\mathbb{E}(V_t^2)}{Var(V_t)}.
\end{aligned}$$

(b) If V is a GARCH diffusion with $\pi(\iota^2) < \infty$, we have

$$\begin{aligned}
\hat{\beta}_{1,v} + \hat{\beta}_{2,v} - 1 &\sim_p -\Delta\kappa, \\
\hat{\beta}_{1,x} + \hat{\beta}_{2,x} - 1 &\sim_p -\Delta \frac{5}{6}\kappa, \\
\hat{\beta}_{1,y} + \hat{\beta}_{2,y} - 1 &\sim_p -\Delta \frac{5}{6}\kappa - \frac{\delta}{\Delta}\frac{2\kappa}{\sigma^2}.
\end{aligned}$$

4.3. Multi-factor Variance

In this section, we assume that the variance process V has multiple factors such that $V = \sum_{k=1}^{K} V_k$ with

$$dV_{k,t} = \mu_k(V_{k,t})dt + \sigma_k(V_{k,t})dW_{k,t},$$

and V_ks are independent each other.

The primary asymptotics of the multi-factor variance can be obtained similarly as those of the one-factor variance, and we have

Proposition 4.6. *Let V_k satisfy Assumption 3.1 for all $k = 1, \cdots, K$. Then Lemma 3.1 and Propositions 3.2-3.4 are still valid for the K-factor variance V.*

Now we consider the long span asymptotics of $\hat{\beta}_z$ and $t(\hat{\beta}_z)$ for $z = v, x, y$. For simplicity, we consider a two-factor variance V, and let $K = 2$.

Theorem 4.7. *Let both V_1 and V_2 satisfy Assumptions 3.1 and 4.1.*

(a) Let both V_1 and V_2 be stationary with $\mathbb{E}(\sigma_k^2(V_{k,t})), \mathbb{E}(V_{k,t}^2) < \infty$ for $k = 1, 2$. Then Theorem 4.3 holds with $a_T = b_T = T$ and $P = \mathbb{E}(\sigma_1^2(V_{1,t})) + \mathbb{E}(\sigma_2^2(V_{2,t}))$, $Q = Var(V_{1,t}) + Var(V_{2,t})$ and $Q_2 = \mathbb{E}(V_{1,t}^2) + \mathbb{E}(V_{2,t}^2)$.

(b) Let both V_1 and V_2 be stationary with $a_{2T}/a_{1T} \to 0$ and $b_{2T}/b_{1T} \to 0$. Then Theorem 4.3 holds with (a_T, b_T, P, Q, Q_2) being obtained from V_1.

In Theorem 4.7, we do not consider the following two cases; (i) both V_1 and V_2 are stationary with $a_{2T}/a_{1T} \to 0$ and $b_{1T}/b_{2T} \to 0$, and (ii) one or both of V_1 and V_2 are nonstationary. The long span asymptotics of the these two cases can be easily obtained at the cost of more involved analysis. Moreover, Theorem 4.7 can be easily extended to the general K-factor model.

5. Simulations

In this section, we show by simulations that our limit theory provides a good approximation for the distribution of OLS estimates in a realistic situation. For our simulation, we use the GARCH diffusion with two sets of parameters. The first one is $(\kappa_0, \mu_0, \sigma_0^2) = (0.0350, 0.6360, 0.0207)$, corresponding to $\psi_0 = \sigma_0^2/(2\kappa_0) = 0.296$. This set of parameters was used by Andersen and Bollerslev (1998) as implied from the (weak) daily GARCH(1,1) model estimates for the DM/dollar from 1987 through 1992 using the temporal aggregation results of Drost and Nijman (1993) and Drost and Werker (1996); the same parameters were used by Andersen et al. (2004). Because $\psi_0 < 1$, the second moment of V_t is bounded.

To consider a process with an unbounded variance, we consider a second set of parameters by keeping the same κ_0 and μ_0, while we multiply σ_0^2 by 4, corresponding to $\psi_0 = 1.183$, that is $(\kappa_0, \mu_0, \sigma_0^2) = (0.0350, 0.6360, 0.0828)$.

The simulation samples are generated by the Euler discretization at 10 seconds ($\delta_0/\Delta = 1/8640$) for $T = 1, 2, 4$ years of sample span with 250 days per year. We assume that the market is open 24 hours. For each day ($\Delta = 1$), we set the daily spot variance as the spot variance at the end of the day, while we compute the integrated variance by the numerical integration of the simulated spot variance process at 10 seconds. As for the realized variance, we analyze the frequency effects by considering three different frequencies: 10 minutes ($\delta/\Delta = 1/144$), 5 minutes ($\delta/\Delta = 1/288$) and

1 minute ($\delta/\Delta = 1/1440$). For each design, we get rid of the first five days to reduce the effect of the initial value, and we do 10,000 replications.

Under the stationarity, Theorem 4.3 implies that

$$\hat{\beta}_v - 1 \sim_d -\Delta \frac{[V]_T}{2\int_0^T (V_t - \overline{V}_T)^2 dt}$$
$$\hat{\beta}_x - 1 \sim_d -\Delta \frac{[V]_T}{3\int_0^T (V_t - \overline{V}_T)^2 dt}$$
$$\hat{\beta}_y - 1 \sim_d -\Delta \frac{[V]_T}{3\int_0^T (V_t - \overline{V}_T)^2 dt} - \frac{\delta}{\Delta} \frac{2\int_0^T V_t^2 dt}{\int_0^T (V_t - \overline{V}_T)^2 dt},$$

whereas Proposition 3.4 implies that $\hat{\beta}_{1,z} + \hat{\beta}_{2,z} - 1 \sim_d \hat{\beta}_z - 1 - \Delta\Gamma_z$, where $\Gamma_v = 0$ and

$$\Gamma_x \sim_d \frac{[V]_T}{12\int_0^T (V_t - \overline{V}_T)^2 dt}, \quad \Gamma_y \sim_d \frac{[V]_T}{12\int_0^T (V_t - \overline{V}_T)^2 dt} - \frac{\delta}{\Delta^2} \frac{\int_0^T V_t^2 dt}{\int_0^T (V_t - \overline{V}_T)^2 dt}.$$

For each simulation, we approximate the limit distributions of $\hat{\beta}_z$ and $\hat{\beta}_{1,z} + \hat{\beta}_{2,z}$ by replacing $[V]_T$, $\int_0^T V_t^2 dt$ and $\int_0^T (V_t - \overline{V}_T)^2 dt$ in the above relations by their sample proxies

$$\sum_{i=1}^{N} (v_{i+1} - v_i)^2, \quad \sum_{i=1}^{N} v_i^2 \Delta, \quad \sum_{i=1}^{N} (v_i - \overline{v}_N)^2 \Delta,$$

respectively. We compare the approximated limit distribution, say $\tilde{\beta}_z$ and $\tilde{\beta}_{1,z} + \tilde{\beta}_{2,z}$, of $\hat{\beta}_z$ and $\hat{\beta}_{1,z} + \hat{\beta}_{2,z}$ to the finite sample distribution of $\hat{\beta}_z$ and $\hat{\beta}_{1,z} + \hat{\beta}_{2,z}$.

In Table 2, we report the root mean squared differences between the finite sample distribution and the approximated distribution. We can easily see that the differences become smaller as T increases in each case. Moreover, the differences become smaller as δ decreases in realized variances. These two results are quite natural since our approximation is based on the asymptotic distribution obtained under $\delta/\Delta, \Delta \to 0$ and $T \to \infty$. The differences are larger in the unbounded variance cases than in the bounded variance cases.

Figures 1-8 show the empirical distribution of $\hat{\beta}_z - 1$ and the approximated limit distribution $\tilde{\beta}_z - 1$. It is easy to see that our approximated limit distribution provides a good approximation for the finite sample distribution of $\hat{\beta}_z - 1$. Moreover, as our

AR(1)	1 Year	2 Year	4 Year
SPOT	0.0068	0.0035	0.0018
IV	0.0081	0.0044	0.0025
RV (01 min)	0.0087	0.0048	0.0029
RV (05 min)	0.0128	0.0077	0.0050
RV (10 min)	0.0218	0.0135	0.0091
AR(2)	1 Year	2 Year	4 Year
SPOT	0.0063	0.0038	0.0024
IV	0.0047	0.0026	0.0016
RV (01 min)	0.0050	0.0028	0.0019
RV (05 min)	0.0086	0.0052	0.0034
RV (10 min)	0.0158	0.0096	0.0065

AR(1)	1 Year	2 Year	4 Year
SPOT	0.0128	0.0077	0.0041
IV	0.0169	0.0104	0.0071
RV (01 min)	0.0173	0.0107	0.0074
RV (05 min)	0.0188	0.0123	0.0086
RV (10 min)	0.0210	0.0143	0.0103
AR(2)	1 Year	2 Year	4 Year
SPOT	0.0139	0.0095	0.0071
IV	0.0097	0.0063	0.0046
RV (01 min)	0.0100	0.0065	0.0048
RV (05 min)	0.0110	0.0074	0.0057
RV (10 min)	0.0130	0.0090	0.0069

Table 2: The root mean squared differences between the finite sample distribution and the approximated distribution (Left=Finite Variance, Right=Infinite Variance)

theory expected (see Remark 3.2 (a)), we tend to have $\hat{\beta}_y < \hat{\beta}_x$ and $\hat{\beta}_v < \hat{\beta}_x$. In particular, the gap between $\hat{\beta}_y$ and $\hat{\beta}_x$ decreases as δ decreases. We also note that there are no qualitative differences between the bonded and unbounded variance cases.

6. Conclusion

Fat tails are a well-known empirical fact of financial returns. Surprisingly, the realized volatility literature ignored this fact. After proving empirically that the second moment of the realized variance is probably unbounded, we studied theoretically the limiting behavior of the OLS estimator of simple auto-regressions of spot, integrated and realized variances. We proved that when the second moment of the spot variance is unbounded, the OLS estimators converge to random variables. Our theory is also valid when the second moment of the spot variance is bounded. In this case, the OLS estimates converge to finite and deterministic quantities which are the same ones derived by Andersen et al. (2004) in population regressions. Likewise, our theory allows for nonstationary volatility. Our theoretical results are based on asymptotic approximations. Both the simulations and the comparison with the results in Andersen et al. (2004) when the spot variance has a finite second moment corroborate the good quality of our approach.

There are at least two important questions that should be addressed. The first one is to provide consistent estimators of the regression coefficients when the second moment of the spot variance in unbounded. Typically, one could use signed power variation which can be view as an instrumental variable estimation where one uses a signed power of the regressor as an instrument in order to reduce the magnitude of the tails; see Samorodnitsky et al. (2007). The second question is more important and concerns the forecast that one should compute under fat tails. Various approaches could be considered like different loss functions or non-linear transforms of the variable of interest. The two questions are currently under investigation.

Appendix

A. Useful Lemmas for Integrated Variance

Lemma A.1. *If f is continuously differentiable, then we have*

$$\sup_{1\le i\le N}\left|f(x_i)-f(V_{(i-i)\Delta})\right| = O_p\left(\Delta T(f'\mu)\right)+O_p\left(\Delta^{1/2}T(f'\sigma)\sqrt{\log(T/\Delta)}\right).$$

Proof for Lemma A.1. Since V has a continuous sample path, we may deduce from the mean value theorem and Taylor expansion that

$$\begin{aligned}\sup_{1\le i\le N}\left|f(x_i)-f(V_{(i-i)\Delta})\right| &= \sup_{1\le i\le N}\left|f'(V_{k_i})\frac{1}{\Delta}\int_{(i-1)\Delta}^{i\Delta}(V_t-V_{(i-1)\Delta})dt\right| \\ &\le T(f')\sup_{1\le i\le N}\left|\frac{1}{\Delta}\int_{(i-1)\Delta}^{i\Delta}(V_t-V_{(i-1)\Delta})dt\right| \qquad (A.1)\end{aligned}$$

for some $k_i\in[(i-1)\Delta, i\Delta]$. Moreover, we have

$$\begin{aligned}\int_{(i-1)\Delta}^{i\Delta}(V_t-V_{(i-1)\Delta})dt &= \int_{(i-1)\Delta}^{i\Delta}\int_{(i-1)\Delta}^{t}\mu(V_s)dsdt+\int_{(i-1)\Delta}^{i\Delta}\int_{(i-1)\Delta}^{t}\sigma(V_s)dW_sdt \\ &= O_p\left(\Delta^2T(\mu)\right)+O_p\left(\Delta^{3/2}T(\sigma)\sqrt{\log(T/\Delta)}\right) \qquad (A.2)\end{aligned}$$

uniformly in $1\le i\le N$, since

$$\sup_{1\le i\le N}\left|\int_{(i-1)\Delta}^{i\Delta}\int_{(i-1)\Delta}^{t}\mu(V_s)dsdt\right| = O_p\left(\Delta^2T(\mu)\right)$$

and

$$\sup_{1\le i\le N}\left|\int_{(i-1)\Delta}^{i\Delta}\int_{(i-1)\Delta}^{t}\sigma(V_s)dW_sdt\right| = O_p\left(\Delta^{3/2}T(\sigma)\sqrt{\log(T/\Delta)}\right)$$

due to the global modulus of continuity for Brownian motion (see, e.g., Theorem 1 of Kanaya et al. (2016) and Lemma B2 of Kim and Park (2017)). The stated result follows immediately from (A.1) and (A.2). □

Lemma A.2. *If f is twice continuously differentiable, then we have*

$$\Delta \sum_{i=1}^{N} f(x_i) = \int_0^T f(V_t)dt + O_p\left(\Delta T(f'\mu)T\right) + O_p\left(\Delta T(f''\sigma)T\right) + O_p\left(\Delta^{1/2} T(f'\sigma)T\sqrt{\log(T/\Delta)}\right) + O_p\left(\Delta T(f'\sigma)T^{1/2}\right).$$

Proof for Lemma A.2. Due to Lemma A.1, we have

$$\Delta \sum_{i=1}^{N} f(x_i) = \Delta \sum_{i=1}^{N} f\left(V_{(i-1)\Delta}\right) + O_p\left(\Delta T(f'\mu)T\right) + O_p\left(\Delta^{1/2} T(f'\sigma)T\sqrt{\log(T/\Delta)}\right). \tag{A.3}$$

Moreover, by Lemma B1 of Kim and Park (2017), we have

$$\Delta \sum_{i=1}^{N} f\left(V_{(i-1)\Delta}\right) = \int_0^T f(V_t)dt + O_p\left(\Delta T(f'\mu)T\right) + O_p\left(\Delta T(f''\sigma)T\right) + O_p\left(\Delta T(f'\sigma)T^{1/2}\right),$$

from which, together with (A.3), the stated result follows immediately. □

Lemma A.3. *If f is twice continuously differentiable, then we have*

$$\frac{1}{\Delta^2} \sum_{i=1}^{N} \int_{i\Delta}^{(i+1)\Delta} ((i+1)\Delta - s)^2 (f(V_s) - f(V_{i\Delta}))ds = O_p\left(\Delta T(f'\mu)T\right) + O_p\left(\Delta T(f''\sigma^2)T\right) + O_p\left(\Delta T(f'\sigma)T^{1/2}\right).$$

Proof for Lemma A.3. Due to Ito's lemma, we have

$$\frac{1}{\Delta^2} \sum_{i=1}^{N} \int_{i\Delta}^{(i+1)\Delta} ((i+1)\Delta - s)^2 (f(V_s) - f(V_{i\Delta}))ds = A_T + B_T, \tag{A.4}$$

where

$$A_T = \frac{1}{\Delta^2} \sum_{i=1}^{N} \int_{i\Delta}^{(i+1)\Delta} ((i+1)\Delta - s)^2 \left(\int_{i\Delta}^{s} (f'\mu + f''\sigma^2/2)(V_t)dt\right) ds,$$

$$B_T = \frac{1}{\Delta^2} \sum_{i=1}^{N} \int_{i\Delta}^{(i+1)\Delta} ((i+1)\Delta - s)^2 \left(\int_{i\Delta}^{s} (f'\sigma)(V_t)dW_t\right) ds.$$

For A_T, it is easy to see that

$$A_T = O_p\left(\Delta T(f'\mu)T\right) + O_p\left(\Delta T(f''\sigma^2)T\right). \tag{A.5}$$

As for B_T, we have

$$\begin{aligned} B_T &= \frac{1}{\Delta^2}\sum_{i=1}^{N}\int_{i\Delta}^{(i+1)\Delta}(f'\sigma)(V_t)\left(\int_t^{(i+1)\Delta}((i+1)\Delta - s)^2 ds\right) dW_t \\ &= O_p\left(\Delta T(f'\sigma)T^{1/2}\right), \end{aligned} \tag{A.6}$$

where the first line is due to the changing the order of integrals, and the second line can be deduced from the proof of Lemma B1 in Kim and Park (2017). The stated result follows immediately from (A.4)-(A.6).

□

Lemma A.4. *If* $\Delta^{1/2}T(\omega^{3/2})T\sqrt{\log(T/\Delta)} \to_p 0$, *we have*

$$\sum_{i=1}^{N}(x_{i+1} - x_i)^2 = \frac{2}{3}[V]_T + o_p(1).$$

Proof for Lemma A.4. We write

$$\begin{aligned} \sum_{i=1}^{N}(x_{i+1} - x_i)^2 &= \frac{1}{\Delta^2}\sum_{i=1}^{N}\left(\int_{i\Delta}^{(i+1)\Delta}(V_t - V_{i\Delta})dt + \int_{(i-1)\Delta}^{i\Delta}(V_{i\Delta} - V_t)dt\right)^2 \\ &= A_T + B_T + R_T, \end{aligned} \tag{A.7}$$

where

$$\begin{aligned} A_T &= \frac{1}{\Delta^2}\sum_{i=1}^{N}\left(\int_{i\Delta}^{(i+1)\Delta}(V_t - V_{i\Delta})dt\right)^2 \\ B_T &= \frac{1}{\Delta^2}\sum_{i=1}^{N}\left(\int_{(i-1)\Delta}^{i\Delta}(V_{i\Delta} - V_t)dt\right)^2 \\ R_T &= \frac{2}{\Delta^2}\sum_{i=1}^{N}\left(\int_{i\Delta}^{(i+1)\Delta}(V_t - V_{i\Delta})dt\right)\left(\int_{(i-1)\Delta}^{i\Delta}(V_{i\Delta} - V_t)dt\right). \end{aligned}$$

Due to (A.7), the stated result follows immediately if we show

$$A_T, B_T = \frac{1}{3}[V]_T + O_p\left(\Delta^{1/2}T(\omega^{3/2})T\sqrt{\log(T/\Delta)}\right), \tag{A.8}$$

$$R_T = O_p\left(\Delta^{1/2}T(\omega^{3/2})T\sqrt{\log(T/\Delta)}\right). \tag{A.9}$$

Proof for (A.8). We will only prove the result for A_T, since the proof of the result for B_T is entirely analogous. For the proof, we write A_T as

$$\begin{aligned} A_T &= \frac{1}{\Delta^2}\sum_{i=1}^{N}\left(\int_{i\Delta}^{(i+1)\Delta}\int_{i\Delta}^{t}\mu(V_s)dsdt + \int_{i\Delta}^{(i+1)\Delta}\int_{i\Delta}^{t}\sigma(V_s)dW_sdt\right)^2 \\ &= A_{1,T} + A_{2,T} + A_{3,T}, \end{aligned} \tag{A.10}$$

where

$$\begin{aligned} A_{1,T} &= \frac{1}{\Delta^2}\sum_{i=1}^{N}\left(\int_{i\Delta}^{(i+1)\Delta}\int_{i\Delta}^{t}\mu(V_s)dsdt\right)^2 = \frac{1}{\Delta^2}\sum_{i=1}^{N}\left(\int_{i\Delta}^{(i+1)\Delta}((i+1)\Delta) - s)\mu(V_s)ds\right)^2 \\ A_{2,T} &= \frac{1}{\Delta^2}\sum_{i=1}^{N}\left(\int_{i\Delta}^{(i+1)\Delta}\int_{i\Delta}^{t}\sigma(V_s)dW_sdt\right)^2 = \frac{1}{\Delta^2}\sum_{i=1}^{N}\left(\int_{i\Delta}^{(i+1)\Delta}((i+1)\Delta) - s)\sigma(V_s)dW_s\right)^2 \\ A_{3,T} &= \frac{2}{\Delta^2}\sum_{i=1}^{N}\left(\int_{i\Delta}^{(i+1)\Delta}\int_{i\Delta}^{t}\mu(V_s)dsdt\right)\left(\int_{i\Delta}^{(i+1)\Delta}\int_{i\Delta}^{t}\sigma(V_s)dW_sdt\right) \\ &= \frac{2}{\Delta^2}\sum_{i=1}^{N}\left(\int_{i\Delta}^{(i+1)\Delta}((i+1)\Delta) - s)\mu(V_s)ds\right)\left(\int_{i\Delta}^{(i+1)\Delta}((i+1)\Delta) - s)\sigma(V_s)dW_s\right) \end{aligned}$$

by changing the order of integrals.

For $A_{1,T}$, we have

$$|A_{1,T}| \leq \sum_{i=1}^{N}\left(\int_{i\Delta}^{(i+1)\Delta}\sup_{i\Delta\leq s\leq(i+1)\Delta}|\mu(V_s)|ds\right)^2 = O_p\left(\Delta T(\mu^2)T\right). \tag{A.11}$$

On the other hand, we can deduce from Lemma B2 of Kim and Park (2017) that

$A_{3,T}$ satisfies

$$|A_{3,T}| \leq 2T(\mu) \sum_{i=1}^{N} \left| \int_{i\Delta}^{(i+1)\Delta} ((i+1)\Delta) - s)\sigma(V_s) dW_s \right| = O_p \left(\Delta^{1/2} T(\mu\sigma) T \sqrt{\log(T/\Delta)} \right). \tag{A.12}$$

As for $A_{2,T}$, we define a continuous martingale M as

$$M_t = \sum_{i=1}^{j-1} \int_{i\Delta}^{(i+1)\Delta} ((i+1)\Delta - s)\sigma(V_s) dW_s + \int_{j\Delta}^{t} ((j+1)\Delta - s)\sigma(V_s) dW_s$$

for $t \in [j\Delta, (j+1)\Delta)$, $j = 1, 2, \cdots, N$, so that we have

$$\begin{aligned} A_{2,T} &= \frac{1}{\Delta^2} \sum_{i=1}^{N} \left(M_{(i+1)\Delta} - M_{i\Delta} \right)^2 \\ &= \frac{1}{\Delta^2} [M]_T + \frac{2}{\Delta^2} \sum_{i=1}^{N} \int_{i\Delta}^{(i+1)\Delta} (M_t - M_{i\Delta}) dM_t, \end{aligned} \tag{A.13}$$

where the last line follows from Ito's lemma.

For the second term of (A.13), we can deduced from Lemma B5 of Kim and Park (2017) that

$$\begin{aligned} &\frac{2}{\Delta^2} \sum_{i=1}^{N} \int_{i\Delta}^{(i+1)\Delta} (M_t - M_{i\Delta}) dM_t \\ &= \frac{2}{\Delta^2} \sum_{i=1}^{N} \int_{i\Delta}^{(i+1)\Delta} \left(\int_{i\Delta}^{t} ((i+1)\Delta - s)\sigma(V_s) dW_s \right) ((i+1)\Delta) - t)\sigma(V_t) dW_t \\ &= O_p \left(\Delta^{1/2} T(\sigma^2) T^{1/2} \sqrt{\log(T/\Delta)} \right). \end{aligned}$$

For the first term of (A.13), we have

$$\begin{aligned}\frac{1}{\Delta^2}[M]_T &= \frac{1}{\Delta^2}\sum_{i=1}^{N}\int_{i\Delta}^{(i+1)\Delta}((i+1)\Delta - s)^2\sigma^2(V_s)ds \\ &= \frac{1}{\Delta^2}\sum_{i=1}^{N}\sigma^2(V_{i\Delta})\int_{i\Delta}^{(i+1)\Delta}((i+1)\Delta - s)^2 ds + S_T \\ &= \frac{1}{3}\sum_{i=1}^{N}\sigma^2(V_{i\Delta})\Delta + S_T, \end{aligned} \tag{A.14}$$

where

$$\begin{aligned} S_T &= \frac{1}{\Delta^2}\sum_{i=1}^{N}\int_{i\Delta}^{(i+1)\Delta}((i+1)\Delta - s)^2(\sigma^2(V_s) - \sigma^2(V_{i\Delta}))ds \\ &= O_p\left(\Delta T(\sigma^{2'}\mu)T\right) + O_p\left(\Delta T(\sigma^{2''}\sigma^2)T\right) + O_p\left(\Delta T(\sigma^{2'}\sigma)T^{1/2}\right) \end{aligned} \tag{A.15}$$

by Lemma A.3. Moreover, it follows from Lemma B1 of Kim and Park (2017) that

$$\begin{aligned}\sum_{i=1}^{N}\sigma^2(V_{i\Delta})\Delta = \int_0^T \sigma^2(V_t)dt + O_p\left(\Delta T(\sigma^{2'}\mu)T\right) + O_p\left(\Delta T(\sigma^{2''}\sigma^2)T\right) \\ + O_p\left(\Delta T(\sigma^{2'}\sigma)T^{1/2}\right), \end{aligned}$$

from which, toghether with (A.14) and (A.15), we have

$$\begin{aligned} A_{2,T} =& \frac{1}{3}[V]_T + O_p\left(\Delta T(\sigma^{2'}\mu)T\right) + O_p\left(\Delta T(\sigma^{2''}\sigma^2)T\right) + O_p\left(\Delta T(\sigma^{2'}\sigma)T^{1/2}\right) \\ &+ O_p\left(\Delta^{1/2}T(\sigma^2)T^{1/2}\sqrt{\log(T/\Delta)}\right). \end{aligned} \tag{A.16}$$

Therefore, we can obtain (A.8) by applying (A.11), (A.12) and (A.16) to (A.10).

Proof for (A.9). We write

$$R_T = R_{1,T} + R_{2,T} + R_{3,T} + R_{4,T}, \tag{A.17}$$

where

$$R_{1,T} = \frac{2}{\Delta^2} \sum_{i=1}^{N} \left(\int_{(i-1)\Delta}^{i\Delta} \int_{(i-1)\Delta}^{t} \mu(V_s) ds dt \right) \left(\int_{i\Delta}^{(i+1)\Delta} \int_{i\Delta}^{t} \mu(V_s) ds dt \right)$$

$$R_{2,T} = \frac{2}{\Delta^2} \sum_{i=1}^{N} \left(\int_{(i-1)\Delta}^{i\Delta} \int_{(i-1)\Delta}^{t} \sigma(V_s) dW_s dt \right) \left(\int_{i\Delta}^{(i+1)\Delta} \int_{i\Delta}^{t} \mu(V_s) ds dt \right)$$

$$R_{3,T} = \frac{2}{\Delta^2} \sum_{i=1}^{N} \left(\int_{(i-1)\Delta}^{i\Delta} \int_{(i-1)\Delta}^{t} \mu(V_s) ds dt \right) \left(\int_{i\Delta}^{(i+1)\Delta} \int_{i\Delta}^{t} \sigma(V_s) dW_s dt \right)$$

$$R_{4,T} = \frac{2}{\Delta^2} \sum_{i=1}^{N} \left(\int_{(i-1)\Delta}^{i\Delta} \int_{(i-1)\Delta}^{t} \sigma(V_s) dW_s dt \right) \left(\int_{i\Delta}^{(i+1)\Delta} \int_{i\Delta}^{t} \sigma(V_s) dW_s dt \right).$$

We can easily show that

$$R_{1,T} = O_p \left(\Delta T(\mu^2) T \right). \tag{A.18}$$

Similarly as in (A.12), we have

$$R_{2,T}, R_{3,T} = O_p \left(\Delta^{1/2} T(\mu\sigma) T \sqrt{\log(T/\Delta)} \right). \tag{A.19}$$

By changing the order of integrals, we rewrite $R_{4,T}$ as

$$R_{4,T} = \frac{2}{\Delta^2} \sum_{i=1}^{N} \left(\int_{(i-1)\Delta}^{i\Delta} (i\Delta - s)\sigma(V_s) dW_s \right) \left(\int_{i\Delta}^{(i+1)\Delta} ((i+1)\Delta - s)\sigma(V_s) dW_s \right)$$

and define a continuous martingale M as

$$\begin{aligned} M_t =& \frac{2}{\Delta^2} \sum_{i=1}^{j-1} \left(\int_{(i-1)\Delta}^{i\Delta} (i\Delta - s)\sigma(V_s) dW_s \right) \left(\int_{i\Delta}^{(i+1)\Delta} ((i+1)\Delta - s)\sigma(V_s) dW_s \right) \\ &+ \frac{2}{\Delta^2} \left(\int_{(j-1)\Delta}^{j\Delta} (j\Delta - s)\sigma(V_s) dW_s \right) \left(\int_{j\Delta}^{t} ((j+1)\Delta - s)\sigma(V_s) dW_s \right) \end{aligned}$$

for $t \in [j\Delta, (j+1)\Delta)$, $j = 1, 2, \cdots, N$, so that we have $M_T = R_{4,T}$. Then we have

$$
\begin{aligned}
[M]_T &= \frac{4}{\Delta^4} \sum_{i=1}^{N} \left(\int_{(i-1)\Delta}^{i\Delta} (i\Delta - s)\sigma(V_s) dW_s \right)^2 \left(\int_{i\Delta}^{(i+1)\Delta} ((i+1)\Delta - s)^2 \sigma^2(V_s) ds \right) \\
&= O_p\left(\Delta T(\sigma^4) T \log(T/\Delta)\right)
\end{aligned}
$$

since

$$
\sup_{1 \le i \le N} \left(\int_{i\Delta}^{(i+1)\Delta} ((i+1)\Delta - s)^2 \sigma^2(V_s) ds \right) = O_p\left(\Delta^3 T(\sigma^2)\right)
$$

and

$$
\sum_{i=1}^{N} \left(\int_{(i-1)\Delta}^{i\Delta} (i\Delta - s)\sigma(V_s) dW_s \right)^2 = O_p\left(\Delta^2 T(\sigma^2) T \log(T/\Delta)\right),
$$

similarly as in (A.12). Therefore, we have $R_{4,T} = O_p\left(\Delta^{1/2} T(\sigma^2) T^{1/2} \sqrt{\log(T/\Delta)}\right)$, from which, together with (A.17)-(A.19), we have (A.9). □

Lemma A.5. *If* $\Delta^{1/2} T(\omega^{3/2}) T \sqrt{\log(T/\Delta)} \to_p 0$, *then we have*

$$
\sum_{i=k+1}^{N} (V_{i\Delta} - V_{(i-1)\Delta})(V_{(i-k)\Delta} - V_{(i-k-1)\Delta}) = o_p(1)
$$

for any positive integer $k \ge 1$.

Proof for Lemma A.5. We have

$$
\sum_{i=k+1}^{N} (V_{i\Delta} - V_{(i-1)\Delta})(V_{(i-k)\Delta} - V_{(i-k-1)\Delta}) = A_T + B_T + C_T + D_T, \tag{A.20}
$$

where

$$
\begin{aligned}
A_T &= \sum_{i=j+1}^{N} \left(\int_{(i-1)\Delta}^{i\Delta} \mu(V_s) ds \right) \left(\int_{(i-k-1)\Delta}^{(i-k)\Delta} \mu(V_s) ds \right) \\
B_T &= \sum_{i=j+1}^{N} \left(\int_{(i-1)\Delta}^{i\Delta} \sigma(V_s) dW_s \right) \left(\int_{(i-k-1)\Delta}^{(i-k)\Delta} \mu(V_s) ds \right) \\
C_T &= \sum_{i=j+1}^{N} \left(\int_{(i-1)\Delta}^{i\Delta} \mu(V_s) ds \right) \left(\int_{(i-k-1)\Delta}^{(i-k)\Delta} \sigma(V_s) dW_s \right) \\
D_T &= \sum_{i=j+1}^{N} \left(\int_{(i-1)\Delta}^{i\Delta} \sigma(V_s) dW_s \right) \left(\int_{(i-k-1)\Delta}^{(i-k)\Delta} \sigma(V_s) dW_s \right).
\end{aligned}
$$

For A_T, we have

$$
A_T = O_p\left(\Delta T(\mu^2) T\right). \tag{A.21}
$$

Moreover, we have

$$
B_T, C_T = O_p\left(\Delta^{1/2} T(\mu\sigma) T \sqrt{\log(T/\Delta)}\right). \tag{A.22}
$$

similarly as in (A.12).

As for D_T, we may show that

$$
D_T = O_p\left(\Delta^{1/2} T(\sigma^2) T^{1/2} \sqrt{\log(T/\Delta)}\right) \tag{A.23}
$$

similarly as in the proof for $R_{4,T}$ in (A.17). The stated result follows immediately from (A.20)-(A.23). □

Lemma A.6. *If* $\Delta^{1/2} T(\omega^{3/2}) T \sqrt{\log(T/\Delta)} \to_p 0$, *we have for* $k \geq 0$

$$
\sum_{i=k+1}^{N} (x_{i+1} - x_{i-k})^2 = \left(\frac{2}{3} + k \right) [V]_T + o_p(1).
$$

Proof for Lemma A.6. We write

$$\begin{aligned}
&\sum_{i=k+1}^{N}(x_{i+1}-x_{i-k})^2\\
&=\frac{1}{\Delta^2}\sum_{i=k+1}^{N}\left(\int_{i\Delta}^{(i+1)\Delta}(V_t-V_{i\Delta})dt+\int_{(i-k-1)\Delta}^{(i-k)\Delta}(V_{(i-k)\Delta}-V_t)dt+\Delta(V_{i\Delta}-V_{(i-k)\Delta})\right)^2\\
&=A_T+B_T+C_T+R_{1,T}+R_{2,T}+R_{3,T}, \qquad \text{(A.24)}
\end{aligned}$$

where

$$\begin{aligned}
A_T&=\frac{1}{\Delta^2}\sum_{i=k+1}^{N}\left(\int_{i\Delta}^{(i+1)\Delta}(V_t-V_{i\Delta})dt\right)^2\\
B_T&=\frac{1}{\Delta^2}\sum_{i=k+1}^{N}\left(\int_{(i-k-1)\Delta}^{(i-k)\Delta}(V_{(i-k)\Delta}-V_t)dt\right)^2\\
C_T&=\sum_{i=k+1}^{N}(V_{i\Delta}-V_{(i-k)\Delta})^2\\
R_{1,T}&=\frac{2}{\Delta^2}\sum_{i=k+1}^{N}\left(\int_{i\Delta}^{(i+1)\Delta}(V_t-V_{i\Delta})dt\right)\left(\int_{(i-k-1)\Delta}^{(i-k)\Delta}(V_{(i-k)\Delta}-V_t)\right)\\
R_{2,T}&=\frac{2}{\Delta}\sum_{i=k+1}^{N}\left(\int_{i\Delta}^{(i+1)\Delta}(V_t-V_{i\Delta})dt\right)(V_{i\Delta}-V_{(i-k)\Delta})\\
R_{3,T}&=\frac{2}{\Delta}\sum_{i=k+1}^{N}(V_{i\Delta}-V_{(i-k)\Delta})\left(\int_{(i-k-1)\Delta}^{(i-k)\Delta}(V_{(i-k)\Delta}-V_t)\right).
\end{aligned}$$

Similarly as in the proofs of (A.8) and (A.9) in Lemma A.4, we may show that

$$A_T,B_T=\frac{1}{3}[V]_T+O_p\left(\Delta^{1/2}T(\omega^{3/2})T\sqrt{\log(T/\Delta)}\right) \qquad \text{(A.25)}$$

$$R_{1,T},R_{2,T},R_{3,T}=O_p\left(\Delta^{1/2}T(\omega^{3/2})T\sqrt{\log(T/\Delta)}\right). \qquad \text{(A.26)}$$

As for C_T, we have

$$C_T = \sum_{j=0}^{k-1} \sum_{i=k+1}^{N} (V_{(i-j)\Delta} - V_{(i-j-1)\Delta})^2 + o_p(1) = k[V]_T + o_p(1), \tag{A.27}$$

where the first equality is due to Lemma A.5, and the last equality follows from Lemma A11 of Kim and Park (2016). The stated result is then follows from (A.24)-(A.27). □

B. Useful Lemmas for Realized Variance

Lemma B.1. *We have*

$$\sup_{1\le i\le N} |y_i - x_i| = O_p\left(\sqrt{(\delta/\Delta)T(\iota^2)\log(T/\Delta)}\right)$$
$$\sup_{1\le i\le N} |y_i^2 - x_i^2| = O_p\left((\delta/\Delta)T(\iota^2)\log(T/\Delta)\right) + O_p\left(\sqrt{(\delta/\Delta)T(\iota^4)\log(T/\Delta)}\right).$$

Proof for Lemma B.1. We have

$$\begin{aligned}\sup_{1\le i\le N} |y_i - x_i| = \sup_{1\le i\le N} |\eta_i g_i| &\le \sqrt{\frac{2\delta}{\Delta^2}} \sup_{1\le i\le N} \left(\frac{1}{\Delta}\int_{(i-1)\Delta}^{i\Delta} V_t^2 dt\right)^{1/2} |G_{i\Delta} - G_{(i-1)\Delta}| \\ &\le \sqrt{\frac{2\delta T(\iota^2)}{\Delta^2}} \sup_{1\le i\le N} |G_{i\Delta} - G_{(i-1)\Delta}| \\ &= O_p\left(\sqrt{(\delta/\Delta)T(\iota^2)\log(T/\Delta)}\right),\end{aligned} \tag{B.1}$$

where the last equality follows from the global modulus of continuity for Brownian motion.

On the other hand, we have

$$\sup_{1\le i\le N} |y_i^2 - x_i^2| \le \sup_{1\le i\le N} |\eta_i^2 g_i^2| + \sup_{1\le i\le N} |2x_i\eta_i g_i|,$$

and

$$\sup_{1\le i\le N}|\eta_i^2 g_i^2| = O_p\left((\delta/\Delta)T(\iota^2)\log(T/\Delta)\right),$$
$$\sup_{1\le i\le N}|2x_i\eta_i g_i| = O_p\left(\sqrt{(\delta/\Delta)T(\iota^4)\log(T/\Delta)}\right)$$

due to (B.1) and

$$\sup_{1\le i\le N}|x_i| = O_p(T(\iota)) + O_p\left(\Delta T(\mu)\right) + O_p\left(\Delta^{1/2}T(\sigma)\sqrt{\log(T/\Delta)}\right).$$

□

Lemma B.2. *We have*

$$\Delta\sum_{i=1}^N y_i = \Delta\sum_{i=1}^N x_i + O_p\left(\sqrt{(\delta/\Delta)\log(T/\Delta)}T\right)$$
$$\Delta\sum_{i=1}^N y_i^2 = \Delta\sum_{i=1}^N x_i^2 + O_p\left((\delta/\Delta)T(\iota^2)\log(T/\Delta)T\right) + O_p\left(\sqrt{(\delta/\Delta)T(\iota^4)\log(T/\Delta)}T\right).$$

Proof for Lemma B.2. The stated results follow immediately from Lemma B.1. □

Lemma B.3. *We have*

$$\sum_{i=1}^N (x_{i+1}-x_i)(\eta_{i+1}g_{i+1}-\eta_i g_i) = O_p\left((\delta/\Delta)^{1/2}T(\iota\sigma)T^{1/2}\right).$$

Proof for Lemma B.3. We write

$$\sum_{i=1}^N (x_{i+1}-x_i)(\eta_{i+1}g_{i+1}-\eta_i g_i) = A_T - B_T,$$

where

$$A_T = \sum_{i=1}^N \eta_{i+1}(x_{i+1}-x_i)g_{i+1} \quad \text{and} \quad B_T = \sum_{i=1}^N \eta_i(x_{i+1}-x_i)g_i.$$

In the following, we are to show

$$A_T, B_T = O_p\left((\delta/\Delta)^{1/2}\, T(\iota\sigma)T^{1/2}\right). \tag{B.2}$$

For A_T, we have

$$A_T = \frac{\sqrt{2\delta}}{\Delta}\sum_{i=1}^{N}\left(\frac{1}{\Delta}\int_{i\Delta}^{(i+1)\Delta} V_u^2 du\right)^{1/2}\left(\frac{1}{\Delta}\int_{i\Delta}^{(i+1)\Delta}(V_u - V_{u-\Delta})du\right)(G_{(i+1)\Delta} - G_{i\Delta})$$

We define a process M as

$$\begin{aligned} M_t =& \frac{\sqrt{2\delta}}{\Delta}\sum_{i=1}^{j-1}\left(\frac{1}{\Delta}\int_{i\Delta}^{(i+1)\Delta} V_u^2 du\right)^{1/2}\left(\frac{1}{\Delta}\int_{i\Delta}^{(i+1)\Delta}(V_u - V_{u-\Delta})du\right)(G_{(i+1)\Delta} - G_{i\Delta}) \\ &+ \frac{\sqrt{2\delta}}{\Delta}\int_{j\Delta}^{t}\left(\frac{1}{\Delta}\int_{j\Delta}^{t} V_u^2 du\right)^{1/2}\left(\frac{1}{\Delta}\int_{j\Delta}^{t}(V_u - V_{u-\Delta})du\right)dG_s \end{aligned}$$

for $t \in [j\Delta, (j+1)\Delta)$, $j = 1, 2, \cdots, N$, so that $A_T = M_T$. Since G and V are independent each other, M is a continuous martingale with a quadratic variation $[M]$ satisfying

$$\begin{aligned} [M]_T &= \frac{2\delta}{\Delta}\sum_{i=1}^{N}\left(\frac{1}{\Delta}\int_{i\Delta}^{(i+1)\Delta} V_u^2 du\right)\left(\frac{1}{\Delta}\int_{i\Delta}^{(i+1)\Delta}(V_u - V_{u-\Delta})du\right)^2 \\ &= O_p\left((\delta/\Delta)T(\iota^2)T(\sigma^2)T\right) \end{aligned} \tag{B.3}$$

due, in particular, to Lemma A.4. The desired result (B.2) for A_T follows immediately from (B.3) since $A_T = M_T$. The proof for B_T is entirely identical to the proof for A_T and omitted here. □

Lemma B.4. *We have*

$$\sum_{i=1}^{N}\eta_i g_i \eta_{i+1} g_{i+1} = O_p\left((\delta/\Delta^{3/2})T(\iota^2)\sqrt{T\log(T/\Delta)}\right)$$

Proof for Lemma B.4. We define a continuous martingale M as

$$\begin{aligned} M_t =& \frac{2\delta}{\Delta^2}\sum_{i=1}^{j-1}\left(\frac{1}{\Delta}\int_{(i-1)\Delta}^{i\Delta} V_u^2 du\right)^{1/2}(G_{i\Delta}-G_{(i-1)\Delta})\left(\frac{1}{\Delta}\int_{i\Delta}^{(i+1)\Delta} V_u^2 du\right)^{1/2}(G_{(i+1)\Delta}-G_{i\Delta}) \\ &+ \frac{2\delta}{\Delta^2}\int_{j\Delta}^{t}\left(\frac{1}{\Delta}\int_{(j-1)\Delta}^{t-\Delta} V_u^2 du\right)^{1/2}(G_{j\Delta}-G_{(j-1)\Delta})\left(\frac{1}{\Delta}\int_{j\Delta}^{t} V_u^2 du\right)^{1/2} dG_s \end{aligned}$$

for $t \in [j\Delta, (j+1)\Delta)$, $j = 1, 2, \cdots, N$, so that $M_T = \sum_{i=1}^N \eta_i g_i \eta_{i+1} g_{i+1}$. Using the global modulus of continuity for Brownian motion, we may show that the quadratic variation process $[M]$ satisfies

$$[M]_T = O_p\left((\delta^2/\Delta^3)T(\iota^4)T\log(T/\Delta)\right),$$

from which we have the stated result.

□

Lemma B.5. *Assume that*

$$\Delta T(\omega^2)\log(T/\Delta) \to_p 0, \quad (\delta/\Delta^{3/2})T(\omega^3)T\sqrt{\log(T/\Delta)} \to_p 0.$$

Then we have

$$\sum_{i=1}^N (y_{i+1}-y_i)^2 = \sum_{i=1}^N (x_{i+1}-x_i)^2 + \frac{4\delta}{\Delta^2}\int_0^T V_t^2 dt + o_p(1).$$

Proof for Lemma B.5. Due to Lemma B.3, we have

$$\sum_{i=1}^N (y_{i+1}-y_i)^2 = \sum_{i=1}^N (x_{i+1}-x_i)^2 + \sum_{i=1}^N (\eta_{i+1}g_{i+1}-\eta_i g_i)^2 + O_p\left((\delta/\Delta)^{1/2}T(\iota\sigma)T^{1/2}\right). \tag{B.4}$$

For the second term of (B.4), we write

$$\begin{aligned}\sum_{i=1}^{N}(\eta_{i+1}g_{i+1}-\eta_i g_i)^2 &= 2\sum_{i=1}^{N}\eta_i^2 g_i^2 + \left(\eta_{N+1}^2 g_{N+1}^2 - \eta_1^2 g_1^2\right) + 2\sum_{i=1}^{N}\eta_{i+1}g_{i+1}\eta_i g_i \\ &= 2\sum_{i=1}^{N}\eta_i^2 g_i^2 + o_p(1), \qquad \text{(B.5)}\end{aligned}$$

where the last line follows from Lemma B.4 with the condition in this lemma since we have

$$\sup_{1\le i\le N}|\eta_i g_i|^2 \le \sup_{1\le i\le N}\left|\frac{1}{\Delta}\int_{(i-1)\Delta}^{i\Delta} V_t^2 dt\right| \sup_{1\le i\le N}\left|G_{i\Delta}-G_{(i-1)\Delta}\right|^2 = O_p\left(\Delta T(\iota^2)\log(T/\Delta)\right)$$

due to the global modulus of continuity for Brownian motion.

For the first term of (B.5), we have

$$\begin{aligned}\sum_{i=1}^{N}\eta_i^2 g_i^2 &= \frac{2\delta}{\Delta^2}\sum_{i=1}^{N}\left(\frac{1}{\Delta}\int_{(i-1)\Delta}^{i\Delta} V_t^2 dt\right)(G_{i\Delta}-G_{(i-1)\Delta})^2 \\ &= \frac{2\delta}{\Delta^2}\sum_{i=1}^{N}\left(\frac{1}{\Delta}\int_{(i-1)\Delta}^{i\Delta} V_t^2 dt\right)(G_{i\Delta}-G_{(i-1)\Delta})^2 \\ &= \frac{2\delta}{\Delta^2}\sum_{i=1}^{N} V_{(i-1)\Delta}^2\Delta + A_T + B_T, \qquad \text{(B.6)}\end{aligned}$$

where

$$\begin{aligned}A_T &= \frac{2\delta}{\Delta^2}\sum_{i=1}^{N}\left(\frac{1}{\Delta}\int_{(i-1)\Delta}^{i\Delta}(V_t^2 - V_{(i-1)\Delta}^2)dt\right)\Delta \\ B_T &= \frac{4\delta}{\Delta^2}\sum_{i=1}^{N}\left(\frac{1}{\Delta}\int_{(i-1)\Delta}^{i\Delta} V_t^2 dt\right)\int_{(i-1)\Delta}^{i\Delta}(G_t - G_{(i-1)\Delta})dG_t.\end{aligned}$$

For A_T, we use Lemma A.1 to have

$$A_T = O_p\left((\delta/\Delta)T(\iota\mu)T\right) + O_p\left((\delta/\Delta^{3/2})T(\iota\sigma)T\sqrt{\log(T/\Delta)}\right). \qquad \text{(B.7)}$$

As for B_T, we define a continuous martingale M as

$$M_t =\frac{4\delta}{\Delta^2}\sum_{i=1}^{j-1}\left(\frac{1}{\Delta}\int_{(i-1)\Delta}^{i\Delta} V_u^2 du\right)\int_{(i-1)\Delta}^{i\Delta}(G_u - G_{(i-1)\Delta})dG_u + \frac{4\delta}{\Delta^2}\int_{j\Delta}^{t}\left(\frac{1}{\Delta}\int_{(j-1)\Delta}^{j\Delta} V_u^2 du\right)(G_t - G_{(j-1)\Delta})dG_s$$

for $t \in [j\Delta, (j+1)\Delta)$, $j = 1, 2, \cdots, N$, so that $A_T = M_T$. As in the proof of Lemma B.4, we may show that

$$[M]_T = O_p\left((\delta^2/\Delta^3)T(\iota^4)T\log(T/\Delta)\right),$$

from which we have

$$B_T = O_p\left((\delta/\Delta^{3/2})T(\iota^2)\sqrt{T\log(T/\Delta)}\right). \tag{B.8}$$

The stated result is then follows from (B.4)-(B.8) with Lemma B1 of Kim and Park (2017). □

Lemma B.6. *Let the conditions in Lemma B.5 hold. Then we have*

$$\sum_{i=k+1}^{N}(y_{i+1} - y_{i-k})^2 = \sum_{i=k+1}^{N}(x_{i+1} - x_{i-k})^2 + \frac{4\delta}{\Delta^2}\int_0^T V_t^2 dt + o_p(1)$$

Proof of Lemma B.6. Similarly as in (B.4), we have

$$\sum_{i=k+1}^{N}(y_{i+1} - y_{i-k})^2 = \sum_{i=k+1}^{N}(x_{i+1} - x_{i-k})^2 + \sum_{i=k+1}^{N}(\eta_{i+1}g_{i+1} - \eta_{i-k}g_{i-k})^2 + o_p(1). \tag{B.9}$$

Moreover, the second term in (B.9) satisfies

$$\sum_{i=k+1}^{N}(\eta_{i+1}g_{i+1} - \eta_{i-k}g_{i-k})^2 = 2\sum_{i=k+1}^{N}\eta_i^2 g_i^2 + o_p(1) = \frac{4\delta}{\Delta}\int_0^T V_t^2 dt + o_p(1).$$

similarly as in the proof of Lemmas B.5, from which, jointly with (B.9), we have the stated result. □

C. Proofs for Main Results

Proof for Lemma 3.1. The stated results for $z = v$ can be found in Lemma 3.1 of Kim and Park (2017), and hence, we prove this lemma for $z = x, y$ below. The part (a) follows from Lemmas A.4 and B.5, respectively, for $z = x$ and y. The part (b) can be obtained from Lemmas A.1-A.2 and Lemmas B.1-B.2, respectively, for $z = x$ and y. Moreover, the part (c) follows immediately from Lemmas A.2 and B.2, respectively, for $z = x$ and y. □

Proof for Proposition 3.2. The stated result in part (a) can be deduced from Lemma 3.1 with Ito's lemma since

$$\sum_{i=1}^{N}(z_i - \overline{z}_N)\triangle x_i = \frac{1}{2}\left(z_N^2 - z_1^2 - \overline{x}_N(z_N - z_1)\right) - \frac{1}{2}\sum_{i=1}^{N}(z_{i+1} - z_i)^2.$$

For $\hat{\tau}_z^2$ in the part (b), we write

$$\begin{aligned}(T/\Delta)\hat{\tau}_z^2 &= \sum_{i=1}^{N}\left(\triangle z_i - \hat{\alpha}_z - (\hat{\beta}_z - 1)z_i\right)^2 \\ &= \sum_{i=1}^{N}(\triangle z_i)^2 - N\left(\overline{\triangle z}_N\right)^2 - \frac{\left(\sum_{i=1}^{N}(z_i - \overline{z}_N)\triangle z_i\right)^2}{\sum_{i=1}^{N}(z_i - \overline{z}_N)^2}. \qquad \text{(C.1)}\end{aligned}$$

For the second term of (C.1), we may show that

$$\overline{\triangle z}_N = \frac{1}{N}(V_T - V_0) + R_{z,T},$$

where

$$\begin{aligned}R_{x,T} &= O_p\left(\Delta^2 T(\mu)/T\right) + O_p\left(\Delta^{3/2}T(\sigma)\sqrt{\log(T/\Delta)}/T\right), \\ R_{y,T} &= R_{x,T} + O_p\left(\sqrt{\delta\Delta\log(T/\Delta)}/T\right)\end{aligned}$$

due to Lemmas A.1 and B.1, and hence,

$$N\left(\overline{\triangle z}_N\right)^2 = O_p\left(\Delta T(\iota^2)/T\right). \qquad \text{(C.2)}$$

For the last term of (C.1), we note that

$$\Delta \sum_{i=1}^{N} (z_i - \overline{z}_N)^2 = O_p(b_T), \tag{C.3}$$

where b_T is defined in Section 4 and satisfies $T/b_T = O(1)$. Moreover, we have

$$\sum_{i=1}^{N} (z_i - \overline{z}_N) \triangle z_i \sim_p \begin{cases} \int_0^T (V_t - \overline{V}_T) dV_t, & \text{if } z = v \\ \int_0^T (V_t - \overline{V}_T) dV_t + (1/6)[V]_T, & \text{if } z = x \\ \int_0^T (V_t - \overline{V}_T) dV_t + (1/6)[V]_T - (2\delta/\Delta^2) \int_0^T V_t^2 dt, & \text{if } z = y \end{cases} \tag{C.4}$$

by the part (a) of this proposition. For each component in (C.4), we have $[V]_T = O_p(T(\sigma^2)T)$, $(2\delta/\Delta^2)\int_0^T V_t^2 dt = O_p\left((\delta/\Delta^2)T(\iota^2)T\right)$ and

$$\begin{aligned} \int_0^T (V_t - \overline{V}_T) dV_t &= \frac{1}{2}\left((V_T - \overline{V}_T)^2 - (V_0 - \overline{V}_T)^2 - \int_0^T \sigma^2(V_t) dt \right) \\ &= O_p\left(T(\iota^2)\right) + O_p\left(T(\sigma^2)T\right), \end{aligned}$$

from which, jointly with (C.3) and (C.4), we have

$$\frac{\left(\sum_{i=1}^{N} (z_i - \overline{z}_N) \triangle z_i\right)^2}{\sum_{i=1}^{N} (z_i - \overline{z}_N)^2} = \begin{cases} O_p\left(\Delta T(\omega^4)T\right), & \text{if } z = v, x \\ O_p\left(\Delta T(\omega^4)T\right) + O_p\left((\delta^2/\Delta^3)T(\omega^4)T\right), & \text{if } z = y. \end{cases} \tag{C.5}$$

The stated result for $\hat{\tau}_z^2$ follows from (C.1), (C.2) and (C.5) with Assumption 3.1.

As for R_z^2 in the part (b), we have

$$R_z^2 = \hat{\beta}_z^2 \frac{\sum_{i=1}^{N} (z_i - \overline{z}_N)^2}{\sum_{i=1}^{N} (z_{i+1} - \overline{z}_N)^2} = \hat{\beta}_z^2 \frac{\sum_{i=2}^{N} (z_i - \overline{z}_N)^2 + (z_1 - \overline{z}_N)^2}{\sum_{i=2}^{N} (z_i - \overline{z}_N)^2 + (z_{N+1} - \overline{z}_N)^2} \sim_p \hat{\beta}_z^2$$

as desired.

The stated result in part (c) follows immediately from Lemma 3.1 and Proposition 3.2 (a)-(b). □

Proof for Proposition 3.3. The part (a) follows immediately from Lemma B.6. We may show parts (b) and (c) similarly as in the proofs for parts (a) and (b) of Proposition 3.2. □

Proof for Proposition 3.4. The parts (a) and (b) follow immediately from Propositions 3.2 and 3.3 since

$$\frac{1}{\sum_{i=2}^N \tilde{z}_i^2}\left(\sum_{i=2}^N w_i \tilde{z}_{i+1}\right) = \frac{1}{\sum_{i=2}^N \tilde{z}_i^2}\begin{pmatrix} \sum_{i=2}^N \tilde{z}_i \tilde{z}_{i+1} \\ \sum_{i=2}^N \tilde{z}_{i-1}\tilde{z}_{i+1} \end{pmatrix},$$
$$\frac{1}{\sum_{i=2}^N \tilde{z}_i^2}\left(\sum_{i=2}^N w_i w_i'\right) = \frac{1}{\sum_{i=2}^N \tilde{z}_i^2}\begin{pmatrix} \sum_{i=2}^N \tilde{z}_i^2 & \sum_{i=2}^N \tilde{z}_i \tilde{z}_{i-1} \\ \sum_{i=2}^N \tilde{z}_i \tilde{z}_{i-1} & \sum_{i=2}^N \tilde{z}_{i-1}^2 \end{pmatrix}.$$

As for the part (c), we have

$$\hat{\beta}_{1,z} + \hat{\beta}_{2,z} \sim_p 2\frac{\hat{\beta}_z}{\hat{\beta}_z + 1} + \frac{\hat{\beta}_x - 1}{\hat{\beta}_x + 1}\frac{\hat{\beta}_v - 1}{\hat{\beta}_x - 1} \tag{C.6}$$

due to the part (b) of this proposition with Proposition 3.3 (b). In what follows, we prove the part (b) only for $z = x$, since the proofs for $z = v, y$ are entirely analogous.

Let $z = x$. It then follows from Proposition 3.2 and Taylor expansion that

$$\frac{\hat{\beta}_x}{\hat{\beta}_x + 1} \sim_p \frac{1 + \Delta(\gamma_v + \gamma_x)}{2 + \Delta(\gamma_v + \gamma_x)} \sim_p \frac{1}{2} + \frac{1}{4}\Delta(\gamma_v + \gamma_x),$$
$$\frac{\hat{\beta}_v - 1}{\hat{\beta}_x - 1} \sim_p \frac{\gamma_v}{\gamma_v + \gamma_x},$$
$$\frac{\hat{\beta}_x - 1}{\hat{\beta}_x + 1} \sim_p \frac{\Delta(\gamma_v + \gamma_x)}{2 + \Delta(\gamma_v + \gamma_x)} \sim_p \frac{1}{2}\Delta(\gamma_v + \gamma_x)$$

since $\Delta(\gamma_v + \gamma_x) \to_p 0$ under Assumption 3.1, from which, together with (C.6), we have the stated result in part (c) for $z = x$.

As for the part (d), we may show that

$$
\begin{aligned}
R_z^2 &\sim_p \hat{\beta}_{1,z}^2 + \hat{\beta}_{2,z}^2 + 2\hat{\beta}_{1,z}\hat{\beta}_{2,z}\frac{\sum_{i=2}^N (z_i - \overline{z}_N)(z_{i-1} - \overline{z}_N)}{\sum_{i=2}^N (z_{i+1} - \overline{z}_N)^2} \\
&\sim_p (\hat{\beta}_{1,z} + \hat{\beta}_{2,z})^2 + 2\hat{\beta}_{1,z}\hat{\beta}_{2,z}\frac{\sum_{i=2}^N (z_{i-1} - \overline{z}_N)\triangle z_{i-1}}{\sum_{i=2}^N (z_{i+1} - \overline{z}_N)^2} \\
&\sim_p (\hat{\beta}_{1,z} + \hat{\beta}_{2,z})^2 + 2\hat{\beta}_{1,z}\hat{\beta}_{2,z}(\hat{\beta}_z - 1)
\end{aligned}
$$

which completes the proof.

□

Proofs for Lemmas 4.1 and 4.2. See Lemmas 3.2 and 3.3 of Kim and Park (2016). □

Proofs for Theorems 4.3 and 4.4. The stated results follow immediately from Proposition 3.2 with Lemmas 4.1 and 4.2. □

Proof for Theorem 4.5. The stated results follow immediately from Proposition 3.4 with Lemmas 4.1 and 4.2. □

Proof for Proposition 4.6. The proof is essentially identical to the proofs for Lemma 3.1 and Propositions 3.2-3.4, and is omitted here. □

Proof for Theorem 4.7. The stated results can be deduced from Proposition 4.6 with Lemmas 4.1 and 4.2. □

References

Andersen, T. and T. Bollerslev (1998): "Answering the Skeptics: Yes, Standard Volatility Models Do Provide Accurate Forecasts," *International Economic Review*, 39, 885–905.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys (2001): "The Distribution of Realized Exchange Rate Volatility," *Journal of the American Statistical Association*, 96, 42–55.

Andersen, T. G., T. Bollerslev, and N. Meddahi (2004): "Analytical Evaluation of Volatility Forecasts," *International Economic Review*, 45, 1079–1110.

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.

Barndorff-Nielsen, O. E. and N. Shephard (2002): "Econometric Analysis of Realized Volatility and Its Use in Estimating Stochastic Volatility Models," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 64, 253–280.

——— (2004): "Econometric Analysis of Realized Covariation: High Frequency Based Covariance, Regression, and Correlation in Financial Economics," *Econometrica*, 72, 885–925.

——— (2006): "Impact of Jumps on Returns and Realized Variances: Econometric Analysis of Time-Deformed Lévy Processes," *Journal of Econometrics*, 131, 217–252.

Bingham, N. H., C. M. Goldie, and J. L. Teugels (1993): *Regular Variation*, Cambridge, UK: Cambridge University Press.

Bollerslev, T. (1986): "Generalized Autoregressive Conditional Heteroskedasticity," *Journal of Econometrics*, 31, 307–327.

Chambers, M. J. (2004): "Testing for Unit Roots with Flow Data and Varying Sampling Frequency," *Journal of Econometrics*, 119, 1–18.

Drost, F. C. and T. E. Nijman (1993): "Temporal Aggregation of Garch Processes," *Econometrica*, 61, 909–927.

Drost, F. C. and B. J. Werker (1996): "Closing the GARCH Gap: Continuous Time GARCH Modeling," *Journal of Econometrics*, 74, 31 – 57.

Engle, R. (1982): "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation," *Econometrica*, 50, 987–1007.

Hill, B. M. (1975): "A Simple General Approach to Inference About the Tail of a Distribution," *The Annals of Statistics*, 3, 1163–1174.

Hill, J. B. (2010): "On Tail Index Estimation for Dependent, Heterogeneous Data," *Econometric Theory*, 26, 1398–1436.

Hsing, T. (1991): "On Tail Index Estimation Using Dependent Data," *The Annals of Statistics*, 19, 1547–1569.

Ibragimov, M., R. Ibragimov, and J. Walden (2015): *Heavy-Tailed Distributions and Robustness in Economics and Finance*, Springer, New York: Lecture Notes in Statistics 214.

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

Kanaya, S., J. Kim, and J. Y. Park (2016): "Moduli of Continuity of Brownian Motion and Gaussian Processes over an Expanding Time Interval with an Application to Jump Thresholding," Working Paper, Indiana University.

Karatzas, I. and S. E. Shreve (1991): *Brownian Motion and Stochastic Calculus*, New York, New York: Springer-Verlag.

Kim, J. and J. Y. Park (2016): "Mean Reversion and Stationarity in Financial Time Series Generated from Diffusion Models," Working Paper.

——— (2017): "Asymptotics for Recurrent Diffusions with Application to High Frequency Regression," *Journal of Econometrics*, 196, 27–54.

Kurtz, T. G. and P. Protter (1991): "Weak Limit Theorems for Stochastic Integrals and Stochastic Differential Equations," *The Annals of Probability*, 19, 1035–1070.

Meddahi, N. (2001): "An Eigenfunction Approach for Volatility Modeling," CIRANO Working Paper.

——— (2003): "ARMA Representation of Integrated and Realized Variances," *Econometrics Journal*, 6, 335–356.

Mikosch, T. and C. Starica (2000): "Limit Theory for the Sample Autocorrelations and Extremes of a GARCH (1, 1) Process," *The Annals of Statistics*, 28, 1427–1451.

Nelson, D. B. (1990): "ARCH Models as Diffusion Approximations," *Journal of Econometrics*, 45, 7–38.

Resnick, S. and C. Stărică (1995): "Consistency of Hill's Estimator for Dependent Data," *Journal of Applied Probability*, 32, 139–167.

——— (1998): "Tail Index Estimation for Dependent Data," *The Annals of Applied Probability*, 8, 1156–1183.

Samorodnitsky, G., S. T. Rachev, J.-R. Kurz-Kim, and S. V. Stoyanov (2007): "Asymptotic Distributions of Unbiased Linear Estimators in the Presence of Heavy-Tailed Stochastic Regressors and Residuals," *Probability and Mathematical Statistics*, 27, 275–302.

Zhang, L., P. A. Mykland, and Y. Aït-Sahalia (2011): "Edgeworth Expansions for Realized Volatility and Related Estimators," *Journal of Econometrics*, 160, 190–203.

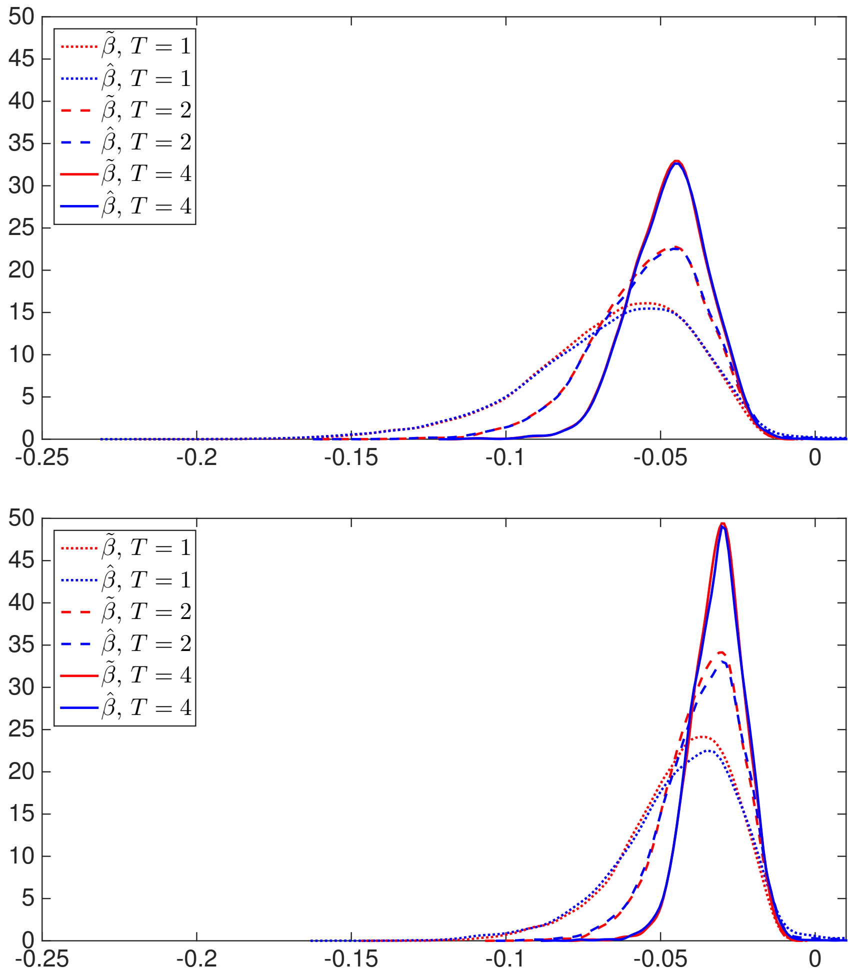

Fig. 1. (AR(1) with Finite Variance) The empirical distributions of $\hat{\beta}_z - 1$ (blue lines) and $\tilde{\beta}_z - 1$ (red lines), where $\tilde{\beta}$ is the approximated limit distribution. The dotted, dashed and solid lines represent respectively $T = 1, 2$ and 4. The upper and lower panels represent respectively the spot and integrated variances.

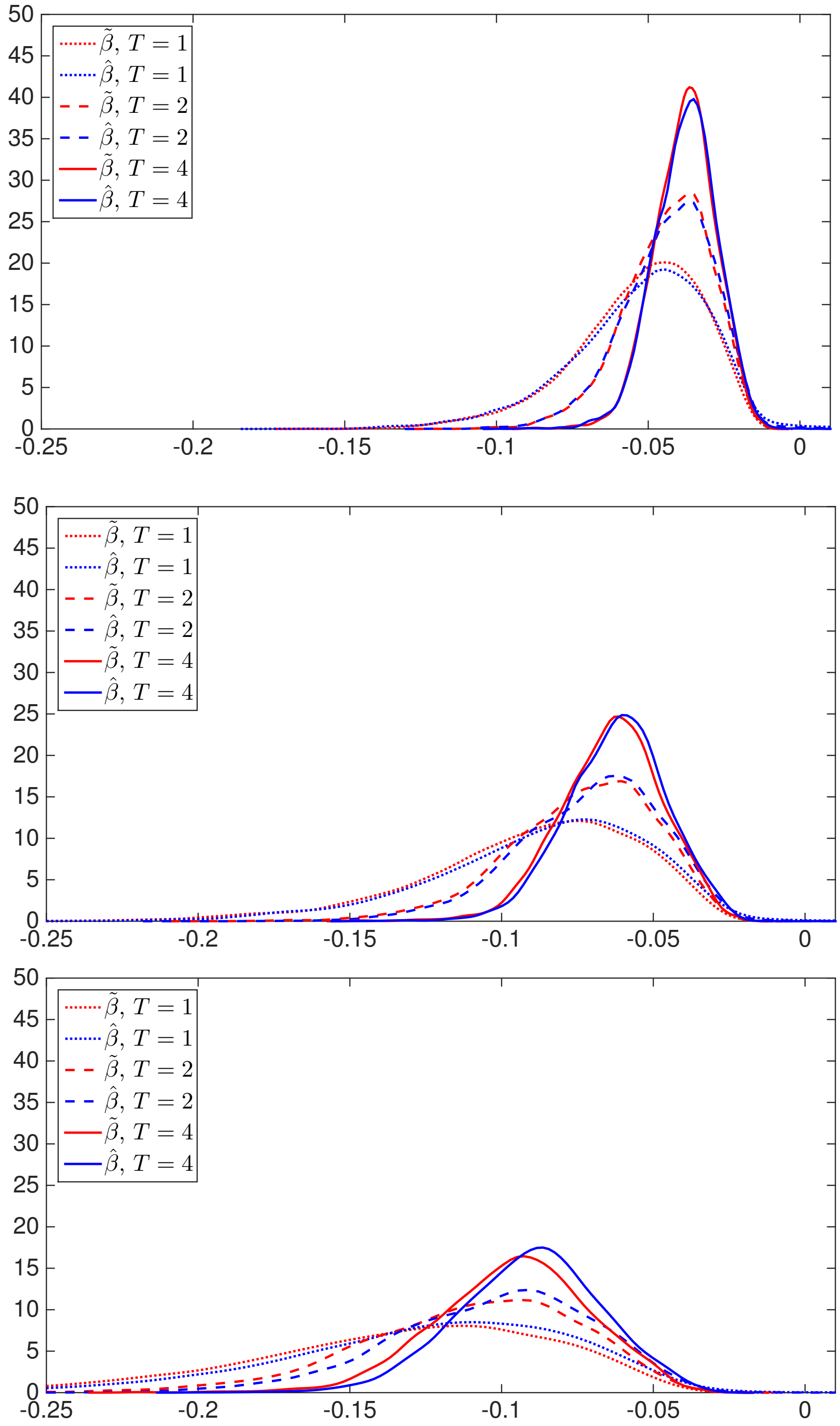

Fig. 2. (AR(1) with Finite Variance) The empirical distributions of $\hat{\beta}_z - 1$ (blue lines) and $\tilde{\beta}_z - 1$ (red lines), where $\tilde{\beta}$ is the approximated limit distribution. The dotted, dashed and solid lines represent respectively $T = 1, 2$ and 4. The upper, middle and lower panels represent the realized variances obtained from, respectively, 1, 5 and 10 minutes data.

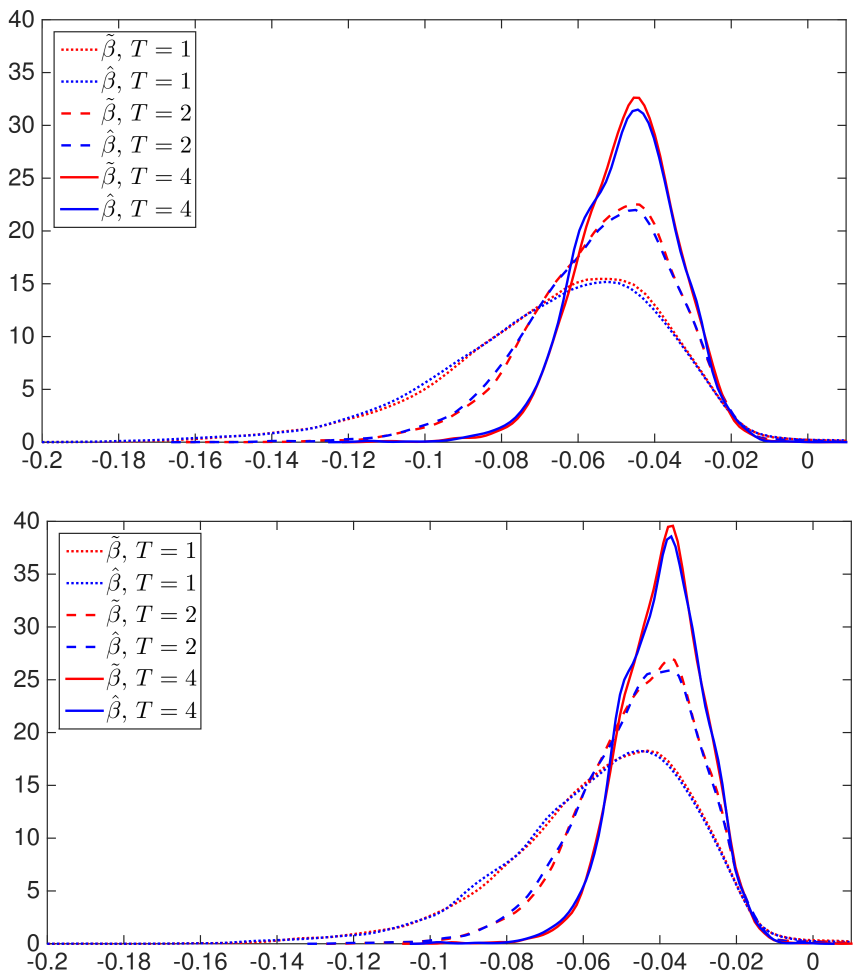

Fig. 3. (AR(2) with Finite Variance) The empirical distributions of $\hat{\beta}_z - 1$ (blue lines) and $\tilde{\beta}_z - 1$ (red lines), where $\tilde{\beta}$ is the approximated limit distribution. The dotted, dashed and solid lines represent respectively $T = 1, 2$ and 4. The upper and lower panels represent respectively the spot and integrated variances.

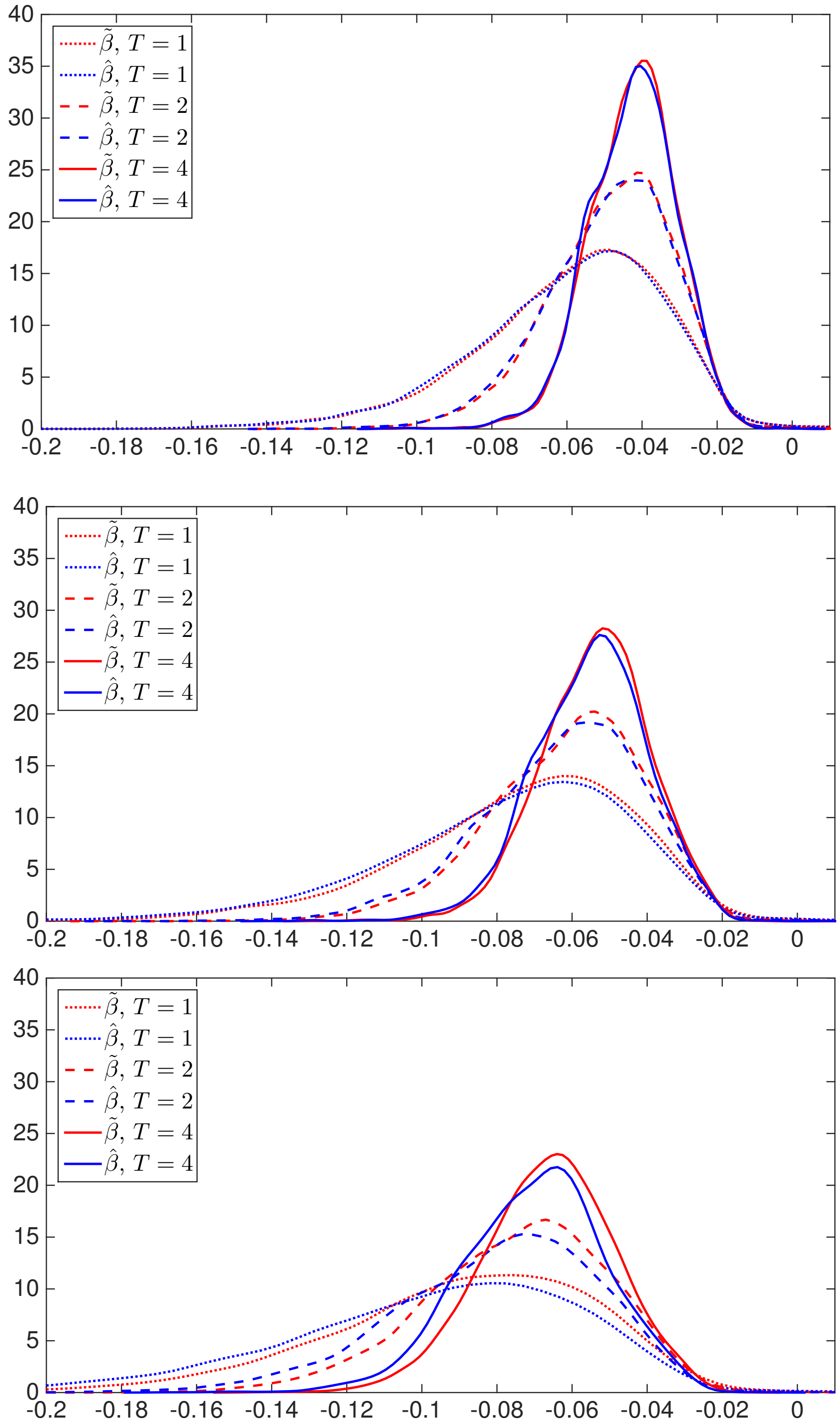

Fig. 4. (AR(2) with Finite Variance) The empirical distributions of $\hat{\beta}_z - 1$ (blue lines) and $\tilde{\beta}_z - 1$ (red lines), where $\tilde{\beta}$ is the approximated limit distribution. The dotted, dashed and solid lines represent respectively $T = 1, 2$ and 4. The upper, middle and lower panels represent the realized variances obtained from, respectively, 1, 5 and 10 minutes data.

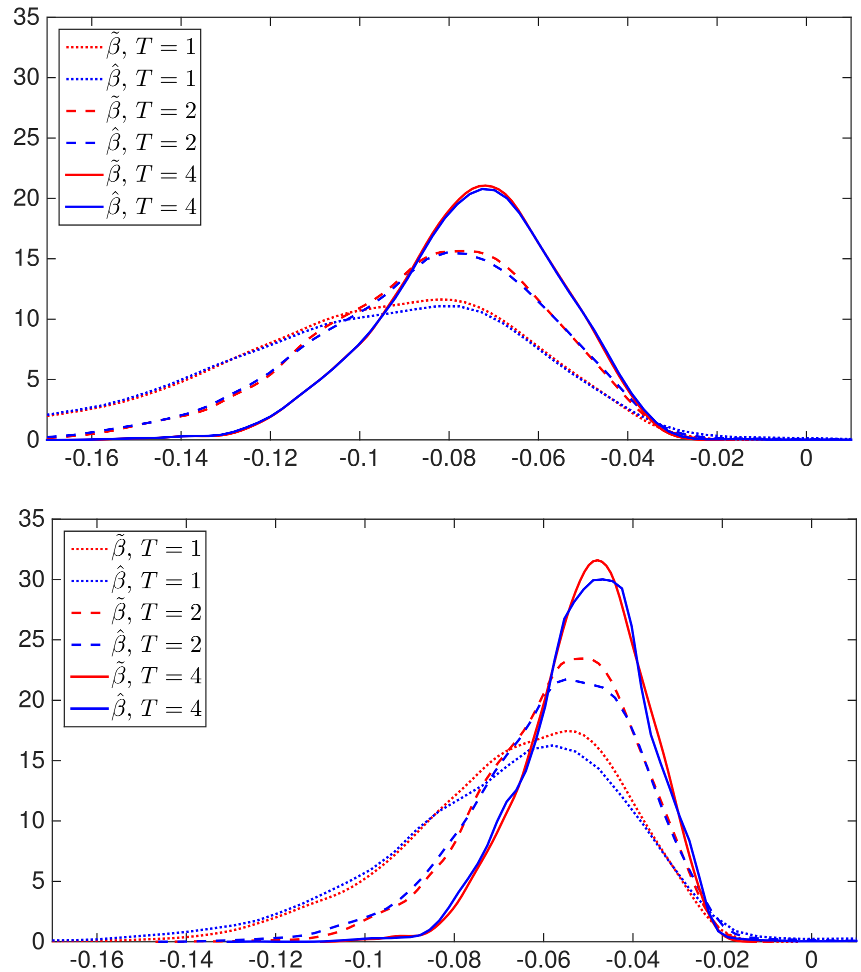

Fig. 5. (AR(1) with Ininite Variance) The empirical distributions of $\hat{\beta}_z - 1$ (blue lines) and $\tilde{\beta}_z - 1$ (red lines), where $\tilde{\beta}$ is the approximated limit distribution. The dotted, dashed and solid lines represent respectively $T = 1, 2$ and 4. The upper and lower panels represent respectively the spot and integrated variances.

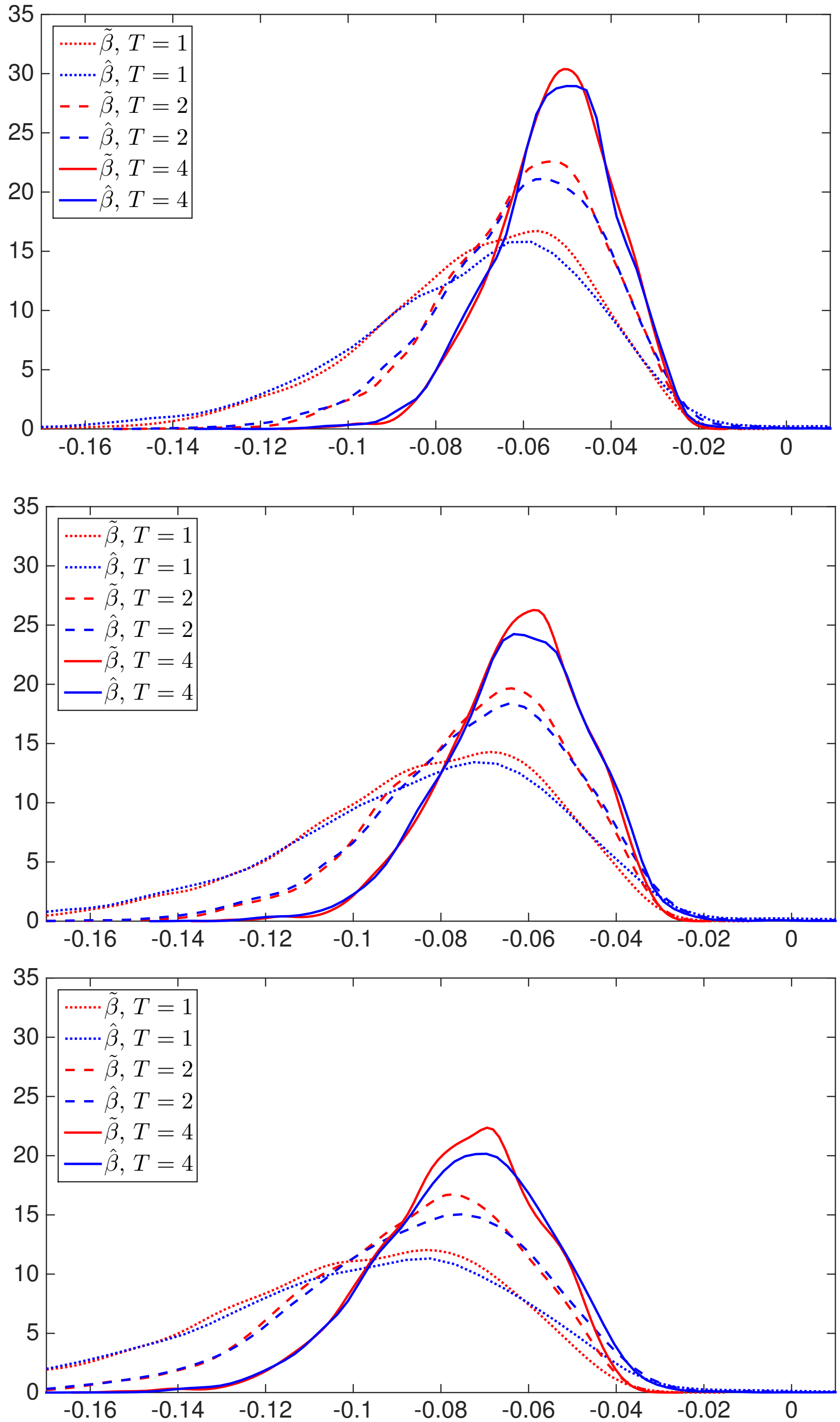

Fig. 6. (AR(1) with Ininite Variance) The empirical distributions of $\hat{\beta}_z - 1$ (blue lines) and $\tilde{\beta}_z - 1$ (red lines), where $\tilde{\beta}$ is the approximated limit distribution. The dotted, dashed and solid lines represent respectively $T = 1, 2$ and 4. The upper, middle and lower panels represent the realized variances obtained from, respectively, 1, 5 and 10 minutes data.

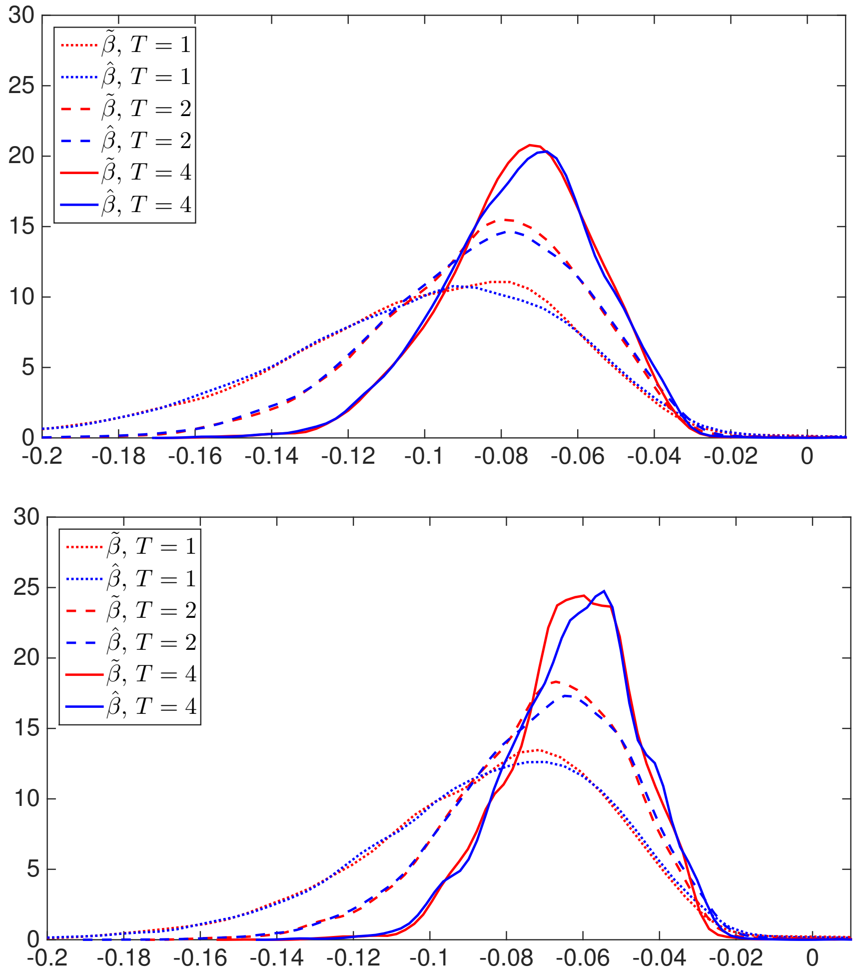

Fig. 7. (AR(2) with Ininite Variance) The empirical distributions of $\hat{\beta}_z - 1$ (blue lines) and $\tilde{\beta}_z - 1$ (red lines), where $\tilde{\beta}$ is the approximated limit distribution. The dotted, dashed and solid lines represent respectively $T = 1, 2$ and 4. The upper and lower panels represent respectively the spot and integrated variances.

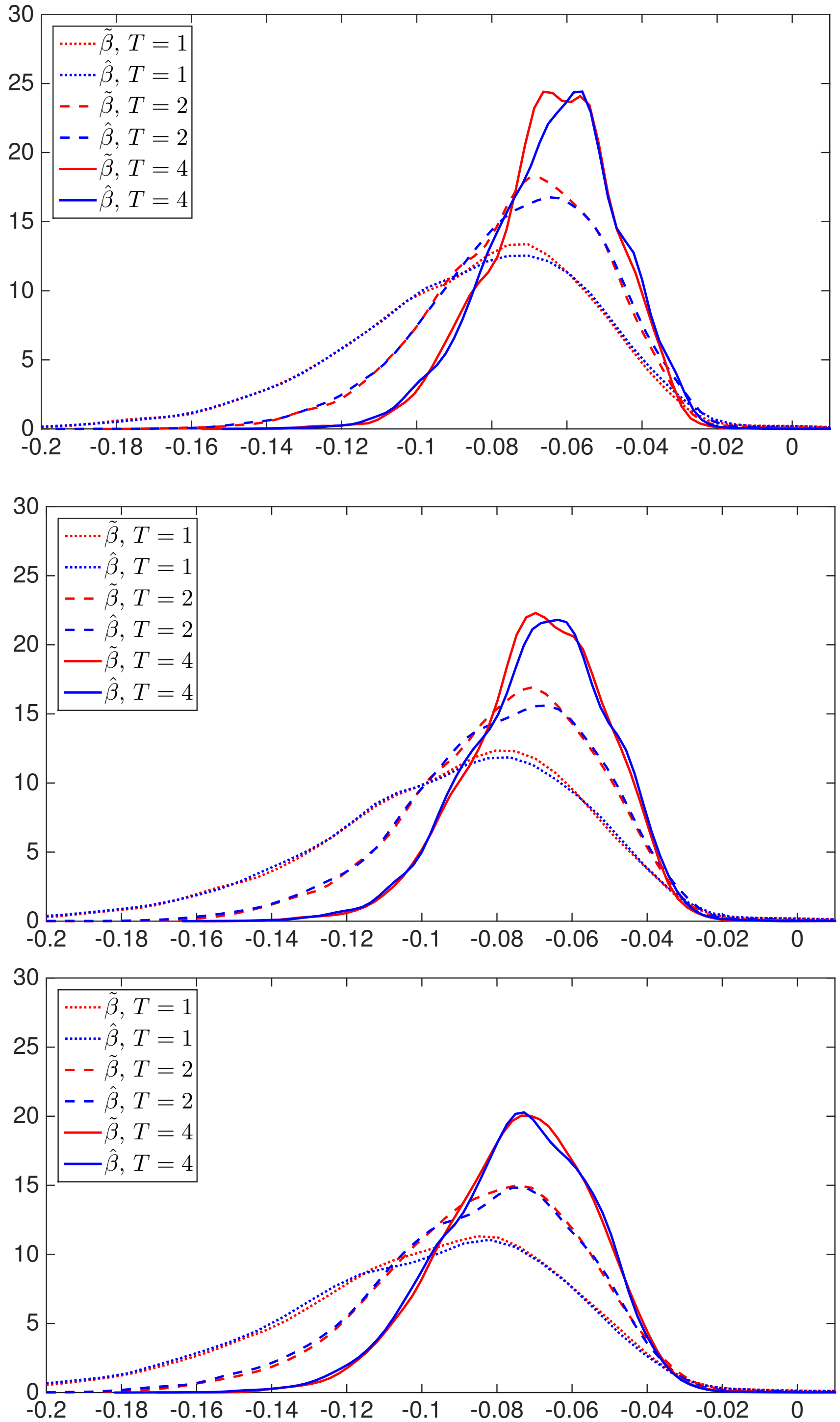

Fig. 8. (AR(2) with Ininite Variance) The empirical distributions of $\hat{\beta}_z - 1$ (blue lines) and $\tilde{\beta}_z - 1$ (red lines), where $\tilde{\beta}$ is the approximated limit distribution. The dotted, dashed and solid lines represent respectively $T = 1, 2$ and 4. The upper, middle and lower panels represent the realized variances obtained from, respectively, 1, 5 and 10 minutes data.