Journal of Business & Economic Statistics

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/ubes20

The Grid Bootstrap for Continuous Time Models

Yiu Lim Lui, Weilin Xiao & Jun Yu

To cite this article: Yiu Lim Lui, Weilin Xiao & Jun Yu (2022) The Grid Bootstrap for Continuous Time Models, Journal of Business & Economic Statistics, 40:3, 1390-1402, DOI: 10.1080/07350015.2021.1930014

To link to this article: https://doi.org/10.1080/07350015.2021.1930014

JOURNAL OF BUSINESS & ECONOMIC STATISTICS
2022, VOL. 40, NO. 3, 1390–1402
https://doi.org/10.1080/07350015.2021.1930014

Taylor & Francis
Taylor & Francis Group

The Grid Bootstrap for Continuous Time Models

Yiu Lim Lui[a], Weilin Xiao[b], and Jun Yu[c]

[a]Institute for Advanced Economic Research, Dongbei University of Finance and Economics, Dalian, China; [b]School of Management, Zhejiang University, Hangzhou, China; [c]School of Economics, Lee Kong Chian School of Business, Singapore Management University, Singapore, Singapore

ABSTRACT
This article proposes the new grid bootstrap to construct confidence intervals (CI) for the persistence parameter in a class of continuous-time models. It is different from the standard grid bootstrap of Hansen in dealing with the initial condition. The asymptotic validity of the CI is discussed under the in-fill scheme. The modified grid bootstrap leads to uniform inferences on the persistence parameter. Its improvement over in-fill asymptotics is achieved by expanding the coefficient-based statistic around its in-fill asymptotic distribution that is non-pivotal and depends on the initial condition. Monte Carlo studies show that the modified grid bootstrap performs better than Hansen's grid bootstrap. Empirical applications to the U.S. interest rates and volatilities suggest significant differences between the two bootstrap procedures when the initial condition is large.

ARTICLE HISTORY
Received March 2020
Accepted May 2021

KEYWORDS
Continuous-time models; Distributional expansion; Grid bootstrap; In-fill asymptotics; Probabilistic expansion; Uniform inference

1. Introduction

A popular model to describe the evolution of an economic time series $y(t)$ is given by the following Ornstein–Uhlenbeck (OU) process:

$$dy(t) = \kappa(\mu - y(t))\mathrm{d}t + \sigma \mathrm{d}W(t), y(0) = y_0, \tag{1}$$

where $\kappa \in [0, \infty)$, $\mu \in (-\infty, \infty)$, and $\sigma \in (0, \infty)$ are all constants, y_0 is the initial condition, and $W(t)$ is a standard Brownian motion (BM). In this model, κ captures the persistence of $y(t)$ and is the parameter of interest in the present article. Consider the case when a discrete sample of observations for $y(t)$ is available as y_{th} with $t = 1, 2, \ldots, T\ (:= N/h)$, where h is the sample interval and T is the sample size. Clearly, N is the time span over which the discrete-sampled data is available.

The exact discrete-time model corresponding to Equation (1) is given by

$$y_{th} = \rho_h(\kappa) y_{(t-1)h} + \mu\left(1 - e^{-\kappa h}\right) + \sqrt{(1 - e^{-2\kappa h})/(2\kappa)}\varepsilon_t, \tag{2}$$

where $\rho_h(\kappa) = e^{-\kappa h}, \varepsilon_t \overset{\text{iid}}{\sim} N(0, \sigma^2)$. The discrete-time representation in Equation (2) is a first-order autoregressive (AR(1)) model with the AR coefficient $\rho_h(\kappa) = e^{-\kappa h} = e^{-\kappa N/T} := e^{c/T}$ where $c = -\kappa N$. When $h \to 0$, $\rho_h(\kappa) \to 1$ and hence, Equation (2) is closely related to the following local-to-unity model studied in Chan and Wei (1987) and Phillips (1987),[1]

$$y_t = \rho_T y_{t-1} + \varepsilon_t, \rho_T = 1 + c/T, y_0 \sim O_p(1). \tag{3}$$

Hansen (1999) proposed the grid bootstrap to construct a confidence interval (CI) for ρ_T in model (3) and shows that as $T \to \infty$, the bootstrap CI (BCI) for ρ_T has an asymptotically correct coverage. Mikusheva (2007) strengthened Hansen's result by showing that the BCI for the AR coefficient is valid uniformly over the parameter space of ρ_T.

Since our continuous-time model is closely related to the local-to-unity model, the results in Hansen (1999) and Mikusheva (2007) motivated us to make use of the grid bootstrap to construct a CI for κ after κ is estimated by least square (LS). Unfortunately, the standard grid bootstrap procedure of Hansen cannot be directly applied to the continuous-time model due to a key difference between the continuous-time model and the local-to-unity model considered in Hansen (1999) and Mikusheva (2007). In particular, to make model (2) and model (3) comparable, both sides of model (2) must be divided by $\sqrt{(1 - e^{-2\kappa h})/(2\kappa)}$. Consequently, the initial condition in model (2) becomes $O_p\left(1/\sqrt{h}\right)$ as $h \to 0$, which is larger than those considered in Hansen (1999) and Mikusheva (2007). As a result, we have to modify Hansen's grid bootstrap by carefully dealing with the initial condition and obtain its asymptotic justification. The asymptotic justification of the modified bootstrap is made under the in-fill scheme, that is, by assuming $h \to 0$ and fixed N. It is shown that the BCI for κ obtained by the modified grid bootstrap has an asymptotically correct coverage uniformly over the parameter space for $\kappa \in [0, \infty)$. Unless $y_0 = 0$ in Equation (1), the standard grid bootstrap procedure does not necessarily lead to an asymptotically correct coverage as $h \to 0$. Moreover, we show that the modified grid bootstrap provides the second-order improvement compared to the in-fill asymptotic distribution that depends on the initial condition. This finding

[1]Phillips (1987) assumed that $\rho_T = \exp(c/T) = 1 + c/T + O(T^{-2}), y_0 = O_p(1)$, and $\{\varepsilon_t\}$ is a strong mixing sequence. Chan and Wei (1987) assumed that $y_0 = 0$ and $\{\varepsilon_t\}$ is a martingale difference sequence.

CONTACT Weilin Xiao wlxiao@zju.edu.cn School of Management, Zhejiang University, Hangzhou, 310058, China.
Supplementary materials for this article are available online. Please go to www.tandfonline.com/UBES.

is interesting as the in-fill asymptotic distribution has been found to outperform the long-span asymptotic distribution (i.e., when $N \to \infty$ and h is fixed) and the double asymptotic distribution (i.e., when $N \to \infty$ and $h \to 0$) in finite samples.

Our setup and approach have a few other attractive features. First, our method can be used to test for the unit root as well as for a stationary root. This is in sharp contrast to the approaches based on the long-span asymptotic scheme where the test statistics and their asymptotic distributions under the unit root null hypothesis (such as the Dickey–Fuller test and the Phillips–Perron test) are very different from those under the stationary null hypothesis (such as the KPSS test of Kwiatkowski et al. (1992) and the test proposed in Chang, Cheng, and Yao 2019). Second, as a by-product, the modified bootstrap method obtains an approximate median unbiased estimator of κ.

We organize the article as follows. Section 2 reviews some relevant results in the literature on the continuous-time model given by Equation (1) and relates some of them to those in the discrete-time AR(1) model. In Section 3, a more general continuous-time model is introduced. The LS estimator of κ and its in-fill asymptotic distribution are also discussed. In Section 4, we first propose the modified grid bootstrap to construct CIs for κ and provide the asymptotic justification to the proposed procedure. We then establish a probabilistic expansion that uses the in-fill asymptotic distribution as the leading term and explain how an approximate median unbiased estimator of κ is obtained as a by-product. Section 5 discusses how to implement the modified grid bootstrap procedure. Simulation studies, which aim to check the finite-sample performance of the modified bootstrap, are carried out in Section 6. Section 7 reports CIs for κ based on the U.S. interest rates and volatilities. Section 8 concludes. Proofs of Lemma 3.1, Remark 4.1, Theorem 4.1, Lemma 4.1, Theorem 4.2, and Remark 4.6 are given in the appendix. Proofs of other lemmas and lemmas that are used to prove Theorem 4.1 and Theorem 4.2 are given in the online supplement.

We use the following notations throughout the article: "$\Rightarrow$" means weak convergence in distribution; "$\to$" means convergence in real sequence; "$\sim$" means asymptotic equivalence; "$\overset{d}{=}$" means distributional equivalence; "$\to_p$, ""$\to_d$"and "$\to_{a.s.}$" mean convergence in probability, distribution, and almost surely, respectively.

2. A Literature Review

Let $Y := \{y_{th}\}_{t=1}^{T}$ be data generated from the continuous-time model given by Equation (1) with the exact discrete-time model given by Equation (2). Clearly, the sample size T can be made to go to infinity by either increasing N (the long-span scheme) or decreasing h (the in-fill scheme) or both (the double scheme). Dividing both sides of Equation (2) by $\sqrt{\frac{\sigma^2(1-e^{-2\kappa h})}{2\kappa}}$ gives rise to

$$x_{th} = \rho_h(\kappa)x_{(t-1)h} + \frac{\mu\left(1-e^{-\kappa h}\right)}{\sqrt{\sigma^2(1-e^{-2\kappa h})/(2\kappa)}} + \varepsilon_t, x_0 = \frac{y_0}{\sqrt{\sigma^2(1-e^{-2\kappa h})/(2\kappa)}}, \tag{4}$$

where $x_{th} = y_{th}/\sqrt{\sigma^2(1-e^{-2\kappa h})/(2\kappa)}$. Model (4) is an AR(1) process with $\rho_h(\kappa) = e^{-\kappa h}$. If $\kappa > 0$, then $0 < \rho_h(\kappa) < 1$. If $\kappa = 0$, then $\rho_h(\kappa) = 1$, implying the presence of a unit root. If $h \to 0$ but N is finite, then $\rho_h(\kappa) = e^{-\kappa h} = 1 - \kappa h + o(h) = 1 - \kappa N/T + o(h)$. So the in-fill asymptotic scheme implies that model (4) has a root that is local-to-unity with the local parameter $c = -\kappa N$ and the initial condition $x_0 \sim O_p\left(1/\sqrt{h}\right)$ that diverges as $h \to 0$ if $y_0 \sim O_p(1)$ but not 0. Let the LS estimator of $\rho_h(\kappa)$ be $\widehat{\rho}_h$ and the LS estimator of κ be $\widehat{\kappa} = -\ln\left(\widehat{\rho}_h\right)/h$.

The long-span, in-fill, and double asymptotic distributions for $\widehat{\kappa}$ have been derived in the literature. For example, when $\kappa > 0$, $\sqrt{N}(\widehat{\kappa} - \kappa) \Rightarrow N\left(0, \left(\exp(2\kappa h) - 1\right)/h\right)$ as $T \to \infty$; see Tang and Chen (2009). When $\kappa = 0$, $N\widehat{\kappa} \Rightarrow -\int_0^1 \overline{W}(r)dW(r)/\int_0^1 \overline{W}(r)^2dr$ with $\overline{W}(r) = W(r) - \int_0^1 W(s)ds$ as $T \to \infty$. The discontinuity in the long-span asymptotic distribution for κ echoes that for ρ in the discrete-time AR(1) model. Yu (2014) and Zhou and Yu (2015) developed the in-fill asymptotic distribution for $\widehat{\kappa}$ when μ is known and unknown, respectively. The in-fill asymptotic distribution is continuous in κ when κ passes zero. Unless $y_0 = 0$, the in-fill asymptotic distribution explicitly depends on the initial condition, y_0.

When κ is positive but reasonably close to zero, Yu (2014), Zhou and Yu (2015), and Bao, Ullah, and Wang (2017) found that the in-fill distribution is much closer to the finite-sample distribution than both the long-span and the double asymptotic distributions, even when 10 years or 50 years of monthly data are used. The superiority of the in-fill distribution over the long-span distribution is not surprising as the in-fill distribution depends explicitly on the initial condition and is asymmetric. While these two features can be found in the finite sample distribution, they are lost in the long-span and double asymptotic distributions. Unfortunately, the in-fill distribution depends on unknown parameters that cannot be consistently estimated under the in-fill scheme in general.

In the discrete-time literature on local-to-unity, the initial condition is typically assumed to be $O_p(1)$ and the corresponding long-span asymptotic distribution involves functionals of the OU process but is independent of the initial condition.[2] Phillips (1987) developed the in-fill asymptotic distribution for the LS estimator of AR coefficient when $y_0 = 0$ and μ is known (i.e., $\mu = 0$). In the same article, Phillips (1987) showed that this in-fill asymptotic distribution is the same as the long-span asymptotic distribution in the local-to-unity model with the initial condition of $O_p(1)$. Perron (1991) extended the results in Phillips (1987) by allowing for a general initial condition y_0. Alternatively, bootstrap methods have been proposed to make inference about the AR coefficient in the AR(1) model.

[2] From Mikusheva (2015), it can be easily shown that as $T \to \infty$, in the local-to-unity model with intercept, $T(\widehat{\rho}-\rho) \Rightarrow \int_0^1 \bar{J}_c(r)dW(r)/\int_0^1 \bar{J}_c(r)^2dr$ where $\bar{J}_c(r) = J_c(r) - \int_0^1 J_c(s)ds$ is the demeaned OU process with $J_c(r) = \int_0^r \exp(c(r-s))dW(s)$.

When the AR(1) model has a unit root, the long-span asymptotic distribution is nonstandard. Basawa et al. (1991) and Park (2003) introduced bootstrap procedures which improve upon the long-span asymptotic theory. In an important study, Park (2003) justified the bootstrap procedure by obtaining a probabilistic expansion for the Dickey–Fuller t statistic around the Dickey–Fuller distribution and shows that the bootstrap offers a second-order refinement for the Dickey–Fuller test. Under the local-to-unity model, Hansen (1999) introduced the grid bootstrap approach. Mikusheva (2007) showed that Hansen's grid bootstrap is uniformly valid. More recently, Andrews, Cheng, and Guggenberger (2020) showed that Hansen's grid bootstrap approach is uniformly valid for unknown innovation distributions. Mikusheva (2015) obtained a probabilistic expansion of the t statistic around the local-to-unity asymptotic distribution and shows that Hansen's grid bootstrap leads to a second-order improvement in the local-to-unity asymptotic approach. The results of Mikusheva (2015) are important because, when the AR(1) coefficient is less than but close to one, the local-to-unity asymptotic distribution tends to give better approximations to the finite-sample distribution than the normal distribution when the sample size is small or moderately large. However, since the initial condition is assumed to be $O_p(1)$ in the model of Mikusheva (2015), the local-to-unity asymptotic distribution is independent of the initial condition. It is unknown if the bootstrap distribution continues to offer improvement when the initial condition is larger.

We now review the concept of CI. Let ρ denote the parameter of interest in a statistical model and $t_T(Y,\rho)$ denote a test statistic with sampling distribution $F_T(x|\rho) = \Pr(t_T(Y,\rho) < x|\rho)$. For $q \in (0,1)$, let $c_T(q|\rho)$ be the quantile function of $t_T(Y,\rho)$, that is, $F_T(c_T(q|\rho)|\rho) = q$. Define a q -level CI for ρ by

$$\mathrm{CI}_q := \{\rho \in R : c_T(x_1|\rho) \le t_T(Y,\rho) \le c_T(x_2|\rho)\}, \tag{5}$$

where $x_1 = (1-q)/2$ and $x_2 = 1-(1-q)/2$. If ρ_0 is the true parameter value of ρ, by definition, $\Pr(\rho_0 \in \mathrm{CI}_q) = q$, and hence, the coverage probability is exactly q, the intended level. When $c_T(q|\rho)$ is replaced with the quantile function of a pivotal asymptotic distribution of $t_T(Y,\rho)$, the asymptotic CI has a correct probability coverage asymptotically. For example, if the asymptotic distribution is $N(0,1)$, then a 95% asymptotic CI is $\mathrm{CI}^A_{95\%} = \{\rho \in R : -1.96 \le t_T(Y,\rho) \le 1.96\}$. If $c_T(q|\rho)$ is replaced with the quantile function of a bootstrap distribution, denoted by $c^*_T(q|\rho)$, then the CI is a bootstrap confidence interval (BCI). There are some advantages of using BCIs over the asymptotic CIs. First, BCIs are obtained by re-sampling the data. Although asymptotic justification of bootstrap methods requires the knowledge of asymptotic theory, generating a BCI does not require an asymptotic scheme. Second, bootstrap methods are known to provide a finite sample refinement to asymptotic theory in the sense that the bootstrap distribution provides better approximations to the finite sample distribution than asymptotic distributions; see Chang and Hall (2015). Not surprisingly, BCIs often have a more accurate coverage than the asymptotic CIs.

3. The Model and In-Fill Asymptotic Theory

3.1. The Model and Estimator

Following Wang and Yu (2016), we consider the following continuous-time model:

$$dy(t) = \kappa(\mu - y(t))\mathrm{d}t + \sigma \mathrm{d}L(t), y(0) = y_0 = O_p(1), \tag{6}$$

where $\kappa \in [0,\infty)$, $\mu \in (-\infty,\infty)$, and $\sigma \in (0,\infty)$ are all constants, $L(t)$ is a Lévy process defined on a probability space $\left(\Sigma, \mathcal{F}, \{\mathcal{F}_t\}_{t\ge 0}, P\right)$, with $L(0) = 0$ a.s., $\mathcal{F}_t = \sigma\left\{\{y(s)\}_{s=0}^t\right\}$. The generalization from $W(t)$ to $L(t)$ is important in empirical applications for many financial variables; see Madan and Seneta (1990) for equity prices, Bai and Ng (2005) for interest rates, and Aït-Sahalia andJacod (2014) for an excellent textbook explanation of why $L(t)$ is important. The generalization makes the analytical approach of Bao, Ullah, and Wang (2017) infeasible.

In this article, we are interested in obtaining CIs for κ from discrete-sampled observations Y. Other parameters, such as μ,σ and parameters in $L(t)$, are treated as nuisance parameters. The exact discrete-time representation of Equation (6) is

$$\begin{aligned} y_{th} &= \rho_h(\kappa) y_{(t-1)h} + \mu(1-\exp(-\kappa h)) \\ &\quad + \sigma \int_{(t-1)h}^{th} \exp(-\kappa(th-s))\mathrm{d}L(s). \end{aligned} \tag{7}$$

The Lévy process makes $\left\{\sigma \int_{(t-1)h}^{th} \exp(-\kappa(th-s))\mathrm{d}L(s)\right\}_{t=1}^{N/h}$ an iid sequence with the distribution depending on the specification of the Lévy measure. Let the characteristic function of $L(t)$ be of the form of $E(\exp\{isL(t)\}) = \exp\{-t\psi(s)\}$, where $i = \sqrt{-1}$ and the function $\psi : R \to C$ is the Lévy exponent of $L(t)$.

Assuming that $L(t)$ is square-integrable, the error term in Equation (7) has the following moments:

$$\begin{aligned} &E\left(\sigma \int_{(t-1)h}^{th} \exp(-\kappa(th-s))\mathrm{d}L(s)\right) \\ &\quad = \sigma i\psi'(0)\frac{1-\exp(-\kappa h)}{\kappa}, \\ &\mathrm{var}\left(\sigma \int_{(t-1)h}^{th} \exp(-\kappa(th-s))\mathrm{d}L(s)\right) \\ &\quad = \sigma^2\psi''(0)\frac{1-\exp(-2\kappa h)}{2\kappa}. \end{aligned}$$

To simplify notations, let

$$\begin{aligned} \lambda_h(\kappa) &= \sqrt{\frac{1-e^{-2\kappa h}}{2\kappa}}, \sigma_\psi^2 = \sigma^2\psi''(0), \\ \mu^* &= \mu + \frac{\sigma i\psi'(0)}{\kappa}, g_h(\kappa) \\ &= \mu^*(1-\rho_h(\kappa)), \\ u_{th}(\kappa) &= (\sigma_\psi\lambda_h)^{-1}\Bigg(\sigma \int_{(t-1)h}^{th} \exp(-\kappa(th-s))dL(s) \\ &\quad -\sigma i\psi'(0)\frac{1-\rho_h(\kappa)}{\kappa}\Bigg). \end{aligned} \tag{8}$$

Note that $\{u_{th}\}_{t=1}^T$ is a sequence of iid variables with zero-mean and unit variance. When there is no confusion, we omit h in y_{th}

and u_{th}. Using notations in Equation (8), we rewrite Equation (7) as follows:

$$y_t = \rho_h(\kappa)y_{t-1} + g_h(\kappa) + \epsilon_t, \epsilon_t = \sigma_\psi \lambda_h(\kappa)u_t, y(0) = y_0 = O_p(1). \tag{9}$$

Dividing both sides of Equation (9) by $\sigma_\psi \lambda_h(\kappa)$ and letting $x_t = y_t/\left(\sigma_\psi \lambda_h(\kappa)\right)$, we have

$$x_t = \rho_h(\kappa)x_{t-1} + g_h(\kappa)/\left(\sigma_\psi \lambda_h(\kappa)\right) + u_t, x_0 = O_p(\lambda_h^{-1}(\kappa)). \tag{10}$$

As in Equation (4), the in-fill asymptotic scheme implies that Equation (10) is a local-to-unity model with the local parameter $c = -\kappa N$ and the initial condition $x_0 \sim O_p\left(1/\sqrt{h}\right)$ that diverges as $h \to 0$ when $y_0 \sim O_p(1)$ but not 0.

In model (7), let $\widehat{\rho}_h$ be the LS estimator of $\rho_h(\kappa)$ and $\widehat{\kappa}_h = -\ln(\widehat{\rho}_h)/h$. The coefficient-based statistic and t statistic for $\rho_h(\kappa)$ are, respectively,

$$z(Y, \rho, T) = T\left(\widehat{\rho}_h - \rho_h(\kappa)\right) \text{ and } t(Y, \rho, T) = \frac{\widehat{\rho}_h - \rho_h(\kappa)}{\widehat{\sigma}_{\widehat{\rho}_h}},$$

where $\widehat{\sigma}_{\widehat{\rho}_h} = \sqrt{\frac{1}{T}\sum_{t=1}^{T}(y_t - \widehat{g}_h - \widehat{\rho}_h y_{t-1})^2 \times \left(\sum_{t=1}^{T} y_{t-1}^2 - \frac{1}{T}\left(\sum_{t=1}^{T} y_{t-1}\right)^2\right)^{-1}}$. The normalization in $z(Y, \rho, T)$ is T not $\sqrt{T}$; see Phillips (1987). Following Perron (1991) and Zhou and Yu (2015), we define the coefficient-based statistic for κ as

$$z(Y, \kappa, h) = N\left(\widehat{\kappa}_h - \kappa\right). \tag{11}$$

3.2. In-Fill Asymptotic Theory

The in-fill asymptotic theory has gained much prominence in the recent years. Studies that have developed in-fill asymptotics for different econometric models include Li and Xiu (2016), Jiang, Wang, and Yu (2018, 2020). In this section, we extend the in-fill asymptotic result of Zhou and Yu (2015) to model (6).

Lemma 3.1. For model (6), define $z(Y, \kappa, h)$ by Equation (11) and let $\theta = (\mu^*, \sigma_\psi)$. Then, as $h \to 0$,

$$z(Y, \kappa, h) \Rightarrow -\frac{\Upsilon_3 - \Upsilon_2 W(1)}{\Upsilon_1 - \Upsilon_2^2} := z^{y_0}(\kappa, \theta), \tag{12}$$

where

$$\begin{aligned}\Upsilon_1 = &\frac{\exp(2c) - 4\exp(c) + 2c + 3}{2c^3}b^2 \\ &+ \frac{2b}{c}\int_0^1 (\exp(rc) - 1)J_c(r)\mathrm{d}r + \int_0^1 J_c^2(r)\mathrm{d}r \\ &+ \frac{\exp(2c) - 2\exp(c) + 1}{c^2}b\gamma_0 \\ &+ 2\gamma_0\int_0^1 \exp(rc)J_c(r)\mathrm{d}r + \gamma_0^2\frac{\exp(2c) - 1}{2c};\end{aligned}$$

$$\begin{aligned}\Upsilon_2 &= \frac{\exp(c) - c - 1}{c^2}b + \int_0^1 J_c(r)\mathrm{d}r + \frac{\exp(c) - 1}{c}\gamma_0; \\ \Upsilon_3 &= \frac{b}{c}\int_0^1 (\exp(rc) - 1)\mathrm{d}W(r) + \int_0^1 J_c(r)\mathrm{d}W(r) \\ &\quad + \gamma_0\int_0^1 \exp(rc)\mathrm{d}W(r); \\ J_c(r) &= \int_0^r \exp(c(r - s))\mathrm{d}W(s); \ \gamma_0 = \frac{y_0}{\sigma_\psi\sqrt{N}}; \\ b &= \mu^*\frac{\sqrt{-c\kappa}}{\sigma_\psi}; \ c = -\kappa N.\end{aligned}$$

If θ can be consistently estimated (say by $\widehat{\theta} = (\widehat{\sigma}_\psi, \widehat{\mu}^*)$), then $z(Y, \kappa, h)$ can be inverted to construct a feasible CI for κ as

$$\mathrm{CI}_q^A = \{\kappa \in R : c(x_1|\kappa) \le z(Y, \kappa, h) \le c(x_2|\kappa)\}, \tag{13}$$

where $c(\cdot|\kappa)$ is the quantile function of $z^{y_0}\left(\kappa, \widehat{\theta}\right)$. This method is related to the inversion method of Stock (1991). Its validity depends crucially on consistent estimation of θ.

While σ_ψ can be consistently estimated under the in-fill scheme (see Lemma 9.1 in the online supplement), it is unclear how to estimate μ^* consistently under the in-fill scheme. A natural estimator of μ^* is $\widehat{\mu}^* = \frac{\widehat{g}_h}{1-\widehat{\rho}_h}$, where $\widehat{g}_h$ is the LS estimator of g_h. The following lemma shows that $\widehat{\mu}^*$ is inconsistent as $h \to 0$. As a result, the inversion method based on $\widehat{\mu}^*$ cannot generate the asymptotically correct coverage as $h \to 0$.

Lemma 3.2. Let $\widehat{\mu}^* = \frac{\widehat{g}_h}{1-\widehat{\rho}_h}$, where $\widehat{g}_h$ is the LS estimator of the intercept in model (9). Then, as $h \to 0$, $\widehat{\mu}^* \sim O_p(h^{-2})$.

Remark 3.1. If model (6) is driven by the standard BM (i.e., $L(t) = W(t)$), then $\psi'(0) = 0$, $\psi''(0) = 1$, and the in-fill distribution of $\widehat{\kappa}_h$ given in Equation (12) is the same as that obtained from Zhou and Yu (2015). In addition, if μ is known and equal to 0, the in-fill distribution of $\widehat{\kappa}_h$ is identical to that in Perron (1991). By further assuming $y_0 = 0$, the in-fill distribution of $\widehat{\kappa}_h$ is the same as that in Phillips (1987).

Remark 3.2. If model (6) is driven by a standard BM, unless $y_0 = 0$, then the in-fill distribution of $\widehat{\kappa}$ depends on the initial condition via γ_0. If $y_0 = 0$ and $\mu = 0$, then γ_0 and b are both equal to 0 in Lemma 3.1. In this case, Lemma 3.1 implies that

$$\begin{aligned}z^{y_0}(\kappa, \theta) &= -\frac{\int_0^1 J_c(r)\mathrm{d}W(r) - \int_0^1 J_c(r)\mathrm{d}r W(1)}{\int_0^1 J_c^2(r)\mathrm{d}r - \left(\int_0^1 J_c(r)\mathrm{d}r\right)^2} \\ &= -\frac{\int_0^1 \bar{J}_c(r)\mathrm{d}W(r)}{\int_0^1 \bar{J}_c(r)^2\mathrm{d}r}.\end{aligned} \tag{14}$$

The initial condition disappears in Equation (14) and the in-fill asymptotic distribution is the mirror image of the long-span asymptotic distribution of $T(\widehat{\rho} - \rho)$ in the local-to-unity model when the initial condition is $O_p(1)$; see Remark 3 in Mikusheva (2015). If we further impose $\kappa = 0$, then $z^{y_0}(\kappa, \theta) = -\int_0^1 \overline{W}(r)\mathrm{d}W(r)/\int_0^1 \overline{W}(r)^2\mathrm{d}r$ where $\overline{W}(r)$ is the demeaned BM. When $y_0 \ne 0$, our in-fill asymptotic theory for $\widehat{\kappa}_h$ corresponds to the long-span asymptotic theory for $\widehat{\rho}$

in the local-to-unity model with the initial condition $O_p(\sqrt{T})$. In this case, the initial condition explicitly enters the in-fill asymptotic distribution via y_0, which is expected to outperform $-\int_0^1 \bar{J}_c(r)dW(r)/\int_0^1 \bar{J}_c(r)^2 \mathrm{d}r$.

4. Modified Grid Bootstrap

4.1. Modified Grid Bootstrap

The standard grid bootstrap of Hansen (1999) is for the local-to-unity AR(1) model with $y_0 \sim O_p(1)$. According to of Hansen (1999, sec. 4), the bootstrap initialization y_0^* is set to 0 when $\rho \geq 1$ and to the fitted value of y_0 that is based on the LS estimates otherwise. Moreover, y_0^* is then fixed in all bootstrap replications. This choice of initialization is made to avoid the dependence of the initialization. Since under the in-fill scheme the initial condition explicitly enters the in-fill asymptotic distribution in our model, the dependence of the initialization is needed. As a result, we have to modify the grid bootstrap procedure when generating a bootstrap sample. Consider generating the following AR(1) pseudo time series $\{y_t^*\}_{t=0}^T$ with errors $\{\varepsilon_t^*\}$ conditional on κ:

$$y_t^* = \rho_h(\kappa)y_{t-1}^* + \tilde{g}_h(\kappa) + \varepsilon_t^*, y^*(0) = y_0 = O_p(1), \tag{15}$$

where $\tilde{g}_h(\kappa)$ is obtained by regressing $y_t - \rho_h(\kappa)y_{t-1}$ on a constant. This way of obtaining $\tilde{g}_h(\kappa)$ is crucial since $g_h(\kappa)$ explicitly depends on κ in our model, unlike the usual discrete-time AR(1) model with intercept where the estimator of the intercept does not depend on the nuisance parameter. More importantly, when generating bootstrap samples, we explicitly retain the initial condition by letting $y^*(0) = y_0$, regardless of the value of $\rho_h(\kappa)$ (and hence, κ). This is critically different from the standard grid bootstrap that sets y_0^* to 0 when $\kappa \leq 0$ (or equivalently $\rho_h(\kappa) \geq 1$).

The error ε_t^* is generated in the following way. First, let $\{e_t\}_{t=1}^T$ be the LS residuals when y_t is regressed on y_{t-1} and a constant by LS. Second, we independently draw ε_t^* from $\{e_t\}_{t=1}^T$ with replacement and obtain a bootstrap sample $\{y_t^*\}_{t=1}^T$ $(:= Y^*)$ from Equation (15). Third, based on LS we obtain $\widehat{\rho}^*$, $\widehat{\kappa}_h^* = -\ln(\widehat{\rho}^*)/h$, $z(Y^*, \kappa, h) = N\left(\widehat{\kappa}_h^* - \kappa\right)$ from Y^*. Fourth, we repeat the above steps for B times to obtain the bootstrap distribution of $z(Y^*, \kappa, h)$. Finally, the BCI for κ is obtained as

$$\mathrm{CI}_q^* = \{\kappa \in R : c_T^*(x_1|\kappa) \leq z(Y, \kappa, h) \leq c_T^*(x_2|\kappa)\}, \tag{16}$$

where $c_T^*(\cdot|\kappa)$ is the quantile function of $z(Y^*, \kappa, h)$.

4.2. Asymptotic Validity of the Modified Bootstrap

We now provide asymptotic justification of the modified grid bootstrap under the in-fill scheme.

Theorem 4.1. Let κ_0 be the true value of κ. Assume that

1. $\kappa_0 \in K = [0, \infty)$, $\mu \in (-\infty, \infty)$, $\sigma \in (0, \infty)$.
2. The increment of the Lévy process $L(t+h) - L(t)$ has a finite rth absolute moment with some $r > 2$.

Under these two assumptions, as $h \to 0$,

$$\sup_{\kappa \in K} \sup_{x} \left|\Pr\{z(Y, \kappa, h) < x|\kappa\} - \Pr{}^*\{z(Y^*, \kappa, h) < x|\kappa, Y\}\right| \to 0, \quad P - a.s., \tag{17}$$

$$\lim_{h \to 0} \inf_{\kappa_0 \in K} \Pr\{\kappa_0 \in \mathrm{CI}_q^*\} = \lim_{h \to 0} \sup_{\kappa_0 \in K} \Pr\{\kappa_0 \in \mathrm{CI}_q^*\} = q, \tag{18}$$

where $\Pr^*$ is the probability with respect to the bootstrap distribution and $P -$ a.s. means "for almost all realizations of Y."

The two assumptions are primitive and easy to verify. Assumption 1 of Theorem 4.1 requires the parameter space of κ be the nonnegative half-line. Assumption 2, Theorem 4.1 is a typical moment restriction that enables us to apply the invariance principle. The result in Equation (17) shows that the distribution of the bootstrap statistic is closer to the finite sample distribution uniformly over the parameter space K, when h is closer to 0. In the limit, the bootstrap statistic behaves like a random variable whose distribution is the in-fill asymptotic distribution. The results in Equation (18) suggest that the coverage probability of CI_q^* converges to q when $h \to 0$.

Unfortunately, the grid bootstrap of Hansen does not necessarily have an asymptotically correct coverage when $h \to 0$. This problem arises because, by setting $y_0^* = 0$ for $\kappa \leq 0$, a bootstrap sample cannot approximate the original data with $y_0 \neq 0$ at nonpositive grid points. This result is reported below.

Lemma 4.1. Under the assumptions specified in Theorem 4.1, if we generate a bootstrap sample by letting $y_0^* = 0$ for $\kappa \leq 0$ and obtain CI_q^* as in Equation (16), then when $y_0 \neq 0$, there exists $0 < x_1 < x_2 < 1$, as $h \to 0$,

$$\Pr\{\kappa_0 \in \mathrm{CI}_q^*\} \nrightarrow x_2 - x_1.$$

Remark 4.1. Although we have used the coefficient-based statistic for κ to construct CIs, we can also construct CI for κ from a carefully defined t statistic. Note that in the proof of Lemma 3.1, as $h \to 0$, $N(\widehat{\kappa}_h - \kappa) = -T(\widehat{\rho}_h - \rho_h(\kappa)) + o_p(1)$. Therefore, we can express

$$\widehat{\kappa}_h - \kappa = -h^{-1}(\widehat{\rho}_h - \rho_h(\kappa)) + o_p(1). \tag{19}$$

From Equation (19), we may define the standard error of $\widehat{\kappa}_h$ by $\widehat{\sigma}_{\widehat{\kappa}_h} = h^{-1}\widehat{\sigma}_{\widehat{\rho}_h}$ and the t statistic for $\widehat{\kappa}_h$ by $t(Y, \kappa, T) = (\widehat{\kappa}_h - \kappa)/\widehat{\sigma}_{\widehat{\kappa}_h}$. It can be shown that, as $h \to 0$,

$$t(Y, \kappa, h) \Rightarrow -\frac{\Upsilon_3 - \Upsilon_2 W(1)}{\sqrt{\Upsilon_1 - \Upsilon_2^2}} := t^{y_0}(\kappa, \theta).$$

Remark 4.2. Since $t(Y^*, \kappa, h) = -t(Y^*, \rho, h) + o_p(1)$ under the in-fill scheme, following the approach of Mikusheva (2007) and the proof of Theorem 4.1, we can also justify the use of $t(Y^*, \kappa, h)$ to construct a CI for κ.

4.3. Asymptotic Expansion and the Second-Order Improvement

An important advantage of bootstrap methods over asymptotic distributions is that bootstrap often provides refinements in finite samples. This feature also holds true in our model. Park (2003) and Mikusheva (2015) justified the bootstrap approach by developing the second-order probabilistic expansions of the statistics of interest. The expansions are obtained in Park (2003) for both the t statistic and the coefficient-based statistic around their respective Dickey–Fuller distributions that are pivotal. The expansion is obtained in Mikusheva (2015) for the t statistic around $\int_0^1 J_c(r)\mathrm{d}W/\sqrt{\int_0^1 J_c(r)^2\mathrm{d}r}$ that is nonpivotal but independent of the initial condition. Our leading term is the in-fill asymptotic distribution, which is not only nonpivotal but also dependent on the initial condition.

Theorem 4.2. Assume that the assumptions specified in Theorem 4.1 hold, and additionally, the increment of Lévy process $L(t+h) - L(t)$ has a bounded rth moment for some $r \geq 8$. Then

$$z(Y,\kappa,h) = z^{y_0}(\kappa,\theta) + T^{-1/4}A + T^{-1/2}B + o_p(T^{-1/2}), \quad (20)$$

where the leading term $z^{y_0}(\kappa,\theta)$ is the in-fill asymptotic distribution given in Equation (12), and the full expressions of the higher order terms A and B, which are all $O_p(1)$, are provided in the appendix. Furthermore, the modified grid bootstrap leads to the distributional expansion:

$$\sup_x |\Pr{}^*(z(Y^*,\kappa,h) < x|\kappa, Y) - \Pr(z(Y,\kappa,h) < x|\kappa)| = o(T^{-1/2}),\ P-\text{a.s.} \quad (21)$$

Remark 4.3. Following Mikusheva (2015), one can expand $z(Y^*,\kappa,h)$ and show that the difference between $z(Y^*,\kappa,h)$ and $z(Y,\kappa,h)$ is $o_p(T^{-1/2})$. Hence, compared with $z^{y_0}(\kappa,\theta)$, $z(Y^*,\kappa,h)$ is closer to $z(Y,\kappa,h)$. The intuition for the improvement is that the bootstrap distribution but not the limit distribution depends on the distribution of error terms.

Remark 4.4. When $\psi'(0) = 0$, $\psi''(0) = 1$, $y_0 = \mu$, $\kappa = 0$, $z^{y_0}(\kappa,\theta) = -\int_0^1 \overline{W}(r)\mathrm{d}W(r)/\int_0^1 \overline{W}(r)^2\mathrm{d}r$. Equation (20) extends the result in Park (2003) from the unit root model without intercept to the unit root model with intercept. When $\psi'(0) = 0$, $\psi''(0) = 1$, $y_0 = \mu$, $z^{y_0}(\kappa,\theta) = -\int_0^1 \bar{J}(r)\mathrm{d}W(r)/\int_0^1 \bar{J}(r)^2\mathrm{d}r$. Equation (20) extends the result in Mikusheva (2015) from the local-to-unity model with the negligible initial condition to the local-to-unity model with the divergent initial condition.

Remark 4.5. According to Equation (20), we have

$$\Pr(z(Y,\kappa,h) < x|\kappa) = \Pr(z^{y_0}(\kappa,\theta) < x|\kappa) + O(T^{-1/4}) \quad (22)$$

uniformly in x. Let $\xi = z^{y_0}(\kappa,\theta) + T^{-1/4}A + T^{-1/2}B$. Theorem 4.2 suggests that ξ provides the second-order improvement to the in-fill asymptotic distribution since

$$\Pr(z(Y,\kappa,h) < x|\kappa) = \Pr(\xi < x|\kappa) + o(T^{-1/2}).$$

Remark 4.6. We can obtain a second-order expansion for $t(Y,\kappa,h)$, defined in Equation (20), as

$$t(Y,\kappa,h) = t^{y_0}(\kappa,\theta) + T^{-1/4}C + T^{-1/2}D + o_p(T^{-1/2}),$$

where the full expressions of C and D, which are all $O_p(1)$, are provided in the appendix.

4.4. Median Unbiased Estimator

The bootstrap sample can be used to construct not only an asymptotically valid CI for κ but also an approximate median unbiased estimator of κ. Define $\tilde{\kappa}_h$ as

$$\tilde{\kappa}_h = \begin{cases} \left(m_T^*(\kappa)\right)^{-1}|_{\kappa=\widehat{\kappa}_h} & \text{if } \left(m_T^*(\kappa)\right)^{-1}|_{\kappa=\widehat{\kappa}_h} \geq 0 \\ 0 & \text{otherwise} \end{cases}, \quad (23)$$

where $m_T^*(\kappa)$ is the median of the bootstrap distribution of $\widehat{\kappa}_h$ conditional on κ, and $\widehat{\kappa}_h$ is the LS estimator from the original sample. Then $\tilde{\kappa}_h$ is approximately median-unbiased.

Andrews (1993) constructed an exact median unbiased estimator for the AR parameter ρ under the assumption of Gaussian errors in the AR(1) model. In our case, $\tilde{\kappa}_h$ is not exact median unbiased because $m_T^*(\widehat{\kappa}_h|\kappa)$ is not the exact but an approximate median of the finite-sample distribution of $\widehat{\kappa}_h$. Since Equation (17) is satisfied under the in-fill scheme, it is immediate that, as $h \to 0$, $c_T^*(q|\kappa) \to c(q|\kappa)$ for any q. As $m_T^*(\kappa)$ is an affine transformation of $c_T^*(0.5|\kappa)$, $m_T^*(\kappa) \to m(\kappa)$ as $h \to 0$, where $m(\kappa)$ is the median function of finite sample distribution of $\widehat{\kappa}_h$. Consequently, $\tilde{\kappa}_h$ is approximately median unbiased.

5. Implementation

In this section, we discuss in practice how we can construct the modified grid bootstrap CI for κ and the approximate median unbiased estimator $\tilde{\kappa}_h$. We introduce the seven-step procedure.

1. Given the data $\{y_{th}\}_{t=0}^T$, we run the following LS regression:
$$y_{th} = \widehat{\rho}_h y_{(t-1)h} + \widehat{g}_h + e_{th},$$
where e_{th} is the LS residual. Use $\{e_{th}\}_{t=1}^T$ to construct $\widehat{\sigma}^2 = \frac{1}{T}\sum_{t=1}^T e_{th}^2$.
2. Construct a grid of ρ_h, $A_J = \{\rho_{h1}, \rho_{h2}, ...\rho_{hJ}\}$, centered at $\widehat{\rho}_h$, with the first and last grid points being $\widehat{\rho}_h \pm 7 \times se(\widehat{\rho}_h)$.
3. Given a point in the grid ($\rho_{hj} \in A_J, j \in \{1, ..., J\}$), perform the second LS regression:
$$y_{th} - \rho_{hj} y_{(t-1)h} = \tilde{g}_h + v_t,$$
where v_t is the residual. Note that $\tilde{g}_h$ is a function of ρ_{hj}.
4. We independently draw $\{\varepsilon_{th}^*\}_{t=1}^T$ from the empirical distribution of $\{e_{th}\}$ with replacement. We then generate a bootstrap sample $\{y_{th}^{*b}\}_{t=1}^T$ based on $\{\varepsilon_{th}^*\}_{t=1}^T$ and the same initial condition as the observed data, that is,[3]
$$y_{th}^{*b} = \rho_{hj} y_{(t-1)h}^{*b} + \tilde{g}_h + \varepsilon_{th}^*, y_0^{*b} = y_0.$$
5. Generate B sets of bootstrap sample, that is, $\left\{\{y_{th}^{*b}\}_{t=1}^T\right\}_{b=1}^B$. For every set of bootstrap sample, obtain the LS estimate of κ (denoted by $\widehat{\kappa}_h^*$) and calculate $z(Y^*,\kappa_j,h) = N(\widehat{\kappa}_h^* - \kappa_j)$ where $\kappa_j = -\ln(\rho_{hj})/h$. Calculate the xth quantile of $\{z(Y^*,\kappa_j,h)\}_{b=1}^B$ and $\{\widehat{\kappa}_h^*\}_{b=1}^B$, denoted by $c_T^*(x|\kappa_j)$ and $q_T^*(x|\kappa_j)$, respectively.

[3] In practice, if we suspects that the error term has unconditional heteroscedasticity, one may combine the wild bootstrap with the grid method, that is, $y_{th}^{*b} = \rho_{hj} y_{(t-1)h}^{*b} + \tilde{g}_h + e_{th}z_{th}^*, y_0^b = y_0, z_{th}^* \overset{\text{iid}}{\sim} N(0,1)$.

6. Estimate the quantile functions $c_T^*(x|\kappa)$ and $q_T^*(x|\kappa_G)$ by kernel regression,

$$c_T^*(x|\kappa) = \frac{\sum_{j=1}^J K\left(\frac{\kappa-\kappa_j}{\delta}\right) c_T^*(x|\kappa_j)}{\sum_{j=1}^J K\left(\frac{\kappa-\kappa_j}{\delta}\right)},$$

$$q_T^*(x|\kappa) = \frac{\sum_{j=1}^J K\left(\frac{\kappa-\kappa_j}{\delta}\right) q_T^*(x|\kappa_j)}{\sum_{j=1}^J K\left(\frac{\kappa-\kappa_j}{\delta}\right)}$$

   where $K(\cdot)$ is a kernel function and δ is a bandwidth. In the simulation and empirical studies that will be reported later, we use the Epanechnikov kernel ($K(x) = \frac{3}{4}(1-x^2)1(|x| \le 1)$) and choose the bandwidth by LS cross-validation.
7. The CI for κ is obtained by:

$$\mathrm{CI}_q^B = \left\{\kappa \in R : c_T^*(x_1|\kappa) \le z(Y,\kappa,h) \le c_T^*(x_2|\kappa)\right\}.$$

   The approximate median unbiased estimate $\tilde{\kappa}_h$ is calculated as in (23) by letting $m_T^*(\kappa) = q_T^*(0.5|\kappa)$.

6. Simulation Studies

To evaluate the performance of the modified bootstrap method in the continuous-time model, we construct CIs with the 95% coverage using three alternative methods: the in-fill asymptotic distribution when it is feasible, the standard grid bootstrap of Hansen (1999), and our modified grid bootstrap method in Section 5. The two bootstrap CIs only differ in the choice of initialization. We consider three data-generating processes (DGPs) (denoted DGP1 to DGP3) and simulate discrete observations with sampling interval h from model (6) where the L évy process is the variance gamma process with $\nu = 0.5$. In all three DGPs, we set $\kappa \in \{0.01, 0.1\}$, $h = 1/12$, $N = 5$, $\sigma = 1$, and $y_0 \in \{0.5, 1, 2\}$. In DGP1, we set $\mu = 0$, $\psi'(0) = 0$, $\psi''(0) = 1$. In DGP2, we set $\mu = 0.1$, $i\psi'(0) = 0$, and $\psi''(0) = 1$. In DGP3 we set $\mu = 0.1$, $i\psi'(0) = 0.05$, and $\psi''(0) = 1$. For DGP1, we assume that μ^* is known and equal to 0. This setting allows us to invert the quantiles based on $z^{y_0}(\kappa, 0, \widehat{\sigma}_\psi)$ in (12). For DGPs 2-3 since there is no consistent estimator for μ^*, the in-fill asymptotic distribution is not feasible and hence we only compare the two bootstrap methods. The number of replications is set at 2500. To calculate BCIs, we set the number of bootstrap iterations to $B = 399$ and the grid size to $J = 50$.

The Monte Carlo average is used to calculate the empirical coverage of the true value (κ_0), that is, $\frac{1}{2500}\sum_{m=1}^{2500} 1\left(\kappa_L^{(m)} \le \kappa_0 \le \kappa_U^{(m)}\right)$, where $\kappa_L^{(m)}$ and $\kappa_U^{(m)}$ are the bounds of a CI in the m th replication, and $1(\cdot)$ is the indicator that equals one if κ_0 is contained in the interval. The closer the empirical coverage to 95%, the better the performance of the method. Table 1 reports the empirical coverage and the median length of the CIs (in the parentheses) for alternative methods. Numbers in boldface indicate that the best performing method (in terms of the absolute difference) in each of the parameter settings.

Several interesting conclusions can be drawn from Table 1. First, it can be seen that the feasible in-fill CI is outperformed by the modified bootstrap. This is expected as we showed in Theorem 4.2 that our method is likely to produce a better finite-sample performance due to the refinement. Second, for the standard bootstrap, the empirical coverage is not close to the nominal one when the initial condition is large. It tends to lead to a too-small coverage probability and a too-narrow CI. This finding is particularly striking when the initial condition is larger. For example, for DGP3, when $y_0 = 2$ and $\kappa = 0.01$, the coverage of the standard bootstrap is only 88.4%. Finally, our modified bootstrap always performs the best with coverage always close to 95%. Regardless of y_0 , it tends to outperform the other CIs in all DGPs, consistent with the prediction of Theorem 4.2 and Lemma 4.1.

To evaluate the finite sample performance of $\tilde{\kappa}_h$, Table 2 reports the mean square error (MSE) of $\widehat{\kappa}_h$ and $\tilde{\kappa}_h$ under DGP1-DGP3. From Table 2, it is clear that $\tilde{\kappa}_h$ has a lower MSE under all

Table 2. MSE of $\widehat{\kappa}_h$ and $\tilde{\kappa}_h$.

			$y_0 = 0.5$	$y_0 = 1$	$y_0 = 2$
DGP1	$\kappa = 0.01$	$\widehat{\kappa}_h$	2.336	2.338	2.336
		$\tilde{\kappa}_h$	1.248	1.229	1.248
	$\kappa = 0.1$	$\widehat{\kappa}_h$	2.408	2.247	1.764
		$\tilde{\kappa}_h$	1.275	1.175	0.912
DGP2	$\kappa = 0.01$	$\widehat{\kappa}_h$	2.336	2.338	2.336
		$\tilde{\kappa}_h$	1.252	1.228	1.249
	$\kappa = 0.1$	$\widehat{\kappa}_h$	2.421	2.290	1.813
		$\tilde{\kappa}_h$	1.283	1.203	0.961
DGP3	$\kappa = 0.01$	$\widehat{\kappa}_h$	2.211	2.227	2.259
		$\tilde{\kappa}_h$	1.119	1.153	1.157
	$\kappa = 0.1$	$\widehat{\kappa}_h$	2.392	2.415	2.049
		$\tilde{\kappa}_h$	1.227	1.274	1.055

Table 1. 95% confidence intervals by alternative methods.

			$y_0 = 0.5$		$y_0 = 1$		$y_0 = 2$	
DGP1	$\kappa = 0.01$	In-fill distribution	0.936	(2.330)	0.929	(2.062)	0.890	(1.728)
		Hansen's bootstrap	0.940	(2.365)	0.925	(2.308)	0.887	(2.104)
		Modified bootstrap	**0.951**	(2.433)	**0.950**	(2.438)	**0.948**	(2.435)
	$\kappa = 0.1$	In-fill distribution	0.934	(2.486)	0.928	(2.146)	0.904	(2.104)
		Hansen's bootstrap	0.944	(2.536)	0.942	(2.418)	0.917	(2.138)
		Modified bootstrap	**0.950**	(2.593)	**0.954**	(2.516)	**0.962**	(2.312)
DGP2	$\kappa = 0.01$	Hansen's bootstrap	0.940	(2.367)	0.924	(2.307)	0.888	(2.104)
		Modified bootstrap	**0.951**	(2.430)	**0.950**	(2.435)	**0.949**	(2.434)
	$\kappa = 0.1$	Hansen's bootstrap	0.943	(2.550)	0.939	(2.429)	0.918	(2.153)
		Modified bootstrap	**0.948**	(2.603)	**0.952**	(2.521)	**0.959**	(2.335)
DGP3	$\kappa = 0.01$	Hansen's bootstrap	0.941	(2.355)	0.926	(2.299)	0.884	(2.071)
		Modified bootstrap	**0.956**	(2.454)	**0.952**	(2.460)	**0.953**	(2.442)
	$\kappa = 0.1$	Hansen's bootstrap	0.945	(2.563)	0.929	(2.464)	0.904	(2.187)
		Modified bootstrap	**0.953**	(2.630)	**0.948**	(2.606)	**0.958**	(2.446)

parameter settings. This finding echoes the results in Andrews (1993) and Andrews and Chen (1994) that the median unbiased estimator performs better than the LS estimator.

7. Empirical Studies

In this section, we apply the modified grid bootstrap method to construct BCIs for κ and $\tilde{\kappa}_h$ to estimate κ in model (1) and model (6) using the U.S. monthly interest rates with different maturities, the Chicago Board Options Exchange logarithmic volatility index (VIX), and the logarithmic S&P 100 volatility index (VXO). The methods used to generate CIs are identical to those in Section 6.

We obtain seven interest rates from the Federal Reserve Bank of St. Louis, including the Federal fund effective rate (FEDFUNDS), 3-month t-bill rate (TB3MS), 3-month, 1-year, 3-year, 5-year, and 10-year treasury constant maturity rates (GS3M, GS, GS3, GS5 and GS10). Our studies cover two time periods, from December 1994 to December 1999 and from January 2003 to January 2008. Table 3 reports the initial condition, $\widehat{\mu}$, $\widehat{\rho}_h(\kappa)$, $\widehat{\kappa}_h$, $\tilde{\kappa}_h$, and the two BCIs generated by the two grid bootstrap procedures.

Several interesting observations may be found from Table 3. First, the standard grid bootstrap always gives a narrower CI than the modified bootstrap. Second, the differences between the two sets of CIs are more noticeable in Panel A than those in Panel B. This finding is not surprising because the initial conditions in Panel A are generally larger. The first two observations agree with our simulation studies where it is found that the modified grid tends to lead to a wider BCI with a large initial condition and that the standard bootstrap method tends to lead to a too-small coverage probability. While the two sets of CIs are less different in Panel B than those in Panel A, we can still see the two CIs for GS10 are remarkably different from each other due to a larger initial condition. Based on the simulation studies and Lemma 4.1, the modified grid bootstrap method should produce a more accurate probability coverage. Third, all the CIs include zero, suggesting that we cannot reject the unit root null hypothesis. Interestingly, $\tilde{\kappa}_h$ in Panel B suggests all the interest rate data except GS10 have the median being exactly zero. This finding of unit root for the interest rate echoes the empirical studies in Romero-Ávila (2007). However, it is important to point out that the linear specification in the term structure and interest rate dynamics has been rejected by nonparametric methods; see Aït-Sahalia (1996). Hence, the failure of rejecting unit root in the interest rate dynamics could be due to nonlinear dynamics.

The daily observations from 2 February 2001 to 29 August 2008 and from 4 January 2010 to 30 December 2016 for CBOE VIX and VXO data are collected from the Federal Reserve Bank of St. Louis. Table 4 reports the initial condition, $\widehat{\mu}$, $\widehat{\rho}_h(\kappa)$, $\widehat{\kappa}_h$, $\tilde{\kappa}_h$, and the two BCIs. As in Table 3, the standard bootstrap method leads to narrower CIs in both cases. Unlike Table 3, all the CIs in two panels exclude zero, suggesting the evidence of stationarity in log-VIX and log-VXO.

Table 3. Empirical results for the interest rates data.

Panel A: December 1994 to December 1999 ($h = 1/12, N = 5$)							
	y_0	$\widehat{\mu}$	$\widehat{\rho}_h(\kappa)$	$\widehat{\kappa}_h$	$\tilde{\kappa}_h$	Standard 95% BCI	Modified 95% BCI
FEDFUNDS	5.45	5.350	0.931	0.858	0.045	(−0.329, 1.126)	(−0.768, 1.643)
TB3MS	5.60	4.935	0.900	1.269	0.729	(−0.316, 1.714)	(−0.600, 2.216)
GS3M	5.76	5.078	0.900	1.265	0.725	(−0.332, 1.711)	(−0.600, 2.209)
GS1	7.14	5.310	0.861	1.799	1.472	(−0.356, 2.335)	(−0.435, 2.752)
GS3	7.71	5.617	0.872	1.640	1.314	(−0.568, 2.179)	(−0.470, 2.637)
GS5	7.78	5.689	0.887	1.445	1.136	(−0.686, 1.953)	(−0.503, 2.344)
GS10	7.81	5.815	0.904	1.217	0.939	(−0.721, 1.642)	(−0.529, 2.026)
Panel B: January 2003 to January 2008 ($h = 1/12, N = 5$)							
	y_0	$\widehat{\mu}$	$\widehat{\rho}_h(\kappa)$	$\widehat{\kappa}_h$	$\tilde{\kappa}_h$	Standard 95% BCI	Modified 95% BCI
FEDFUNDS	1.24	7.052	0.989	0.139	0.0	(−0.433, 0.417)	(−0.496, 0.479)
TB3MS	1.17	4.468	0.983	0.206	0.0	(−0.434, 0.593)	(−0.495, 0.644)
GS3M	1.19	4.581	0.983	0.207	0.0	(−0.444, 0.594)	(−0.495, 0.645)
GS1	1.36	4.361	0.980	0.244	0.0	(−0.422, 0.631)	(−0.460, 0.691)
GS3	2.18	3.749	0.967	0.405	0.0	(−0.459, 0.855)	(−0.551, 0.882)
GS5	3.05	3.901	0.946	0.671	0.0	(−0.463, 1.146)	(−0.620, 1.337)
GS10	4.05	4.364	0.857	1.854	1.059	(−0.046, 2.456)	(−0.399, 3.237)

Table 4. Empirical results for the interest rates data for the log volatility indices.

Panel A: 2 February 2001 to 29 August 2008 ($N = 7.55, h = 1/252$)							
	y_0	$\widehat{\mu}$	$\widehat{\rho}_h(\kappa)$	$\widehat{\kappa}_h$	$\tilde{\kappa}_h$	Standard 95% BCI	Modified 95% BCI
Log-VIX	3.089	2.901	0.986	3.621	3.346	(2.078, 4.757)	(1.071, 5.204)
Log-VXO	3.209	2.929	0.987	3.219	2.700	(1.630, 4.348)	(0.731, 4.751)
Panel B: 4 January 2010 to 30 December 2016 ($N = 6.99, h = 1/252$)							
	y_0	$\widehat{\mu}$	$\widehat{\rho}_h(\kappa)$	$\widehat{\kappa}_h$	$\tilde{\kappa}_h$	Standard 95% BCI	Modified 95% BCI
Log-VIX	2.998	2.837	0.964	9.186	8.723	(6.600, 11.413)	(5.390, 11.967)
Log-VXO	2.878	2.800	0.960	10.277	7.787	(7.369, 12.755)	(6.285, 13.255)

8. Conclusion

In this article, we discuss the advantages and drawbacks of using three asymptotic distributions obtained from the long-span, double, and in-fill schemes for constructing CIs of persistence parameter κ under a Lévy-driven OU model. Although the in-fill asymptotic distribution is closer to the finite sample distribution than the long-span and double asymptotic distributions and is continuous in κ, it is generally infeasible in practice.

Since the discrete-time representation of our continuous-time model is similar to the local-to-unity model, for which the asymptotic justification of the standard grid bootstrap has been provided in the literature, it is natural to consider the use of the grid BCI for the continuous-time model. However, since the initial conditions in the exact discretization of the continuous-time model under the in-fill scheme is larger than those in the local-to-unity model typically assumed in the literature, the standard grid bootstrap may fail to provide the asymptotically correct coverage under the in-fill scheme. In this article, we propose a modified grid bootstrap method to cater a larger initial condition.

We show that the modified bootstrap leads to uniform inferences on the persistence parameter and provides a probabilistic expansion of the coefficient-based statistic around its in-fill asymptotic distribution. The probabilistic expansion allows us to establish a second-order approximation by the bootstrap distribution of the coefficient-based statistic to its finite sample distribution. While our parameter of interest is κ in this article, we expect our grid bootstrap performs better than the standard bootstrap for the AR coefficient in the AR(1) model when the initial condition is not zero.

Monte Carlo studies reveal several important results. First, the BCIs implied by the standard grid bootstrap generally suffer from an under-coverage problem. The under-coverage problem is more significant when the initial condition becomes larger. On the other hand, the BCI implied by the modified grid gives a probability coverage that is very close to the nominal value. Second, the modified grid bootstrap performs better than the in-fill asymptotic distribution, even when the latter is feasible.

The empirical application to the U.S. interest rate data under different maturities shows that the unit root hypothesis cannot be rejected by the modified grid bootstrap, suggesting non-stationarity in the interest rate data. Our BCIs are wider than those implied by the standard grid bootstrap. The empirical application to CBOE's VIX and VXO data shows that the unit root hypothesis is rejected by the modified grid bootstrap.

While the proposed modified grid bootstrap method performs better than the standard grid bootstrap method in a class of simple continuous-time models, it is not clear how to implement it for models with a more complicated dynamic behavior. An interesting class of model is the generalized local-to-unity (GLTU) model proposed recently by Dou and Müller (2021) whose limit is the continuous-time ARMA model (see Chambers and Thornton 2012). How to implement the grid bootstrap method for the GLTU model or the continuous-time ARMA model is an open question.

Acknowledgments

We thank to Peter Phillips for helpful discussions and comments.

Funding

Xiao gratefully acknowledges the financial support of the National Natural Science Foundation of China (No. 71871202). Yu acknowledges the financial support from Singapore Ministry of Education Tier 2 grant under grant MOE2018-T2-2-169.

References

Aït-Sahalia, Y. (1996), "Testing Continuous-Time Models of the Spot Interest Rate," *Review of Financial Studies*, 9, 385–426. [1397]
Aït-Sahalia, Y., and Jacod, J. (2014), *High-Frequency Financial Econometrics*, Princeton, NJ: Princeton University Press. [1392]
Andrews, D. W. K. (1993), "Exactly Median-Unbiased Estimation of First Order Autoregressive/Unit Root Models," *Econometrica*, 61, 139–165. [1395,1397]
Andrews, D. W. K., and Chen, H. (1994), "Approximately Median-Unbiased Estimation of Autoregressive Models," *Journal of Business & Economic Statistics*, 2, 187–204. [1397]
Andrews, D. W. K., Cheng, X., and Guggenberger, P. (2020), "Generic Results for Establishing the Asymptotic Size of Confidence Sets and Tests," *Journal of Econometrics*, 218, 496-531. [1392]
Bai, J., and Ng, S. (2005), "Tests for Skewness, Kurtosis, and Normality for Time Series Data," *Journal of Business & Economic Statistics*, 23, 49–60. [1392]
Bao, Y., Ullah, A., and Wang, Y. (2017), "Distribution of the Mean Reversion Estimator in the Ornstein–Uhlenbeck Process," *Econometric Reviews*, 36, 1039–1056. [1391,1392,1400]
Basawa, I. V., Mallik, A. K., McCormick, W. P., Reeves, J. H., and Taylor, R. L. (1991), "Bootstrapping Unstable First-Order Autoregressive Processes," *Annals of Statistics*, 19, 1098–1101. [1392]
Chambers, M., and Thornton, M. (2012), "Discrete Time Represenation of Continuous Time ARMA Processes," *Econometric Theory*, 28, 219–238. [1398]
Chan, N., and Wei, C. Z. (1987), "Asymptotic Inference for Nearly Nonstationary AR(1) Processes," *Annals of Statistics*, 15, 1050–1063. [1390]
Chang, J., Cheng, G., and Yao, Y. (2019), "A Power One Test for Unit Roots Based on Sample Autocovariances," London School of Economics and Political Science, Working paper. [1391]
Chang, J., and Hall, P. (2015), "Double-Bootstrap Methods that use a Single Double-Bootstrap Simulation," *Biometrika*, 102, 203–214. [1392]
Dou, L., and Müller, U. K. (forthcoming), "Generalized Local-to-Unity Models," *Econometrica*. [1398]
Hansen, B. E. (1999), "The Grid Bootstrap and the Autoregressive Model," *Review of Economics and Statistics*, 81, 594–607. [1390,1392,1394,1396]
Jiang, J., Wang, X., and Yu, J. (2018), "New Distribution Theory for the Estimation of Structural Break Point in Mean," *Journal of Econometrics*, 205, 156–176. [1393]
——— (2020), "In-Fill Asymptotic Theory for Structural Break Point in Autoregression," *Econometric Reviews*. [1393]
Kwiatkowski, D., Phillips, P. C. B., Schmidt, P., and Shin, Y. (1992), "Testing the Null Hypothesis of Stationarity Against the Alternative of a Unit Root," *Journal of Econometrics*, 54, 159–178. [1391]
Li, J., and Xiu, D. (2016), "Generalized Method of Integrated Moments for High-Frequency Data," *Econometrica*, 84, 1613–1633. [1393]
Mikusheva, A. (2007), "Uniform Inference in Autoregressive Models," *Econometrica*, 75, 1411–1452. [1390,1392,1394]
——— (2015), "Second Order Expansion of t-Statistic in Autoregressive Models," *Econometric Theory*, 31, 426–448. [1391,1392,1393,1395,1402]
Madan, D., and Seneta, E. (1990), "The Variance Gamma (V.G.) Model for Share Market Returns," *Journal of Business*, 63, 511–524. [1392]
Park, J. Y. (2003), "Bootstrap Unit Root Tests," *Econometrica*, 71, 1845–1895. [1392,1395]
Perron, P. (1991), "A Continuous-Time Approximation to the Unstable First-Order Autoregressive Process: The Case Without an Intercept," *Econometrica*, 59, 211–236. [1391,1393]

Phillips, P. C. B. (1987), "Towards a Unified Asymptotic Theory for Autoregression," *Biometrika*, 74, 535–547. [1390,1391,1393]
Romero-Ávila, D. (2007), "Unit Roots and Persistence in the Nominal Interest Rate: A Confirmatory Analysis Applied to the OECD," *Canadian Journal of Economics*, 40, 980–1007. [1397]
Stock, J. H. (1991), "Confidence Intervals for the Largest Autoregressive Root in US Macroeconomic Time Series," *Journal of Monetary Economics*, 28, 435–459. [1393]
Tang, C.Y., and Chen, S. X. (2009), "Parameter Estimation and Bias Correction for Diffusion Processes," *Journal of Econometrics*, 149, 65–81. [1391]
Wang, X., and Yu, J. (2016), "Double Asymptotics for Explosive Continuous-Time Models," *Journal of Econometrics*, 193, 35–53. [1392]
Yu, J. (2014), "Econometric Analysis of Continuous Time Models: A Survey of Peter Phillips's Work and Some New Results," *Econometric Theory*, 30, 737–774. [1391]
Zhou, Q., and Yu, J. (2010), "Asymptotic Distributions of the Least Squares Estimator for Diffusion Processes," Working Paper No. 02-2010, Singapore Management University. [1399]
——— (2015), "Asymptotic Theory for Linear Diffusions Under Alternative Sampling Schemes," *Economics Letters*, 128, 1–5. [1391,1393]

Appendix

Proof of Lemma 3.1 and Remark 4.1

Proof. In the following proof, although $\rho_h(\kappa), \widehat{\rho}_h(\kappa), g_h(\kappa), \widehat{g}_h(\kappa), \sigma_\psi(\kappa), \lambda_h(\kappa)$ are all dependent on κ, we omit this dependence to keep notations simple. We can follow Zhou and Yu (2010) to prove Lemma 3.1 and Remark 4.1. The only difference is that in Zhou and Yu (2010) $L(t) = W(t)$. If we divide equation (9) by $\sigma_\psi\lambda_h$, then we have $x_t = \rho_h x_{t-1} + \breve{g}_h + u_t$, where $\breve{g}_h = g_h/\left(\sigma_\psi\lambda_h\right)$. Under the in-fill scheme, we have

$$\frac{1}{T^2}\sum_{t=1}^{T} x_{t-1}^2 \Rightarrow \Upsilon_1,\ \frac{1}{T^{3/2}}\sum_{t=1}^{T} x_t \Rightarrow \Upsilon_2,\ \frac{1}{T}\sum_{t=1}^{T} x_{t-1}u_t \Rightarrow \Upsilon_3. \tag{A.1}$$

Let $S(T,\kappa) = \frac{1}{\widehat{\sigma}^2 T}\sum_{t=1}^{T} y_{t-1}\epsilon_t - \frac{1}{\widehat{\sigma}T}\sum_{t=1}^{T} y_{t-1}\frac{1}{\widehat{\sigma}T}\sum_{t=1}^{T}\epsilon_t$, and $R(T,\kappa) = \frac{1}{\widehat{\sigma}^2 T^2}\sum_{t=1}^{T} y_{t-1}^2 - \left(\frac{1}{\widehat{\sigma}T^{\frac{3}{2}}}\sum_{t=1}^{T} y_{t-1}\right)^2$, where $\widehat{\sigma}^2 = \frac{1}{T}\sum_{t=1}^{T}\left(y_t - \widehat{g}_h - \widehat{\rho}_h y_{t-1}\right)^2$. By construction,

$$T(\widehat{\rho}_h - \rho_h) = \frac{S(T,\kappa)}{R(T,\kappa)},\ t\,(Y,\rho,T) = \frac{S(T,\kappa)}{\sqrt{R(T,\kappa)}}.$$

Hence,

$$T(\widehat{\rho}_h - \rho_h) = \frac{\frac{1}{T}\sum_{t=1}^{T} x_{t-1}u_t - \frac{1}{\sqrt{T}}\sum_{t=1}^{T} u_t\frac{1}{T^{3/2}}\sum_{t=1}^{T} x_{t-1}}{\frac{1}{T^2}\sum_{t=1}^{T} x_{t-1}^2 - \left(\frac{1}{T^{3/2}}\sum_{t=1}^{T} x_{t-1}\right)^2}. \tag{A.2}$$

Letting $\varsigma_h(\cdot) = -\ln(\cdot)/h$, we have

$$\widehat{\kappa}_h - \kappa = \varsigma_h(\widehat{\rho}_h) - \varsigma_h(\rho_h) = \varsigma_h'(\widetilde{\rho}_h)(\widehat{\rho}_h - \rho_h),$$

where $\widetilde{\rho}_h(\kappa)$ is a value between $\widehat{\rho}_h$ and ρ_h. Therefore, we can write

$$\frac{T}{\varsigma_h'(\rho_h(\kappa))}(\widehat{\kappa}_h - \kappa) = \left(1 + \frac{\varsigma_h'(\widetilde{\rho}_h) - \varsigma_h'(\rho_h)}{\varsigma_h'(\rho_h)}\right)T(\widehat{\rho}_h - \rho_h). \tag{A.3}$$

Given $T = N/h$ and $N(\widehat{\kappa} - \kappa) = z\,(Y,\kappa,h)$, Equation (A.3) implies

$$z\,(Y,\kappa,h) = \xi_h(\kappa)z\,(Y,\rho,T)\,, \tag{A.4}$$

where $\xi_h(\kappa) = h\varsigma_h'(\rho_h)\left(1 + \frac{\varsigma_h'(\widetilde{\rho}_h) - \varsigma_h'(\rho_h)}{\varsigma_h'(\rho_h)}\right)$.

Since $\widehat{\kappa}_h = \frac{-\ln(\widehat{\rho}_h)}{h}$, applying the generalized Delta method and using the relationship in Equation (A.4), $Th = N$, $\left(1 + \frac{\varsigma_h'(\widetilde{\rho}_h) - \varsigma_h'(\rho_h)}{\varsigma_h'(\rho_h)}\right) \to_p 1$, $h\varsigma_h'(\widetilde{\rho}_h) \to_p -1$, and $\xi_h(\kappa) \to_p -1$, we obtain $z\,(Y,\kappa,h) \Rightarrow -\frac{\Upsilon_3 - \Upsilon_2 W(1)}{\Upsilon_1 - \Upsilon_2^2}$. This completes the proof of Lemma 3.1.

For $t\,(Y,\rho,T)$, we have

$$t\,(Y,\rho,T) = \frac{\sum_{t=1}^{T} y_{t-1}\epsilon_t - \frac{1}{T}\sum_{t=1}^{T} y_{t-1}\frac{1}{T}\sum_{t=1}^{T}\epsilon_t}{\sqrt{\widehat{\sigma}^2\left(\sum_{t=1}^{T} y_{t-1}^2 - \frac{1}{T}\left(\sum_{t=1}^{T} y_{t-1}\right)^2\right)}} \tag{A.5}$$
$$= \frac{\sigma_\psi\lambda_h}{\widehat{\sigma}_c\sqrt{h}}\left[\frac{\frac{1}{T}\sum_{t=1}^{T} x_{t-1}u_t - \frac{1}{T^{3/2}}\sum_{t=1}^{T} x_{t-1}\frac{1}{\sqrt{T}}\sum_{t=1}^{T} u_t}{\sqrt{\frac{1}{T^2}\sum_{t=1}^{T} x_{t-1}^2 - \left(\frac{1}{T^{3/2}}\sum_{t=1}^{T} x_{t-1}\right)^2}}\right],$$

where $\widehat{\sigma}_c = \frac{1}{T\lambda_h^2}\sum_{t=1}^{T}\left(y_t - \widehat{g}_h - \widehat{\rho}_h y_{t-1}\right)^2$. Note that $\lambda_h/h \to 1$. By Lemma 9.1 in the online supplement, we have $\frac{\sigma_\psi\lambda_h}{\widehat{\sigma}_c\sqrt{h}} \to_p 1$. Applying the results in Equation (A.1), we can obtain the limit of $t\,(Y,\rho,T)$.

To obtain the limit of $t(Y,\kappa,h)$, similar to Equation (A.4), we have

$$t\,(Y,\kappa,h) = \xi_h(\kappa)t\,(Y,\rho,T)\,. \tag{A.6}$$

Since $\xi_h(\kappa) \to_p -1$, we have $t\,(Y,\kappa,h) = -t\,(Y,\rho,T) + o_p(1)$ under the in-fill scheme, giving the result in Remark 4.1. □

Proof of Lemma 3.2

Proof. Note that $\widehat{\mu}^* = \frac{\widehat{g}_h}{1-\widehat{\rho}_h}$ and

$$\frac{\widehat{g}_h}{1-\widehat{\rho}_h} = \frac{T\widehat{g}_h}{T\left[1 - \rho_h + \left(\rho_h - \widehat{\rho}_h\right)\right]}. \tag{A.7}$$

For the denominator in Equation (A.7), applying Equations (A.1) and (A.2) in the main article, as $h \to 0$, we have

$$T\left[1 - \rho_h + \left(\rho_h - \widehat{\rho}_h\right)\right] = T(1 - \exp(c/T)) + T\left(\rho_h - \widehat{\rho}_h\right)$$
$$= -c + O_p(1). \tag{A.8}$$

For the numerator in Equation (A.7), letting $\widehat{\breve{g}}_h = \widehat{g}_h/\left(\sigma_\psi\lambda_h\right)$ and $\breve{g}_h = g_h/\left(\sigma_\psi\lambda_h\right)$, we have

$$T\widehat{g}_h = T\sigma_\psi\lambda_h\left(\widehat{\breve{g}}_h - \breve{g}_h + \breve{g}_h\right) = Tg_h + T\sigma_\psi\lambda_h\left(\widehat{\breve{g}}_h - \breve{g}_h\right)$$
$$= \frac{T\mu}{1-\rho_h(\kappa)} + \sigma_\psi\lambda_h O_p(T^{1/2})$$
$$= -\frac{T^2\mu}{c} + o(1) + \sigma_\psi O_p(N^{1/2}) = O_p(T^2). \tag{A.9}$$

Equations (A.7)–(A.9) give the result in Lemma 3.2 and the proof is completed. □

Proof of Theorem 4.1

Proof. First, let us define $(S(T,\kappa), R(T,\kappa))$ and $(S^*(T,\kappa), R^*(T,\kappa))$ as

$$S(T,\kappa) = \frac{1}{T}\left(\frac{1}{\widehat{\sigma}^2}\sum_{t=1}^{T} y_{t-1}\epsilon_t - \frac{1}{\widehat{\sigma}T}\sum_{t=1}^{T} y_{t-1}\frac{1}{\widehat{\sigma}}\sum_{t=1}^{T}\epsilon_t\right),$$

$$R(T,\kappa) = \frac{1}{T^2}\left(\frac{1}{\widehat{\sigma}^2}\sum_{t=1}^{T} y_{t-1}^2 - \left(\frac{1}{\widehat{\sigma}T}\sum_{t=1}^{T} y_{t-1}\right)^2\right),$$

$$S^*(T,\kappa) = \frac{1}{T}\left(\frac{1}{\widehat{\sigma}^2}\sum_{t=1}^{T} y_{t-1}^*\epsilon_t^* - \frac{1}{\widehat{\sigma}T}\sum_{t=1}^{T} y_{t-1}^*\frac{1}{\widehat{\sigma}}\sum_{t=1}^{T}\epsilon_t^*\right),$$

$$R^*(T,\kappa) = \frac{1}{T^2}\left(\frac{1}{\widehat{\sigma}^2}\sum_{t=1}^{T} y_{t-1}^{*2} - \left(\frac{1}{\widehat{\sigma}T}\sum_{t=1}^{T} y_{t-1}^*\right)^2\right).$$

Note that as in Equation (A.4), we can also express

$$z(Y,\kappa,h) = \xi_h(\kappa)\frac{S(T,\kappa)}{R(T,\kappa)}, z\left(Y^*,\kappa,h\right) = \xi_h^*(\kappa)\frac{S^*(T,\kappa)}{R^*(T,\kappa)}, \quad \text{(A.10)}$$

where $\xi_h(\kappa) = h\varsigma_h'(\rho_h)\left(1+\frac{\varsigma_h'(\tilde{\rho}_h)-\varsigma_h'(\rho_h)}{\varsigma_h'(\rho_h)}\right)$ and $\xi_h^*(\kappa) = h\varsigma_h'(\rho_h)\left(1+\frac{\varsigma_h'(\tilde{\rho}_h^*)-\varsigma_h'(\rho_h)}{\varsigma_h'(\rho_h)}\right)$.

We are now in the position to show

$$\sup_{\kappa\in K}\Pr\{|z(Y,\kappa,h) - z\left(Y^*,\kappa,h\right)| > \epsilon\} \to 0. \quad \text{(A.11)}$$

Note that for any $\kappa \in K$,

$$\begin{aligned}&\sup_{\kappa\in K}\Pr\left\{\left|\frac{S(T,\kappa)}{R(T,\kappa)} - \frac{S^*(T,\kappa)}{R^*(T,\kappa)}\right| > 2\epsilon\right\}\\ &\le \sup_{\kappa\in K}\Pr\left\{C\left|S(T,\kappa)-S^*(T,\kappa)\right| > \frac{\epsilon}{2}\right\}\\ &\quad+\sup_{\kappa\in K}\Pr\left\{C\left|R(T,\kappa)-R^*(T,\kappa)\right| > \frac{\epsilon}{2}\right\} \to 0,\end{aligned}$$

where the convergence can be established using items 5 to 8 of Lemma 9.4 in the online supplement. We can express

$$\begin{aligned}&\sup_{\kappa\in K}\Pr\left\{|z(Y,\kappa,h) - z\left(Y^*,\kappa,h\right)| > \epsilon\right\} \quad \text{(A.12)}\\ &= \sup_{\kappa\in K}\Pr\left\{\left|\xi_h(\kappa)\frac{S(T,\kappa)}{R(T,\kappa)} - \xi_h^*(\kappa)\frac{S^*(T,\kappa)}{R^*(T,\kappa)}\right| > \epsilon\right\}.\end{aligned}$$

For all value of $\kappa \in K$, we have

$$\frac{\varsigma_h'\left(\widehat{\rho}_h\right) - \varsigma_h'\left(\rho_h\right)}{\varsigma_h'\left(\rho_h\right)} = -\frac{\frac{1}{h\rho_h} - \frac{1}{h\widehat{\rho}_h}}{\frac{1}{h\rho_h}} = \frac{\rho_h - \widehat{\rho}_h}{\widehat{\rho}_h} \to_p 0, \quad \text{(A.13)}$$

$$\varsigma_h'(\rho_h)h = -\frac{h}{h\rho_h} \to -1. \quad \text{(A.14)}$$

Combing (A.12), (A.13) and (A.14), we obtain

$$\begin{aligned}&\sup_{\kappa\in K}\Pr\left\{|z(Y,\kappa,h) - z\left(Y^*,\kappa,h\right)| > \epsilon\right\}\\ &= \sup_{\kappa\in K}\Pr\left\{(1+o_p(1))\left|\frac{S(T,\kappa)}{R(T,\kappa)} - \frac{S^*(T,\kappa)}{R^*(T,\kappa)}\right| > \epsilon\right\} \to 0,\end{aligned}$$

This proves Equation (A.11).

We now proceed to show that $z\left(Y^*,\kappa,h\right)$ has a uniformly continuous probability distribution in the sense that as $\epsilon \to 0$,

$$\sup_{\kappa\in K}\sup_{x}\Pr{}^*\left\{x-\epsilon < z\left(Y^*,\kappa,h\right) < x+\epsilon\right\} \to 0,\ P-\text{a.s.} \quad \text{(A.15)}$$

For $\kappa \in K$, let $\{x_t^*\}$ be a bootstrap sample generated from

$$x_t^* = \rho_h x_{t-1}^* + \breve{g}_h + u_t, \quad \text{(A.16)}$$

where $x_t^* = y_t^*/(\sigma_\psi\lambda_h)$ and $\breve{g}_h = \tilde{g}_h/(\sigma_\psi\lambda_h)$. Let

$$\tilde{x}_t^* = \rho_h\tilde{x}_{t-1}^* + \breve{g}_h + z_t, z_t \overset{\text{iid}}{\sim} N(0,1). \quad \text{(A.17)}$$

Let $\widehat{\rho}_h^*$ and $\widehat{\rho}_h^{*,z}$ be the LS estimator of ρ_h in model (A.16) and model (A.17), respectively. Using Equation (A.17) and following the similar steps that proves Lemma 9.4 in the online supplement, we can show

$$T\left(\widehat{\rho}_h^* - \rho_h\right) = T\left(\widehat{\rho}_h^{*,z} - \rho_h\right) + o_{a.s.}(1).$$

Moreover, for model (A.17), Bao, Ullah, and Wang (2017) derived the exact cumulative distribution function of $\widehat{\rho}_h^{*,z} - \rho_h$, that is uniformly continuous. Hence, $T\left(\widehat{\rho}_h^*(\kappa) - \rho_h(\kappa)\right)$ is asymptotically uniformly continuous for any $\kappa \in K$.

Equations (A.15) and (A.11) imply that as $T \to 0$, we have

$$\begin{aligned}&\sup_{\kappa\in K}\sup_{x}\left|\Pr\{z(Y,\kappa,h) < x|\kappa\} - \Pr{}^*\{z\left(Y^*,\kappa,h\right) < x|Y,\kappa\}\right|\\ &\quad\to 0,\ P-a.s.\end{aligned}$$

The proof of the first part of Theorem 4.1 is completed.

To prove the second part of Theorem 4.1, following notations of Mikusheva (2007) by letting $F_{T,\kappa}(x) = \Pr\{z(Y,\kappa,h) < x|\kappa\}$, $G_{T,\kappa}(x) = \Pr{}^*\{z(Y^*,\kappa,h) < x|\kappa, Y\}$, $q_x^G(h,\kappa)$ be the xth quantile of $G_{T,\kappa}(x)$, we have

$$\begin{aligned}\Pr\{\kappa_0 \in \mathrm{CI}_q^*\} &= F_{T,\kappa}\left(q_{(1+q)/2}^G(h,\kappa)\right) - F_{T,\kappa}(q_{(1-q)/2}^G(h,\kappa))\\ &= F_{T,\kappa}\left(q_{(1+q)/2}^G(h,\kappa)\right) - G_{T,\kappa}\left(q_{(1+q)/2}^G(h,\kappa)\right)\\ &\quad -F_{T,\kappa}(q_{(1-q)/2}^G(h,\kappa)) + G_{T,\kappa}(q_{(1-q)/2}^G(h,\kappa))\\ &\quad +G_{T,\kappa}\left(q_{(1+q)/2}^G(h,\kappa)\right) - G_{T,\kappa}(q_{(1-q)/2}^G(h,\kappa))\\ &= (1+q)/2 - (1-q)/2\\ &\quad +\left[F_{T,\kappa}\left(q_{(1+q)/2}^G(h,\kappa)\right) - G_{T,\kappa}\left(q_{(1+q)/2}^G(h,\kappa)\right)\right]\\ &\quad +\left[G_{T,\kappa}(q_{(1-q)/2}^G(h,\kappa)) - F_{T,\kappa}(q_{(1-q)/2}^G(h,\kappa))\right]\\ &= q + \left[F_{T,\kappa}\left(q_{(1+q)/2}^G(h,\kappa)\right)\right.\\ &\quad\left. -G_{T,\kappa}\left(q_{(1+q)/2}^G(h,\kappa)\right)\right]\\ &\quad +\left[G_{T,\kappa}(q_{(1-q)/2}^G(h,\kappa)) - F_{T,\kappa}(q_{(1-q)/2}^G(h,\kappa))\right].\end{aligned}$$

Moreover, we can express

$$\begin{aligned}&\lim_{h\to 0}\sup_{\kappa\in K}\Pr\{\kappa_0 \in \mathrm{CI}_q^*\}\\ &= q + \lim_{h\to 0}\sup_{\kappa\in K}\left[F_{T,\kappa}\left(q_{(1+q)/2}^G(h,\kappa)\right)\right.\\ &\quad\left. -G_{T,\kappa}\left(q_{(1+q)/2}^G(h,\kappa)\right)\right]\\ &\quad + \lim_{h\to 0}\sup_{\kappa\in K}\left[G_{T,\kappa}(q_{(1-q)/2}^G(h,\kappa)) - F_{T,\kappa}(q_{(1-q)/2}^G(h,\kappa))\right]\\ &\le q + 2\lim_{h\to 0}\sup_{\kappa\in K}\sup_{x}|F_{T,\kappa}(x) - G_{T,\kappa}(x)| = q,\end{aligned}$$

where the last equality is from Equation (17). Note that

$$\begin{aligned}\Pr\{\kappa_0 \in \mathrm{CI}_q^*\} &= q - \left[G_{T,\kappa}\left(q_{(1+q)/2}^G(h,\kappa)\right)\right.\\ &\quad\left. -F_{T,\kappa}\left(q_{(1+q)/2}^G(h,\kappa)\right)\right]\\ &\quad -\left[F_{T,\kappa}(q_{(1-q)/2}^G(h,\kappa)) - G_{T,\kappa}(q_{(1-q)/2}^G(h,\kappa))\right]\\ &\ge q - 2\sup_{x}|F_{T,\kappa}(x) - G_{T,\kappa}(x)|.\end{aligned}$$

Hence, we have

$$\begin{aligned}&\lim_{h\to 0}\inf_{\kappa\in K}\Pr\{\kappa_0 \in \mathrm{CI}_q^*\} \ge q - 2\\ &\lim_{h\to 0}\sup_{\kappa\in K}\sup_{x}|F_{T,\kappa}(x) - G_{T,\kappa}(x)| = q,\end{aligned}$$

$$q \geq \lim_{h\to 0}\sup_{\kappa\in K}\Pr\{\kappa_0 \in \mathrm{CI}_q^*\} \geq \lim_{h\to 0}\inf_{\kappa\in K}\Pr\{\kappa_0\in \mathrm{CI}_q^*\} \geq q.$$

This completes the proof of Theorem 4.1. □

Proof of Lemma 4.1

Proof. To prove Lemma 4.1, we only need to find an example where the coverage probability is incorrect asymptotically. To simplify notations in various limits and without loss of generality, assume $\mu^* = 0$. We first study the limits of $\frac{1}{\hat\sigma T}\sum_{t=1}^T y_{t-1}^*\epsilon_t^*$, $\frac{1}{\hat\sigma^2 T^2}\sum_{t=1}^T y_{t-1}^{*2}$, and $\frac{1}{\hat\sigma T^{3/2}}\sum_{t=1}^T y_{t-1}^*$ when y_0^* is set to 0. From Equation (A.1) and Lemma 9.1, Lemma 9.4.1 of the online supplement, when $y_0^*=0$, we have,

$$\frac{1}{\hat\sigma^2T^2}\sum_{t=1}^T y_{t-1}^{*2} \Rightarrow \Upsilon_1^*,$$
$$\frac{1}{\hat\sigma T^{3/2}}\sum_{t=1}^T y_{t-1}^* \Rightarrow \Upsilon_2^*,$$
$$\frac{1}{\hat\sigma T}\sum_{t=1}^T y_{t-1}^*\epsilon_t^* \Rightarrow \Upsilon_3^*,$$

where

$$\Upsilon_1^* = \int_0^1 J_c^2(r)dr,\ \Upsilon_2^* = \int_0^1 J_c(r)dr,\ \Upsilon_3^* = \int_0^1 J_c(r)dW(r).$$

We can re-express Υ_1^*, Υ_2^* and Υ_3^* as

$$\Upsilon_1^* \overset{d}{=} \Upsilon_1 - 2\gamma_0\int_0^1 \exp(rc)J_c(r)dr - \gamma_0^2\frac{\exp(2c)-1}{2c},$$
$$\Upsilon_2^* \overset{d}{=} \Upsilon_2 - \gamma_0\frac{\exp(c)-1}{c},\ \Upsilon_3^* \overset{d}{=} \Upsilon_3 - \gamma_0\int_0^1\exp(rc)dW(r),$$

where Υ_1, Υ_2 and Υ_3 are defined in Equation (12) with $b=0$ since $\mu^*=0$. Note that Υ_1, Υ_2 and Υ_3 all depend on the initial condition y_0 via γ_0. The coefficient-based statistics of the bootstrap sample has the following limit:

$$z(Y^*,\kappa,h) \Rightarrow -\frac{\Upsilon_3^* - \Upsilon_2^* W(1)}{\Upsilon_1^* - \Upsilon_2^{*2}} \overset{d}{=} -\frac{\Upsilon_3 - \Upsilon_2 + E_{1,\gamma_0}}{\Upsilon_1 - \Upsilon_2^2 + E_{2,\gamma_0}} := z^0(\kappa,\theta), \tag{A.18}$$

where $\theta = (0,\sigma_\psi)$, and

$$E_{1,\gamma_0} = \gamma_0\frac{\exp(c)-1}{c}\int_0^1 dW(r) - \gamma_0\int_0^1\exp(rc)dW(r),$$
$$E_{2,\gamma_0} = 2\Upsilon_2\gamma_0\frac{\exp(c)-1}{c} - 2\gamma_0\int_0^1\exp(rc)J_c(r)dr - \gamma_0^2\frac{\exp(2c)-1}{2c} - \left(\gamma_0\frac{\exp(c)-1}{c}\right)^2.$$

Equation (A.18) allows us to write

$$z^0(\kappa,\theta) = -\frac{\Upsilon_3-\Upsilon_2}{\Upsilon_1-\Upsilon_2^2} + E_{3,\gamma_0} = z^{y_0}(\kappa,\theta) + E_{3,\gamma_0},$$

where $E_{3,\gamma_0} = -\frac{E_{1,\gamma_0}(\Upsilon_1-\Upsilon_2^2)-(\Upsilon_3-\Upsilon_2)E_{2,\gamma_0}}{(\Upsilon_1-\Upsilon_2^2)(\Upsilon_1-\Upsilon_2^2+E_{2,\gamma_0})} \neq 0$ almost surely when $y_0\neq 0$.

When $y_0 \neq 0$, $\Pr(z^0(\kappa,\theta)\leq x|\kappa) \neq \Pr(z^{y_0}(\kappa,\theta)\leq x|\kappa)$. Hence, we have

$$G_{T,\kappa}(x) = F_{T,\kappa}(x) + \left(G_{T,\kappa}(x) - F_{T,\kappa}(x)\right) \to_p F_{\infty,\kappa}(x) + E_{\gamma_0},$$

where $G_{T,\kappa}(x) = \Pr{}^*\{z(Y^*,\kappa,h) < x|Y,\kappa\}$, $F_{T,\kappa}(x) = \Pr(z(Y,\kappa,h)\leq x|\kappa)$, $F_{\infty,\kappa}(x) = \Pr(z^{y_0}(\kappa,\theta)\leq x|\kappa)$, $E_{\gamma_0} = G_{\infty,\kappa}(x) - F_{\infty,\kappa}(x)$ and $G_{\infty,\kappa}(x) = \Pr(z^0(\kappa,\theta)\leq x|\kappa)$. The BCI from the bootstrap sample that is initialized at $y_0^*=0$ is

$$\begin{aligned}\mathrm{CI}_q^* &= \{\kappa\in R: q_{x_1}^G(h,\kappa) \leq z(Y,\kappa,h) \leq q_{x_2}^G(h,\kappa)\}\\ &= \{\kappa\in R: G_{T,\kappa}\left(q_{x_1}^G(h,\kappa)\right) \leq G_{T,\kappa}(z(Y,\kappa,h))\\ &\quad \leq G_{T,\kappa}(q_{x_2}^G(h,\kappa))\\ &= \{\kappa\in R: x_1 \leq G_{T,\kappa}(z(Y,\kappa,h)) \leq x_2\}.\end{aligned} \tag{A.19}$$

Note that $G_{T,\kappa}(x) \to_p G_{\infty,\kappa}(y) = \Pr(z^0(\kappa,\theta)\leq y|\kappa)$, $F_{T,\kappa}(x) \to F_{\infty,\kappa}(y) = \Pr(z^{y_0}(\kappa,\theta)\leq y)$. Hence,

$$\begin{aligned}\Pr(G_{\infty,\kappa}(y)\leq x) &= \Pr(F_{\infty,\kappa}(y) + E\gamma_0 \leq x)\\ &= \Pr(F_{\infty,\kappa}(y) \leq x - E\gamma_0)\\ &= \Pr(y \leq F_{\infty,\kappa}^{-1}(x-E\gamma_0))\\ &= F_{\infty,\kappa}\left(F_{\infty,\kappa}^{-1}(x-E\gamma_0)\right)\\ &= x - E_{\gamma_0}.\end{aligned}$$

If $E_{\gamma_0}=0$, $G_{\infty,\kappa}(y)$ is an $U[0,1]$ random variable as $\Pr(G_{\infty,\kappa}(y)\leq x) = x$. In this case, the coverage probability of CI_q^* would go to $x_2 - x_1 = q$. However, if $E_{\gamma_0}\neq 0$, for any $x_1 < E\gamma_0$, $\Pr(G_{\infty,\kappa}(y)\leq x_1) = 0$. In this case, we have

$$\begin{aligned}\Pr(x_1 \leq G_{\infty,\kappa}(y)\leq x_2) &= \Pr(G_{\infty,\kappa}(y)\leq x_2)\\ &= x_2 - E\gamma_0 < x_2 - x_1.\end{aligned}$$

This gives an example where the coverage is not well controlled in the limit and hence, proves Lemma 4.1. □

Proof of Theorem 4.2 and Remark 4.6

Proof. To show the probabilistic expansion, we rewrite $z(Y,\rho,T)$ by applying Lemma 9.7 in the online supplement:

$$\begin{aligned}z(Y,\rho,T) &= \frac{\frac{1}{T}\sum_{t=1}^T x_{t-1}u_t - \frac{1}{T^{3/2}}\sum_{t=1}^T x_{t-1}\frac{1}{\sqrt{T}}\sum_{t=1}^T u_t}{\frac{1}{T^2}\sum_{t=1}^T x_{t-1}^2 - \left(\frac{1}{T^{3/2}}\sum_{t=1}^T x_{t-1}\right)^2}\\ &= A_1/B_1\end{aligned} \tag{A.20}$$

where

$$\begin{aligned}A_1 &= \Upsilon_3 - \Upsilon_2 W(1) + \frac{1}{T^{1/4}}(R_{3,T^{-1/4}} - M(V)\Upsilon_2)\\ &+\frac{1}{T^{1/2}}\left(R_{3,T^{-1/2}} - N(V)\Upsilon_2 - R_{2,T^{-1/2}}W(1)\right)\\ &-\frac{1}{T^{3/4}}R_{2,T^{-1/2}}M(V) - \frac{1}{T}R_{2,T^{-1/2}}N(V) + o_p(T^{-1/2}),\end{aligned}$$

$$\begin{aligned}B_1 &= \Upsilon_1 - \Upsilon_2^2 + \frac{1}{T^{1/2}}(R_{1,T^{-1/2}} - 2R_{2,T^{-1/2}})\\ &-\frac{1}{T}R_{2,T^{-1/2}} + o_p(T^{-1/2}).\end{aligned}$$

Expanding $z(Y,\rho,T)$ around its in-fill asymptotic distribution, we obtain

$$z(Y,\rho,T) = \frac{\Upsilon_3 - \Upsilon_2 W(1)}{\Upsilon_1 - \Upsilon_2^2} + \frac{1}{T^{1/4}}\frac{R_{3,T^{-1/4}} - M(V)\Upsilon_2}{\Upsilon_1-\Upsilon_2^2} + \frac{1}{T^{1/2}}\left(\begin{array}{c}\frac{R_{3,T^{-1/2}} - N(V)\Upsilon_2 - R_{2,T^{-1/2}}W(1)}{\Upsilon_1-\Upsilon_2^2}\\ -\frac{\Upsilon_3-\Upsilon_2W(1)}{(\Upsilon_1-\Upsilon_2^2)^2}\left(R_{1,T^{-1/2}} - 2R_{2,T^{-1/2}}\right)\end{array}\right)$$

$$+o_p\left(T^{-1/2}\right)$$
$$= z^{y_0}(\rho,\theta) + T^{-1/4}\tilde{A} + T^{-1/2}\tilde{B} + o_p(T^{-1/2}),$$

where

$$\tilde{A} = \frac{R_{3,T^{-1/4}} - M(V)\Upsilon_2}{\Upsilon_1 - \Upsilon_2^2},$$
$$\tilde{B} = \frac{R_{3,T^{-1/2}} - N(V)\Upsilon_2 - R_{2,T^{-1/2}}W(1)}{\Upsilon_1 - \Upsilon_2^2} - \frac{\Upsilon_3 - \Upsilon_2 W(1)}{(\Upsilon_1 - \Upsilon_2^2)^2}\left(R_{1,T^{-1/2}} - 2R_{2,T^{-1/2}}\right).$$

The expansion of $z(Y,\kappa,h)$ can be obtained from Equation (A.4). With $\xi_h(\kappa) \to_p -1$, we can show

$$z(Y,\kappa,h) = \xi_h(\kappa)\left(z^{y_0}(\rho,\theta) + T^{-1/4}\tilde{A} + T^{-1/2}\tilde{B} + o_p(T^{-1/2})\right)$$
$$= z^{y_0}(\kappa,\theta) + T^{-1/4}A + T^{-1/2}B + o_p(T^{-1/2}),$$

where $A = -\tilde{A}$ and $B = -\tilde{B}$.

Finally, for the last claim in Theorem 4.2, following Theorem 3 in Mikusheva (2015), we can easily show that the difference between $z(Y^*,\kappa,h)$ and $z(Y,\kappa,h)$ is $o_p(T^{-1/2})$. This completes the proof of Theorem 4.2.

For Remark 4.6, from Equation (A.5) note that

$$t(Y,\rho,T) = \frac{\sigma_\psi \lambda_h}{\widehat{\sigma}_c\sqrt{h}}\left[\frac{\frac{1}{T}\sum_{t=1}^T x_{t-1}u_t - \frac{1}{T^{3/2}}\sum_{t=1}^T x_{t-1}\frac{1}{\sqrt{T}}\sum_{t=1}^T u_t}{\sqrt{\frac{1}{T^2}\sum_{t=1}^T x_{t-1}^2 - \left(\frac{1}{T^{3/2}}\sum_{t=1}^T x_{t-1}\right)^2}}\right] = \left(1 + o_p(1)\right)\frac{A_1}{C_1},$$

where $C_1 = \sqrt{\Upsilon_1 - \Upsilon_2^2 + \frac{1}{T^{1/2}}(R_{1,T^{-1/2}} - 2R_{2,T^{-1/2}}) - \frac{1}{T}R_{2,T^{-1/2}}} + o_p(T^{-1/2})$. Expanding $t(Y,\rho,T)$ around its in-fill asymptotic distribution, we obtain

$$t(Y,\rho,T) = \frac{\Upsilon_3 - \Upsilon_2 W(1)}{\sqrt{\Upsilon_1 - \Upsilon_2^2}} + \frac{1}{T^{1/4}}\frac{R_{3,T^{-1/4}} - M(V)\Upsilon_2}{\sqrt{\Upsilon_1 - \Upsilon_2^2}} + \frac{1}{T^{1/2}}\left(\frac{R_{3,T^{-1/2}} - N(V)\Upsilon_2 - R_{2,T^{-1/2}}W(1)}{\sqrt{\Upsilon_1 - \Upsilon_2^2}} - \frac{1}{2}\frac{\Upsilon_3 - \Upsilon_2 W(1)}{(\Upsilon_1 - \Upsilon_2^2)^{2/3}}\left(R_{1,T^{-1/2}} - 2R_{2,T^{-1/2}}\right)\right) + o_p\left(T^{-1/2}\right).$$

From Equation (A.6) and $\xi_h(\kappa) \to_p -1$, we have

$$t(Y,\kappa,h) = t^{y_0}(\kappa,\theta) + T^{-1/4}C + T^{-1/2}D + o_p(T^{-1/2}),$$

where

$$C = -\frac{R_{3,T^{-1/4}} - M(V)\Upsilon_2}{\sqrt{\Upsilon_1 - \Upsilon_2^2}},$$

$$D = -\frac{R_{3,T^{-1/2}} - N(V)\Upsilon_2 - R_{2,T^{-1/2}}W(1)}{\sqrt{\Upsilon_1 - \Upsilon_2^2}} + \frac{1}{2}\frac{\Upsilon_3 - \Upsilon_2 W(1)}{(\Upsilon_1 - \Upsilon_2^2)^{2/3}}\left(R_{1,T^{-1/2}} - 2R_{2,T^{-1/2}}\right).$$

This completes the proof of Remark 4.6. □