3-2015

Limit Theory for Continuous Time Systems with Mildly Explosive Regressors

Peter C. B. PHILLIPS
Singapore Management University, peterphillips@smu.edu.sg

Chen Ye

Jun Yu
Singapore Management University, yujun@smu.edu.sg

SMU ECONOMICS & STATISTICS
WORKING PAPER SERIES

Limit Theory for Continuous Time Systems with Mildly Explosive Regressors

Peter C. B. Phillips, Ye Chen and Jun Yu

March 2015

Paper No. 03-2015

Limit Theory for Continuous Time Systems with Mildly Explosive Regressors*

Peter C. B. Phillips†, Ye Chen†† and Jun Yu††
†Yale University, University of Auckland,
Singapore Management University & University of Southampton
††Singapore Management University

March 13, 2015

Abstract

Limit theory is developed for continuous co-moving systems with mildly explosive regressors. The theory uses double asymptotics with infill (as the sampling interval tends to zero) and large time span asymptotics. The limit theory explicitly involves initial conditions, allows for drift in the system, is provided for single and multiple explosive regressors, and is feasible to implement in practice. Simulations show that double asymptotics deliver a good approximation to the finite sample distribution, with both finite sample and asymptotic distributions showing sensitivity to initial conditions. The methods are implemented in the US real estate market for an empirical application, illustrating the usefulness of double asymptotics in practical work.

Keywords: Cointegrated system; Explosive Process; Moderate Deviations from Unity; Double Asymptotics; Real Estate Market.

JEL classification: C12, C13, C58

1 Introduction

The recent global financial crisis has motivated econometricians to study potentially explosive behavior in financial time series and develop technologies for the detection of bubbles in

*Phillips acknowledges support from the NSF under Grant No. SES 12-58258. Yu acknowledges financial support from Singapore Ministry of Education Academic Research Fund Tier 2 under the grant number MOE2011-T2-2-096.

financial markets. For example, Phillips, Wu and Yu (2011) and Phillips, Shi and Yu (2015a, b) use mildly explosive representations to capture market exuberance in financial time series and recursive regressions to provide dating algorithms. Both these studies use machinery that draws on work of Phillips and Magdalinos (2007) on mildly explosive processes and the limit theory associated with these processes, which have a growing number of applications in economics and finance. Other recent research has focussed on mechanisms for generating financial bubbles rather than reduced form methods. Among his many wide-ranging contributions to econometrics and finance, Christian Gourieroux has recently explored new ways of generating explosive bubbles via non-causal forward-looking processes (Gourieroux and Zakoian, 2013).

Long run equilibrium relationships among nonstationary variables are often modeled in terms of cointegrated systems. In a typical cointegrated system variables are assumed to be integrated $I(1)$ processes and the model is formulated in discrete time. However, financial applications often use continuous time representations, given the presence of high frequency observations, making these representations popular in empirical work. Phillips (1991) showed how to formulate a cointegrated system in continuous time and proposed an inferential procedure for such systems based on frequency domain techniques. That work maintained the usual $I(1)$ process assumption, thereby excluding episodes of exuberance in the data.

Extending the framework of co-movement in data to mildly explosive variables, Magdalinos and Phillips (2009, MP hereafter) developed a generalized cointegrated system with multiple variables that may be mildly explosive, leading to mixed normal limit theory and mildly explosive rates of convergence, just as in the univariate autoregression studied in Phillips and Magdalinos (2007). Like autoregressive roots that are local to unity, mildly explosive roots depend on the sample size but deliver parameterizations that lie in a wider vicinity of unity. The limit theory in such systems is independent of the initial condition when, as is often the case, the initialization is assumed to be asymptotically negligible relative to the order of the sample observations. Other cases, where the initialization is non-negligible and may figure in the limit theory in various ways are considered in other work (Andrews and Guggenberger, 2012; Phillips and Magdalinos, 2009).

In a recent study, Wang and Yu (2014, WY hereafter) developed a double asymptotic theory for an explosive continuous time model, where the sampling interval passes to zero and the time span passes to infinity. In this double asymptotic setting, the explosive continuous time model implies mildly explosive behavior in discrete time but with an autoregressive parameter that depends on the sampling frequency, not the sample size, by virtue of the discrete time solution of the continuous system. 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 sampling frequency often provides greater realism in empirical work where it is necessary to model near unit root phenomena in discrete time. The limit theory in WY contains a term that explicitly depends on initial conditions, thereby differing from the (large span) limit theory in Phillips and Magdalinos (2007). This difference arises from the different order of magnitude implied for the initial conditions in the two approaches. Simulations in WY reveal that double asymptotics involving initial condition dependencies typically outperform in finite samples the asymptotics that are free of the initial condition. The changes in the limit theory induced by these initial condition dependences are sufficient to materially change conclusions in empirical work.

This paper extends work by Phillips (1991) on continuous system cointegration by developing asymptotics for continuous models where the variables are mildly explosive. The model differs from MP's mildly explosive system in three ways. First, our model is formulated and parameterized in continuous time whereas MP uses a discrete time specification. This difference is important because the implied (discrete time) autoregressive parameter of the continuous system depends on the known sampling frequency, not on the sample size in terms of an unknown localizing coefficient. Pivotal limit theory is therefore possible in the continuous time formulation. Second, the initial conditions in the two models are different. Third, the continuous time model allows for a drift in the regressor, which affects the limit theory. In developing double asymptotics, we utilize the limit theory of MP while adjusting for the initial condition, the drift, and the autoregressive specification, all of which affect the resulting limit distribution.

There are good reasons for extending discrete time cointegrated systems to continuous time. Continuous time models now enjoy a wide range of empirical applications both in macroeconomics and financial economics. They provide for discrete sampling at any frequency, including intermittent random sampling, and they allow for convenient handling of both stock variables and flow variables, the latter by simple time aggregation. Importantly in the present setting, the use of a continuous time framework readily accommodates initial condition and drift effects, with a limit theory that is easy to implement in practice with no nuisance parameters. In particular, the limit theory in the continuous system here depends on a persistence parameter (κ) which is consistently estimable. By 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.

The paper is organized as follows. Section 2 introduces the model and gives our main results, providing connections between the continuous time framework considered here and the discrete time cointegrated systems in MP. The limit theory of MP is modified to allow for a discrete time model with initial condition and drift induced by the continuous system, which assists in delivering double asymptotics for the least squares estimator in the continuous system. Section 3 extends the limit results to the multivariate setting. Section 4 reports simulations studying the finite sample performance of the methods. An empirical application of the methodology to US real estate data is given in Section 4. Section 5 concludes. Proofs are given in the Appendix.

2 Continuous Systems with a Mildly Explosive Regressor

We start our investigation with the following scalar continuous time model in the two variates $y\left(t\right)$ and $x\left(t\right)$

$$dy(t) = \beta dx(t) + \sigma_{00} dB_0(t), \tag{2.1}$$

$$dx\left(t\right) = \kappa\left(\mu - x\left(t\right)\right)dt + \sigma_{xx} dB_x(t),\ x\left(0\right) = x_0 = O_p\left(1\right),\ \kappa < 0, \tag{2.2}$$

where $B_0(t)$ and $B_x(t)$ are two correlated standard Brownian motions. The parameter of central interest for inference is the coefficient β which captures the co-movement between $y(t)$ and $x(t)$. The driver process $x(t)$ follows an Ornstein–Uhlenbeck equation with persistence parameter κ. For $\kappa > 0$ the process $x(t)$ is stationary, for $\kappa = 0$ it is Brownian motion, and for $\kappa < 0$ it is explosive. For data over a large time span several different regimes of κ might be contemplated, possibly with break points separating the regimes. The present paper focuses on the explosive case of $\kappa < 0$. The scalar model is important particularly in financial applications and leads to simple results that avoid some of the complications of systems with multiple explosive regressors, which are considered in the next section.

Suppose data are recorded at N equally spaced points, $\{th\}_{t=1}^{N}$, over a time interval $[0,T]$, with sampling interval h and overall time span T so that $N = T/h$. To develop asymptotics we assume that both $h \to 0$ and $T \to \infty$. The exact discrete time representation of (2.1)-(2.2) is (Phillips, 1972)

$$y_{th} = \beta x_{th} + u_{0,th}, \tag{2.3}$$

$$x_{th} = a_h\left(\kappa\right) x_{(t-1)h} + g_h + u_{x,th},\ x_{0h} = x_0 = O_p\left(1\right), \tag{2.4}$$

where

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

The autoregressive parameter $a_h(\kappa) = \exp(-\kappa h)$ depends directly on the sampling frequency h. Indirectly, h and $a_h(\kappa)$ are both related to the sample size N. When T is fixed, $h = T/N = O\left(N^{-1}\right) \to 0$, and when $T \to \infty$, $h = O(T/N) \to 0$. Gaussianity follows from the Brownian motion driver processes in (2.1)-(2.2). The standard error of $u_{x,th}$ is $\lambda_h = \sigma_{xx}\sqrt{\left(1 - e^{-2\kappa h}\right)/2\kappa} \sim \sigma_{xx}\sqrt{h} \to 0$, concordant with the sample path continuity of $x(t)$.

Re-standardizing the equations (2.3)-(2.4) by λ_h we have

$$\widetilde{y}_{th} = \beta\widetilde{x}_{th} + \widetilde{u}_{0,th}, \tag{2.5}$$

$$\widetilde{x}_{th} = a_h(\kappa)\widetilde{x}_{(t-1)h} + \widetilde{g}_h + \widetilde{u}_{x,th}, \quad \widetilde{x}_{0h} = x_{0h}/\lambda_h, \ \widetilde{u}_{x,th} \overset{iid}{\sim} N(0,1), \tag{2.6}$$

where $\widetilde{y}_{th} = y_{th}/\lambda_h$, $\widetilde{x}_{th} = x_{th}/\lambda_h$, $\widetilde{g}_h = g_h/\lambda_h$, $\widetilde{u}_{0,th} = u_{0,th}/\lambda_h \overset{d}{=} N\left(0, \widetilde{\sigma}_{00}^2\right)$, and $\widetilde{\sigma}_{00}^2 = h\sigma_{00}^2/\lambda_h^2$. Clearly, $\widetilde{\sigma}_{00}^2 \to \sigma_{00}^2/\sigma_{xx}^2$ as $h \to 0$. When $T \to \infty$ and $h \to 0$, we have

$$\frac{1}{Nh} = \frac{1}{T} \to 0, \ a_h(\kappa) = \exp\{-\kappa h\} = 1 - \kappa h + O\left(h^2\right) \to 1.$$

Hence $\widetilde{x}_{th}$ in (2.6) is a mildly explosive process as in Phillips and Magdalinos (2007). Furthermore, since $\kappa < 0$, when $h \to 0$ we have

$$\begin{aligned}
\widetilde{x}_{0h} &= x_{0h}/\lambda_h = O_p\left(h^{-1/2}\right) \text{ since } \lambda_h \sim \sqrt{h}, \\
(a_h(\kappa))^{-N} &= \exp\{\kappa h N\} = \exp\{\kappa T\} = o(1/T), \\
\widetilde{g}_h &\sim O\left(\sqrt{h}\right).
\end{aligned}$$

Thus, in the standardized discrete system (2.5)-(2.6) the order of magnitude of the initial condition is $\widetilde{x}_{0h} \sim O_p\left(h^{-1/2}\right)$ while in the original system (2.1)-(2.2) it is $x_0 \sim O_p(1)$. In addition, the order of magnitude of the drift is $O\left(\sqrt{h}\right)$ in model (2.6) but is $O_p(1)$ in (2.2).

MP (2009) analyzed the triangular system

$$y_t = Ax_t + u_{0t}, \tag{2.7}$$

$$x_t = R_N x_{t-1} + u_{xt}, \quad x_0 = o_p(N^{\alpha/2}), \tag{2.8}$$

where $R_n = I_K + \frac{C}{N^\alpha}$, $\alpha \in (0,1)$, $C = \text{diag}\,(c_1, \ldots c_K)$, and discrete observations $\{y_t, x_t\}_{t=0}^N$ are available. In this system, A is the matrix of cointegrating (or, more specifically in the development below, co-mildly explosive) coefficients; R_N represents moderate deviations from a unit root in the sense of Phillips and Magdalinos (2007); x_t is a moderately integrated time series as $N^\alpha \to \infty$ when $N \to \infty$. If $C > 0$, x_t is a mildly explosive time series. The vector (u_{0t}, u_{xt}) is a sequence of zero mean, weakly dependent linear process errors which satisfy certain standard regularity conditions. The analysis of MP covers both cases $C > 0$ and $C < 0$, our focus here is on asymptotics for the mildly explosive case $C > 0$.

There are some common features in model (2.5)-(2.6) and the MP model (2.7)-(2.8): both systems imply co-movement between y and x, and in both models x_t may be mildly explosive. But there are also important differences between these discrete time systems. First, the moderate deviations from unity in the autoregressive coefficient take different forms: in (2.6) the autoregressive coefficient is a function of the sampling interval h, whereas in (2.8) it is formulated as a function of the overall sample size N. A second difference is that, while in (2.8) the initial condition for x_t is assumed to be $o_p(N^{\alpha/2}) = o_p\left(h^{-1/2}\right)$, in (2.6) it is $O_p\left(h^{-1/2}\right)$, which translates to $x_0 = O_p(1)$ in the original continuous time system (2.1)-(2.2). So, the initial condition in (2.6) has the larger order of magnitude $O_p\left(h^{-1/2}\right)$, which corresponds to a distant past initialization in the terminology of Phillips and Magdalinos (2009), where it is shown that such initializations do affect the limit theory. The third difference in the models occurs in the drift. In (2.8) there is no intercept, and if a constant intercept were present it would typically dominate the asymptotics. By contrast, in (2.6) the intercept

$$\widetilde{g}_h = g_h/\lambda_h = \mu\left(1 - e^{-\kappa h}\right)/\lambda_h = \frac{\mu\left(1 - e^{-\kappa h}\right)}{\sigma_{xx}\sqrt{\left(1 - e^{-2\kappa h}\right)/2\kappa}} = \frac{\mu\kappa h}{\sigma_{xx} h^{1/2}}\left\{1 + o\left(h\right)\right\} = O\left(\sqrt{h}\right)$$

is asymptotically negligible as $h \to 0$, so the intercept does not affect the double asymptotics. The limit theory of MP (2009) is readily modified to take into account this new initial condition and drift.

To fix ideas, consider the modified MP model

$$y_t = Ax_t + u_{0t}, \tag{2.9}$$

$$x_t = \mu + R_N x_{t-1} + u_{xt}, x_0 = x_{0N} = O_p\left(N^{\alpha/2}\right), \mu = O_p\left(N^{-\alpha/2}\right). \tag{2.10}$$

Let $\widetilde{x}_0 = x_{0N} N^{-\alpha/2} \Rightarrow X^*$ and $\widetilde{\mu} = N^{\alpha/2}\mu \Rightarrow \mu^*$. The error $u_t = [u_{0t}, u_{xt}]'$ is an i.i.d. sequence with mean zero and covariance $\begin{bmatrix} \sigma_{00}^2 & \sigma_{0x} \\ \sigma_{0x} & \sigma_{xx}^2 \end{bmatrix}$. This model extends (2.7)-(2.8) by allowing for a larger initial condition and a (local to zero) drift. The following theorem gives the limit theory for the LS estimator of A in (2.9) for the case of a single scalar regressor x_t.

Theorem 2.1 *For the discrete time system (2.9)-(2.10) with* $R_N = 1 + \frac{c}{N^\alpha}$, $\alpha \in (0,1)$, *and* $c > 0$, *when* $N \to \infty$, *we have*

(i) $\left(R_N^N N^\alpha\right)^{-1} \sum_{t=1}^N x_t u_{0t} \Rightarrow \frac{\sigma_{00}}{2c} U_0 \left(\sigma_{xx} U_x + (2c)^{1/2} D\right)$,

(ii) $\left(R_N^N N^\alpha\right)^{-2} \sum_{t=1}^N x_t^2 \Rightarrow \left(\frac{1}{2c}\right)^2 \left(\sigma_{xx} U_x + (2c)^{1/2} D\right)^2$,

where $(U_0, U_x) \overset{d}{=} N(0, I_2)$, $D = X^* + \frac{\mu^*}{c}$, *and so*

$$R_N^N N^\alpha \left(\widehat{A} - A\right) \Rightarrow 2c \frac{\sigma_{00} U_0}{\sigma_{xx} U_x + (2c)^{1/2} D}. \tag{2.11}$$

Remark 1 *If* $\widetilde{x}_0 = -\frac{\widetilde{\mu}}{c}$, *then* $D = 0$ *and the limit (2.11) is simply*

$$R_N^N N^\alpha \left(\widehat{A} - A\right) \Rightarrow 2c \frac{\sigma_{00} U_0}{\sigma_{xx} U_x} = 2c \frac{\sigma_{00}}{\sigma_{xx}} \mathbb{C}, \tag{2.12}$$

where $\mathbb{C}$ *is a standard Cauchy variate. This limit distribution is the same as that given by MP (2009, p. 496) and depends on the localizing coefficient c, although the standardized estimation error satisfies*

$$\frac{R_N^N}{R_N^2 - 1} \left(\widehat{A} - A\right) \Rightarrow \frac{\sigma_{00}}{\sigma_{xx}} \mathbb{C}, \tag{2.13}$$

when $D = 0$ *and this limit does not depend on c. In the general case where* $D \neq 0$

$$\frac{R_N^N}{R_N^2 - 1} \left(\widehat{A} - A\right) \Rightarrow \frac{\sigma_{00} U_0}{\sigma_{xx} U_x + (2c)^{1/2} D}. \tag{2.14}$$

Remark 2 *The limit distribution of* $\widehat{\mu}$, *the LS estimator of the intercept parameter* μ, *follows simply as*

$$\begin{aligned}
& \sqrt{N}\left(\widehat{\mu} - \mu\right) \\
= \quad & \sqrt{N} \frac{\left(\sum_{t=1}^N x_{t-1}^2\right)\left(\sum_{t=1}^N u_{xt}\right) - \left(\sum_{t=1}^N x_{t-1}\right)\left(\sum_{t=1}^N x_{t-1} u_{xt}\right)}{N\left(\sum_{t=1}^N x_{t-1}^2\right) - \left(\sum_{t=1}^N x_{t-1}\right)^2} \\
= \quad & \sqrt{N} \frac{\left(R_N^N N^\alpha\right)^{-2} \sum_{t=1}^N x_{t-1}^2 \left(\frac{1}{N} \sum_{t=1}^N u_{xt}\right) - \frac{1}{N}\left\{\left(R_N^N N^\alpha\right)^{-1} \sum_{t=1}^N x_{t-1}\right\}\left\{\left(R_N^N N^\alpha\right)^{-1} \sum_{t=1}^N x_{t-1} u_{xt}\right\}}{\frac{N}{N}\left\{\left(R_N^N N^\alpha\right)^{-2} \sum_{t=1}^N x_{t-1}^2\right\} - \frac{1}{N}\left\{\left(R_N^N N^\alpha\right)^{-1} \sum_{t=1}^N x_{t-1}\right\}^2} \\
= \quad & \frac{1}{\sqrt{N}} \sum_{t=1}^N u_{xt} + o_p(1) \Rightarrow N\left(0, \sigma_{xx}^2\right).
\end{aligned}$$

This result is useful in testing for $\mu = 0$ *in the modified MP model.*

Remark 3 *Self normalized statistics based on* $\widehat{A}$ *have a much simpler limit theory that is convenient for inference. For instance, defining the regression residuals* $\hat{u}_{0t} = y_t - \widehat{A} x_t$ *and* *noting that the residual variance estimate* $s_0^2 = N^{-1}\sum_{t=1}^N \hat{u}_{0t}^2 \xrightarrow{p} \sigma_{00}^2$*, it follows immediately from Theorem 2.1 that the usual t statistic for testing* $\mathbb{H}_0 : A = A^0$ *satisfies*

$$\begin{aligned} t_A &= \frac{\widehat{A}-A^0}{s_A} = \frac{\widehat{A}-A^0}{\left\{s_0^2\left(\sum_{t=1}^N x_t^2\right)^{-1}\right\}^{1/2}} = \frac{R_N^N N^\alpha\left(\widehat{A}-A^0\right)}{\left\{s_0^2\left(\frac{1}{R_N^{2N}N^{2\alpha}}\sum_{t=1}^N x_t^2\right)^{-1}\right\}^{1/2}} \\ &\Rightarrow \frac{2c\frac{\sigma_{00}U_0}{\sigma_{xx}U_x+(2c)^{1/2}D}}{\sigma_{00}\left\{\left(\frac{1}{2c}\right)^2\left(\sigma_{xx}U_x+(2c)^{1/2}D\right)^2\right\}^{-1/2}} = U_0 \stackrel{d}{=} N(0,1), \end{aligned} \tag{2.15}$$

and standard methods of inference apply.

Remark 4 *Let* $\widehat{R}_N$ *be the LS estimator of* R_N *and* $\widehat{c} = N^\alpha\left(\widehat{R}_N - 1\right)$*. The limit theory for* $\widehat{R}_N$ *and* $\widehat{c}$ *follows from Remark 3. Defining the regression residuals* $\hat{u}_{xt} = x_t - \widehat{R}_N x_{t-1} - \widehat{\mu}$ *and noting that the residual variance estimate* $s_x^2 = N^{-1}\sum_{t=1}^N \hat{u}_{xt}^2 \xrightarrow{p} \sigma_{xx}^2$*, we have the following result for the t statistic for testing* $\mathbb{H}_0 : R_N = R_N^0$*,*

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

Similarly, given $R_N = 1 + \frac{c}{N^\alpha}$*, we have* $s_c = N^\alpha s_{R_N}$ *and* $\widehat{c} - c^0 = \left(\widehat{R}_N - R_N^0\right)N^\alpha$*. Hence, if* α *is known, the t statistic for testing* $\mathbb{H}_0 : c = c^0$ *is*

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

However, if α *is unknown, the standard error* $s_c = N^\alpha s_{R_N}$ *is unavailable and inference using this limit theory for* $\widehat{c}$ *is infeasible. As discussed below, this infeasible feature of the discrete time case is quite different in continuous time.*

Remark 5 *The limit distribution (2.11) is a ratio of two independent Gaussian variates and has heavy tails, just as the Cauchy limit in the special case (2.13) where* $D = 0$. *Observe that*

$$R_N^N N^\alpha \left(\widehat{A} - A\right) \Rightarrow 2c \frac{\sigma_{00} U_0}{\sigma_{xx} U_x + (2c)^{1/2} D} = b \frac{U_0}{U_x + d}, \quad b = \frac{2c\sigma_{00}}{\sigma_{xx}}, \quad d = \frac{(2c)^{1/2}}{\sigma_{xx}} D.$$

When $D \geq 0$, *the density of* $U = \frac{U_0}{U_x + d}$ *is,*

$$p_U(u) = \frac{e^{-\frac{1}{2}d^2}}{\pi(1+u^2)} \left[1 + \frac{q}{\varphi(q)} \int_0^q \varphi(y)\, dy\right], \quad \varphi(y) = \frac{e^{-y^2/2}}{\sqrt{2\pi}}, \quad q = \frac{d}{\sqrt{1+u^2}},$$

(e.g., see Marsaglia, 1965) and has Cauchy-like tails.

We have the following expression for the LS estimator $\widehat{\beta}$ of the slope coefficient in the continuous time model (2.1), which is given by

$$\widehat{\beta} - \beta = \left(\sum_{t=1}^N x_{th}^2\right)^{-1} \left(\sum_{t=1}^N x_{th} u_{0,th}\right) = \left(\sum_{t=1}^N \widetilde{x}_{th}^2\right)^{-1} \left(\sum_{t=1}^N \widetilde{x}_{th} \widetilde{u}_{0,th}\right). \tag{2.16}$$

The associated limit theory is given in the following theorem.

Theorem 2.2 *For the continuous time system (2.3)-(2.4) with* $\kappa < 0$, *when* $h \to 0$ *and* $T \to \infty$, *we have*

$$\frac{a_h^N}{h}\left(\widehat{\beta} - \beta\right) \Rightarrow (-2\kappa) \frac{\sigma_{00} U_0}{\sigma_{xx} U_x + (-2\kappa)^{1/2} (x_0 - \mu)}. \tag{2.17}$$

Remark 6 *The limit distribution (2.17) follows directly from (2.11) by replacing* σ_{00}, σ_{xx}, *and* D *in Theorem 2.1 with* $\frac{\sigma_{00}}{\sigma_{xx}}$, 1, *and* $D^* = \frac{x_0 - \mu}{\sigma_{xx}}$ *respectively, giving the stated result.*

Remark 7 *The continuous time counterpart of* N^α *is* $1/h$ *which is known for any given data, so there is no need to estimate the rate parameter* α. *The continuous time counterpart of* c *is* $-\kappa$ *which can be consistently estimated by the least squares method as long as* $T \to \infty$. *Analogous to (2.15), self normalized statistics are free of nuisance parameters and hypothesis testing about* β *can be conducted using the residual variance estimate* $s_0^2 = N^{-1} \sum_{t=1}^N u_{0,th}^2$, *which satisfies* $h^{-1} s_0^2 \xrightarrow{p} \sigma_{00}^2$. *Theorem 2.2 and results (8.11) and (8.12) in the Appendix then give the following double asymptotics for the usual* t *statistic for testing* $\mathbb{H}_0 : \beta = \beta^0$

$$t_\beta = \frac{\widehat{\beta} - \beta^0}{s_\beta} = \frac{\widehat{\beta} - \beta^0}{\left\{ s_0^2 \left(\sum_{t=1}^N x_{th}^2\right)^{-1}\right\}^{1/2}} = \frac{\left(\widehat{\beta} - \beta^0\right) a_h^N / h}{\left\{ s_0^2 \left(\frac{a_h^{2N}}{h^2} \sum_{t=1}^N \widetilde{x}_{th}^2 \sigma_{xx}^2 h\right)^{-1}\right\}^{1/2}}$$

$$
\begin{aligned}
&= \frac{\left(\widehat{\beta}-\beta^0\right) a_h^N/h}{\left\{h^{-1}s_0^2\left(\frac{a_h^{2N}}{h^2}\sum_{t=1}^N \widetilde{x}_{th}^2\sigma_{xx}^2\right)^{-1}\right\}^{1/2}} \\
&\Rightarrow \frac{(-2\kappa)\frac{\sigma_{00}U_0}{\sigma_{xx}U_x+(-2\kappa)^{1/2}(x_0-\mu)}}{\sigma_{00}\left\{\left(\frac{1}{-2\kappa}\right)^2\left(U_x+(-2\kappa)^{1/2}D^*\right)^2\sigma_{xx}^2\right\}^{-1/2}} = U_0 \overset{d}{=} N(0,1),
\end{aligned}
$$

which leads to feasible inference concerning the slope coefficient β in continuous time, just as in (2.15) for the coefficient A in the modified MP model.

Remark 8 *Following Remark 7, we can obtain the double asymptotic distributions for $\widehat{a}_h$ and $\widehat{\kappa}$. Defining $s_x^2 = N^{-1}\sum_{t=1}^N u_{x,th}^2$, which satisfies $h^{-1}s_x^2 \overset{p}{\to} \sigma_{xx}^2$, the t statistic for $\widehat{a}_h$ is:*

$$
\begin{aligned}
t_{a_h} &= \frac{\widehat{a}_h - a_h^0}{s_{a_h}} = \frac{\widehat{a}_h - a_h^0}{\left\{s_x^2\left(\sum_{t=1}^N x_{(t-1)h}^2 - \frac{1}{N}\left(\sum_{t=1}^N x_{(t-1)h}\right)^2\right)^{-1}\right\}^{1/2}} \\
&= \frac{\left(\widehat{a}_h - a_h^0\right)a_h^N/h}{\left\{s_x^2\left(\frac{1}{a_h^{2N}/h^2}\sum_{t=1}^N x_{(t-1)h}^2 - \frac{1}{N}\left(\frac{1}{a_h^N/h}\sum_{t=1}^N x_{(t-1)h}\right)^2\right)^{-1}\right\}^{1/2}} \\
&\Rightarrow \frac{-2\kappa\frac{\sigma_{xx}Ux}{\sigma_{xx}U_x+(-2\kappa)^{1/2}D}}{\sigma_{xx}\left\{\left(\frac{1}{-2\kappa}\right)^2\left(\sigma_{xx}U_x+(-2\kappa)^{1/2}D\right)^2\right\}^{-1/2}} = U_x \overset{d}{=} N(0,1).
\end{aligned}
$$

Similarly, given $a_h = \exp(-kh)$, we have $hs_\kappa = s_{a_h} + o_p(h)$. Following Wang and Yu (2014), we have

$$
a_h^N\left(\widehat{\kappa}-\kappa^0\right) \Rightarrow 2\kappa\frac{\sigma_{xx}U_x}{\sigma_{xx}U_x+(-2\kappa)^{1/2}D},
$$

and

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

Clearly, t_κ is a feasible statistic for testing $\mathbb{H}_0 : \kappa = \kappa^0$ in contrast to the discrete time case where the test statistic relies on the unknown rate parameter α.

Remark 9 *If $x_0 = \mu$, we have $D^* = 0$ and*

$$\frac{a_h^N}{h}\left(\widehat{\beta} - \beta\right) \Rightarrow (-2\kappa)\frac{\sigma_{00}U_0}{\sigma_{xx}U_x} = (-2\kappa)\frac{\sigma_{00}}{\sigma_{xx}}\mathbb{C}. \tag{2.18}$$

3 Continuous Systems with Multiple Explosive Regressors

This section extends the results above to continuous time systems with more than one mildly explosive regressor. We allow for regressors with multiple forms of explosive behavior using the approach developed in MP (2009) for discrete systems. As above, we establish the limit theory for a modified MP model that incorporates an intercept term and allows for a larger initial condition. This theory is applied to the continuous system by assuming $T \to \infty$ and $h \to 0$. Following MP, two different cases will be examined which lead to somewhat different limit behavior: (i) when all the regressors have distinct explosive roots; and (ii) when all the regressors share the same explosive root.

3.1 Limit Results in the Discrete Time Framework

We start with the following system with multiple mildly explosive regressors, based on MP (2009),

$$\begin{aligned} y_t &= Ax_t + u_{0t}, & (3.1)\\ x_t &= \mu + R_N x_{t-1} + u_{xt}, \text{ with } x_0 = x_{0N} = O_p\left(N^{\alpha/2}\right) \text{ and } \mu = O_p\left(N^{-\alpha/2}\right). & (3.2)\end{aligned}$$

In this case, y_t and x_t are $m \times 1$ and $K \times 1$ vector respectively, and A is a $m \times K$ matrix of coefficients. In addition, $R_N = I_K + C/N^\alpha$ is a $K \times K$ matrix with $C = diag\left(c_1, c_2, \ldots, c_K\right) > 0$. We assume that the errors satisfy

$$u_t = \left[u_{0t}',\ u_{xt}'\right]' \overset{iid}{\sim} (0, \Omega) \text{ with } \Omega = \begin{bmatrix} \Omega_{00} & \Omega_{0x} \\ \Omega_{0x} & \Omega_{xx} \end{bmatrix}.$$

Let the standardized initialization and intercept satisfy $\widetilde{x}_0 = x_{0N}N^{-\alpha/2} \Rightarrow X^*$ and $\widetilde{\mu} = N^{\alpha/2}\mu \Rightarrow \mu^*$. The model now modifies MP (2009) in two ways: (i) the initial value for x is $O_p\left(N^{\alpha/2}\right)$ which is larger than the $o_p\left(N^{\alpha/2}\right)$ initialization in MP; (ii) a non-zero drift term of order $O_p\left(N^{-\alpha/2}\right)$ is included. Following closely the approach of MP, we obtain the limit theory for the LS estimator $\widehat{A}$ under two scenarios: (i) where C has distinct diagonal elements, i.e., $c_i \neq c_j$ for $i \neq j$; (ii) where C is a scalar matrix and does not have distinct diagonal elements, so that $c_i = c_j$, for all i, j. In what follows we will frequently use a zero affix to denote the true value of the associated element or matrix.

3.1.1 $c_i \neq c_j$ for $i \neq j$

Theorem 3.1 *For the discrete time system (3.1)-(3.2) with $R_N = I_K + C/N^\alpha$, $\alpha \in (0,1)$, $C = diag(c_1, c_2, \ldots, c_K) > 0$ and $c_i \neq c_j$ for $i \neq j$, when $N \to \infty$, we have*

(i) $vec\left(\frac{1}{N^\alpha}\sum_{t=1}^N u_{0t}x_t' R_N^{-N}\right) \Rightarrow \left(\int_0^\infty e^{-pC}\left(D+\widetilde{U}_x\right)\left(D+\widetilde{U}_x\right)' e^{-pC}dp \otimes \Omega_{00}\right)^{1/2} W_0,$

(ii) $\frac{1}{N^{2\alpha}}\sum_{t=1}^N R_N^{-N}x_t x_t' R_N^{-N} = \int_0^\infty e^{-pC}\left(D+\widetilde{U}_x\right)\left(D+\widetilde{U}_x\right)' e^{-pC}dp + o_p(1),$

(iii)

$$\begin{aligned}
&vec\left\{N^\alpha\left(\widehat{A}-A\right)R_N^N\right\} \Rightarrow \left\{\left(\int_0^\infty e^{-pC}\left(D+\widetilde{U}_x\right)\left(D+\widetilde{U}_x\right)' e^{-pC}dp\right)^{-1/2} \otimes \Omega_{00}^{1/2}\right\} W_0 \\
\overset{d}{=}\ & MN\left(0, \left(\int_0^\infty e^{-pC}\left(D+\widetilde{U}_x\right)\left(D+\widetilde{U}_x\right)' e^{-pC}dp\right)^{-1} \otimes \Omega_{00}\right), \qquad (3.3)
\end{aligned}$$

where $W_0 \overset{d}{=} N(0, I_{mK})$, $\widetilde{U}_x = \left(\int_0^\infty e^{-pC}\Omega_{xx}e^{-pC}dp\right)^{1/2} U_x$, $U_x \overset{d}{=} N(0, I_K)$, $D = X^* + C^{-1}\mu^*$, MN represents a mixed normal distribution.

Remark 10 *If $\widetilde{x}_0 = -C^{-1}\widetilde{\mu}$, then $D = 0$ and the limit (3.3) becomes*

$$vec\left\{N^\alpha\left(\widehat{A}-A\right)R_N^N\right\} \Rightarrow MN\left(0, \left(\int_0^\infty e^{-pC}\widetilde{U}_x\widetilde{U}_x' e^{-pC}dp\right)^{-1} \otimes \Omega_{00}\right).$$

This limit distribution corresponds to that in Theorem 4.1 of MP (2009, p. 496).

Remark 11 *The limit distribution of $\widehat{\mu}$ is*

$$\begin{aligned}
&\sqrt{N}\left(\widehat{\mu}-\mu\right) \\
=\ & \sqrt{N}\left\{\begin{array}{c} \frac{1}{N}\sum_{t=1}^N u_{xt} - \frac{1}{N}\left(\frac{1}{N^\alpha}\sum_{t=1}^N u_{xt}x_{t-1}' R_N^{-N}\right) \\ \left(\frac{1}{N^{2\alpha}}\sum_{t=1}^N R_N^{-N}x_{t-1}x_{t-1}'R_N^{-N}\right)^{-1}\left(\frac{1}{N^\alpha}\sum_{t=1}^N R_N^{-N}x_{t-1}\right)\end{array}\right\} \\
& \left\{\frac{N}{N} - \frac{1}{N}\left(\frac{1}{N^\alpha}\sum_{t=1}^N x_{t-1}'R_N^{-N}\right)\left(\frac{1}{N^{2\alpha}}\sum_{t=1}^N R_N^{-N}x_{t-1}x_{t-1}'R_N^{-N}\right)^{-1}\left(\frac{1}{N^\alpha}\sum_{t=1}^N R_N^{-N}x_{t-1}\right)\right\}^{-1} \\
=\ & \frac{1}{\sqrt{N}}\sum_{t=1}^N u_{xt} + o_p(1) \Rightarrow N(0, \Omega_{xx}).
\end{aligned}$$

Remark 12 *Let A_j and $\widehat{A}_j$ denote the jth $m \times 1$ column of A and $\widehat{A}$, and x_{jt} the jth element of x_t. Define the equation residuals $\hat{u}_{0t} = y_t - \widehat{A}x_t$, the error variance matrix estimate*

$$S_{00} = N^{-1} \sum_{t=1}^{N} \hat{u}_{0t}\hat{u}_{0t}' \xrightarrow{p} \Omega_{00},$$

and the corresponding estimate of the variance matrix of $\widehat{A}_j$

$$S_{A_jA_j} = \left(\sum_{t=1}^{N} x_{jt}^2 \right)^{-1} S_{00}.$$

The limit distribution of $\widehat{A}_j$ is given by

$$N^{\alpha} \left(\widehat{A}_j - A_j \right) \rho_j^N \Rightarrow MN \left(0, \left(\int_0^{\infty} e^{-2pc_j} \left(D_j + \widetilde{U}_{x_j} \right)^2 dp \right)^{-1} \Omega_{00} \right) \overset{d}{=} MN \left(0, \frac{2c_j\Omega_{00}}{\left(D_j + \widetilde{U}_{x_j} \right)^2} \right)$$

where $\rho_j = 1 + \frac{c_j}{N^{\alpha}}$, and D_j and $\widetilde{U}_{jx}$ are the jth element of D and $\widetilde{U}_x$ for $j = 1, \ldots, K$. Using Theorem 3.1, we obtain the following limit distribution for the Wald statistic for testing $\mathbb{H}_0 : Q_jA_j = Q_jA_j^0 = q_j$, where Q_j is a $g \times m$ restriction matrix of full row rank $g \leq m$ and q_j is a given $g \times 1$ vector,

$$\begin{aligned} W_{A_j} &:= \left\{ Q_j\widehat{A}_j - q_j \right\}' \left(Q_jS_{A_jA_j}Q_j' \right)^{-1} \left\{ Q_j\widehat{A}_j - q_j \right\} \\ &= \left\{ Q_jN^{\alpha} \left(\widehat{A}_j - A_j \right) \rho_j^N \right\}' \left(\left(N^{-2\alpha}\rho_j^{-2N} \sum_{t=1}^{N} x_{jt}^2 \right)^{-1} Q_jS_{00}Q_j' \right)^{-1} \left\{ Q_jN^{\alpha} \left(\widehat{A}_j - A_j \right) \rho_j^N \right\} \\ &\Rightarrow \chi_g^2, \end{aligned}$$

where χ_g^2 denotes a chi-squared variate with g degrees of freedom.

Remark 13 *Let R_{jN} and $\widehat{R}_{jN}$ denote the jth $K \times 1$ column of R_N and $\widehat{R}_N$, and define $\widehat{C}_j = \left(\widehat{R}_j - e_j \right) N^{\alpha}$, where e_j is the jth unit vector with unity in the jth position and zeros elsewhere. Setting $\hat{u}_{xt} = x_t - \widehat{R}_N x_{t-1} - \widehat{\mu}$, the residual second moment matrix is*

$$S_{xx} = N^{-1} \sum_{t=1}^{N} \hat{u}_{xt}\hat{u}_{xt}' \xrightarrow{p} \Omega_{xx},$$

and the corresponding estimate of the variance matrix of $\widehat{R}_{jN}$ is

$$S_{R_jR_j} = \left(\sum_{t=1}^{N} x_{jt-1}^2 - \frac{1}{N} \left(\sum_{t=1}^{N} x_{jt-1} \right)^2 \right)^{-1} S_{xx}.$$

The Wald statistic for testing the (full rank) restrictions $\mathbb{H}_0 : Q_j R_{jN} = Q_j R_{jN}^0 = q_j$, *where* Q_j *is a* $g \times K$ *restriction matrix of full row rank* $g \leq K$ *and* q_j *is a given* $g \times 1$ *vector, is:*

$$W_{R_{jN}} := \left\{Q_j \widehat{R}_{jN} - q_j\right\}' \left(Q_j S_{R_j R_j} Q_j'\right)^{-1} \left\{Q_j \widehat{R}_{jN} - q_j\right\} \Rightarrow \chi_g^2,$$

under the null. Similarly, given $R_N = I_K + C/N^\alpha$, *we can set* $S_{C_j C_j} = N^{2\alpha} S_{R_j R_j}$ *where* C_j *is the* j*th column of* C, $S_{C_j C_j}$ *is the covariance matrix of* $\widehat{C}_j$. *Further* $\widehat{C}_j - C_j^0 = \left(\widehat{R}_{jN} - R_{jN}^0\right) N^\alpha$ *leads to the following limit theory for* $\widehat{C}_j$

$$\left(\widehat{C_j} - C_j^0\right) \rho_j^N \Rightarrow MN\left(0, \left(\int_0^\infty e^{-2pc_j} \left(D_j + \widetilde{U}_{x_j}\right)^2 dp\right)^{-1} \Omega_{xx}\right) \stackrel{d}{=} MN\left(0, \frac{2c_j \Omega_{xx}}{\left(D_j + \widetilde{U}_{x_j}\right)^2}\right).$$

Hence, if α *is known, we have the corresponding feasible Wald statistic for testing the restrictions* $\mathbb{H}_0 : Q_j C_j = Q_j C_j^0 = q_j$,

$$W_{C_j} := \left\{Q_j \widehat{C}_j - q_j\right\}' \left(Q_j S_{C_j C_j} Q_j'\right)^{-1} \left\{Q_j \widehat{C}_j - q_j\right\} \Rightarrow \chi_g^2,$$

under the null and with full row rank Q_j. *If* α *is unknown, just as in the scalar model, the estimated variance matrix* $S_{C_j C_j} = N^{2\alpha} S_{R_j R_j}$ *is unavailable and inference using this limit theory for* $\widehat{C}_j$ *is infeasible. Note that under the null* $C_j^0 = c_j^0 e_j$. *Imposing this (maintained) restriction on the form of* C_j^0 *implies that the null can be rewritten as* $\mathbb{H}_0 : c_j = c_j^0$ *and a test analogous to the scalar case can be mounted using the* j*th diagonal element of the (unrestricted) estimate* $\widehat{C}$ *or a similar estimate obtained by imposing the maintained restriction on* C_j *and estimating the system as a seemingly unrelated regression (SUR). Similar constraints on inference due to the infeasibility of the tests apply in each of these cases.*

3.1.2 $c_i = c_j$ for all i, j

When $c_i = c_j = c$, for all i, j, the limiting standardized form of the signal matrix $\sum_{i=1}^N x_t x_t'$ is singular due to commonality in the explosive behavior of the components of x_t. Let $R_N = \rho_N I_K$ with $\rho_N = 1 + c/N^\alpha$. Following MP (2009), we rotate regression coordinates to address the singularity using an orthogonal random matrix $H_N = [H_{cN}, H_{\perp N}]$ where $H_{cN} = \frac{x_N}{\left(x_N' x_N\right)^{1/2}}$ and $H_{\perp N}$ is a $K \times (K-1)$ orthogonal complement matrix such that $H_{\perp N}' H_{cN} \stackrel{a.s.}{=} 0$. Then, $H_{\perp N}' H_{\perp N} = I_{K-1}$ and $H_{\perp N} H_{\perp N}' \stackrel{a.s.}{=} I_K - H_{cN} H_{cN}'$. The limit of $H_{\perp N}$ is denoted as $H_\perp$, which satisfies $H_\perp H_\perp' = I_K - X_c X_c'$ where X_c is defined in (3.4) in Theorem 3.2 below. Next, rotate the regressor x_t by H_N and transform to the variate $z_t = H_N' x_t = [H_{cN}' x_t, H_{\perp N}' x_t] =: [z_{1t}', z_{2t}']'$. The following result gives the required limit theory for the LS estimator $\widehat{A}$ in this case.

Theorem 3.2 *For the discrete time system (3.1)-(3.2) with* $R_N = I_K + C/N^{\alpha}$, $C = diag\,(c_1, c_2, \ldots, c_K)$ *with* $c_i = c > 0$ *for* $i = 1, \ldots, K$, *when* $N \to \infty$, *we have,*

(i) $\frac{1}{N^{1+\alpha}} \sum_{t=1}^{N} z_{2t} z_{2t}' \Rightarrow M$, *with* $M = H_{\perp}' \left(\frac{\mu^*}{c} \frac{\mu^{*\prime}}{c} + \frac{1}{2c} \Omega_{xx} \right) H_{\perp}$,

(ii) $\frac{1}{N^{(1+\alpha)/2}} vec \left(\sum_{t=1}^{N} u_{0t} z_{2t}' \right) \Rightarrow \{ M \otimes \Omega_{00} \}^{1/2} \times N\left(0, I_{m(K-1)}\right)$, *where* $H_{\perp}$ *is a* $K \times (K-1)$ *random matrix that is an orthogonal complement to*

$$X_c = \left(D + \tilde{U}_x \right) / \left\{ \left(D + \tilde{U}_x \right)' \left(D + \tilde{U}_x \right) \right\}^{1/2}, \tag{3.4}$$

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

$$\widetilde{U}_x \equiv \left(\int_0^{\infty} e^{-pc} \Omega_{xx} e^{-pc} dp \right)^{1/2} U_x = \Omega_{xx}^{1/2} U_x / (2c)^{1/2}, \text{ and } D = X^* + \mu^*/c,$$

(iii)

$$\begin{aligned} N^{(1+\alpha)/2} vec \left(\widehat{A} - A \right) &\Rightarrow \left(H_{\perp} M^{-1/2} \otimes \Omega_{00}^{1/2} \right) \times N\left(0, I_{mK}\right) \\ &\overset{d}{=} MN\left(0, H_{\perp} M^{-1} H_{\perp}' \otimes \Omega_{00}\right). \end{aligned} \tag{3.5}$$

Remark 14 *The limit distribution of* $\widehat{\mu}$ *is obtained as follows:*

$$\begin{aligned} & \sqrt{N} \left(\widehat{\mu} - \mu \right) \\ = & \sqrt{N} \left\{ \sum_{t=1}^{N} u_{xt} - \left(\sum_{t=1}^{N} u_{xt} x_{t-1}' H_N \right) \left(\sum_{t=1}^{N} H_N' x_{t-1} x_{t-1}' H_N \right)^{-1} \left(\sum_{t=1}^{N} H_N' x_{t-1} \right) \right\} \\ & \left\{ N - \left(\sum_{t=1}^{N} u_{xt} x_{t-1}' H_N \right) \left(\sum_{t=1}^{N} H_N x_{t-1} x_{t-1}' H_N' \right)^{-1} \left(\sum_{t=1}^{N} H_N' x_{t-1} \right) \right\}^{-1} \\ = & \sqrt{N} \left\{ \frac{1}{N} \sum_{t=1}^{N} u_{xt} - \frac{1}{N} \left(\frac{1}{N^{(1+\alpha)/2}} \sum_{t=1}^{N} u_{xt} z_{t-1}' \right) \left(\frac{1}{N^{(1+\alpha)}} \sum_{t=1}^{N} z_{t-1} z_{t-1}' \right)^{-1} \left(\frac{1}{N^{(1+\alpha)/2}} \sum_{t=1}^{N} z_{t-1} \right) \right\} \\ & \left\{ \frac{N}{N} - \frac{1}{N} \left(\frac{1}{N^{(1+\alpha)/2}} \sum_{t=1}^{N} u_{xt} z_{t-1}' \right) \left(\frac{1}{N^{(1+\alpha)}} \sum_{t=1}^{N} z_{t-1} z_{t-1}' \right)^{-1} \left(\frac{1}{N^{(1+\alpha)/2}} \sum_{t=1}^{N} z_{t-1} \right) \right\}^{-1} \\ = & \frac{1}{\sqrt{N}} \sum_{t=1}^{N} u_{xt} + o_p(1) \Rightarrow N\left(0, \Omega_{xx}\right). \end{aligned}$$

Remark 15 *Let* $\hat{u}_{0t} = y_t - \widehat{A} x_t$, *the estimate of the error variance matrix be*

$$S_{00} = N^{-1} \sum_{t=1}^{N} \hat{u}_{0t} \hat{u}_{0t}' \xrightarrow{p} \Omega_{00},$$

and the estimated variance matrix of $\widehat{A}_j$ be

$$S_{A_jA_j} = \left(\sum_{t=1}^{N} x_{jt}^2\right)^{-1} S_{00}.$$

Following Theorem 3.2, we have the following limit theory for the Wald statistic for testing $\mathbb{H}_0 : Q_j A_j = Q_j A^0 = q_j$

$$W_{A_j} := \left\{Q_j\widehat{A}_j - q_j\right\}' \left(Q_j S_{A_jA_j} Q_j'\right)^{-1} \left\{Q_j\widehat{A}_j - q_j\right\} \Rightarrow \chi_g^2.$$

Remark 16 *Let $\hat{u}_{xt} = x_t - \widehat{R}_N x_{t-1} - \widehat{\mu}$, giving the error variance matrix estimate*

$$S_{xx} = N^{-1}\sum_{t=1}^{N} \hat{u}_{xt}\hat{u}_{xt}' \xrightarrow{p} \Omega_{xx},$$

and the corresponding estimate of the variance matrix of $\widehat{R}_N$ (in column vector form)

$$S_{RR} = \left(\sum_{t=1}^{N} x_{t-1}x_{t-1}' - \frac{1}{N}\left(\sum_{t=1}^{N} x_{t-1}\right)\left(\sum_{t=1}^{N} x_{t-1}\right)'\right)^{-1} \otimes S_{xx}.$$

Then the Wald statistic for testing $\mathbb{H}_0 : Qvec\left(R_N\right) = Qvec\left(R_N^0\right) = r$, where Q is a $g \times mK$ restriction matrix of full row rank $g \leq mK$, is

$$W_{R_N} := \left\{Qvec\left(\widehat{R}_N - R_N^0\right)\right\}' \left(QS_{RR}Q'\right)^{-1} \left\{Qvec\left(\widehat{R}_N - R_N^0\right)\right\} \Rightarrow \chi_g^2,$$

since

$$N^{(1+\alpha)/2} Qvec\left(\widehat{R}_N - R_N^0\right) \Rightarrow MN\left(0, QH_\perp M^{-1}H_\perp' \otimes \Omega_{xx}Q'\right),$$

and

$$\begin{aligned} N^{1+\alpha}QS_{RR}Q' &= Q\left(\frac{1}{N^{1+\alpha}}\sum_{t=1}^{N} x_{t-1}x_{t-1}' - \frac{1}{N^{2+\alpha}}\left(\sum_{t=1}^{N} x_{t-1}\right)\left(\sum_{t=1}^{N} x_{t-1}\right)'\right)^{-1} \otimes S_{xx}Q' \\ &= QH_N\left(\frac{1}{N^{1+\alpha}}\sum_{t=1}^{N} z_{t-1}z_{t-1}'\right)^{-1} H_N' \otimes S_{xx}Q' + o_p(1) \\ &\Rightarrow Q\left(H_\perp M^{-1}H_\perp' \otimes \Omega_{xx}\right)Q'. \end{aligned}$$

Similarly, given $R_N = 1 + \frac{C}{N^\alpha}$, we have $S_{CC} = N^{2\alpha}S_{RR}$ and

$$vec\left(\widehat{C} - C^0\right) N^{\frac{1-\alpha}{2}} \Rightarrow MN\left(0, H_\perp M^{-1}H_\perp' \otimes \Omega_{xx}\right).$$

Hence, given α, we have the following feasible Wald test,

$$W_C := \left\{ Qvec\left(\widehat{C} - C^0\right)\right\}' \left(QS_{CC}Q'\right)^{-1} \left\{ Qvec\left(\widehat{C} - C^0\right)\right\} \Rightarrow \chi^2_g.$$

Again as in the scalar model, if α is unknown (which is the usual situation in practical work), the estimated variance matrix $S_{CC} = N^{2\alpha} S_{RR}$ is unavailable and inference using this limit theory for $\widehat{C}$ is infeasible.

Importantly, for the common explosive root case when α is known, we are able to perform statistical inference concerning the full matrix coefficients R_N and C using Wald tests because the normalization factor $N^{1+\alpha}$ is common and thereby commutable with the restriction matrix Q. However, for the distinct explosive roots case, we can only perform statistical inference about individual column vectors of R_N and C, as demonstrated in Remark 13. The same phenomenon applies for tests involving the matrix A. As shown below, these features carry over to inference in the continuous time system although in this case the sampling frequency is known so there is no difficulty relating to an unknown rate parameter α.

3.2 Limit Results in the Continuous Time Framework

The above results apply to the multivariate continuous time system

$$\begin{aligned} dy(t) &= \beta dx(t) + \Omega_{00}^{1/2} dB_0(t), && (3.6)\\ dx(t) &= \kappa(\mu - x(t))\,dt + \Omega_{xx}^{1/2} dB_x(t),\ x(0) = x_0 = O_p(1),\ \kappa < 0, && (3.7) \end{aligned}$$

where $B_0(t)$ and $B_x(t)$ are $m-$ and $K-$ vectors of standard Brownian motion. The driver process $x(t)$ follows 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. We focus on the explosive case where $\kappa_i < 0$ for $i = 1, \ldots, K$. As in the discrete time case, we are interested in β, an $m \times K$ matrix of coefficients which captures co-movement between $y(t)$ and $x(t)$.

The exact discrete time representation of (3.6)-(3.7) is given by (see Phillips, 1972)

$$\begin{aligned} y_{th} &= \beta x_{th} + u_{0,th}, && (3.8)\\ x_{th} &= a_h(\kappa)\, x_{(t-1)h} + g_h + u_{x,th},\ x_{0h} = x_0 = O_p(1), \end{aligned}$$

where

$$\begin{aligned} a_h(\kappa) &= \exp(-\kappa h),\\ g_h &= \kappa^{-1}\left(I_K - e^{-\kappa h}\right)\kappa\mu,\\ u_{x,th} &= \int_{(t-1)h}^{th} e^{-\kappa(th-s)} \Omega_{xx} dB_x(s) \sim N(0, \Omega_{xx} h), \end{aligned}$$

since

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

Thus, upon restandardization by $\sqrt{h}$, the system becomes

$$\widetilde{y}_{th} = \beta\widetilde{x}_{th} + \widetilde{u}_{0,th}, \tag{3.9}$$

$$\widetilde{x}_{th} = a_h(\kappa)\widetilde{x}_{(t-1)h} + \widetilde{g}_h + \widetilde{u}_{x,th},\ \widetilde{x}_{0h} = h^{-1/2}x_{0h},\ \widetilde{u}_{x,th} \overset{iid}{\sim} N(0,\Omega_{xx}), \tag{3.10}$$

where $\widetilde{y}_{th} = h^{-1/2}y_{th}$, $\widetilde{x}_{th} = h^{-1/2}x_{th}$, $\widetilde{g}_h = h^{-1/2}g_h$, $\widetilde{u}_{0,th} = h^{-1/2}u_{0,th} \overset{d}{=} N(0,\Omega_{00})$, and $\widetilde{u}_{x,th} = h^{-1/2}u_{x,th} \overset{d}{=} N(0,\Omega_{xx})$. As in the univariate case, the order of the initial value $\widetilde{x}_{0h} = h^{-1/2}x_{0h}$ is $O_p\left(h^{-1/2}\right)$, and the order for the drift term $\widetilde{g}_h$ is $O_p\left(h^{1/2}\right)$.

For the continuous time system (3.9)-(3.10), the double asymptotic theory for the LS estimator of the coefficient matrix β when κ has distinct diagonal elements (i.e., $\kappa_i \neq \kappa_j$ for $i \neq j$) is given in the following theorem.

Theorem 3.3 *For the continuous time system (3.9)-(3.10) with $\kappa = diag(\kappa_1, \kappa_2, \ldots, \kappa_K)$ with $\kappa_i \neq \kappa_j$, for $i \neq j$, and $\kappa_i < 0$ for $i = 1, \ldots, K$, when $h \to 0$ and $T \to \infty$, we have*

$$vec\left\{\frac{1}{h}\left(\widehat{\beta} - \beta\right)a_h^N\right\} \Rightarrow \left[\left(\int_0^\infty e^{p\kappa}\left(D + \widetilde{U}_x\right)\left(D + \widetilde{U}_x\right)' e^{p\kappa}dp\right)^{-1} \otimes \Omega_{00}\right]^{1/2} \times N\left(0, I_{mK}\right). \tag{3.11}$$

Remark 17 *The double asymptotic distribution (3.11) follows directly from (3.3) with $\mu^* = \kappa\mu$, $C = -k$ and $D = x_0 - \mu$. To enhance readability in terms of the relationship between the systems, we provide in the following Table 1 the correspondence between the models, variables and parameters in the discrete and continuous time cases.*

Remark 18 *The LS estimator of κ is consistent since h is known. Let $S_{00} = N^{-1}\sum_{t=1}^N u_{0,th}u_{0,th}'$, which satisfies*

$$h^{-1}S_{00} \xrightarrow{p} \Omega_{00},$$

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

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

The Wald statistic for testing the full rank restrictions $\mathbb{H}_0 : Q_j\beta_j = Q_j\beta_j^0 = r_j$ is then

$$W_{\beta_j} = \left\{Q_j\left(\widehat{\beta}_j - \beta_j^0\right)\right\}'\left(Q_jS_{\beta_j\beta_j}Q_j'\right)^{-1}\left\{Q_j\left(\widehat{\beta}_j - \beta_j^0\right)\right\} \Rightarrow \chi_g^2,$$

leading to feasible inference about β_j in the continuous time framework.

Table 1: Correspondence between systems (3.1)-(3.2) and (3.9)-(3.10).

Discrete Time	Continuous Time
$y_t = Ax_t + u_{0t}$ $x_t = \mu + R_N x_{t-1} + u_{xt}$ $x_0 = x_{0N} = O_p\left(N^{\alpha/2}\right)$ $\mu = O_p\left(N^{-\alpha/2}\right)$	$\widetilde{y}_{th} = \beta\widetilde{x}_{th} + \widetilde{u}_{0,th}$ $\widetilde{x}_{th} = a_h(\kappa)\,\widetilde{x}_{(t-1)h} + \widetilde{g}_h + \widetilde{u}_{x,th}$ $\widetilde{x}_{0h} = h^{-1/2}x_{0h} = O_p\left(h^{-1/2}\right)$ $\widetilde{g}_h = O_p\left(h^{1/2}\right)$
C	$-\kappa$
μ with $N^{\alpha/2}\mu \to \mu^*$	$\widetilde{g}_h$ with $h^{-1/2}\widetilde{g}_h \to \kappa\mu$
$X^* + C^{-1}\mu^*$	$x_0 - \mu$

Remark 19 *Let a_j be the jth column of $a_h(\kappa)$. The Wald statistic for testing the full rank restrictions $\mathbb{H}_0 : Q_j a_j = Q_j a_j^0 = q_j$ for given (Q_j, q_j) has the following chi-squared limit*

$$W_{a_j} := \left\{Q_j\left(\widehat{a}_j - a_j^0\right)\right\}'\left(Q_j S_{a_j a_j} Q_j'\right)^{-1}\left\{Q_j\left(\widehat{a}_j - a_j^0\right)\right\} \Rightarrow \chi_g^2,$$

where $S_{a_j a_j} = \left(\sum_{t=1}^N x_{j(t-1)h}^2 - \frac{1}{N}\left(\sum_{t=1}^N x_{j(t-1)h}\right)^2\right)^{-1} S_{xx}$ and $S_{xx} = N^{-1}\sum_{t=1}^N \widehat{u}_{x,th}\widehat{u}_{x,th}'$ satisfying $h^{-1}S_{xx} \xrightarrow{p} \Omega_{xx}$ where $\widehat{u}_{x,th} = x_{th} - \widehat{a}_h x_{(t-1)h} - \hat{g}_h$ are regression residuals. 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 $h^2 S_{\kappa^j\kappa^j} = S_{a_j a_j} + o(h)$ and so

$$\left(\widehat{\kappa}^j - \kappa^{j0}\right)e^{-k_j N} \Rightarrow MN\left(0, \left(\int_0^\infty e^{2p\kappa_j}\left(D_j + \widetilde{U}_{jx}\right)^2 dp\right)^{-1}\Omega_{xx}\right) \overset{d}{=} MN\left(0, \frac{-2\kappa_j\Omega_{xx}}{\left(D_j + \widetilde{U}_{jx}\right)^2}\right).$$

Then the Wald statistic for testing the (full rank) restrictions $\mathbb{H}_0 : Q_j\kappa^j = Q_j\kappa^{j0} = q_j$ satisfies

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

Remark 20 *The OLS estimates, $\widehat{a}_h$ and $\widehat{\kappa}$, do not take account of the diagonal structure of a_h and κ. If the known diagonal structure is imposed, we can use either SUR estimation or restricted OLS (in which only the diagonal elements of the original OLS estimates are employed). The simulation section below explores the finite sample performance of these three estimates.*

Now we consider the case where the localizing explosive coefficients are identical, so that $\kappa_i = \kappa$ for $i = 1, \ldots, K$.

Theorem 3.4 *For the continuous time system (3.9)-(3.10) with* $\kappa = diag(\kappa_1, \kappa_2, \ldots, \kappa_K)$ *and* $\kappa_i = \overline{\kappa} < 0$ *for* $i = 1, \ldots, K$, *when* $h \to 0$ *and* $T \to \infty$, *we have*

$$vec\left\{\sqrt{N/h}\left(\widehat{\beta}-\beta\right)\right\} \Rightarrow \left[H_\perp\left\{H_\perp'\left(\mu\mu'+\frac{1}{-2\overline{\kappa}}\Omega_{xx}\right)H_\perp\right\}^{-1/2}\otimes\Omega_{00}^{1/2}\right]\times N\left(0, I_{mK}\right). \tag{3.12}$$

Remark 21 *The double asymptotic distribution (3.12) follows directly from (3.5). with* $\mu^* = \kappa\mu$ *and* $c = -k$.

Remark 22 *The Wald statistic for testing* $\mathbb{H}_0 : Qvec(\beta) = Qvec\left(\beta^0\right) = q$ *for full row rank* (Q, q) *is then*

$$W_\beta := \left\{Q\widehat{\beta}-q\right\}'\left(QS_{\beta\beta}Q'\right)^{-1}\left\{Q\widehat{\beta}-q\right\} \Rightarrow \chi^2_g,$$

leading 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 (3.12).*

Remark 23 *The Wald statistics for testing full rank restrictions on* a_h *and* κ *such as* $\mathbb{H}_0 : Qvec(a_h) = Qvec\left(a_h^0\right) = q$ *and* $\mathbb{H}_0 : Qvec(\kappa) = Qvec\left(\kappa^0\right) = q$ *are defined in a similar way and have the following chi-squared limits:*

$$W_{a_h} := \left\{Qvec\left(\widehat{a}_h\right)-q\right\}'\left(QS_{aa}Q'\right)^{-1}\left\{Qvec\left(\widehat{a}_h\right)-q\right\} \Rightarrow \chi^2_g,$$

and

$$W_\kappa := \left\{Qvec\left(\widehat{\kappa}\right)-q\right\}'\left(QS_{\kappa\kappa}Q'\right)^{-1}\left\{Qvec\left(\widehat{\kappa}\right)-q\right\} \Rightarrow \chi^2_g,$$

again leading to feasible inference on a_h *and* κ *because of the common factorization convergence rate.*

Remark 24 *When* x_t *has a common explosive root, OLS estimation by* $\widehat{a}_h$ *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.*

Table 2: Comparison of the finite sample and double asymptotic distributions of $\hat{\beta}$, when the initial value is $x_0 = 0$.

	Time Span	T = 4						T = 10					
Frequency	C test	1%	2.50%	10%	90%	97.50%	99%	1%	2.50%	10%	90%	97.50%	99%
Daily	new	-32.233	-13.222	-3.027	3.063	12.063	29.452	-32.233	-13.222	-3.027	3.063	12.063	29.452
(h=1/252)	Finite Sample	-36.346	-13.501	-3.271	2.870	11.647	30.428	-28.528	-12.100	-2.889	3.166	13.304	32.007
Weekly	new	-32.233	-13.222	-3.027	3.063	12.063	29.452	-32.233	-13.222	-3.027	3.063	12.063	29.452
(h=1/52)	Finite Sample	-31.973	-11.329	-3.006	2.973	11.752	27.031	-32.775	-12.919	-3.145	3.018	12.735	35.756
Monthly	new	-32.233	-13.222	-3.027	3.063	12.063	29.452	-32.233	-13.222	-3.027	3.063	12.063	29.452
(h=1/12)	Finite Sample	-34.772	-12.327	-2.889	2.793	12.284	27.255	-31.209	-11.698	-2.922	2.842	11.947	32.185
	t test												
Daily	new	-2.326	-1.960	-1.282	1.282	1.960	2.326	-2.326	-1.960	-1.282	1.282	1.960	2.326
(h=1/252)	Finite Sample	-2.305	-1.950	-1.271	1.253	1.915	2.256	-2.380	-1.976	-1.277	1.301	2.015	2.359
Weekly	new	-2.326	-1.960	-1.282	1.282	1.960	2.326	-2.326	-1.960	-1.282	1.282	1.960	2.326
(h=1/52)	Finite Sample	-2.334	-2.004	-1.292	1.318	1.956	2.340	-2.375	-2.000	-1.286	1.248	1.936	2.285
Monthly	new	-2.326	-1.960	-1.282	1.282	1.960	2.326	-2.326	-1.960	-1.282	1.282	1.960	2.326
(h=1/12)	Finite Sample	-2.452	-2.072	-1.326	1.331	2.057	2.463	-2.434	-1.977	-1.301	1.306	1.969	2.301

4 Monte Carlo Studies

This section examines the performance of the double asymptotic limit theory in simulations. We generate data from model (2.3)-(2.4) with $\kappa = -2, \sigma_{00} = \sigma_{xx} = 1$, $\mu = 0$, and consider three sampling intervals, $h = 1/12$, $1/52, 1/252$, corresponding to monthly, weekly and daily frequencies. The initial value x_0 is set at $(0, 3, 10)$ and time spans of $T = 4$ and $T = 10$ years are considered. We report percentiles at levels $\{1\%, 2.5\%, 10\%, 90\%, 97.5\%, 99\%\}$ in the limit distribution (2.17) and the finite sample distribution of the coefficient based test (called the C test hereafter) $\frac{a^N}{-2\kappa h}\left(\widehat{\beta} - \beta\right)$ and t_β from Remark 7. In addition, we provide comparisons of the densities of the limit distributions and finite sample distributions of the C test statistic and t_β statistic. The number of replications is set at 10,000.

Tables 2, 3, and 4 report the percentiles when $x_0 = 0, 3, 10$ by using the true values of κ and μ. It can be seen that the double asymptotic distribution and the finite sample distribution are both sensitive to changes in initial condition. In all cases the new limit distribution provides a good approximation to the finite sample distribution.

Figure 1, 2, and 3 plot the densities of the C test statistic and t test statistic when $T = 4$. The result is similar to the case of $T = 10$, which is not reported. These plots show the limit

Table 3: Comparison of the finite sample and double asymptotic distributions of $\hat{\beta}$, when the initial value is $x_0 = 3$.

	Time Span	T = 4						T = 10					
Frequency	C test	1%	2.50%	10%	90%	97.50%	99%	1%	2.50%	10%	90%	97.50%	99%
Daily	new	-0.427	-0.348	-0.220	0.219	0.347	0.426	-0.427	-0.348	-0.220	0.219	0.347	0.426
(h=1/252)	Finite Sample	-0.416	-0.338	-0.215	0.215	0.341	0.412	-0.426	-0.350	-0.220	0.218	0.346	0.418
Weekly	new	-0.427	-0.348	-0.220	0.219	0.347	0.426	-0.427	-0.348	-0.220	0.219	0.347	0.426
(h=1/52)	Finite Sample	-0.413	-0.338	-0.215	0.215	0.337	0.414	-0.411	-0.341	-0.218	0.212	0.336	0.413
Monthly	new	-0.427	-0.348	-0.220	0.219	0.347	0.426	-0.427	-0.348	-0.220	0.219	0.347	0.426
(h=1/12)	Finite Sample	-0.385	-0.324	-0.204	0.205	0.321	0.386	-0.391	-0.318	-0.200	0.207	0.322	0.391
	t test												
Daily	new	-2.326	-1.960	-1.282	1.282	1.960	2.326	-2.326	-1.960	-1.282	1.282	1.960	2.326
(h=1/252)	Finite Sample	-2.275	-1.924	-1.271	1.249	1.947	2.283	-2.369	-1.989	-1.290	1.285	1.993	2.358
Weekly	new	-2.326	-1.960	-1.282	1.282	1.960	2.326	-2.326	-1.960	-1.282	1.282	1.960	2.326
(h=1/52)	Finite Sample	-2.331	-1.976	-1.309	1.300	1.985	2.340	-2.352	-1.965	-1.285	1.249	1.950	2.320
Monthly	new	-2.326	-1.960	-1.282	1.282	1.960	2.326	-2.326	-1.960	-1.282	1.282	1.960	2.326
(h=1/12)	Finite Sample	-2.439	-2.072	-1.325	1.327	2.050	2.469	-2.397	-1.986	-1.293	1.311	1.963	2.364

Table 4: Comparison of the finite sample and double asymptotic distributions of $\hat{\beta}$, when the initial value is $x_0 = 10$.

	Time Span	T = 4						T = 10					
	C test	1%	2.50%	10%	90%	97.50%	99%	1%	2.50%	10%	90%	97.50%	99%
Daily	new	-0.118	-0.099	-0.064	0.064	0.099	0.117	-0.118	-0.099	-0.064	0.064	0.099	0.117
(h=1/252)	Finite Sample	-0.114	-0.096	-0.063	0.063	0.098	0.115	-0.119	-0.100	-0.064	0.064	0.099	0.117
Weekly	new	-0.118	-0.099	-0.064	0.064	0.099	0.117	-0.118	-0.099	-0.064	0.064	0.099	0.117
(h=1/52)	Finite Sample	-0.114	-0.097	-0.064	0.063	0.097	0.116	-0.116	-0.096	-0.063	0.061	0.094	0.115
Monthly	new	-0.118	-0.099	-0.064	0.064	0.099	0.117	-0.118	-0.099	-0.064	0.064	0.099	0.117
(h=1/12)	Finite Sample	-0.109	-0.092	-0.059	0.060	0.092	0.108	-0.110	-0.091	-0.059	0.060	0.090	0.110
	t test												
Daily	new	-2.326	-1.960	-1.282	1.282	1.960	2.326	-2.326	-1.960	-1.282	1.282	1.960	2.326
(h=1/252)	Finite Sample	-2.274	-1.924	-1.271	1.249	1.947	2.283	-2.369	-1.989	-1.290	1.285	1.993	2.358
Weekly	new	-2.326	-1.960	-1.282	1.282	1.960	2.326	-2.326	-1.960	-1.282	1.282	1.960	2.326
(h=1/52)	Finite Sample	-2.331	-1.976	-1.309	1.300	1.985	2.340	-2.352	-1.965	-1.285	1.249	1.950	2.320
Monthly	new	-2.326	-1.960	-1.282	1.282	1.960	2.326	-2.326	-1.960	-1.282	1.282	1.960	2.326
(h=1/12)	Finite Sample	-2.439	-2.072	-1.325	1.327	2.050	2.469	-2.397	-1.986	-1.293	1.311	1.963	2.364

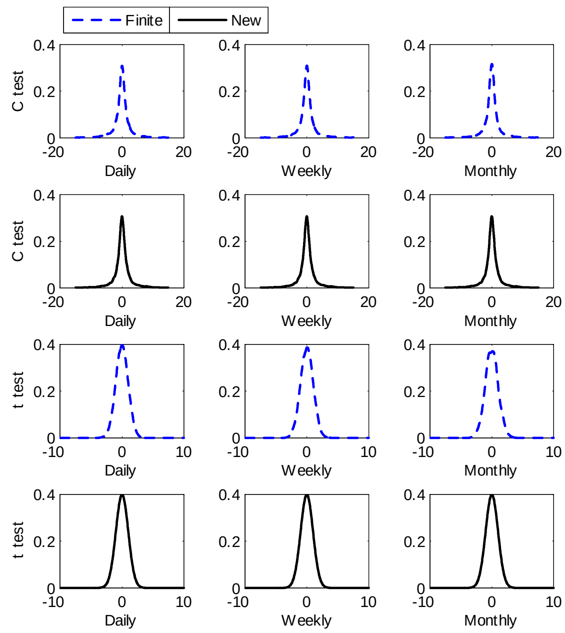

Figure 1: Density comparision between C test and t test for both finite sample distribution and limit distribution, when the initial value is $x_0 = 0$.

distribution well approximates the finite sample distribution for both tests.

To examine the performance of the limit theory in the multivariate setup, we consider a bivariate model using monthly data ($h = \frac{1}{12}$) with time span $T = 20$. Data are generated from the continuous time system (3.8), with $\beta = [1, 1]'$, $x(0) = [3, 1]'$,

$$\Omega = \begin{bmatrix} 1.5 & -0.9 & -0.8 \\ -0.9 & 2 & 0.8 \\ -0.8 & 0.8 & 1 \end{bmatrix},$$

$\mu = [1, 1]'$, $vec(\kappa) = [\kappa_1, 0, 0, \kappa_2]'$ with $\kappa_1 = -0.2$ and $\kappa_2 = -0.4$ in the first case and $\kappa_1 = \kappa_2 = -0.2$ in the second case. Therefore, $a_1 = \exp(-\kappa_1 h) = 1.0168$ and $a_2 = 1.0339$ in the first case and $a_1 = a_2 = 1.0168$ in the second case. In Table 5, we report the percentiles of the finite sample distribution with those of the limit distribution for W_β, W_κ^{OLS}, W_κ^{ReOLS}, and W_κ^{SUR}, where W indicates the Wald test statistic for the parameter of interest, $r_\beta = [1, 1]'$, $r_{a_h} = [1.0168, 1.0339]'$, $r_\kappa = [-0.2, -0.4]'$ in the first case, $r_{a_h} = [1.0168, 1.0168]'$, $r_\kappa = [-0.2, -0.2]'$ in the second case. In addition, OLS corresponds to the OLS estimates, $ReOLS$ to the estimates based on the diagonal elements of OLS estimates, and SUR to the

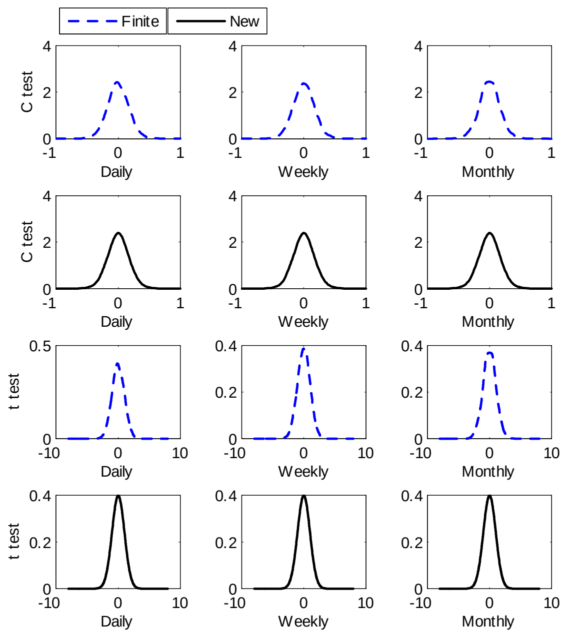

Figure 2: Density comparision between C test and t test for both finite sample distribution and limit distribution, when the initial value is $x_0 = 3$.

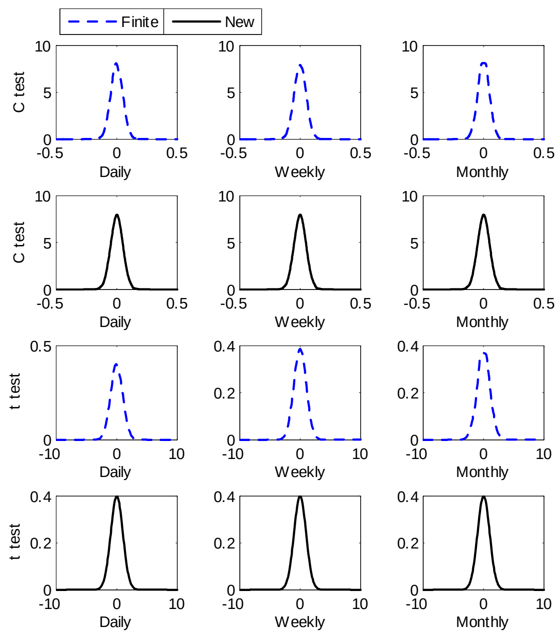

Figure 3: Density comparision between C test and t test for both finite sample distribution and limit distribution, when the initial value is $x_0 = 10$.

Table 5: Comparison of the finite sample distribution and the double asymptotic distributions of the Wald tests related to $\hat{\beta}_{2\times 1}$ and $\widehat{\kappa}_{2\times 1}$.

		Case 1: κ_1 = -0.2, κ_2 = -0.4						Case 2: κ_1 = -0.2, κ_2 = -0.2					
Wald Test	Percentile	1%	2.50%	10%	90%	97.50%	99%	1%	2.50%	10%	90%	97.50%	99%
Asymptotic	χ_g^2	0.020	0.051	0.211	4.605	7.378	9.210	0.020	0.051	0.211	4.605	7.378	9.210
W_β	Finite	0.023	0.057	0.206	4.626	7.373	9.354	0.021	0.055	0.226	4.831	7.871	9.760
W_κ^{OLS}	Finite	0.020	0.048	0.209	5.264	9.215	12.752	7.645	8.953	11.391	20.142	24.165	26.482
W_κ^{ReOLS}	Finite	0.000	0.000	0.000	0.044	1.466	7.980	5.035	6.104	8.082	17.906	22.410	24.899
W_κ^{SUR}	Finite	0.020	0.050	0.201	4.589	7.509	9.688	0.021	0.050	0.225	5.087	8.215	10.347

estimates based on the seemingly unrelated regressions.

Several conclusions can be drawn from this Monte Carlo study. First, for β, our limit distribution well approximates the finite sample distribution in both cases. Second, for κ, the limit distribution is closer to the finite sample distribution based on SUR than those based on OLS or $ReOLS$. While in Case 1 where the explosive roots are distinct, the three finite sample distributions are very close to each other, in Case 2 where there is a common explosive root, the the limit distribution is much closer to the finite sample distribution based on SUR than to those based on either OLS or $ReOLS$, suggesting that one should use the limit distribution to make inference about κ based on SUR.

To understand why SUR provides much better results than OLS for testing hypotheses about κ in Case 2, Table 6 reports the mean and variance of the two sets of estimates of κ in both cases. While SUR produces slightly better estimates than OLS in Case 1, it yields much better estimates of κ in Case 2. As shown in Phillips and Magdalinos (2013), due to the endogeneity problem in the VAR models when there is a common explosive root, the OLS estimate of the common explosive autoregression parameter is biased downward, suggesting that the estimate of κ is biased upward. Naturally this bias distorts the asymptotic appraximation of the Wald statistic.

5 Empirical Illustration for the US Real Estate Market

This section illustrates use of the limit theory in an empirical study of the relationship between the U.S. nationwide real estate market and 13 metropolitan real estate markets respectively between January 2000 and April 2006. We apply the limit theory for univariate co-moving system (2.1)-(2.2) to real estate data using the S&P/Case-Shiller home price composite 20-

Table 6: Finite sample comparison of $\hat{\beta}_{2\times1}$ and $\widehat{\kappa}_{2\times1}$ for the OLS and SUR estimates.

		Case 1: κ_1 = -0.2, κ_2 = -0.4					Case 2: κ_1 = -0.2, κ_2 = -0.2			
Method		OLS		SUR			OLS		SUR	
Parameter	TRUE	Mean	VAR	Mean	VAR	TRUE	Mean	VAR	Mean	VAR
κ_1	-0.200	-0.187	4.60E-03	-0.197	7.16E-04	-0.200	0.042	9.03E-02	-0.197	7.50E-04
κ_2	-0.400	-0.400	1.38E-05	-0.400	2.17E-07	-0.200	0.200	1.05E-01	-0.195	1.12E-03

city index and 13 metropolitan area indices. The S&P/Case-Shiller home price indices are the leading measures of U.S. residential real estate prices, tracking changes in the value of residential real estate nationwide. Monthly data for these indices between February 2000 and August 2014 were downloaded from the St. Louis Fed.[1]

Similar to the capital asset pricing model CAPM, we use the composite 20 index to measure overall market movements. A multi-equation continuous time system (2.1)-(2.2) is estimated with x_t as the composite 20 index and each y_t being one of the 13 metropolitan area indices. The coefficient β then measures the co-movement of each metropolitan area index with the nationwide index. With monthly data, the sampling interval is set to $h = 1/12$. The initial value in each equation of the system is set to the composite 20 index in January 2000, i.e., $x_0 = 100.59$.

We focus on the sample period between January 2000 and April 2006 (in this case $T =$ 6.25). The choice of the sample period is guided by a recent work (Phillips and Yu, 2011) documenting the presence of the explosive behavior in the U.S. real estate market over much of this period. Before estimating the main model (2.1), we examine for the presence of explosive behavior in x_{th} and y_{th} by estimating κ and κ_y. The LS estimate of κ is -0.1187, with the estimated standard error of 0.00045, and the t statistics is -7.147, confirming explosive behavior over this period, consistent with the results in Phillips and Yu (2011). For other city indices y_{th}, we report their estimator of κ_y in the second block of Table 7. In addition, we report the estimated standard error and the t statistic. The second block of Table 7 reports these estimates. The results indicate that all of the 13 metropolitan area indices are explosive over this sample period. Estimates of β for the 13 areas together with 99% and 90% confidence intervals using both the C test and t test are reported in Table 7. The results reported in Table 7 indicate that the coefficient based test produces tighter confidence bands than the t test. Hence, there is some empirical advantage to using the C test. Table 7 shows that the

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

Table 7: Estimated autoregressive coefficients in y_t and confidence intervals for β in U.S. real estate data

City	κ_y	Se(κ_y)	t(κ_y)	β	Se(β)	C test 99% CI		C test 90% CI		t test 99% CI		t test 90% CI	
LA	-0.205	8.414E-04	-7.062	1.170	2.287E-03	1.143	1.198	1.152	1.188	1.047	1.308	1.109	1.257
LasVegas	-0.168	4.127E-03	-2.610	1.045	2.057E-03	1.020	1.072	1.028	1.062	0.929	1.169	0.987	1.124
Miami	-0.320	4.858E-04	-14.518	1.130	2.095E-03	1.104	1.157	1.113	1.147	1.012	1.256	1.071	1.210
Phoenix	-0.455	3.227E-03	-8.018	0.935	1.443E-03	0.914	0.958	0.921	0.949	0.837	1.022	0.886	0.990
DC	-0.167	8.481E-04	-5.725	1.132	1.011E-03	1.114	1.151	1.120	1.144	1.051	1.193	1.092	1.171
Chicago	-0.106	9.811E-04	-3.397	0.886	6.231E-04	0.872	0.901	0.877	0.896	0.822	0.924	0.854	0.910
Boston	0.152	1.236E-03	4.311	1.006	8.130E-04	0.990	1.023	0.995	1.016	0.932	1.055	0.969	1.037
Portland	-0.484	1.004E-03	-15.269	0.832	5.133E-04	0.819	0.845	0.823	0.840	0.773	0.862	0.802	0.851
Dallas	0.233	1.245E-02	2.092	0.753	2.628E-03	0.724	0.783	0.734	0.772	0.621	0.911	0.687	0.853
Detroit	0.213	6.056E-03	2.738	0.772	2.178E-03	0.745	0.799	0.754	0.789	0.652	0.903	0.712	0.855
Seattle	-0.402	7.602E-04	-14.583	0.848	5.522E-04	0.835	0.862	0.839	0.857	0.787	0.881	0.818	0.869
Tampa	-0.368	9.694E-04	-11.819	1.015	5.329E-04	1.002	1.028	1.006	1.023	0.955	1.047	0.985	1.035
NY	-0.111	4.169E-04	-5.432	1.047	7.883E-05	1.042	1.053	1.044	1.051	1.024	1.052	1.036	1.050

90% confidence intervals are quite tight and comfortably reject the null hypothesis $\mathbb{H}_0 : \beta = 0$ 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, $\beta < 1$) the index of the associated metropolitan area moves faster (slower) than the nationwide index, giving a useful perspective on the relationship of different metropolitan area indices to the national index. The results show that LA, Las Vegas, Miami, DC, Boston, Tampa and NY have more "aggressive" real estate markets in the U.S. than the nation as a whole. The epithet "aggressive" is interpreted in the sense that the index for these cities moves more than the countrywide index.

6 Conclusion

This paper studies co-moving systems with explosive regressors in a continuous time framework. The exact discretized model corresponds to a modified version of the discrete time model of Magdalinos and Phillips (2009) but allows for larger initial condition effects and an asymptotically negligible intercept. The limit theory is developed for this modified model, enabling us to obtain double asymptotic limit theory for a continuous time system in which the span $T \to \infty$ and the sampling interval $h \to 0$. The extensions have some important implications for practical work. First, the limit distribution depends explicitly on the initial condition. This dependence mimics a corresponding property in the finite sample distribution and thereby improves the quality of the double asymptotic limit theory as a finite sample approximation. Second, the localized coefficient c, whose counterpart in continuous time is $-\kappa$, is consistently estimable in continuous time using the LS estimator, facilitating a coefficient based test for mildly explosive behavior. Finally, pivotal inference is facilitated in the continuous time case because the sampling interval is known whereas in discrete time system the corresponding localizing rate parameter is unknown.

The double asymptotic limit theory is developed for univariate and multivariate systems in continuous time. Simulations suggest that for the coefficient based test and the t test statistics, these asymptotics well approximate the finite sample distributions. An empirical illustration with US real estate prices at national and various metropolitan areas shows how the methods assist in identifying regions where real estate markets are more aggressive than others.

7 References

Andrews, D. W., and Guggenberger, P. (2012). Asymptotics for LS, GLS, and feasible GLS statistics in an AR (1) model with conditional heteroskedasticity. *Journal of Econometrics*, 169(2), 196-210.

Gourieroux, C. and Zakoian, J. M. (2013). Explosive bubble modeling by non-causal process. CREST Working Paper.

Magdalinos, T. and Phillips, P. C. B. (2009). Limit Theory for Cointegrated Systems with Moderately Integrated and Moderately Explosive Regressors, *Econometric Theory*, 482-526.

Marsaglia, G. (1965). Ratios of normal variables and ratios of sums of uniform variables. *Journal of the American Statistical Association*, 60(309), 193-204.

Phillips, P. C. B. (1972). The structural estimation of a stochastic differential equation system. *Econometrica, 40*, 1021-1041.

Phillips, P. C. B. (1991), Error Correction and Long Run Equilibria in Continuous Time, *Econometrica*, 59, 967-980.

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

Phillips, P. C. B., and Magdalinos, T. (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 Magdalinos, T. (2013). Inconsistent VAR Regression with Common Explosive Roots. *Econometric Theory*, 29(04), 808-837.

Phillips, P. C. B., Shi, S., and Yu, J., (2014). Specification sensitivity in right-tailed unit root testing for explosive behaviour, *Oxford Bulletin of Economics and Statistics*, 76, 315-333.

Phillips, P. C. B., Shi, S., and Yu, J., (2015a). Testing for multiple bubbles: Historical episodes of exuberance and collapse in the S&P 500, *International Economic Review*, forthcoming.

Phillips, P. C. B., Shi, S., and Yu, J., (2015b), Testing for Multiple Bubbles: Limit Theory of Dating Algorithms, *International Economic Review*, forthcoming.

Phillips, P. C. B., Wu, Y., and Yu, J., (2011) Explosive behavior in the 1990s NASDAQ: When did exuberance escalate asset values? *International Economic Review*,, 52: 201-226.

Phillips, P. C. B., and Yu, J., (2011) Dating the timeline of financial bubbles during the subprime crisis, *Quantitative Economics*, 2, 455-491.

Wang, X., and Yu, J., (2014). Double Asymptotics for an Explosive Continuous Time Model, Working Paper, Singapore Management University.

8 Appendix

Proof of Theorem 2.1. The arguments here and in much of what follows closely mirror those of MP (2009) in the mildly explosive case. We therefore provide only the main new details here. The limit theory of $\sum_{t=1}^{N} x_t^2$ and $\sum_{t=1}^{N} x_t u_{0t}$ is obtained using split sample arguments replacing summations in $\sum_{t=1}^{N}$ by $\left(\sum_{t=1}^{m_N} + \sum_{t=m_N+1}^{N}\right)$ where m_N is such that $\frac{m_N}{N^\alpha} + \frac{N}{m_N} \to \infty$ so that with $c > 0$ and $\alpha \in (0, 1)$ we have

$$R_N^{-m_N} \sim \left(1 + \frac{c}{N^\alpha}\right)^{-m_N} \to 0, \ \frac{N^\alpha}{R_N^{N-m_N}} \to 0. \tag{8.1}$$

(i) Start by writing x_t in (2.9) as:

$$x_t = R_N^t x_0 + \sum_{j=1}^{t} R_N^{t-j} u_{xj} + \frac{1 - R_N^t}{1 - R_N}\mu, \tag{8.2}$$

so the standardized numerator can be decomposed as

$$
\begin{aligned}
& \left(R_N^N N^\alpha\right)^{-1} \sum_{t=1}^N x_t u_{0t} \\
= \; & \left(R_N^N N^\alpha\right)^{-1} \sum_{t=1}^N R_N^t u_{0t} \left(x_0 - \frac{\mu}{1-R_N}\right) + \left(R_N^N N^\alpha\right)^{-1} \sum_{t=1}^N u_{0t} \frac{\mu}{1-R_N} \\
& + \frac{R_N^{-N}}{\sqrt{N^\alpha}} \left(\sum_{t=1}^{m_N} + \sum_{t=m_N+1}^{N}\right) u_{0t} \left(\frac{1}{\sqrt{N^\alpha}} \sum_{j=1}^t R_N^{t-j} u_{xj}\right). \qquad (8.3)
\end{aligned}
$$

For the first term on the right hand side of (8.3), since $\widetilde{x}_0 = x_{0N} N^{-\alpha/2} \Rightarrow X^*$ and $\widetilde{\mu} = N^{\alpha/2}\mu \Rightarrow \mu^*$, we have

$$
D_N = N^{-\alpha/2} x_0 - \frac{N^{-\alpha/2}\mu}{1-R_N} = N^{-\alpha/2} x_0 - \frac{N^{\alpha/2}\mu}{-c+o(1)} \Rightarrow X^* + \frac{\mu^*}{c} = D. \qquad (8.4)
$$

setting $D_N = N^{-\alpha/2}\left(x_0 - \frac{\mu}{1-R_N}\right)$, we then have,

$$
\begin{aligned}
& \left(R_N^N N^\alpha\right)^{-1} \sum_{t=1}^N R_N^t u_{0t} \left(x_0 - \frac{\mu}{1-R_N}\right) \\
= \; & N^{-\alpha/2} \sum_{t=1}^N R_N^{-(N-t)} u_{0t} \left\{N^{-\alpha/2}\left(x_0 - \frac{\mu}{1-R_N}\right)\right\} = D_N N^{-\alpha/2} \sum_{t=1}^N R_N^{-(N-t)} u_{0t} \\
= \; & D N^{-\alpha/2} \sum_{t=m_N+1}^N R_N^{-(N-t)} u_{0t} + o_p(1), \text{ with } D := X^* + \frac{\mu^*}{c}. \qquad (8.5)
\end{aligned}
$$

where we assume the probability space is expanded in such a way so that the weak convergence $\Rightarrow$ can be replaced by $\xrightarrow{p}$. Also note that

$$
\begin{aligned}
\mathbb{E}\left\{N^{-\alpha/2} \sum_{t=1}^{m_N} R_N^{-(N-t)} u_{0t}\right\}^2 &= \sigma_{00} N^{-\alpha} \sum_{t=1}^{m_N} R_N^{-2(N-t)} = \sigma_{00} N^{-\alpha} R_N^{-2N+1} \frac{1-R_N^{2m_N}}{1-R_N^2} \\
&= \sigma_{00} \frac{R_N^{-2N+1} - R_N^{-2N+1+2m_N}}{-2c} = o(1),
\end{aligned}
$$

so that

$$
N^{-\alpha/2} \sum_{t=1}^{m_N} R_N^{-(N-t)} u_{0t} = o_p(1), \qquad (8.6)
$$

and then

$$
\left(R_N^N N^\alpha\right)^{-1} \sum_{t=1}^{m_N} R_N^t u_{0t} \left(x_0 - \frac{\mu}{1-R_N}\right) = o_p(1).
$$

Hence, for the first term of (8.3) we have by virtue of the martingale central limit theorem (MCLT), as in Phillips and Magdalinos (2007),

$$
\begin{aligned}
\left(R_N^N N^{\alpha}\right)^{-1} \sum_{t=1}^{N} R_N^t u_{0t} \left(x_0 - \frac{\mu}{1-R_N}\right) &= D_N \left(R_N^N N^{\alpha/2}\right)^{-1} \sum_{t=1}^{N} R_N^t u_{0t} \\
= \quad D \frac{1}{\sqrt{N^{\alpha}}} \sum_{t=m_N+1}^{N} R_N^{-(N-t)} u_{0t} + o_p(1) &= D \left(\frac{1}{\sqrt{N^{\alpha}}} \sum_{k=0}^{N-m_N-1} R_N^{-k} u_{0N-k} \right) \\
= \quad D \left(\frac{1}{\sqrt{N^{\alpha}}} \sum_{k=0}^{N-m_N-1} R_N^{-k} u'_{0k} \right) &\Rightarrow \frac{D\sigma_{00}}{(2c)^{1/2}} U_0,
\end{aligned}
$$

where $u'_{0k} := u_{0N-k} \overset{iid}{\sim} \left(0, \sigma_{00}^2\right)$ and $U_0 = N(0,1)$ since $\frac{1}{N^{\alpha}} \sum_{k=0}^{N-m_N-1} R_N^{-2k} = \frac{1}{N^{\alpha}} \frac{1-R_N^{-2N}}{1-R_N^{-2}} \sim \frac{1}{N^{\alpha}} \frac{1}{R_N^2 - 1} \to \frac{1}{2c}$.

For the second term on the right hand side of (8.3), noting that $R_N^{-N}\sqrt{N} = \left(1+\frac{c}{N^{\alpha}}\right)^{-N}\sqrt{N} = O\left(e^{-c\frac{N}{N^{\alpha}}}\sqrt{N}\right) = o(1)$ for all $\alpha \in (0,1)$ we obtain

$$
\left(R_N^N N^{\alpha}\right)^{-1} \sum_{t=1}^{N} u_{0t} \frac{\mu}{1-R_N} \sim R_N^{-N}\sqrt{N} \frac{\mu}{-c} \frac{1}{\sqrt{N}} \sum_{t=1}^{N} u_{0t} = O\left(e^{-c\frac{N}{N^{\alpha}}} \sqrt{\frac{N}{N^{\alpha}}}\right) O_p(1) = o_p(1). \tag{8.7}
$$

The third term on the right hand side of (8.3) is

$$
\begin{aligned}
& \frac{R_N^{-N}}{\sqrt{N^{\alpha}}} \left(\sum_{t=1}^{m_N} + \sum_{t=m_N+1}^{N} \right) u_{0t} \left(\frac{1}{\sqrt{N^{\alpha}}} \sum_{j=1}^{t} R_N^{t-j} u_{xj} \right) \\
= \quad & \frac{1}{\sqrt{N^{\alpha}}} \sum_{t=1}^{m_N} R_N^{-(N-t)} u_{0t} \left(\frac{1}{\sqrt{N^{\alpha}}} \sum_{j=1}^{t} R_N^{-j} u_{xj} \right) + \frac{1}{\sqrt{N^{\alpha}}} \sum_{t=m_N+1}^{N} R_N^{-(N-t)} u_{0t} \left(\frac{1}{\sqrt{N^{\alpha}}} \sum_{j=1}^{t} R_N^{-j} u_{xt} \right) \\
= \quad & \frac{1}{\sqrt{N^{\alpha}}} \sum_{t=m_N+1}^{N} R_N^{-(N-t)} u_{0t} \left(\frac{1}{\sqrt{N^{\alpha}}} \sum_{j=1}^{m_N} R_N^{-j} u_{xj} \right) + o_p(1),
\end{aligned}
$$

where we use the fact that $N^{-\alpha/2} \sum_{t=1}^{m_N} R_N^{-(N-t)} u_{0t} = o_p(1)$ from (8.6). We now use a joint MCLT for the components

$$
\begin{aligned}
(U_{0N}, U_{xN}) \quad &= \quad \left(\frac{1}{\sqrt{N^{\alpha}}} \sum_{t=m_N+1}^{N} R_N^{-(N-t)} u_{0t}, \frac{1}{\sqrt{N^{\alpha}}} \sum_{j=1}^{m_N} R_N^{-j} u_{xj} \right) \\
&= \quad \left(\frac{1}{\sqrt{N^{\alpha}}} \sum_{t=1}^{N} R_N^{-(N-t)} u_{0t}, \frac{1}{\sqrt{N^{\alpha}}} \sum_{j=1}^{N} R_N^{-j} u_{xj} \right) + o_p(1)
\end{aligned}
$$

$$\Rightarrow \left(\frac{\sigma_{00}}{(2c)^{1/2}} U_0, \frac{\sigma_{xx}}{(2c)^{1/2}} U_x \right) \text{ with } \left(U_0', U_x' \right)' \sim N(0, I_2),$$

just as in Phillips and Magdalinos (2007) and MP (2009), using the fact that the limit variates (U_0, U_x) are independent because

$$\mathbb{E}\left\{ \left(\frac{1}{\sqrt{N^\alpha}} \sum_{t=1}^{N} R_N^{-(N-t)} u_{0t} \right) \left(\frac{1}{\sqrt{N^\alpha}} \sum_{j=1}^{N} R_N^{-j} u_{xj} \right) \right\} = \frac{N^{1-\alpha}}{R_N^N} \sigma_{0x} \to 0.$$

Hence

$$\begin{aligned} \frac{R_N^{-N}}{\sqrt{N^\alpha}} \left(\sum_{t=1}^{m_N} + \sum_{t=m_N+1}^{N} \right) u_{0t} \left(\frac{1}{\sqrt{N^\alpha}} \sum_{j=1}^{t} R_N^{t-j} u_{xj} \right) &= \frac{1}{\sqrt{N^\alpha}} \sum_{t=m_N+1}^{N} R_N^{-(N-t)} u_{0t} \left(\frac{1}{\sqrt{N^\alpha}} \sum_{j=1}^{m_N} R_N^{-j} u_{xj} \right) + o_p(1) \\ &\Rightarrow \left(\frac{\sigma_{00}}{(2c)^{1/2}} U_0 \right) \left(\frac{\sigma_{xx}}{(2c)^{1/2}} U_x \right) = \frac{\sigma_{00}\sigma_{xx}}{2c} U_0 U_x. \end{aligned}$$

Combining the above results and using (8.4) we obtain

$$\begin{aligned} & \left(R_N^N N^\alpha \right)^{-1} \sum_{t=1}^{N} x_t u_{0t} \\ = \;& \left(R_N^N N^\alpha \right)^{-1} \sum_{t=1}^{N} R_N^t u_{0t} \left(x_0 - \frac{\mu}{1 - R_N} \right) + \left(R_N^N N^\alpha \right)^{-1} \sum_{t=1}^{N} u_{0t} \frac{\mu}{1 - R_N} \\ & + \frac{R_N^{-N}}{\sqrt{N^\alpha}} \sum_{t=m_N+1}^{N} u_{0t} \left(\frac{1}{\sqrt{N^\alpha}} \sum_{j=1}^{t} R_N^{t-j} u_{xj} \right) + o_p(1) \\ = \;& \left(R_N^N N^\alpha \right)^{-1} \sum_{t=1}^{N} R_N^t u_{0t} \left(x_0 - \frac{\mu}{1 - R_N} \right) + \frac{R_N^{-N}}{\sqrt{N^\alpha}} \sum_{t=m_N+1}^{N} u_{0t} \left(\frac{1}{\sqrt{N^\alpha}} \sum_{j=1}^{t} R_N^{t-j} u_{xj} \right) + o_p(1) \\ \Rightarrow \;& \frac{D\sigma_{00}}{(2c)^{1/2}} U_0 + \frac{\sigma_{00}\sigma_{xx}}{2c} U_0 U_x = \frac{\sigma_{00}}{(2c)^{1/2}} U_0 \left(D + \frac{\sigma_{xx}}{(2c)^{1/2}} U_x \right), \end{aligned} \tag{8.8}$$

giving the limit of the numerator.

(ii) From the identity

$$x_t^2 = R_N^2 x_{t-1}^2 + \mu^2 + u_{xt}^2 + 2R_N \mu x_{t-1} + 2R_N x_{t-1} u_{xt} + 2\mu u_{xt},$$

we have

$$\left(R_N^2 - 1 \right) \sum_{t=1}^{N} x_t^2 = R_N^2 x_N^2 - R_N^2 x_0^2 - 2R_N \mu \sum_{t=1}^{N} x_{t-1} - 2R_N \sum_{t=1}^{N} x_{t-1} u_{xt} - N\mu^2 - \sum_{t=1}^{N} u_{xt}^2 - 2\mu \sum_{t=1}^{N} u_{xt}. \tag{8.9}$$

We show in the following that each of the following standardized terms

$$\frac{R_N^2 x_0^2}{R_N^{2N} N^\alpha}, \frac{N\mu^2}{R_N^{2N} N^\alpha}, \frac{2\mu \sum_{t=1}^N u_{xt}}{R_N^{2N} N^\alpha}, \frac{\sum_{t=1}^N u_{xt}^2}{R_N^{2N} N^\alpha}, \frac{R_N \mu \sum_{t=1}^N x_{t-1}}{R_N^{2N} N^\alpha}, \ \frac{R_N \sum_{t=1}^N x_{t-1} u_{xt}}{R_N^{2N} N^\alpha}$$

are asymptotically negligible. In particular, since the standardized initial condition and drift satisfy $\widetilde{x}_0 = x_{0N} N^{-\alpha/2} \Rightarrow X^*$ and $\widetilde{\mu} = N^{\alpha/2}\mu \Rightarrow \mu^*$ we find that

$$\frac{R_N^2 x_0^2}{R_N^{2N} N^\alpha} = O_p\left(\left(\frac{x_0}{N^{\alpha/2}}\right)^2 \frac{1}{R_N^{2N}}\right) = o_p(1),$$

$$\frac{N\mu^2}{R_N^{2N} N^\alpha} = O_p\left(\frac{N^{1-2\alpha}}{R_N^{2N}}\right) = o_p(1),$$

$$\frac{2\mu \sum_{t=1}^N u_{xt}}{R_N^{2N} N^\alpha} = \left(\frac{1}{R_N^{2N}}\right)\left(\frac{2\mu\sqrt{N}}{N^\alpha}\right)\left(\frac{1}{\sqrt{N}}\sum_{t=1}^N u_{xt}\right) = O_p\left(\frac{1}{R_N^{2N}}\right) \times O_p\left(N^{\frac{1}{2}-\frac{3}{2}\alpha}\right) \times O_p(1) = o_p(1),$$

$$\frac{\sum_{t=1}^N u_{xt}^2}{R_N^{2N} N^\alpha} = \frac{N}{R_N^{2N} N^\alpha} \frac{1}{N} \sum_{t=1}^N u_{xt}^2 = O_p\left(\frac{N^{1-\alpha}}{R_N^{2N}}\right) \times O_p(1) = o_p(1),$$

$$\frac{R_N \sum_{t=1}^N x_{t-1} u_{xt}}{R_N^{2N} N^\alpha} = O_p\left(\frac{\sum_{t=1}^N x_{t-1} u_{xt}}{R_N^N N^\alpha}\right) \times O_p\left(\frac{1}{R_N^N}\right) = o_p(1),$$

since $\left(R_N^N N^\alpha\right)^{-1} \sum_{t=1}^N x_{t-1} u_{xt} = O_p(1)$ just as in the analysis of $\left(R_N^N N^\alpha\right)^{-1} \sum_{t=1}^N x_t u_{0t}$ in part (i); and finally

$$\frac{R_N \mu \sum_{t=1}^N x_t}{R_N^{2N} N^\alpha} = \frac{\mu}{N^{\alpha/2} R_N^N} \sum_{t=1}^N \frac{x_t}{N^{\alpha/2} R_N^t} \frac{R_N^t}{R_N^N} = O_p\left(\frac{1}{N^\alpha R_N^N}\right) \times O_p(N) = o_p(1).$$

Hence, from (8.9) and (8.2) we deduce that

$$\begin{aligned}
\frac{\left(R_N^2 - 1\right) \sum_{t=1}^N x_t^2}{R_N^{2N} N^\alpha} &= \frac{R_N^2 x_N^2}{R_N^{2N} N^\alpha} \{1 + o_p(1)\} = \left(\frac{x_N}{R_N^N N^{\alpha/2}}\right)^2 \{1 + o_p(1)\} \\
&= \left\{\frac{x_0}{N^{\alpha/2}} + \frac{1}{N^{\alpha/2}} \sum_{j=1}^N R_N^{-j} u_{xj} + \frac{N^{\alpha/2}}{c}\mu\right\}^2 \{1 + o_p(1)\} \\
&\Rightarrow \left(\frac{\sigma_{xx}}{(2c)^{1/2}} U_x + D\right)^2. \qquad (8.10)
\end{aligned}$$

(iii) Combining the results (8.8) and (8.10), we have

$$\frac{\left(R_N^N N^\alpha\right)^{-1} \sum_{t=1}^N x_t u_{xt}}{\frac{\left(R_N^2-1\right)}{R_N^{2N} N^\alpha} \sum_{t=1}^N x_t^2} \sim \frac{\left(R_N^N N^\alpha\right)^{-1} \sum_{t=1}^N x_t u_{xt}}{\frac{2c}{R_N^{2N} N^{2\alpha}} \sum_{t=1}^N x_t^2} \Rightarrow \frac{\frac{\sigma_{00}}{(2c)^{1/2}} U_0 \left(D + \frac{\sigma_{xx}}{(2c)^{1/2}} U_x\right)}{\left(\frac{\sigma_{xx}}{(2c)^{1/2}} U_x + D\right)^2}$$

$$= \frac{\frac{\sigma_{00}}{(2c)^{1/2}} U_0}{\frac{\sigma_{xx}}{(2c)^{1/2}} U_x + D} = \frac{\sigma_{00} U_0}{\sigma_{xx} U_x + (2c)^{1/2} D}.$$

Therefore,

$$R_N^N N^\alpha \left(\widehat{A} - A \right) = \frac{\left(R_N^N N^\alpha \right)^{-1} \sum_{t=1}^N x_t u_{xt}}{\left(R_N^{2N} N^{2\alpha} \right)^{-1} \sum_{t=1}^N x_t^2} \Rightarrow 2c \frac{\sigma_{00} U_0}{\sigma_{xx} U_x + (2c)^{1/2} D},$$

giving the stated result. ■

Proof of Theorem 2.2. The proof follows from Theorem 2.1 by noting the mappings

$$\begin{aligned} \sigma_{00}^2 &\mapsto \tilde{\sigma}_{00}^2 = \frac{\sigma_{00}^2}{\sigma_{xx}^2}, \sigma_{xx}^2 \mapsto 1, R_N \mapsto a_h = e^{-\kappa h}, X^* \mapsto \frac{x_0}{\sigma_{xx}}, \mu \mapsto \frac{\mu\kappa}{\sigma_{xx}} h^{1/2}, \mu^* \mapsto \frac{\mu\kappa}{\sigma_{xx}}, \\ D_N &\mapsto D_h = \tilde{x}_{0h} h^{1/2} - \frac{h^{-1/2} \tilde{g}_h}{\kappa} \to D^* = \frac{x_0}{\sigma_{xx}} - \frac{\mu}{\sigma_{xx}}, \end{aligned}$$

with $h = 1/N^\alpha$. It follows that

$$a_h^{-N} h \sum_{t=1}^N \tilde{x}_{th} \tilde{u}_{0,th} \Rightarrow \frac{\tilde{\sigma}_{00}}{-2\kappa} U_0 \left(U_x + (-2\kappa)^{1/2} D^* \right), \tag{8.11}$$

$$a_h^{-2N} h^2 \sum_{t=1}^N \tilde{x}_{th}^2 \Rightarrow \left(\frac{1}{-2\kappa} \right)^2 \left(U_x + (-2\kappa)^{1/2} D^* \right)^2, \tag{8.12}$$

and hence

$$\frac{a_h^N}{h} \left(\widehat{\beta} - \beta \right) \Rightarrow (-2\kappa) \frac{\widetilde{\sigma}_{00} U_0}{U_x + (-2\kappa)^{1/2} D^*} = (-2\kappa) \frac{\sigma_{00} U_0}{\sigma_{xx} U_x + (-2\kappa)^{1/2} (x_0 - \mu)}.$$

■

Proof of Theorem 3.1. First, we rewrite x_t by backward recursion as,

$$x_t = (I - R_N)^{-1} \mu + R_N^t \left(x_0 - (I - R_N)^{-1} \mu \right) + \sum_{j=1}^t R_N^{t-j} u_{xj}.$$

(i) With this expression, we have

$$\begin{aligned} & vec \left(\frac{1}{N^\alpha} \sum_{t=1}^N u_{0t} x_t' R_N^{-N} \right) \qquad (8.13) \\ = \ & vec \left(\frac{1}{N^\alpha} \sum_{t=1}^N u_{0t} \left\{ (I - R_N)^{-1} \mu + R_N^t \left(x_0 - (I - R_N)^{-1} \mu \right) + \sum_{j=1}^t R_N^{t-j} u_{xj} \right\}' R_N^{-N} \right) \\ = \ & \frac{1}{N^\alpha} vec \left(\sum_{t=1}^N u_{0t} \left\{ x_0 - (I - R_N)^{-1} \mu \right\}' R_N^{t-N} \right) + \frac{1}{N^\alpha} vec \left(\sum_{t=1}^N u_{0t} \mu' (I - R_N)^{-1} R_N^{-N} \right) \end{aligned}$$

$$+\frac{1}{N^{\alpha}}vec\left(\left(\sum_{t=1}^{m_N}+\sum_{t=m_N+1}^{N}\right)u_{0t}\left(\sum_{j=1}^{t}R_N^{t-j}u_{xj}\right)'R_N^{-N}\right).$$

For the first item on the right side of (8.13), letting $D_N = N^{-\alpha/2}\left(x_0 + N^{\alpha}C^{-1}\mu\right)$, we obtain

$$\begin{aligned}
&\frac{1}{N^{\alpha}}vec\left(\sum_{t=1}^{N}u_{0t}\left\{x_0-(I-R_N)^{-1}\mu\right\}'R_N^{t-N}\right)\\
=\;&\frac{1}{N^{\alpha/2}}vec\left(\sum_{t=1}^{N}u_{0t}D_N'R_N^{t-N}\right)=\frac{1}{N^{\alpha/2}}\sum_{t=1}^{N}\left(R_N^{t-N}\otimes u_{0t}\right)vec\left(D_N'\right)\\
=\;&\frac{1}{N^{\alpha/2}}\sum_{t=m_N+1}^{N}\left(R_N^{t-N}\otimes u_{0t}\right)D\;+o_p(1)\;\text{ with } D:=X^*+C^{-1}\mu^*,
\end{aligned}$$

since $D_N = N^{-\alpha/2}x_0 + N^{\alpha/2}C^{-1}\mu \Rightarrow X^* + C^{-1}\mu^* = D$ and by replacing the weak convergence with convergence in probability in an expanded space for the final step. In addition, we have

$$\begin{aligned}
&\mathbb{E}\left\|\frac{1}{N^{\alpha/2}}\sum_{t=1}^{m_N}\left(R_N^{t-N}\otimes u_{0t}\right)\right\|^2\\
=\;&N^{-\alpha}\sum_{t=1}^{m_N}\left\|R_N^{2(t-N)}\right\|\mathbb{E}\left\|u_{0t}\right\|^2=\frac{N^{-\alpha}\left\|R_N\right\|^{-2(N-1)}\left(1-\left\|R_N\right\|^{2m_N}\right)}{1-\left\|R_N\right\|^2}\mathbb{E}\left\|u_{0t}\right\|^2\\
=\;&\frac{\mathbb{E}\left\|u_{0t}\right\|^2}{-\max_{1\le i\le K}c_i}o(1)=o(1)\;\text{ assuming }\mathbb{E}\left\|u_{0t}\right\|^2<\infty.
\end{aligned}$$

The result implies

$$\frac{1}{N^{\alpha/2}}\sum_{t=1}^{m_N}\left(R_N^{t-N}\otimes u_{0t}\right)=o_p(1)\;\text{ and }\;\frac{1}{N^{\alpha/2}}\sum_{t=1}^{m_N}\left(R_N^{t-N}\otimes u_{0t}\right)D=o_p(1).$$

Hence, for the first item of (8.13), we have

$$\begin{aligned}
&\frac{1}{N^{\alpha}}vec\left(\sum_{t=1}^{N}u_{0t}\left\{x_0-(I-R_N)^{-1}\mu\right\}'R_N^{t-N}\right)\\
=\;&\frac{1}{N^{\alpha/2}}\sum_{t=m_N+1}^{N}\left(R_N^{t-N}\otimes u_{0t}\right)D+o_p(1)=\frac{1}{N^{\alpha/2}}\sum_{j=0}^{N-m_N-1}\left(R_N^{-j}D\otimes u_{0N-j}\right)\\
=\;&\frac{1}{N^{\alpha/2}}\sum_{j=0}^{N-m_N-1}\left(R_N^{-j}D\otimes\widetilde{u}_{0j}\right),
\end{aligned}$$

where $\widetilde{u}_{0j} = u_{0N-j} \overset{d}{=} N(0,\Omega_{00})$.

For the second item of (8.13), we have

$$
\begin{aligned}
& \frac{1}{N^{\alpha}} vec\left(\sum_{t=1}^{N} u_{0t}\mu^{'}\left(I-R_N\right)^{-1} R_N^{-N}\right) \\
= \; & \frac{-1}{N^{\alpha}} vec\left(\sum_{t=1}^{N} u_{0t}\mu^{'} C^{-1} R_N^{-N} N^{\alpha}\right) = vec\left(-\sum_{t=1}^{N} \frac{u_{0t}}{\sqrt{N}}\mu^{'} C^{-1} R_N^{-N}\sqrt{N}\right) = o_p\left(1\right),
\end{aligned}
$$

since $R_N^{-N}\sqrt{N} = O\left(e^{-CN^{1-\alpha}}\sqrt{N}\right) = o_p\left(1\right)$.

For the third item of (8.13), we have

$$
\begin{aligned}
& \frac{1}{N^{\alpha}} vec\left(\left(\sum_{t=1}^{m_N} + \sum_{t=m_N+1}^{N}\right) u_{0t}\left(\sum_{j=1}^{t} R_N^{t-j} u_{xj}\right)^{'} R_N^{-N}\right) \\
= \; & \frac{1}{N^{\alpha}} vec\left(\left(\sum_{t=1}^{m_N} + \sum_{t=m_N+1}^{N}\right) u_{0t}\left(\sum_{j=1}^{N} R_N^{t-j} u_{xj}\right)^{'} R_N^{-N}\right) + o_p\left(1\right) \\
= \; & \left(\frac{1}{N^{\alpha/2}}\sum_{t=1}^{m_N} R_N^{-(N-t)} \otimes u_{0t}\right) vec\left(\frac{1}{N^{\alpha/2}}\sum_{j=1}^{N} R_N^{-j} u_{xj}\right)^{'} + \\
& \left(\frac{1}{N^{\alpha/2}}\sum_{t=m_N+1}^{N} R_N^{-(N-t)} \otimes u_{0t}\right) vec\left(\frac{1}{N^{\alpha/2}}\sum_{j=1}^{N} R_N^{-j} u_{xj}\right)^{'} \\
= \; & \left(\frac{1}{N^{\alpha/2}}\sum_{t=m_N+1}^{N} R_N^{-(N-t)} \otimes u_{0t}\right)\left(\frac{1}{N^{\alpha/2}}\sum_{j=1}^{m_N} R_N^{-j} u_{xt}\right) + o_p\left(1\right) \\
= \; & \frac{1}{N^{\alpha/2}}\sum_{t=m_N+1}^{N} R_N^{-(N-t)} U_{xN} \otimes u_{0t} + o_p\left(1\right), \text{ with } U_{xN} = \sum_{j=1}^{m_N} R_N^{-j} u_{xt},
\end{aligned}
$$

since we have shown $\sum_{t=1}^{m_N} R_N^{-(N-t)} \otimes u_{0t} = o_p\left(1\right)$. Note that, from MP

$$
U_{xN} \Rightarrow \left(\int_0^{\infty} e^{-pC}\Omega_{xx}e^{-pC}dp\right)^{1/2} U_x =: \widetilde{U}_x,
$$

with $U_x = N\left(0, I_K\right)$. Hence the third item has the following form:

$$
\begin{aligned}
& \frac{1}{N^{\alpha}} vec\left(\sum_{t=1}^{m_N} + \sum_{t=m_N+1}^{N}\right) u_{0t}\left(\sum_{j=1}^{t} R_N^{t-j} u_{xj}\right)^{'} R_N^{-N} \\
= \; & \frac{1}{N^{\alpha/2}}\sum_{t=m_N+1}^{N} R_N^{-(N-t)} U_{xN} \otimes u_{0t} + o_p\left(1\right).
\end{aligned}
$$

Combining the above results and using (8.13), we have the limit result for the numerator as,

$$
\begin{aligned}
& vec\left(\frac{1}{N^{\alpha}}\sum_{t=1}^{N} u_{0t}x_t^{'}R_N^{-N}\right) \\
= \; & \frac{1}{N^{\alpha/2}}\sum_{t=m_N+1}^{N} R_N^{-(N-t)}\left(U_{xN}+D\right)\otimes u_{0t}+o_p\left(1\right) \\
\Rightarrow \; & \left(\int_0^{\infty} e^{-pC}\left(D+\widetilde{U}_x\right)\left(D+\widetilde{U}_x\right)' e^{-pC}dp\otimes\Omega_{00}\right)^{1/2} W_0 \\
\overset{d}{=} \; & MN\left(0,\int_0^{\infty} e^{-pC}\left(D+\widetilde{U}_x\right)\left(D+\widetilde{U}_x\right)' e^{-pC}dp\otimes\Omega_{00}\right), \qquad (8.14)
\end{aligned}
$$

where $W_0 = N\left(0, I_{mK}\right)$. Due to the sample splitting at $t = m_N$, as $N \to \infty$ the limit variate W_0 is independent of the limit variate U_x.

(ii) From the identity

$$x_t x_t' = \mu\mu' + R_N x_{t-1}\mu' + u_{xt}\mu^{'} + \mu x_{t-1}' R_N + R_N x_{t-1}x_{t-1}'R_N' + u_{xt}x_{t-1}'R_N + \mu u_{xt}^{'} + R_N x_{t-1}u_{xt}^{'} + u_{xt}u_{xt}',$$

we have

$$
\begin{aligned}
& \left(R_N\otimes R_N - I_{K\times K}\right)\sum_{t=1}^{N} vec\left(x_t x_t'\right) \qquad (8.15) \\
= \; & \left(R_N\otimes R_N\right) vec\left(x_N x_N'\right) - \left(R_N\otimes R_N\right) vec\left(x_0 x_0'\right) - N vec\left(\mu\mu'\right) - \sum_{t=1}^{N} vec\left(R_N x_{t-1}\mu'\right) - \sum_{t=1}^{N} vec\left(u_{xt}\mu^{'}\right) \\
& - \sum_{t=1}^{N} vec\left(\mu x_{t-1}' R_N\right) - \sum_{t=1}^{N} vec\left(u_{xt}x_{t-1}'R_N\right) - \sum_{t=1}^{N} vec\left(\mu u_{xt}^{'}\right) - \sum_{t=1}^{N} vec\left(R_N x_{t-1}u_{xt}^{'}\right) - \sum_{t=1}^{N} vec\left(u_{xt}u_{xt}'\right).
\end{aligned}
$$

We show in the following that each of the following terms standardized by $N^{-\alpha}\left(R_N^{-N}\otimes R_N^{-N}\right)$

$$\left(R_N\otimes R_N\right) vec\left(x_0 x_0'\right), N vec\left(\mu\mu'\right), \sum_{t=1}^{N} vec\left(R_N x_{t-1}\mu'\right), \sum_{t=1}^{N} vec\left(\mu u_{xt}'\right), \sum_{t=1}^{N} vec\left(u_{xt}x_{t-1}'R_N\right), \sum_{t=1}^{N} vec\left(u_{xt}u_{xt}'\right)$$

are asymptotically negligible. In particular, we have

$$
\begin{aligned}
& N^{-\alpha}\left(R_N^{-N}\otimes R_N^{-N}\right)\left(R_N\otimes R_N\right) vec\left(x_0 x_0'\right) \\
= \; & \left(R_N^{-(N-1)}\otimes R_N^{-(N-1)}\right) vec\left(N^{-\alpha/2}x_0 x_0' N^{-\alpha/2}\right) = o_p\left(1\right),
\end{aligned}
$$

$$N^{-\alpha}\left(R_N^{-N}\otimes R_N^{-N}\right) N vec\left(\mu\mu'\right) = O_p\left(N^{1-2\alpha}R_N^{-N}\otimes R_N^{-N}\right) = o_p\left(1\right),$$

$$N^{-\alpha}\left(R_N^{-N}\otimes R_N^{-N}\right)\sum_{t=1}^{N} vec\left(\mu u_{xt}'\right)$$

$$
\begin{aligned}
&= N^{\frac{1}{2}-\alpha}\left(R_N^{-N}\otimes R_N^{-N}\right)(I_K\otimes\mu)\frac{1}{\sqrt{N}}\sum_{t=1}^{N}u_{xt} \\
&= O_p\left(N^{\frac{1-3\alpha}{2}}\right)O_p\left(R_N^{-N}\otimes R_N^{-N}\right)O_p(1)=o_p(1),
\end{aligned}
$$

$$
\begin{aligned}
& N^{-\alpha}\left(R_N^{-N}\otimes R_N^{-N}\right)\sum_{t=1}^{N}vec\left(u_{xt}x_{t-1}'R_N\right) \\
&= N^{-\alpha}\left(R_N^{-N}\otimes R_N^{-N}\right)(R_N\otimes I_K)\sum_{t=1}^{N}vec\left(u_{xt}x_{t-1}'\right) \\
&= R_N^{-(N-1)}\otimes R_N^{-N}N^{-\alpha}\sum_{t=1}^{N}vec\left(u_{xt}x_{t-1}'\right)=O_p\left(R_N^{-(N-1)}\right)\otimes O_p(1)=o_p(1),
\end{aligned}
$$

since $R_N^{-N}N^{-\alpha}\sum_{t=1}^{N}vec\left(u_{xt}x_{t-1}'\right)=O_p(1)$ following the same argument that $R_N^{-N}N^{-\alpha}\sum_{t=1}^{N}vec\left(u_{0t}x_t'\right)=O_p(1)$. Finally,

$$
\begin{aligned}
& N^{-\alpha}\left(R_N^{-N}\otimes R_N^{-N}\right)\sum_{t=1}^{N}vec\left(R_Nx_{t-1}\mu'\right) \\
&= N^{-\alpha}\left(R_N^{-N}\otimes R_N^{-N}\right)(\mu\otimes R_N)\sum_{t=1}^{N}x_{t-1}=N^{-\alpha}\left(R_N^{-N}\mu\otimes R_N^{-(N-1)}\right)\sum_{t=1}^{N}x_{t-1} \\
&= R_N^{-N}\mu N^{-\alpha/2}\otimes\sum_{t=1}^{N}R_N^{-(N-1-t)}N^{-\alpha/2}R_N^{-t}x_{t-1}=R_N^{-N}\mu N^{-\alpha/2}\otimes O_p\left(N^{\alpha}\right).
\end{aligned}
$$

Therefore, for the denominator, we have from (8.15) that

$$
\begin{aligned}
& N^{-\alpha}\left(R_N\otimes R_N-I_{K\times K}\right)\left(R_N^{-N}\otimes R_N^{-N}\right)\sum_{t=1}^{N}vec\left(x_tx_t'\right) \\
&= N^{-\alpha}\left(R_N^{-(N-1)}\otimes R_N^{-(N-1)}\right)vec\left(x_Nx_N'\right)+o_p(1) \\
&= vec\left\{\left(N^{-\alpha/2}R_N^{-(N-1)}x_N\right)\left(N^{-\alpha/2}R_N^{-(N-1)}x_N\right)'\right\}+o_p(1) \\
&\Rightarrow vec\left(\left(D+\widetilde{U}_x\right)\left(D+\widetilde{U}_x\right)'\right),
\end{aligned}
$$

since

$$
\begin{aligned}
& N^{-\alpha/2}R_N^{-N}x_N \\
&= N^{-\alpha/2}R_N^{-N}\left((I-R_N)^{-1}\mu+R_N^N\left(x_0-(I-R_N)^{-1}\mu\right)+\sum_{j=1}^{N}R_N^{N-j}u_{xj}.\right)
\end{aligned}
$$

$$= \quad R_N^{-N} C^{-1} N^{\alpha/2}\mu + N^{-\alpha/2}x_0 + N^{\alpha/2}C^{-1}\mu + N^{-\alpha/2}\sum_{j=1}^{N} R_N^{-j} u_{xj}$$

$$\Rightarrow \quad D + \left(\int_0^\infty e^{-pC}\Omega_{xx}e^{-pC}dp\right)^{1/2} U_x = D + \widetilde{U}_x.$$

Hence, we have

$$\begin{aligned} & N^{-2\alpha}\left(R_N^{-N}\otimes R_N^{-N}\right)\sum_{t=1}^{N} vec\left(x_t x_t'\right) \\ = \quad & \left(C\otimes I_K + I_K\otimes C\right)^{-1} vec\left(\left(D+\widetilde{U}_x\right)\left(D+\widetilde{U}_x\right)'\right) + o_p(1) \\ = \quad & vec\left(\int_0^\infty e^{-pC}\left(D+\widetilde{U}_x\right)\left(D+\widetilde{U}_x\right)' e^{-pC}dp\right) + o_p(1). \end{aligned} \tag{8.16}$$

(iii) Combining results from (8.14) and (8.16), we obtain

$$\begin{aligned} & vec\left\{N^{\alpha}\left(\widehat{A}-A\right)R_N^N\right\} \\ = \quad & vec\left\{N^{\alpha}\left(\sum_{t=1}^{N} u_{0t}x_t^{'}\right)\left(\sum_{t=1}^{N} x_t x_t'\right)^{-1} R_N^N\right\} \\ = \quad & \left[\left\{N^{-2\alpha}\sum_{t=1}^{N} R_N^{-N}x_t x_t' R_N^{-N}\right\}^{-1}\otimes I_m\right] vec\left\{N^{-\alpha}\sum_{t=1}^{N} u_{0t}x_t^{'}R_N^{-N}\right\} \\ \Rightarrow \quad & \left[\left(\int_0^\infty e^{-pC}\left(D+\widetilde{U}_x\right)\left(D+\widetilde{U}_x\right)' e^{-pC}\right)^{-1}\otimes I_m\right]\left(\int_0^\infty e^{-pC}\left(D+\widetilde{U}_x\right)\left(D+\widetilde{U}_x\right)' e^{-pC}dp\otimes\Omega_{00}\right)^{1/2} W_0 \\ = \quad & \left[\left(\int_0^\infty e^{-pC}\left(D+\widetilde{U}_x\right)\left(D+\widetilde{U}_x\right)' e^{-pC}\right)^{-1/2}\otimes\Omega_{00}^{1/2}\right] W_0, \end{aligned}$$

giving the stated result. ■

Proof of Theorem 3.2. Given the following limit result obtained in the proof of Theorem 3.1

$$N^{-\alpha/2}R_N^{-N}x_N \Rightarrow D + \left(\int_0^\infty e^{-pC}\Omega_{xx}e^{-pC}dp\right)^{1/2} U_x = D + \widetilde{U}_x,$$

we have

$$\begin{aligned} H_{\perp N}H_{\perp N}' \quad = \quad & I_K - \frac{x_N x_N'}{x_N' x_N} \Rightarrow I_K - \frac{\left(D+\widetilde{U}_x\right)\left(D+\widetilde{U}_x\right)'}{\left(D+\widetilde{U}_x\right)'\left(D+\widetilde{U}_x\right)} \\ = \quad & I_K - X_c X_c' := H_\perp H_\perp', \end{aligned}$$

where $X_c := \left(D+\widetilde{U}_x\right)/\left\{\left(D+\widetilde{U}_x\right)'\left(D+\widetilde{U}_x\right)\right\}^{1/2}$ and $D+\widetilde{U}_x$ is the same as the limit given in Theorem 3.1 but with $C=cI_K$. The subvector z_{2t} can be written as

$$z_{2t}=-H_{\perp N}'\mu\sum_{j=1}^{N-t}\rho_N^{-j}-H_{\perp N}'\sum_{j=1}^{N-t}\rho_N^{-j}u_{xt+j}, \tag{8.17}$$

by the reverse autoregression

$$z_{2t}=-\rho_N^{-1}H_{\perp N}'\mu+\rho_N^{-1}z_{2t+1}-\rho_N^{-1}H_{\perp N}'u_{xt+1}.$$

Using the following expression for the scaled error in the LS estimator of A

$$N^{(1+\alpha)/2}\left(\widehat{A}_N-A\right)=\left(N^{-(1+\alpha)/2}\sum_{t=1}^{N}u_{0t}z_t'\right)\left(\sum_{t=1}^{N}z_tz_t'\right)^{-1}H_N',$$

we can write the expression in component form as

$$\begin{aligned}&\left(\frac{1}{N^{(1+\alpha)/2}}\sum_{t=1}^{N}u_{0t}z_t'\right)\begin{pmatrix}O_p\left(N^{1-\alpha}\rho_N^{-2N}\right) & O_p\left(\rho_N^{-N}\right)\\ O_p\left(\rho_N^{-N}\right) & \left(N^{-(1+\alpha)}\sum_{t=1}^{N}z_{2t}z_{2t}'\right)^{-1}+o_p(1)\end{pmatrix}H_N'\\ =\;&\left(\frac{1}{N^{(1+\alpha)/2}}\sum_{t=1}^{N}u_{0t}z_{2t}'\right)\left(\frac{1}{N^{1+\alpha}}\sum_{t=1}^{N}z_{2t}z_{2t}'\right)^{-1}H_{\perp N}'+o_p(1).\end{aligned}$$

(i) Letting $Z_1=[z_{11},z_{12},\dots z_{1N}]'$ and $Z_2=[z_{11},z_{12},\dots z_{1N}]'$, we have

$$\left(\frac{1}{N^{1+\alpha}}\sum_{t=1}^{N}z_tz_t'\right)^{-1}=\begin{pmatrix}\left(\frac{Z_1'Z_1}{N^{1+\alpha}}\right)^{-1}+\Pi_{1N}\left(\frac{Z_2'Q_1Z_2}{N^{1+\alpha}}\right)^{-1}\Pi_{1N}' & -\Pi_{1N}\left(\frac{Z_2'Q_1Z_2}{N^{1+\alpha}}\right)^{-1}\\ -\left(\frac{Z_2'Q_1Z_2}{N^{1+\alpha}}\right)^{-1}\Pi_{1N}' & \left(\frac{Z_2'Q_1Z_2}{N^{1+\alpha}}\right)^{-1}\end{pmatrix},$$

with $Q_1=I_N-Z_1\left(Z_1'Z_1\right)^{-1}Z_1'$ and $\Pi_{1N}=\left(Z_1'Z_1\right)^{-1}Z_1'Z_2$. We show in the following that

$$\left(\frac{1}{N^{1+\alpha}}\sum_{t=1}^{N}z_tz_t'\right)^{-1}=\begin{pmatrix}O_p\left(N^{1-\alpha}\rho_N^{-2N}\right) & O_p\left(\rho_N^{-N}\right)\\ O_p\left(\rho_N^{-N}\right) & \left(\frac{1}{N^{1+\alpha}}\sum_{t=1}^{N}z_{2t}z_{2t}'\right)^{-1}+o_p(1)\end{pmatrix}. \tag{8.18}$$

First, $Z_1'Z_1=O_p\left(\rho_N^{2N}N^{2\alpha}\right)$ since

$$\left\|\frac{\rho_N^{-2N}}{N^{2\alpha}}Z_1'Z_1\right\|=\frac{\rho_N^{-2N}}{N^{2\alpha}}\left\|\sum_{t=1}^{N}z_{1t}z_{1t}'\right\|=\frac{\rho_N^{-2N}}{N^{2\alpha}}\left\|H_{cN}\right\|\left\|H_{cN}'\right\|\left\|\sum_{t=1}^{N}x_tx_t'\right\|=O_p(1), \tag{8.19}$$

and $\sum_{t=1}^{N}x_tx_t'=O_p\left(\rho_N^{2N}N^{2\alpha}\right)$.

Second, we show $Z_1' Z_2 = O_p\left(\rho_N^N N^{2\alpha}\right)$. Using (8.17) and

$$z_{1t} = H_{cN}' \left\{ \frac{N^\alpha}{-c}\mu + \rho_N^t \left(x_0 + \frac{N^\alpha}{c}\mu \right) + \sum_{j=1}^{t} \rho_N^{t-j} u_{xj} \right\},$$

we have the following representation for $Z_1' Z_2$:

$$\begin{aligned} \sum_{t=1}^N z_{1t} z_{2t}' &= \sum_{t=1}^N H_{cN}' \left\{ \begin{array}{c} \left(\frac{N^\alpha}{-c}\mu + \rho_N^t \left(x_0 + \frac{N^\alpha}{c}\mu \right) + \sum_{j=1}^t \rho_N^{t-j} u_{xj} \right) \\ \left(-\mu' \frac{1-\rho_N^{t-N}}{\rho_N - 1} - \sum_{j=t+1}^N \rho_N^{-(j-t)} u_{xj}' \right) \end{array} \right\} H_{\perp N} \\ &= \sum_{t=1}^N H_{cN}' \left\{ \begin{array}{c} \frac{N^\alpha}{c} \frac{1-\rho_N^{t-N}}{\rho_N - 1} \mu\mu' - \frac{\rho_N^t - \rho_N^{2t-N}}{\rho_N - 1} \left(x_0 + \frac{N^\alpha}{c}\mu \right) \mu' \\ - \frac{1-\rho_N^{t-N}}{\rho_N - 1} \sum_{j=1}^t \rho_N^{t-j} u_{xj} \mu' + \frac{N^\alpha}{c}\mu \sum_{j=t+1}^N \rho_N^{-(j-t)} u_{xj}' \\ -\rho_N^t \left(x_0 + \frac{N^\alpha}{c}\mu \right) \sum_{j=t+1}^N \rho_N^{-(j-t)} u_{xj}' - \sum_{j=1}^t \rho_N^{t-j} u_{xj} \sum_{j=t+1}^N \rho_N^{-(j-t)} u_{xj}' \end{array} \right\} H_{\perp N}. \end{aligned}$$

Hence,

$$\begin{aligned} & \left\| \frac{\rho_N^{-N}}{N^{2\alpha}} \sum_{t=1}^N z_{1t} z_{2t}' \right\| \leqslant \frac{\rho_N^{-N}}{N^{2\alpha}} \left\| H_{cN}' \right\| \left\| H_{\perp N} \right\| \\ & \left\{ \begin{array}{c} \left(\frac{N^\alpha}{c} \right)^2 \left\| \sum_{t=1}^N \left(1 - \rho_N^{t-N} \right) \mu\mu' \right\| + \left\| \sum_{t=1}^N \frac{\rho_N^t - \rho_N^{2t-N}}{\rho_N - 1} \left(x_0 + \frac{N^\alpha}{c}\mu \right) \mu' \right\| \\ + \left\| \sum_{t=1}^N \frac{1-\rho_N^{t-N}}{\rho_N - 1} \sum_{j=1}^t \rho_N^{t-j} u_{xj} \mu' \right\| + \left\| \sum_{t=1}^N \frac{N^\alpha}{c}\mu \sum_{j=t+1}^N \rho_N^{-(j-t)} u_{xj}' \right\| \\ + \left\| \sum_{t=1}^N \rho_N^t \left(x_0 + \frac{N^\alpha}{c}\mu \right) \sum_{j=t+1}^N \rho_N^{-(j-t)} u_{xj}' \right\| + \left\| \sum_{t=1}^N \sum_{j=1}^t \rho_N^{t-j} u_{xj} \sum_{j=t+1}^N \rho_N^{-(j-t)} u_{xj}' \right\| \end{array} \right\} \\ = \ & O_p(1), \end{aligned}$$

since

$$\begin{aligned} & \frac{\rho_N^{-N}}{N^{2\alpha}} \left(\frac{N^\alpha}{c} \right)^2 \left\| \sum_{t=1}^N \left(1 - \rho_N^{t-N} \right) \mu\mu' \right\| \\ = \ & \frac{\rho_N^{-N}}{c^2} \left(N - \frac{\rho_N^{-(N-1)} - \rho_N}{1 - \rho_N} \right) \left\| \mu\mu' \right\| \\ = \ & \frac{\rho_N^{-N}}{c^2} \left(N - \frac{\rho_N^{-(N-1)} - \rho_N}{1 - \rho_N} \right) N^{-\alpha} \left\| \mu^* \mu^{*\prime} \right\| + o_p(1) = o_p(1), \end{aligned}$$

$$\begin{aligned} & \frac{\rho_N^{-N}}{N^{2\alpha}} \left\| \sum_{t=1}^N \frac{\rho_N^t - \rho_N^{2t-N}}{\rho_N - 1} \left(x_0 + \frac{N^\alpha}{c}\mu \right) \mu' \right\| \\ = \ & \frac{\rho_N^{-N}}{N^{2\alpha}(\rho_N - 1)} \left(\frac{\rho_N \left(1 - \rho_N^N \right)}{1 - \rho_N} - \frac{\rho_N^{-N} \rho_N^2 \left(1 - \rho_N^{2N} \right)}{1 - \rho_N^2} \right) \left\| \left(x_0 - \frac{N^\alpha}{c}\mu \right) \mu' \right\| \end{aligned}$$

$$= \quad \frac{2\rho_N + \rho_N^2}{2c^2} \left\| x^* \mu^{*\prime} - \frac{1}{c} \mu^* \mu^{*\prime} \right\| = O_p(1),$$

and

$$\frac{\rho_N^{-N}}{N^{2\alpha}} \left\| \sum_{t=1}^{N} \frac{1 - \rho_N^{t-N}}{\rho_N - 1} \sum_{j=1}^{t} \rho_N^{t-j} u_{xj} \mu' \right\| = o_p(1),$$

as

$$\frac{\rho_N^{-N}}{N^{2\alpha}} \mathbb{E} \left\| \sum_{t=1}^{N} \frac{1 - \rho_N^{t-N}}{\rho_N - 1} \sum_{j=1}^{t} \rho_N^{t-j} u_{xj} \mu' \right\| = 0.$$

Similarly,

$$\left\| \sum_{t=1}^{N} \frac{N^\alpha}{c} \mu \sum_{j=t+1}^{N} \rho_N^{-(j-t)} u'_{xj} \right\| = o_p(1) \text{ and } \left\| \sum_{t=1}^{N} \rho_N^t \left(x_0 - \frac{N^\alpha}{c} \mu \right) \sum_{j=t+1}^{N} \rho_N^{-(j-t)} u'_{xj} \right\| = o_p(1).$$

Further, we have from MP (2009) that

$$\frac{\rho_N^{-N}}{N^{2\alpha}} \mathbb{E} \left\| \sum_{t=1}^{N} \sum_{j=1}^{t} \rho_N^{t-j} u_{xj} \sum_{j=t+1}^{N} \rho_N^{-(j-t)} u'_{xj} \right\| \leqslant B, \text{ where } B \text{ is some constant.}$$

In summary, combining the above results, we have

$$\begin{aligned} Z_1' Z_2 &= O_p\left(\rho_N^N N^{2\alpha}\right), Z_1' Z_1 = O_p\left(\rho_N^{2N} N^{2\alpha}\right), \Pi_{1N} = \left(Z_1' Z_2\right)^{-1} \left(Z_1' Z_1\right) = O_p\left(\rho_N^{-N}\right), \\ \frac{Z_2' Q_1 Z_2}{N^{1+\alpha}} &= \frac{Z_2' Z_2}{N^{1+\alpha}} + O_p\left(N^{\alpha-1}\right). \end{aligned}$$

Next, we derive the limit distribution for $\frac{Z_2' Z_2}{N^{1+\alpha}}$. Considering that $z_{2N} = 0$ by construction, we have

$$Z_2' Z_2 = \sum_{t=1}^{N-1} \left(-\rho_N^{-1} H'_{\perp N} \mu + \rho_N^{-1} z_{2t+1} - \rho_N^{-1} H'_{\perp N} u_{xt+1} \right) \left(-\rho_N^{-1} H'_{\perp N} \mu + \rho_N^{-1} z_{2t+1} - \rho_N^{-1} H'_{\perp N} u_{xt+1} \right)',$$

which leads to

$$\begin{aligned} &\frac{\rho_N^2 - 1}{N} \sum_{t=1}^{N-1} z_{2t} z'_{2t} \\ = \; &\frac{z_{2N} z'_{2N}}{N} - \frac{z_{0N} z'_{0N}}{N} + \frac{1}{N} H_{\perp N} \sum_{t=1}^{N-1} \mu \mu' H'_{\perp N} - \frac{1}{N} H'_{\perp N} \sum_{t=1}^{N-1} \mu z'_{2t+1} \\ &+ \frac{1}{N} H'_{\perp N} \sum_{t=1}^{N-1} \mu u'_{xt+1} H_{\perp N} - \frac{1}{N} \sum_{t=1}^{N-1} z_{2t+1} \mu' H_{\perp N} - \frac{1}{N} \sum_{t=1}^{N-1} z_{2t+1} u'_{xt+1} H_{\perp N} \end{aligned}$$

$$
\begin{aligned}
& +\frac{1}{N} H_{\perp N} \sum_{t=1}^{N-1} u_{xt+1} \mu' H_{\perp N}' - \frac{1}{N} H_{\perp N}' \sum_{t=1}^{N-1} u_{xt+1} z_{2t+1}' + \frac{1}{N} H_{\perp N}' \sum_{t=1}^{N-1} u_{xt+1} u_{xt+1}' H_{\perp N} \\
= & \frac{1}{N} H_{\perp N}' \sum_{t=1}^{N-1} u_{xt+1} u_{xt+1}' H_{\perp N} - \frac{1}{N} \sum_{t=1}^{N-1} z_{2t+1} \mu' H_{\perp N} - \frac{1}{N} H_{\perp N}' \sum_{t=1}^{N-1} \mu z_{2t+1}' + o_p(1) \\
= & \frac{1}{N} H_{\perp N}' \Omega_{xx} H_{\perp N} + \frac{2}{c} H_{\perp N}' \mu^* \mu^{*\prime} H_{\perp N} + o_p(1),
\end{aligned}
$$

since the following hold:

(1)

$$
\frac{z_{2N} z_{2N}'}{N} = O_p\left(N^{-1}\right) = o_p(1),
$$

(2)

$$
\frac{z_{0N} z_{0N}'}{N} = O_p\left(N^{\alpha-1}\right) = o_p(1),
$$

(3)

$$
\begin{aligned}
& \frac{1}{N} H_{\perp N}' \sum_{t=1}^{N-1} u_{xt+1} z_{2t+1}' \\
= & \frac{1}{N} H_{\perp N}' \sum_{t=1}^{N-1} u_{xt+1}' \left\{ -\mu \sum_{j=1}^{N-1-t} \rho_N^{-j} - \sum_{j=1}^{N-1-t} \rho_N^{-j} u_{xt+1+j} \right\}' H_{\perp N} \\
= & o_p(1),
\end{aligned}
$$

(4)

$$
\begin{aligned}
\frac{1}{N} H_{\perp N}' \sum_{t=1}^{N-1} \mu z_{2t+1}' &= \frac{1}{N} H_{\perp N}' \mu \sum_{t=1}^{N-1} \left\{ -\mu \sum_{j=1}^{N-t-1} \rho_N^{-j} - \sum_{j=1}^{N-t-1} \rho_N^{-j} u_{xt+1+j} \right\}' H_{\perp N} \\
&= -\frac{1}{N} H_{\perp N}' \sum_{t=1}^{N-1} \sum_{j=1}^{N-1-t} \rho_N^{-j} \mu \mu' H_{\perp N} + o_p(1) \\
&= -\frac{1}{N^{1+\alpha}} H_{\perp N}' \sum_{t=1}^{N-1} \frac{\rho_N^{-1} \left(1 - \rho_N^{-(N-t-1)}\right)}{1 - \rho_N^{-1}} \mu^* \mu^{*\prime} H_{\perp N} + o_p(1) \\
&= -\frac{1}{c} H_{\perp N}' \mu^* \mu^{*\prime} H_{\perp N} + o_p(1),
\end{aligned}
$$

(5)

$$
\frac{1}{N} H_{\perp N}' \sum_{t=1}^{N-1} u_{xt+1} u_{xt+1}' H_{\perp N} \Longrightarrow \frac{1}{N} H_{\perp N}' \Omega_{xx} H_{\perp N},
$$

and

(6)

$$\frac{1}{N}H_{\perp N}\sum_{t=1}^{N-1}\mu\mu' H_{\perp N}' = o_p(1).$$

Hence, by the same argument as in Lemma 4.3 of MP, we have

$$\begin{aligned}
&\frac{\rho_N^2-1}{N}\sum_{t=1}^{N-1} z_{2t}z_{2t}' \\
= &\ H_{\perp N}'\Omega_{xx}H_{\perp N} + \frac{2}{c}H_{\perp N}'\mu^*\mu^{*\prime}H_{\perp N} + o_p(1) \\
\Rightarrow &\ H_\perp'\Omega_{xx}H_\perp + \frac{2}{c}H_\perp'\mu^*\mu^{*\prime}H_\perp,
\end{aligned}$$

where $H_\perp$ is a $K\times(K-1)$ matrix (an orthogonal complement of the vector X_c) satisfying

$$H_\perp H_\perp' = I_K - \frac{\left(D+\widetilde{U}_x\right)\left(D+\widetilde{U}_x\right)'}{\left(D+\widetilde{U}_x\right)'\left(D+\widetilde{U}_x\right)}.$$

Therefore,

$$\frac{1}{N^{1+\alpha}}\sum_{t=1}^{N} z_{2t}z_{2t}' \Rightarrow H_\perp'\left(\frac{1}{2c}\Omega_{xx}+\frac{\mu^*}{c}\frac{\mu^{*\prime}}{c}\right)H_\perp.$$

(ii) Normalizing by N^{-1}, the component $N^{-1}\sum_{t=1}^N u_{0t}z_{2t}'$ is asymptotically negligible, since

$$\frac{1}{N}\sum_{t=1}^{N} u_{0t}z_{2t}' = \frac{1}{N}\sum_{t=1}^{N}u_{0t}\left(-H_{\perp N}'\mu\sum_{j=1}^{N-t}\rho_N^{-j} - H_{\perp N}'\sum_{j=1}^{N-t}\rho_N^{-j}u_{xt+j}\right)' = o_p(1).$$

Hence, when normalized by $\frac{1}{N^{(1+\alpha)/2}}$, we have

$$\begin{aligned}
&\frac{1}{N^{(1+\alpha)/2}}vec\left(\sum_{t=1}^{N}u_{0t}z_{2t}'\right) \qquad (8.20)\\
= &\ -\frac{1}{N^{(1+\alpha)/2}}vec\left(\sum_{t=1}^{N}u_{0t}\mu'\sum_{j=1}^{N-t}\rho_N^{-j}H_{\perp N}\right) - \frac{1}{N^{(1+\alpha)/2}}vec\left(\sum_{t=1}^{N}u_{0t}\sum_{j=1}^{N-t}\rho_N^{-j}u_{xt+j}'H_{\perp N}\right).
\end{aligned}$$

For the first item on the right side of (8.20), we have

$$\begin{aligned}
&-\frac{1}{N^{(1+\alpha)/2}}vec\left(\sum_{t=1}^{N}u_{0t}\mu'\sum_{j=1}^{N-t}\rho_N^{-j}H_{\perp N}\right) \\
= &\ -\frac{1}{N^{(1+\alpha)/2}}vec\left(\sum_{t=1}^{N}u_{0t}\mu'\sum_{j=t+1}^{N}\rho_N^{-(j-t)}H_{\perp N}\right)
\end{aligned}$$

$$
\begin{aligned}
&= -\frac{1}{N^{(1+\alpha)/2}} vec\left(\sum_{t=1}^{N} u_{0t}\left(1-\rho_N^{-(N-t)}\right)\mu' H_{\perp N}\frac{N^{\alpha}}{c}\right) \\
&= -\frac{1}{c}\frac{1}{\sqrt{N}} vec\left(\sum_{t=1}^{N} u_{0t}\mu^{*\prime} H_{\perp N}\right) + \frac{1}{c}\frac{1}{\sqrt{N}} vec\left(\sum_{t=1}^{N} \rho_N^{-(N-t)} u_{0t}\mu^{*} H_{\perp N}\right) \\
&= -H_{\perp N}'\frac{\mu^*}{c}\otimes\left(\frac{1}{\sqrt{N}} vec\left(\sum_{t=1}^{N} u_{0t}\right)\right) \\
&\Rightarrow H_{\perp}'\frac{\mu^*}{c}\otimes N\left(0,\Omega_{00}\right) \overset{d}{=} N\left(0, H_{\perp}'\frac{\mu^*}{c}\frac{\mu^{*\prime}}{c}H_{\perp}\otimes\Omega_{00}\right).
\end{aligned} \tag{8.21}
$$

The second item on the right handside of (8.20) is asymptotically negligible, since

$$
-\frac{1}{N^{(1+\alpha)/2}}\sum_{t=1}^{N} u_{0t}\sum_{j=1}^{N-t}\rho_N^{-j}u_{xt+j}'H_{\perp N} = -\frac{1}{N^{(1+\alpha)/2}}\sum_{t=m_N+1}^{N} u_{0t}\sum_{j=t+1}^{N}\rho_N^{-(j-t)}u_{xt+j}'H_{\perp N} + o_p(1), \tag{8.22}
$$

and

$$
-\frac{1}{N^{(1+\alpha)/2}} vec\left(\sum_{t=m_N+1}^{N} u_{0t}\sum_{j=t+1}^{N}\rho_N^{-(j-t)}u_{xt+j}'H_{\perp N}\right) \Rightarrow N\left(0, \frac{1}{2c}H_{\perp N}'\Omega_{xx}H_{\perp N}'\otimes\Omega_{00}\right). \tag{8.23}
$$

(1) For equation (8.22), we have

$$
\begin{aligned}
&\mathbb{E}\left\|\frac{1}{N^{(1+\alpha)/2}}\sum_{t=1}^{m_N} u_{0t}\sum_{j=t+1}^{N}\rho_N^{-(j-t)}u_{xt+j}'\right\|^2 \\
&= \frac{\mathbb{E}\left\|u_{0t}\right\|^2}{N^{1+\alpha}}\mathbb{E}\left\|\sum_{t=1}^{m_N}\sum_{j=t+1}^{N}\rho_N^{-(j-t)}u_{xt+j}'\right\| \\
&= \frac{\mathbb{E}\left\|u_{01}\right\|^2\mathbb{E}\left\|u_{x1}\right\|^2}{N^{1+\alpha}}\sum_{t=1}^{m_N}\sum_{j=t+1}^{N}\rho_N^{-2(j-t)} = O\left(N^{\alpha-1}\right).
\end{aligned}
$$

(2) The result (8.23) follows from Lemma 4.4 of MP.

Combining (8.21) and (8.23), the limit distribution of $\frac{1}{N^{(1+\alpha)/2}}\sum_{t=1}^{N} u_{0t}z_{2t}'$ is

$$
\begin{aligned}
&\frac{1}{N^{(1+\alpha)/2}} vec\left(\sum_{t=1}^{N} u_{0t}z_{2t}'\right) \\
\Rightarrow\ &\left(H_{\perp}'\frac{\mu^*}{c}\frac{\mu^{*\prime}}{c}H_{\perp}\otimes\Omega_{00}\right)^{1/2}\times N\left(0, I_{m\times(K-1)}\right) + \left(\frac{1}{2c}H_{\perp}'\Omega_{xx}H_{\perp}\otimes\Omega_{00}\right)^{1/2}\times N\left(0, I_{m\times(K-1)}\right) \\
=\ &\left(H_{\perp}'\left(\frac{\mu^*}{c}\frac{\mu^{*\prime}}{c}+\frac{1}{2c}\Omega_{xx}\right)H_{\perp}\otimes\Omega_{00}\right)^{1/2}\times N\left(0, I_{m\times(K-1)}\right),
\end{aligned}
$$

since we have the following independent structure asymptotically

$$\lim_{N\to\infty}\mathbb{E}\left\{\left(\frac{1}{\sqrt{N}}\sum_{t=1}^{N}u_{0t}\mu^{*\prime}\right)\left(\frac{1}{N^{(1+\alpha)/2}}\sum_{t=m_N+1}^{N}u_{0t}\sum_{j=t+1}^{N}\rho_N^{-(j-t)}u_{xt+j}^{\prime}\right)^{\prime}\right\}=0.$$

(iii) Using the results from (i) and (ii), we obtain

$$\begin{aligned}
& N^{(1+\alpha)/2}vec\left(\widehat{A}-A\right) \\
= \;& vec\left\{N^{(1+\alpha)/2}\left(\sum_{t=1}^{N}u_{0t}z_{2t}^{'}\right)\left(\sum_{t=1}^{N}z_{2t}z_{2t}^{'}\right)^{-1}H_{\perp N}^{\prime}\right\}+o_p(1) \\
= \;& \left[\left\{H_{\perp N}\left(\frac{1}{N^{1+\alpha}}\sum_{t=1}^{N}z_{2t}z_{2t}^{'}\right)^{-1}\right\}\otimes I_m\right]vec\left\{\frac{1}{N^{(1+\alpha)/2}}\sum_{t=1}^{N}u_{0t}z_{2t}^{'}\right\}+o_p(1) \\
\Rightarrow \;& \left[H_{\perp}\left\{H_{\perp}^{\prime}\left(\frac{\mu^*}{c}\frac{\mu^{*\prime}}{c}+\frac{1}{2c}\Omega_{xx}\right)H_{\perp}\right\}^{-1}\otimes I_m\right]\left\{H_{\perp}^{\prime}\left(\frac{\mu^*}{c}\frac{\mu^{*\prime}}{c}+\frac{1}{2c}\Omega_{xx}\right)H_{\perp}\otimes\Omega_{00}\right\}^{1/2}\times N\left(0,I_{m\times(K-1)}\right) \\
= \;& \left[H_{\perp}\left\{H_{\perp}^{\prime}\left(\frac{\mu^*}{c}\frac{\mu^{*\prime}}{c}+\frac{1}{2c}\Omega_{xx}\right)H_{\perp}\right\}^{-1/2}\otimes\Omega_{00}^{1/2}\right]\times N\left(0,I_{m\times K}\right) \\
\overset{d}{=} \;& MN\left(0,H_{\perp}\left\{H_{\perp}^{\prime}\left(\frac{\mu^*}{c}\frac{\mu^{*\prime}}{c}+\frac{1}{2c}\Omega_{xx}\right)H_{\perp}\right\}^{-1}H_{\perp}^{\prime}\otimes\Omega_{00}\right),
\end{aligned}$$

giving the stated result. ■