Journal of Economic Theory and Econometrics, Vol. 33, No. 2, Jun. 2022, 75–95

Nonstationary Volatility Regressions*

Jihyun Kim†

Abstract A popular approach to forecast variance is to use the fitted value of a simple OLS autoregression of realized variance measures. However, many financial returns are known to have highly persistent and possibly nonstationary volatilities. Under the nonstationarity, the asymptotic behaviors of the OLS estimators are unclear. We consider the autoregressions with spot, integrated, and realized variance measures when the spot variance process is nonstationary, and derive the asymptotic properties of the OLS estimators of the autoregressions. In particular, the asymptotic biases of the OLS estimators for the regressions with the integrated and realized variances are obtained. We then consider a feasible instrumental variable (IV) approach to reduce the bias of the OLS estimator, where the instrument equals the lagged value of the variable of interest, and show that the feasible IV estimator obtained from the realized variance is asymptotically equivalent to the infeasible OLS estimator obtained from the regression with the spot variance. Simulation results corroborate the theoretical findings of the paper.

Keywords volatility; autoregression; nonstationarity

JEL Classification C13, C22

*The author thank the two anonymous reviewers for their thoughtful comments and suggestions. The author is also grateful to Nour Meddahi, Changsik Kim, Heejoon Han, and Boreum Kwak for their helpful suggestions and discussions.

†Deparment of Economics, Sungkyunkwan University, E-mail: kim.jihyun@skku.edu

Received March 02, 2022, Revised April 09, 2022, Accepted April 19, 2022

1. INTRODUCTION

This paper considers a regression-based forecasting of future variance using high frequency variance measures. In practice, a popular approach is to approximate the latent spot or integrated variance by realized variance measures using high frequency data such as 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 a simple autoregressive regression of these realized measures by OLS to get a forecast of the spot or integrated variance. This approach based on autoregressive regressions is often misspecified because the dynamics of the spot and integrated variances are more complex than simple autoregressive processes. For instance, if the true spot variance follows a square-root process, then the integrated and realized variances become ARMA (1,1) processes (Barndorff-Nielsen and Shephard (2002); Meddahi (2003)). Still, even if the autoregressive model is not correctly specified, it provides a very accurate forecast because the integrated variance as well as high frequency-based realized measures are persistent and therefore, only few lags are sufficient to predict well future variance (Andersen *et al.*, 2003, 2004).

The properties of volatility regressions are well known under the assumption that the spot variance process is stationary. Andersen *et al.* (2004) consider autoregressions of the variance processes in population when the spot variance is stationary and has a finite second moment. The results of Andersen *et al.* (2004) are revisited by Kim and Meddahi (2020), in which empirical autoregressions of the variance processes are considered when the spot variance is stationary, but may have unbounded second moment. Both papers show that the OLS estimator of the feasible autoregressions with the realized variance measures is (asymptotically) biased, and may yield imprecise forecasts for the underlying variance measures such as spot or integrated variances. To solve the bias issue in the OLS-based forecasts, Kim and Meddahi (2020) propose an instrumental variable (IV) approach and show that the IV estimator has a smaller asymptotic bias compared to the OLS estimator under the stationarity.

Many key financial returns are known to have highly persistent and possibly nonstationary volatilities (see, e.g., Bollerslev and Wright (2000), Hansen and Lunde (2014) and Zhang and Han (2014)). It is also a well-known stylized fact that GARCH models fit to stock return data yield parameter estimates reflecting high persistence as they (nearly) violate stationarity conditions, and this phenomenon is often referred to as (near-)integrated GARCH (see, e.g., (Engle and Rangel, 2008)). However, the properties of the nonstationary volatility regressions are largely unknown. In particular, it is completely unknown about the asymptotic biases of the OLS estimators applied to the nonstationary volatility regressions with various variance measures as well as the validity of the IV approach of Kim and Meddahi (2020).

The goal of this paper is to fill the gaps between stationary and nonstationary volatility regressions by providing the asymptotic properties of the OLS and IV estimators for the nonstationary volatility regressions. First, we develop the asymptotics of the OLS estimators for the nonstationary volatility regressions, and provide the exact forms of the asymptotic biases of the OLS estimators. We then revisit the IV approach of Kim and Meddahi (2020), and show that the IV approach obtained from the realized variance remains valid and provides a robust estimation of the OLS regression estimator for the volatility regression with the spot variance. Simulation results corroborate the theoretical findings of the paper.

The paper is organized as follows. Section 2 provides the setup, various regressions, and the primary asymptotics. In Section 3, we provide the main results including the long span asymptotics of the OLS and IV estimators as well as simulations to assess the finite sample properties of the estimators. Section 4 concludes the paper, and all the proofs are provided in the Appendix.

Throughout the paper, we use "$P_T \sim Q_T$" to denote $P_T = Q_T(1+o(1))$. Similarly, "$P_T \sim_p Q_T$" and "$P_T \sim_d Q_T$" mean $P_T = Q_T(1+o_p(1))$ and $P_T =_d Q_T(1+o_p(1))$, respectively. These notations, as well as other standard notations used in asymptotics, will be used frequently throughout the paper without further references.

2. MODEL AND PRELIMINARIES

2.1. SPOT, INTEGRATED AND REALIZED VARIANCES

Consider a price process $(P_t, 0 \leq t \leq T)$ defined on a filtered probability space $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\geq 0}, \mathbb{P})$. Assume that P_t is a Brownian semimartingale with the following form:

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

where W_t^P is a Brownian motion, D_t and V_t are adapted processes with càdlàg paths. For a Δ-interval, we define the spot variance (v_i), integrated variance (x_i) and realized variance (y_i) of the price process (P_t) as

$$v_i = V_{i\Delta}, \qquad x_i = \frac{1}{\Delta}\int_{(i-1)\Delta}^{i\Delta} V_t dt, \qquad y_i = \frac{1}{\Delta}\sum_{j=1}^{n}\left(r^{(\delta)}_{(i-1)\Delta + j\delta}\right)^2, \tag{1}$$

for $i = 1,\cdots,N$ with $N\Delta = T$, where $r^{(\delta)}$ is the δ-period return defined as $r^{(\delta)}_{(i-1)\Delta+j\delta} = \log(P_{(i-1)\Delta+j\delta}) - \log(P_{(i-1)\Delta+(j-1)\delta})$ for $j = 1,\cdots,n$ with $n\delta = \Delta$. It is well known that the realized variance y is a noisy measure of the integrated variance x, and satisfies

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

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., **?**. For a fixed $T = N\Delta$, the convergence (2) holds jointly for $i = 1,\cdots,N$ (see, e.g., (Jacod and Protter, 1998)).

In this paper, we analyze the asymptotic properties of the least squares estimator for the volatility regression with a nonstationary variance process V. Specifically, we consider the autoregression

$$z_{i+1} = \alpha_z + \beta_z^{(k)} z_{i-k} + u_{i+1} \quad \text{with} \quad k \geq 0 \tag{3}$$

for $z = v,x,y$, and estimate the slope coefficient $\beta_z^{(k)}$ using the OLS estimator. Our asymptotics for the spot variance v and the integrated variance y involve two parameters, the sampling interval Δ and the time span T, and the asymptotic properties are developed under the assumption that $\Delta \to 0$ and $T \to \infty$ simultaneously. On the other hand, the asymptotics for the realized variance involve 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 are developed under the assumption that $\delta/\Delta \to 0$, $\Delta \to 0$ and $T \to \infty$ simultaneously.

As for the large T asymptotics in Section 3, we assume that the underlying variance process V is a diffusion process on $\mathscr{D} = (\underline{v},\overline{v}) \subset \mathbb{R}$ driven by

$$dV_t = \mu(V_t)dt + \sigma(V_t)dW_t, \tag{4}$$

where W is a Brownian motion, and μ and σ are respectively drift and diffusion functions of V. To obtain more explicit asymptotic results, we mainly consider a pure diffusion V without having leverage effects, i.e., each of V and D is independent of W^P.

We let s be the scale function defined as

$$s(v) = \int_y^v \exp\left(-\int_y^x \frac{2\mu(z)}{\sigma^2(z)}dz\right)dx, \tag{5}$$

where the lower limits of the integrals can be arbitrarily chosen to be any point $y \in \mathscr{D}$. 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

$$m(v)=\frac{1}{(\sigma^2 s')(v)} \tag{6}$$

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

Throughout this paper, we assume

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

Assumption 1 provides a simple sufficient set of conditions to ensure that a weak solution to the stochastic differential equation (4) exists uniquely up to an explosion time. See, e.g., Theorem 5.5.15 in Karatzas and Shreve (1991). Note, under Assumption 1, that both the scale function s and speed density m are well defined, and that the scale function is strictly increasing, on $\mathscr{D}$. Consequently, the natural scale diffusion V^s of V, where $V^s=s(V)$, is well defined with speed density $m_s=(m/s')\circ s^{-1}$. Moreover, under Assumption 1, the diffusion V is recurrent if and only if the scale function s is unbounded at both boundaries, i.e., $s(\underline{v})=-\infty$ and $s(\overline{v})=\infty$. A diffusion which is not recurrent is said to be transient. Furthermore, the recurrent diffusion V becomes positive or null recurrent depending upon

$$m(\mathscr{D})<\infty \quad \text{or} \quad m(\mathscr{D})=\infty.$$

2.2. PRIMARY ASYMPTOTICS

The OLS estimator $\hat{\beta}_z^{(k)}$ for $\beta_z^{(k)}$ in the autoregression (3) is given by

$$\hat{\beta}_z^{(k)}=\frac{\sum_{i=k+1}^{N-1}(z_{i-k}-\bar{z}_N)z_{i+1}}{\sum_{i=k+1}^{N-1}(z_{i-k}-\bar{z}_N)^2}$$

where $\bar{z}_N$ is the sample mean of $(z_{i-k}: i=k+1,\cdots,N-1)$. For $k=0$, we simply write $\beta_z^{(0)}=\beta_z$ and $\hat{\beta}_z^{(0)}=\hat{\beta}_z$.

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 moderately long span.

In our asymptotics, various functional transforms of D and V over time interval $[0,T]$ need to be properly controlled. To effectively handle such functional transforms of D and V, we define

$$T_D = \max_{0\le t\le T} |D_t| \quad \text{and} \quad T_V(f) = \max_{0\le t\le T} |f(V_t)|$$

for some function $f : \mathscr{D} \to \mathbb{R}$. We also denote by ι the identity function on $\mathscr{D}$, and $\iota(v) = v$ for all $v \in \mathscr{D}$. Consequently, we have $T_V(\iota) = \max_{0\le t\le T} |V_t|$ for the identity function. Obviously, T_D and $T_V(\iota)$ are the asymptotic orders of extremal process of D and V, respectively. The order of the extremal process is known for a wide class of diffusions (see, e.g., Kim and Park (2017) for a nonstationary diffusion, and Davis (1982) for a stationary diffusion). More generally, we may obtain the exact rate of $T_V(f)$ from the asymptotic behavior of extremal process. In particular, if f is regularly varying and c_T is the order of the extremal process, then the asymptotic order of $T_V(f)$ is given by $O_p(f(c_T))$.

Assumption 2. *(a) σ^2 is twice continuously differentiable on $\mathscr{D}$, and (b) for $f = \mu$, σ^2, $\sigma^{2'}$, $\sigma^{2''}$ and ι, there is a locally bounded function $\omega : \mathscr{D} \to \mathbb{R}$ such that $|f(v)| \le \omega(v)$ for all $v \in \mathscr{D}$.*

The differentiability condition of σ^2 in Assumption 2 (a) is routinely assumed in the study of diffusion models. Under Assumption 2 (a), the majorizing function ω in Assumption 2 (b) always exists as long as μ is locally bounded.

Assumption 3. *For ω in Assumption 2, $\Delta T_V(\omega^8)T^2 \log(T/\Delta) \to_p 0$.*

Assumption 4. *For ω in Assumption 2, $(\delta/\Delta)T_V(\omega^8)T^2 \log^3(T/\delta) \to_p 0$.*

Assumption 5. $(\delta/\Delta)T_D^4 T \to_p 0$.

Assumption 6. $(\delta/\Delta^2) = O(1)$.

Assumption 3 is similar to Assumption 5.1 in Kim and Park (2017), and provides a sufficient condition for our primary asymptotics of spot variance (v_i) and integrated variance (x_i). On the other hand, the asymptotics of realized variance (y_i) involve three parameters, δ, Δ and T, and require Assumptions 4-6 in addition to Assumption 3. Those assumptions are introduced in Kim and Meddahi (2020). The role of Assumption 4 is to analyze the asymptotic effect of the errors $(x_i - y_i)$ in the OLS estimates. On the other hand, Assumption 5 is a condition to handle the effects from the drift part (D_t) in (P_t) so that (D_t) has only asymptotically negligible effects to the asymptotics of the OLS estimates with (y_i). Lastly, Assumption 6 is to exclude less interesting situations where the errors $(x_i - y_i)$ dominate the signals (x_i) in the OLS estimates with (y_i). In particular, if $\delta/\Delta^2 \to \infty$, then the error components may have bigger stochastic order than the signals.

Under Assumptions 2-6, Kim and Meddahi (2020) develop the primary asymptotics for $\hat{\beta}_z^{(k)}$ as follows.

Proposition 1. *Let Assumptions 2-6 hold. As $\Delta \to 0$ and $\delta/\Delta \to 0$, we have*

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

and $\hat{\beta}_z^{(k)} - 1 \sim_p (\hat{\beta}_z - 1) + k(\hat{\beta}_v - 1)$ for all large T, where $([V]_t, 0 \le t \le T)$ is the quadratic variation process of $(V_t, 0 \le t \le T)$, and $\overline{V}_T = T^{-1} \int_0^T V_t dt$.

The primary asymptotics in Proposition 1 do not require the long span assumption of $T \to \infty$. In particular, if T is fixed, then $(N/T)(\hat{\beta}_z - 1) = (1/\Delta)(\hat{\beta}_z - 1)$ is random for all $z = v, x, y$, and is characterized by a particular realization of the underlying variance process V. Under the fixed T asymptotic scheme, the law of motion of the variance process V is less important in the asymptotics of $\hat{\beta}_z$ for all $z = v, x, y$. Importantly, the primary asymptotics in Proposition 1 require neither certain moment conditions nor stationarity. However, the underlying probabilistic structure of V becomes crucial in the development of the large T asymptotics of $\hat{\beta}_z$. In particular, the large T asymptotics depend heavily on the stationarity of V. In the following section, we develop the large T asymptotics of $\hat{\beta}_z$ when V is nonstationary.[1]

[1] The reader is referred to Kim and Meddahi (2020) for the large T asymptotics of $\hat{\beta}_z$ when V is stationary.

3. MAIN RESULTS

3.1. LONG SPAN ASYMPTOTICS

In our large T asymptotics, we only consider a recurrent nonstationary diffusion V to effectively analyze consequences of the nonstationarity in the volatility regressions. Accordingly, we assume that the scale function s in (5) and the speed density m in (6) satisfy the following conditions:

Assumption 7. *(a) $s(\underline{v}) = -\infty$ and $s(\overline{v}) = \infty$, and (b) $\int_{\mathscr{D}} m(v)dv = \infty$.*

Assumption 7 (a) implies that the diffusion process V is recurrent. Moreover, under Assumption 7 (b), jointly with (a), V becomes null recurrent and is nonstationary. A prime example of null recurrent diffusions is a Brownian motion. Null recurrent diffusions do not have time invariant distributions. Note, in particular, that Kim and Meddahi (2020) assume $\int_{\mathscr{D}} m(v)dv < \infty$, which implies that V is positive recurrent and has a time invariant distribution.[2]

In our asymptotics, we frequently deal with nonintegrable functions with respect to the speed density m of V. To effectively analyze m-nonintegrable functions, we need some regularity conditions. Following Kim and Park (2017), it will be maintained throughout the paper that all m-nonintegrable functions f are m-regularly varying, i.e., mf is regularly varying on $\mathscr{D}$. For a m-nonintegrable function f, we say that f is m-strongly nonintegrable if $f\ell$ is not m-integrable for any slowly varying function ℓ on $\mathscr{D}$. On the other hand, we say that f is m-nearly integrable if $f\ell$ is m-integrable for some slowly varying function ℓ on $\mathscr{D}$. Moreover, we let $f_s = f \circ s^{-1}$ for any function f on $\mathscr{D}$ other than m.[3] 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 $v \neq 0$.

We assume that

Assumption 8. *(a) s' is regularly varying or rapidly varying with index $c \neq -1$, (b) σ^2 is regularly varying, (c) σ^2 is m-strongly nonintegrable, and (d) m is strongly nonintegrable.*

Assumption 8 (a) and (b) appear in Kim and Park (2018), and are mild enough to include most diffusion processes used in practice. Under Assumption

[2] The time invariant density of the positive recurrent diffusion V is given by $\pi(v) = m(v)/\int_{\mathscr{D}} m(v)dv$. Positive recurrent diffusions become stationary if they are started from the time invariant distributions.

[3] In Section 2.1, m_s is defined as $m_s = (m/s') \circ s^{-1}$ which is the speed density of natural scale diffusion $V^s = s(V)$ of the underlying diffusion V.

8 (a), in particular, s^{-1} is regularly varying and its limit homogeneous function $\overline{s^{-1}}$ is well defined. The reader is also referred to Bingham *et al.* (1993) for more discussions about the regularly and rapidly varying functions. In Assumption 8 (c), we assume that σ^2 is m-strongly nonintegrable. This assumption is a technical condition to simplify our discussions below. Our subsequent theory can also be developed under the m-near integrablility and m-integrability at the cost of more involved analysis (see Kim and Park (2017, 2018) for the related discussions). Assumption 8 (d) is to exclude the barely nonstationary case introduced in Kim and Park (2017). We note that the asymptotics of a barely nonstationary process are similar to those of a stationary process.[4]

We let (λ_T) be the normalizing sequence satisfying $T = \lambda_T^2 m_s(\lambda_T)$, where m_s is the speed density of natural scale diffusion $V^s = s(V)$ of the underlying diffusion V. We define V^T by $V_t^T = V_{Tt}/s^{-1}(\lambda_T)$ for $t \in [0,1]$ with the normalizing sequence λ_T. It then follows from Proposition 3.2 of Kim and Park (2017) that $V^T \to_d V^\circ$ 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 V° as

$$V_t^\circ = \overline{s^{-1}}(B \circ \overline{A}_t), \quad \text{with} \quad \overline{A}_t = \inf\left\{ s \Big| \int_{\mathbb{R}} L(s,x)\overline{m}_s(dx) > t \right\},$$

where, in particular, $\overline{s^{-1}}$ is well defined under Assumption 8 (a), and $\overline{m}_s$ is well defined under Assumption 7 (b) and Assumption 8 (a) and (d).

The long span asymptotics for $\hat{\beta}_z$ can be obtained by letting $T \to \infty$ to the primary asymptotics in Proposition 1.

Theorem 1. *Let Assumptions 2-8 hold. As $\delta/\Delta, \Delta \to 0$ and $T \to \infty$, we have*

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

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

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

where $([V^\circ]_t, 0 \le t \le 1)$ is the quadratic variation process of $(V_t^\circ, 0 \le t \le 1)$, and $\overline{V}_1^\circ = \int_0^1 V_t^\circ dt$.

[4]The reader is referred to Sections 2 and 3 of Kim and Park (2017) for a precise definition of the barely nonstationarity and related discussions.

Remark 2. *The stationary and nonstationary volatility regressions have some similarities and differences when they are estimated by the OLS estimator.*

(a) As in the stationary volatility regressions in Kim and Meddahi (2020), the OLS estimators in the nonstationary regressions satisfy $\hat{\beta}_v < \hat{\beta}_x$ *and* $\hat{\beta}_y < \hat{\beta}_x$ *with probability approaching one. As noted in Kim and Meddahi (2020), the integrated variance* (x_i) *has a smoother sample path than that of the spot variance* (v_i)*, which yields* $\hat{\beta}_v < \hat{\beta}_x$*. On the other hand, the realized variance* (y_i) *is a noisy measure of* (x_i)*, and the noise induces downward bias with* $\hat{\beta}_y < \hat{\beta}_x$*.*

(b) The limit of normalized $\hat{\beta}_z - 1$ *is always random in the nonstationary volatility regressions. In the stationary volatility regressions, however, the normalized* $\hat{\beta}_z - 1$ *has a non-random limit as long as some moment conditions are satisfied. As an illustration, we consider a stationary Ornstein-Uhlenbeck process* V*, given as*

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

for which we have $\mathbb{E}(\sigma^2(V_t)) = \sigma^2$*,* $Var(V_t) = \sigma^2/(2\kappa)$ *and* $\mathbb{E}(V_t^2) = \sigma^2/(2\kappa) + \mu^2$*. It is shown in Kim and Meddahi (2020) (see Remark 3.3 (b)) that*

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

However, if V *is a Brownian motion with* $V = \sigma W$*, then* V *is nonstationary and*

$$N(\hat{\beta}_v - 1) \sim_d \frac{\int_0^1 (W_t - \overline{W}_1)dW_t}{\int_0^1 (W_t - \overline{W}_1)^2 dt}, \qquad N(\hat{\beta}_x - 1) \sim_d \frac{\int_0^1 (W_t - \overline{W}_1)dW_t + (1/6)}{\int_0^1 (W_t - \overline{W}_1)^2 dt},$$
$$N(\hat{\beta}_y - 1) \sim_d \frac{\int_0^1 (W_t - \overline{W}_1)dW_t + (1/6) + 2(\delta/\Delta)\int_0^1 W_t^2 dt}{\int_0^1 (W_t - \overline{W}_1)^2 dt}.$$

Clearly, the limit of normalized $\hat{\beta}_z - 1$ *is non-random for the stationary Ornstein-Uhlenbeck process* V*, whereas it becomes random for the Brownian* V*. Moreover, the OLS estimators in the nonstationary volatility regressions have faster rates of convergence than those of the stationary volatility regressions. We also note that* $N(\hat{\beta}_v - 1)$ *converges to the Dickey-Fuller distribution, whereas* $N(\hat{\beta}_x - 1)$ *and* $N(\hat{\beta}_y - 1)$ *are asymptotically biased from* $N(\hat{\beta}_v - 1)$ *and do not converge to the Dickey-Fuller distribution.*

In the volatility regressions, a feasible regression is the regression with the realized variance (y_i) since neither the spot variance (v_i) nor the integrated variance (x_i) is observable in reality. Moreover, researchers and financial practitioners are often interested in the regression with the spot or integrated variance depending upon their objectives. Our asymptotic results in Theorem 1 for the nonstationary volatility regressions imply that the feasible regression with (y_i) is asymptotically biased and may provide misleading forecasts for (v_i) or (x_i) as in the stationary volatility regression (see Theorem 3.4 of Kim and Meddahi (2020)). In Section 3.2, we show that the IV approach proposed by Kim and Meddahi (2020) reduces the asymptotic bias of the OLS estimator for the nonstationary volatility regressions. Combined with the results in Kim and Meddahi (2020), we may say that the IV approach is a robust method for the bias correction which can be applicable for both stationary and nonstationary volatility regressions.

3.2. A ROBUST ESTIMATION: KIM AND MEDDAHI'S IV APPROACH

The bias issue of stationary volatility regressions is well known in the literature (see Andersen *et al.* (2004); Kim and Meddahi (2020)). To overcome the bias issue, an instrumental variable (IV) approach has been proposed by Kim and Meddahi (2020). However, the validity of the IV method has been unknown for the nonstationary volatility regressions. Below we show that the IV approach proposed by Kim and Meddahi (2020) remains valid for the nonstationary volatility regressions.

Consider an IV estimator

$$\check{\beta}_z = \frac{\sum_{i=2}^{N-1}(z_{i-1}-\bar{z}_N)(z_{i+1}-\bar{z}_N)}{\sum_{i=2}^{N-1}(z_{i-1}-\bar{z}_N)(z_i-\bar{z}_N)},$$

which is the IV estimator with an instrument $z_{i-1}-\bar{z}_N$ for $z_i-\bar{z}_N$. Note that Kim and Meddahi (2020) consider a more general class of IV estimators with the instrument being nonlinear transformations of the lagged value of the variable of interest, and the IV estimator $\check{\beta}_z$ above is a special case of Kim and Meddahi (2020)'s nonlinear IV approach with the transformation being the identity function. Below we show that the IV estimator $\check{\beta}_z$ has a smaller asymptotic bias than the OLS estimator $\hat{\beta}_z$ for the nonstationary volatility regressions. It is interesting to consider the asymptotic behavior of the general nonlinear IV approach and to analyze the efficiency comparisons among various choices of transformations. However, the analysis of the efficiency or optimality is beyond of the scope of this paper, and we leave them for future research.

For the asymptotics of $\check{\beta}_z$, we write

$$\begin{aligned}\check{\beta}_z - 1 &= \frac{\sum_{i=2}^{N-1}(z_{i-1}-\bar{z}_N)(z_{i+1}-z_i)}{\sum_{i=2}^{N-1}(z_{i-1}-\bar{z}_N)(z_i-\bar{z}_N)} \\ &= \frac{\sum_{i=2}^{N-1}(z_{i-1}-\bar{z}_N)(z_{i+1}-z_{i-1})}{\sum_{i=2}^{N-1}(z_{i-1}-\bar{z}_N)(z_i-\bar{z}_N)} - \frac{\sum_{i=2}^{N-1}(z_{i-1}-\bar{z}_N)(z_i-z_{i-1})}{\sum_{i=2}^{N-1}(z_{i-1}-\bar{z}_N)(z_i-\bar{z}_N)} \equiv \phi_z - \psi_z.\end{aligned}$$

The following proposition provides the asymptotic properties of ϕ_z and ψ_z as well as $\hat{\beta}_z - 1$ for $z = v, x, y$.

Proposition 2. *Let Assumptions 2-8 hold, and let $\delta/\Delta, \Delta \to 0$ and $T \to \infty$. For $z = v, x, y$, we have $\phi_z \sim_p \hat{\beta}_z^{(1)} - 1$ and $\psi_z \sim_p \hat{\beta}_z - 1$ with $\hat{\beta}_z^{(1)} - 1 \sim_p (\hat{\beta}_z - 1) + (\hat{\beta}_v - 1)$, and therefore,*

$$\check{\beta}_z - 1 \sim_p \hat{\beta}_v - 1 \sim_d \frac{\Delta}{T} \frac{\int_0^1 (V_t^\circ - \overline{V}_1^\circ) dV_t^\circ}{\int_0^1 (V_t^\circ - \overline{V}_1^\circ)^2 dt}.$$

Unlike $\hat{\beta}_z - 1$, the limits of $\check{\beta}_z - 1$ are given by $\hat{\beta}_v - 1$ (up to the order of Δ/T) for all $z = v, x, y$. In particular, Proposition 2 provides the asymptotic equivalence between the feasible OLS estimator $\check{\beta}_y$ and the infeasible OLS estimator $\hat{\beta}_v$ up to the order of Δ/T. Moreover, we may expect that the feasible IV estimator $\check{\beta}_y$ is more closer to the infeasible OLS estimator $\hat{\beta}_v$, which is the object of interests in our volatility regressions, than the feasible OLS estimator $\hat{\beta}_y$.

To see the performance of the IV approach, we simulate nonstationary volatility regression models with no drift $D_t = 0$ for all t. As for the underlying variance process, we consider two designs

$$\begin{aligned}&\text{Model BM: } V_t = W_t, \\ &\text{Model HK: } dV_t = aV_t(c + V_t^2)^{b-1} dt + (c + V_t^2)^{b/2} dW_t,\end{aligned}$$

where $\mathbb{E}[dW_t dW_t^P] = 0$. Model BM is a Brownian variance process, and Model HK is the generalized Höpfner-Kutoyants model proposed by Kim and Park (2018). In particular, Model HK encompasses several diffusion models introduced in the literature. If, for instance, $a = 0$ or $b = 0$, Model HK reduces to the diffusion considered by Chen *et al.* (2010) or Höpfner and Kutoyants (2003), respectively.[5] In our simulations of Model HK, we let (a, b, c) = (-1/4, 1/10, 1/100), for which the process becomes null recurrent satisfying Assumption 8.[6]

[5] If $a = b = 0$, then Model HK becomes Model BM.

[6] The reader is referred to Kim and Park (2018) for more detailed discussions about the generalized Höpfner-Kutoyants model.

As an illustration, Figure 1 depicts three sample paths of Model HK with the parameter values used in our simulation.[7]

The simulation samples are generated by the Euler discretization at 5 minutes for $T = 500$ days corresponding to 2 years, and the number of simulations is 5,000. 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 the integrated variance is computed by the numerical integration of the simulated spot variance process at 5 minutes. The realized variance is also computed using the simulated returns at the frequency of 5 minutes with $\delta/\Delta = 1/288$.

Table 1: Mean Squared Deviations ($\times 10,000$) of $\hat{\beta}_z$ and $\check{\beta}_z$ for $z = x, y$

	MSD($\hat{\beta}_x$)	MSD($\hat{\beta}_y$)	MSD($\check{\beta}_x$)	MSD($\check{\beta}_y$)
Model BM	0.667	6.092	0.044	0.146
Model HK	2.418	2.886	0.243	0.329

Figure 2 depicts the empirical distributions of $\hat{\beta}_v - \hat{\beta}_z$ (blue) and $\hat{\beta}_v - \check{\beta}_z$ (red) for $z = x, y$. The empirical distributions are coherent with the theory, and the IV approach reduces significantly the bias of the OLS estimators. In particular, the IV approach performs well in the feasible volatility regression with the realized varaince. We also compute the mean squared deviations (MSD) of $\hat{\beta}_z$ and $\check{\beta}_z$ from $\hat{\beta}_v$, and the results are summarized in Tables 1 and 2. Clearly, the IV estimator $\check{\beta}_z$ has a smaller MSD compared to the corresponding OLS estimator $\hat{\beta}_z$ for $z = x, y$. Importantly, in the feasible volatility regression with the realized variance (y_i), the MSDs of the feasible IV estimators $\check{\beta}_y$ are only 2.4% for Model BM and 11.4% for Model HK of those of the feasible OLS estimators $\hat{\beta}_y$ (see Table 2).

Table 2: Ratios of Mean Squared Deviations (in %): MSD($\check{\beta}_z$)/MSD($\hat{\beta}_z$) for $z = x, y$

	MSD($\check{\beta}_x$)/MSD($\hat{\beta}_x$)	MSD($\check{\beta}_y$)/MSD($\hat{\beta}_y$)
Model BM	6.6	2.4
Model HK	10.0	11.4

[7] We leave the modeling of nonstationary volatility processes for future research. Even though the goal of this paper is not to propose new models for nonstationary volatility processes, Model HK seems to be able to generate persistent and realistic dynamics as shown in Figure 1.

4. CONCLUSION

Persistence and (near-)nonstationarity are a well-known stylized fact of financial volatilites. The nonstationarity has been often ignored in the literature on volatility regressions. We study the limiting behavior of the OLS estimators of the volatility regressions of the spot, integrated and realized variances. We show that the OLS estimators converge to random variables, and the feasible regression with the realized variance is biased asymptotically. We then revisit the IV approach of Kim and Meddahi (2020) and show that the feasible IV estimator obtained from the realized variance is asymptotically equivalent to the infeasible OLS estimator obtained from the spot variance. Our simulation studies show that the IV estimator has good performances in finite samples.

APPENDIX: PROOFS

Proof of Proposition 1. The stated results follow immediately from Proposition 3.2 of Kim and Meddahi (2020). □

Proof of Theorem 1. The stated results can be obtained by letting $T \to \infty$ to the primary asymptotics of $\hat{\beta}_z$ for each $z = v, x, y$ obtained in Proposition 1. As for the large T asymptotics, we let (λ_T) be the normalizing sequence satisfying $T = \lambda_T^2 m_s(\lambda_T)$, where m_s is the speed density of natural scale diffusion $V^s = s(V)$ of the underlying diffusion V. We define V^T by $V_t^T = V_{Tt}/s^{-1}(\lambda_T)$ for $t \in [0,1]$ with the normalizing sequence λ_T. We note that Assumption 8 (a) implies that s^{-1} is regularly varying on $\mathbb{R}$ and its limit homogeneous function $\overline{s^{-1}}$ is well defined. Moreover, the limit homogeneous function $\overline{m}_s$ of m_s is also well defined under Assumption 7 (b) and Assumption 8 (a) and (d). It then follows from Proposition 3.2 of Kim and Park (2017) that

$$V^T \to_d V^\circ \tag{7}$$

as $T \to \infty$ in the space $C[0,1]$ of continuous functions with uniform topology, where

$$V_t^\circ = \overline{s^{-1}}(B \circ \overline{A}_t), \quad \text{with} \quad \overline{A}_t = \inf\left\{ s \Big| \int_{\mathbb{R}} L(s,x)\overline{m}_s(dx) > t \right\}$$

with Brownian motion B and its local time L.

Now we are going to show that

$$\frac{1}{T(s^{-1}(\lambda_T))^2}\int_0^T V_t^2 dt \to_d \int_0^1 V_t^{\circ 2} dt, \tag{8}$$

$$\frac{1}{T(s^{-1}(\lambda_T))^2}\int_0^T (V_t - \overline{V}_T)^2 dt \to_d \int_0^1 (V_t^\circ - \overline{V}_1^\circ)^2 dt, \tag{9}$$

$$\frac{1}{(s^{-1}(\lambda_T))^2}[V]_T \to_d [V^\circ]_1, \tag{10}$$

$$\frac{1}{(s^{-1}(\lambda_T))^2}\int_0^T (V_t - \overline{V}_T) dV_t \to_d \int_0^1 (V_t^\circ - \overline{V}_T^\circ) dV_t^\circ. \tag{11}$$

For (8), we have

$$\frac{1}{T(s^{-1}(\lambda_T))^2}\int_0^T V_t^2 dt = \int_0^1 (V_t^T)^2 dt \to_d \int_0^1 V_t^{\circ 2} dt,$$

where the first equality follows from the construction of V^T and the change of variables, and the last convergence holds due to (7) and the continuous mapping theorem. In a similar manner, we can show for (9) that

$$\begin{aligned}
&\frac{1}{T(s^{-1}(\lambda_T))^2}\int_0^T (V_t-\overline{V}_T)^2 dt \\
&= \frac{1}{T(s^{-1}(\lambda_T))^2}\int_0^T V_t^2 dt - \frac{1}{(Ts^{-1}(\lambda_T))^2}\left(\int_0^T V_t dt\right)^2 \\
&= \int_0^1 (V_t^T)^2 dt - \left(\int_0^1 V_t^T dt\right)^2 \\
&\to_d \int_0^1 V_t^{\circ 2} dt - \left(\int_0^1 V_t^\circ dt\right)^2 = \int_0^1 (V_t^\circ - \overline{V}_1^\circ)^2 dt
\end{aligned}$$

as desired. Moreover, the convergences (10) and (11) are developed in Lemma 3.3 of Kim and Park (2018) under the assumptions that V is null recurrent satisfying, in particular, Assumption 8.

Finally, we can obtain the stated results immediately by applying the convergences (8)-(11) to the primary asymptotics in Proposition 1. □

Proof of Proposition 2. Due to Proposition 1, we have $\hat{\beta}_z^{(1)} - 1 \sim_p (\hat{\beta}_z - 1) + (\hat{\beta}_v - 1)$. Thus if we show

$$\phi_z \sim_p \hat{\beta}_z^{(1)} - 1 \quad \text{and} \quad \psi_z \sim_p \hat{\beta}_z - 1, \tag{12}$$

then we have $\check{\beta}_z - 1 = \phi_z - \psi_z \sim_p \hat{\beta}_v - 1$ and

$$\check{\beta}_z - 1 \sim_p \hat{\beta}_v - 1 \sim_d \frac{\Delta}{T}\frac{\int_0^1 (V_t^\circ - \overline{V}_1^\circ) dV_t^\circ}{\int_0^1 (V_t^\circ - \overline{V}_1^\circ)^2 dt}$$

due, in particular, to Theorem 1.

To show (12), it suffice to show that

$$\begin{aligned}
&\sum_{i=2}^{N-1}(z_{i-1}-\overline{z}_N)(z_i-\overline{z}_N)\Delta - \sum_{i=2}^{N-1}(z_{i-1}-\overline{z}_N)^2\Delta \\
&= \sum_{i=2}^{N-1}(z_{i-1}-\overline{z}_N)(z_i-z_{i-1})\Delta = o_p(1)
\end{aligned} \tag{13}$$

since

$$\phi_z - (\hat{\beta}_z^{(1)} - 1) = \Delta \frac{\sum_{i=2}^{N-1} (z_{i-1} - \bar{z}_N)(z_{i+1} - z_{i-1})}{\sum_{i=2}^{N-1} (z_{i-1} - \bar{z}_N)(z_i - \bar{z}_N)\Delta} - \Delta \frac{\sum_{i=2}^{N-1} (z_{i-1} - \bar{z}_N)(z_{i+1} - z_{i-1})}{\sum_{i=2}^{N-1} (z_{i-1} - \bar{z}_N)^2 \Delta},$$

$$\psi_z - (\hat{\beta}_z - 1) = \Delta \frac{\sum_{i=2}^{N-1} (z_{i-1} - \bar{z}_N)(z_i - z_{i-1})}{\sum_{i=2}^{N-1} (z_{i-1} - \bar{z}_N)(z_i - \bar{z}_N)\Delta} - \Delta \frac{\sum_{i=2}^{N-1} (z_{i-1} - \bar{z}_N)(z_i - z_{i-1})}{\sum_{i=2}^{N-1} (z_{i-1} - \bar{z}_N)^2 \Delta}.$$

For (13), we have

$$\begin{aligned}
&\Delta \sum_{i=2}^{N-1} (z_{i-1} - \bar{z}_N)(z_i - z_{i-1}) \\
&\quad = \frac{\Delta}{2}\{z_N^2 - z_1^2 - \bar{z}_N(z_N - z_1)\} - \frac{\Delta}{2}\sum_{i=2}^{N-1} (z_i - z_{i-1})^2 \\
&\quad \sim_p \frac{\Delta}{2}\{V_T^2 - V_0^2 - \overline{V}_T(V_T - V_0)\} - \begin{cases} (\Delta/2)[V]_T, & \text{for } z = v \\ (\Delta/3)[V]_T, & \text{for } z = x \\ (\Delta/3)[V]_T + (2\delta/\Delta)\int_0^T V_t^2 dt, & \text{for } z = y \end{cases} \\
&\quad = \begin{cases} O_p(\Delta T_V(\iota^2)T) + O_p(\Delta T_V(\sigma^2)T), & \text{for } z = v, x \\ O_p(\Delta T_V(\iota^2)T) + O_p(\Delta T_V(\sigma^2)T) + O_p((\delta/\Delta)T_V(\iota^2)T), & \text{for } z = y \end{cases} \\
&\quad = o_p(1),
\end{aligned}$$

where the second line holds due to the summation by parts, and the third line can be deduced from Lemma 3.1 of Kim and Meddahi (2020), and the fourth line follows from the construction of T_V, and the last line follows immediately from Assumptions 3 and 4. This completes the proof. □

REFERENCES

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, F. X. Diebold, and P. Labys (2001). "Modeling and Forecasting Realized Volatility," *Econometrica* 71, 579-625.

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.

Barndorff-Nielsen, O. E. and N. Shephard (2002). "Econometric Analysis of Realized Covariation: High Frequency Based Covariance, Regression, and Correlation in Financial Economics," *Econometrica* 72, 885-925.

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

Bollerslev, T. and J. H. Wright (2000). "Semiparametric Estimation of Long-Memory Volatility Dependencies: The Role of High-Frequency Data," *Journal of Econometrics* 98, 81-106.

Chen, X., L. P. Hansen, and M. Carrasco (2010). "Nonlinearity and Temporal Dependence," *Journal of Econometrics* 155, 155-169.

Davis, R. A. (1982). "Maximum and Minimum of One-Dimensional Diffusions," *Stochastic Processes and their Applications* 13, 1-9.

Engle, R. and J. Rangel (2008). "The Spline-GARCH Model for Low-Frequency Volatility and Its Global Macroeconomic Causes," *Review of Financial Studies* 21, 1187-1222.

Hansen, P. R. and A. Lunde (2014). “Estimating the Persistence and the Autocorrelation Function of a Time Series that is Measured with Error,” *Econometric Theory* 30, 60-93.

Höpfner, R. and Y. Kutoyants (2003). “On a Problem of Statistical Inference in Null Recurrent Diffusions,” *Statistical Inference for Stochastic Processes* 6, 25-42.

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 (2998). “Asymptotic Error Distributions for the Euler Method for Stochastic Differential Equations,” *Annals of Probability* 26, 267-307.

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

Kim, J. and N. Meddahi (2020). “Volatility Regressions with Fat Tails,” *Journal of Econometrics* 218, 690-713.

Kim, J. and J. Y. Park (2017). “Asymptotics for Recurrent Diffusions with Application to High Frequency Regression,” *Journal of Econometrics* 196, 27-54.

Kim, J. and J. Y. Park (2018). “Unit Root, Mean Reversion and Nonstationarity in Financial Time Series,” Working Paper, Indiana University.

Meddahi, N. (2003). “ARMA Representation of Integrated and Realized Variances,” *Econometrics Journal* 6, 335-356.

Zhang, S. and H. Han (2014). “Semiparametric ARCH-X Model for Leverage Effect and Long Memory in Stock Return Volatility,” *Journal of Economic Theory and Econometrics*, 25, 81-100.

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

Figure 1: Three Sample Paths of Model HK

Model HK: $dV_t = aV_t(c+V_t^2)^{b-1}dt + (c+V_t^2)^{b/2}dW_t$ with $(a,b,c) = (-1/4, 1/10, 1/100)$
The simulation samples are generated by the Euler discretization at 5 minutes for $T = 500$ days corresponding to 2 years.

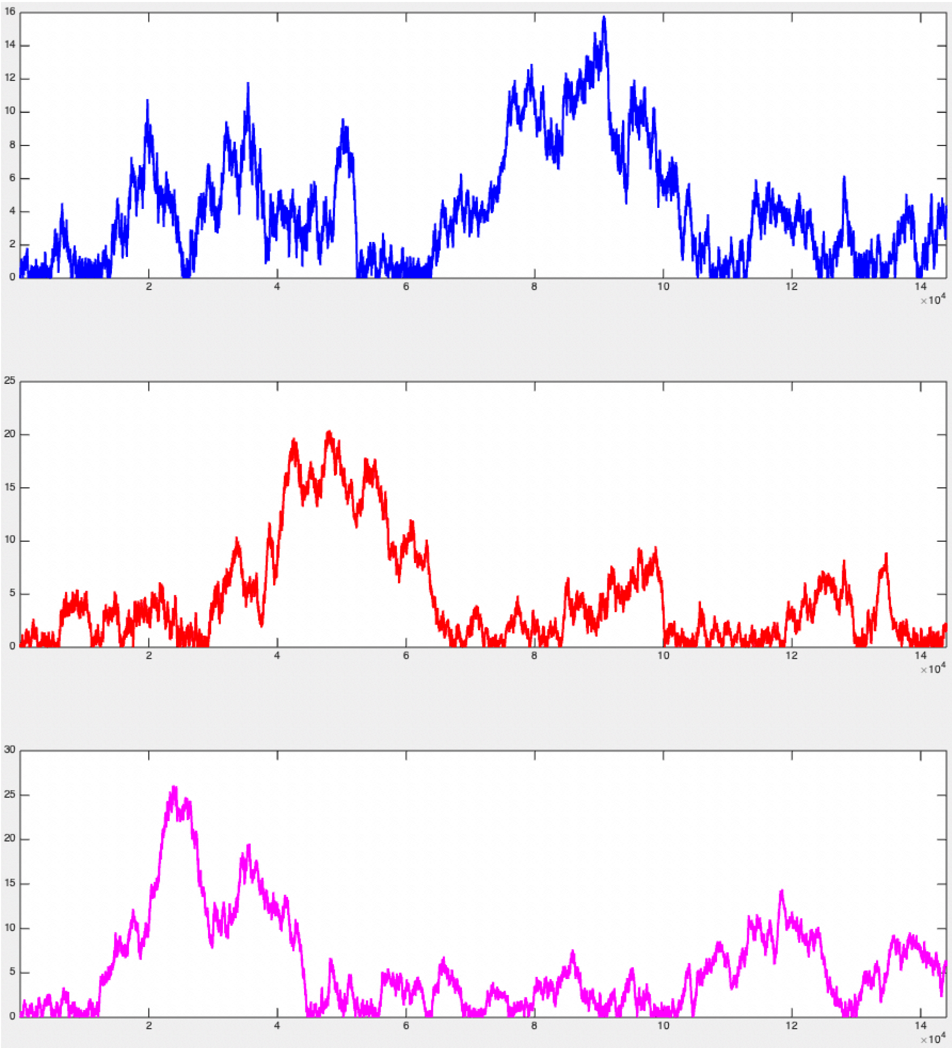
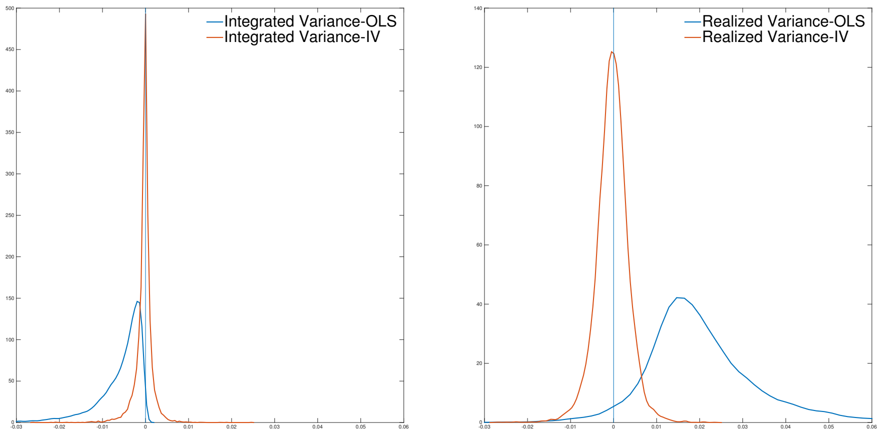
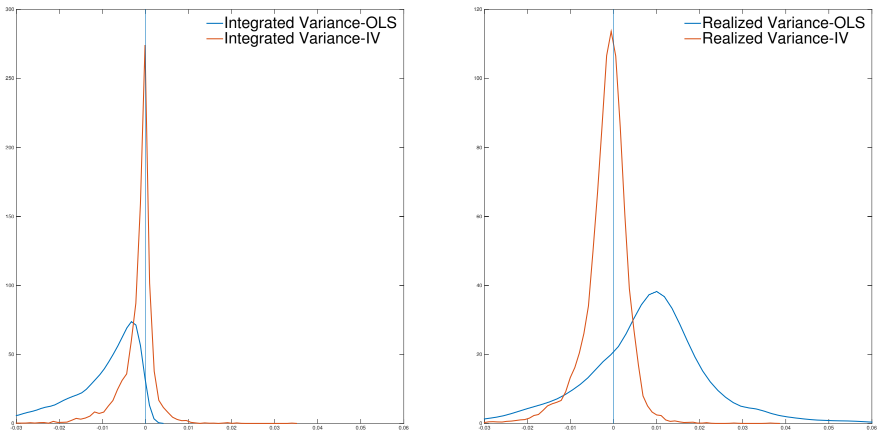


Figure 2: Empirical Distributions of $\hat{\beta}_v - \hat{\beta}_z$ and $\hat{\beta}_v - \check{\beta}_z$ for $z = x, y$

(a) Model BM



(b) Model HK



Left: $\hat{\beta}_v - \hat{\beta}_x$ (blue) and $\hat{\beta}_v - \check{\beta}_x$ (red); Right: $\hat{\beta}_v - \hat{\beta}_y$ (blue) and $\hat{\beta}_v - \check{\beta}_y$ (red)