Asymptotic Properties of the Least Squares Estimator in Local to Unity Processes with Fractional Gaussian Noise[1]

Xiaohu Wang
Fudan University

Weilin Xiao
Zhejiang University

Jun Yu
Singapore Management University

October 19, 2022

[1]Xiaohu Wang, School of Economics, Fudan University, Shanghai, China, and Shanghai Institute of International Finance and Economics, Shanghai, China, Email: wang_xh@fudan.edu.cn. Weilin Xiao, School of Management, Zhejiang University, Hangzhou, 310058, China. Email: wlxiao@zju.edu.cn. Jun Yu, School of Economics and Lee Kong Chian School of Business, Singapore Management University, 90 Stamford Rd, Singapore. Email for Jun Yu: yujun@smu.edu.sg.

Wang acknowledges the support by Shanghai Pujiang Program under No.22PJC022. Yu would like to acknowledge that this research/project is supported by the Ministry of Education, Singapore, under its Academic Research Fund (AcRF) Tier 2 (Award Number MOE-T2EP402A20-0002).

Abstract

This paper derives asymptotic properties of the least squares estimator of the autoregressive parameter in local to unity processes with errors being fractional Gaussian noise with the Hurst parameter $H \in (0,1)$. It is shown that the estimator is consistent for all values of $H \in (0,1)$. Moreover, the rate of convergence is n^{-1} when $H \in [0.5, 1)$. The rate of convergence is n^{-2H} when $H \in (0, 0.5)$. Furthermore, the limiting distribution of the centered least squares estimator depends on H. When $H = 0.5$, the limiting distribution is the same as that obtained in Phillips (1987a) for the local to unity model with errors for which the standard functional central limit theorem is applicable. When $H > 0.5$ or when $H < 0.5$, the limiting distributions are new to the literature. The asymptotic properties of the least squares estimator with fitted intercept are also derived. Simulation studies are performed to check the reliability of the asymptotic approximation for different values of sample size.

JEL classification: C22

Keywords: Least squares, Local to unity, Fractional Brownian motion, Fractional Ornstein-Uhlenbeck process.

1 Introduction

In this paper, we consider the following model:

$$X_t = \rho_n X_{t-1} + \varepsilon_t, \rho_n = \exp(-c/n), t = 1, ..., n, \tag{1}$$

where c is a constant, $\varepsilon_t = \sigma u_t$, u_t is a fractional Gaussian noise (FGN) that has mean zero, variance one, and covariance function as

$$\gamma_u(k) := \mathbb{E}(u_t u_s) = \frac{1}{2}\left[(k+1)^{2H} + (k-1)^{2H} - 2k^{2H}\right] \text{ with } k = |t-s|, \tag{2}$$

and $H \in (0,1)$. The parameter H is known as the Hurst parameter in the literature. When $H = 0.5$, it has $\gamma_u(k) = 0$ for any $k \neq 0$, in which case $\{u_t\}$ form a sequence of independent and identically distributed (i.i.d.) variables with the standard normal distribution $N(0,1)$. However, when $H \neq 0.5$, it has $\gamma_u(k) \neq 0$ for any k, meaning that $\{u_t\}$ have serial dependence. Moreover, it has

$$\gamma_u(k) \sim H(2H-1)k^{2H-2}, \text{ for large } k. \tag{3}$$

That is, $\gamma_u(k)$ decays at a hyperbolic rate as k goes to infinity. As a result, for the case of $H > 0.5$, it has $\gamma_u(k) > 0$ and $\sum_{k=-\infty}^{\infty} \gamma_u(k) = \infty$, giving rise to the terminology of 'long-range-dependent' errors. In contrast, for the case of $H < 0.5$, it has $\gamma_u(k) < 0$ for $k \neq 0$ and $\sum_{k=-\infty}^{\infty} \gamma_u(k) = 0$, giving rise to the terminology of 'anti-persistent' errors.

The FGN u_t has the same distribution as the increment of the fractional Brownian motion (fBm) $B^H(t)$ that is a zero-mean Gaussian process with the covariance function

$$Cov\left(B^H(t), B^H(s)\right) = \frac{1}{2}\left(|t|^{2H} + |s|^{2H} - |t-s|^{2H}\right) \quad \forall t, s \geq 0. \tag{4}$$

That is, $u_t \sim B^H(t) - B^H(t-1)$ where $\sim$ stands for equivalence in distribution.

Model (1) is related to the local to unity model of Phillips (1987a) and Chan and Wei (1987) by replacing the noise where the classical central limit theorem is applicable with FGNs. Model (1) is also related to the fractional unit root model of Sowell (1990) by replacing the autoregressive (AR) coefficient of unity with the AR coefficient of local to unity. Although we replace the $I(d)$ noise of Sowell (1990) with the FGN, the results in this paper also apply to $I(d)$ errors as it will become clear later. Model (1) is also related to the model of Park (2003) where $\rho_n = 1 - m/n$ if we assume m is fixed in his model.

We consider two regressions to estimate the AR root ρ_n in Model (1). The first is an AR regression without intercept fitted, which leads to the least squares (LS) estimator of ρ_n as

$$\widehat{\rho}_n = \sum_{t=1}^{n} X_{t-1}X_t / \sum_{t=1}^{n} X_{t-1}^2 = \rho_n + \sum_{t=1}^{n} X_{t-1}\varepsilon_t / \sum_{t=1}^{n} X_{t-1}^2. \tag{5}$$

The second is an AR regression with intercept fitted, giving the following LS estimator of ρ_n

$$\begin{aligned} \widetilde{\rho}_n &= \sum_{t=1}^{n} \left(X_{t-1} - \overline{X}_{-1}\right) X_{t-1} / \sum_{t=1}^{n} \left(X_{t-1} - \overline{X}_{-1}\right)^2 \\ &= \rho_n + \sum_{t=1}^{n} \left(X_{t-1} - \overline{X}_{-1}\right) \varepsilon_t / \sum_{t=1}^{n} \left(X_{t-1} - \overline{X}_{-1}\right)^2, \end{aligned} \tag{6}$$

where $\overline{X}_{-1} = \frac{1}{n}\sum_{t=1}^{n} X_{t-1}$.

The goal of this paper is to derive the asymptotic properties of the two estimators $\widehat{\rho}_n$ and $\widetilde{\rho}_n$ under the assumption of $n \to \infty$. As it is well expected for local to unity model, the initial value of X_t significantly affects the finite sample distribution of $\widehat{\rho}_n$ and $\widetilde{\rho}_n$. To capture the impact of the initial value on asymptotics, we set the initial value of X_t to be $X_0 = O_p\left(n^H\right)$ and

$$n^{-H}\frac{X_0}{\sigma} \xrightarrow{p} \pi_0,$$

where π_0 is a constant (such as zero) or $O_p(1)$.

The rest of the paper is organized as follows. Section 2 reviews the results in the literature. The asymptotic properties of the normalized $\widehat{\rho}_n - \rho_n$ are developed in Section 3. Section 4 extends the results to the case when the intercept is fitted. Section 5 examines the finite sample properties of the normalized $\widehat{\rho}_n - \rho_n$ and $\widetilde{\rho}_n - \rho_n$. Section 6 concludes. The Appendix collects proofs of the main results.

Throughout the paper, we use $\xrightarrow{p}$, $\xrightarrow{d}$, $\Rightarrow$, and $\sim$ to denote convergence in probability, convergence in distribution, convergence in functional space, and equivalence in distribution, respectively. The notation $[nr]$ represents the integer part of nr.

2 A Literature Review

Phillips (1987a) considers the following local to unity model

$$X_t = \rho_n X_{t-1} + v_t,\ \rho_n = \exp(-c/n),\ X_0 = O_p(1), \tag{7}$$

where $\{v_t\}$ is a strong mixing sequence with mixing coefficients α_m that satisfies $\sum_{m=1}^{\infty} \alpha_m^{1-2/\beta} < \infty$ and $\sup_t |v_t|^{\beta+\delta} < \infty$ for some $\beta > 2$ and $\delta > 0$. There are two important features in Model (7). First, since $\rho_n = 1 - c/n + O(n^{-2})$, the autoregressive coefficient depends on n and converges to unity as $n \to \infty$. Second, the functional central limit theorem is applicable to $\{v_t\}$. An interesting special case of Model (7) is when $\{v_t\}$ are i.i.d. with $E|v_t|^{\beta} < \infty$ for some $\beta > 2$. In this case, as $n \to \infty$, it has

$$n(\widehat{\rho}_n - \rho_n) \xrightarrow{d} \frac{\int_0^1 J_c(r)dW(r)}{\int_0^1 J_c(r)^2 dr} = \frac{\left\{J_c(1)^2 + 2c\int_0^1 J_c(r)^2 dr - 1\right\}/2}{\int_0^1 J_c(r)^2 dr}. \tag{8}$$

where $J_c(r)$ denotes an Ornstein-Uhlenbeck (OU) process defined by the stochastic differential equation

$$dJ_c(r) = -cJ_c(r)dr + dW(r),\ J_c(0) = 0, \tag{9}$$

with $W(r)$ being a standard Brownian motion.

Sowell (1990) considers the following unit root model with $\rho = 1$:

$$X_t = \rho X_{t-1} + \sigma v_t,\ v_t = (1-L)^{-d}\epsilon_t,\ \epsilon_t \overset{i.i.d.}{\sim} (0,1),\ X_0 = O_p(1), \tag{10}$$

where L is the lag operator with $(1-L)^{-d}$ defined as

$$(1-L)^{-d} = \sum_{j=0}^{\infty} \frac{\Gamma(j+d)}{\Gamma(d)\Gamma(j+1)} L^j \quad \text{for} \quad d \in (-0.5, 0.5).$$

In this model, the error term v_t is assumed to follow a fractionally integrated process of order d, which is referred to as an $I(d)$ process in the literature. With $\widehat{\rho}$ being the LS estimator of ρ, Sowell (1990) and Marinucci and Robinson (1999) show that, as $n \to \infty$,

$$n(\widehat{\rho} - 1) \xrightarrow{d} \frac{\int_0^1 W(r)dW(r)}{\int_0^1 W(r)^2 dr}, \quad \text{if } d = 0, \tag{11}$$

$$n\left(\widehat{\rho}-1\right) \xrightarrow{d} \frac{\frac{1}{2}B^H(1)^2}{\int_0^1 B^H(r)^2 dr}, \quad \text{if } d>0, \tag{12}$$

$$n^{2H}\left(\widehat{\rho}-1\right) \xrightarrow{d} -\frac{H\frac{\Gamma(0.5+H)}{\Gamma(1.5-H)}}{\int_0^1 B^H(r)^2 dr}, \quad \text{if } d<0, \tag{13}$$

where $H = d + 0.5$.[1]

Setting $c = 0$ in (8) or setting $d = 0$ in (11) can lead to the well-known result for the unit root model obtained in Phillips (1987b) as

$$n\left(\widehat{\rho}-1\right) \xrightarrow{d} \frac{\int_0^1 W(r)dW(r)}{\int_0^1 W(r)^2 dr} = \frac{\frac{1}{2}\left(W(1)^2-1\right)}{\int_0^1 W(r)^2 dr}.$$

3 Asymptotic Properties

To develop the asymptotic properties of the centered LS estimator $\widehat{\rho}_n - \rho_n$ defined in (5), we first introduce the limit behavior of the partial sum process $\sum_{t=1}^{[nr]} u_t$ for any $r \in [0,1]$. As $u_t = B^H(t) - B^H(t-1)$, we have

$$\begin{aligned} n^{-H}\sum_{t=1}^{[nr]} u_t &\sim n^{-H}\sum_{t=1}^{[nr]}\left\{B^H(t) - B^H(t-1)\right\} \\ &= n^{-H}B^H([nr]) \\ &\sim B^H\left(\frac{[nr]}{n}\right) \\ &\Rightarrow B^H(r), \quad \text{as } n \to \infty, \end{aligned} \tag{14}$$

where equivalence in distribution comes from the self-similarity property of the fBm $B^H(t)$. Note that the sample path of $n^{-H}\sum_{t=1}^{[nr]} u_t$ is a function of $r \in [0,1]$ that is right-continuous with left limits. Hence, the convergence result of the partial sum sequence is built up in the space of $D[0,1]$, which is the space of all real valued functions on $[0,1]$ that are right-continuous with finite left limits, equipped with the Skorokhod topology. The convergence results obtained in the rest of the paper are all considered in the space of $D[0,1]$ with the same topology.

[1] Equations (11)-(13) are different from those reported in Theorem 3 in Sowell (1990). This is because, as remarked in Section 3 of Marinucci and Robinson (1999), the partial sum of an $I(d)$ process, adjusted an appropriate normalizing term, should converge to the Type I fBm denoted by $B^H(t)$ in the present paper, not to the Type II fBm adopted in Sowell (1990).

The convergence result in (14) is the source of the asymptotic theory developed in the present paper. Sowell (1990) gives a similar weak convergence result for the partial sum process $\sum_{t=1}^{[nr]} u_t$ when $u_t \sim I(d)$; see also Marinucci and Robinson (1999). Therefore, all the results in our paper applies to the case where $u_t \sim I(d)$. It is important to note that Sowell (1990) uses the result of Davydov (1970) to establish the weak convergence while we do not need to resort to Davydov (1970) as our errors are normally distributed.

The result in (14) compares with Donsker's functional central limit theorem, which states that,

$$n^{-0.5} \sum_{t=1}^{[nr]} \epsilon_t \Rightarrow W(r) = B^{0.5}(r), \quad \text{as } n \to \infty, \tag{15}$$

where ϵ_t is a sequence of i.i.d. random variables with mean zero and variance one.

Define a fractional OU (fOU) process through the following stochastic differential equation

$$dJ_c^H(t) = -cJ_c^H(t)dt + dB^H(t), \text{ with } J_c^H(0) = 0. \tag{16}$$

Cheridito et al. (2003) proved that, for $t > 0$, the differential equation (16) has a unique solution, taking the form of

$$J_c^H(t) = \int_0^t e^{-c(t-s)} dB^H(s),$$

where the integral is a path-wise Riemann-Stieltjes integral. It is worthwhile to mention that, when $H = 0.5$, $J_c^H(t)$ becomes the traditional OU process considered in Phillips (1987a). If in addition, $c = 0$, the process $J_c^H(t)$ becomes a standard Brownian motion.

Lemma 1 *Let $\{X_t\}$ be the time series generated by Model (1). Then, as $n \to \infty$,*

1. $n^{-H} X_{[nr]} \Rightarrow \sigma J_c^H(r) + e^{-cr}\sigma\pi_0$;
2. $n^{-1-H} \sum_{t=1}^n X_t \Rightarrow \sigma \int_0^1 \left[J_c^H(r) + e^{-cr}\pi_0\right] dr$;
3. $n^{-1-2H} \sum_{t=1}^n X_t^2 \Rightarrow \sigma^2 \int_0^1 \left[J_c^H(r) + e^{-cr}\pi_0\right]^2 dr$;
4. $n^{-2H} \sum_{t=1}^n X_{t-1}\varepsilon_t$
$$\Rightarrow \begin{cases} \sigma^2 \left([J_c(1) + e^{-c}\pi_0]^2 - \pi_0^2 + 2c\int_0^1 [J_c(r) + e^{-cr}\pi_0]^2 dr - 1 \right)/2, & \text{if } H = 0.5 \\ \sigma^2 \left(\left[J_c^H(1) + e^{-c}\pi_0\right]^2 - \pi_0^2 + 2c\int_0^1 \left[J_c^H(r) + e^{-cr}\pi_0\right]^2 dr \right)/2, & \text{if } H > 0.5 \end{cases};$$

5. $n^{-1} \sum_{t=1}^{n} X_{t-1}\varepsilon_t \xrightarrow{p} -\sigma^2/2$, *if* $H < 0.5$.

Remark 1 *This lemma is related to Lemma 1 in Phillips (1987a) with several differences. First, the initial condition* π_0*, which is the limit of* $n^{-H}X_0/\sigma$*, plays explicit roles in all the limits except the last one. Second, compared with Lemma 1.a-1.c of Phillips (1987a),* $J_c(r)$ *is replaced with* $J_c^H(r)$ *in our Lemma, and the orders of* $X_{[nr]}$*,* $\sum_{t=1}^{n} X_t$*, and* $\sum_{t=1}^{n} X_t^2$ *becomes* n^H*,* n^{-1+H}*, and* n^{1+2H}*, respectively. Third, both the order and the limit of* $\sum_{t=1}^{n} X_{t-1}\varepsilon_t$ *depend on* H*. When* $H \geq 0.5$*, the order of* $\sum_{t=1}^{n} X_{t-1}\varepsilon_t$ *is* n^{2H}*. Whereas, when* $H < 0.5$*, the order becomes* n*. In addition, the limit of* $\sum_{t=1}^{n} X_{t-1}\varepsilon_t$ *has one more term (i.e.,* $-\sigma^2/2$*) when* $H = 0.5$ *than when* $H > 0.5$*, and three more terms than* $H < 0.5$*. These differences root in the distinct properties of the FGN,* u_t*, when the Hurst parameter* H *takes different values. For example, when* $H = 0.5$*, the limit of* $n^{-2H}\sum_{t=1}^{n} \varepsilon_t^2$ *is* σ^2*. Whereas, when* $H > 0.5$*, the limit of* $n^{-2H}\sum_{t=1}^{n} \varepsilon_t^2$ *is zero.*

Remark 2 *When* $H = 0.5$*, it has* $J_c^H(r) = J_c(r)$*. If we further let* $\pi_0 = 0$*, the results in Parts 1-3 of Lemma 1 above becomes exactly the same as those in Lemma 1.a-1.c in Phillips (1987a). Moreover, the result in Part 4 of Lemma 1 above can be written as*

$$n^{-1}\sum_{t=1}^{n} X_{t-1}\varepsilon_t \Rightarrow \sigma^2 \left(J_c(1)^2 + 2c\int_0^1 J_c(r)^2 dr - 1 \right)/2 = \sigma^2 \int_0^1 J_c(r)dW(r),$$

which is the same as that in Lemma 1.d of Phillips (1987a).

Remark 3 *The convergence result in Part 1 of Lemma 1 is the key to the development of the results in the rest of the Lemma. With slight adjustments, the result in Part 1 can be extended to the case where* u_t *becomes an* $I(d)$ *process. When* $u_t \sim I(d)$*,* $n^{-H}\left(\frac{\Gamma(1-2d)}{(1+2d)\Gamma(1+d)\Gamma(1-d)}\right)^{-1/2}\sum_{t=1}^{[nr]} u_t \Rightarrow B^H(r)$ *as* $n \to \infty$*. This is a special case of a more general result obtained in Taqqu (1975). Consequently, with the use of the continuous mapping theorem, it can be proved easily that*

$$n^{-H}\left(\frac{\Gamma(1-2d)}{(1+2d)\Gamma(1+d)\Gamma(1-d)}\right)^{-1/2} X_{[nr]}$$
$$\Rightarrow \ \sigma J_c^H(r) + \left(\frac{\Gamma(1-2d)}{(1+2d)\Gamma(1+d)\Gamma(1-d)}\right)^{-1/2} e^{-cr}\sigma\pi_0.$$

Theorem 2 *Let* $\{X_t\}$ *be the time series generated by (1) and (2). Then, as* $n \to \infty$*, if* $H = 0.5$*,*

$$n\left(\widehat{\rho}_n - \rho_n\right) \Rightarrow \frac{\left(\left[J_c(1) + e^{-c}\pi_0\right]^2 - \pi_0^2 + 2c\int_0^1 \left[J_c(r) + e^{-cr}\pi_0\right]^2 dr - 1\right)/2}{\int_0^1 \left[J_c(r) + e^{-cr}\pi_0\right]^2 dr}; \tag{17}$$

if $H > 0.5$

$$n\left(\widehat{\rho}_n - \rho_n\right) \Rightarrow \frac{\left(\left[J_c^H\left(1\right) + e^{-c}\pi_0\right]^2 - \pi_0^2 + 2c\int_0^1 \left[J_c^H\left(r\right) + e^{-cr}\pi_0\right]^2 dr\right)/2}{\int_0^1 \left[J_c^H(r) + e^{-cr}\pi_0\right]^2 dr}; \tag{18}$$

if $H < 0.5$,

$$n^{2H}\left(\widehat{\rho}_n - \rho_n\right) \Rightarrow \frac{-1/2}{\int_0^1 \left[J_c^H(r) + e^{-cr}\pi_0\right]^2 dr}. \tag{19}$$

Remark 4 *When we compare Theorem 2 in our paper to Theorem 1 in Phillips (1987a), we have a few observations. First, the initial condition* π_0 *plays significant roles in all the limits in Theorem 2. Second, when* $H = 0.5$ *and* $\pi_0 = 0$, $\widehat{\rho}_n - \rho_n$ *has the same convergence rate and the same limiting distribution as those in Phillips (1987a). Third, there is a discontinuity in our limit theory when* H *passes* 0.5. *When* $H > 0.5$, *the convergence rate of* $\widehat{\rho}_n - \rho_n$ *is* n, *which is the same as that when* $H = 0.5$. *However, the numerator of the limit has one term less comparing to the case of* $H = 0.5$. *Furthermore, When* $H < 0.5$, *the rate of convergence of* $\widehat{\rho}_n - \rho_n$ *becomes* n^{2H}, *which is slower than that when* $H \geq 0.5$. *The numerator in the limit has three terms less than that when* $H = 0.5$.

Remark 5 *If* $c = 0$, *then* $\rho_n = \exp(-c/n) = 1$. *In this case, Model (1) gives a unit root process with FGNs. With the further assumption of* $X_0 = 0$ *that leads to* $\pi_0 = 0$, *the results in Theorem 2 become*

$$n\left(\widehat{\rho}_n - \rho_n\right) \Rightarrow \frac{\frac{1}{2}B^H(1)^2}{\int_0^1 B^H(r)^2 dr}, \quad \textit{when} \quad H > 0.5, \tag{20}$$

and

$$n^{2H}\left(\widehat{\rho}_n - \rho_n\right) \Rightarrow \frac{-1/2}{\int_0^1 B^H(r)^2 dr}, \quad \textit{when} \quad H < 0.5. \tag{21}$$

The result in (20) is the same as that developed in Sowell (1990) and Marinucci and Robinson (1999) for the unit root process with $I(d)$ *errors when* $d = H - 1/2 > 0$. *However, when* $H < 0.5$ *our limiting result in (21) is slightly different with that obtained in Sowell (1990) and Marinucci and Robinson (1999) when* $d = H - 1/2 < 0$; *see (13) in the present paper. The difference arises because the* $I(d)$ *process used in Sowell (1990) has different variance and long-run variance from those of the FGN. The variance and the long-run variance of an* $I(d)$ *process is* $\frac{\Gamma(1-2d)}{\Gamma(1-d)^2}$ *and* $O(n^{2H})\frac{\Gamma(1-2d)}{(1+2d)\Gamma(1+d)\Gamma(1-d)}$, *respectively. The ratio of* $\frac{\Gamma(1-2d)}{\Gamma(1-d)^2}$ *and* $\frac{\Gamma(1-2d)}{(1+2d)\Gamma(1+d)\Gamma(1-d)}$, *divided by 2, gives*

$$\frac{(1+2d)\Gamma(1+d)}{2\Gamma(1-d)} = \frac{H\Gamma(0.5+H)}{\Gamma(1.5-H)},$$

which is the numerator of the limit in (12) that has been derived by Marinucci and Robinson (1999).

4 Asymptotic Properties with Fitted Intercept

In this section, we assume that, while the data are generated from Model (1), it is not known apriori that the intercept is zero. Hence, an AR regression with an intercept is estimated as

$$X_t = \alpha + \rho_n X_{t-1} + \varepsilon_t, \tag{22}$$

which leads to the LS estimator of ρ_n as in (6).

Theorem 3 presents the large sample theory of $\widetilde{\rho}_n$ for various values of H.

Theorem 3 *Let $\{X_t\}$ be the time series generated by Model (1) and $\widetilde{\rho}_n$ be the estimator of the AR root from Model (22) with fitted intercept. Then, as $n \to \infty$,*
if $H = 0.5$,

$$\begin{aligned} & n\left(\widetilde{\rho}_n - \rho_n\right) \\ \Rightarrow \quad & \frac{\left(\left[J_c\left(1\right) + e^{-c}\pi_0\right]^2 - \pi_0^2 + 2c\int_0^1 \left[J_c\left(r\right) + e^{-cr}\pi_0\right]^2 dr - 1\right)/2 - B^H(1)\int_0^1 \left[J_c(r) + e^{-cr}\pi_0\right] dr}{\int_0^1 \left[J_c(r) + e^{-cr}\pi_0\right]^2 dr - \left(\int_0^1 \left[J_c(r) + e^{-cr}\pi_0\right] dr\right)^2}; \end{aligned} \tag{23}$$

if $H > 0.5$,

$$\begin{aligned} & n\left(\widetilde{\rho}_n - \rho_n\right) \\ \Rightarrow \quad & \frac{\left(\left[J_c^H\left(1\right) + e^{-c}\pi_0\right]^2 - \pi_0^2 + 2c\int_0^1 \left[J_c^H\left(r\right) + e^{-cr}\pi_0\right]^2 dr\right)/2 - B^H(1)\int_0^1 \left[J_c^H(r) + e^{-cr}\pi_0\right] dr}{\int_0^1 \left[J_c^H(r) + e^{-cr}\pi_0\right]^2 dr - \left(\int_0^1 \left[J_c^H(r) + e^{-cr}\pi_0\right] dr\right)^2}; \end{aligned} \tag{24}$$

if $H < 0.5$,

$$n^{2H}\left(\widetilde{\rho}_n - \rho_n\right) \Rightarrow \frac{-1/2}{\int_0^1 \left[J_c^H(r) + e^{-cr}\pi_0\right]^2 dr - \left(\int_0^1 \left[J_c^H(r) + e^{-cr}\pi_0\right] dr\right)^2}. \tag{25}$$

Remark 6 *When $H = 0.5$ and $\pi_0 = 0$, it has $B^H(r) = W(r)$ and*

$$n(\widetilde{\rho}_n - \rho_n) \Rightarrow \frac{\left(J_c(1)^2 + 2c\int_0^1 J_c(r)^2 dr - 1\right)/2 - B^H(1)\int_0^1 J_c(r)dr}{\int_0^1 [J_c(r)]^2 dr - \left(\int_0^1 J_c(r)dr\right)^2}$$

Appling a result in Phillips (1987a) that

$$\left(J_c(1)^2 + 2c\int_0^1 J_c(r)^2 dr - 1\right)/2 = \int_0^1 J_c(r)dW(r),$$

we obtain the limiting distribution of $n(\widetilde{\rho}_n - \rho_n)$ as

$$n(\widetilde{\rho}_n - \rho_n) \Rightarrow \frac{\int_0^1 J_c(r)dW(r) - B^H(1)\int_0^1 J_c(r)dr}{\int_0^1 [J_c(r)]^2 dr - \left(\int_0^1 J_c(r)dr\right)^2} = \frac{\int_0^1 \overline{J}_c(r)dW(r)}{\int_0^1 \overline{J}_c(r)^2 dr},$$

where $\overline{J}_c(r) = J_c(r) - \int_0^1 J_c(s)ds$ is the de-meaned OU process. This limiting distribution is the same as that given in Remark 3 of Mikusheva (2015) for the local-to-unity model with weakly dependent errors.

Remark 7 *Theorem 3 shows that the large sample theory of the centered LS estimator $\widetilde{\rho}_n - \rho_n$ not only depends on the values of H, but also on the initial condtion π_0.*

In Corrollary 4 below, it is shown that, when $c = 0$ that makes the Model (1) a unit root process, the limiting distributions of $\widetilde{\rho}_n - \rho_n$ becomes independent of the initial condition.

Corollary 4 *Let $\{X_t\}$ be the time series generated from Model (1) with $c = 0$. In this case, it has $J_c^H(r) = B^H(r)$. Then, as $n \to \infty$,*
if $H = 0.5$,

$$n(\widetilde{\rho}_n - \rho_n) \Rightarrow \frac{(W(1)^2 - 1)/2 - W(1)\int_0^1 W(r)dr}{\int_0^1 W(r)^2 dr - \left(\int_0^1 W(r)dr\right)^2}; \tag{26}$$

if $H > 0.5$

$$n(\widetilde{\rho}_n - \rho_n) \Rightarrow \frac{B^H(1)^2/2 - B^H(1)\int_0^1 B^H(r)dr}{\int_0^1 B^H(r)^2 dr - \left(\int_0^1 B^H(r)dr\right)^2}; \tag{27}$$

if $H < 0.5$,

$$n^{2H}(\widetilde{\rho}_n - \rho_n) \Rightarrow \frac{-1/2}{\int_0^1 B^H(r)^2 dr - \left(\int_0^1 B^H(r)dr\right)^2}. \tag{28}$$

Remark 8 *The large sample result in (26) of Corollary 4 is well-known in the literature; see, for example, Equation (17.4.28) in Hamilton (1994). Loosely speaking, the large sample results in (27) and (28) of Corollary 4 extend those of Sowell (1990) from the estimated model without fitted intercept to the estimated model with fitted intercept.*

5 Monte Carlo Studies

To check how well the limit distribution perform in finite sample, we carry out several Monte Carlo studies. In all studies, we simulate data from Model (1). Four different sample sizes are considered, namely, $n = 32, 512, 2048, 8192$. Three values are considered for H, namely $H = 0.5, 0.9, 0.1$.[2] Two values are considered for c, namely, $c = 10, 5$.

5.1 Without fitted intercept

For each time series simulated, we estimate ρ_n by $\widehat{\rho}_n$ and calculate $n\left(\widehat{\rho}_n - \rho_n\right)$ when $H \geq 0.5$ and $n^{2H}\left(\widehat{\rho}_n - \rho_n\right)$ when $H < 0.5$. The $200,000$ replications are used to obtain density of $n\left(\widehat{\rho}_n - \rho_n\right)$ or $n^{2H}\left(\widehat{\rho}_n - \rho_n\right)$.

Figures 1-2 display the density of $n\left(\widehat{\rho}_n - \rho_n\right)$ when $H = 0.5$ and $c = 10, 5$. In each of the two values of c, the densities are almost identical for all n, suggesting the limit distribution provides accurate approximations to the finite sample distribution when the sample size is as small as 32. In all cases, the density is left-skewed.

Figures 3-4 display the density of $n\left(\widehat{\rho}_n - \rho_n\right)$ when $H = 0.9$ and $c = 10, 5$. For both values of c, the density when $n = 32$ is very different from that when $n = 8192$. The density for $n = 2048$ is very close to that for $n = 8192$. For small values of n, the density is left-skewed. Interestingly, the density becomes right-skewed when n is larger. Although the same rate applies to $H = 0.5$ and to $H > 0.5$, the convergence of the density is much slower when $H > 0.5$ than that when $H = 0.5$. This study indicates that the asymptotic distribution approximates the finite sample distribution less accurately when $H > 0.5$ than when $H = 0.5$ if n is small.

Figures 5-6 display the density of $n^{2H}\left(\widehat{\rho}_n - \rho_n\right)$ when $H = 0.1$ and $c = 10, 5$. For

[2] The choice of $H = 0.1$ is empirically relevant for modeling logarithmic realized volatility, as found in Gatheral et al. (2018) and Wang et al. (2021).

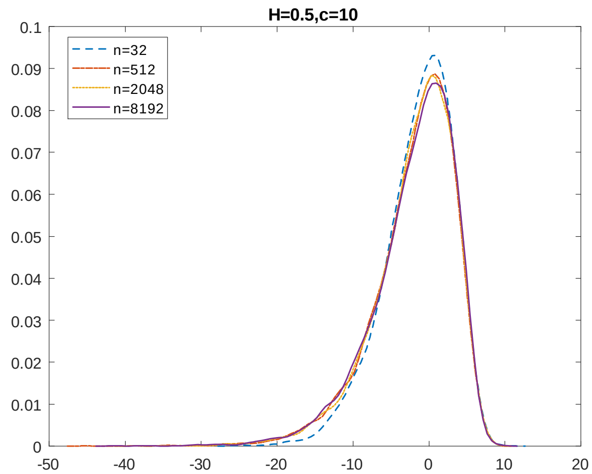

Figure 1: The density of $n\left(\widehat{\rho}_n - \rho_n\right)$.

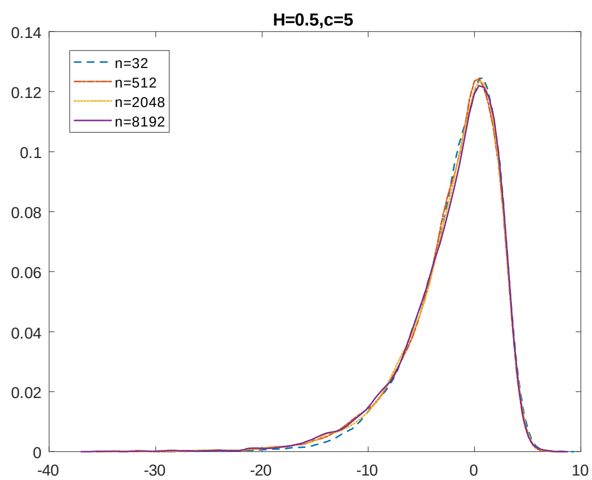

Figure 2: The density of $n\left(\widehat{\rho}_n - \rho_n\right)$.

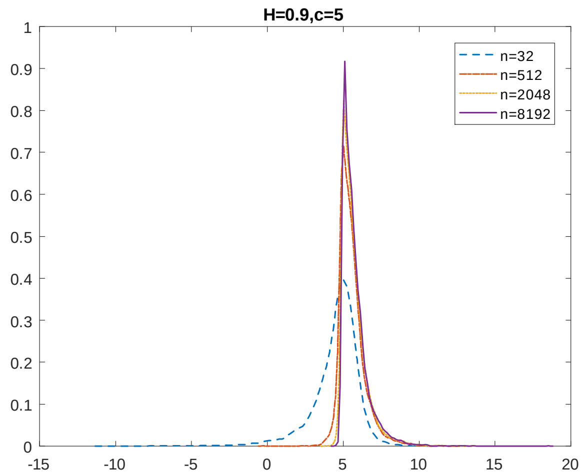

Figure 3: The density of $n\,(\widehat{\rho}_n - \rho_n)$.

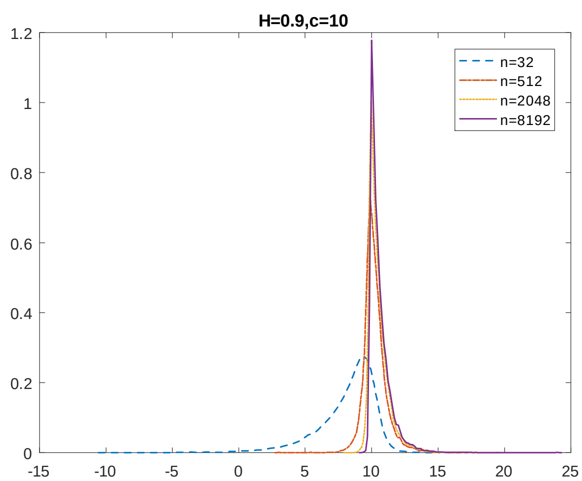

Figure 4: The density of $n\,(\widehat{\rho}_n - \rho_n)$.

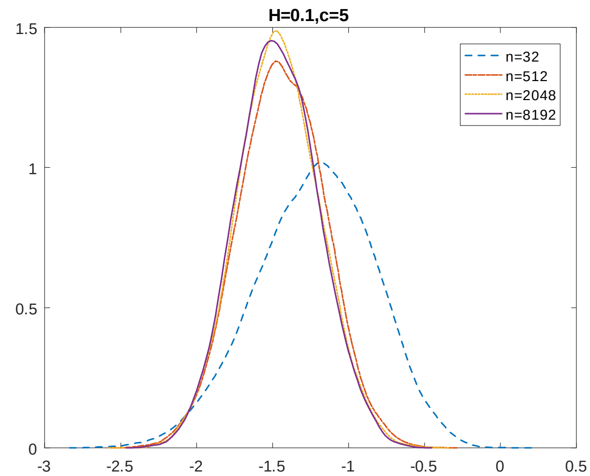

Figure 5: The density of $n^{0.2}(\widehat{\rho}_n - \rho_n)$.

both values of c, the density when $n = 32$ is hugely different from those for other values of n, suggesting one would make a terrible mistake by using the limit distribution to approximate the finite sample distribution when $n = 32$. However, the densities for $n = 2048, 8192$ are close to each other. For all values of n, the density of $n^{2H}(\widehat{\rho}_n - \rho_n)$ is symmetric.

5.2 With fitted intercept

For each time series simulated, we now estimate ρ_n by $\widetilde{\rho}_n$ and calculate $n(\widetilde{\rho}_n - \rho_n)$ when $H \geq 0.5$ and $n^{2H}(\widetilde{\rho}_n - \rho_n)$ when $H < 0.5$. The $200,000$ replications are used to obtain density of $n(\widetilde{\rho}_n - \rho_n)$ or $n^{2H}(\widetilde{\rho}_n - \rho_n)$.

Figures 7-8 display the densities of $n(\widetilde{\rho}_n - \rho_n)$ when $H = 0.5$ and $c = 10, 5$. For every value of c, the densities are almost identical for all n, suggesting that the limiting distribution provides accurate approximations to the finite sample distribution when the sample size is as small as 32. In all cases, the density is left-skewed. Compared with Figures 1-2, the densities in Figures 7-8 are more spread. This is expected as the intercept is also fitted.

Figures 9-10 display the densities of $n(\widetilde{\rho}_n - \rho_n)$ when $H = 0.9$ and $c = 10, 5$. For both values of c, the density when $n = 32$ is very different from that when $n = 8192$. The density for $n = 2048$ is very close to that for $n = 8192$. The densities are right-

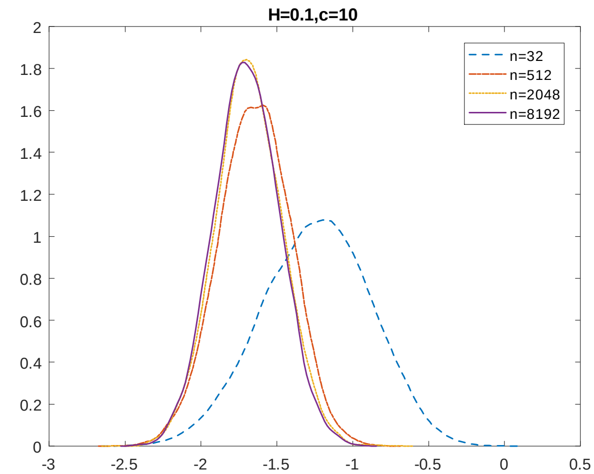

Figure 6: The density of $n^{0.2}\,(\widehat{\rho}_n - \rho_n)$.

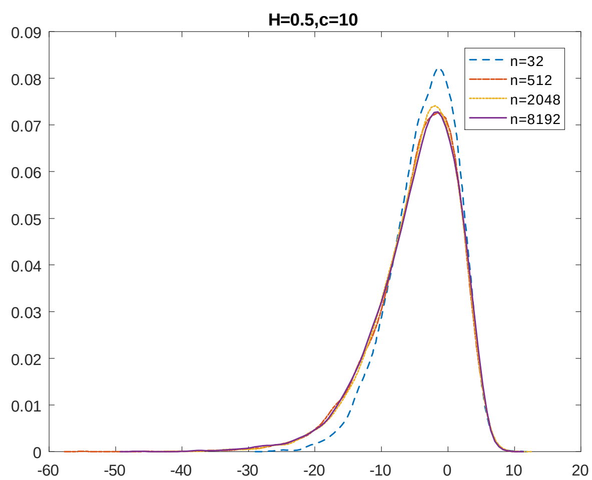

Figure 7: The density of $n\,(\widetilde{\rho}_n - \rho_n)$.

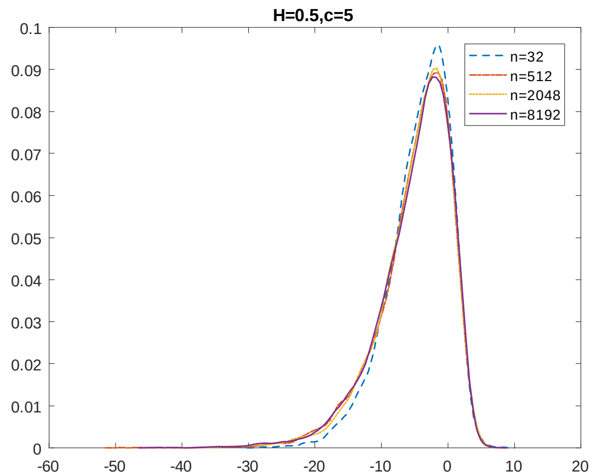

Figure 8: The density of $n\left(\widetilde{\rho}_n - \rho_n\right)$.

skewed for all n. Although the same rate applies to $H = 0.5$ and $H > 0.5$, the convergence in density is much slower when $H > 0.5$ than when $H = 0.5$. This study indicates that the asymptotic distribution approximates the finite sample distribution less accurately when $H > 0.5$ than when $H = 0.5$ if n is small. Compared with Figures 3-4, the densities in Figures 9-10 are more spread, as expected.

Figures 11-12 display the density of $n^{2H}\left(\widetilde{\rho}_n - \rho_n\right)$ when $H = 0.1$ and $c = 10, 5$. For both values of c, the density when $n = 32$ is hugely different from those for other values of n, suggesting that one would make a terrible mistake by using the limiting distribution to approximate the finite sample distribution when $n = 32$. However, the densities for $n = 2048, 8192$ are nearly identical. The density is symmetric for all n. Compared with Figures 5-6, the densities in Figures 11-12 are more spread, as expected.

6 Conclusions

In this paper, we study the properties of the LS estimator (with and without the intercept fitted) of the autoregressive parameter in local to unity processes when errors are assumed to be FGNs with the Hurst parameter H. It is shown that the estimator is consistent when $H \in (0, 1)$. Moreover, the rate of convergence is n when $H \in [0.5, 1)$, whereas the rate of convergence is n^{2H} when $H \in (0, 0.5)$. This result suggests that the

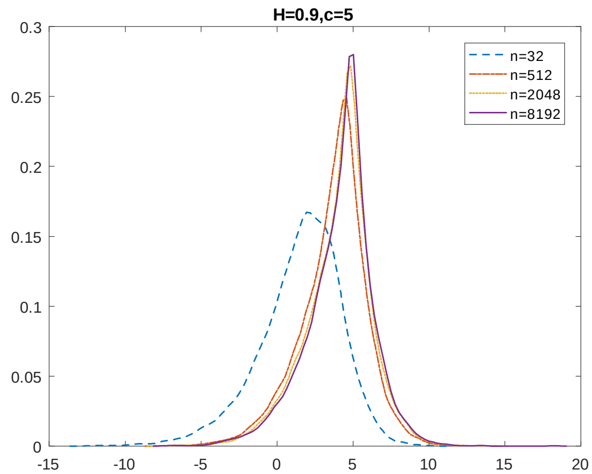

Figure 9: The density of $n\,(\widetilde{\rho}_n - \rho_n)$.

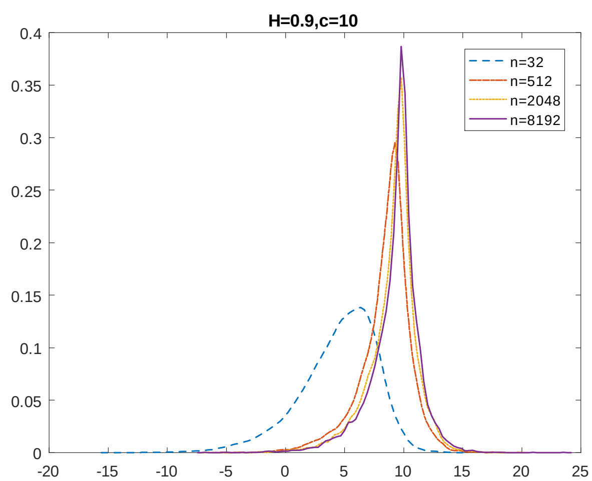

Figure 10: The density of $n\,(\widetilde{\rho}_n - \rho_n)$.

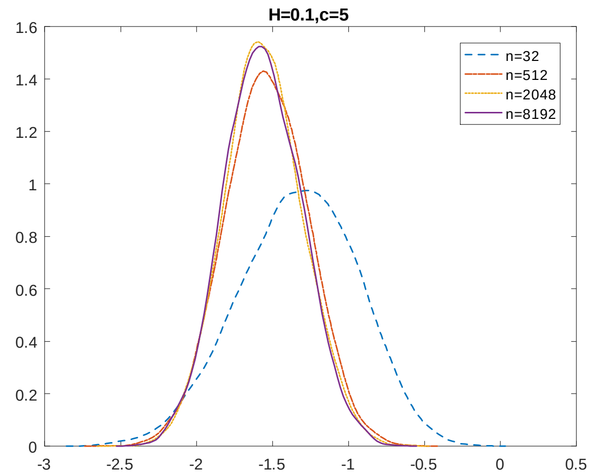

Figure 11: The density of $n^{0.2}\,(\widetilde{\rho}_n - \rho_n)$.

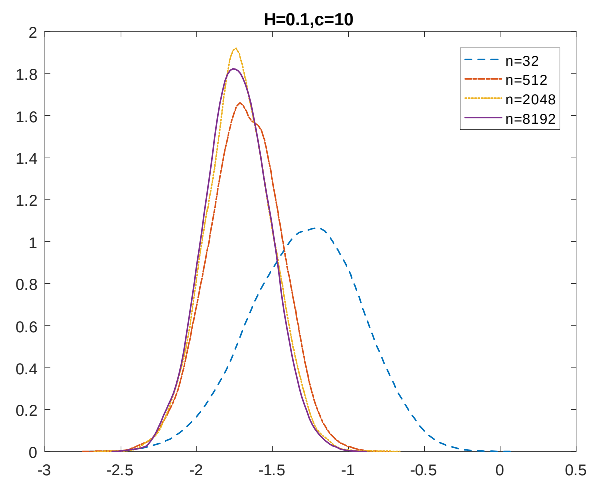

Figure 12: The density of $n^{0.2}\,(\widetilde{\rho}_n - \rho_n)$.

estimator has a slower rate of consistency when $H \in (0, 0.5)$ than when $H \in [0.5, 1)$.

Furthermore, the limiting distribution of the centered LS estimator depends on H. When $H = 0.5$, the limiting distribution is the same as that obtained in Phillips (1987a) for the local to unity model with errors for which the standard functional central theorem is applicable. When $H > 0.5$ or when $H < 0.5$, the limiting distributions are new to the literature. The limiting distribution for $H > 0.5$ has one term less than that for $H = 0.5$. The limiting distribution for $H < 0.5$ has three terms less than that for $H = 0.5$. Simulation studies are performed to check the reliability of the asymptotic approximation. When $H > 0.5$, a large sample size is needed for the limiting distribution to provide an accurate approximation to the finite sample distribution. When $H = 0.5$, a small sample size is enough for the limiting distribution to provide an accurate approximation to the finite sample distribution. When $H < 0.5$, a moderate sample size is needed for the limiting distribution to approximate the finite sample distribution accurately.

Appendix

Proof of Lemma 1. To prove Lemma 1.1, we first note that

$$\begin{aligned}
X_t &= \rho_n X_{t-1} + \varepsilon_t = \rho_n^t X_0 + \sum_{j=0}^{t-1} \rho_n^j \varepsilon_{t-j} = \rho_n^t X_0 + \sum_{s=1}^{t} \rho_n^{t-s} \varepsilon_s \\
&\sim \rho_n^t X_0 + \sigma \sum_{s=1}^{t} \rho_n^{t-s} \left[B^H(s) - B^H(s-1) \right] \\
&\sim \rho_n^t X_0 + n^H \sigma \sum_{s=1}^{t} \rho_n^{t-s} \left[B^H\left(\frac{s}{n}\right) - B^H\left(\frac{s-1}{n}\right) \right] \\
&= \rho_n^t X_0 + n^H \sigma \sum_{s=1}^{t} \rho_n^{t-s} \int_{(s-1)/n}^{s/n} dB^H(r),
\end{aligned}$$

where the fifth equation is from the similarity property of the fBm. We then have

$$\begin{aligned} n^{-H} X_t &\sim e^{-ct/n} \frac{X_0}{n^H} + \sigma \sum_{s=1}^{t} \int_{(s-1)/n}^{s/n} e^{-c(t-s)/n} dB^H(r) \\ &= e^{-ct/n} \frac{X_0}{n^H} + \sigma \sum_{s=1}^{t} \int_{(s-1)/n}^{s/n} e^{-c(t/n-r)} e^{-c(r-s/n)} dB^H(r) \\ &= e^{-ct/n} \frac{X_0}{n^H} + \sigma \sum_{s=1}^{t} \int_{(s-1)/n}^{s/n} e^{-c(t/n-r)} \left[1 + O\left(1/n\right)\right] dB^H(r) \\ &= e^{-ct/n} \frac{X_0}{n^H} + \sigma \int_0^{t/n} e^{-c(t/n-r)} dB^H(r) + O_p\left(1/n\right) \\ &= \sigma e^{-ct/n} \left[\pi_0 + o_p\left(1\right)\right] + \sigma \int_0^{t/n} e^{-c(t/n-r)} dB^H(r) + O_p\left(1/n\right) \\ &= e^{-ct/n} \sigma \pi_0 + \sigma J_c^H\left(t/n\right) + o_p\left(1\right) \end{aligned}$$

where the third equation is from the Taylor expansion of $e^{-c(r-s/n)}$ and the last equation comes from the definition of the fOU process $J_c^H\left(t/n\right)$ given in (16). Hence, for any $r \in [0,1]$.

$$n^{-H} X_{[nr]} \sim \exp\left\{-c\frac{[nr]}{n}\right\} \sigma\pi_0 + \sigma J_c^H\left(\frac{[nr]}{n}\right) + o_p\left(1\right) \Rightarrow e^{-cr}\sigma\pi_0 + \sigma J_c^H\left(r\right), \quad \text{as} \ \ n \to \infty.$$

Since $n^{-H} X_{[nr]}$ is a Gaussian process with a finite first-order absolute moment, it is easy to show that the above result holds uniformly in $r \in [0,1]$ under the Skorokhod topology. This proves Lemma 1.1.

Then, the convergence results in Lemma 1.2-1.3 can be obtained straightforwardly by using the continuous mapping theorem (Billingsley, 1968, p. 30).

To prove the results in Lemma 1.4-1.5, we first have

$$\begin{aligned} X_t^2 &= \left(\rho_n X_{t-1} + \varepsilon_t\right)^2 = \rho_n^2 X_{t-1}^2 + 2\rho_n X_{t-1}\varepsilon_t + \varepsilon_t^2 \\ &= X_{t-1}^2 + \left(\rho_n^2 - 1\right) X_{t-1}^2 + 2\rho_n X_{t-1}\varepsilon_t + \varepsilon_t^2, \end{aligned}$$

and

$$\sum_{t=1}^{n} X_{t-1}\varepsilon_t = \frac{1}{2\rho_n}\left\{X_n^2 - X_0^2 - \left(\rho_n^2 - 1\right)\sum_{t=1}^{n} X_{t-1}^2 - \sum_{t=1}^{n} \varepsilon_t^2\right\}.$$

From the results in Lemma 1.1-1.3, we have

$$\frac{X_n^2 - X_0^2}{n^{2H}} \Rightarrow \sigma^2 \left\{\left[J_c^H\left(1\right) + e^{-c}\pi_0\right]^2 - \pi_0^2\right\}$$

and

$$n^{-2H}\left(\rho_n^2-1\right)\sum_{t=1}^{n}X_{t-1}^2 \Rightarrow -2c\sigma^2\int_0^1\left[J_c^H(r)+e^{-cr}\pi_0\right]^2 dr.$$

It is crucially important to note that $\sum_{t=1}^{n}\varepsilon_t^2=O_p(n)$ for all values of $H\in(0,1)$ and

$$\begin{aligned}\frac{1}{n}\sum_{t=1}^{n}\varepsilon_t^2 &\sim n^{-1}\sigma^2\sum_{t=1}^{n}\left[B^H(t)-B^H(t-1)\right]^2 \\ &\sim n^{-1+2H}\sigma^2\sum_{t=1}^{n}\left[B^H\left(\frac{t}{n}\right)-B^H\left(\frac{t-1}{n}\right)\right]^2 \xrightarrow{p}\sigma^2,\end{aligned}$$

where the convergence result is from Proposition 4.2 in Vittasaari (2019). As a result,

$$n^{-2H}\sum_{t=1}^{n}\varepsilon_t^2 \xrightarrow{p}\begin{cases}0 & \text{when } H>0.5\\ \sigma^2 & \text{when } H=0.5\\ +\infty & \text{when } H=0.5\end{cases}.$$

This is the reason why $\sum_{t=1}^{n}X_{t-1}\varepsilon_t$ having distinct asymptotic behaviors when H takes various values.

Now, it can be seen clearly that, when $H=0.5$, the four items in the decomposition of $\sum_{t=1}^{n}X_{t-1}\varepsilon_t$ have the same order and, as $n\to\infty$,

$$\begin{aligned}&n^{-2H}\sum_{t=1}^{n}X_{t-1}\varepsilon_t\\ =\;&\frac{1}{2\rho_n}\left\{\frac{X_n^2-X_0^2}{n^{2H}}-n\left(\rho_n^2-1\right)\frac{1}{n^{1+2H}}\sum_{t=1}^{n}X_{t-1}^2-n^{-2H}\sum_{t=1}^{n}\varepsilon_t^2\right\}\\ \Rightarrow\;&\frac{\sigma^2}{2}\left\{\left[J_c^H(1)+e^{-c}\pi_0\right]^2-\pi_0^2+2c\int_0^1\left[J_c^H(r)+e^{-cr}\pi_0\right]^2 dr-1\right\}.\end{aligned}$$

Whereas, when $H>0.5$, it has $2H>1$. Thus, $\sum_{t=1}^{n}\varepsilon_t^2=O_p(n)$ is asymptotically dominated by the other terms in the decomposition of $\sum_{t=1}^{n}X_{t-1}\varepsilon_t$. Hence, it disappears in the limit of $n^{-2H}\sum_{t=1}^{n}X_{t-1}\varepsilon_t$ that takes the form of

$$\begin{aligned}n^{-2H}\sum_{t=1}^{n}X_{t-1}\varepsilon_t &= \frac{1}{2\rho_n}\left\{\frac{X_n^2-X_0^2}{n^{2H}}-n\left(\rho_n^2-1\right)\frac{1}{n^{1+2H}}\sum_{t=1}^{n}X_{t-1}^2-n^{-2H}\sum_{t=1}^{n}\varepsilon_t^2\right\}\\ &\Rightarrow \frac{\sigma^2}{2}\left\{\left[J_c^H(1)+e^{-c}\pi_0\right]^2-\pi_0^2+2c\int_0^1\left[J_c^H(r)+e^{-cr}\pi_0\right]^2 dr\right\}.\end{aligned}$$

In contrast, when $H < 0.5$, $2H < 1$, thereby, $\sum_{t=1}^{n} \varepsilon_t^2 = O_p(n)$ asymptotically dominates the other terms in the decomposition of $\sum_{t=1}^{n} X_{t-1}\varepsilon_t$. Hence,

$$n^{-1}\sum_{t=1}^{n} X_{t-1}\varepsilon_t = \frac{1}{2\rho_n}\left\{o_p(1) - n^{-1}\sum_{t=1}^{n}\varepsilon_t^2\right\} \xrightarrow{p} -\frac{\sigma^2}{2}.$$

The proof of Lemma 1 is completed.

Proof of Theorem 2: The theorem is the direct consequence of Lemma 1.3-1.5. In particular, (17) and (18) follow from Lemma 1.3-1.4 and (19) follow from Lemma 1.3 and Lemma 1.5.

Proof of Theorem 3: The centered LS estimator given in (6) has the following representation,

$$\widetilde{\rho}_n - \rho_n = \frac{\sum_{t=1}^{n} X_{t-1}\varepsilon_t - n^{-1}\left(\sum_{t=1}^{n} X_{t-1}\right)\left(\sum_{t=1}^{n}\varepsilon_t\right)}{\sum_{t=1}^{n} X_{t-1}^2 - n^{-1}\left(\sum_{t=1}^{n} X_{t-1}\right)^2}.$$

From the results in Lemma 1.2-1.3, when $n \to \infty$, the large sample theory of the denominator of $\widetilde{\rho}_n - \rho_n$ is obtained as

$$\begin{aligned} & n^{-1-2H}\sum_{t=1}^{n} X_{t-1}^2 - n^{-2-2H}\left(\sum_{t=1}^{n} X_{t-1}\right)^2 \\ \Rightarrow\ & \sigma^2\int_0^1 \left[J_c^H(r) + e^{-cr}\pi_0\right]^2 dr - \sigma^2\left(\int_0^1 \left[J_c^H(r) + e^{-cr}\pi_0\right] dr\right)^2. \end{aligned}$$

The numerator of $\widetilde{\rho}_n - \rho_n$ has two components, $\sum_{t=1}^{n} X_{t-1}\varepsilon_t$ and $-n^{-1}\left(\sum_{t=1}^{n} X_{t-1}\right)\left(\sum_{t=1}^{n}\varepsilon_t\right)$, respectively. The second term, based on the results in Equation (14) and in Lemma 1.2, has the order of n^{2H} and the following large sample property as $n \to \infty$:

$$-n^{-1-2H}\left(\sum_{t=1}^{n} X_{t-1}\right)\left(\sum_{t=1}^{n}\varepsilon_t\right) \Rightarrow -\sigma^2\int_0^1 \left[J_c^H(r) + e^{-cr}\pi_0\right] dr \cdot B^H(1).$$

Whereas, the order of the first term, that is $\sum_{t=1}^{n} X_{t-1}\varepsilon_t$, is n^{2H} when $H \geq 0.5$, and n when $H < 0.5$, as proved in Lemma 1.4-1.5. Therefore, when $H \geq 0.5$, the two terms in the numerator of $\widetilde{\rho}_n - \rho_n$ have the same magnitude and are equally important as

$n \to \infty$. In this case, the numerator has the order of n^{2H} and the limit can be obtained straightforwardly from Lemma 1.4. Note that the limits of $\sum_{t=1}^{n} X_{t-1}\varepsilon_t$ are different with each other when $H = 0.5$ and $H > 0.5$.

In contrast, when $H < 0.5$, the first term in the numerator of $\widetilde{\rho}_n - \rho_n$ dominates the second term. In this case, the numerator has the order of n and the following limit as $n \to \infty$:

$$
\begin{aligned}
& n^{-1}\left[\sum_{t=1}^{n} X_{t-1}\varepsilon_t - n^{-1}\left(\sum_{t=1}^{n} X_{t-1}\right)\left(\sum_{t=1}^{n}\varepsilon_t\right)\right] \\
= \ & n^{-1}\sum_{t=1}^{n} X_{t-1}\varepsilon_t + o_p(1) \xrightarrow{p} -\sigma^2/2,
\end{aligned}
$$

where the last limit comes from the result in Lemma 1.5.

From the limits obtained above of the numerator and the denominator of the estimator $\widetilde{\rho}_n - \rho_n$, the large sample theory presented in Theorem 3 can be obtained straightforwardly. The proof of Theorem 3 is completed.

References

Billingsley, P., 1968. *Convergence of Probability Measures.* Wiley, New York.

Chan, N. H, , Wei, C. Z., 1987. Asymptotic Inference for Nearly Nonstationary AR(1) Processes. *Annals of Statistics*, 15(3), 1050-1063

Cheridito, P., Kawaguchi, H., Maejima, M., 2003. Fractional Ornstein–Uhlenbeck processes. *Electronical Journal Probability* 8(3), 1-14.

Davydov, Y. A., 1970. The Invariance Principle for Stationary Processes, *Theory of Probability and Its Applications* 15, 487-489.

Gatheral, J., T. Jaisson, and Rosenbaum, M., 2018. Volatility is rough. *Quantitative Finance* 18 (6), 933-949.

Hamilton, J., 1994. *Time Series Analysis*, Princeton University Press.

Marinucci, D., and Robinson, P. M., 1999. Alternative forms of fractional Brownian motion. *Journal of Statistical Planning and Inference*, 80, 111-122.

Mikusheva, A., 2015. Second order expansion of t-statistic in autoregressive models. *Econometric Theory*, 31(3), 426-448.

Park, J. Y., 2003. Weak unit roots. Working Paper, Department of Economics, Rice University.

Phillips, P. C. B., 1987a. Toward a unified asymptotic theory for autoregression. *Biometrika* 74, 533-547.

Phillips, P. C. B., 1987b. Time series regression with a unit root. *Econometrica* 55, 277-301.

Sowell, F., 1990. The Fractional Unit Root Distribution, *Econometrica* 58, 495-505.

Taqqu, M., 1975. Weak Convergence to Fractional Brownian Motion and to the Rosenblatt Process, *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, 31, 287-302.

Wang, X., Xiao, W., and Yu, J., 2021. Modeling and Forecasting Realized Volatility with the Fractional Ornstein-Uhlenbeck Process. *Journal of Econometrics, forthcoming.*

Viitasaari, L., 2019. Necessary and sufficient conditions for limit theorems for quadratic variations of Gaussian sequences. *Probability Survey* 16, 62–98.