Econometric Analysis of Nonstationary Continuous-Time Models

YE CHEN *Capital University of Economics and Business*

This chapter discusses the nonstationary continuous-time models, including unit root and explosive regressors. The contents cover estimation methods, inferential theory, and empirical examples demonstrating the use of these models. It starts with a univariate framework and extends to multivariate cases for generality.

6.1 Introduction

A covariance stationary process is a stochastic process whose mean, variance, and covariance do not change when shifted in time. In contrast, a nonstationary process has a time-varying mean, variance, or covariance. It may exhibit a trend, a structural break, cycles, or some combination thereof. While stationarity makes the econometric analysis of time series simpler, many economic and financial time series are nonstationary.

In finance, stock prices, earning-price ratios, and dividend-price ratios are commonly believed to be integrated of order one (Alessi and Kerssenfischer, 2019; Stambaugh, 1999; Welch and Goyal, 2008; Cai et al., 2015; Phillips, 2015). The time series properties of those data series are fundamental to the efficient market hypothesis (Lo and MacKinlay (1988)) and have been of significant interest. In the housing market, in addition to housing prices, some fundamentals such as price-to-income ratios and rent-to-price ratios are generally found to be nonstationary (Case and Shiller, 2003; McCarthy et al., 2004; Ambrose et al., 2013). In particular, house prices follow an explosive process in the presence of speculative bubbles. Evidence of bubble presence has been documented for many housing markets; see, for example, Greenaway-McGrevy and Phillips (2016); Shi et al. (2016); Hu and Oxley (2018); Pan (2019); Chen et al. (2022). When analyzing long-run purchasing power parity, the nonstationarity of exchange rates is studied. For example, Boswijk and Zu (2022) investigated the cointegration hypothesis of exchanges rates for the United Kingdom, Germany, and Japan relative to the United States. Furthermore, many macroeconomic variables such as GDP, inflation, money demand and interest rates are commonly believed to be nonstationary (Zivot and Andrews, 2002; Henry and Shields, 2004; Bae and De Jong, 2007). Hence, it is of great importance to take nonstationarity into consideration in modeling.

We can model nonstationary time series in both discrete time and continuous time. While discrete-time models are more popular in practical applications, there are several advantages of using continuous-time models in econometric modelling.

First, continuous-time models enable us to sample discretely at any frequency; hence, parameters are not subject to the time-aggregation problem. Due to the cost of collecting and measuring variables, data are usually collected at different frequencies (Ghysels and Miller, 2015). Hence, the frequency of these nonstationary data can vary considerably. For example, the DGP is usually collected at a quarterly frequency, and stock returns can be sampled yearly, quarterly, monthly, daily and even every second. Second, the continuous-time model enables the convenient handling of not only stock variables (CPI, population, and money supply) but also flow variables such as GDP, income, and exports by simple time aggregation. Last, the continuous-time model enables us to study the impact of the initial conditions on parameter estimation and inference, which is one of focuses of this chapter. In particular, the initial value enters into the limiting distribution of each estimated parameter. By taking into account the initial values, test statistics that are based on the new limiting distributions are expected to result in better finite sample performance.

This chapter discusses nonstationary continuous-time models, including models with unit root regressors and explosive regressors. Many empirical works in finance and economics have identified the I(1) property and explosiveness in time series data. Hence, the continuous-time models that are introduced in this chapter provide basic econometric tools for modelling, inference, and forecasting economic and finance time series. In particular, we aim to provide a selective review of estimation methods and inferential theory and to present an empirical example that demonstrates the usefulness of nonstationary continuous-time models. For both unit root and explosive processes, we begin our discussions with the univariate analysis framework, followed by multivariate models for generality.

The remainder of this chapter is organized as follows. Section 2 presents an overview. Section 3 introduces univariate continuous-time models, including stationary, unit root and explosive Ornstein-Uhlenbeck (OU hereafter) models, and discusses the estimation method and the limit theory. Section 4 introduces multivariate continuous-time models and discusses the estimation method and the limit theory. Section 5 presents Monte Carlo simulations. Section 6 provides an empirical study. Section 7 presents the conclusions of this study.

6.2 Overview

The random process that is studied in this chapter is the OU process. The OU process has a wide range of empirical applications in the practice. Vasicek (1977) assumes the spot rate process follows the OU process. Burgess (2014) provides an overview of the Vasicek model and its advantages and extensions. Barndorff-Nielsen and Shephard (2001) considers the non-Gaussian processes of OU type, where the deviation from Gaussianity allows for flexibility in modelling the dependence structures in the data. The model has been applied to study the instantaneous volatility of exchange rates (Barndorff-Nielsen and Shephard, 2003) and stock returns (Roberts et al., 2004).

The OU process is the solution of the following stochastic differential equation:

$$dx\left(t\right)=\kappa\left(\mu-x\left(t\right)\right)dt+dB_x(t), \tag{6.2.1}$$

with $B_x(t)=\sigma_{xx}W_x\left(t\right)$, where W_x is a standard Brownian motion. In a more general case, we allow $B_x(t)$ to be a Lévy process. The initial value is set to $x\left(0\right)=x_0=O_p\left(1\right)$.

Parameter μ is the drift of the process. Parameter κ determines the level of persistence in $x(t)$. For $\kappa>0$, the process is stationary when the initial condition comes from the infinite past ($x\left(0\right)=\int_{-\infty}^{0}e^{\kappa s}dB_x(s)$). In this case, μ is the unconditional mean of $x(t)$. κ determines the speed of mean-reversion. In particular, the process $x\left(t\right)$ is asymptotically stationary for $\kappa>0$. $x\left(t\right)$ is a Brownian motion for $\kappa=0$. $x\left(t\right)$ is explosive for $\kappa<0$. For data over a large time span, several different regimes of κ might be contemplated, possibly with break points separating the regimes. In this chapter, we focus on when $\kappa=0$ and when $\kappa<0$.

6.2.1 Exact discrete-time representation

When only discrete observations are available, discretization of the continuous-time model is necessary for estimating the parameters of interest, namely, the persistence parameter κ (defined in equation (6.2.1)) and the co-movement parameters B (defined in equation (6.4.1) for cointegration models) and β (defined in equation (6.4.8) for co-explosive models).

Suppose data are recorded at N equally spaced points over a time interval $[0,T]$. Denote the sampling interval by Δ such that we have $N=T/\Delta$ observations. In practical applications in economics, T measures the number of years. Typical values for T are not very large (between 1 and 50). In some cases, even if T is large, a smaller value for T may be used to avoid possible structural breaks (e.g., Zhou and Yu (2015)). In empirical studies, three sampling intervals are often considered, namely, $\Delta=1/12,\ 1/52, 1/252$, which correspond to monthly, weekly and daily frequencies.

Following Phillips (1972), the exact discrete-time representation of (6.2.1) is as follows:

$$x_{t\Delta}=a_\Delta\left(\kappa\right)x_{(t-1)\Delta}+g_\Delta+u_{x,t\Delta}, \tag{6.2.2}$$

where

$$\begin{aligned}
a_\Delta\left(\kappa\right) &= \exp\left(-\kappa\Delta\right),\\
g_\Delta &= \mu\left(1-e^{-\kappa\Delta}\right),\\
u_{x,t\Delta} &= \sigma_{xx}\int_{(t-1)\Delta}^{t\Delta}e^{-\kappa(t\Delta-s)}dB_x\left(s\right)\overset{d}{=}\mathcal{N}\left(0,\frac{\sigma_{xx}^2}{2\kappa}\left(1-e^{-2\kappa\Delta}\right)\right).
\end{aligned}$$

The initial value is $x_{0\Delta}=x_0=O_p\left(1\right)$. The autoregressive parameter $a_\Delta\left(\kappa\right)$ depends directly on the sampling frequency Δ and persistency parameter κ. Noted that Δ is related to the sample size N. The standard error of $u_{x,t\Delta}$ is $\lambda_\Delta=\sqrt{\frac{\sigma_{xx}^2}{2\kappa}\left(1-e^{-2\kappa\Delta}\right)}\sim\sigma_{xx}\sqrt{\Delta}\to0$ as $\Delta\to0$.

Re-standardizing equation (6.2.2) by λ_Δ, equation (6.2.2) becomes

$$\widetilde{x}_{t\Delta} = a_\Delta(\kappa)\,\widetilde{x}_{(t-1)\Delta} + \widetilde{g}_\Delta + \widetilde{u}_{x,t\Delta}, \tag{6.2.3}$$

where $\widetilde{x}_{t\Delta} = x_{t\Delta}/\lambda_\Delta$, $\widetilde{g}_\Delta = g_\Delta/\lambda_\Delta$ and $\widetilde{u}_{x,t\Delta} = u_{x,t\Delta}/\lambda_\Delta$. In addition, we have,

$$\begin{aligned} a_\Delta(\kappa) &= 1 - \kappa\Delta + O\left(\Delta^2\right) \to 1, \\ \widetilde{g}_\Delta &\sim O\left(\sqrt{\Delta}\right), \\ \widetilde{u}_{x,t\Delta} &\sim {}_{iid}\mathcal{N}(0,1), \end{aligned}$$

as $\Delta \to 0$. The initial value then become $\widetilde{x}_{0\Delta} = x_{0\Delta}/\lambda_\Delta = O_p\left(\Delta^{-1/2}\right)$.

Figure 6.1 presents one typical simulated path of the data generating process (6.2.1) using the exact discrete-time representation of (6.2.3). The parameter configuration is $T = 6$, $\Delta = 1/12$, $\mu = 1$, and $\sigma_{xx} = 1$. It corresponds to monthly data for a six-year time span. From the left to the right panel, the parameter κ is set to -0.3, 0, and 3. These values correspond to $a_\Delta(\kappa) = 1.0253$, 1, and 0.7788, respectively. The initial value x_0 is set to 0. Figure 6.1 shows that the explosive series demonstrates a rapid escalation, the unit root process wanders around and demonstrates a stochastic trend, and the stationary series exhibits a mean-reversion behavior.

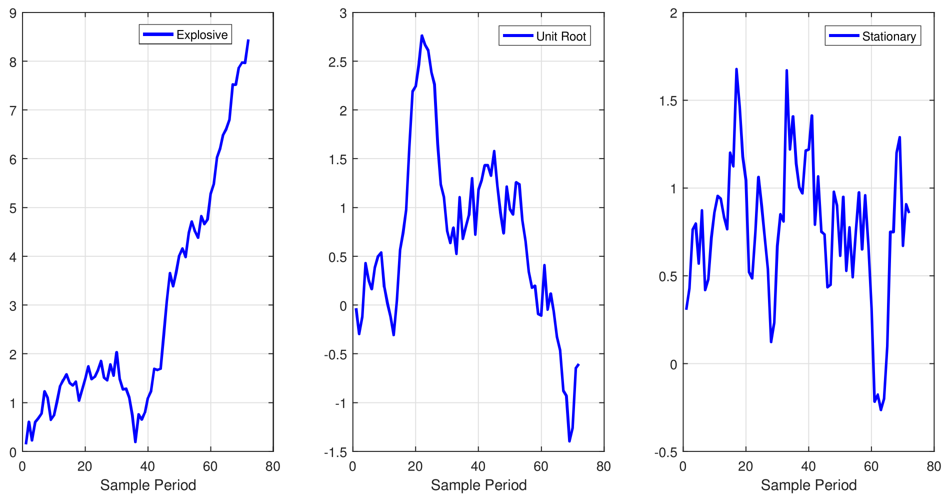

Figure 6.1: A typical realization of the data generating process (6.2.1).

6.2.2 Estimator and test statistic

For the OU process, the parameters of interest are g_Δ and $a_\Delta(\kappa)$ (a_Δ hereafter). The corresponding least squares (LS) estimators are as follows:

$$\begin{bmatrix} \widehat{g}_\Delta \\ \widehat{a}_\Delta \end{bmatrix} = \begin{bmatrix} T & \sum_{t=1}^N x_{(t-1)\Delta} \\ \sum_{t=1}^N x_{(t-1)\Delta} & \sum_{t=1}^N x_{(t-1)\Delta}^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum_{t=1}^N x_{t\Delta} \\ \sum_{t=1}^N x_{t\Delta} x_{(t-1)\Delta} \end{bmatrix}, \tag{6.2.4}$$

and hence,

$$\begin{bmatrix} \widehat{g}_\Delta - g_\Delta \\ \widehat{a}_\Delta - a_\Delta \end{bmatrix} = \begin{bmatrix} T & \sum_{t=1}^N \widetilde{x}_{(t-1)\Delta} \\ \sum_{t=1}^N \widetilde{x}_{(t-1)\Delta} & \sum_{t=1}^N \widetilde{x}_{(t-1)\Delta}^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum_{t=1}^N \widetilde{u}_{x,t\Delta} \\ \sum_{t=1}^N \widetilde{x}_{(t-1)\Delta}\widetilde{u}_{x,t\Delta} \end{bmatrix}. \tag{6.2.5}$$

The LS estimator of κ, which is denoted by $\hat{\kappa}$, is $-\frac{1}{\Delta}\ln\left(\widehat{a}_\Delta\right)$.

Defining $s_x^2 = N^{-1}\sum_{t=1}^{N}\widehat{u}_{x,t\Delta}^2$, we have

$$s_{a_\Delta}^2 = s_x^2\left(\sum_{t=1}^{N}\widetilde{x}_{(t-1)\Delta}^2\sigma_{xx}^2\Delta - \frac{1}{N}\left(\sum_{t=1}^{N}\widetilde{x}_{(t-1)\Delta}\sigma_{xx}\sqrt{\Delta}\right)^2\right)^{-1}.$$

The t statistic for $\widehat{a}_\Delta$ is

$$t_{a_\Delta} = \frac{\widehat{a}_\Delta - a_\Delta}{s_{a_\Delta}} = \frac{a_\Delta^N/\Delta\left(\widehat{a}_\Delta - a_\Delta\right)}{\left\{s_x^2\left(a_\Delta^{-2N}\Delta^2\sum_{t=1}^{N}\widetilde{x}_{(t-1)\Delta}^2\sigma_{xx}^2\Delta - \frac{1}{N}\left(a_\Delta^{-N}\Delta\sum_{t=1}^{N}\widetilde{x}_{(t-1)\Delta}\sigma_{xx}\sqrt{\Delta}\right)^2\right)^{-1}\right\}^{1/2}}. \tag{6.2.6}$$

Similarly, given $a_\Delta = \exp\left(-k\Delta\right)$, we have

$$\Delta s_\kappa = s_{a_\Delta} + o_p\left(\Delta\right),$$

and the t statistic for $\widehat{\kappa}$ is

$$t_\kappa = \frac{\widehat{\kappa} - \kappa}{s_\kappa} = \frac{a_\Delta^N\left(\widehat{\kappa} - \kappa\right)}{\left\{s_x^2\left(\frac{1}{a_\Delta^{2N}/\Delta^2}\sum_{t=1}^{N}x_{(t-1)\Delta}^2 - \frac{1}{N}\left(\frac{1}{a_\Delta^N/\Delta}\sum_{t=1}^{N}x_{(t-1)\Delta}\right)^2\right)^{-1}\right\}^{1/2}}. \tag{6.2.7}$$

6.2.3 Asymptotics

Three sampling schemes are considered in this Chapter:

1. the long-span asymptotics: $T \to \infty$, Δ is fixed, and hence, $N \to \infty$;
2. the double asymptotics: $T \to \infty$, $\Delta \to 0$, and hence, $N \to \infty$;
3. the infill asymptotics: $\Delta \to 0$, T is fixed, and hence, $N \to \infty$.

The first scheme (the long-span asymptotics) assumes that the sampling interval Δ is fixed and the time span $T \to \infty$ and hence the sample size N goes to infinity as $T \to \infty$. This is the usual asymptotic scheme that is considered in the discrete-time literature. The limiting distribution is not continuous in κ (Tang and Chen (2009), Yu (2012), and Yu (2014a)), and hence, disjoint pieces of the confidence interval are generated.

The third scheme (the infill asymptotics) enables the sample size to go to infinity by decreasing the sampling interval but keeping the time span fixed. Perron (1991a) provides the infill asymptotics regarding the persistence parameter κ for the OU process (6.2.1) without an intercept. The limiting distribution is continuous in κ for all values of κ, and hence, it provides a unified framework for making inferences regarding parameter κ. In addition, the limiting distribution of κ depends on the initial values, and it provides a good approximation of the finite sample distribution. In addition, the infill asymptotic theory with a known intercept is presented in Yu (2014b), and the infill asymptotic theory with a more general initial condition is provided in Perron (1991b).

The second scheme (the double asymptotics) combines the long-span scheme and the infill scheme and is referred to as the double asymptotics. There is a discontinuity of the limiting distribution of the persistence parameter κ (Zhou and Yu (2015)).

6.3 Univariate continuous-time models

In this section, we discuss univariate continuous-time models. We start with the univariate OU process with $\kappa > 0$ under three sampling schemes. Among the three asymptotics, the infill asymptotics provide the best approximation to the finite sample distribution (Zhou and Yu (2015)). However, the infill distribution depends on unknown parameters that cannot be consistently estimated. Lui et al. (2021) suggested using the new grid bootstrap for inference. Hence, we further discuss the local-to-unity model under the infill scheme since the discretized model is closely related to the local-to-unity model under the infill sampling scheme as $\Delta \to 0$. We also discuss the unit root continuous-time model. Finally, we consider the univariate OU process with $\kappa < 0$ (the explosive case). The use of a continuous-time framework with double asymptotics readily accommodates initial condition and drift effects, with a limit theory that is easy to implement in practice with no nuisance parameters (Wang and Yu (2016) and Chen et al. (2017)). In contrast, discrete-time models with local-to-unity and mildly integrated or mildly explosive autoregressive parameters typically involve localizing coefficients that enter the limit theory as nuisance parameters and are not generally consistently estimable, thereby complicating inference.

6.3.1 Univariate OU process with $\kappa > 0$

When $\kappa > 0$, $x(t)$ as defined in (6.2.1) is a stationary process. Zhou and Yu (2015) summarized the limiting distribution of the LS estimator of $\widehat{\kappa}$ under the three sampling schemes as follows.

The long-span asymptotics of $\widehat{\kappa}$ are

$$T\left(\widehat{\kappa}-\kappa\right) \xrightarrow{d} N\left(0, \frac{e^{2\kappa\Delta}-1}{\Delta}\right), \tag{6.3.1}$$

as T goes to ∞.

Under the double asymptotics, as $T \to \infty$ and $\Delta \to 0$, the limiting result is

$$T\left(\widehat{\kappa}-\kappa\right) \xrightarrow{d} N\left(0, 2\kappa\right),$$

as $T \to \infty$ and $\Delta \to 0$.

For model (6.2.1), if $\kappa < 0$, the infill asymptotic distribution of $\widehat{\kappa}$ is

$$T\left(\widehat{\kappa}-\kappa\right) \xrightarrow{d} \frac{A\left(\gamma_0, c\right)}{B\left(\gamma_0, c\right)},$$

where $c = -\kappa T$, $c_1 = e^{rc} - 1$, $c_2 = \frac{e^c - c - 1}{c^2}$, $c_3 = \frac{e^{2c} - 4e^c + 2c + 3}{2c^3}$, $c_4 = \frac{e^c - 1}{c}$, $J_c(r) = \int_0^r e^{c(r-s)} dW(s)$, $b = \mu\sqrt{-c\kappa}/\sigma_{xx}$, $\gamma_0 = X_0/\left(\sigma_{xx}\sqrt{T}\right)$ and

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

The infill asymptotics apply for all values of κ. Zhou and Yu (2015) showed that the infill asymptotic distribution provides much more accurate approximations to the finite sample distribution in a simulation study. In an empirical application, they applied the three alternative asymptotic theories to real monthly short-term interest rates from July 1954 to June 2002. The confidence intervals using all three alternative sampling schemes suggested a unit root in both the 90% and 95% confidence intervals. However, the confidence intervals that were implied by both the long-span asymptotics and the double asymptotics suggested that there was no unit root in the series.

6.3.2 Local-to-unity model under the infill scheme

When $\kappa \neq 0$, the autoregressive parameter in the exact discrete-time representation of (6.2.1) is $a_\Delta = \exp(-\kappa\Delta)$. We rewrite it as

$$a_\Delta = \exp(-\kappa\Delta) = \exp\left(\frac{-\kappa T}{N}\right).$$

Under the infill asymptotic scheme ($\Delta \to 0$, T is fixed, and hence, $N \to \infty$), the discretized model has a root that is local-to-unity (Phillips (1987b)). Correspondingly, Phillips (1987b) studied a near-integrated random process whose autoregressive parameter is defined by $\exp(c/N)$, where c is the local-to-unity parameter and measures the deviation from the unit root case and N is the sample size. When $c > 0$, the process is said to be locally explosive, and when $c < 0$, the process is said to be locally stationary.

The asymptotics of the estimated autoregressive parameter in discrete time are provided in Phillips (1987b). For continuous-time models, the infill asymptotic theory with a more general initial condition is provided in Perron (1991b). The infill asymptotic theory with a known intercept is presented in Yu (2014b). There is a discontinuity in the long-span asymptotics and the double asymptotics of $\widehat{\kappa}$ as κ passes through zero. However, the infill asymptotic distribution is continuous in κ. Hence, it provides a unified framework for statistical inference about κ.

Lui et al. (2021) considered the infill asymptotic distribution of $\widehat{\kappa}$. Since there are unknown parameters in the limiting result of $\widehat{\kappa}$ (Lemma 3.2 in Lui et al. (2021)), the CIs are infeasible. In particular, Lui et al. (2021) studied the following Lévy-driven models:

$$dx(t) = \kappa(\mu - x(t))\,dt + \sigma dL(t), x(0) = x_0 = O_p(1), \tag{6.3.2}$$

where $L(t)$ is a Lévy process that is defined on a filtered probability space $\left(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geqslant 0}, P\right)$ with $L(0) = 0$ a.s.. $F_t = \sigma\left\{[x(s)]_{s=0}^{t}\right\}$ and satisfies the following three properties:

1. Independent increments: for every increasing sequence of times $t_0, \ldots .t_N$, the random variables $L(t_0)$, $L(t_1) - L(t_0)$, ..., $L(t_N) - L(t_{N-1})$ are independent;

2. Stationary increments: the law of $L(t+\Delta) - L(t)$ is independent of t;

3. Stochastic continuity: for any $\varepsilon > 0$ and $t \geqslant 0$, $\lim_{\Delta\to 0} P(|L(t+\Delta) - L(t)| \geqslant \varepsilon) = 0$. The characteristic function of $L(t)$ takes the form $E(isL(t)) = \exp\{-t\psi(s)\}$, where i is the imaginary unit and the function $\psi(\cdot)\mathbb{R} \to \mathbb{C}$ is the Lévy exponent of $L(t)$.

The initial value x_0 is assumed to be independent of $L(t)$. The exact discrete-time representation of (6.3.2) is

$$x_{t\Delta} = a_\Delta x_{(t-1)\Delta} + g_\Delta + \lambda_\Delta u_{x,t\Delta},\ x_{0\Delta} = x_0 = O_p(1), \tag{6.3.3}$$

or

$$\widetilde{x}_{t\Delta} = a_\Delta \widetilde{x}_{(t-1)\Delta} + \widetilde{g}_\Delta + u_{x,t\Delta},\ \widetilde{x}_{0\Delta} = x_0/\lambda_\Delta, \tag{6.3.4}$$

where

$$\begin{aligned} a_\Delta &= \exp(-\kappa\Delta), \\ g_\Delta &= \left(\mu + \frac{\sigma i\psi'(0)}{\kappa}\right)\left(1 - e^{-\kappa\Delta}\right), \end{aligned}$$

$$\begin{aligned}
\lambda_{\Delta} &= \sigma\sqrt{\psi''(0)\left(1-e^{-2\kappa\Delta}\right)/2\kappa}, \\
u_{x,t\Delta} &= \lambda_{\Delta}^{-1}\left(\sigma\int_{(t-1)\Delta}^{t\Delta} e^{-\kappa(t\Delta-s)}dL(s) - \sigma i\psi'(0)\frac{1-e^{-\kappa\Delta}}{\kappa}\right).
\end{aligned}$$

Lemma 3.1 of Lui et al. (2021) provides the infill asymptotic theory that is stated below:

Lemma 6.3.1 *For model (6.3.2), as* $\Delta \to 0$,

$$N\left(\widehat{\kappa}-\kappa\right) \Rightarrow -\frac{\Upsilon_3 - \Upsilon_2 W(1)}{\Upsilon_1 - \Upsilon_2^2}, \tag{6.3.5}$$

where

$$\begin{aligned}
\Upsilon_1 &= \frac{\exp(2c)-4\exp(c)+2c+3}{2c^3}b^2 + \frac{2b}{c}\int_0^1(\exp(rc)-1)J_c(r)dr + \int_0^1 J_c^2(r)dr \\
&\quad + \frac{\exp(2c)-2\exp(c)+1}{c^2}b\gamma_0 + 2\gamma_0\int_0^1\exp(rc)J_c(r)dr + \gamma_0^2\frac{\exp(2c)-1}{2c}; \\
\Upsilon_2 &= \frac{\exp(c)-c-1}{c^2}b + \int_0^1 J_c(r)dr + \frac{\exp(c)-1}{c}\gamma_0; \\
\Upsilon_3 &= \frac{b}{c}\int_0^1(\exp(rc)-1)dW(r) + \int_0^1 J_c(r)dW(r) + \gamma_0\int_0^1\exp(rc)dW(r),
\end{aligned}$$

with

$$\begin{aligned}
J_c(r) &= \int_0^r \exp(c(r-s))dW(s);\ \gamma_0 = \frac{x_0}{\sigma_\psi\sqrt{N}};\ \sigma_\psi^2 = \sigma^2\psi''(0); \\
b &= \left(\mu + \frac{\sigma i\psi'(0)}{\kappa}\right)\frac{\sqrt{-c\kappa}}{\sigma_\psi};\ c = -\kappa N.
\end{aligned}$$

The limiting result facilitates the inversion of the coefficient-based statistic and the construction of confidence intervals (CIs) for κ. Since there are unknown parameters that cannot be consistently estimated in the limiting result, the CIs are infeasible. Hence, Lui et al. (2021) suggested using the new grid bootstrap for inference.

According to Lui et al. (2021), the following steps are needed to construct a grid bootstrap CI for κ:

1. Given the data $\{x_{t\Delta}\}_{t=0}^{N}$, the LS estimates of $\widehat{g}_{\Delta}$ and $\widehat{a}_{\Delta}$ are computed, and the residual

$$\widehat{u}_{x,t\Delta} = x_{t\Delta} - \widehat{a}_{\Delta}x_{(t-1)\Delta} - \widehat{g}_{\Delta},$$

   and its variance estimator $s_x^2 = N^{-1}\sum_{t=1}^{N}\widehat{u}_{x,t\Delta}^2$ are obtained;

2. A grid of a_{Δ}, $A_G = \{a_{\Delta 1}, a_{\Delta 2}, \ldots, a_{\Delta G}\}$ that is centered at $\widehat{a}_{\Delta}$ is constructed, with the first and last grid points being calculated from $\widehat{a}_{\Delta} \pm 7 \times s.e.(\widehat{a}_{\Delta})$. Lui et al. (2021) set $G = 50$ in the simulation;

3. Given a point in the grid $(a_{\Delta G} \in A_G)$, a second regression is performed:

$$x_{t\Delta} - a_{\Delta G}x_{(t-1)\Delta} = \widetilde{g}_{\Delta} + v_t,$$

   where v_t is the residual of the second regression. $\widetilde{g}_{\Delta}$ is a function of $a_{\Delta G}$;

4. $u^*_{x,t\Delta}$ (independently drawn from the empirical distribution of $\{\widehat{u}_{x,t\Delta}\}_{t=0}^{N}$) is an $i.i.d.$ random variable for $t = 1, \ldots, N$. Bootstrap data $\{x^*_{t\Delta}\}_{t=1}^{N}$ are generated based on $\left\{u^*_{x,t\Delta}\right\}_{t=1}^{\mathcal{N}}$ and the same initial condition as that of the observed data, namely,

$$x^*_{t\Delta} = a_{\Delta G} x^*_{(t-1)\Delta} + \widetilde{g}_{\Delta} + s_x \lambda_{\Delta} u^*_{x,t\Delta}, x^*_{0\Delta} = x_0;$$

5. B sets of bootstrap data of $\left\{\{x^{*B}_{t\Delta}\}_{t=0}^{N}\right\}_{b=1}^{B}$ are generated. For every set of bootstrap data, the LS estimator of κ (denoted by $\widehat{\kappa}^*$) is obtained, and the bootstrap coefficient-based statistic $T\left(\widehat{\kappa}^* - \kappa_G\right)$, where $\kappa_G = -\frac{\ln(a_{\Delta G})}{\Delta}$, is calculated. The $y^{t\Delta}$ quantile of the bootstrap statistic $T\left(\widehat{\kappa}^* - \kappa_G\right)$ is determined, which is denoted by $c^*_N\left(y|\kappa_G\right)$. Lui et al. (2021) set $B = 399$ in the simulation;

6. Following Hansen (1999), the quantile function $c^*_N\left(y|\kappa\right)$ is estimated by applying kernel regression:

$$c^*_N\left(y|\kappa\right) = \frac{\sum_{g=1}^{G} K\left(\frac{\kappa-\kappa_G}{\delta}\right) c^*_N\left(y|\kappa_G\right)}{\sum_{g=1}^{G} K\left(\frac{\kappa-\kappa_G}{\delta}\right)},$$

where $K\left(\cdot\right)$ is a kernel function and δ is the bandwidth. Lui et al. (2021) suggested using the Epanechnokov kernel such that $K\left(x\right) = \frac{3}{4}\left(1 - x^2 1\left(|x| \leqslant 1\right)\right)$ and choosing the bandwidth by LS cross-validation;

7. The CI for κ is obtained by inverting

$$CI^B_q = \left\{\kappa \in R : c^*_N\left(y_1|\kappa\right) \leqslant T\left(\widehat{\kappa} - \kappa_G\right) \leqslant c^*_N\left(y_2|\kappa\right)\right\}.$$

Step 4 is critically different from that in Hansen (1999) as the initial condition in the bootstrap samples is always set to $x^*_{0\Delta} = x_0$.

The grid bootstrap procedure that was proposed in Lui et al. (2021) is different from the standard grid bootstrap procedure that was proposed by Hansen (1999), which was developed for the local-to-unity AR(1) model when the initial condition is $O_p(1)$. In the procedure of Hansen (1999), the bootstrap initialization x^*_0 is set to 0 when the AR coefficient is not smaller than one and to the fitted value of x_0 that is based on the LS estimates when the AR coefficient is smaller than one. This choice of initialization is made to avoid dependence on the initialization. Since under the infill scheme the initial condition explicitly enters the infill asymptotic distribution, the initialization is needed when generating the bootstrapped samples. As a result, Lui et al. (2021) modified the grid bootstrap procedure when generating bootstrap samples by setting $x^*_0 = x_0$ regardless of the AR coefficient. The modified grid bootstrap leads to uniform inferences on the persistence parameter. Simulation studies showed that this improves the infill asymptotics in terms of empirical coverage.

6.3.3 Unit Root Models in Continuous Time

When $\kappa = 0$, $x(t)$ as defined in (6.2.1) is an $I\left(1\right)$ process. Zhou and Yu (2015) summarized the limiting distribution of the LS estimator of $\widehat{\kappa}$ under the three sampling schemes as follows.

The long-span asymptotics of $\widehat{\kappa}$ are

$$T\left(\widehat{\kappa} - 0\right) \xrightarrow{d} -\frac{\int_0^1 W\left(r\right) dW\left(r\right) - W\left(1\right) \int_0^1 W\left(r\right) dr}{\int_0^1 W\left(r\right)^2 dr - \left(\int_0^1 W\left(r\right) dr\right)^2}, \tag{6.3.6}$$

as T goes to ∞. The result in (6.3.6) was developed in Phillips (1987a) and Phillips (1987b).

Under the double asymptotics, as $T \to \infty$ and $\Delta \to 0$, the obtained limiting result is the same as the long-span asymptotics, as stated in (6.3.6).

6.3.4 Explosive Continuous-Time Models

Explosive OU processes are able to capture market exuberance in financial time series. In empirical work, the value of the autoregressive coefficient is also often taken to depend on the frequency of observation. This is because the use of higher-frequency data typically leads to a more persistent autoregressive coefficient estimate and expectations do not change over short time horizons as much as they do over long horizons. For these reasons, dependence of the autoregressive parameter on the sampling frequency often provides greater realism in empirical work where it is necessary to model near-unit root phenomena in continuous time.

Link with the discrete-time models

When $\kappa < 0$, model (6.2.1) is an explosive OU model. Its standardized discrete model (6.2.3) corresponds to the model of Phillips and Magdalinos (2007). In particular, Phillips and Magdalinos (2007) analyzed the following triangular model:

$$x_t = R_N x_{t-1} + u_{xt}, \quad x_0 = o_p(N^{\alpha/2}), \tag{6.3.7}$$

The autoregressive coefficient has the form $R_N = I_K + \frac{C}{N^\alpha}$, $\alpha \in (0,1)$, and $C = \text{diag}\,(c_1, \dots c_K)$, where $N^\alpha \to \infty$ when $N \to \infty$. x_t hence is a moderately integrated time series as R_N involves moderate deviations from a unit root in the sense of Phillips and Magdalinos (2007). Moreover, since $C > 0$, x_t is a mildly explosive time series. The vector (u_{xt}) is a sequence of random variables with zero mean and a finite second moment.

Justification of the moderate deviations from a unit root from the double asymptotic scheme in continuous-time systems is stated in Chen et al. (2017). First, the autoregressive setup with $R_N = I_K + \frac{C}{N^\alpha}$, $\alpha \in (0,1)$ corresponds to the scheme of $\Delta \to 0$ and $T \to \infty$. In a similar way of justification, Boswijk (2001) and Phillips et al. (2001) introduce the block-local-to-unity concept. They formulate the block-local-to-unity model using $1 + c/m$, where m denotes the number of observations such that $m = \frac{N}{M}$ for M blocks. When M is fixed and $m \to \infty$, the autoregressive coefficient parameter corresponds to the standard local-to-unity model, where $1 + \frac{c}{m}$ has the same order of magnitude as the local-to-unity parameter. When $m \to \infty$ and then $M \to \infty$, the autoregressive coefficient parameter is further from unity compared with the standard local-to-unity model. By setting $m = N^\gamma$ and $M = N^{1-\gamma}$, $\gamma \in (0,1)$, the latter case corresponds to the standard local-to-unity model. In this case, both m and M are decided by N and γ. By setting $\Delta = \frac{1}{m}$, and $T = M$, the latter case corresponds to the model Wang and Yu (2016) and Chen et al. (2017). However, unlike in Phillips et al. (2001) and Boswijk (2001), N is not predetermined in Wang and Yu (2016) or Chen et al. (2017) because N is determined after Δ and T are chosen in the setup.

By setting $\Delta = \frac{1}{m}$, and $T = M$, links between the notations (m, M, N) in Boswijk (2001) and those in Wang and Yu (2016) and Chen et al. (2017) (Δ, T, N) are established. Clearly, the double asymptotics ($T \to \infty$ and $\Delta \to 0$) correspond to case (2) in Phillips et al. (2001) and Boswijk (2001). However, unlike in Phillips et al. (2001) and Boswijk (2001), N is not predetermined in Wang and Yu (2016) or Chen et al. (2017) because N is determined after Δ and T are chosen in our setup.

Asymptotics

Consider the following modified model of (6.3.7):

$$x_t = \mu + R_N x_{t-1} + u_{xt}, x_0 = O_p\left(N^{\alpha/2}\right), \mu = O_p\left(N^{-\alpha/2}\right). \tag{6.3.8}$$

The error u_{xt} is an iid sequence with mean zero and variance σ^2_{xx}. Compared to the model in Phillips and Magdalinos (2007), model of (6.3.7) allows for a larger initial condition and a local (to zero) drift. The limit distribution of $\widehat{\mu}$, which is the LS estimator of the intercept parameter μ, follows simply as

$$\begin{aligned}
&\sqrt{N}\left(\widehat{\mu}-\mu\right) \\
&= \sqrt{N}\frac{\left[\begin{array}{c}\left(R_N^N N^{\alpha}\right)^{-2}\sum_{t=1}^{N}x_{t-1}^2\left(\frac{1}{N}\sum_{t=1}^{N}u_{xt}\right)- \\ \frac{1}{N}\left\{\left(R_N^N N^{\alpha}\right)^{-1}\sum_{t=1}^{N}x_{t-1}\right\}\left\{\left(R_N^N N^{\alpha}\right)^{-1}\sum_{t=1}^{N}x_{t-1}u_{xt}\right\}\end{array}\right]}{\left(R_N^N N^{\alpha}\right)^{-2}\sum_{t=1}^{N}x_{t-1}^2-\frac{1}{N}\left\{\left(R_N^N N^{\alpha}\right)^{-1}\sum_{t=1}^{N}x_{t-1}\right\}^2} \\
&= \frac{1}{\sqrt{N}}\sum_{t=1}^{N}u_{xt}+o_p(1)\xrightarrow{d}\mathcal{N}\left(0,\sigma_{xx}^2\right). \qquad (6.3.9)
\end{aligned}$$

This result is useful in testing for $\mu=0$ in the model of (6.3.7).

We let $\widetilde{x}_0=x_0N^{-\alpha/2}\Rightarrow X^*$, $\widetilde{\mu}=N^{\alpha/2}\mu\Rightarrow\mu^*$ and $D=X^*+\frac{\mu^*}{c}$. The LS estimator of R_N is denoted by $\widehat{R}_N$, and hence, $\widehat{c}=N^{\alpha}\left(\widehat{R}_N-1\right)$. The limit theory for $\widehat{R}_N$ follows from the proof of Theorem 2.1 in Chen et al. (2017):

$$R_N^N N^{\alpha}\left(\widehat{R}_N-R_N\right)\xrightarrow{d}2c\frac{\sigma_{xx}U_x}{\sigma_{xx}U_x+(2c)^{1/2}D}, \qquad (6.3.10)$$

where $U_x\stackrel{d}{=}\mathcal{N}(0,1)$. Given $\widehat{c}-c=\left(\widehat{R}_N-R_N\right)N^{\alpha}$, the limit theory for $\widehat{c}$ is

$$R_N^N\left(\widehat{c}-c\right)\xrightarrow{d}2c\frac{\sigma_{xx}U_x}{\sigma_{xx}U_x+(2c)^{1/2}D}.$$

Defining the regression residuals $\hat{u}_{xt}=x_t-\widehat{R}_Nx_{t-1}-\widehat{\mu}$ and considering that the residual variance estimate satisfies $s_x^2=N^{-1}\sum_{t=1}^{N}\hat{u}_{xt}^2\xrightarrow{p}\sigma_{xx}^2$, the associated t statistic is as follows:

$$\begin{aligned}
t_{R_N} &= \frac{\widehat{R}_N-R_N}{s_{R_N}}=\frac{R_N^N N^{\alpha}\left(\widehat{R}_N-R_N\right)}{\left\{s_x^2\left(\frac{1}{R_N^{2N}N^{2\alpha}}\sum_{t=1}^{N}x_{t-1}^2-\frac{1}{N}\left(\frac{1}{R_N^N N^{\alpha}}\sum_{t=1}^{N}x_{t-1}\right)^2\right)^{-1}\right\}^{1/2}} \\
&\xrightarrow{d} \frac{2c\frac{\sigma_{xx}U_x}{\sigma_{xx}U_x+(2c)^{1/2}D}}{\sigma_{xx}\left\{\left(\frac{1}{2c}\right)^2\left(\sigma_{xx}U_x+(2c)^{1/2}D\right)^2\right\}^{-1/2}}=U_x\stackrel{d}{=}\mathcal{N}(0,1).
\end{aligned}$$

Similarly, given $R_N=1+\frac{c}{N^{\alpha}}$($\alpha$ is known), we have $s_c=N^{\alpha}s_{R_N}$, and the associated t test statistic is

$$t_c=\frac{\widehat{c}-c^0}{s_c}=\left(\widehat{R}_N-R_N\right)N^{\alpha}N^{-\alpha}s_{R_N}^{-1}\xrightarrow{d}U_x\stackrel{d}{=}\mathcal{N}(0,1).$$

From the mappings

$$\begin{aligned}
\sigma_{xx}^2 &\mapsto 1, R_N\mapsto a_{\Delta}=e^{-\kappa\Delta}, X^*\mapsto\frac{x_0}{\sigma_{xx}}, \mu\mapsto\frac{\mu\kappa}{\sigma_{xx}}\Delta^{1/2}, \mu^*\mapsto\frac{\mu\kappa}{\sigma_{xx}}, \\
D_{\mathcal{N}} &\mapsto D_{\Delta}=\tilde{x}_{0\Delta}\Delta^{1/2}-\frac{\Delta^{-1/2}\tilde{g}_{\Delta}}{\kappa}\to D^*=\frac{x_0}{\sigma_{xx}}-\frac{\mu}{\sigma_{xx}},
\end{aligned}$$

with $\Delta=1/N^{\alpha}$, it follows that

$$\frac{a_{\Delta}^N}{\sqrt{\Delta}}\left(\widehat{a}_{\Delta}-a_{\Delta}\right)\xrightarrow{d}(-2\kappa)\frac{\sigma_{xx}U_x}{U_x+(-2\kappa)^{1/2}D^*}. \qquad (6.3.11)$$

and

$$a_{\Delta}^{N}\left(\widehat{\kappa}-\kappa\right)\xrightarrow{d}2\kappa\frac{\sigma_{xx}U_x}{\sigma_{xx}U_x+\left(-2\kappa\right)^{1/2}D}.$$

Defining $s_x^2=N^{-1}\sum_{t=1}^{N}\widehat{u}_{x,t\Delta}^2$, which satisfies $\Delta^{-1}s_x^2\xrightarrow{p}\sigma_{xx}^2$, the t statistic for $\widehat{a}_{\Delta}$ is

$$\begin{aligned}
t_{a_{\Delta}} &= \frac{\widehat{a}_{\Delta}-a_{\Delta}}{s_{a_{\Delta}}}\\
&= \frac{a_{\Delta}^{N}/\Delta\left(\widehat{a}_{\Delta}-a_{\Delta}\right)}{\left\{s_x^2\left(a_{\Delta}^{-2N}\Delta^2\sum_{t=1}^{N}\widetilde{x}_{(t-1)\Delta}^2\sigma_{xx}^2\Delta-\frac{1}{N}\left(a_{\Delta}^{-N}\Delta\sum_{t=1}^{N}\widetilde{x}_{(t-1)\Delta}\sigma_{xx}\sqrt{\Delta}\right)^2\right)^{-1}\right\}^{1/2}}\\
&\xrightarrow{d} \frac{-2\kappa\frac{\sigma_{xx}Ux}{\sigma_{xx}U_x+(-2\kappa)^{1/2}D}}{\sigma_{xx}\left\{\left(\frac{1}{-2\kappa}\right)^2\left(\sigma_{xx}U_x+\left(-2\kappa\right)^{1/2}D\right)^2\right\}^{-1/2}}=U_x\overset{d}{=}\mathcal{N}\left(0,1\right),
\end{aligned}$$

and the t statistic for $\widehat{\kappa}$ is

$$\begin{aligned}
t_{\kappa} &= \frac{\widehat{\kappa}-\kappa}{s_{\kappa}}=\frac{a_{\Delta}^{N}\left(\widehat{\kappa}-\kappa\right)}{\left\{s_x^2\left(\frac{1}{a_{\Delta}^{2N}/\Delta^2}\sum_{t=1}^{N}x_{(t-1)\Delta}^2-\frac{1}{N}\left(\frac{1}{a_{\Delta}^{N}/\Delta}\sum_{t=1}^{N}x_{(t-1)\Delta}\right)^2\right)^{-1}\right\}^{1/2}}\\
&\xrightarrow{d} \frac{-2\kappa\frac{\sigma_{xx}Ux}{\sigma_{xx}U_x+(-2\kappa)^{1/2}D}}{\sigma_{xx}\left\{\left(\frac{1}{-2\kappa}\right)^2\left(\sigma_{xx}U_x+\left(-2\kappa\right)^{1/2}D\right)^2\right\}^{-1/2}}=U_x\overset{d}{=}\mathcal{N}\left(0,1\right),
\end{aligned}$$

where $\Delta s_{\kappa}=s_{a_{\Delta}}+o_p\left(\Delta\right)$ given $a_{\Delta}=\exp\left(-k\Delta\right)$. Clearly, t_{κ} is a feasible statistic, in contrast to the discrete-time case, where the test statistic relies on the unknown rate parameter α. Notably, if α is unknown, then both the estimate $\widehat{c}$ and the standard error $s_c=N^{\alpha}s_{R_N}$ are unavailable, and inference using this limit theory for $\widehat{c}$ is infeasible.

There are also important differences between the discrete-time model (6.2.3) and the MP model (6.3.7). First, the autoregressive coefficient in (6.2.3) is determined by the sampling interval Δ. However, the autoregressive coefficient in (6.3.7) is formulated as a function of the overall sample size N. Second, the initial condition for x_t is assumed to be $o_p(N^{\alpha/2})=o_p\left(\Delta^{-1/2}\right)$ in (6.3.7). However, in (6.2.3), it has a larger order of magnitude of $O_p\left(\Delta^{-1/2}\right)$. This result corresponds to a distant past initialization introduced in Phillips and Magdalinos (2009). These initializations do affect the limit theory due to its larger order property. Third, the presence of the constant intercept in (6.3.7) dominates the asymptotics. However, the intercept in (6.2.3)

$$\begin{aligned}
\widetilde{g}_{\Delta} &= g_{\Delta}/\lambda_{\Delta}=\mu\left(1-e^{-\kappa\Delta}\right)/\lambda_{\Delta}\\
&= \frac{\mu\left(1-e^{-\kappa\Delta}\right)}{\sigma_{xx}\sqrt{\left(1-e^{-2\kappa\Delta}\right)/2\kappa}}\\
&= \frac{\mu\kappa\Delta}{\sigma_{xx}\Delta^{1/2}}\left\{1+o\left(\Delta\right)\right\}=O\left(\sqrt{\Delta}\right)
\end{aligned}$$

is asymptotically negligible as $\Delta\to 0$; hence, it does not affect the double asymptotics for the model.

Lévy-driven models

Wang and Yu (2016) studied Lévy-driven models (6.3.2) with $\kappa<0$. The LS estimator of $a_{\Delta}\left(\kappa\right)$ and g_{Δ} is presented in (6.2.4). Wang and Yu (2016) developed the simultaneous double asymptotics for the estimator $\widehat{a}_\Delta(\kappa)$ and $\widehat{g}_\Delta$. The result is

$$\begin{aligned}\frac{a_\Delta^N}{a_\Delta(\kappa)^2-1}\left(\widehat{a}_\Delta-a_\Delta\right) &\overset{d}{\to} \frac{\xi}{\eta+D},\\ \sqrt{T\kappa}\left(\widehat{g}_\Delta-g_\Delta\right) &\overset{d}{\to} \sigma\sqrt{\psi''(0)}Z,\end{aligned}$$

where $D=\sqrt{2}\left(\kappa\mu+\sigma i\psi'(0)-\kappa x_0\right)/\left(\sigma\sqrt{-\kappa\psi''(0)}\right)$ and (ξ,η,Z) are independent standard normal random variables. They also considered two types of sequential double asymptotics: (i) $T\to\infty$ followed by $\Delta\to 0$ and (ii) $\Delta\to 0$ followed by $T\to\infty$. The two sequential asymptotics are the same as the simultaneous double asymptotics.

6.4 Multivariate Continuous-Time Models

This section is the multivariate counterpart of Section 3. In particular, we discuss both the cointegration and co-explosive systems.

6.4.1 Cointegration system and error-correction models

One of the most researched relationships for empirical work is cointegration. The model measures the long-run relationship, namely, equilibrium, between underlying variables. When the underlying variables deviate from the long-run relationship in the short term dynamics, we use an error correction model to measure the deviation. In discrete time, Granger (1981) and Engle and Granger (1987) studied the connection between cointegration systems and error correction models. Later, Phillips (1991) showed how to formulate a cointegration system and error correction models in continuous time and proposed an inferential procedure for such systems based on frequency-domain techniques.

Let $y(t)$ be an $m\times 1$ $I(1)$ process in continuous time and $u(t)$ be an $m\times 1$ stationary time series. $y(t)$ is further partitioned into two subvectors: an $m_1\times 1$ vector $y_1(t)$ and an $m_2\times 1$ vector $y_2(t)$. $u(t)$ is further partitioned into two subvectors: an $m_1\times 1$ vector $u_1(t)$ and an $m_2\times 1$ vector $u_2(t)$. We let $D=d/dt$ represent the mean square differential operator with respect to continuous time. The model that Phillips (1991) studied is as follows:

$$y_1(t) = By_2(t)+u_1(t), \tag{6.4.1}$$
$$Dy_2(t) = u_2(t), \tag{6.4.2}$$

where B is an $m_1\times m_2$ matrix of unknown coefficients. Equation (6.4.1) expresses the long-run equilibrium relationship between variables $y_1(t)$ and $y_2(t)$. This long-run relationship is perturbed by a stationary deviation $u_1(t)$. The error-correction model (ECM hereafter) of the system (6.4.1)-(6.4.2) is of the form

$$Dy_1(t)=-[I,-B]\,y(t)+u_1(t)+Bu_2(t)+Du_1(t). \tag{6.4.3}$$

Combining equations (6.4.2) and (6.4.3), the error correction model (6.4.4) is obtained for the cointegration system (6.4.1)-(6.4.2):

$$Dy(t)=-EAy(t)+w(t), \tag{6.4.4}$$

where

$$E=\begin{bmatrix} I_{m_1\times m_1}\\ 0_{m_1\times m_2}\end{bmatrix},\ A=\left[I_{m_1\times m_1},-B\right],\ w(t)=\begin{bmatrix} u_1(t)+Bu_2(t)+Du_1(t)\\ u_2(t)\end{bmatrix}.$$

The exact discrete model of (6.4.4) is

$$\begin{aligned} y(n) &= \exp(-EA)\, y(n-1) + \varepsilon(n), \\ \varepsilon(n) &= \int_0^1 \exp(-sEA)\, w(n-s)\, ds. \end{aligned} \tag{6.4.5}$$

Since $AE = I$ and $\exp(-EA) = I - \frac{e-1}{e}EA$, (6.4.5) can be rewritten in the following triangular system ECM format:

$$\begin{aligned} \Delta y(n) &= -EAy(n-1) + x(n), \\ x(n) &= \varepsilon(n) + (1/e)\, EAy(n-1), \end{aligned} \tag{6.4.6}$$

where $x(n) = I(0)$.

The coefficient matrix B, which is a submatrix of A, is to be estimated. Phillips (1991) suggested the Hannan efficient and band spectral estimators for estimating the coefficient matrix B under the following assumptions:

1. The residual process $x(n)$ is stationary with spectral matrix $f_{xx}^{d}(\lambda) > 0$ that is continuous at the origin $\lambda = 0$.
2. We set $\Omega = 2\pi\ f_{xx}^{d}(\lambda)$, decompose the long-run covariance matrix as $\Omega = \Sigma + \Lambda + \Lambda'$, where $\Sigma = E\left(x(0)\, x(0)'\right)$ and $\Lambda = \sum_{k=1}^{\infty} E\left(x(0)\, x(k)'\right)$, and define $\Delta = \Sigma + \Lambda$. Ω is partitioned as $\begin{bmatrix} \Omega_{11} & \Omega_{21}' \\ \Omega_{21} & \Omega_{22} \end{bmatrix}$, and the dimensions of Ω_{11}, Ω_{21} and Ω_{21} are $m_1 \times m_1$, $m_2 \times m_1$, and $m_2 \times m_2$, respectively. We define $\Omega_{11 \cdot 2} = \Omega_{11} - \Omega_{21}'\Omega_{22}^{-1}\Omega_{21}$.

In addition, we define

$$\begin{aligned} y_*(n)' &= \left(y_1(n)', \Delta y_2(n)'\right), w_*(\lambda)' = (2\pi T)^{-1/2} \sum_{n=1}^{T} y_*(n)\, e^{in\lambda}, \\ w_2(\lambda) &= (2\pi T)^{-1/2} \sum_{n=1}^{T} y_2(n)\, e^{in\lambda} \text{ for } \lambda \in [-\pi, \pi]. \end{aligned}$$

We let $\widehat{B}$ denote the LS estimator of B and let $\omega_j = \pi j/M$ for $j = -M+1, \ldots, M$, where M is an integer (this requires $M = o\left(T^{1/2}\right)$). The smoothed periodogram estimate is

$$\widehat{f}_{xx}(\omega_j) = \frac{M}{T} \sum_{\lambda_s \in \mathcal{B}_j} \left[w_*(\lambda_s) - E\widehat{B}w_2(\lambda_s)\right]\left[w_*(\lambda_s) - E\widehat{B}w_2(\lambda_s)\right]$$

with

$$\begin{aligned} \lambda_s &= 2\pi s/T, \\ \mathcal{B}_j &= (\omega_j - \pi/2M < \lambda \leqslant \omega_j + \pi/2M). \end{aligned}$$

Hence, the Hannan efficient estimator of B is

$$\begin{aligned} vec\left(\widetilde{B}\right) = &\left[\frac{1}{2M} \sum_{j=-M+1}^{M} E'\widehat{f}_{xx}(\omega_j)^{-1} E \otimes \widehat{f}_{22}(\omega_j)\right]^{-1} \times \\ &\left[\frac{1}{2M} \sum_{j=-M+1}^{M} \left(E'\widehat{f}_{xx}(\omega_j)^{-1} \otimes I_{m_2}\right) vec\left(\widehat{f}_{*2}(\omega_j)\right)\right], \end{aligned} \tag{6.4.7}$$

where

$$\widehat{f}_{22}(\omega_j) = l^{-1}\sum_{\mathcal{B}_j} w_2(\lambda_2)\, w_2(\lambda_s)^*,$$
$$\widehat{f}_{*2}(\omega_j) = l^{-1}\sum_{\mathcal{B}_j} w_*(\lambda_2)\, w_2(\lambda_s)^*,$$

for $l = [T/M]$. The band spectral estimator $\widetilde{B}_0$ is similarly defined but based on the spectral estimates at the origin.

Phillips (1991) also considered the subsystem band spectral estimator of B, which is expressed as follows:

$$B_0^+ = \left[\widehat{f}_{12}(0) - \widehat{f}_{1\Delta}(0)\,\widehat{f}_{\Delta\Delta}(0)^{-1}\,\widehat{f}_{\Delta 2}(0)\right]\left[\widehat{f}_{22}(0) - \widehat{f}_{2\Delta}(0)\,\widehat{f}_{\Delta\Delta}(0)^{-1}\,\widehat{f}_{\Delta 2}(0)\right]^{-1},$$

where $\widehat{f}_{\Delta\Delta}(0)$ denotes the estimated spectral matrix of $\Delta y_2(n)$ and $\widehat{f}_{12}(0)$, $\widehat{f}_{1\Delta}(0)$ and $\widehat{f}_{2\Delta}(0)$ denote the estimated cross spectral matrices of $(y_1(n), y_2(n))$, $(y_1(n), \Delta y_2(n))$ and $(y_2(n), \Delta y_2(n))$, respectively.

Following the approach in Phillips and Perron (1988), the three estimators $\widetilde{B}$, $\widetilde{B}_0$, and B_0^+ are asymptotically equivalent, and the following limiting result is obtained:

$$T\left(\widetilde{B} - B\right), T\left(\widetilde{B}_0 - B\right), T\left(B_0^+ - B\right) \xrightarrow{d} \left(\int_0^1 dS_{1\cdot 2} S_2'\right)\left(\int_0^1 S_2 S_2'\right)^{-1},$$

where

$$\begin{bmatrix} S_{1\cdot 2} \\ S_2 \end{bmatrix} = BM\left(\begin{bmatrix} \Omega_{11\cdot 2} & 0 \\ 0 & \Omega_{22} \end{bmatrix}\right).$$

The finite sample performance of the Hannan efficient estimator (6.4.7) was evaluated in Chambers (2001). Simulation studies in Chambers (2001) showed that the choices of bandwidth parameter and kernel function impact the finite sample performance. In particular, in terms of mean squared error, they found that the band spectral estimator $\widetilde{B}_0$ with the Parzen kernel performed the best in the studied cases.

6.4.2 Explosive Continuous-Time Systems

Chen et al. (2017) studied the following scalar continuous-time model in two variates, $y(t)$ and $x(t)$. Here, $x(t)$ follows an Ornstein–Uhlenbeck process, and the stochastic process $y(t)$ co-moves with $x(t)$ as

$$y(t) = \beta x(t) + u_0(t), \tag{6.4.8}$$
$$dx(t) = \kappa(\mu - x(t))\,dt + dB_x(t),\; x(0) = x_0 = O_p(1),\; \kappa < 0, \tag{6.4.9}$$

where $u_0(t)$ is Gaussian pure noise – a generalized stochastic process in continuous time (see Hannan (1970) and Phillips (1991)). and $B_x(t) = \sigma_{xx} W_x(t)$ with W_x being a standard Brownian motion that may be correlated with W_0. The parameter of central interest for inference is the coefficient β, which captures the co-movement between $y(t)$ and $x(t)$.

Let $u_0(t) = DB_0(t)$, where $B_0(t) = \sigma_{00} W_0(t)$ with W_0 being the standard Brownian motion, and $D = d/dt$ is the mean square differential operator. Consider the case where $u_0(t)$ is unrealizable as a covariance stationary stochastic proces, the corresponding discrete-time process of $y(t)$ is realizable. In particular, $y(t)$ takes the form of a pure noise process of independent identically distributed (iid) $\mathrm{N}(0, \sigma_{00}^2)$ errors. This formulation is extensively used in modeling microstructure noise effects in the measurement of efficient financial asset prices. This formulation corresponds to the discrete-time system (6.4.12). Phillips and Yu (2006) discussed such models with standard formulation of an efficient price subject to unobserved microstructure noise. Alternate specifications is to set $u_0(t) = 0$ a.s.. This corresponds to the limit form of a discrete-time cointegrated system, as discussed in Remark 12 of Chen et al. (2017).

The continuous-time system (6.4.8)-(6.4.9) is related to the market microstructure literature. For example, Zhang et al. (2005), Aït-Sahalia et al. (2005)(AMZ, hereafter) and Bandi and Russell (2006) assume that the observed transaction price is equal to the sum of the efficient price and an iid noise component. In particular, AMZ assumes that the observed logarithmic price of a security follows

$$\widetilde{X}(t\Delta) = X(t\Delta) + U(t\Delta), \tag{6.4.10}$$

where the logarithmic efficient price $X(t)$ is

$$dX(t) = \mu(X(t))\,dt + \sigma(X(t))\,dW(t), \tag{6.4.11}$$

and the iid error $\{U(t\Delta)\}$ has a zero mean and a finite variance σ_u^2. $U(t\Delta)$ is independent of $X(t\Delta)$. Formulating of systems (6.4.10)-(6.4.11) in continuous time have been discussed in the market microstructure noise literature, see Hansen and Lunde (2006) and Phillips and Yu (2006).

The exact discrete-time representation

Following Phillips (1972), the exact discrete time representation of (6.4.8)-(6.4.9) is

$$\begin{aligned} y_{t\Delta} &= \beta x_{t\Delta} + u_{0,t\Delta}, & (6.4.12)\\ x_{t\Delta} &= a_\Delta(\kappa)\, x_{(t-1)\Delta} + g_\Delta + u_{x,t\Delta},\ x_{0\Delta} = x_0 = O_p(1), & (6.4.13) \end{aligned}$$

where

$$\begin{aligned} a_\Delta(\kappa) &= \exp(-\kappa\Delta), \\ g_\Delta &= \mu\left(1 - e^{-\kappa\Delta}\right), \\ u_{x,t\Delta} &= \sigma_{xx}\int_{(t-1)\Delta}^{t\Delta} e^{-\kappa(t\Delta - s)} dB_x(s) \overset{d}{=} \mathcal{N}\left(0, \frac{\sigma_{xx}^2}{2\kappa}\left(1 - e^{-2\kappa\Delta}\right)\right), \\ u_{0,t\Delta} &\overset{d}{=} \mathcal{N}\left(0, \sigma_{00}^2\right). \end{aligned}$$

The autoregressive parameter $a_\Delta(\kappa)$ is determined by the sampling frequency Δ, and hence both Δ and $a_\Delta(\kappa)$ are related to the sample size N. The standard error of $u_{x,t\Delta}$ is $\lambda_\Delta \sim \sigma_{xx}\sqrt{\Delta} \to 0$. But the variance of $u_{0,t\Delta}$ does not depend on the sampling interval Δ. $u_{0,t\Delta}$ in discrete time corresponds to the generalized stochastic process $u_0(t)$ in continuous time. Gaussianity follows from the Brownian motion driver processes in (6.4.8)-(6.4.9).

Standardizing the equation (6.4.13) by λ_Δ gives

$$\begin{aligned} y_{t\Delta} &= \beta x_{t\Delta} + u_{0,t\Delta}, & (6.4.14)\\ \widetilde{x}_{t\Delta} &= a_\Delta(\kappa)\, \widetilde{x}_{(t-1)\Delta} + \widetilde{g}_\Delta + \widetilde{u}_{x,t\Delta},\ \widetilde{u}_{x,t\Delta} \sim_{iid} \mathcal{N}(0,1), & (6.4.15) \end{aligned}$$

where $\widetilde{x}_{t\Delta} = x_{t\Delta}/\lambda_\Delta$ ($\widetilde{x}_{0\Delta} = x_{0\Delta}/\lambda_\Delta$), $\widetilde{g}_\Delta = g_\Delta/\lambda_\Delta$ as $\Delta \to 0$. Clearly, we have $\frac{1}{N\Delta} = \frac{1}{T} \to 0$, and $a_\Delta(\kappa) = 1 - \kappa\Delta + O\left(\Delta^2\right) \to 1$ when $T \to \infty$ and $\Delta \to 0$.

Since $\kappa < 0$, $\widetilde{x}_{t\Delta}$ in (6.4.15) is a mildly explosive process, as defined in Phillips and Magdalinos (2007). Further as $\Delta \to 0$, it follows that $\widetilde{x}_{0\Delta} = x_{0\Delta}/\lambda_\Delta = O_p\left(\Delta^{-1/2}\right)$. Thus, in the standardized discrete system (6.4.14)-(6.4.15), the initial condition is $\widetilde{x}_{0\Delta} \sim O_p\left(\Delta^{-1/2}\right)$, while in the original system (6.4.8)-(6.4.9), it is $x_0 \sim O_p(1)$. In addition, the order of magnitude of the drift is $O\left(\sqrt{\Delta}\right)$ in model (6.4.15) but is $O_p(1)$ in (6.4.9).

The standardized discrete system (6.4.14)-(6.4.15) corresponds to the modified MP model

$$y_t = Ax_t + u_{0t}, \tag{6.4.16}$$

$$x_t = \mu + R_N x_{t-1} + u_{xt}, x_0 = x_{0\mathcal{N}} = O_p\left(\mathcal{N}^{\alpha/2}\right), \mu = O_p\left(\mathcal{N}^{-\alpha/2}\right). \quad (6.4.17)$$

where the error $u_t = [u_{0t}, u_{xt}]'$ is an iid sequence with mean zero and covariance $\begin{bmatrix} \sigma_{00}^2 & \sigma_{0x} \\ \sigma_{0x} & \sigma_{xx}^2 \end{bmatrix}$. This model extends systems in MP by allowing for a larger initial condition and a (local to zero) drift.

Double asymptotics for the co-explosive system

The LS estimator $\widehat{\beta}$ of the slope coefficient in the continuous-time model (6.4.8) is

$$\widehat{\beta} - \beta = \left(\sum_{t=1}^{N} x_{t\Delta}^2\right)^{-1} \left(\sum_{t=1}^{N} x_{t\Delta} u_{0,t\Delta}\right) = \frac{1}{\lambda_\Delta}\left(\sum_{t=1}^{N} \widetilde{x}_{t\Delta}^2\right)^{-1} \left(\sum_{t=1}^{N} \widetilde{x}_{t\Delta} u_{0,t\Delta}\right). \quad (6.4.18)$$

Chen et al. (2017) presents the associated limit theory, according to which

$$\frac{a_\Delta^N}{\sqrt{\Delta}}\left(\widehat{\beta} - \beta\right) \xrightarrow{d} (-2\kappa)\frac{\sigma_{00} U_0}{\sigma_{xx} U_x + (-2\kappa)^{1/2}(x_0 - \mu)}, \quad (6.4.19)$$

under the condition that $\Delta^{1-\alpha} T^\alpha \to 1$ as $T \to \infty$. The limit result can be extended to the case where $u_0(t)$ is weakly dependent.

Next, we let the residual variance estimate be $s_0^2 = N^{-1}\sum_{t=1}^{N} \widehat{u}_{0,t\Delta}^2$, which satisfies $s_0^2 \xrightarrow{p} \sigma_{00}^2$. The usual t statistic is

$$\begin{aligned} t_\beta &= \frac{\widehat{\beta} - \beta}{s_\beta} = \frac{\left(\widehat{\beta} - \beta\right) a_\Delta^N / \sqrt{\Delta}}{\left\{ s_0^2 \left(a_\Delta^{-2N} \Delta \sum_{t=1}^{N} \widetilde{x}_{t\Delta}^2 \sigma_{xx}^2 \Delta \right)^{-1} \right\}^{1/2}} \\ &\xrightarrow{d} \frac{(-2\kappa)\frac{\sigma_{00} U_0}{\sigma_{xx} U_x + (-2\kappa)^{1/2}(x_0 - \mu)}}{\sigma_{00}\left\{ \left(\frac{1}{-2\kappa}\right)^2 \left(U_x + (-2\kappa)^{1/2} D^* \right)^2 \sigma_{xx}^2 \right\}^{-1/2}} = U_0 \overset{d}{=} \mathcal{N}(0,1), \end{aligned}$$

where $D^* = X^* - \mu^*$ with $\widetilde{x}_0 = x_{0N} N^{-\alpha/2} \Rightarrow X^*$ and $\widetilde{\mu} = N^{\alpha/2}\mu \Rightarrow \mu^*$. This result leads to feasible inference concerning the slope coefficient β in continuous time.

When the errors are weakly dependent such that $u_{0,t\Delta}$ is a sequence of zero mean and weakly dependent errors, the limiting results of (6.4.19) holds. In particular, we assume $u_{0,t\Delta}$ satisfying Assumption LP as defined in MP, such that

$$N^{-1/2}\sum_{t=1}^{N} u_{0,t\Delta}^2 \Rightarrow \mathcal{N}\left(0, \omega_{00}^2\right).$$

where the long-run variance ω_{00}^2 can be decomposed as

$$\omega_{00}^2 = \sigma_{00}^2 + 2\lambda_{00},$$

with variance component σ_{00}^2 and one-sided long-run variance

$$\lambda_{00} := \sum_{i=1}^{\infty} E\left(u_{0,t\Delta} u_{0,(t-i)\Delta}\right).$$

Then we obtain

$$\frac{a_\Delta^N}{\sqrt{\Delta}}\left(\widehat{\beta} - \beta\right) \xrightarrow{d} (-2\kappa)\frac{\omega_{00} U_0}{\sigma_{xx} U_x + (-2\kappa)^{1/2}(x_0 - \mu)}.$$

follows the limit results from Theorem 4.1 in MP and (6.4.19),

Alternatively, we set $u_0(t) \overset{a.s.}{=} 0$ in (6.4.8). This formulation corresponds to a discrete system (6.4.14) with $u_{0,t\Delta} \sim_{iid} \mathcal{N}\left(0, \sigma_{00}^2 \Delta\right)$. Hence, the relationship between $x(t)$ and $y(t)$ is exact in the limit, which is analogous to the relationship of limit Brownian motion processes (B_y, B_x) of cointegrated discrete series (y_t, x_t) where $x_t = x_{t-1} + u_{xt}$ and $y_t = \beta x_t + u_{0t}$ with (u_{0t}, u_{xt}) stationary and limiting linear relation $B_y(t) = \beta B_x(t)$. In this case, it follows that the limit distribution of $\widehat{\beta}$ is

$$\frac{a_\Delta^N}{\Delta}\left(\widehat{\beta} - \beta\right) \xrightarrow{d} (-2\kappa) \frac{\sigma_{00} U_0}{\sigma_{xx} U_x + (-2\kappa)^{1/2}(x_0 - \mu)}, \tag{6.4.20}$$

when $\Delta \to 0$ and $T \to \infty$. In view of the scaling effect in the discrete time error, there is a faster rate of convergence in the estimation of β.

Double asymptotics for the cointegration system When $\kappa = 0$, the model (6.4.8)-(6.4.9) corresponds to the cointegration system in continuous time. Its discretized system corresponds to the following cointegration model:

$$y_t = Ax_t + u_{0t}, \tag{6.4.21}$$

$$x_t = Rx_{t-1} + u_{xt}, x_0 = O_p\left(N^{1/2}\right). \tag{6.4.22}$$

where the error $u_t = [u_{0t}, u_{xt}]'$ is an iid sequence with mean zero and covariance $\begin{bmatrix} \sigma_{00}^2 & \sigma_{0x} \\ \sigma_{0x} & \sigma_{xx}^2 \end{bmatrix}$. We have assumed that $x_0 N^{-\frac{1}{2}} \Rightarrow X^*$.

We have the following asymptotics.

Theorem 6.4.1 *For the discrete time system (6.4.21)-(6.4.22) with $R = 1$, when $N \to \infty$,*

(i)

$$N\left(\widehat{R} - 1\right) \Rightarrow \frac{X^* \sigma_{xx} W_x(1) + \sigma_{xx}^2 \int_0^1 W_x(r)\, dW_x(r)}{X^{*2} + \sigma_{xx}^2 \int_0^1 W_x(r)^2 dr + 2X^* \sigma_{xx} \int_0^1 W_x(r)\, dr}$$

(ii)

$$N\left(\widehat{A} - A\right) \Rightarrow \frac{X^* \sigma_{00} W_0(1) + \sigma_{xx}\sigma_{00} \int_0^1 W_x(r)\, dW_0(r)}{X^{*2} + \sigma_{xx}^2 \int_0^1 W_x(r)^2 dr + 2X^* \sigma_{xx} \int_0^1 W_x(r)\, dr}.$$

Corollary 6.4.2 provides the asymptotics for the LS estimates of β for the continuous-time system (6.4.8)-(6.4.9) with $\kappa = 0$.

Corollary 6.4.2 *We let $\Delta = N^{-\alpha}$, where $\alpha \in (0,1)$. For the continuous-time system (6.4.8)-(6.4.9) with $\kappa = 0$, we assume that there exists $\alpha \in (0,1)$ such that $h^{1-\alpha} T^{\alpha} \to 1$ as $T \to \infty$. Then, we have*

(i)

$$N\sqrt{\Delta}\left(\widehat{a}_\Delta - 1\right) \Rightarrow \frac{\frac{x_0}{\sigma_{xx}} W_x(1) + \int_0^1 W_x(r)\, dW_x(r)}{\sigma_{xx}\left(\left(\frac{x_0}{\sigma_{xx}}\right)^2 + \int_0^1 W_x(r)^2 dr + 2\frac{x_0}{\sigma_{xx}} \int_0^1 W_x(r)\, dr\right)}.$$

(ii)

$$N\sqrt{\Delta}\left(\widehat{\beta} - \beta\right) \Rightarrow \frac{\frac{\sigma_{00}}{\sigma_{xx}}\left(\frac{x_0}{\sigma_{xx}} W_0(1) + \int_0^1 W_x(r)\, dW_0(r)\right)}{\left(\left(\frac{x_0}{\sigma_{xx}}\right)^2 + \int_0^1 W_x(r)^2 dr + 2\frac{x_0}{\sigma_{xx}} \int_0^1 W_x(r)\, dr\right)}.$$

Multivariate co-explosive systems in continuous time

In continuous-time systems with more than one mildly explosive regressor, two cases are examined in Chen et al. (2017): (i) when all the regressors have distinct explosive roots and (ii) when all the regressors share the same explosive root. The limit behavior is different for these two cases.

The multivariate continuous-time system is

$$y(t) = \beta x(t) + u_0(t), \tag{6.4.23}$$
$$dx(t) = \kappa(\mu - x(t))\,dt + \Omega_{xx}^{1/2} dB_x(t),\ x(0) = x_0 = O_p(1),\ \kappa < 0, \tag{6.4.24}$$

where $u_0(t) \sim \mathcal{N}(0, \Omega_{00})$ is Gaussian noise. The driver process $x(t)$ is a multivariate Ornstein–Uhlenbeck process with persistence matrix κ, where $\kappa = diag(\kappa_1, \kappa_2, \ldots \kappa_K)$ is a $K \times K$ diagonal matrix and $\kappa_i < 0$ for $i = 1, \ldots, K$. In addition, $B_x(t) = \Omega_{xx}^{1/2} W_x(t)$, where W_x is standard vector Brownian motion that may be correlated with W_0. Our interest is the co-movement coefficient mtrix β, which is an $m \times K$ matrix.

The exact discrete time representation of (6.4.23)-(6.4.24) is (see Phillips (1972))

$$y_{t\Delta} = \beta x_{t\Delta} + u_{0,t\Delta}, \tag{6.4.25}$$
$$x_{t\Delta} = a_\Delta(\kappa)\, x_{(t-1)\Delta} + g_\Delta + u_{x,t\Delta},\ x_{0\Delta} = x_0 = O_p(1),$$

where

$$a_\Delta(\kappa) = \exp(-\kappa\Delta),$$
$$g_\Delta = \kappa^{-1}\left(I_K - e^{-\kappa\Delta}\right)\kappa\mu,$$
$$u_{x,t\Delta} = \int_{(t-1)\Delta}^{t\Delta} e^{-\kappa(t\Delta - s)}\Omega_{xx} dB_x(s) \overset{d}{=} \mathcal{N}(0, \Omega_{xx}\Delta),$$
$$u_{0,t\Delta} \overset{d}{=} \mathcal{N}(0, \Omega_{00}),$$

since

$$\mathbb{E}\left(u_{x,t\Delta}u_{x,t\Delta}'\right) = \int_{(t-1)\Delta}^{t\Delta} e^{-2\kappa(t\Delta - s)}\Omega_{xx} ds = \frac{1}{2}\kappa^{-1}\left(I_K - e^{-2\kappa\Delta}\right)\Omega_{xx}.$$

Re-standardized it by $\sqrt{\Delta}$, the system becomes

$$y_{t\Delta} = \beta x_{t\Delta} + u_{0,t\Delta}, \tag{6.4.26}$$
$$\widetilde{x}_{t\Delta} = a_\Delta(\kappa)\,\widetilde{x}_{(t-1)\Delta} + \widetilde{g}_\Delta + \widetilde{u}_{x,t\Delta},\ \widetilde{x}_{0\Delta} = \Delta^{-1/2}x_{0\Delta},\ \widetilde{u}_{x,t\Delta} \sim_{iid} \mathcal{N}(0, \Omega_{xx}) \tag{6.4.27}$$

where $\widetilde{x}_{t\Delta} = \Delta^{-1/2}x_{t\Delta}$, $\widetilde{g}_\Delta = \Delta^{-1/2}g_\Delta$ and $\widetilde{u}_{x,t\Delta} = \Delta^{-1/2}u_{x,t\Delta} \overset{d}{=} \mathcal{N}(0, \Omega_{xx})$. The order of the initial value $\widetilde{x}_{0\Delta} = \Delta^{-1/2}x_{0\Delta}$ is $O_p\left(\Delta^{-1/2}\right)$, and the order of the drift term $\widetilde{g}_\Delta$ is $O_p\left(\Delta^{1/2}\right)$.

Distinct explosive roots We let $U_x \overset{d}{=} \mathcal{N}(0, I_K)$, $\widetilde{U}_x = \left(\int_0^\infty e^{p\kappa}\Omega_{xx}e^{p\kappa}dp\right)^{1/2} U_x$, and $D = x_0 - \mu$. For the continuous-time system (6.4.26)-(6.4.27), the double asymptotic theory for the LS estimator of the coefficient matrix β when κ has distinct diagonal elements ($\kappa_i \neq \kappa_j$ for $i \neq j$) is expressed as follows:

$$vec\left\{\frac{1}{\sqrt{\Delta}}\left(\widehat{\beta} - \beta\right)a_\Delta^N\right\} \overset{d}{\to} \left[\left(\int_0^\infty e^{p\kappa}\left(D + \widetilde{U}_x\right)\left(D + \widetilde{U}_x\right)' e^{p\kappa}dp\right)^{-1} \otimes \Omega_{00}\right]^{1/2} \times \mathcal{N}(0, I_{mK}), \tag{6.4.28}$$

assuming that there exists $\alpha \in (0,1)$ such that $\Delta^{1-\alpha}T^\alpha \to 1$ as $T \to \infty$.

The LS estimator of κ is consistent since Δ is known. We let $S_{00} = N^{-1}\sum_{t=1}^N \widehat{u}_{0,t\Delta}\widehat{u}_{0,t\Delta}'$, which satisfies

$$S_{00} \overset{p}{\to} \Omega_{00},$$

and the corresponding estimate of the covariance matrix of $\widehat{\beta}_j$ is

$$S_{\beta_j\beta_j} = \left(\sum_{t=1}^{N} x_{j,t\Delta}^2\right)^{-1} S_{00}.$$

Then, the Wald statistic for testing the full rank restrictions $\mathbb{H}_0 : Q_j\beta_j = r_j$ is

$$W_{\beta_j} = \left\{Q_j\left(\widehat{\beta}_j - \beta_j\right)\right\}'\left(Q_j S_{\beta_j\beta_j} Q_j'\right)^{-1}\left\{Q_j\left(\widehat{\beta}_j - \beta_j\right)\right\} \xrightarrow{d} \chi_g^2,$$

which leads to feasible inference about β_j in the continuous-time framework.

Let a_j be the jth column of $a_\Delta(\kappa)$. The Wald statistic for testing the full rank restrictions $\mathbb{H}_0 : Q_j a_j = q_j$ for given (Q_j, q_j) has the following Chi-squared limit:

$$W_{a_j} := \left\{Q_j\left(\widehat{a}_j - a_j\right)\right\}'\left(Q_j S_{a_ja_j} Q_j'\right)^{-1}\left\{Q_j\left(\widehat{a}_j - a_j\right)\right\} \xrightarrow{d} \chi_g^2,$$

where $S_{a_ja_j} = \left(\sum_{t=1}^N x_{j(t-1)\Delta}^2 - \frac{1}{N}\left(\sum_{t=1}^N x_{j(t-1)\Delta}\right)^2\right)^{-1} S_{xx}$, with $S_{xx} = N^{-1}\sum_{t=1}^N \widehat{u}_{x,t\Delta}\widehat{u}_{x,t\Delta}'$ satisfying $\Delta^{-1}S_{xx} \xrightarrow{p} \Omega_{xx}$, where $\widehat{u}_{x,t\Delta} = x_{t\Delta} - \widehat{a}_\Delta x_{(t-1)\Delta} - \widehat{g}_\Delta$ are regression residuals. We let κ^j denote the jth column of κ. Given the matrix exponential relation, we have the covariance matrix of $\widehat{\kappa}^j$, which satisfies $\Delta^2 S_{\kappa^j\kappa^j} = S_{a_ja_j} + o(\Delta)$. Thus,

$$\left(\widehat{\kappa}^j - \kappa^j\right)e^{-k_jN} \xrightarrow{d} M\mathcal{N}\left(0, \left(\int_0^\infty e^{2p\kappa_j}\left(D_j + \widetilde{U}_{jx}\right)^2 dp\right)^{-1}\Omega_{xx}\right) \overset{d}{=} M\mathcal{N}\left(0, \frac{-2\kappa_j\Omega_{xx}}{\left(D_j + \widetilde{U}_{jx}\right)^2}\right).$$

Then, the Wald statistic for testing the (full rank) restrictions $\mathbb{H}_0 : Q_j\kappa^j = q_j$ satisfies

$$W_{\kappa^j} := \left\{Q_j\widehat{\kappa}^j - q_j\right\}'\left(Q_j S_{\kappa_j\kappa_j} Q_j'\right)^{-1}\left\{Q_j\widehat{\kappa}^j - q_j\right\} \xrightarrow{d} \chi_g^2.$$

Common explosive roots Now we consider the case where the localizing explosive coefficients are identical, namely, $\kappa_i = \kappa$ for $i = 1, \ldots, K$. Let $H_\perp$ be the $K \times (K-1)$ random matrix that is an orthogonal complement to

$$X_c = \left(D + \tilde{U}_x\right) / \left\{\left(D + \tilde{U}_x\right)'\left(D + \tilde{U}_x\right)\right\}^{1/2}, \tag{6.4.29}$$

satisfying $H_\perp H_\perp' = I_K - X_c X_c'$ and with

$$\widetilde{U}_x \equiv \left(\int_0^\infty e^{p\kappa}\Omega_{xx}e^{p\kappa}dp\right)^{1/2} U_x = \Omega_{xx}^{1/2}U_x/(2c)^{1/2}, U_x \overset{d}{=} \mathcal{N}(0, I_K), \text{ and } D = x_0 - \mu.$$

The LS estimator of the coefficient matrix β when κ has common diagonal elements ($\kappa_i = \kappa$ for $i = 1, \ldots, K$.) is expressed as follows:

$$vec\left\{\sqrt{N}\left(\widehat{\beta} - \beta\right)\right\} \xrightarrow{d} \left[H_\perp\left\{H_\perp'\left(\mu\mu' + \frac{1}{-2\kappa}\Omega_{xx}\right)H_\perp\right\}^{-1/2} \otimes \Omega_{00}^{1/2}\right] \times \mathcal{N}(0, I_{mK}), \tag{6.4.30}$$

assuming that there exists $\alpha \in (0,1)$, such that $\Delta^{1-\alpha}T^\alpha \to 1$ as $T \to \infty$.

Then, the Wald statistic for testing $\mathbb{H}_0 : Q\left[vec(\beta)\right] = q$ for full row rank (Q, q) is

$$W_\beta := \left\{Q\widehat{\beta} - q\right\}'\left(QS_{\beta\beta}Q'\right)^{-1}\left\{Q\widehat{\beta} - q\right\} \xrightarrow{d} \chi_g^2,$$

which leads to feasible inference about the matrix coefficient β in the continuous-time framework. Inference about the full matrix β is possible in this case because of the common factorization convergence rate in (6.4.30). The Wald statistics for testing full rank restrictions on $a_\triangle$ and κ such as $\mathbb{H}_0 : Q[vec(a_\triangle)] = q$ and $\mathbb{H}_0 : Q[vec(\kappa)] = q$ are defined in a similar way and have the following chi-squared limits:

$$W_{a_\triangle} := \{Q[vec(\widehat{a}_\triangle)] - q\}' \left(QS_{aa}Q'\right)^{-1} \{Q[vec(\widehat{a}_\triangle)] - q\} \xrightarrow{d} \chi^2_g,$$

and

$$W_\kappa := \{Q[vec(\widehat{\kappa})] - q\}' \left(QS_{\kappa\kappa}Q'\right)^{-1} \{Q[vec(\widehat{\kappa})] - q\} \xrightarrow{d} \chi^2_g,$$

which again lead to feasible inference on $a_\triangle$ and κ because of the common factorization convergence rate. When x_t has a common explosive root, LS estimation by $\widehat{a}_\triangle$ and $\widehat{\kappa}$ produces biased estimates due to endogeneity in the regressor, as shown in Phillips and Magdalinos (2013). The bias distorts the Wald test statistics, and the distortion will be demonstrated in the Monte Carlo simulation below.

6.5 Monte Carlo studies

This section examines the performance of the double asymptotic limit theory for both cointegration ($\kappa = 0$) and co-explosive ($\kappa = -2$) systems in simulations.

Data are generated from the model (6.4.12)-(6.4.13) with $\sigma_{00} = \sigma_{xx} = 1$ and $\mu = 0$, and three sampling intervals are considered, namely, $\Delta = 1/12$, $1/52, 1/252$, which correspond to monthly, weekly and daily frequencies, respectively. The initial value x_0 is set to 3.5, and time spans of $T = 4$ and $T = 10$ years are considered. The percentiles at levels $\{1\%, 2.5\%, 10\%, 90\%, 97.5\%, 99\%\}$ are reported in the limit distribution (6.4.19) and the finite sample distribution of the coefficient-based test (called the C test hereafter) for both the cointegration ($N\sqrt{\Delta}\left(\widehat{\beta} - \beta\right)$) and co-explosive systems ($\frac{a^N}{-2\kappa\sqrt{\Delta}}\left(\widehat{\beta} - \beta\right)$). In addition, the densities of the limit distributions and finite sample distributions of the C test statistic are compared. The number of replications is 10,000.

Table 6.1 [1] reports the percentiles when $x_0 = 3.5$ by using the true values of κ and μ. Both the double asymptotic distribution and the finite sample distribution are sensitive to changes in the initial condition for the C test. In all cases, the double asymptotic limit distribution provides a good approximation to the finite sample distribution.

Figure 6.2 plots the densities of the C test statistic when $T = 4$.

The first row corresponds to the C test for cointegration cases. The second row corresponds to the C test for co-explosive cases. These plots show that the limit distribution well approximates the finite sample distribution for C tests.

6.6 Empirical illustrations

We conduct an empirical study of the relationship between the US nationwide real estate market and metropolitan real estate markets during an explosive period and a normal period by applying the limit theory for univariate co-moving systems (6.4.8)-(6.4.9) for $\kappa < 0$ (coexplosive period) and $\kappa = 0$ (cointegration period) to real estate data using the S&P/Case-Shiller home price composite 20-city index and metropolitan area indices. The S&P/Case-Shiller home price indices are the leading measures of U.S. residential real estate prices and track changes in the value of residential real estate nationwide. Monthly data for the composite 20-city index and 20 metropolitan area indices were downloaded from the St. Louis Fed.[2] We apply the logarithmic transformation to all data. With monthly data, the sampling interval is set to $\Delta = 1/12$.

[1]The result in Table 6.1 for the co-explosive system is from Chen et al. (2017).

[2]http://research.stlouisfed.org/fred2/release?rid=199

	Time Span	T = 4						T = 10					
$\kappa = -2$	C test	1%	2.50%	10%	90%	97.50%	99%	1%	2.50%	10%	90%	97.50%	99%
Daily	Double Asymptotics	-1.0810	-0.9316	-0.6634	0.6676	1.3653	1.8028	-2.1942	-1.7964	-1.1937	0.5512	1.2872	1.7229
(h=1/252)	Finite Sample	-1.0913	-0.9366	-0.6625	0.6609	1.3540	1.7802	-2.1824	-1.7224	-1.1627	0.5783	1.3165	1.8168
Weekly	Double Asymptotics	-1.1036	-0.9676	-0.6810	0.7114	1.4448	1.9358	-2.0883	-1.7241	-1.1884	0.6124	1.3744	1.8624
(h=1/52)	Finite Sample	-1.0913	-0.9366	-0.6625	0.6609	1.3540	1.7802	-2.1824	-1.7224	-1.1627	0.5783	1.3165	1.8168
Monthly	Double Asymptotics	-1.1164	-0.9601	-0.6819	0.6699	1.3325	1.8141	-2.1003	-1.7457	-1.2049	0.5910	1.3528	1.9057
(h=1/12)	Finite Sample	-1.0913	-0.9366	-0.6625	0.6609	1.3540	1.7802	-2.1824	-1.7224	-1.1627	0.5783	1.3165	1.8168
$\kappa = 0$	C test												
Daily	Double Asymptotics	-0.3537	-0.2943	-0.1874	0.1866	0.2952	0.3562	-0.3537	-0.2943	-0.1874	0.1866	0.2952	0.3562
(h=1/252)	Finite Sample	-0.3576	-0.2817	-0.1826	0.1818	0.2897	0.3558	-0.3748	-0.3025	-0.1890	0.1865	0.2924	0.3485
Weekly	Double Asymptotics	-0.3537	-0.2943	-0.1874	0.1866	0.2952	0.3562	-0.3537	-0.2943	-0.1874	0.1866	0.2952	0.3562
(h=1/52)	Finite Sample	-0.3498	-0.2870	-0.1833	0.1835	0.2920	0.3464	-0.3385	-0.2864	-0.1859	0.1834	0.2858	0.3505
Monthly	Double Asymptotics	-0.3537	-0.2943	-0.1874	0.1866	0.2952	0.3562	-0.3537	-0.2943	-0.1874	0.1866	0.2952	0.3562
(h=1/12)	Finite Sample	-0.3317	-0.2828	-0.1766	0.1706	0.2668	0.3186	-0.3157	-0.2595	-0.1686	0.1783	0.2738	0.3337

Table 6.1: Comparison of the finite sample and double asymptotic distributions of $\hat{\beta}$ when the initial value is $x_0 = 3.5$.

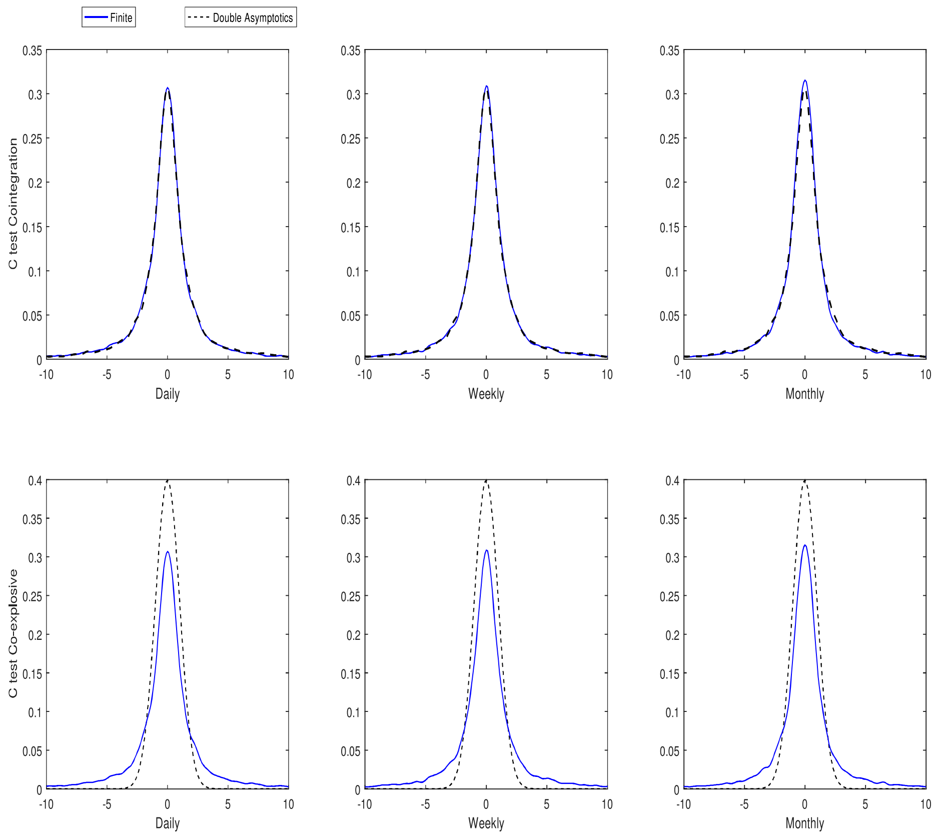

Figure 6.2: Density comparison of the C test for both finite sample and limit distributions when the initial value is $x_0 = 3.5$.

Co-explosive sample period:

The selected explosive sample period is from April 2000 and April 2006. In this case, $N = 73$, $\Delta = 1/12$, and hence, $T = 5.25$. The initial value in each equation of the system is set to the value of the composite 20 index in January 2000, namely, $x_0 = \log(162.01) = 5.0877$.

Cointegration sample period:

The selected cointegration sample period is from August 2008 and August 2014. In this case, $N = 73$, $\Delta = 1/12$, and hence, $T = 5.25$. The initial value in each equation of the system is set to the value of the composite 20 index in August 2008, namely, $x_0 = \log(104) = 4.6444$.

A multi-equation continuous-time system (6.4.8)-(6.4.9) is estimated with x_t as the logarithmic composite 20 index (called the countrywide index in the present paper) and each y_t being one of the logarithmic metropolitan area indices. The coefficient β then measures the co-movement of each metropolitan area index with the countrywide index.

The steps for the empirical studies are as follows:

Step 1:

The occurrence of explosive/normal behavior in the market index ($x_{t\Delta}$) and in the individual area indices ($y_{t\Delta}$) is investigated by estimating κ and κ_y and obtaining the 99% and 90% confidence intervals based on the C test confidence intervals.

Step 2:

For the areas that exhibit explosive/normal behavior, the possible co-movement with the market index is studied further. Allowing for possible weak dependence in u_{0t}, the variance and long-run variances of u_{0t} are estimated by

$$\begin{aligned}\widehat{\sigma}_{00}^2 &= \frac{1}{N}\sum_{\Delta=1}^{N}\widehat{u}_{0t}\widehat{u}_{0t}', \quad \widehat{\varpi}_{00}^2 = \widehat{\sigma}_{00}^2 + 2\widehat{\lambda}_{00}^2,\\ \widehat{\lambda}_{00}^2 &= \frac{1}{N}\sum_{\Delta=1}^{M_n}\left(1-\frac{\Delta}{M_n+1}\right)\sum_{t=\Delta+1}^{N}\widehat{u}_{0t}\widehat{u}_{0t-\Delta}',\end{aligned}$$

with truncation lag $M_n = N^{1/3}$.

Results from Step 1 are reported in Table 6.2 [3]. First, in the explosive period, for the countrywide index, the LS estimate of κ is -0.1184. Its 90% confidence interval is $[-0.1202, -0.1168]$ based on the C test. The results confirm explosive behavior in $x_{t\Delta}$ over this period. For all individual area indices, explosive behavior is found based on both the 90% and the 99% confidence intervals by the C test. Second, in the normal period, for the countrywide index, the LS estimate of κ is -0.0088. Its 90% confidence interval is $[-0.0144, -0.0022]$ based on the C test. The results confirm unit root behavior in $x_{t\Delta}$ over this period. The unit root behavior is found for all individual area indices based on both the 90% and the 99% confidence intervals by the C test.

Now we study the possible co-movement with the market index. Based on the C test, Table 6.3 reports the LS estimates of β and the 99% and 90% confidence intervals. The null hypothesis $H_0 : \beta = 0$ is comfortably rejected in all cases. The confidence intervals can also be used to assess whether $\beta = 1$ versus $\beta < 1$ or $\beta > 1$. If $\beta > 1$ (respectively, $\beta < 1$), we claim that the index of the associated metropolitan area moves faster (slower) than the national index, which provides a useful perspective on the relationships of different metropolitan area indices to the national index. The results show that cities such as LA, Miami, SanDiego, DC, Boston, and NY have more "aggressive" real estate markets than the whole U.S. in both the explosive sample period and in the normal sample period.

[3] The result in Table 6.2 for the co-explosive system is from Chen et al. (2017).

	Co-explosive Periods					Cointegration Periods				
Place	κ or κ_y	C test 99% CI		C test 90% CI		κ or κ_y	C test 99% CI		C test 90% CI	
Market	-0.1184	-0.1212	-0.1159	-0.1202	-0.1168	-0.0088	-0.0172	0.0025	-0.0144	-0.0022
SF	-0.0024	-0.0044	-0.0017	-0.0034	-0.0019	-0.0011	-0.0150	0.0213	-0.0106	0.0112
LA	-0.0036	-0.0050	-0.0029	-0.0043	-0.0031	-0.0007	-0.0115	0.0147	-0.0080	0.0082
LasVegas	-0.0031	-0.0119	-0.0020	-0.0056	-0.0023	0.0006	-0.0135	0.0235	-0.0090	0.0132
Miami	-0.0040	-0.0053	-0.0034	-0.0047	-0.0036	-0.0001	-0.0109	0.0155	-0.0074	0.0089
Phoenix	-0.0036	-0.0083	-0.0027	-0.0054	-0.0029	-0.0001	-0.0142	0.0228	-0.0097	0.0125
SanDiego	-0.0027	-0.0055	-0.0019	-0.0040	-0.0021	-0.0008	-0.0114	0.0143	-0.0080	0.0079
Denver	-0.0009	-0.0012	-0.0008	-0.0011	-0.0008	-0.0006	-0.0064	0.0063	-0.0045	0.0036
DC	-0.0032	-0.0044	-0.0025	-0.0038	-0.0027	-0.0003	-0.0076	0.0090	-0.0052	0.0053
Chicago	-0.0017	-0.0019	-0.0015	-0.0018	-0.0016	0.0007	-0.0077	0.0118	-0.0049	0.0072
Boston	-0.0016	-0.0023	-0.0012	-0.0020	-0.0013	-0.0003	-0.0061	0.0067	-0.0041	0.0040
Charlotte	-0.0007	-0.0008	-0.0006	-0.0008	-0.0006	0.0001	-0.0078	0.0104	-0.0051	0.0062
Portland	-0.0022	-0.0031	-0.0017	-0.0027	-0.0019	0.0001	-0.0085	0.0115	-0.0057	0.0068
Dallas	-0.0006	-0.0007	-0.0005	-0.0007	-0.0006	-0.0006	-0.0065	0.0066	-0.0045	0.0038
Detroit	-0.0007	-0.0008	-0.0006	-0.0008	-0.0007	-0.0002	-0.0125	0.0186	-0.0086	0.0103
Atlanta	-0.0009	-0.0010	-0.0008	-0.0009	-0.0009	0.0003	-0.0110	0.0168	-0.0074	0.0097
Seattle	-0.0021	-0.0027	-0.0017	-0.0025	-0.0018	0.0002	-0.0084	0.0116	-0.0056	0.0069
Minneapolis	-0.0017	-0.0020	-0.0015	-0.0019	-0.0015	0.0001	-0.0114	0.0170	-0.0077	0.0096
Tampa	-0.0035	-0.0048	-0.0028	-0.0042	-0.0030	0.0004	-0.0094	0.0140	-0.0062	0.0083
Cleveland	-0.0006	-0.0007	-0.0006	-0.0007	-0.0006	0.0001	-0.0058	0.0073	-0.0038	0.0045
NY	-0.0026	-0.0029	-0.0023	-0.0028	-0.0024	0.0004	-0.0048	0.0065	-0.0030	0.0042

Table 6.2: Estimated persistence parameter values in y_{th} and x_{th} and confidence intervals for the persistence parameter in the U.S. logarithmic real estate market.

	Co-explosive Periods					Cointegration Periods				
City	β	C test 99% CI		C test 90% CI		β	C test 99% CI		C test 90% CI	
SF	1.0697	1.0684	1.0710	1.0688	1.0705	0.9863	0.9856	0.9871	0.9858	0.9869
LA	1.1731	1.1683	1.1781	1.1700	1.1763	1.2134	1.2129	1.2139	1.2131	1.2138
LasVegas	1.0467	1.0420	1.0515	1.0436	1.0497	0.7354	0.7348	0.7360	0.7350	0.7358
Miami	1.1329	1.1284	1.1376	1.1300	1.1359	1.0371	1.0368	1.0374	1.0369	1.0373
Phoenix	0.9341	0.9304	0.9380	0.9317	0.9366	0.7993	0.7987	0.7999	0.7989	0.7996
SanDiego	1.2087	1.2048	1.2127	1.2062	1.2112	1.1099	1.1094	1.1104	1.1096	1.1102
Denver	0.8288	0.8237	0.8339	0.8255	0.8320	0.8856	0.8854	0.8859	0.8855	0.8858
DC	1.1350	1.1318	1.1383	1.1329	1.1371	1.2530	1.2526	1.2534	1.2527	1.2533
Chicago	0.8843	0.8818	0.8868	0.8826	0.8859	0.8129	0.8124	0.8134	0.8126	0.8132
Boston	1.0058	1.0029	1.0088	1.0039	1.0076	1.0510	1.0508	1.0513	1.0509	1.0512
Charlotte	0.7220	0.7174	0.7266	0.7190	0.7249	0.7905	0.7902	0.7908	0.7903	0.7907
Portland	0.8285	0.8263	0.8307	0.8271	0.8298	0.9888	0.9885	0.9891	0.9886	0.9889
Dallas	0.7481	0.7430	0.7533	0.7448	0.7514	0.8135	0.8133	0.8136	0.8134	0.8136
Detroit	0.7676	0.7629	0.7724	0.7646	0.7706	0.5244	0.5241	0.5247	0.5242	0.5246
Atlanta	0.7732	0.7688	0.7777	0.7704	0.7760	0.7014	0.7010	0.7018	0.7011	0.7016
Seattle	0.8450	0.8427	0.8474	0.8436	0.8465	0.9917	0.9914	0.9920	0.9915	0.9919
Minneapolis	0.9437	0.9408	0.9468	0.9418	0.9456	0.8289	0.8287	0.8290	0.8288	0.8290
Tampa	1.0154	1.0131	1.0177	1.0139	1.0168	0.9442	0.9439	0.9446	0.9440	0.9444
Cleveland	0.7441	0.7396	0.7487	0.7412	0.7469	0.6867	0.6864	0.6871	0.6865	0.6870
NY	1.0483	1.0475	1.0492	1.0478	1.0489	1.1363	1.1357	1.1370	1.1359	1.1367

Table 6.3: Estimated β coefficients and confidence intervals for β in the U.S. logarithmic real estate market

6.7 Conclusions

This chapter reviewed recent developments in nonstationary continuous-time models. For I(1) continuous-time models, the main focus has been on the I(1) OU process, the associated cointegrated system, and the error correction model. Several estimators have been introduced in the literature, and their finite sample performance and the applications have been reported. For explosive continuous-time systems, the focus has been on the explosive OU model and the co-explosive model. The limit distributions of the parameters of interest depend explicitly on the initial conditions. This dependence mimics a corresponding property in the finite sample distribution and, thus, improves the quality of the double asymptotic limit theory as a finite sample approximation. The localized coefficient c in a discrete time explosive model, whose counterpart in continuous time is $-\kappa$, is consistently estimable in continuous time using the LS estimator, thereby facilitating the use of a coefficient-based test for mildly explosive behavior. The co-explosive model is useful in studying the co-movement behavior in different financial markets. These constitute promising directions for future empirical research. For future theoretical research, the development of nonstationary continuous-time models with mixed frequency data is a possible direction.

6.8 Appendix

This Online Supplement provides proofs of the main results in this Chapter.

6.8.1 Proof of Theorem 6.4.2

Proof. (i) We start by expressing x_t in (4.22) as

$$x_t = x_0 + \sum_{j=1}^{t} u_{xj}. \tag{6.8.1}$$

Then, the standardized numerator can be decomposed as

$$N^{-1}\sum_{t=1}^{N} x_t u_{xt} = N^{-1}\sum_{t=1}^{N} u_{xt}x_0 + N^{-1}\sum_{t=1}^{N} u_{xt}\left(\sum_{j=1}^{t} u_{xj}\right). \tag{6.8.2}$$

For the first term on the right-hand side of (6.8.2), since $x_0 N^{\frac{-1}{2}} \Rightarrow X^*$, we have

$$N^{-1}\sum_{t=1}^{N} u_{xt}x_0 \Rightarrow X^*\sigma_{xx}W_x(1), \tag{6.8.3}$$

from Hamilton (2020) (Proposition 17.1 (a)). For the second term on the right hand side of (6.8.2),

$$N^{-1}\sum_{t=1}^{N} u_{xt}\left(\sum_{j=1}^{t} u_{xj}\right) \Rightarrow \sigma_{xx}^2 \int_0^1 W_x(r)\, dW_x(r), \tag{6.8.4}$$

from Hamilton (2020) (Proposition 17.1 (b)). Combining the results of (6.8.3) and (6.8.4) yields

$$N^{-1}\sum_{t=1}^{N} x_t u_{xt} \Rightarrow X^*\sigma_{xx}W_x(1) + \sigma_{xx}^2 \int_0^1 W_x(r)\, dW_x(r).$$

Next, we consider the standardized denominator

$$N^{-2}\sum_{t=1}^{N} x_t^2 = N^{-1}x_0^2 + N^{-2}\sum_{t=1}^{N}\left(\sum_{j=1}^{t} u_{xj}\right)^2 + 2N^{-2}x_0\sum_{t=1}^{N}\left(\sum_{j=1}^{t} u_{xj}\right). \tag{6.8.5}$$

For the first term on the right-hand side of (6.8.5), since $\widetilde{x}_0 = x_{0N}N^{-1/2} \Rightarrow X^*$, we have

$$N^{-1}x_0^2 \Rightarrow X^{*2}. \tag{6.8.6}$$

For the second term on the right-hand side of (6.8.5), we then have

$$N^{-2}\sum_{t=1}^{N}\left(\sum_{j=1}^{t} u_{xj}\right)^2 \Rightarrow \sigma_{xx}^2 \int_0^1 W_x(r)^2\, dr, \tag{6.8.7}$$

from Hamilton (2020) (Proposition 17.1 (e)). For the third term on the right hand side of (6.8.5), we have

$$2N^{-2}x_0\sum_{t=1}^{N}\left(\sum_{j=1}^{t} u_{xj}\right) \Rightarrow 2X^*\sigma_{xx}\int_0^1 W_x(r)\, dr, \tag{6.8.8}$$

from Hamilton (2020) (Proposition 17.1 (d)). Combining the results of (6.8.6)-(6.8.8) yields

$$N^{-2}\sum_{t=1}^{N} x_t^2 \Rightarrow X^{*2} + \sigma_{xx}^2 \int_0^1 W_x(r)^2\, dr + 2X^*\sigma_{xx}\int_0^1 W_x(r)\, dr.$$

Hence, we obtain

$$N\left(\widehat{R}-1\right) \Rightarrow \frac{X^*\sigma_{xx}W_x(1) + \sigma_{xx}^2 \int_0^1 W_x(r)\, dW_x(r)}{X^{*2} + \sigma_{xx}^2 \int_0^1 W_x(r)^2\, dr + 2X^*\sigma_{xx}\int_0^1 W_x(r)\, dr}$$

(ii) By the same argument, we obtain

$$N\left(\widehat{A}-A\right) \Rightarrow \frac{X^*\sigma_{00}W_0(1) + \sigma_{xx}\sigma_{00} \int_0^1 W_x(r)\, dW_0(r)}{X^{*2} + \sigma_{xx}^2 \int_0^1 W_x(r)^2\, dr + 2X^*\sigma_{xx}\int_0^1 W_x(r)\, dr}.$$

■

6.8.2 Proof of Corollary 6.4.2

The proof follows from Theorem 6.4.2 by considering the mappings

$$\sigma_{00}^2 \mapsto \sigma_{00}^2, \sigma_{xx}^2 \mapsto 1, X^* \mapsto \frac{\widetilde{x}^*}{\sigma_{xx}},$$

with $\Delta = N^{-\alpha}$.

For (i), it follows that

$$\begin{aligned} N^{-1}\sum_{t=1}^{N} \tilde{x}_{t\Delta}u_{x,t\Delta} &\Rightarrow \frac{\widetilde{x}^*}{\sigma_{xx}}W_x(1) + \int_0^1 W_x(r)\, dW_x(r), \\ N^{-2}\sum_{t=1}^{N} \tilde{x}_{t\Delta}^2 &\Rightarrow \left(\frac{\widetilde{x}^*}{\sigma_{xx}}\right)^2 + \int_0^1 W_x(r)^2\, dr + 2\frac{\widetilde{x}^*}{\sigma_{xx}}\int_0^1 W_x(r)\, dr, \end{aligned}$$

and hence,

$$\begin{aligned} & N\sqrt{\Delta}\left(\widehat{a}_\Delta(\kappa)-1\right) \\ = & \sqrt{\Delta}\frac{N^{-1}\sum_{t=1}^{N} x_{t\Delta}u_{x,t\Delta}}{N^{-2}\sum_{t=1}^{N} x_{t\Delta}^2} = \frac{N^{-1}\sum_{t=1}^{N} \tilde{x}_{t\Delta}u_{x,t\Delta}}{N^{-2}\sum_{t=1}^{N} \tilde{x}_{t\Delta}^2}\frac{\sqrt{\Delta}}{\lambda_\Delta} \\ \Rightarrow & \frac{\frac{\widetilde{x}^*}{\sigma_{xx}}W_x(1) + \int_0^1 W_x(r)\, dW_x(r)}{\sigma_{xx}\left(\left(\frac{\widetilde{x}^*}{\sigma_{xx}}\right)^2 + \int_0^1 W_x(r)^2\, dr + 2\frac{\widetilde{x}^*}{\sigma_{xx}}\int_0^1 W_x(r)\, dr\right)}. \end{aligned}$$

For (ii), it follows that

$$N^{-1}\sum_{t=1}^{N} \tilde{x}_{t\Delta}u_{0,t\Delta} \Rightarrow \frac{\widetilde{x}^*}{\sigma_{xx}}\sigma_{00}W_0(1) + \sigma_{00}\int_0^1 W_x(r)\, dW_0(r),$$

and hence,

$$\begin{aligned} & N\sqrt{\Delta}\left(\widehat{\beta}-\beta\right) \\ = & \sqrt{\Delta}\frac{N^{-1}\sum_{t=1}^{N} x_{t\Delta}u_{0,t\Delta}}{N^{-2}\sum_{t=1}^{N} x_{t\Delta}^2} = \frac{N^{-1}\sum_{t=1}^{N} \tilde{x}_{t\Delta}u_{0,t\Delta}}{N^{-2}\sum_{t=1}^{N} \tilde{x}_{t\Delta}^2}\frac{\sqrt{\Delta}}{\lambda_\Delta} \\ \Rightarrow & \frac{\frac{\sigma_{00}}{\sigma_{xx}}\left(\frac{\widetilde{x}^*}{\sigma_{xx}}W_0(1) + \int_0^1 W_x(r)\, dW_0(r)\right)}{\left(\left(\frac{\widetilde{x}^*}{\sigma_{xx}}\right)^2 + \int_0^1 W_x(r)^2\, dr + 2\frac{\widetilde{x}^*}{\sigma_{xx}}\int_0^1 W_x(r)\, dr\right)}. \end{aligned}$$

Bibliography

Aït-Sahalia, Y., P. A. Mykland, and L. Zhang (2005). How often to sample a continuous-time process in the presence of market microstructure noise. *The Review of Financial Studies 18*(2), 351–416.

Alessi, L. and M. Kerssenfischer (2019). The response of asset prices to monetary policy shocks: stronger than thought. *Journal of Applied Econometrics 34*(5), 661–672.

Ambrose, B. W., P. Eichholtz, and T. Lindenthal (2013). House prices and fundamentals: 355 years of evidence. *Journal of Money, Credit and Banking 45*(2-3), 477–491.

Bae, Y. and R. M. De Jong (2007). Money demand function estimation by nonlinear cointegration. *Journal of Applied Econometrics 22*(4), 767–793.

Bandi, F. M. and J. R. Russell (2006). Separating microstructure noise from volatility. *Journal of Financial Economics 79*(3), 655–692.

Barndorff-Nielsen, O. E. and N. Shephard (2001). Non-gaussian ornstein–uhlenbeck-based models and some of their uses in financial economics. *Journal of the Royal Statistical Society: Series B (Statistical Methodology) 63*(2), 167–241.

Barndorff-Nielsen, O. E. and N. Shephard (2003). Integrated ou processes and non-gaussian ou-based stochastic volatility models. *Scandinavian Journal of statistics 30*(2), 277–295.

Boswijk, H. P. (2001). Block local to unity and continuous record asymptotics. Tinbergen Institute Discussion Paper, No. 01-078/4.

Boswijk, H. P. and Y. Zu (2022). Adaptive testing for cointegration with nonstationary volatility. *Journal of Business & Economic Statistics 40*(2), 744–755.

Burgess, N. (2014). An overview of the vasicek short rate model. *Available at SSRN 2479671*.

Cai, Z., Y. Wang, and Y. Wang (2015). Testing instability in a predictive regression model with nonstationary regressors. *Econometric Theory 31*(5), 953–980.

Case, K. E. and R. J. Shiller (2003). Is there a bubble in the housing market? *Brookings papers on economic activity 2003*(2), 299–362.

Chambers, M. J. (2001). Temporal aggregation and the finite sample performance of spectral regression estimators in cointegrated systems: a simulation study. *Econometric Theory 17*(3), 591–607.

Chen, Y., P. C. B. Phillips, and S. Shi (2022, 01). Common bubble detection in large dimensional financial systems. *Journal of Financial Econometrics*. nbab027.

Chen, Y., P. C. B. Phillips, and J. Yu (2017). Inference in continuous systems with mildly explosive regressors. *Journal of Econometrics 201*(2), 400–416.

Engle, R. F. and C. W. Granger (1987). Co-integration and error correction: representation, estimation, and testing. *Econometrica: journal of the Econometric Society 55*(2), 251–276.

Ghysels, E. and J. I. Miller (2015). Testing for cointegration with temporally aggregated and mixed-frequency time series. *Journal of Time Series Analysis 36*(6), 797–816.

Granger, C. W. (1981). Some properties of time series data and their use in econometric model specification. *Journal of Econometrics 16*(1), 121–130.

Greenaway-McGrevy, R. and P. C. B. Phillips (2016). Hot property in new zealand: Empirical evidence of housing bubbles in the metropolitan centres. *New Zealand Economic Papers 50*(1), 88–113.

Hamilton, J. D. (2020). *Time series analysis*. Princeton university press.

Hannan, E. J. (1970). *Multiple Time Series*. New York: Wiley.

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

Hansen, P. R. and A. Lunde (2006). Realized variance and market microstructure noise. *Journal of Business and Economic Statistics 24*(2), 127–161.

Henry, O. T. and K. Shields (2004). Is there a unit root in inflation? *Journal of Macroeconomics 26*(3), 481–500.

Hu, Y. and L. Oxley (2018). Bubbles in us regional house prices: evidence from house price–income ratios at the state level. *Applied Economics 50*(29), 3196–3229.

Lo, A. W. and A. C. MacKinlay (1988). Stock market prices do not follow random walks: Evidence from a simple specification test. *The review of financial studies 1*(1), 41–66.

Lui, Y. L., W. Xiao, and J. Yu (2021). The grid bootstrap for continuous time models. *Journal of Business & Economic Statistics 00*(0), 1–13.

McCarthy, J., R. Peach, et al. (2004). Are home prices the next bubble? *Economic Policy Review 10*(3), 1–17.

Pan, W.-F. (2019). Detecting bubbles in china's regional housing markets. *Empirical Economics 56*(4), 1413–1432.

Perron, P. (1991a). 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. (1991b). 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.

Phillips, P. C. B. (1972). The structural estimation of a stochastic differential equation system. *Econometrica: Journal of the Econometric Society 40*(6), 1021–1041.

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. (1991). Error correction and long-run equilibrium in continuous time. *Econometrica 59*(4), 967–980.

Phillips, P. C. B. (2015). Halbert white jr. memorial jfec lecture: Pitfalls and possibilities in predictive regression. *Journal of Financial Econometrics 13*(3), 521–555.

Phillips, P. C. B. and T. Magdalinos (2007). Limit theory for moderate deviations from a unit root. journal of econometrics. *Journal of Econometrics 136*(1), 115–130.

Phillips, P. C. B. and T. Magdalinos (2009). Unit root and cointegrating limit theory when initialization is in the infinite past. *Econometric Theory 25*(06), 1682–1715.

Phillips, P. C. B. and T. Magdalinos (2013). Inconsistent var regression with common explosive roots. *Econometric Theory 29*(4), 808–837.

Phillips, P. C. B., H. R. Moon, and Z. Xiao (2001). How to estimate autoregressive roots near unity. *Econometric Theory 17*(1), 29–69.

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. and J. Yu (2006). Comments on 'Realized Variance and Market Microstructure Noise' by P. R. Hansen and A. Lunde. *Journal of Business and Economic Statistics 24*, 202–208.

Roberts, G. O., O. Papaspiliopoulos, and P. Dellaportas (2004). Bayesian inference for non-gaussian ornstein–uhlenbeck stochastic volatility processes. *Journal of the Royal Statistical Society: Series B (Statistical Methodology) 66*(2), 369–393.

Shi, S., A. Valadkhani, R. Smyth, and F. Vahid (2016). Dating the timeline of house price bubbles in australian capital cities. *Economic Record 92*(299), 590–605.

Stambaugh, R. F. (1999). Predictive regressions. *Journal of Financial Economics 54*(3), 375–421.

Tang, C. Y. and S. X. Chen (2009). Parameter estimation and bias correction for diffusion processes. *Journal of Econometrics 149*(1), 65–81.

Vasicek, O. (1977). An equilibrium characterization of the term structure. *Journal of Financial Economics 5*(2), 177–188.

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

Welch, I. and A. Goyal (2008). A comprehensive look at the empirical performance of equity premium prediction. *The Review of Financial Studies 21*(4), 1455–1508.

Yu, J. (2012). Bias in the estimation of the mean reversion parameter in continuous time models. *Journal of Econometrics 169*(1), 114–122.

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

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

Zhang, L., P. A. Mykland, and Y. Aït-Sahalia (2005). A tale of two time scales: Determining integrated volatility with noisy high-frequency data. *Journal of the American Statistical Association 100*(472), 1394–1411.

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

Zivot, E. and D. W. K. Andrews (2002). Further evidence on the great crash, the oil-price shock, and the unit-root hypothesis. *Journal of business & economic statistics 20*(1), 25–44.