In-fill Asymptotic Theory and Applications in Financial Econometrics

YIU LIM LUI *Dongbei University of Finance and Economics*

Recent years have seen a surge in the econometric development of in-fill asymptotic theory. Unlike the traditional large sample theory, which assumes that an increasing sample size is due to an increasing time span (denoted as the long-span asymptotic theory in this chapter), in-fill asymptotic theory assumes that the sample size increases because the sampling frequency is shrinking towards zero. Not surprisingly, the development of in-fill asymptotic theory is closely linked to the increased popularity of continuous time models. This chapter reviews the literature on the in-fill asymptotic theory and its applications in unit root testing and structural break models. In many applications, a non-standard limiting distribution will arise. In some cases, the initial condition shows up in the limiting distributions. Monte Carlo studies are carried out to check the performance of the in-fill asymptotic theory relative to the long-span asymptotic theory.

5.1 Introduction

Consider the following first-order autoregressive (AR (1)) model:

$$X_t = \rho X_{t-1} + \varepsilon_t,\ \varepsilon_t \sim_{iid} \left(0, \sigma^2\right), t = 1, 2, ..., T, X_0 = O_p(1), \tag{5.1.1}$$

where ρ is the AR coefficient, $\{\varepsilon_t\}$ is a sequence of a independent and identically distributed error term, T is the number of observations, and X_0 is the initialization of the time series.

Different values of ρ have a different implication for the dynamic behavior of the time series X_t. When $|\rho| < 1$, X_t is a stationary time series, which features a mean reversion; when $\rho = 1$, X_t is a unit root time series and follows a random walk; when $\rho > 1$, X_t is an explosive time series that can produce an upward trajectory. Not surprisingly, the model defined in (5.1.1) has elicited a wide range of interest in econometrics with studies focusing on the estimation and inference of ρ. An incomplete list of references includes Phillips (1987a,b), Phillips and Perron (1988) and Kwiatkowski et al. (1992), where the authors test whether a time series is stationary or not. The studies reported in Andrews (1993), Andrews and Chen (1994), Hansen (1999) and Mikusheva (2015), consider how to construct a better estimate and inference of ρ via bootstrapping. On the other hand, Diba and Grossman (1988), Evans (1991) and Phillips and Yu (2011); Phillips et al. (2015), explore the existence of explosiveness in the time series data.

Suppose ρ is estimated by least squares (LS) (denote as $\hat{\rho}$), the inference of ρ usually relies on asymptotic analysis where the number of observations T in model (5.1.1) is assumed to diverge to infinity. As discussed in Phillips (1987a), limiting distribution of $\hat{\rho}$ depends critically on the value of ρ,. When $|\rho| < 1$, upon normalization, $\hat{\rho}$ asymptotically has a Gaussian distribution, while when $\rho = 1$, $\hat{\rho}$ asymptotically has a nonstandard Dickey-Fuller distribution. If $\rho > 1$ and ε_t follows a Gaussian distribution in (5.1.1), White (1958) showed a normalized $\hat{\rho}$ follows a Cauchy distribution.

Since the value of ρ is not known a priori, the usually used limiting Gaussian distribution which assumes $|\rho| < 1$ is not always valid. Likewise, the use of the Dickey-Fuller distribution, which assumes $\rho = 1$, is incorrect when $\rho \neq 1$. Clearly, the discontinuity in ρ of the limiting distributional function and ignoring the value of ρ make the inference of ρ difficult because we need some knowledge of ρ to carry out its inference. However, difficulties may arise even if $|\rho|$ is known to be less than 1. For example, when ρ is close to 1 (say $\rho = 0.98$), it is well known that Dickey-Fuller distribution will provide a better finite sample approximation than the Gaussian distribution. However, as ρ moves further closer to the stationary region (say $\rho = 0.9$), Dickey-Fuller distribution starts to provide a less accurate finite sample approximation. Chan and Wei (1987) and Phillips (1987b) introduced a local to unit unity (LTU) model to cater the case where ρ is in the neighborhood of 1, and they assumed ρ has the form $\rho = 1 + c/T$ and $\rho = \exp(c/T)$, respectively, where c is a localizing constant. They found that the asymptotic distribution of $\hat{\rho}$ is nonstandard and continuous in $c$and provides a good finite sample performance when ρ is near 1. As expected, the LTU distribution outperforms the Gaussian distribution and Dickey-Fuller distribution for the finite sample approximation of $\hat{\rho}$ when ρ is close to 1.

Although the LTU distribution has a good finite sample performance, it does share the same drawback as the limiting Gaussian distribution and Dickey-Fuller distribution in that the initialization X_0 is not taken into account in the distributional functions. Since the finite sample distribution of $\hat{\rho}$ is dependent on X_0, none of the three asymptotic distributions of $\hat{\rho}$ may approximate the finite sample distribution well when the magnitude of X_0 is large.

In empirical studies, while we usually observe economic data at some discrete points (e.g., monthly and weekly data), we can think of the data series as a continuous time process being observed at some sampling interval Δ and over a time span N. More precisely, we assume a continuous time series $X(t)$ with T observations that are sampled in a discrete time interval at $[0, N]$ and therefore $t = 0, \Delta, 2\Delta..., T\Delta = N$. Monthly data in one year is a special case with $\Delta = 1/12$, $N = 1$ and $T = 12$. This setting allows us to obtain limiting distributions with an increasing T under one of the following three asymptotic schemes:

1. (Long span) Δ is fixed and $N \to \infty$.
2. (In-fill) N is fixed and $\Delta \to 0$.
3. (Double) $N \to \infty$ and $\Delta \to 0$.

The long-span scheme assumes that the sampling interval is fixed and the time span increases. This scheme corresponds to the ones used by White (1958), Phillips (1987a,b) and Phillips and Perron (1988). In-fill scheme assumes the time span is fixed and the sampling interval shrinks to zero. Finally, the double scheme assumes the span goes to infinity and the sampling interval shrinks to zero. Phillips (1987b), Perron (1991) and Zhou and Yu (2015) showed that the in-fill distributions of $\hat{\rho}$ under different AR(1) models provide a better finite sample approximation than the long-span distribution in three aspects. First, the in-fill distributions do not have a discontinuity in the value of ρ.. Second, X_0 is preserved at the limit. Third, the in-fill distribution is not symmetric. Eventually, the in-fill distribution outperforms the long-span distribution by exhibiting these three features and they can also be found in finite sample distributions of $\hat{\rho}$.

We have just mentioned that under an AR model, the in-fill distribution of $\hat{\rho}$ can outperform those under the long-span scheme. Interestingly, the AR model is not the only example in which the in-fill scheme provides a better finite sample performance. Suppose that we have the following time series with a structural break in mean:

$$\begin{aligned} X_t &= \begin{cases} \mu + u_t & \text{if } t \leq k_0 \\ \mu + \delta + u_t & \text{if } t > k_0 \end{cases}, t = 1, ..., T \text{ , } u_t \sim_{iid} (0, \sigma^2), \\ 1 &\leq k_0 < T. \end{aligned} \tag{5.1.2}$$

X_t has mean μ before the break date k_0, and the mean increases by δ after k_0. Suppose the goal is to estimate the structural break date k_0. Jiang et al. (2018) pointed out that when the break size δ is small, the finite sample distribution of the LS estimator of the break date is often asymmetric and has tri-modality. These two features are not captured by the limiting distribution developed under the long-span asymptotic scheme. Consequently, the long-span asymptotic distribution might poorly approximate the finite sample distribution. Jiang et al. (2018) then constructed a continuous-time model equivalent to the structure break model in (5.1.2) and introduced an in-fill asymptotic scheme. They proved that their in-fill asymptotic distribution has asymmetry and tri-modality similar to the finite sample distribution and can provide an excellent approximation. Jiang et al. (2021) extend the in-fill asymptotic theory to the AR(1) model with a structural break in the AR root.

This chapter consists of two parts. In the first part, we will discuss the long-span distribution of the LS estimator under the discrete time AR models. Then, corresponding continuous time models and their discrete time representations and in-fill theory of the LS estimator are introduced. In addition, we will discuss the application of the in-fill theory in unit root test and double asympotics for explosive time series. Finally, we briefly discuss some difficulties in applying the in-fill asymptotics and review the grid bootstrap confidence interval method that can be used to form an inference for the AR parameter.

In the second part of the chapter, we will introduce two structural break models in which one has a break in mean and the other has a break in the AR parameter. After discussing the finite sample distribution and its long-span distribution of the break point estimator, we will review the corresponding continuous time models and the in-fill asymptotic distribution of the break date estimator. Finally, we will show that in-fill theory outperforms long-span theory in finite sample approximation in both AR and structural break models.

It is worth mentioning that while we will introduce continuous time models when we study the in-fill asymptotics, we are primarily interested in the estimation and inference of the parameters in the discrete time models. The continuous time models are only used to obtain an in-fill distribution that eventually give us a better finite sample approximation. The chapter is organized as follows.

Section 2 reviews the unit root tests in the literature under continuous-time models. Section 3 reviews the structural break models. Section 4 concludes.

5.2 The AR model

Consider the following model:

$$X_t = d + \rho X_{t-1} + \varepsilon_t, \varepsilon_t \sim_{iid} \left(0, \sigma^2\right), t = 1, 2, ..., T, X_0 = O_p(1), \tag{5.2.1}$$

and the following two regressions:

$$X_t = \hat{\rho} X_{t-1} + \hat{e}_t, \tag{5.2.2}$$

$$X_t = \hat{d} + \tilde{\rho} X_{t-1} + \tilde{e}_t, \tag{5.2.3}$$

where $\tilde{\rho}$ and $\hat{\rho}$ are the LS estimators of ρ with and without intercept, respectively, $\hat{d}$ is the LS estimator of the intercept d in (5.2.1), $\hat{e}_t$ and $\tilde{e}_t$ are LS residuals. In this section, we analyze the limiting distribution of the LS estimators based on different asymptotic schemes.

5.2.1 Long-span asymptotics

The asymptotic distributions of $\hat{\rho}$ and $\tilde{\rho}$ where we let $T \to \infty$ are listed in the following theorems. Later, when we introduce the discrete time representation of a continuous time model, it will be clear that the long-span distribution where the $N \to \infty$ with a fixed Δ corresponds to $T \to \infty$ under model (5.2.1). For the moment, we will just call the asymptotic distributions of $\hat{\rho}$ and $\tilde{\rho}$ under model (5.2.1) the long-span asymptotic distribution. As will be shown, the long-span distributions of the LS estimators are crucially dependent on the value of AR parameter ρ.

The long-span distribution of the LS estimator $\hat{\rho}$ from regression (5.2.2) is summarized as follows.

Theorem 5.2.1 *Under model (5.2.1) with $d = 0$, the LS estimator $\hat{\rho}$ has the following limits as $T \to \infty$.*

(a) If $|\rho| < 1$,

$$\sqrt{T}(\hat{\rho} - \rho) \Rightarrow N\left(0, 1 - \rho^2\right), \tag{5.2.4}$$

(b) If $\rho = 1$,

$$T(\hat{\rho} - 1) \Rightarrow \frac{\int_0^1 W(r) dW(r)}{\int_0^1 W^2(r) dr}, \tag{5.2.5}$$

where $W(r)$ is a standard Brownian motion (BM).

(c) If $\rho > 1$, $\varepsilon_t \sim_{iid} N\left(0, \sigma^2\right)$ and $X_0 = 0$,

$$\frac{\rho^T}{\rho^2 - 1}(\hat{\rho} - \rho) \Rightarrow C, \tag{5.2.6}$$

where C is a standard Cauchy variate.

We also summarize the long-span distribution of the LS estimator of ρ from regression (5.2.3) as follows.

Theorem 5.2.2 *Under model (5.2.1), the LS estimator $\tilde{\rho}$ has the following limit as $T \to \infty$.*

(a) If $|\rho| < 1$,

$$\sqrt{T}(\tilde{\rho} - \rho) \Rightarrow N\left(0, 1 - \rho^2\right). \tag{5.2.7}$$

(b) If $\rho = 1$ and $d = 0$,

$$T(\tilde{\rho} - 1) \Rightarrow \frac{\int_0^1 \bar{W}(r)dW(r)}{\int_0^1 \bar{W}^2(r)dr}, \tag{5.2.8}$$

where $\bar{W}(r) = W(r) - \int_0^1 W(s)ds$ is the demeaned BM.

(c) If $\rho > 1$, $\varepsilon_t \sim i.i.d.N\left(0, \sigma^2\right)$,

$$\frac{\rho^T}{\rho^2 - 1}(\tilde{\rho} - \rho) \Rightarrow \frac{y}{z + \rho d/(\rho - 1)}, \tag{5.2.9}$$

where $y = N\left(0, \rho^2\sigma^2/(\rho^2 - 1)\right)$, $z = N\left(0, \rho^2\sigma^2/(\rho^2 - 1)\right) + \rho X_0$, and y and z are independently distributed.

The results (5.2.4) and (5.2.7) are standard. (5.2.5) and (5.2.8), which can be obtained from Phillips (1987a), are usually called the Dickey-Fuller distributions. (5.2.6) and (5.2.9) were derived by White (1958) and Wang and Yu (2015), respectively. The asymptotic results under $\rho > 1$, require that ε_t follows a Gaussian distribution, Therefore, (5.2.6) and (5.2.9) are not obtained from an invariance principle. Theorem 5.2.1 and 5.2.1 show that magnitude of ρ determines the form of the asymptotic distribution. When $|\rho| < 1$,, the LS estimates are asymptotically Gaussian. When $\rho = 1$, Dickey-Fuller distribution shows up at the limit. When $\rho > 1$, $\hat{\rho}$ asymptotically follow a Cauchy distribution and the limiting distribution of $\tilde{\rho}$ is nonstandard. Clearly, the asymptotic distributional functions of the LS estimators are not continuous in ρ.

Aside from the settings in Theorem 5.2.1 and 5.2.1, Phillips (1987b) and Chan and Wei (1987) introduced the local to unity (LTU) AR model where ρ is in the vicinity of 1. Phillips (1987b) assumed $\rho = \exp\left(\frac{c}{T}\right)$, $-\infty < c < \infty$, and it is called LTU because $\rho = \exp\left(\frac{c}{T}\right) = 1 + \frac{c}{T} + O(T^{-2}) \to 1$ as $T \to \infty$, and if $c = 0$, the AR model reduces to a unit root model. Under model (5.2.1) with $d = 0$ and $\rho = \exp\left(\frac{c}{T}\right)$, using repeated backward substitution, X_t can be expressed as a partial sum:

$$X_t = \sum_{j=1}^{t} \exp\left(\frac{c}{T}\right)^{t-j} \varepsilon_j + \exp\left(\frac{c}{T}\right)^t X_0.$$

Suppose $t = [Tr]$ with $r \in [0, 1]$, $W_n(r) = T^{-1/2}\sum_{j=1}^{[Tr]} \varepsilon_j$, pre-multiply X_t by $T^{-1/2}$, Phillips (1987b) showed that as $T \to \infty$,

$$\begin{aligned}
T^{-1/2}X_{[Tr]} &= \sum_{j=1}^{[Tr]} \exp\left(\frac{c}{T}\right)^{[Tr]-j} \frac{\varepsilon_j}{T^{1/2}} + \exp\left(\frac{c}{T}\right)^{[Tr]} \frac{X_0}{T^{1/2}} \\
&= \sigma \sum_{j=1}^{[Tr]} \exp\left(\frac{c}{T}\right)^{[Tr]-j} \int_{(j-1)/T}^{j/T} dW_n(s) + O_p(T^{-1/2}) \\
&= \sigma \int_0^r \exp(c(r-s))dW_n(s) + O_p(T^{-1/2}) \\
&\Rightarrow \sigma \int_0^r \exp\left(c(r-s)\right) dW(s) := J_c(r),
\end{aligned} \tag{5.2.10}$$

where we have applied the functional central limit theorem $W_n(s) \Rightarrow W(s)$ [1]to obtain the convergence (5.2.10). $J_c(r) = \int_0^r e^{c(r-s)}dW(r)$ is called an Ornstein-Uhlenbeck (OU) process and is generated by the stochastic process

$$dJ_c(r) = cJ_c(r)dr + dW(r)$$

[1]The functional central limit theorem ensures that $W_n(r)$ converges weakly to $W(r)$. For a detailed explanation, please see Phillips (1987a,b) and chapter 17 in Davidson (1994).

with $J_c(r) = 0$.

It is noteworthy that the way we normalize $X_{[Tr]}$ makes the initial condition X_0 irrelevant asympotically because $X_0/T^{-1/2} = O_p(T^{-1/2})$. Consequently, the LTU distributions of $\hat{\rho}$ and $\tilde{\rho}$ do not include the initialization.

We are now ready to discuss the following asymptotic distributions for the LS estimators of ρ under the LTU model.

Theorem 5.2.3 *Under (5.2.1) with $d = 0$, $\rho = \exp\left(\frac{c}{T}\right)$ and $-\infty < c < \infty$, the LS estimators $\hat{\rho}$ and $\tilde{\rho}$ have the following limits as $T \to \infty$,*

$$T(\hat{\rho} - \rho) \quad \Rightarrow \quad \frac{\int_0^1 J_c(r)dW(r)}{\int_0^1 J_c^2(r)dr}, \tag{5.2.11}$$

$$T(\tilde{\rho} - \rho) \quad \Rightarrow \quad \frac{\int_0^1 \bar{J}_c(r)dW(r)}{\int_0^1 \bar{J}_c^2(r)dr}, \tag{5.2.12}$$

where $\bar{J}_c(r) = \bar{J}_c(r) - \int_0^1 \bar{J}_c(s)ds$ is the demean OU process.

The proofs of (5.2.11) and (5.2.12) can be found in Phillips (1987b) and Mikusheva (2007), respectively, and are called the LTU distributions. When $c = 0$, LTU asymptotic distributions of $\hat{\rho}$ and $\tilde{\rho}$ become the Dickey-Fuller distributions. The limit distributions of $\hat{\rho}$ and $\tilde{\rho}$ under an LTU model are nonstandard, dependent on the localizing parameter c and the Ornstein-Uhlenbeck (OU) process $J_c(r)$. Phillips (1987b) showed the LTU distribution (5.2.11) is continuous in c, and when $c \to -\infty$ and $c \to +\infty$, the LTU asymptotic distribution of a suitably normalized $\hat{\rho}$ becomes the limiting Gaussian distribution and Cauchy distribution, respectively. Therefore, LTU asymptotic theory bridges the limiting Gaussian distribution, Dickey-Fuller distribution and Cauchy distribution.

From Theorem 5.2.1, 5.2.1 and 5.2.3, it can be seen that all asymptotic distributions except (5.2.9) are independent of the initial condition X_0, and the distributions are discontinuous in ρ except those under the LTU models. In simulations, Zhou and Yu (2015) showed that the asymmetry of the distribution of $\hat{\rho}$, the dependence on X_0, and the continuity of ρ can be seen in the finite sample. Since these three features are not captured by these asymptotic distributions, they may not provide a good finite sample approximation of the distributional functional of $\hat{\rho}$ and $\tilde{\rho}$ when ρ is close to 1 and X_0 is large.

5.2.2 Continuous time model and in-fill asymptotics

We now introduce a continuous time model that is closely related to the discrete time model (5.1.1). Consider the following stochastic differential equation

$$dX(t) = -\kappa X(t)dt + \sigma dW(t),\ X(0) = X_0 = O_p(1), t \in [0, N], \tag{5.2.13}$$

where $\kappa \in (-\infty, +\infty)$ measures the persistence of $X(t)$, $\sigma > 0$ measures the spot volatility and $W(t)$ is a standard BM.

Given a sampling interval Δ,, equation (5.2.13) has an exact discrete time first order autoregressive (AR(1)) representation:

$$X_{t\Delta} = \rho_\kappa X_{(t-1)\Delta} + \varepsilon_{t\Delta},\ t = 1, 2, ..., T = N/\Delta, \tag{5.2.14}$$

where $\rho_\kappa = \exp(-\kappa\Delta)$, $\varepsilon_{t\Delta} = \sigma \int_{(t-1)\Delta}^{t\Delta} \exp(-\kappa(t\Delta - s))dW(s) \sim_{iid} N\left(0, \sigma_\kappa^2\right)$ and $\sigma_\kappa^2 = \sigma^2 \frac{1-\exp(-2\kappa\Delta)}{2\kappa}$.

Equation (5.2.14) shows that the sign of κ (or equivalently ρ_κ) determines the dynamics of $X(t)$: if $\kappa > 0$, $\rho_\kappa < 1$ and $X(t)$ is a stationary time series; if $\kappa = 0$, $X(t)$ follows a random walk; if $\kappa < 0$, $\rho_\kappa > 1$ and $X(t)$ is an explosive time series.

When Δ is fixed, model (5.2.14) is equivalent to model (5.1.1) with $\rho = \rho_\kappa$, $X_t = X_{t\Delta}$ and σ in (5.1.1) becomes σ_κ. Therefore, the limiting results in Theorem 5.2.1 correspond to those under long-span schemes in model (5.2.14).

Under in-fill asymptotics, model (5.2.13) and Equation (5.2.14) are closely related to the LTU model studied in Phillips (1987b). To see this, since $T = N/\Delta$, $T \to \infty$ when $\Delta \to 0$, letting $c = -\kappa N$ gives

$$\rho_\kappa = \exp\left(\frac{c}{T}\right) = 1 + \frac{c}{T} + O(T^{-2}) \to 1 \text{ as } \Delta \to 0.$$

Dividing both sides of (5.2.14) by σ_κ, letting $Y_{t\Delta} = X_{t\Delta}/\sigma_\kappa$ and $\varepsilon_{t\Delta} = \varepsilon_{t\Delta}/\sigma_\kappa$ gives a discrete time representation:

$$\begin{aligned} Y_{t\Delta} &= \rho_\kappa Y_{(t-1)\Delta} + \varepsilon_{t\Delta}, t = 1, 2, ..., T, \\ Y_0 &= X_0/\sigma_\kappa = O_p(\sigma_\kappa^{-1}), \\ \varepsilon_{t\Delta} &\sim {}_{iid} N(0, 1). \end{aligned} \tag{5.2.15}$$

Model (5.2.15) therefore shows that $\{Y_{t\Delta}\}_{t=1}^T$ is an LTU model with an AR coefficient ρ_κ that has a localizing parameter $c = -\kappa N$. When $X_0 = 0$, Phillips (1987b) showed the convergence (5.2.10) holds for $Y_{t\Delta}$ upon normalization by $T^{-1/2}$.

However, when $X_0 = O_p(1)$, the initial condition in model (5.2.15) is diverging since $Y_0 = O_p(\sigma_\kappa^{-1})$ and

$$\sigma_\kappa^{-1} = \left(\frac{2\kappa}{\sigma\left(1 - \exp(-2\kappa\Delta)\right)}\right)^{-1} = \left(\frac{2\kappa}{\sigma(2\kappa\Delta + o(1))}\right)^{-1} = O\left(\frac{1}{\sqrt{\Delta}}\right) \to \infty \text{ as } \Delta \to 0.$$

Perron (1991) generalized the results of Phillips (1987b) by allowing for a nonzero X_0. Since $Y_0 = O_p(1/\sqrt{\Delta})$, upon a suitable normalization, the initial condition appears at the in-fill limit unlike that under the LTU model and we have

$$\exp\left(\frac{c}{T}\right)^{[Tr]} \frac{Y_0}{T^{1/2}} = \exp\left(\frac{c}{T}\right)^{[Tr]} \frac{X_0}{T^{1/2}\sigma_\kappa^2} \to e^{cr}\gamma_0,$$

where $\gamma_0 = \frac{X_0}{\sigma\sqrt{N}}$.

The following theorem gives the limit expression of $Y_{t\Delta}$ and its sample moments.

Theorem 5.2.4 *Under model (5.2.15), suppose $X_0 = 0$, as $\Delta \to 0$,*

$$T(\hat{\rho}_\kappa - \rho_\kappa) \Rightarrow \frac{\int_0^1 J_c(r)dW(r)}{\int_0^1 J_c^2(r)dr}. \tag{5.2.16}$$

Suppose $X_0 = O_p(1)$, as $\Delta \to 0$,

$$T(\hat{\rho}_\kappa - \rho_\kappa) \Rightarrow \frac{A(\gamma_0, c)}{B(\gamma_0, c)}, \tag{5.2.17}$$

where $\hat{\rho}_\kappa$ is the LS estimator of ρ_κ,

$$\begin{aligned} A(\gamma_0, c) &= \int_0^1 J_c(r)dW(r) + \gamma_0 \int_0^1 e^{cr} dW(r), \\ B(\gamma_0, c) &= \int_0^1 J_c^2(r)dr + 2\gamma_0 \int_0^1 e^{cr} J_c(r)dr + \gamma_0^2 \frac{e^{2c} - 1}{2c}, \end{aligned}$$

and $c = -\kappa N$.

(5.2.16) and (5.2.17) were proven in Phillips (1987b) and Perron (1991), respectively. Theorem 5.2.4 shows that the in-fill distribution of $\hat{\rho}_\kappa$ is the same as the LTU distribution (5.2.11) when the initialization $X_0 = 0$, and the in-fill distribution contains the initialization X_0 via γ_0 when $X_0 = O_p(1)$.

Zhou and Yu (2015) later generalized the model (5.2.13) in Phillips (1987b) and Perron (1991) by adding an additional parameter $\mu \in (-\infty, \infty)$ such that

$$dX(t) = \kappa\left(\mu - X(t)\right)dt + \sigma dW(t), X(0) = X_0 = O_p(1). \tag{5.2.18}$$

Under model (5.2.18), when $\kappa > 0$, $X(t)$ is stationary and mean-reverting, μ is the long-run mean of the model. Analogous to (5.2.14), Zhou and Yu (2015) used the exact discrete time representation of (5.2.18), which has the following form

$$X_{t\Delta} = \rho_\kappa X_{(t-1)\Delta} + g_\kappa + \varepsilon_{t\Delta},\ t = 1, 2, ..., T,\ g_\kappa = \mu(1 - e^{-\kappa\Delta}). \tag{5.2.19}$$

and is analogous to model (5.2.1).

Similar to (5.2.15), dividing both sides of (5.2.19) by σ_κ gives

$$Y_{t\Delta} = \rho_\kappa Y_{(t-1)\Delta} + \tilde{g}_\kappa + \varepsilon_{t\Delta},\ t = 1, 2, ..., T,\ \breve{g}_\kappa = g_\kappa/\sigma_\kappa, \tag{5.2.20}$$

Under in-fill asymptotic, the intercept term $\breve{g}_\kappa$ is not asymptotically negligible upon normalization. To see this, using backward substitution, $Y_{t\Delta}$ can be expressed as:

$$Y_{t\Delta} = \rho_\kappa^t Y_0 + \breve{g}_\kappa \sum_{j=1}^{t} \rho_\kappa^{t-j} + \sum_{j=1}^{t} \rho_\kappa^{t-j}\varepsilon_{j\Delta}.$$

Note that $\sum_{j=1}^{t} \rho_\kappa^{t-j}\varepsilon_{j\Delta} = O_p(T^{1/2})$ for $t = [Tr]$ from (5.2.10), $\sum_{j=1}^{t} \rho_\kappa^{t-j} = \frac{\rho_\kappa^{t+1}-1}{\rho_\kappa - 1} = \frac{\exp\left(\frac{c(t+1)}{T}\right)-1}{\exp\left(\frac{c}{T}\right)-1}$. If $t = [Tr]$ with $r \in [0, 1]$, we can see $\frac{\exp\left(\frac{c(t+1)}{T}\right)-1}{T\left[\exp\left(\frac{c}{T}\right)-1\right]} \to \frac{e^{cr}-1}{c}$ as $T \to 0$. Consequently, $\sum_{j=1}^{t} \rho_\kappa^{t-j} = O(T) = O(\Delta^{-1})$. Moreover, as $\Delta \to 0$, $N/T = \Delta$

$$\begin{aligned}
\breve{g}_\kappa &= g_\kappa/\sigma_\kappa = \mu(1 - e^{-\kappa\Delta}) \times \left(\frac{2\kappa}{\sigma^2\left(1 - \exp(-2\kappa\Delta)\right)}\right)^{1/2} \\
&= \mu(\kappa\Delta) \times \left(\frac{1}{\sigma^2\Delta}\right)^{1/2} + o(1) \\
&= \frac{\mu\kappa}{\sigma}\sqrt{\Delta} + o(1) \\
&= \frac{\mu\kappa}{\sigma}\sqrt{\frac{N}{T}} + o(1).
\end{aligned} \tag{5.2.21}$$

This implies $\breve{g}_\kappa \sum_{j=1}^{t} \rho_\kappa^{t-j} = O(T^{-1/2})O(T) = O(T^{1/2})$, and $T^{-1/2}\breve{g}_\kappa \sum_{j=1}^{t} \rho_\kappa^{t-j} \to \frac{\mu\kappa\sqrt{N}}{\sigma}\frac{e^{cr}-1}{c}$ which is dependent on μ.

Zhou and Yu (2015) showed that the introduction of $\mu \neq 0$ in (5.2.18) produces an in-fill asymptotic expression of $T^{-1/2}Y_{(t-1)\Delta}$ which consists of μ, γ_0 and the OU process $J_c(r)$. Additionally, the LS estimator of ρ_κ has the in-fill asymptotic distribution stated in the following theorem.

Theorem 5.2.5 *Under model (5.2.20), the LS estimator (with intercept)* $\tilde{\rho}_\kappa = \frac{\sum_{t=1}^{T}\left(Y_{(t-1)\Delta}-\bar{Y}\right)\left(Y_{t\Delta}-\bar{Y}\right)}{\sum_{t=1}^{T}\left(Y_{t\Delta}-\bar{Y}\right)^2}$ *where* $\bar{Y} = \frac{1}{T}\sum_{t=1}^{T} Y_{(t-1)\Delta}$ *has the following asymptotic distribution as* $\Delta \to 0$,

$$T(\tilde{\rho}_\kappa - \rho_\kappa) \Rightarrow \frac{A_2(\gamma_0, c)}{B_2(\gamma_0, c)}, \tag{5.2.22}$$

Table 5.1: Percentiles of $T(\tilde{\rho}_\kappa - \rho_\kappa)$ when $\kappa = -0.01$ and $\rho_\kappa = 1.0008$.

Percentile	1%	5%	10%	50%	90%	95%	99%
Finite sample	-19.26	-13.28	-10.70	-4.29	-0.83	-0.12	1.30
In-fill	-21.96	-13.66	-10.82	-3.92	-0.81	0.00	1.10
Long-span	-3.25	-0.59	-0.29	0.00	0.28	0.57	3.10

Table 5.2: Percentiles of $T(\tilde{\rho}_\kappa - \rho_\kappa)$ when $\kappa = 0.01$ and $\rho_\kappa = 0.9992$.

Percentile	1%	5%	10%	50%	90%	95%	99%
Finite sample	-18.75	-13.28	-10.78	-4.31	-0.82	-0.09	1.31
In-fill	-22.03	-14.91	-11.13	-4.12	-0.77	-0.02	1.06
Long-span	-0.74	-0.52	-0.41	0.00	0.41	0.52	0.74

where

$$A_2(\gamma_0, c) = \frac{b}{c}\int_0^1 c_1 dW(r) + \int_0^1 J_c(r)dW(r) + \gamma_0 \int_0^1 e^{cr} dW(r) - W(1)\left(c_2 b + \int_0^1 J_c(r)dr + c_4\gamma_0\right),$$

$$\begin{aligned} B_2(\gamma_0, c) &= c_3 b^2 + \frac{2b}{c}\int_0^1 c_1 J_c(r)dr + \int_0^1 J_c^2(r)dr + c_4^2 b\gamma_0 + 2\gamma_0 \int_0^1 e^{cr} J_c(r)dr \\ &\quad + \gamma_0^2 \frac{e^{2c}-1}{2c} - \left(c_2 b + \int_0^1 J_c(r)dr + c_4\gamma_0\right)^2, \end{aligned}$$

$c_1 = e^{rc} - 1, c_2 = \frac{e^c - c - 1}{c^2}, c_3 = \frac{e^{2c} - 4e^c + 2c + 3}{2c^3}, c_4 = \frac{e^c - 1}{c}$ *and* $b = \mu\frac{\sqrt{-c\kappa}}{\sigma}$.

Some interesting findings can be seen from the limit (5.2.5) and Theorem 5.2.1. First, the limit expressions of the long-span asymptotic distributions are critically dependent on the value of the AR coefficient, while the in-fill distribution has a uniform expression for any ρ_κ. Second, unlike the long-span asymptotics, the initial condition always shows up at the in-fill limit.

To know which limiting distributions produce the most accurate finite sample approximation, we follow Zhou and Yu (2015) and conduct a small Monte Carlo simulation.

Table 5.1 and 5.2 report the simulated percentiles of the finite sample distribution and the limiting distribution of $T(\tilde{\rho}_\kappa - \rho_\kappa)$ under the in-fill and long-span asymptotic schemes and model (5.2.19) with the following parameters: $N = 5, \Delta = 1/12$, $\mu = 0.1$, $X_0 = 5$. And we set $\kappa = -0.01$ and $\rho_\kappa = 1.0008$ in Table 1 and $\kappa = 0.01$ and $\rho_\kappa = 0.9992$ in Table 2. The number of replications is set to 2,500. This parameter setting allows us to compare the finite sample performance of the limiting distributions (5.2.7) or (5.2.9) with (5.2.22).

From Table 2, it can be seen that the percentiles of the finite sample distribution and the in-fill distribution of $T(\tilde{\rho}_\kappa - \rho_\kappa)$ do not vary much when we slightly change the value of ρ_κ. However, for the long-span distribution, the percentiles change drastically. Second, the percentiles of the in-fill distribution are very close to those of the finite sample distribution, while the long-span asymptotic distributions do not approximate the finite sample very well. Third, both the finite sample distribution and in-fill distribution are not symmetric. These simulation results echo the findings of Yu (2014), Zhou and Yu (2015) and Bao et al. (2017). They showed that in simulations, the in-fill distribution approximates the finite sample percentile the best among the two schemes. They attributed the superiority of the in-fill distribution to the inclusion of the initial condition via γ_0 for all values of ρ_κ, the continuity in ρ_κ of the in-fill distribution and the asymmetry of the in-fill distribution. All these three features can be seen in the finite sample distribution but are lost under the long-span asymptotics.

Dickey-Fuller t statistic $t_T(\tilde{\rho}) = \frac{\tilde{\rho}-1}{se(\tilde{\rho})}$ is often used to detect unit roots in time series. Laurent and Shi (2022) derived the in-fill distribution of the Dickey-Fuller t statistic $t_T(\tilde{\rho})$ under the model

(5.2.18) in Zhou and Yu (2015). They showed that the test statistic $t_T(\tilde{\rho})$ has the following the in-fill distribution:

Theorem 5.2.6 *Under the model (5.2.18) and (5.2.19) in Zhou and Yu (2015), let* $t_T(\tilde{\rho}) = \frac{\tilde{\rho}-1}{se(\tilde{\rho})}$, *as* $\Delta \to 0$, *we have*

$$t_T(\tilde{\rho}) \Rightarrow \begin{cases} \frac{\frac{1}{2}\left((W(1))^2-1\right)-W(1)\int_0^1 W(s)ds}{\left(\int_0^1 (W(s))^2 ds-\left(\int_0^1 W(s)ds\right)^2\right)^{1/2}} & \text{if } \rho_\kappa = 1, \text{ or equivalently } \kappa = 0 \\ \frac{-\Xi_2 W(1)-\Xi_1}{(\Xi_3-\Xi_2^2)^{1/2}} + c\left(\Xi_3 - \Xi_2^2\right)^{1/2} & \text{if } \rho_\kappa \neq 1, \text{ or equivalently } \kappa \neq 0 \end{cases} \tag{5.2.23}$$

where

$$\begin{aligned} \Xi_1 &\equiv \frac{b}{c}\int_0^1 (e^{cr}-1)dW(r) + \int_0^1 J_c(r)dW(r) + \gamma_0 \int_0^1 e^{cr}dW(r) \\ \Xi_2 &\equiv \frac{e^c - c - 1}{c^2} b + \int_0^1 J_c(r)dr + \gamma_0 \frac{e^c-1}{c}, \end{aligned}$$

$$\begin{aligned} \Xi_3 \equiv\ & \frac{e^{2c}-4e^c+2c+3}{2c^3}b^2 + \frac{2b}{c}\int_0^1 (e^{cr}-1)J_c(r)dr + \int_0^1 J_c^2(r)dr \\ & + \frac{e^{2c}-2e^c+1}{c^2} b\gamma_0 + 2\gamma_0 \int_0^1 e^{cr} J_c(r)dr + \gamma_0^2 \frac{e^{2c}-1}{2c}. \end{aligned}$$

The in-fill distribution (5.2.23) shows that imposing the condition $\rho_\kappa = 1$ removes the parameters c, μ and γ_0 in the limiting distribution. Therefore, (5.2.23) provides a feasible test under the unit root hypothesis. The authors show in simulation that the test has bad asymptotic power against the mean reversion alternative $\rho_\kappa < 1$ when the time span N is small and ρ_κ is close to 1 but generally has good asymptotic power against the explosive alternative $\rho_\kappa > 1$.

5.2.3 Double asymptotics

Aside from performing unit root tests, the in-fill theory can also be used to obtain the double asymptotic distribution when we first let $\Delta \to 0$ followed by $N \to \infty$. Wang and Yu (2016) derived the sequential asymptotics for the centered LS estimator under the model (5.2.15) with $\kappa < 0$[2]. Note that under this parameter setting, the model has an explosive AR parameter as $\rho_\kappa = \exp(-\kappa\Delta) > 1$. More precisely, they showed

$$\lim_{N\to\infty}\lim_{\Delta\to 0} \frac{e^{-\kappa N}(\hat{\rho}-\rho_\kappa)}{-2\kappa\Delta} = \lim_{c\to\infty} \frac{(2c)e^{-c}A(\gamma_0, c)}{(2c)^2 e^{-2c} B(\gamma_0, c)} = \frac{\xi}{\lambda_0 + \eta}, \tag{5.2.24}$$

where $\lambda_0 = Y_0\sqrt{-2\kappa}/\sigma$, ξ and η are two independent $N(0,1)$ random variables, $A(\gamma_0, c)$ and $B(\gamma_0, c)$ are defined in **(5.2.17).**

Similar to the in-fill distribution (5.2.17) and the long span distribution (5.2.9), the double asymptotic distribution preserves the initial condition via λ_0. Using the double asymptotic distribution, Wang and Yu (2016) formed a confidence interval for ρ_κ and confirmed the explosiveness in the monthly log Nasdaq real price data between 1990 and 2000.

5.2.4 Grid bootstrap confidence interval

The in-fill asymptotic theory is shown to approximate the finite sample distribution well but it does have an obvious drawback. As the localizing parameter c is not consistently estimable under the

[2]Wang and Yu (2016) also derived the double asymptotic distribution where $\Delta \to 0$ and $N \to \infty$ simultaneously.

in-fill scheme, this feature prevents us from directly inferring the AR coefficient based on the LTU distributions (5.2.11), (5.2.12), the in-fill distributions (5.2.16), (5.2.17) and (5.2.22).

Hansen (1999) proposed the grid bootstrap confidence interval (CI) method to perform a feasible inference for ρ under the discrete time LTU model. This method first defines a set of $G \equiv [\rho_L, \rho_R]$, generates bootstrap samples at each $\rho \in G$ under the hypothesis that ρ is the true value[3]and the coefficient-based statistic or test statistic is inverted to obtain an q-level grid bootstrap CI

$$CI_q = \{\rho \in R : c_T^*(q_1|\rho) \leq t_T(\hat{\rho}, \rho) \leq c_T^*(q_2|\rho)\} \tag{5.2.25}$$

where $t_T(\hat{\rho}, \rho)$ can be $T(\hat{\rho} - \rho)$ or $(\hat{\rho} - \rho)/se(\hat{\rho})$ and the qth quantile of the bootstrap distribution of $t_T(\hat{\rho}, \rho)$ is $c_T^*(q|\rho)$.

Hansen showed that the grid bootstrap CI in (5.2.25) produces a valid inference under the local-to-unit-root model by showing the coverage probability is asymptotically correct as $T \to \infty$,

$$\Pr\{\rho_0 \in CI_q\} \to q_2 - q_1, \tag{5.2.26}$$

where ρ_0 is the true value of ρ.

The key idea to obtain (5.2.26) is that, given the value of the localized parameter c,, the bootstrap data have the same limiting distribution as the original LTU model so that the analog of (5.2.10holds:

$$T^{-1/2}Y_{[Tr]}^* = \sum_{j=1}^{[Tr]} \exp\left(\frac{c}{T}\right)^{[Tr]-j} \frac{\varepsilon_j^*}{T^{1/2}} \Rightarrow J_c(r), \tag{5.2.27}$$

where $Y_t^* = \sum_{j=1}^{t} \exp\left(\frac{c}{T}\right)^{t-j} \varepsilon_j^*$, $\{\varepsilon_j^*\}_{j=1}^T$ is the usual LS nonparametric bootstrap residuals and $J_c(r)$ is defined in (5.2.10). Mikusheva (2007, 2015) later showed that Hansen's method is uniformly valid and provide a second order refinement to the LTU distribution.

Noting the similarity between the local to unit root model and the discrete time representation (5.2.15) under an in-fill asymptotic scheme, Lui et al. (2022) proposed constructing a grid bootstrap CI in the spirit of Hansen's method. Since the initial condition shows up at the in-fill limit, they modify the grid bootstrap procedure by letting the bootstrap initialization X_0^* to be equal to the initial condition X_0 in (5.2.18). They showed that the modified approach can produce a CI that is asymptotically and uniformly valid for any $\kappa \geq 0$ under the in-fill scheme. Moreover, they found that the modified grid bootstrap CI provides a second-order refinement to the in-fill distribution. Therefore, this method provides a valid CI under model (5.2.18). The authors used the CI to check the stationarity of the daily log realized volatility and the monthly nominal interest rate data under different maturities and concluded that log realized volatility is stationary while the stationarity of nominal interest rates are inconclusive.

5.3 Structural break models

Structural break models are widely used to study unit root problems and the parameter instability in models. Perron (1989, 1997) and Perron and Vogelsang (1992) studied the unit root test problem when a structural break is allowed in the model and the estimation of the break date. In empirical studies, Stock and Watson (1996) found evidence of non-constant parameters in commodity prices and interest rates. Bai (1997b) showed interest rate response to the change in the discount rate has a structural break. Paye and Timmermann (2006) showed that the instability of the parameters has implications for return prediction models. How to locate the break time and make statistical inferences about the break time is therefore crucial. Hence, it is well studied in the literature. Among all the possible estimations methods, the LS method of the break date was popularized by

[3]A fine grid for ρ containing 200 grid points is used by Hansen (1999).

Bai (1994, 1997a,b) and Chong (2001). These authors derived the long-span asymptotic distributions of the LS estimator and showed that they are pivotal and symmetric. Consequently, structural break tests based on LS can be easily applied.

Recent studies by Jiang et al. (2018, 2021) (hereafter JWYa and JWYb) showed that the finite sample distribution of the LS estimator is not symmetric and has three modes. As a result, they suggested that the inference based on the popularly used long-span distribution, which has a single mode and is symmetric, can potentially be misleading. However, the in-fill distribution of the LS estimator displays trimodality and asymmetry. Therefore, it offers a better finite sample approximation than that under the long-span scheme.

In this section, we review two structural break models. The first model allows for a structural break in the mean, while the second model allows for a break in the AR parameter. The organization of this section closely follows the previous one. We first introduce the discrete time models followed by the continuous time models and their discrete time representations. We then show that the in-fill asymptotic provides a more reliable approximation than the long-span asymptotic to the finite sample distribution of the LS estimator.

5.3.1 Structural break in mean

Hinkley (1970), Hawkins et al. (1986), Yao (1987) and Bai (1994, 1997a,b) studied time series models with a structural break and developed the asymptotic theory for the estimator of the break date under long-span scheme. Consider the structural break in the mean model (5.1.2),

$$\begin{aligned} X_t &= \begin{cases} \mu + u_t & \text{if } t \le k_0 \\ \mu + \delta + u_t & \text{if } t > k_0 \end{cases}, t = 1, ..., T\ , u_t \sim_{iid} (0, \sigma^2), \\ 1 &\le k_0 < T, \end{aligned} \tag{5.3.1}$$

we call k_0 and δ the structural break point and the break size, respectively.

Yao (1987) assumed that the error term u_t follows a Gaussian process and proposed to estimate k_0 by maximum likelihood estimation (MLE). We will focus on the LS method proposed by Bai (1994), as he showed that the LS method provided the same limit distribution as that obtained by MLE in Yao (1987). However, unlike Yao, Bai did not impose the Gaussian assumption on u_t.

Let $\tau_0 = k_0/T$ be the fractional break point, $\hat{k}_T^{LS}$ and $\hat{\tau}_T^{LS}$ be the LS estimators of k_0 and τ_0, respectively, such that

$$\begin{aligned} \hat{k}_T^{LS} &= \arg \min_{k=1,\dots,T-1} \{S_k^2\}, \\ S_k^2 &= \sum_{t=1}^{k} \left(X_t - \bar{X}_t\right)^2 + \sum_{t=k+1}^{T} \left(X_t - \tilde{X}_t\right)^2, \\ \hat{\tau}_T^{LS} &= \hat{k}_T^{LS}/T, \end{aligned} \tag{5.3.2}$$

where $\bar{X}_t$ and $\tilde{X}_t$ are the sample averages of the first k observations and last $T-k$ observations, respectively.

Bai (1994) assumed $\delta \to 0$ as $T \to \infty$, with the condition $\frac{\sqrt{T}\delta}{\sqrt{\log T}} \to \infty$, and showed that the limit of $\hat{\tau}_T^{LS}$ does not depend on the distributional function of u_t[4] and has the following asymptotic distribution.

Theorem 5.3.1 *Under model (5.3.1) with the conditions $\delta \to 0$ as $T \to \infty$ and $\frac{\sqrt{T}\delta}{\sqrt{\log T}} \to \infty$,*

$$T\frac{\delta_T^2}{\sigma^2}\left(\hat{\tau}_T^{LS} - \tau_0\right) \Rightarrow \arg \max_{u \in (-\infty,\infty)} \left\{ W_2(u) - \frac{1}{2}|u| \right\}, \tag{5.3.3}$$

[4]In contrast, if δ is fixed, Hinkley (1970) showed that the limit distribution of $\hat{\tau}_T^{LS}$ depends on the distributional function of u_t and hence there is no invariance principle.

where $W_2(u)$ is a two-sided BM such that

$$W_2(u) = \begin{cases} W_a(-u) = W(\tau_0) - W(\tau_0 - (-u)) & \text{if } u \leq 0 \\ W_b(-u) = W(\tau_0) - W(\tau_0 + u) & \text{if } u > 0 \end{cases},$$

$W_a(-u)$ and $W_b(-u)$ are two independent BMs.

Theorem 5.3.1 suggests the long-span distribution of $\hat{\tau}_T^{LS} - \tau_0$ is symmetric. JWYa argued that the finite sample distribution of $\hat{\tau}_T^{LS} - \tau_0$ is not symmetric when $\tau_0 \neq 0.5$ as the number of observations before and after the break are unequal. To see the finite sample performance of the long-span limiting distribution (5.3.3), we follow JWYa and conduct a Monte Carlo simulation with $\delta = 0.3$, $u_t \sim_{iid} N(0,1)$, $\tau_0 = 0.75$, $T = 100$. The number of replications is set to $10,000$.

Figure 5.1: The density of $\arg\max_{u\in(-\infty,\infty)} \left\{W_2(u) - \frac{1}{2}|u|\right\}$ and the finite sample distribution of $T\frac{\delta_T^2}{\sigma^2}\left(\hat{\tau}_T^{LS} - \tau_0\right)$.

Figure 5.1 shows the simulated density of the finite sample distribution of $T\frac{\delta_T^2}{\sigma^2}\left(\hat{\tau}_T^{LS} - \tau_0\right)$ and long-span limiting distribution (5.3.3) in blue and red lines, respectively. From Figure 5.1, it can be seen that the long-span distribution (5.3.3) is symmetric, has a single mode centered at 0 and suggests that the estimator is unbiased. On the other hand, the finite sample distribution is asymmetric, has three modes with the centered mode not around zero when $\tau_0 \neq 0.5$ and suggests that the estimator is seriously biased. Therefore, the long-span distribution does not provide a good finite sample approximation. This unsatisfactory performance of the long-span theory motivated JWYa to seek a limit distribution under the in-fill asymptotic scheme, which can provide a better finite sample approximation. To obtain an in-fill distribution, as in the AR(1) models, we need a continuous time model and its discrete time representation, which are closely related to model (5.3.1).

Consider the following continuous time model:

$$dX(t) = \left(\mu + \left(\frac{\delta}{\varepsilon}\right) 1(t > \tau_0)\right) dt + \sigma dW(t) \tag{5.3.4}$$

where $t \in [0,1]$, $1(t > \tau_0)$ is an indicator function, $\sigma > 0$ is the spot volatility, μ, δ, ε and τ_0 are constants. Model (5.3.4) has a structural break at τ_0 for the drift function with the break size δ/ε.

We now introduce the discrete time representation that is used to elicit the in-fill asymptotics. Let $\{X_{t\Delta}\}_{t=1}^{T}$ be the discrete time observations, JWYa showed that the continuous time model (5.3.4) has the following exact discrete time representation:

$$\begin{aligned} X_{t\Delta} - X_{(t-1)\Delta} &= \begin{cases} \mu\Delta + \sqrt{\Delta}\varepsilon_t & \text{if } t = 1, ..., k_0, \\ [\mu + \delta/\varepsilon]\,\Delta + \sqrt{\Delta}\varepsilon_t & \text{if } t = k_0 + 1, ..., T, \end{cases} \\ \varepsilon_t &\sim {}_{iid}N(0, \sigma^2). \end{aligned} \tag{5.3.5}$$

Let $Z_{t\Delta} = \left(X_{t\Delta} - X_{(t-1)\Delta}\right)/\sqrt{\Delta}$, Equation (5.3.5) becomes

$$Z_{t\Delta} = \begin{cases} \mu\sqrt{\Delta} + \varepsilon_t & \text{if } t = 1, ..., k_0, \\ [\mu + \delta/\varepsilon]\sqrt{\Delta} + \varepsilon_t & \text{if } t = k_0 + 1, ..., T. \end{cases} \tag{5.3.6}$$

When Δ is fixed, the discrete time model (5.3.6) is the same as model (5.3.1) with the break size equal to $\frac{\delta}{\varepsilon}\sqrt{\Delta}$.

Theorem 5.3.2 *(a) Under model (5.3.5), the LS estimator in (5.3.2) has the following limit distribution when $\Delta \to 0$,*

$$T\left(\frac{\delta}{\sigma\varepsilon}\sqrt{\Delta}\right)^2 \left(\hat{\tau}_T^{LS} - \tau_0\right) \Rightarrow \left(\frac{\delta}{\sigma\varepsilon}\right)^2 \arg\max_{u\in(-\tau_0,(1-\tau_0))} \left\{\tilde{W}(u)\right\}^2, \tag{5.3.7}$$

where

$$\tilde{W}(u) = \begin{cases} \frac{W(\tau_0+u)-(\tau_0+u)W(1)}{\sqrt{(\tau_0+u)(1-\tau_0-u)}} - \frac{(1-\tau_0)\sqrt{\tau_0+u}}{\sqrt{1-\tau_0-u}}\frac{\delta}{\sigma\varepsilon} & \textit{if } u \le 0 \\ \frac{W(\tau_0+u)-(\tau_0+u)W(1)}{\sqrt{(\tau_0+u)(1-\tau_0-u)}} - \frac{\tau_0\sqrt{1-\tau_0-u}}{\sqrt{\tau_0+u}}\frac{\delta}{\sigma\varepsilon} & \textit{if } u > 0 \end{cases}.$$

(b) Furthermore, when $\Delta \to 0$, $\varepsilon \to 0$ with $\frac{\delta}{\varepsilon}\sqrt{\Delta} \to 0$ simultaneously,

$$T\left(\frac{\delta}{\sigma\varepsilon}\sqrt{\Delta}\right)^2 \left(\hat{\tau}_T^{LS} - \tau_0\right) \Rightarrow \arg\max_{u\in(-\infty,+\infty)} \left\{W_2(u) - \frac{1}{2}|u|\right\}. \tag{5.3.8}$$

Theorem 5.3.2, which gives the in-fill distribution (5.3.7) of the LS estimator was proven in JWYa. It shows that the in-fill distribution shares two important features. It has three modes and is asymmetric. The similarity of the in-fill distribution and finite sample distribution enables the in-fill asymptotic distribution to outperform the long-span distribution (5.3.3) in approximating the finite sample distribution of $\hat{\tau}_T^{LS}$. Note that the interval to locate the argmax is $(-\infty, \infty)$ in the long-span distribution, while it is $u \in (-\tau_0, 1 - \tau_0)$ under the in-fill distribution. This difference, as discussed in detail in JWYa, gives the in-fill distribution the additional two modes at the left and right ends and yields the trimodality. The result (5.3.8) shows that the small ε distribution coincides with the distribution (5.3.3) obtained by Yao (1987) and Bai (1994), respectively.

Using the same data generating process as in Figure 5.1, we report the simulated density of the distribution (5.3.7).

Figure 5.2 complements Figure 5.1 by adding the simulated density of the in-fill distribution (5.3.7) in the green dotted line. The in-fill distribution has three modes and is asymmetric. It suggests that the LS estimator is biased at the in-fill limit. This distribution therefore provides a significantly better finite sample approximation than the long-span distribution.

5.3.2 Structural break in autoregressions

Consider the following AR(1) model:

$$X_t = \begin{cases} \theta_1 X_{t-1} + u_t & \text{if } t \le k_0 \\ \theta_2 X_{t-1} + u_t & \text{if } t > k_0 \end{cases}, t = 1, ..., T\,,\ u_t \sim_{iid} (0, \sigma^2). \tag{5.3.9}$$

Figure 5.2: The density of $\left(\frac{\delta}{\sigma\varepsilon}\right)^2 \arg\max_{u\in(-\tau_0,(1-\tau_0))}\left\{\tilde{W}(u)\right\}^2$, $\arg\max_{u\in(-\infty,\infty)}\left\{W_2(u)-\frac{1}{2}|u|\right\}$ and the finite sample distribution of $T\frac{\delta_T^2}{\sigma^2}\left(\hat{\tau}_T^{LS}-\tau_0\right)$.

Assuming $\theta_1 \neq \theta_2$, model (5.3.9) is a model with a structural break in the AR(1) parameter at k_0 with break size $|\theta_2-\theta_1|$.

Let the LS estimator for k_0 be

$$\begin{aligned}
\hat{k}_T^{LS} &= \arg\min_{k=1,\ldots,T-1}\{S_k^2\}, \\
S_k^2 &= \sum_{t=1}^{k}\left(X_t-\hat{\theta}_1(k)X_{t-1}\right)^2+\sum_{t=k+1}^{T}\left(X_t-\hat{\theta}_2(k)X_{t-1}\right)^2, \qquad (5.3.10)\\
\hat{\theta}_1(k) &= \frac{\sum_{t=1}^{k}X_tX_{t-1}}{\sum_{t=1}^{k}X_{t-1}^2}, \hat{\theta}_2(k)=\frac{\sum_{t=k+1}^{T}X_tX_{t-1}}{\sum_{t=k+1}^{T}X_{t-1}^2},
\end{aligned}$$

and the LS estimator of the fractional break point is $\hat{\tau}_T^{LS}=\hat{k}_T^{LS}/T$.

Chong (2001), Pang et al. (2014) and Liang et al. (2018) derived the long-span asymptotic distributions of the LS estimator $\hat{\tau}_T^{LS}$ under various assumptions of θ_1,θ_2, break size and the initialization X_0. Their results showed that the long-span distributions are critically dependent on these assumptions. We summarize their results in the following theorem.

Theorem 5.3.3 *Under model (5.3.9), $\hat{\tau}_T^{LS}$ has the following limits under different sets of assumptions.*

(a) Suppose that $|\theta_1|<1, |\theta_2|<1$, $\theta_2=\theta_{2,T}$, $\theta_{2,T}-\theta_1\to 0$, $\sqrt{T}|\theta_{2,T}-\theta_1|\to\infty$ and $X_0=O_p(1)$, as $T\to\infty$,

$$\frac{T(\theta_{2,T}-\theta_1)^2}{1-\theta_1^2}\left(\hat{\tau}_T^{LS}-\tau_0\right)\Rightarrow \arg\max_{u\in(-\infty,+\infty)}\left\{W_2(u)-\frac{1}{2}|u|\right\}. \qquad (5.3.11)$$

(b) Suppose that $|\theta_1|<1, \theta_2=1$, $\theta_1=\theta_{1,T}$, $1-\theta_{1,T}\to 0$ with $T(1-\theta_{1,T})\to\infty$ and $X_0=O_p(1)$, as $T\to\infty$,

$$T(1-\theta_{1,T})(\hat{\tau}_T^{LS}-\tau_0)\Rightarrow \arg\max_{u\in(-\infty,+\infty)}\left\{\frac{W_a^*(u)}{R_1}-\frac{1}{2}|u|\right\} \qquad (5.3.12)$$

where

$$W_a^*(u) = \begin{cases} W_a(-u) & \text{if } u \le 0 \\ -W_b(u) - \int_0^u \frac{W_b(s)}{R_1} dW_b(s) - \int_0^u \left(\frac{W_b(s)}{2R_1} + 1\right) W_b(s) ds & \text{if } u > 0 \end{cases},$$

$W_a(u)$ and $W_b(u)$ are two independent BMs, and $R_1 = \int_0^\infty \exp(-s) dW(s)$.

(c) Suppose that $\theta_1 = 1, |\theta_2| < 1, \theta_2 = \theta_{2,T}, \sqrt{T}(1-\theta_{2,T}) \to \infty, T^{3/4}(1-\theta_{2,T}) \to \infty$ and $X_0 = O_p(1)$, as $T \to \infty$,

$$T^2(\theta_{2,T} - 1)^2 \left(\hat{\tau}_T^{LS} - \tau_0\right) \Rightarrow \arg \max_{u \in (-\infty, +\infty)} \left\{ \frac{W_2(u)}{W_c(u)} - \frac{1}{2}|u| \right\}, \tag{5.3.13}$$

where $W_c(u)$ is an independent BM.

(d) Suppose that $\theta_1 = \theta_{1,T}, |\theta_{1,T}| < 1, \theta_2 = \theta_{2,T} = 1 + c/T, |\theta_{2,T} - \theta_1| \to 0, T(\theta_{2,T} - \theta_1) \to \infty$ and $X_0 = o_p(\sqrt{T})$, as $T \to \infty$,

$$T(\theta_{2,T} - \theta_1) \left(\hat{\tau}_T^{LS} - \tau_0\right) \Rightarrow \arg \max_{u \in (-\infty, +\infty)} \left\{ \frac{W_d^*(u)}{R_1} - \frac{1}{2}|u| \right\},$$

where

$$W_d^*(u) = \begin{cases} W_a(-u) & \text{if } u \le 0 \\ -I(W_b, c, \tau_0, u) - \int_0^u \frac{I(W_b, c, \tau_0, u)}{R_1} dI(W_b, c, \tau_0, u) & \\ - \int_0^u \left(\frac{I(W_b, c, \tau_0, s)}{2R_1} + 1\right) I(W_b, c, \tau_0, s) ds & \text{if } u > 0 \end{cases}, \tag{5.3.14}$$

and

$$I(W_b, c, \tau_0, u) = W_b(\tau_0 + u) - W_b(\tau_0) - c \int_{\tau_0}^{\tau_0 + u} e^{-c(\tau_0 + u - r)} (W_b(u) - W_b(\tau_0)) du.$$

(e) Suppose $\theta_1 = \theta_{1,T} = 1 + c/T, \theta_2 = \theta_{2,T}, |\theta_{2,T}| < 1, \sqrt{T}(\theta_{2,T} - \theta_{1,T}) \to 0, T^{3/4}(\theta_{2,T} - \theta_{1,T}) \to \infty$ and $X_0 = o_p(\sqrt{T})$, as $T \to \infty$,

$$T^2(\theta_{2,T} - \theta_{1,T})^2 \left(\hat{\tau}_T^{LS} - \tau_0\right) \Rightarrow \arg \max_{u \in (-\infty, +\infty)} \left\{ \frac{W_2(u)}{\exp(c(1-\tau_0)) G(W_a, c, \tau_0)} - \frac{1}{2}|u| \right\}, \tag{5.3.15}$$

where $G(W_a, c, \tau_0) = \exp(-c(1-\tau_0)) W_a(\tau_0) - c \int_0^{\tau_0} \exp(-c(1-s)) W_a(s) ds$.

(f) Suppose $\theta_1 = \theta_{1,T} = 1 - c/k_T, c > 0, \theta_2 = 1, k_T \to \infty, k_T/T \to 0$ and $X_0 = o_p(\sqrt{k_T})$, as $T \to \infty$,

$$\frac{cT}{k_T} \left(\hat{\tau}_T^{LS} - \tau_0\right) \Rightarrow \arg \max_{u \in (-\infty, +\infty)} \left\{ \frac{W_e^*(u)}{R_c} - \frac{1}{2}|u| \right\}, \tag{5.3.16}$$

where

$$W_e^*(u) = \begin{cases} W_a(-u) & \text{if } u \le 0 \\ -W_b(-u) - \int_0^u \frac{W_b(s)}{R_c} dW_b(s) - \int_0^u \left(\frac{W_b(s)}{2R_c} + 1\right) W_b(s) ds & \text{if } u > 0 \end{cases},$$

and $R_c = \sqrt{c} \int_0^\infty \exp(-cs) dW_a(s)$.

(g) Suppose $\theta_1 = 1, \theta_2 = \theta_{2,T} = 1 - c/k_T, k_T \to \infty, k_T/T \to 0, T^{1/2}/k_T \to 0, T^{3/4}/k_T \to \infty$ and $X_0 = o_p(\sqrt{k_T})$, as $T \to \infty$,

$$\frac{c^2 T^2}{k_T^2} \left(\hat{\tau}_T^{LS} - \tau_0\right) \Rightarrow \arg \max_{u \in (-\infty, +\infty)} \left\{ \frac{W_2(u)}{W_c(\tau_0)} - \frac{1}{2}|u| \right\}. \tag{5.3.17}$$

Theorem 5.3.3 shows that when the AR parameters before and after the break are strictly stationary, the limit has the form (5.3.11). When one of the AR parameters is stationary, but the other is of unit root, (5.3.12) and (5.3.13) become applicable. When one of the AR parameters is local to unity, but the other is stationary, (5.3.14) and (5.3.15) apply. When one of the AR parameters is of unit root, the other is in the further vicinity of unity, (5.3.16) and (5.3.17) are valid. It can be clearly seen that there is no uniform expression for the long-span asymptotic distribution since the limiting distribution varies across different assumptions regarding AR parameters, break size and initial condition.

The results show several weaknesses of long-span asymptotics. First, since the long-span distribution has a discontinuity in the values θ_1 and θ_2, and the AR parameters are not known a priori, it is unclear which limit distribution should be used in practice. Second, the long-span distribution does not approximate the finite sample distribution well. Under the case where limit (5.3.11) is applicable, the long-span distribution is symmetric and has only one mode, while trimodality and asymmetry can be observed in a finite sample distribution under general parameter settings. Thirdly, as in the AR(1) model in Section 5.2, the finite sample distribution also depends on the initial condition X_0, while X_0 is missing in the limit distributions in various settings of Theorem 5.3.3. Finally, when θ_1 and θ_2 are both less than 1, for example, $\theta_1 = 0.85$ and $\theta_2 = 0.90$. It is not clear whether θ_2 is local to unity or not. Consequently, which result ((a), (d) or (g)) under Theorem 5.3.3 should be applied becomes debatable.

JWYb performed extensive simulations to showcase the significant differences between the long-span distributions (5.3.11) to (5.3.17) and the finite sample distributions. They found that long-span distributions (5.3.11) to (5.3.17) only have a single mode, unlike that under finite sample which displays trimodality. These significant differences motivate them to study the in-fill distribution of $\hat{\tau}_T^{LS}$.

To elicit the in-fill asymptotic, we now introduce a continuous time model. Consider the following model:

$$\begin{aligned} dX(t) &= -(\kappa + \delta 1(t > \tau_0))X(t)dt + \sigma dW(t), t \in [0,1], \\ X(0) &= X_0 = O_p(1). \end{aligned} \tag{5.3.18}$$

JWYb showed the exact discrete time representation of model (5.3.18) with sampling interval Δ has the following form:

$$\begin{aligned} X_{t\Delta} &= (\theta_1 1(t \le k_0) + \theta_2 1(t > k_0)) X_{(t-1)\Delta} + \sqrt{\frac{1 - e^{-2(\kappa+\delta 1(t\Delta > \tau_0))\Delta}}{2(\kappa + \delta 1(t\Delta > \tau_0)}}\varepsilon_t, t = 1, ..., T, \\ \varepsilon_t &\sim i.i.d. N(0, \sigma^2), \theta_1 = \exp(-\kappa/T), \theta_2 = \exp(-(\kappa+\delta)/T), \\ \Delta &= 1/T, \\ X_0 &= O_p(1), \end{aligned} \tag{5.3.19}$$

and the fractional break date is $\tau_0 = k_0/T$.

Clearly, the continuous time model (5.3.18) with the exact discrete time representation (5.3.19) generalizes (5.2.13) and (5.2.14) with time span $N = 1$, and both models (5.2.14) and (5.3.19) have local to unit root coefficients as $\Delta \to 0$ with a variance of the error term that is asymptotically equal to Δ.[5] To generalize the Gaussian assumption of the error term in (5.3.19), JWYb studied the following discrete time model:

$$\begin{aligned} X_t &= (\theta_1 1(t \le k_0) + \theta_2 1(t > k_0)) X_{t-1} + \sqrt{\Delta}\varepsilon_t, t = 1, ..., T, \\ \varepsilon_t &\sim i.i.d.(0, \sigma^2), \end{aligned} \tag{5.3.20}$$

where $\theta_1, \theta_2, \Delta$ and X_0 are the same as those in (5.3.19).

[5] $\frac{1-e^{-2\kappa\Delta}}{2\kappa} = \Delta + O(\Delta^2)$ and $\frac{1-e^{-2(\kappa+\delta)\Delta}}{2(\kappa+\delta)} = \Delta + O(\Delta^2)$ as $\Delta \to 0$.

If we let $Y_t = X_t/\sqrt{\Delta}$,, we can write

$$\begin{aligned} Y_t &= (\theta_1 1(t \le k_0) + \theta_2 1(t > k_0)) Y_{t-1} + \varepsilon_t, t = 1, ..., T, \\ Y_0 &= X_0/\sqrt{\Delta} = O_p(T^{1/2}). \end{aligned} \tag{5.3.21}$$

Clearly, when Δ is fixed, model (5.3.21) is the same as model (5.3.9). While under the in-fill asymptotic, the initial condition Y_0 diverges with rates $T^{1/2}$ and θ_1 and θ_2 are both local to unit root AR coefficients in model (5.3.21). Consequently, these two features allow us to obtain the in-fill distribution, which has a uniform expression for different values of κ and an initial condition as $\Delta \to 0$. We list the in-fill asymptotic distribution of $\hat{\tau}_T^{LS}$ in the following theorem.

Theorem 5.3.4 *Suppose the LS estimator is $\hat{\tau}_T^{LS} = \hat{k}_T^{LS}/T$, with $\hat{k}_T^{LS}$ being defined in (5.3.10). Under model (5.3.20), as $\Delta \to 0$, we have*

$$\hat{\tau}_T^{LS} \Rightarrow \arg\max_{\tau \in (0,1)} \left\{ \frac{\left\{ \left(\tilde{J}_{\tau_0}(\tau)\right)^2 - \left(\tilde{J}_{\tau_0}(0)\right)^2 - \tau \right\}^2}{\int_0^\tau \left(\tilde{J}_{\tau_0}(r)\right)^2 dr} + \frac{\left\{ \left(\tilde{J}_{\tau_0}(1)\right)^2 - \left(\tilde{J}_{\tau_0}(\tau)\right)^2 - (1-\tau) \right\}^2}{\int_\tau^1 \left(\tilde{J}_{\tau_0}(r)\right)^2 dr} \right\} \tag{5.3.22}$$

where $\tilde{J}_{\tau_0}(r), r \in (0,1)$ is defined by

$$d\tilde{J}_{\tau_0}(r) = -(\kappa + \delta 1(t > \tau_0))\tilde{J}_{\tau_0}(r)dr + dW(r),$$

with $\tilde{J}_{\tau_0}(0) = X(0)/\sigma$.

Let $\tau = \tau_0 + u$,. We can also obtain an alternative expression for the centered estimator $\hat{\tau}^{ML} - \tau_0$:

$$\hat{\tau}^{ML} - \tau_0 \stackrel{d}{=} \arg\max_{u \in (-\tau_0, 1-\tau_0)} \begin{cases} \delta \left\{ \left(\tilde{J}_{\tau_0}(\tau_0 + u)\right)^2 - u - (2\kappa + \delta) \int_{\tau_0+u}^{\tau_0} \left(\tilde{J}_{\tau_0}(r)\right)^2 dr \right\} & \text{if } u \le 0 \\ \delta \left\{ \left(\tilde{J}_{\tau_0}(\tau_0 + u)\right)^2 - u + (2\kappa + \delta) \int_{\tau_0}^{\tau_0+u} \left(\tilde{J}_{\tau_0}(r)\right)^2 dr \right\} & \text{if } u > 0 \end{cases}. \tag{5.3.23}$$

Theorem 5.3.4 was proven in JWYb. Several observations can be made based on the theorem. First, from the expression (5.3.23), the asymmetry is from the interval to locate the argmax $(-\tau_0, 1-\tau_0)$ and $\tilde{J}_{\tau_0}(\tau)$. Second, the initial condition X_0 enters the in-fill limit via $\tilde{J}_{\tau_0}(\tau)$. Finally, for all values of κ and δ, the limit (5.3.22) is valid. From the simulation, JWYb also showed that the in-fill distribution (5.3.22) displays trimodality.

From Theorem 5.3.4, we can see that the in-fill theory has two major advantages over the long-span theory. First, it is a unified theory where the in-fill distributional form (5.3.22) is valid regardless of the sign and values of κ and δ. This is in sharp contrast to the long-span theory, where the values of θ_1 and θ_2 determine the form of the limiting distribution. These findings indeed echo the results in the AR models, where the in-fill distributions are continuous in the localizing parameter c. These features allow the in-fill distribution to provide a better approximation to the finite sample distribution than those of the long-span distribution. Second, the in-fill theory provides distributions with trimodality in both structural break in the mean and AR coefficient models. It thus provides a better finite sample approximation.

Remark 5.3.5 *The trimodality in the in-fill distribution makes the use of confidence intervals invalid. Therefore, JWYb suggested the use of the highest density region (HDR) to construct an inference for τ.*

Remark 5.3.6 *To avoid the trimodality in the in-fill distribution, Baek (2018) proposed a modified LS estimator, which removes the modes at the two ends of the unit interval. She showed that if the LS objective function is premultiplied by an inverse U-shaped weighting function, one can obtain a modified LS estimator, and the in-fill distribution of the modified estimator has a single mode.*

5.4 Conclusion

This chapter reviews the discrete time AR(1) models and structural breaks models and some applications of in-fill asymptotic distributions. For the AR models, we have reviewed the asymptotic results in Baek (2018), Perron (1991), Zhou and Yu (2015), and Wang and Yu (2015). These results gave the limit distributions of LS estimators under the long-span asymptotics. We showed that these long-span distributions, in general, do not approximate the finite sample distribution well. The finite sample distribution, in simulation, is shown to be continuous in the autoregressive parameter, has a dependence on the initial condition and is asymmetric. These three important features are generally missing in the long-span distributions which results in the poor finite sample performance of long-span distributions.

Phillips (1987b), Perron (1991) and Zhou and Yu (2015) showed that continuous time models and their discrete time representations, which are analogous to the original discrete time AR(1) models can be used to obtain the in-fill asymptotics where the sampling interval is assumed to shrink to zero. The in-fill distributions provide a much better finite sample approximation as it is continuous in the AR parameter and the initial condition is preserved under the in-fill limit. In-fill theory can also be applied to unit root tests as shown by Laurent and Shi (2022). Wang and Yu (2016) showed that in-fill asymptotics provide an intermediate tool to obtain the sequential double asymptotics that first let the sampling interval shrinks to zero, followed by an increasing time span.

The in-fill distribution generally provides a good finite sample Approximation. However, it is not directly implementable due to an inconsistent estimable localizing parameter that appears in the limit. To obtain an inference of the autoregressive parameter, Lui et al. (2022) proposed a grid bootstrap method, which is a modified version of Hansen (1999), to obtain a feasible and asymptotically correct confidence interval under the in-fill asymptotic scheme.

Aside from AR(1) models, two structural break models are reviewed in this chapter. The first structural break model we considered has a break in mean. Bai (1994) showed the long-span distribution of the LS estimator of the break date is symmetric and has one mode. However, the finite sample distribution can be significantly different, as it can be asymmetric and has three modes, as shown by JWYa. The poor finite sample performance of long-span theory motivated JWYa to use the continuous time model with the discrete time representation to develop an in-fill theory. JWYa showed that in-fill asymptotic distribution of the LS estimator well approximates the finite sample distribution, and the distribution also displays the trimodality.

The second structural break model has a break in the AR parameter. Similar to the finding of JWYa, JWYb showed that the long-span distribution cannot well approximate the finite sample. In contrast, they showed the in-fill distribution of the LS estimator of the break date has trimodality, incorporates the initial condition, is asymmetric and valid across all values of the AR coefficients. These features of the in-fill distribution are similar to those found in the finite sample distribution of the LS estimator. Therefore, the in-fill distribution outperforms the long-span distribution in approximating the finite sample. Due to the nonstandard in-fill distribution with three modes, JWYb suggested using the highest density region (HDR) to construct an inference for the break date.

Bibliography

Andrews, D. W. K. (1993). Exactly median-unbiased estimation of first order autoregressive/unit root models. *Econometrica 61*(1), 139–165.

Andrews, D. W. K. and H.-Y. Chen (1994). Approximately median-unbiased estimation of autoregressive models. *Journal of Business & Economic Statistics 12*(2), 187–204.

Baek, Y. (2018). Estimation of a structural break point in linear regression models. *arXiv preprint arXiv:1811.03720*.

Bai, J. (1994). Least squares estimation of a shift in linear processes. *Journal of Time Series Analysis 15*(5), 453–472.

Bai, J. (1997a). Estimating multiple breaks one at a time. *Econometric theory 13*(3), 315–352.

Bai, J. (1997b). Estimation of a change point in multiple regression models. *Review of Economics and Statistics 79*(4), 551–563.

Bao, Y., A. Ullah, and Y. Wang (2017). Distribution of the mean reversion estimator in the ornstein–uhlenbeck process. *Econometric Reviews 36*(6-9), 1039–1056.

Chan, N. H. and C. Z. Wei (1987). Asymptotic inference for nearly nonstationary ar(1) processes. *The Annals of Statistics 15*(3), 1050–1063.

Chong, T. T.-L. (2001). Structural change in ar (1) models. *Econometric Theory 17*(1), 87–155.

Davidson, J. (1994). *Stochastic limit theory: An introduction for econometricians*. OUP Oxford.

Diba, B. T. and H. I. Grossman (1988). Explosive rational bubbles in stock prices? *The American Economic Review 78*(3), 520–530.

Evans, G. W. (1991). Pitfalls in testing for explosive bubbles in asset prices. *The American Economic Review 81*(4), 922–930.

Hansen, B. E. (1999). The grid bootstrap and the autoregressive model. *Review of Economics and Statistics 81*(4), 594–607.

Hawkins, D., A. Gallant, and W. Fuller (1986). A simple least squares method for estimating a change in mean. *Communications in Statistics - Simulation and Computation 15*(3), 523–530.

Hinkley, D. V. (1970). Inference about the change-point in a sequence of random variables. *Biometrika 57*(1), 1–17.

Jiang, L., X. Wang, and J. Yu (2018). New distribution theory for the estimation of structural break point in mean. *Journal of Econometrics 205*(1), 156–176.

Jiang, L., X. Wang, and J. Yu (2021). In-fill asymptotic theory for structural break point in autoregressions. *Econometric Reviews 40*(4), 359–386.

Kwiatkowski, D., P. C. Phillips, P. Schmidt, and Y. Shin (1992). Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? *Journal of Econometrics 54*(1), 159–178.

Laurent, S. and S. Shi (2022). Unit root test with high-frequency data. *Econometric Theory 38*(1), 113–171.

Liang, Y., T. Pang, D. Zhang, and T. Chong (2018). Structural change in mildly integrated ar (1) models. *Econometric Theory 34*(5), 985–1017.

Lui, Y. L., W. Xiao, and J. Yu (2022). The grid bootstrap for continuous time models. *Journal of Business & Economic Statistics 40*(3), 1390–1402.

Mikusheva, A. (2007). Uniform inference in autoregressive models. *Econometrica 75*(5), 1411–1452.

Mikusheva, A. (2015). Second order expansion of the t-statistic in ar (1) models. *Econometric Theory 31*(3), 426–448.

Pang, T., D. Zhang, and T. T.-L. Chong (2014). Asymptotic inferences for an ar (1) model with a change point: stationary and nearly non-stationary cases. *Journal of Time Series Analysis 35*(2), 133–150.

Paye, B. S. and A. Timmermann (2006). Instability of return prediction models. *Journal of Empirical Finance 13*(3), 274–315.

Perron, P. (1989). The great crash, the oil price shock, and the unit root hypothesis. *Econometrica: Journal of the Econometric Society 57*(6), 1361–1401.

Perron, P. (1991). A continuous time approximation to the unstable first-order autoregressive process: the case without an intercept. *Econometrica: Journal of the Econometric Society 59*(1), 211–236.

Perron, P. (1997). Further evidence on breaking trend functions in macroeconomic variables. *Journal of Econometrics 80*(2), 355–385.

Perron, P. and T. J. Vogelsang (1992). Nonstationarity and level shifts with an application to purchasing power parity. *Journal of Business & Economic Statistics 10*(3), 301–320.

Phillips, P. C. B. (1987a). Time series regression with a unit root. *Econometrica: Journal of the Econometric Society 55*(2), 277–301.

Phillips, P. C. B. (1987b). Towards a unified asymptotic theory for autoregression. *Biometrika 74*(3), 535–547.

Phillips, P. C. B. and P. Perron (1988). Testing for a unit root in time series regression. *Biometrika 75*(2), 335–346.

Phillips, P. C. B., S. Shi, and J. Yu (2015). Testing for multiple bubbles: Historical episodes of exuberance and collapse in the s&p 500. *International Economic Review 56*(4), 1043–1078.

Phillips, P. C. B. and J. Yu (2011). Warning signs of future asset bubbles.

Stock, J. H. and M. W. Watson (1996). Evidence on structural instability in macroeconomic time series relations. *Journal of Business & Economic Statistics 14*(1), 11–30.

Wang, X. and J. Yu (2015). Limit theory for an explosive autoregressive process. *Economics Letters 126*, 176–180.

Wang, X. and J. Yu (2016). Double asymptotics for explosive continuous time models. *Journal of Econometrics 193*(1), 35–53.

White, J. S. (1958). The Limiting Distribution of the Serial Correlation Coefficient in the Explosive Case. *The Annals of Mathematical Statistics 29*(4), 1188 – 1197.

Yao, Y.-C. (1987). Approximating the distribution of the maximum likelihood estimate of the change-point in a sequence of independent random variables. *The Annals of Statistics 15*(3), 1321–1328.

Yu, J. (2014). Econometric analysis of continuous time models: A survey of peter phillips's work and some new results. *Econometric Theory 30*(4), 737–774.

Zhou, Q. and J. Yu (2015). Asymptotic theory for linear diffusions under alternative sampling schemes. *Economics Letters 128*, 1–5.