Inference in continuous systems with mildly explosive regressors

Ye Chen a, Peter C.B. Phillips b,c,d,e,*, Jun Yu d

a Capital University of Economics and Business, China
b Yale University, United States
c University of Auckland, New Zealand
d Singapore Management University, Singapore
e University of Southampton, United Kingdom

A R T I C L E I N F O

Article history:
Available online xxxx

JEL classification:
C12
C13
C58

Keywords:
Cointegrated system
Explosive process
Moderate deviations from unity
Double asymptotics
Real estate market

A B S T R A C T

New limit theory is developed for co-moving systems with explosive processes, connecting continuous and discrete time formulations. 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.

© 2017 Elsevier B.V. All rights reserved.

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 financial markets. For example, Phillips et al. (2011) and Phillips et al. (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, PM hereafter) 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 Gouriéroux has recently explored new ways of generating explosive bubbles via non-causal forward-looking processes (Gouriéroux and Zakoian, 2017).

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 PM. 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 (2016, 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 MP. 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 dependencies are sufficient to materially change conclusions in empirical work.

This paper extends work by Philips (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 a way that relates to a continuous time system whereas MP use 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 \( \kappa \) 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 of Theorem 2.1 and Corollary 2.2 are given in Appendix A. Proofs of Theorem 3.1 and 3.2 are provided in an online technical supplement.

2. Continuous systems with a mildly explosive regressor

We start our investigation with the following scalar continuous time model in two variates \( y(t) \) and \( x(t) \). Here \( x(t) \) follows an Ornstein–Uhlenbeck process and the stochastic process \( y(t) \) co-moves with \( x(t) \) as

\[
y(t) = \beta x(t) + u_0(t),
\]

\[
dx(t) = \kappa (\mu - x(t)) dt + dB_t(t),
\]

\[
x(0) = x_0 = O_p(1), \quad \kappa < 0,
\]

where \( u_0(t) \) is Gaussian pure noise – a generalized stochastic process in continuous time (see Hannan, 1970 and Phillips, 1991). Formally, we may write \( u_0(t) = DB_0(t) \), where \( B_0(t) = \sigma_{00} W_0(t) \), \( W_0 \) is standard Brownian motion, and \( D = d/dt \) is the mean square differential operator. While \( u_0(t) \) is unrealizable as a covariance stationary stochastic process in continuous time, the corresponding discrete time process is realizable and takes the form of a pure noise process of independent identically distributed (iid) \( N(0, \sigma_{00}^2) \) errors. This formulation is convenient in that it corresponds to the discrete time system (2.3) that follows and has been extensively used in modeling microstructure noise effects in the measurement of efficient financial asset prices, as discussed below.\(^{2}\) We write \( B_0(t) = \sigma_{00} W_k(t) \) where \( W_k \) is a standard Brownian motion that may be correlated with \( W_0 \).

The parameter of central interest for inference is the coefficient \( \beta \) which captures the co-movement between \( y(t) \) and \( x(t) \). The driver process \( x(t) \) follows an Ornstein–Uhlenbeck equation with persistence parameter \( \kappa \). 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 \( \kappa \) 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, \( \{t_h\}_{h=1}^N \), 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 \( h \rightarrow 0 \) and \( T \rightarrow \infty \). The exact discrete time representation of (2.1)–(2.2) is (Phillips, 1972)

\[
y_{th} = \beta x_{th} + u_{0,th},
\]

\[
x_{th} = a_0(\kappa) x_{(t-1)h} + g_h + u_{th}, \quad x_{0h} = x_0 = O_p(1),
\]

where

\[
a_0(\kappa) = \exp \left( -\kappa h \right), \quad g_h = \mu \left( 1 - e^{-\kappa h} \right),
\]

\(^{1}\) Alternate specifications are possible. For example, as discussed in Remark 12 we might set \( u_0(t) = u_0 \), which corresponds to the limit form of a discrete time cointegrated system. A more general model specification might allow for a realizable covariance stationary error in (2.1). Extending the double limit theory developed in the present paper to cover the case of dependent errors is beyond the scope of the present work and will be explored later.

\(^{2}\) In such models, the standard formulation of an efficient price subject to unobserved microstructure noise leads to a continuous system involving generalized stochastic processes (see the discussion in Phillips and Yu, 2006).
\[ u_{x,th} = \sigma_u \int_{e^{-(t-h)}}^{t} e^{-c(t-h)} d\omega_x(s) = N \left( 0, \frac{\sigma_u^2}{2} \left( 1 - e^{-2c_h} \right) \right), \]

\[ u_{0,th} = N \left( 0, \sigma_\epsilon^2 \right). \]

Unlike the variance of the discrete time error \( u_{x,th} \) for the process \( x_{th} \), the variance of \( u_{0,th} \) does not depend on the sampling interval \( h \) and, as implied in the above discussion, \( u_{0,th} \) therefore corresponds to an \( iid \) Gaussian sequence in discrete time that is associated with the generalized stochastic process \( u_0(t) \) in continuous time. The autoregressive parameter \( a(\cdot) \) depends directly on the sampling frequency \( h \) and, indirectly, \( h \) and \( x_0(\cdot) \) are both related to the sample size \( N \). When \( T \) is fixed, \( h = T/N = O(N^{-1}) \), and when \( T \to \infty \), \( h = 0 \) (\( T/N \to 0 \). Gaussianity follows from the Brownian motion driven processes in (2.1)-(2.2). The standard error of \( u_{x,th} \) is \( \lambda_h \sim \sigma_{\epsilon x} \sqrt{h} \to 0 \), concordant with the sample path continuity of \( x(t) \).

Re-standardizing Eq. (2.4) by \( \lambda_h \) we have

\[ y_{th} = \beta x_{th} + u_{0,th}, \quad (2.5) \]

\[ \tilde{x}_{th} = a_0(\cdot) \tilde{x}_{t-1} + u_{t,x,th}, \quad \tilde{x}_{th} = x_{th}/\lambda_h, \]

\[ \tilde{u}_{x,th} = N \left( 0, 1 \right), \quad (2.6) \]

where \( \tilde{x}_{th} = x_{th}/\lambda_h, \tilde{u}_{th} = u_{th}/\lambda_h \) as \( h \to 0 \). When \( T \to \infty \) and \( h \to 0 \), we have

\[ \frac{1}{Nh} = \frac{1}{T} \to 0, \quad a_k(\cdot) = 1 - k + O(h^2) \to 1. \]

Hence, \( \tilde{x}_{th} \) in (2.6) is a mildly explosive process as in PM. Furthermore, since \( h \sim 0 \), when \( h \to 0 \) we have \( \tilde{x}_{th} = x_{th}/\lambda_h = O_p(h^{-1/2}), (a_0(\cdot) - N = 0(1/T), \tilde{u}_{th} \sim O(\sqrt{h}) \). Thus, in the standardized discrete system (2.5)-(2.6) the order of magnitude of the initial condition is \( \tilde{x}_{th} \sim O_p(h^{-1/2}) \) while in the original system (2.1)-(2.2) it is \( x_{th} \sim O_p(1) \). In addition, the order of magnitude of the drift is \( O(\sqrt{h}) \) in model (2.6) but is \( O_p(1) \) in (2.2).

The continuous time system (2.1)-(2.2) is formulated in a manner similar to models used in the market microstructure literature, where the observed transaction price is often assumed to differ from the efficient price by an \( iid \) noise component—see Zhang et al. (2005), Ait-Sahalia et al. (2005, AMZ hereafter) and Bandi and Russell (2006). In AMZ, the logarithmic efficient price of a security follows

\[ dx(t) = \mu(X(t)) dt + \sigma(X(t)) dW(t), \quad (2.7) \]

where \( W(t) \) is a standard Brownian motion, and the logarithmic price is observed with an \( iid \) error, namely

\[ \tilde{X}(t) = X(t) + U(t), \quad (2.8) \]

where \( U(t) \) is \( iid \) noise with mean zero and a finite variance \( \sigma_u^2 \) that is independent of the sampling interval \( h \), and \( U(t) \) is independent of \( X(t) \). The formulation (2.8) therefore corresponds to (2.3) and is a discrete time version of the continuous process co-movement \( \tilde{X}(t) = X(t) + U(t) \) with a generalized stochastic process error \( U(t) \). Issues of formulating such systems in continuous time have been discussed in this market microstructure noise literature to which readers are referred (e.g. Hansen and Lunde, 2006; Phillips and Yu, 2006).

The continuous time system here also relates to the model in Phillips (1991), who studied the system

\[ y_1(t) = B y_2(t) + u_1(t), \quad D y_2(t) = u_2(t), \]

where \( D = d/dt \) denotes the mean square differential operator and \( u(t) = [u_1(t), u_2(t)]' \) is a stationary time series. In Phillips (1991), both \( y_1(t) \) and \( y_2(t) \) are \( I(1) \) processes and co-movement is a form of continuous time cointegration. In contrast, our focus of attention in the present paper is co-movement among potentially explosive processes.

The standardized discrete system (2.5)-(2.6) is related to MP. MP analyzed the triangular system

\[ y_1(t) = A x_t + \epsilon_{0,t}, \quad (2.9) \]

\[ x_t = \gamma x_{t-1} + \epsilon_{0,t}, \quad x_0 = \theta_0 N^{m/2}, \quad (2.10) \]

where \( \gamma = I_k + \frac{\alpha}{m}, \alpha \in (0, 1), C = diag(c_1, \ldots, c_k) \), and discrete observations \( \{y_t, x_t\} \) are available. In this system, \( A \) is the matrix of cointegrating (or, more specifically in the development below, co-mildly explosive) coefficients; \( \gamma \) represents moderate deviations from a unit root in the sense of PM; \( x_t \) is a moderately integrated time series as \( N^w \to \infty \) when \( N \to \infty \). If \( C > 0 \), \( x_t \) is a mildly explosive time series. The vector \( \{x_0, \epsilon_{0,t}\} \) 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 \( C > 0 \).

The formulation of moderate deviations from a unit root with the setting \( R = 1 + c/N^\alpha, \alpha \in (0, 1) \), can be justified from the double asymptotic \( (h \to 0 \text{ and } T \to \infty) \) point of view in continuous time systems, as argued in BY. This justification is similar to that in Boswijk (2001) who justified the block-local-to-unity concept introduced by Phillips et al. (2001, PMX hereafter) in continuous time systems. In PMX, the block-local-to-unity is formulated as \( (1 + c/m) \) where the full sample consists of \( M \) blocks of \( m \) observations so that \( N = Mm \). In PMX and Boswijk (2001), two kinds of asymptotics were considered, namely, (1) \( M \) is fixed and \( m \to \infty; \quad (2) \quad m \to \infty \text{ and then } M \to \infty. \) In case (1), the autoregressive coefficient parameter \( 1 + \frac{c}{m} \) has the same order of magnitude as the standard local-to-unity model. In case (2), the autoregressive coefficient parameter is further away from unity than that implied by the standard local-to-unity model. An interesting example in case (2) was considered in PMX is \( m = N^\gamma \) and \( M = N^{1-\gamma}, \gamma \in (0, 1) \). In this case, both \( m \) and \( M \) are determined by \( N \).

To formalize the link of the notations (i.e. \( m, M, N \)) in Boswijk (2001) and those in our paper (i.e. \( h, T, N \)) we have \( h = \frac{1}{T}, \quad T = M \). Clearly our double asymptotics \( (T \to \infty \text{ and } h \to 0) \) correspond to case (2) in PMX and Boswijk (2001). However, we do not choose \( N \) as in PMX and Boswijk (2001) because \( N \) is determined after \( h \) and \( T \) are chosen in our setup.

There are some common features in model (2.5)-(2.6) and the MP model (2.9)-(2.10): 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 (2.10) it is formulated as a function of the overall sample size \( N \). A second difference is that, while in (2.10) the initial condition for \( x_0 \) is assumed to be \( \theta_0 N^{m/2} \), in (2.6) it is \( \theta_0 h^{-1/2} \), which translates to \( x_0 = \theta_0 \) in the original continuous time system (2.1)-(2.2). So, the initial condition in (2.6) has the larger order of magnitude \( \theta_0 h^{-1/2} \), 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 (2.10) there is no intercept, and if a constant intercept were present it would typically dominate the asymptotics. By contrast, in (2.6) the
intercept

\[ \hat{\beta}_h = \beta_0/\lambda_h = \mu \left( 1 - e^{-x_h} \right) / \lambda_h = \frac{\mu \left( 1 - e^{-x_h} \right)}{\sigma_{xx} \sqrt{1 - e^{-2x_h}} / 2x} \]

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

To fix ideas, consider the modified MP model

\[ y_t = \lambda_kx_t + w_t, \quad (2.11) \]

\[ x_t = \mu + R_0x_{t-1} + u_{dt}, \quad x_0 = \lambda_0 \left( N^{\alpha/2} \right), \]

\[ \mu = O_p \left( N^{-\alpha/2} \right). \]

Let \( \lambda_0 = x_0(N^{-\alpha/2}) \Rightarrow \lambda^* \) and \( \hat{\mu} = N^{\alpha/2} \mu \Rightarrow \mu^* \). The error \( u_t = [u_{dt}, u_{dt}]^T \) is an iid sequence with mean zero and covariance \( \sigma_0^0 \sigma_0 \sigma_0^0 \). This model extends \((2.9)-(2.10)\) 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 \((2.11)\) for the case of a single scalar regressor \( x_t \).

**Theorem 2.1.** For the discrete time system \((2.11)-(2.12)\) with \( N = 1 + \frac{N^\alpha}{N^\alpha}, \alpha \in (0,1), \) and \( c > 0, \) when \( N \to \infty, \) we have

(i) \[ \left( R_N^N N^{\alpha/2} \right) \sum_{t=1}^N x_t u_{dt} \Rightarrow \int_0^1 \hat{b}_0 u_0 \left( \sigma_{xx} U_x + (2c)^{1/2} D \right), \]

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

where \( \left( U_x, U_{dt} \right) \overset{d}{=} \left( N(0, I), D = X + \frac{\mu}{c} \right) \), and so

\[ R_N^N N^{\alpha/2} \left( \hat{A} - A \right) \Rightarrow 2c \frac{\sigma_0 u_0}{\sigma_{xx} U_x + (2c)^{1/2} D}. \]

**Remark 1.** The limit result in Theorem 2.1 can be extended to the case where the errors are weakly dependent. Let \( u_{dt} \) in the discrete time system \((2.11)-(2.12)\) be a sequence of zero mean and weakly dependent errors, satisfying Assumption LP as defined in MP, so that

\[ N^{-1/2} \sum_{t=1}^N u_{dt} \Rightarrow N(0, \sigma_0^0). \]

where the long-run variance \( \sigma_0^0 \) can be decomposed as

\[ \sigma_0^0 = \sigma_0^2 + 2\lambda \sigma_0, \]

with variance component \( \sigma_0^2 \) and one-sided long-run variance

\[ \lambda \sigma_0 := \sum_{b=1}^\infty E \left( u_{0b} u_{0b-b} \right). \]

Using the limit results from Theorem 4.1 in MP and Theorem 2.1, we obtain

\[ R_N^N N^{\alpha/2} \left( \hat{A} - A \right) \Rightarrow 2c \frac{\sigma_0 u_0}{\sigma_{xx} U_x + (2c)^{1/2} D}. \]

**Remark 2.** If \( \hat{\lambda}_0 = -\hat{\mu} \), then \( D = 0 \) and the limit \((2.13)\) is simply

\[ R_N^N N^{\alpha/2} \left( \hat{A} - A \right) \Rightarrow 2c \frac{\sigma_0 u_0}{\sigma_{xx} U_x}, \]

where \( C \) 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

\[ \frac{R_N^N}{R_N^N - 1} \left( \hat{A} - A \right) \Rightarrow \frac{\sigma_0 u_0}{\sigma_{xx} C}, \]

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

\[ \frac{R_N^N}{R_N^N - 1} \left( \hat{A} - A \right) \Rightarrow \frac{\sigma_0 u_0}{\sigma_{xx} U_x + (2c)^{1/2} D}. \]

**Remark 3.** The limit distribution of \( \hat{\mu} \), the LS estimator of the intercept parameter \( \mu \), follows simply as given in Box 1. This result is useful in testing for \( \mu = 0 \) in the modified MP model.

**Remark 4.** Self-normalized statistics based on \( \hat{A} \) have a much simpler limit theory that is convenient for inference. For instance, the regression residuals \( \hat{u}_{dt} = y_t - \hat{\lambda}_0 x_t \) noting that the residual variance estimate \( s_0^2 = N^{-1} \sum_{t=1}^N \hat{u}_{dt}^2 \Rightarrow \sigma_0^2 \) follows immediately from Theorem 2.1 that the usual t statistic for testing \( H_0 : \lambda = A^0 \) satisfies

\[ t_A = \frac{\hat{A} - A^0}{s_A} = \frac{R_N^N N^{\alpha/2} (\hat{A} - A^0)}{\left( \frac{s_0^2}{\sigma_0^2} \frac{1}{N} \sum_{t=1}^N X_t \right)^{-1/2} \left( \frac{1}{2} \right)^2 \left( \sigma_{xx} U_x + (2c)^{1/2} D \right)^2} \Rightarrow 2c \frac{\sigma_0 u_0}{\sigma_{xx} U_x + (2c)^{1/2} D} \]

\[ \Rightarrow \frac{s_0}{\sigma_{xx} \left( \frac{1}{2} \right)^2 \left( \sigma_{xx} U_x + (2c)^{1/2} D \right)^2} \]

\[ \Rightarrow U_x \overset{d}{=} N(0, 1), \]

and standard methods of inference apply.

**Remark 5.** Let \( \hat{R}_N \) be the LS estimator of \( R_N \) and \( \hat{c} = N^{\alpha} (\hat{R}_N - 1) \). The limit theory for \( \hat{R}_N \) and \( \hat{c} \) follows from the proof of Theorem 1.1 and Remark 3. Defining the regression residuals \( \hat{u}_{dt} = \hat{\lambda}_0 x_t - \hat{\lambda}_0 x_{t-1} - \hat{\lambda}_0 x_{t-2} \) and noting that the residual variance estimate \( s_0^2 = N^{-1} \sum_{t=1}^N \hat{u}_{dt}^2 \Rightarrow \sigma_0^2 \), we have the following result for the t statistic for testing \( H_0 : \lambda = A^0 \)

\[ t_A = \frac{\hat{R}_N - R_0^N}{s_{R_N}} = \frac{R_N^N N^{\alpha/2} (\hat{R}_N - R_0^N)}{\left( \frac{s_0^2}{\sigma_0^2} \frac{1}{N} \sum_{t=1}^N X_t \right)^{-1/2} \left( \frac{1}{2} \right)^2 \left( \sigma_{xx} U_x + (2c)^{1/2} D \right)^2} \Rightarrow 2c \frac{\sigma_0 u_0}{\sigma_{xx} U_x + (2c)^{1/2} D} \]

\[ \Rightarrow \frac{s_0}{\sigma_{xx} \left( \frac{1}{2} \right)^2 \left( \sigma_{xx} U_x + (2c)^{1/2} D \right)^2} \]

\[ \Rightarrow U_x \overset{d}{=} N(0, 1). \]

Similarly, given \( R_N = 1 + \frac{N^\alpha}{N^\alpha}, \) we have \( s_0 = N^{\alpha} s_{R_N} \) and \( \hat{c} - c^0 = (\hat{R}_N - R_0^N) N^{\alpha} \). Hence, if \( \alpha \) is known, the t statistic for testing \( H_0 : c = c^0 \)

\[ t_c = \frac{\hat{c} - c^0}{s_c} \approx (\hat{R}_N - R_0^N) N^{\alpha} \overset{d}{=} U_x \overset{d}{=} N(0, 1). \]

However, if \( \alpha \) is unknown, then both the estimate \( \hat{c} \) and the standard error \( s_0 = N^{\alpha} s_{R_N} \) are unavailable and inference using this limit theory for \( c \) is infeasible. As discussed below, this infeasible feature of the discrete time case is quite different in continuous time.

**Remark 6.** The limit distribution \((2.13)\) is a ratio of two independent Gaussian variates and has heavy tails, just as the Cauchy limit in the special case \((2.17)\) where \( D = 0 \). Observe that

\[ R_N^N N^{\alpha/2} \left( \hat{A} - A \right) \Rightarrow 2c \frac{\sigma_0 u_0}{\sigma_{xx} U_x + (2c)^{1/2} D} \]

\[ = b \frac{U_0}{U_x + d} \]

\[ b = \frac{2c \sigma_0 u_0}{\sigma_{xx}}, \quad d = \frac{(2c)^{1/2}}{D}. \]
\[
\sqrt{N} (\hat{\beta} - \beta) = \frac{1}{\sqrt{N}} \sum_{t=1}^{N} x_{t-1}^2 (1 - \frac{N}{N-1} \sum_{t=1}^{N} x_{t-1} U_{t,0}) \Rightarrow N(0, \sigma^2_{\hat{x}}).
\]

\section*{Box 1}

When \( D \geq 0 \), the density of \( U = \frac{y_t}{\sigma_0 + \hat{\sigma}_0} \) is
\[
p_U(u) = \frac{e^{-u^2/2}}{\sqrt{2\pi}} \left[ 1 + \frac{q}{\psi(q)} \int_0^q \varphi(y) dy \right].
\]
\( \varphi(y) = \frac{e^{-y^2/2}}{\sqrt{\pi}} \), \( q = \frac{d}{\sqrt{1 + u^2}} \).
(E.g., see Marsaglia, 1965) and has Cauchy-like tails.

We have the following expression for the LS estimator \( \hat{\beta} \) of the slope coefficient in the continuous time model (2.1), which is given by
\[
\hat{\beta} - \beta = \left( \sum_{t=1}^{N} x_{t,0}^2 \right)^{-1} \left( \sum_{t=1}^{N} x_{t,0} U_{t,0} \right)
\]
\[
= \frac{1}{\lambda} \left( \sum_{t=1}^{N} x_{t,0}^2 \right)^{-1} \left( \sum_{t=1}^{N} x_{t,0} U_{t,0} \right)^{2}.
\] (2.20)

The associated limit theory is given in the following corollary.

\textbf{Corollary 2.2.} Let \( h = N^{-1/2} \). For the continuous time system (2.1)-(2.2) with \( \kappa < 0 \), assume that there exists \( \alpha \in (0, 1) \), such that \( h^{1/2} T \rightarrow 1 \) as \( T \rightarrow \infty \), we have
\[
\frac{\sigma_0 U_0}{\sqrt{\lambda}} (\hat{\beta} - \beta) \Rightarrow (-2\kappa) \frac{\sigma_0 U_0}{\sigma_0 U_0 + (-2\kappa)^{1/2} (\sigma_0 - \mu)}. \quad (2.21)
\]

\textbf{Remark 7.} The limit result in Corollary 2.2 can be extended to the case where \( u_0(t) \) is weakly dependent in (2.1). Using the limit results from Remark 1, we obtain
\[
\frac{\sigma_0 U_0}{\sqrt{\lambda}} (\hat{\beta} - \beta) \Rightarrow (-2\kappa) \frac{\sigma_0 U_0}{\sigma_0 U_0 + (-2\kappa)^{1/2} (\sigma_0 - \mu)}. \quad (2.22)
\]

\textbf{Remark 8.} After multiplying \( \sqrt{\lambda} \) on both sides of (2.13), the limit distribution (2.21) follows directly from (2.13) by replacing \( \sigma_0 \), \( \kappa \), and \( D \) in Theorem 2.1 with 1, and \( D^* = \frac{\kappa h}{\sigma_0} \), respectively, giving the stated result.

\textbf{Remark 9.} The continuous time counterpart of \( N^* \) is \( 1/h \) which is known for any given data, so there is no need to estimate the rate parameter \( \kappa \). The continuous time counterpart of \( c \) is \( -\kappa \) which can be consistently estimated by the least squares method as long as \( T \rightarrow \infty \). Analogous to (2.19), self-normalized statistics are free of nuisance parameters and hypothesis testing about \( \beta \) can be conducted using the residual variance estimate \( s_0^2 = N^{-1} \sum_{t=1}^{N} x_{t,0}^2 \), which satisfies \( s_0^2 \overset{P}{\rightarrow} \sigma_0^2 \), Corollary 2.2 and results (A.11) and (A.12) in Appendix A then give the following double asymptotics for the usual \( t \) statistic for testing \( H_0 : \beta = \beta^0 \)
\[
t_\beta = \frac{\hat{\beta} - \beta^0}{s_\beta} = \frac{\hat{\beta} - \beta^0}{\sigma_0 U_0 + (-2\kappa)^{1/2} (\sigma_0 - \mu)} \Rightarrow \frac{(-2\kappa) \sigma_0 U_0}{\sigma_0 U_0 + (-2\kappa)^{1/2} (\sigma_0 - \mu)} \sim U_0 \sim N(0, 1).
\]

Similarly, given \( \alpha_h = \exp(-kh) \), we have \( h \alpha_h = s_{h} + \alpha_p(h) \). Following WY, we have
\[
\alpha^N_h (\tilde{\kappa} - \kappa^0) \Rightarrow 2 \frac{\alpha_0 U_0}{\sigma_0 U_0 + (-2\kappa)^{1/2} D}. \quad (2.23)
\]

Clearly, \( t_\kappa \) is a feasible statistic for testing \( H_0 : \kappa = \kappa^0 \) in contrast to the discrete time case where the test statistic relies on the unknown rate parameter \( \alpha \).

\textbf{Remark 11.} If \( x_0 = \mu \), we have \( D^* = 0 \) and
\[
\frac{\alpha^N_h}{\sqrt{h}} (\hat{\beta} - \beta) = (-2\kappa) \frac{\sigma_0 U_0}{\sigma_0 U_0 + (-2\kappa)^{1/2} D} \Rightarrow \frac{(-2\kappa) \sigma_0 U_0}{\sigma_0 U_0 + (-2\kappa)^{1/2} D} \sim U_0 \sim N(0, 1).
\]

Please cite this article in press as: Chen Y., et al., Inference in continuous systems with mildly explosive regressors, Journal of Econometrics (2017), http://dx.doi.org/10.1016/j.jeconom.2017.08.016.
Remark 12. An alternate formulation of (2.1) is to set \( u_0(t) \overset{d}{=} 0 \). This formulation corresponds to a discrete system (2.5) in which \( u_{0t} \overset{d}{=} N(0, \sigma^2_0 h) \), so that the error variance depends on the sampling interval, just as the error \( u_{0t} \), in (2.4). In this case, in view of the scaling effect in the discrete time error, it follows that there is a faster rate of convergence in estimation of \( \beta \) and the limit distribution of \( \beta \) is given by

\[
\frac{\sigma_0^2 u_0}{h} (\hat{\beta} - \beta) \overset{d}{=} (2x) \frac{\sigma_0 u_0}{\sigma_0 u_0 + (2x)^{1/2} (x_0 - \mu)} ,
\]

(2.24)

when \( h \to 0 \) and \( T \to \infty \). In this alternate model, the relationship between \( x(t) \) and \( y(t) \) is exact in the limit, analogous to the relationship of limit Brownian motion processes \( B_t, B_t \) of cointegrated discrete series \( (y_t, x_t) \) where \( x_t = x_{t-1} + u_t \) and \( y_t = \beta x_t + u_t \) with \( (u_0, u_t) \) stationary and limiting linear relation \( B_t(t) = \beta B_t(t) \).

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 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 behaviors: (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,

\[
y_t = Ax_t + u_t ,
\]

(3.1)

\[
x_t = \mu + R_N x_{t-1} + u_t , \quad \text{with } x_0 = x_0 = O_p \left( N^{n/2} \right)
\]

and \( \mu = Op \left( N^{n/2} \right) \).

(3.2)

In this case, \( y_t \) and \( x_t \) are \( m \times 1 \) and \( K \times 1 \) vectors 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 \( C = \text{diag}(c_1, c_2, \ldots, c_K) > 0 \). We assume that the errors satisfy

\[
u_t = \left[ u_{0t}, u_{0t}' \right] \overset{d}{=} N(0, \Omega) ,
\]

\[
\Omega \begin{bmatrix}
\Omega_{00} & \Omega_{0x} \\
\Omega_{0x} & \Omega_{xx}
\end{bmatrix} .
\]

Let the standardized initialization and intercept satisfy \( \tilde{x}_0 = x_0 N^{n/2} \Rightarrow X^* \) and \( \tilde{\mu} = N^{n/2} \mu \Rightarrow \mu^* \). The model now modifies MP in two ways: (i) the initial value for \( x \) is \( \tilde{x}_0 = N^{n/2} \) which is larger than the \( Op(N^{n/2}) \) initialization in MP; (ii) a non-zero drift term of order \( Op(N^{n/2}) \) is included. Following closely the approach of MP, we obtain the limit theory for the LS estimator \( \hat{A} \) 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 affine 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 \( R_N = I_K + C/N^\alpha, \alpha \in (0, 1), C = \text{diag}(c_1, c_2, \ldots, c_K) > 0 \) and \( c_i \neq c_j \) for \( i \neq j \), when \( N \to \infty \), we have

\[
\text{(i)} \quad \text{vec} \left( \frac{1}{N} \sum_{i=1}^N u_{t-1}^t x_t^r \right) \Rightarrow \left( \int_0^\infty e^{-\beta C} \left( D + \tilde{u}_t \right)^* \left( D + \tilde{u}_t \right) \right) e^{-\beta C} dp \otimes \Omega_{00}^{1/2} \Omega_0
\]

\[
\text{(ii)} \quad \frac{1}{N} \sum_{i=1}^N R_N^j x_t^r R_N^j \Rightarrow \int_0^\infty e^{-\beta C} \left( D + \tilde{u}_t \right)^* \left( D + \tilde{u}_t \right) e^{-\beta C} dp + o_p(1)
\]

\[
\text{(iii)} \quad \text{vec} \left\{ N^\alpha \left( \tilde{A} - A \right) \right\} \Rightarrow \left\{ \int_0^\infty e^{-\beta C} \left( D + \tilde{u}_t \right)^* \left( D + \tilde{u}_t \right) e^{-\beta C} dp \right\} \otimes \Omega_{00}^{1/2} \Omega_0
\]

\[
\Rightarrow \left\{ \int_0^\infty e^{-\beta C} \left( D + \tilde{u}_t \right)^* \left( D + \tilde{u}_t \right) e^{-\beta C} dp \right\} \otimes \Omega_{00}^{1/2} \Omega_0
\]

(3.3)

where \( W_0 \overset{d}{=} N \left( 0, I_{mK} \right), \tilde{u}_t = \left( \int_0^\infty e^{-\beta C} \left( D + \tilde{u}_t \right)^* \left( D + \tilde{u}_t \right) e^{-\beta C} dp \right)^{1/2} U_{0t}, U_{0t} \overset{d}{=} N \left( 0, I_{mK} \right), D = X^* + C^{-1} \mu^* \), \( MN \) represents a mixed normal distribution.

Remark 13. Again, we can extend the limit result in Theorem 3.1 to the case where \( u_{0t} \) is weakly dependent. Let \( u_{0t} \) in the discrete time system (3.1)–(3.2) be a sequence of zero mean and weakly dependent errors, satisfying Assumption LP as defined in MP. In particular,

\[
N^{-1/2} \sum_{i=1}^N u_{0t} \Rightarrow N \left( 0, \Omega_{00} \right)
\]

with long-run variance matrix \( \Omega_{00} \) decomposed as

\[
\Omega_{00} = \Sigma_{00} + A_{10} + A_{01}
\]

(3.4)

where

\[
\Sigma = \begin{bmatrix}
\Sigma_{00} & \Sigma_{0x} \\
\Sigma_{0x} & \Sigma_{xx}
\end{bmatrix}
\]

and the one-sided long-run covariance is

\[
A_{10} := \sum_{h=1}^\infty E \left( u_{0t} u_{0t-h} \right).
\]

Using the limit results from Theorem 3.1 and Theorem 4.1 in MP, we obtain

\[
\text{vec} \left\{ N^\alpha \left( \tilde{A} - A \right) \right\} \Rightarrow \left\{ \int_0^\infty e^{-\beta C} \left( D + \tilde{u}_t \right)^* \left( D + \tilde{u}_t \right) e^{-\beta C} dp \right\} \otimes \Omega_{00}^{1/2} \Omega_0
\]

\[
\Rightarrow \left\{ \int_0^\infty e^{-\beta C} \left( D + \tilde{u}_t \right)^* \left( D + \tilde{u}_t \right) e^{-\beta C} dp \right\} \otimes \Omega_{00}^{1/2} \Omega_0
\]

(3.5)

Remark 14. If \( \tilde{x}_0 = -C^{-1} \mu \), then \( D = 0 \) and the limit (3.3) becomes

\[
\text{vec} \left\{ N^\alpha \left( \tilde{A} - A \right) \right\} \Rightarrow MN \left( 0, \left( \int_0^\infty e^{-\beta C} \tilde{u}_t \tilde{u}_t^* e^{-\beta C} dp \right) \otimes \Omega_{00} \right).
\]

This limit distribution corresponds to that in Theorem 4.1 of MP (2009, p. 496).
Remark 15. The limit distribution of $\hat{\mu}$ is
\[
\sqrt{N} (\hat{\mu} - \mu) = \sqrt{N} \left\{ \frac{1}{N} \sum_{t=1}^{N} u_{at} - \frac{1}{N} \left( \frac{1}{N^2} \sum_{t=1}^{N} u_{at} x_{t-1} R_{N}^{-1} \right) \right\},
\]
where $\mu$ is the distribution for the Wald statistic for testing $\hat{R}_{N} = \hat{R}_{0}$. When $Q_{0}$ is a $g \times K$ restriction matrix of full row rank $g \leq K$, and $q_{i}$ is a given $g \times 1$ vector, is
\[
W_{R_{N}} := \left\{ \left( \hat{R}_{N} - q_{i} \right)^{T} (\hat{S}_{R_{N}} q_{i})^{-1} \left( \hat{R}_{N} - q_{i} \right) \right\}^{\frac{1}{2}}.
\]
under the null. Similarly, given $R_{N} = I_{K} + C/N^{a}$, we can set $S_{C_{j}j} := N^{2a} S_{R_{N}}$, where $C_{j}$ is the $j$th column of $C$. Set $C_{j}$ is the covariance matrix of $C_{j}$. Further, $C_{j} \Rightarrow C_{j}^{0} = (R_{N} - R_{0}^{a}) N^{a}$ leads to the following limit theory for $C_{j}$
\[
(C_{j} - C_{j}^{0}) \rho_{j}^{N} \Rightarrow MN \left( 0, \int_{0}^{\infty} e^{-2p\rho_{j}} (D_{j} + \tilde{U}_{j})^{2} dp \right)^{-1} \Omega_{\alpha}.
\]
Hence, if $\alpha$ is known, we have the corresponding feasible Wald statistic for testing the restrictions $H_{0} : Q_{0} C_{j} = \hat{Q}_{0} C_{j} = q_{i}$
\[
W_{C_{j}} := \left\{ \left( \hat{Q}_{0} C_{j} - q_{i} \right)^{T} (\hat{Q}_{0} C_{j})^{-1} \left( \hat{Q}_{0} C_{j} - q_{i} \right) \right\}^{\frac{1}{2}}.
\]
under the null and with full row rank $Q_{0}$. If $\alpha$ is unknown, just as in the scalar model, the estimated variance matrix $S_{C_{j}j} := N^{2a} S_{R_{N}}$ is unavailable and inference using this limit theory for $C_{j}$ 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 $H_{0} : \hat{C}_{j} = C_{j}^{0} e_{j}$ and a test analogous to the scalar case can be mounted using the $j$th diagonal element of the (unrestricted) estimate $\hat{C}_{j}$ 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 appear 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 $\sum_{i=j}^{N} x_{i} x_{i}'$ is singular due to commonality in the explosive behavior of the components of $x_{i}$. Let $R_{K} = \rho_{w} I_{K}$ with $\rho_{w} = 1 + c/N^{a}$. Following MP, we rotate regression coordinates to address the singularity using an orthogonal random matrix $H_{K} = [H_{K}, H_{K} H_{N}^{-1}]$ where
\[
H_{K} H_{K}^{-1} = (\rho_{w} \rho_{w}^{-1})^{1/2} \text{ and } H_{K} H_{K}^{-1} \text{ is a } K \times (K - 1) \text{ orthogonal complement matrix such that } H_{K} H_{K}^{-1} = 0.\text{ Then, } H_{K} H_{K}^{-1} H_{K} H_{N}^{-1} = \rho_{w} I_{K} - H_{K} H_{K}^{-1}. \text{ The limit of } H_{K} H_{K}^{-1} \text{ is denoted as } H_{K} H_{K}^{-1} = \rho_{w} I_{K} - X_{c} X_{c}' \text{ where } X_{c} \text{ is defined in (3.6) in Theorem 3.2. Next, rotate the regressor } x_{i} \text{ by } H_{K} H_{K}^{-1} \text{ and transform to the variate } z_{i} = H_{K} x_{i} = [H_{K} x_{i}, H_{K} H_{N}^{-1}] := \begin{bmatrix} z_{i,1}^{T} & z_{i,2}^{T} \end{bmatrix}. \text{ The following result gives the required limit theory for the LS estimator } \hat{A} \text{ in this case.}
\]

Theorem 3.2. For the discrete time system (3.1)–(3.2) with $R_{N} = I_{K} + C/N^{a}$, $C = \text{diag} (c_{1}, c_{2}, \ldots, c_{K})$ with $c_{i} = c > 0$ for $i = 1, \ldots, K$, when $N \to \infty$, we have
\[
(i) \frac{1}{N^{a}} \sum_{t=1}^{N} z_{i,2} z_{j,2} \Rightarrow M, \text{ with } M = H_{K} (W_{K} / \tau + \frac{1}{\tau} \Omega_{\alpha}) H_{K}.
\]
\[
(ii) \frac{1}{N^{a}} \text{ vec } \left\{ \sum_{t=1}^{N} u_{at} z_{i,2} \right\} \Rightarrow \{M \otimes \Omega_{\alpha} / 2 \times (N, \Omega_{\alpha}) - 1 \times N (0, \Omega_{\alpha}) \}
\]
where $H_{K}$ is a $K \times (K - 1)$ random matrix that is an orthogonal complement to
\[
X_{c} = \left( D + \tilde{U}_{c} \right) / \left( D + \tilde{U}_{c} \right)^{1/2} \left( D + \tilde{U}_{c} \right)^{1/2}.
\]
satisfying $H_1 H_1' = I_k - X_1 X_1'$ and with
\[
\hat{U}_x = \left( \int_0^\infty e^{-r} \Omega_{xx} e^{-r} \, dr \right)^{1/2} U_x = \Omega_{xx}^{1/2} U_x / (2c)^{1/2}, \quad \text{and} \quad D = X' + \mu' / c,
\]
\[
D_1 = \sqrt{\text{S}_1} = \sum_{t=1}^T \hat{u}_t.
\]

Remark 18. When $u_{xt}$ is weakly dependent and satisfies the conditions discussed in earlier remarks, the limit result again follows MP and has the form
\[
N^{1/2} \{ \hat{\Theta} - \Theta \} \overset{d}{\to} \mathcal{N} (0, \Sigma_{xx}).
\]

Remark 19. The limit distribution of $\hat{u}_{0t}$ is obtained as follows:
\[
\sqrt{N} (\hat{u}_{0t} - \mu) = \frac{1}{N} \left[ \frac{1}{N} \sum_{t=1}^N u_{xt} - \frac{1}{N} \left( \frac{1}{N^{1/2}} \sum_{t=1}^N u_{xt} \right) \right] \times \left[ \frac{1}{N^{1/2}} \sum_{t=1}^N z_{t-1} \right]^{-1} \left( \frac{1}{N^{1/2}} \sum_{t=1}^N z_{t-1} \right)
\]
\[
= \frac{1}{\sqrt{N}} \sum_{t=1}^N u_{xt} + o_p (1) \Rightarrow \mathcal{N} (0, \Sigma_{xx}).
\]

Remark 20. Define $\hat{u}_{0t} = y_t - \hat{A}_t x_t$ and let the estimate of the error variance matrix be
\[
\mathbf{S}_{00} = N^{-1} \sum_{t=1}^N \hat{u}_{0t} \hat{u}_{0t}' \Rightarrow \mathcal{O}_{xx},
\]

and the estimated variance matrix of $\hat{A}_t$ be
\[
\mathbf{S}_{A \hat{A}} = \left( \sum_{t=1}^N \hat{A}_t \right)^{-1} \Rightarrow \mathcal{O}_{xx}.
\]

Following Theorem 3.2, we have the following limit theory for the Wald statistic for testing $H_0 : Q \hat{A}_t = Q A_0 = q_t$
\[
W_{A_t} = \left[ Q \hat{A}_t \right] \left( Q \mathbf{S}_{A \hat{A}} Q \right)^{-1} \left( Q \hat{A}_t \right)' \Rightarrow \chi^2.
\]

Remark 21. Let $\hat{u}_{xt} = x_t - \hat{A}_t x_t - \mu$, giving the error variance matrix estimate
\[
\mathbf{S}_{xx} = N^{-1} \sum_{t=1}^N \hat{u}_{xt} \hat{u}_{xt}' \Rightarrow \mathcal{O}_{xx},
\]

and the corresponding estimate of the variance matrix of $\hat{R}_N$ (in column vector form)
\[
\mathbf{S}_{RR} = \left( \sum_{t=1}^N x_{t-1} x_{t-1}' \right) \times \mathcal{O}_{xx}.
\]

Then the Wald statistic for testing $H_0 : Q \hat{c} (\hat{R}_N) = Q \hat{c} (\hat{R}_0) = r$, where $Q$ is a $g \times mK$ restriction matrix of full row rank $g \leq mK$, is
\[
W_{R_N} = \left( Q \hat{c} (\hat{R}_N - R_0) \right) \left( Q \mathbf{S}_{RR} Q \right)^{-1} \left( Q \hat{c} (\hat{R}_N - R_0) \right)' \Rightarrow \chi^2.
\]

3.2. Limit results in the continuous time framework

The above results apply to the multivariate continuous time system
\[
y(t) = \beta x(t) + u_0 (t),
\]
\[
dx(t) = \kappa (\mu - x(t)) \, dt + \Omega_{xx}^{1/2} \, dB_{x}(t),
\]
\[
x(0) = x_0 = O_p (1), \kappa < 0,
\]

where $u_0 (t) \sim \mathcal{N} (0, \Omega_{xx})$ is a Gaussian noise. Let $B_x (t) = \Omega_{xx}^{1/2} W_x (t)$ where $W_x$ is standard vector Brownian motion that may be correlated with $W_0$. The driver process $x(t)$ then follows a
multivariate Ornstein–Uhlenbeck process with persistence matrix \( \kappa \), where \( \kappa = \text{diag}(\kappa_1, \kappa_2, \ldots, \kappa_K) \) 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 \( \beta \), an \( m \times K \) matrix of coefficients which captures co-movement between \( y(t) \) and \( x(t) \).

The exact discrete time representation of \((3.9)-(3.10)\) is given by (see Phillips, 1972)

\[
y_{th} = \beta x_{th} + u_{0,th}, \quad \text{(3.11)}
\]

\[
x_{th} = \phi_0 (k) x_{(t-1)h} + g_0 + u_{x,th}, \quad x_0 = \phi_0 (1), \quad \text{(3.11)}
\]

where

\[
a_n (k) = \exp (-\kappa h), \quad \text{for } i \neq j, \quad \text{(3.9)}
\]

\[
u_{x,th} = \int_{(t-1)h}^{th} e^{-\kappa (t\cdot -1h)} \Omega_{xx} dB_t(s) \overset{d}{=} \mathcal{N}(0, \Omega_{xx}), \quad \text{(3.12)}
\]

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

\[
y_{th} = \beta x_{th} + u_{0,th}, \quad \text{(3.12)}
\]

\[
x_{th} = \phi_0 (k) x_{(t-1)h} + g_0 + u_{x,th}, \quad x_0 = \phi_0 (1), \quad \text{(3.13)}
\]

For the continuous time system \((3.12)-(3.13)\), the double asymptotic property for the LS estimator of the coefficient matrix \( \beta \) when \( \kappa \) has distinct diagonal elements (i.e., \( \kappa_i \neq \kappa_j \) for \( i \neq j \)) is given in the following corollary.

**Corollary 3.3.** Let \( h = N^{-\alpha} \). For the continuous time system \((3.12)-(3.13)\) with \( \kappa = \text{diag}(\kappa_1, \kappa_2, \ldots, \kappa_K) \) with \( \kappa_i \neq \kappa_j \) for \( i \neq j \), assuming that there exists \( \alpha \in (0, 1) \), such that \( h^{1-\alpha} T^\alpha \to 1 \) as \( T \to \infty \), we have

\[
\begin{align*}
\text{vec} \left[ \left( \frac{\sqrt{h}}{h} (\hat{\beta} - \beta) \right)^{\sigma_{\beta}} \right] & \Rightarrow \left[ \left( \int_0^\infty e^{vp} (D + \bar{U}_j) (D + \bar{U}_j)^{\sigma_{\beta}} dp \right)^{-1} \otimes \Omega_{00} \right]^{1/2} \times \mathcal{N}(0, I_{mn}) \quad \text{(3.14)}
\end{align*}
\]

**Remark 22.** By multiplying both sides of \((3.3)\) by \( \sqrt{h} \), the double asymptotic distribution \((3.14)\) follows directly from \((3.3)\) with \( \mu^* = \mu \), \( C = -\kappa \) and \( D = \kappa_0 = \mu \). To enhance readability in terms of the relationship between the systems, we provide in Table 1 the correspondence between the models, variables and parameters in the discrete and continuous time cases.

**Remark 23.** The LS estimator of \( \kappa \) is consistent since \( h \) is known. Let \( S_{00} = N^{-1} \sum_{i=1}^{\infty} \bar{u}_{0,th} \bar{u}_{0,th}^{\top} \), which satisfies \( S_{00} \overset{p}{\to} \Omega_{00} \).

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

\[
W_{ij} := \left\{ Q_j (\hat{\beta}_j - \beta_j^0) \right\} \left( \left\{ Q_j S_{ij} Q_j \right\}^{-1} \left\{ Q_j (\hat{\beta}_j - \beta_j^0) \right\} \right\} \Rightarrow \chi^2_{\nu}.
\]

leading to feasible inference about \( \beta_j \) in the continuous time framework.

**Remark 24.** Let \( q_j \) be the \( j \)th column of \( a_0 (k) \). 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

\[
W_{ij} := \left\{ Q_j (\hat{a}_j - a_j^0) \right\} \left( \left\{ Q_j S_{ij} Q_j \right\}^{-1} \left\{ Q_j (\hat{a}_j - a_j^0) \right\} \right\} \Rightarrow \chi^2_{\nu},
\]

where \( S_{ij} = \left( \sum_{i=1}^{\infty} x_{ij}^2 (t_{ij} - 1) \right) - \frac{1}{n} \left( \sum_{i=1}^{\infty} x_{ij} (t_{ij} - 1) \right)^2 \). Let \( \kappa \) denote the jth column of \( \beta \). Given the matrix exponential relation, we have the covariance matrix of \( \kappa' \) which satisfies \( h^2 S_{ij} \Gamma_j \) is \( \mathbb{H}_0 : \mathbb{Q}_k \beta_j \) where \( \mathbb{Q}_k \beta_j \) is the joint column of \( \beta_j \). Given the known diagonal structure, we can use either SUR estimation or restricted LS (in which only the diagonal elements of the original LS 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 \( i = 1, \ldots, K \).

**Corollary 3.4.** Let \( h = N^{-\alpha} \). For the continuous time system \((3.12)-(3.13)\) with \( \kappa = \text{diag}(\kappa_1, \kappa_2, \ldots, \kappa_K) \) and \( \kappa_i = \kappa < 0 \) for \( i = 1, \ldots, K \), assuming that there exists \( \alpha \in (0, 1) \), such that \( h^{1-\alpha} T^\alpha \to 1 \) as \( T \to \infty \), we have

\[
\begin{align*}
\text{vec} \left[ \sqrt{h} (\beta - \beta) \right] & \Rightarrow \left[ H_{ij}^{1/2} \left( \mu \mu' + \frac{1}{2\kappa} \Omega_{xx} \right) H_{ij}^{1/2} \right]^{-1/2} \otimes \Omega_{00}^{1/2} \times \mathcal{N}(0, I_{mn}) \quad \text{(3.15)}
\end{align*}
\]

**Table 1**

<table>
<thead>
<tr>
<th>Discrete time</th>
<th>Continuous time</th>
</tr>
</thead>
<tbody>
<tr>
<td>( y_i = \beta x_i + u_{i} )</td>
<td>( \bar{y}_n = \beta \bar{x}<em>n + u</em>{in} )</td>
</tr>
<tr>
<td>( x_i = \mu_k + \mu kx_{i-1} + u_{i} )</td>
<td>( \bar{x}_n = a_0 (k) \bar{x}_n-1 + \bar{u}_n )</td>
</tr>
<tr>
<td>( x_0 = \phi_0 (N^{1/2}) )</td>
<td>( \bar{x}<em>0 = h^{-1/2} x</em>{0h} = \phi_0 (h^{-1/2}) )</td>
</tr>
<tr>
<td>( \mu = \phi_k (N^{1/2}) )</td>
<td>( \bar{g}_0 = \phi_0 (h^{-1/2}) )</td>
</tr>
</tbody>
</table>

\( \mu \) with \( N^{1/2} \mu \to \mu \) for \( h^{-1/2} \to 1 \) with \( \mu \) with \( h^{-1/2} \to 1 \)
Remark 26. Both Corollaries 3.3 and 3.4 can be extended in the same way as before to cover the case where $u_0(t)$ is weakly dependent.

Remark 27. The double asymptotic distribution (3.15) follows directly from (3.7) with $\mu = \kappa \mu$ and $c = -k$.

Remark 28. The Wald statistic for testing $H_0 : Q_\nu \vec{c} = Q_\nu \vec{c}^0$ ($\beta^0$) is $q$ for full row rank ($Q, q$) is then

$$W_{\beta} := \left\{Q_\nu \vec{c} - Q_\nu \vec{c}^0 \right\}^{-1} \left\{Q_\nu \vec{c} - q \right\} \Rightarrow \chi^2,$$

leading to feasible inference about the matrix coefficient $\beta$ in the continuous time framework. Inference about the full matrix $\beta$ is possible in this case because of the common factorization convergence rate in (3.15).

Remark 29. The Wald statistics for testing full rank restrictions on $a_0$ and $\kappa$ such as $H_0 : Q_\nu \bar{a}_0 = Q_\nu \bar{a}_0^0 = q$ and $H_0 : Q_\nu \vec{c} = Q_\nu \vec{c}^0 = q$ are defined in a similar way and have the following chi-squared limits:

$$W_{\bar{a}} := \left\{Q_\nu \bar{a}_0 - Q_\nu \bar{a}_0^0 \right\}^{-1} \left\{Q_\nu \bar{a}_0 - q \right\} \Rightarrow \chi^2,$$

and

$$W_{\vec{c}} := \left\{Q_\nu \vec{c} - Q_\nu \vec{c}^0 \right\}^{-1} \left\{Q_\nu \vec{c} - q \right\} \Rightarrow \chi^2,$$

again leading to feasible inference on $a_0$ and $\kappa$ because of the common factorization convergence rate.

Remark 30. When $x_t$ has a common explosive root, LS estimation by $\hat{a}_0$ and $\hat{\kappa}$ 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 shown in the Monte Carlo simulation below.

4. Monte Carlo studies

This section examines the performance of the double asymptotic test in simulations. We generate data from model (2.3)–(2.4) with $\kappa = -2, \sigma_0 = \sigma_\nu = 1, \mu = 0$, 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 (2.21) and the finite sample distribution of the coefficient based test (called the C test hereafter) $-\frac{\sigma_\nu}{2 \kappa \sqrt{n}} (\hat{\beta} - \beta)$ and $t_\beta$ from Remark 9. In addition, we provide comparisons of the densities of the limit distributions and finite sample distributions of the $C$ test statistic and $t_\beta$ statistic. The number of replications is set at $10,000$.

Tables 2–4 report the percentiles when $x_0 = 0, 3, 10$ by using the true values of $\kappa$ and $\mu$. It can be seen that both the double asymptotic distribution and the finite sample distribution are sensitive to changes in the initial condition for the $C$ test. The $C$ test is recommended for empirical studies since it provides a tighter confidence interval than the $t$ test as demonstrated in our empirical study. In all cases the new limit distribution provides a good approximation to the finite sample distribution.

Figs. 1 and 2, and 3 plot the densities of the $C$ test statistic and $t_\beta$ test statistic when $T = 4$. The result is similar to the case of $T = 10$, which is not reported. These plots show the limit 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 (3.11), with $\beta = [1, 1]^T, \chi(0) = [3, 1]^T$,

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

$\mu = [1, 1]^T, \vec{c}(\kappa) = [\kappa_1, 0, 0, \kappa_2]^T$ 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, $a_1 = \exp(-\kappa h) = 1.0168$ 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 $W_{\beta}, W_{\beta}^{OLS}, W_{\beta}^{ReOLS}$, and $W_{\beta}^{SUR}$, where $W$ indicates the Wald test statistic for the parameter of interest, $r_\beta = [1, 1], r_{\nu} = [1.0168, 1.0339], r_\nu, r_{\beta} = [1.0168, 1.0339], r_\nu = [1.0168, 1.0339]$, and $r_{\beta} = [1.0168, 1.0339]$ for full row rank $\kappa$ in the first case, $r_\nu = [1.0168, 1.0339], r_{\beta} = [1.0168, 1.0339]$ in the second case. In addition, LS corresponds to the LS estimates, ReOLS to the estimates based on the diagonal elements of LS estimates, and SUR to the estimates based on the seemingly unrelated regressions.

Several conclusions can be drawn from this Monte Carlo study. First, for $\beta$, our limit distribution well approximates the finite sample distribution in both cases. Second, for $\kappa$, the limit distribution is closer to the finite sample distribution based on SUR than those based on LS or ReOLS. While in Case 1 where the explosive roots are distinct, the finite sample distributions based on LS and SUR are very close to each other, in Case 2 where there is a common explosive root, the limit distribution is much closer to the finite sample distribution based on SUR than to those based on OLS. But in neither case was there any evidence to support the use of ReOLS. The result in Table 5 suggests that it is preferable to use the limit distribution for inference about $\kappa$ based on SUR.

To understand why SUR provides much better results than LS for testing hypotheses about $\kappa$ in Case 2, Table 6 reports the mean and variance of the two sets of estimates of $\kappa$ in both cases. While SUR produces slightly better estimates than LS in Case 1, it yields much better estimates of $\kappa$ 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 LS estimate of the common explosive autoregression parameter is biased downward, suggesting that the estimate of $\kappa$ is biased upward. Naturally this bias distorts the asymptotic approximation 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 US nationwide real estate market and 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-city index and metropolitan area indices. The S&P/Case–Shiller home price indices are the leading measures of the US residential real estate prices, tracking changes in the value of residential real estate nationwide. Monthly data for the composite 20-city index and 20 metropolitan area indices between February 2000 and August 2014 were downloaded from the St. Louis Fed.3 Similar to the capital asset pricing model CAPM, we use the composite 20 index to measure the overall market movements. A multi-equation continuous time system (2.1)–(2.2) is estimated with $x_t$ as the logarithmic composite 20 index (called the countrywide index in the present paper) and each $y_t$ being one of the logarithmic metropolitan area indices that is found to be explosive. The coefficient $\beta$ then measures the co-movement of each metropolitan area index with the countrywide index. With monthly data, the sampling interval is set to $h = \frac{1}{12}$. The initial

---

Fig. 1. Density comparison between the $C$ test and $t$ test for both finite sample and limit distributions, when the initial value is $x_0 = 0$.

Fig. 2. Density comparison between the $C$ test and $t$ test for both finite sample and limit distributions, when the initial value is $x_0 = 3$. 
Table 2
Comparison of the finite sample and double asymptotic distributions of $\hat{\beta}$, when the initial value is $x_0 = 0$.

<table>
<thead>
<tr>
<th>Time Span</th>
<th>T = 4</th>
<th>T = 10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency</td>
<td>C test</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1%</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Time Span</th>
<th>T = 4</th>
<th>T = 10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency</td>
<td>C test</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1%</td>
</tr>
<tr>
<td>Daily</td>
<td>New</td>
<td>−0.427</td>
</tr>
<tr>
<td>(h = 1/252)</td>
<td>Finite Sample</td>
<td>−0.416</td>
</tr>
<tr>
<td>Weekly</td>
<td>New</td>
<td>−0.427</td>
</tr>
<tr>
<td>(h = 1/52)</td>
<td>Finite Sample</td>
<td>−0.413</td>
</tr>
<tr>
<td>Monthly</td>
<td>New</td>
<td>−0.427</td>
</tr>
<tr>
<td>(h = 1/12)</td>
<td>Finite Sample</td>
<td>−0.385</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Time Span</th>
<th>T = 4</th>
<th>T = 10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency</td>
<td>C test</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1%</td>
</tr>
<tr>
<td>Daily</td>
<td>New</td>
<td>−0.118</td>
</tr>
<tr>
<td>(h = 1/252)</td>
<td>Finite Sample</td>
<td>−0.114</td>
</tr>
<tr>
<td>Weekly</td>
<td>New</td>
<td>−0.118</td>
</tr>
<tr>
<td>(h = 1/52)</td>
<td>Finite Sample</td>
<td>−0.114</td>
</tr>
<tr>
<td>Monthly</td>
<td>New</td>
<td>−0.118</td>
</tr>
<tr>
<td>(h = 1/12)</td>
<td>Finite Sample</td>
<td>−0.109</td>
</tr>
</tbody>
</table>

Table 5
Comparison of the finite sample and double asymptotic distributions of the Wald tests related to $\hat{\beta}_{2,1}$ and $x_{2,1}$.

<table>
<thead>
<tr>
<th>Wald test</th>
<th>Percentile</th>
<th>Case 1: $x_1 = -0.2, x_2 = -0.4$</th>
<th>Case 2: $x_1 = -0.2, x_2 = -0.2$</th>
<th>1%</th>
<th>2.50%</th>
<th>10%</th>
<th>90%</th>
<th>97.50%</th>
<th>99%</th>
<th>1%</th>
<th>2.50%</th>
<th>10%</th>
<th>90%</th>
<th>97.50%</th>
<th>99%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asymptotic $W_{\hat{\beta}}$</td>
<td>Finite Sample</td>
<td>0.020</td>
<td>0.051</td>
<td>0.211</td>
<td>4.605</td>
<td>7.378</td>
<td>9.210</td>
<td>0.020</td>
<td>0.051</td>
<td>0.211</td>
<td>4.605</td>
<td>7.378</td>
<td>9.210</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$W_{\hat{\beta}}$</td>
<td>Finite Sample</td>
<td>0.023</td>
<td>0.057</td>
<td>0.206</td>
<td>4.662</td>
<td>7.373</td>
<td>9.354</td>
<td>0.021</td>
<td>0.055</td>
<td>0.226</td>
<td>4.831</td>
<td>7.871</td>
<td>9.760</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$W_{\hat{\beta}}^{\text{KS}}$</td>
<td>Finite</td>
<td>0.020</td>
<td>0.048</td>
<td>0.099</td>
<td>5.264</td>
<td>9.215</td>
<td>12.752</td>
<td>0.029</td>
<td>0.059</td>
<td>0.104</td>
<td>5.503</td>
<td>10.808</td>
<td>17.906</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$W_{\hat{\beta}}^{\text{KS}}$</td>
<td>Finite</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.044</td>
<td>1.466</td>
<td>7.980</td>
<td>0.035</td>
<td>0.104</td>
<td>0.104</td>
<td>5.503</td>
<td>10.808</td>
<td>17.906</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Please cite this article in press as: Chen Y, et al., Inference in continuous systems with mildly explosive regressors. Journal of Econometrics (2017), http://dx.doi.org/10.1016/j.jeconom.2017.08.016.
value in each equation of the system is set to the composite 20 index in January 2000, i.e., \( x_0 \). 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 recent work in Phillips and Yu, (2011) that documents the presence of the explosive behavior in the US real estate market over much of this period. Before estimating the main model (2.1), we first explore the presence of explosive behavior in the market index \( x(t) \) and in the individual area indices \( y(t) \) by estimating \( \kappa \) and \( \kappa_y \) and obtaining the standard errors, the t-statistics, and the 99% and 90% confidence intervals based on the C test and the t test using the double asymptotic theory developed in WY. Results are reported in Table 7. For the countrywide index, the LS estimate of \( \kappa \) is \(-0.0467\), with the estimated standard error of 0.0002 and the t statistic of \(-2.3538\). Its 90% confidence interval is \([-0.0486, -0.0449]\) based on the C test and \([-0.0793, -0.0141]\) based on the t test. All the results confirm the presence of explosive behavior in \( x_0 \) over this period. For 11 out of 20 individual area indices, we find explosive behavior based on both the 90% and the 95% confidence intervals by the C test. Moreover, upon comparison of the confidence intervals reported in Table 7, it is apparent that the C test produces tighter confidence intervals than the t test.

For the 11 areas that exhibit explosive behavior, we further study possible co-movement with the market index. We report the LS estimate of \( \beta \), the estimated standard error, the t-statistic, and the 99% and 90% confidence intervals based on the C test and the t test in Table 8. Allowing for possible weak dependence in \( u_{0t} \) and following Remark 7 we estimate the variance and long-run variances of \( u_{0t} \) by

\[
\widehat{\sigma}^2_{00} = \frac{1}{N} \sum_{h=1}^{N} \widehat{u}_{0t} \widehat{u}_{0t}, \quad \widehat{\sigma}^2_{00} = \widehat{\sigma}^2_{00} + 2\widehat{\lambda}^2_{00},
\]

with truncation lag \( M_n = N^{1/3} \). Evidently, the 90% confidence intervals are quite tight and comfortably reject the null hypothesis \( H_0 : \beta = 0 \) in all cases. The confidence intervals can also be used to assess whether \( \beta \) is 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 countrywide index, giving a useful perspective on the relationship of different metropolitan area indices to the national index. The results show that LA, Miami, DC, and NY have more “aggressive” real estate markets in the US than the whole market, in the sense that the index for these cities moves more than the countrywide index.
Table 7
The estimated persistence parameter in $y_{30}$ and $x_{4}$, the confidence intervals for persistence parameter in the US logarithmic real estate market.

<table>
<thead>
<tr>
<th>Place</th>
<th>$\kappa$ or $x_4$</th>
<th>Std Err</th>
<th>$t$-stat</th>
<th>C test 95% CI</th>
<th>C test 90% CI</th>
<th>$t$ test 95% CI</th>
<th>$t$ test 90% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Market</td>
<td>-0.0467</td>
<td>1.9843E-02</td>
<td>-2.3538</td>
<td>-0.0496</td>
<td>-0.0438</td>
<td>-0.0486</td>
<td>-0.0449</td>
</tr>
<tr>
<td>SF</td>
<td>0.0669</td>
<td>6.8910E-02</td>
<td>0.9705</td>
<td>0.0753</td>
<td>0.0584</td>
<td>0.0723</td>
<td>0.0615</td>
</tr>
<tr>
<td>LA</td>
<td>-0.0671</td>
<td>2.5867E-02</td>
<td>-3.3658</td>
<td>-0.0913</td>
<td>-0.0829</td>
<td>-0.0907</td>
<td>-0.0844</td>
</tr>
<tr>
<td>Las Vegas</td>
<td>-0.1072</td>
<td>5.0086E-02</td>
<td>-1.8146</td>
<td>-0.1148</td>
<td>-0.1027</td>
<td>-0.1101</td>
<td>-0.1043</td>
</tr>
<tr>
<td>Miami</td>
<td>-0.1605</td>
<td>1.8272E-02</td>
<td>-8.8739</td>
<td>-0.1654</td>
<td>-0.1556</td>
<td>-0.1636</td>
<td>-0.1574</td>
</tr>
<tr>
<td>Phoenix</td>
<td>-0.3464</td>
<td>4.9243E-02</td>
<td>-7.0355</td>
<td>-0.3497</td>
<td>-0.3431</td>
<td>-0.3486</td>
<td>-0.3443</td>
</tr>
<tr>
<td>San Diego</td>
<td>0.0526</td>
<td>3.3438E-02</td>
<td>1.5729</td>
<td>0.0587</td>
<td>0.0465</td>
<td>0.0565</td>
<td>0.0487</td>
</tr>
<tr>
<td>Denver</td>
<td>0.4462</td>
<td>5.7549E-02</td>
<td>7.7529</td>
<td>1.0475</td>
<td>0.1569</td>
<td>0.8304</td>
<td>0.0671</td>
</tr>
<tr>
<td>DC</td>
<td>-0.0536</td>
<td>2.3675E-02</td>
<td>2.2647</td>
<td>-0.0568</td>
<td>-0.0556</td>
<td>-0.0556</td>
<td>0.0054</td>
</tr>
<tr>
<td>Chicago</td>
<td>-0.0199</td>
<td>2.8680E-02</td>
<td>-0.6092</td>
<td>-0.0213</td>
<td>-0.0184</td>
<td>-0.0208</td>
<td>-0.0209</td>
</tr>
<tr>
<td>Boston</td>
<td>0.2315</td>
<td>3.2179E-02</td>
<td>7.1950</td>
<td>0.3131</td>
<td>0.1497</td>
<td>0.2837</td>
<td>0.1794</td>
</tr>
<tr>
<td>Charlotte</td>
<td>-0.1667</td>
<td>1.2529E-01</td>
<td>1.3307</td>
<td>-0.1716</td>
<td>-0.1618</td>
<td>-0.1698</td>
<td>-0.1636</td>
</tr>
<tr>
<td>Portland</td>
<td>-0.3691</td>
<td>3.2029E-02</td>
<td>-1.1523</td>
<td>-0.3721</td>
<td>-0.3660</td>
<td>-0.3710</td>
<td>-0.3671</td>
</tr>
<tr>
<td>Dallas</td>
<td>0.2945</td>
<td>1.1385E-01</td>
<td>2.5866</td>
<td>0.4483</td>
<td>0.1402</td>
<td>0.3928</td>
<td>0.1962</td>
</tr>
<tr>
<td>Detroit</td>
<td>0.2508</td>
<td>7.3621E-02</td>
<td>3.4072</td>
<td>0.3506</td>
<td>0.1508</td>
<td>0.3146</td>
<td>0.1871</td>
</tr>
<tr>
<td>Atlanta</td>
<td>0.0453</td>
<td>4.3169E-02</td>
<td>1.0486</td>
<td>0.0502</td>
<td>0.0403</td>
<td>0.0484</td>
<td>0.0421</td>
</tr>
<tr>
<td>Seattle</td>
<td>-0.2937</td>
<td>2.8503E-02</td>
<td>1.0303</td>
<td>-0.2976</td>
<td>-0.2898</td>
<td>-0.2962</td>
<td>-0.2912</td>
</tr>
<tr>
<td>Minneapolis</td>
<td>0.1418</td>
<td>2.6267E-02</td>
<td>5.4002</td>
<td>0.1704</td>
<td>0.1132</td>
<td>0.1601</td>
<td>0.1236</td>
</tr>
<tr>
<td>Tampa</td>
<td>-0.2111</td>
<td>2.9201E-02</td>
<td>7.2302</td>
<td>-0.2158</td>
<td>-0.2064</td>
<td>-0.2141</td>
<td>-0.2081</td>
</tr>
<tr>
<td>Cleveland</td>
<td>0.1000</td>
<td>6.5902E-02</td>
<td>1.5175</td>
<td>0.1155</td>
<td>0.0845</td>
<td>0.1099</td>
<td>0.0901</td>
</tr>
<tr>
<td>NY</td>
<td>-0.0005</td>
<td>1.3207E-02</td>
<td>0.0024</td>
<td>-0.0005</td>
<td>-0.0004</td>
<td>-0.0005</td>
<td>-0.0004</td>
</tr>
</tbody>
</table>

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 MP 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 $c$, is consistently estimable in continuous time using the LS estimator, facilitating a efficient 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 discrete time systems that relate to systems formulated 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.

Appendix A

A.1. Proof of Theorem 2.1

Proof. The arguments here and in much of what follows closely mirror those of MP 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 $\sum_{t=1}^{N}x_{t}u_{t}$ is obtained using split sample arguments replacing summations in $\sum_{t=1}^{N}$ by $(\sum_{t=1}^{N}N_{1} + \sum_{t=1}^{N_{1}}N_{2})$ where $m_{N}$ is such that $m_{N}/N \to c$ so that with $c > 0$ and $\alpha \in (0,1)$ we have

\[ R_{N}^{-m_{N}} \sim \left( \frac{c}{N} \right)^{-m_{N}} \to 0. \]

(A.1)

(i) Start by writing $x_{t}$ in (2.9) as follows:

\[ x_{t} = R_{N}x_{0} + \sum_{j=1}^{t} R_{N}^{-j}u_{t} + \frac{1 - R_{N}}{1 - R_{N}} \mu. \]

(A.2)

so the standardized numerator can be decomposed as

\[ \left( R_{N}^{N}A_{N} \right)^{-1} I_{N} x_{t} u_{t} = \left( R_{N}^{N}A_{N} \right)^{-1} \sum_{t=1}^{N} R_{N}^{-t} u_{t} \left( x_{0} - \frac{\mu}{1 - R_{N}} \right) + \left( R_{N}^{N}A_{N} \right)^{-1} \sum_{t=1}^{N} u_{t} \frac{\mu}{1 - R_{N}}. \]
\[ + \frac{R_N^{N^a}}{\sqrt{N^a}} \left( \sum_{t=1}^{m_N} \sum_{t=m_N+1}^{N} \right) v_{0t} \left[ \frac{1}{\sqrt{N^a}} \sum_{j=1}^{t} \eta_{j}^{N^a} \right]. \] (A.3)

For the first term on the right hand side of (A.3), since \( \bar{X}_0 = x_{0N} N^{-a/2} \Rightarrow X^* \) and \( \bar{\mu} = N^{a/2} \mu \Rightarrow \mu^* \), we have

\[ D_N = N^{-a/2} x_0 - N^{-a/2} \frac{\mu}{1 - R_N} = N^{-a/2} x_0 - N^{a/2} \frac{\mu}{c + o(1)} \]
\[ = X^* + \frac{\mu^*}{c} = D. \] (A.4)

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

\[ \left( R_N^{N^a} \right)^{-1} \sum_{t=1}^{N} R_N^{N^a} v_{0t} \left( x_0 - \frac{\mu}{1 - R_N} \right) \]
\[ = N^{-a/2} \sum_{t=1}^{N} R_N^{N^a} \left( x_0 - \frac{\mu}{1 - R_N} \right) \]
\[ = D_N N^{-a/2} \sum_{t=m_N+1}^{N} R_N^{N^a} v_{0t} + o_p(1) \]
\[ = \frac{1}{\sqrt{N^a}} \sum_{t=m_N+1}^{N} \left( x_0 - \frac{\mu}{1 - R_N} \right) \]
\[ = D \]

where we assume the probability space is expanded in such a way so that the weak convergence \( \Rightarrow \) can be replaced by \( \Rightarrow d \). Also note that

\[ \sum_{t=1}^{m_N} R_N^{N^a} v_{0t} = o_p(1) \]

and then

\[ \left( R_N^{N^a} \right)^{-1} \sum_{t=1}^{m_N} R_N^{N^a} v_{0t} \left( x_0 - \frac{\mu}{1 - R_N} \right) = o_p(1). \]

Hence, for the first term of (A.3) we have by virtue of the martingale central limit theorem (MCLT), as in PM,

\[ \left( R_N^{N^a} \right)^{-1} \sum_{t=1}^{N} R_N^{N^a} v_{0t} \left( x_0 - \frac{\mu}{1 - R_N} \right) \]
\[ = D \]

where \( u_{0t}^* := u_{0N+1} - \frac{1}{\sqrt{N^a}} (0, \sigma^2_0) \) and \( U_0 = (N, 0, 1) \) since

\[ \frac{1}{N^a} \sum_{k=0}^{N-1} R_N^{N^a} u_{0k} \sim \frac{1}{\sqrt{N^a}} \sum_{k=0}^{N-1} \frac{R_N^{N^a}}{1 - R_N} \sim \frac{1}{N^a} \Rightarrow \frac{1}{N^a} \Rightarrow \frac{1}{\sqrt{N^a}}. \]

For the second term on the right hand side of (A.3), noting that \( R_N^{N^a} \sqrt{N^a} = (1 + \frac{1}{2})^{N^a} \sqrt{N^a} = o(1) \) for all \( \alpha \in (0, 1) \) we obtain

\[ \left( R_N^{N^a} \right)^{-1} \sum_{t=1}^{N} u_{0t} \left( \mu - \frac{1}{1 - R_N} \right) \sim \frac{1}{\sqrt{N^a}} \sum_{t=1}^{N} u_{0t} \]
\[ = o \left( e^{-\frac{1}{2\sqrt{N^a}}} \right) \]

The third term on the right hand side of (A.3) is

\[ \left( R_N^{N^a} \right)^{-1} \sum_{t=m_N+1}^{N} R_N^{N^a} v_{0t} \left( x_0 - \frac{\mu}{1 - R_N} \right) \]
\[ = o \left( e^{-\frac{1}{2\sqrt{N^a}}} \right) \]

where we use the fact that \( N^{-a/2} \sum_{t=m_N+1}^{N} R_N^{N^a} v_{0t} = o_p(1) \) from (A.6). We now use a joint MCLT for the components

\[ (U_{0N}, U_{0}) = \left( \frac{1}{\sqrt{N^a}} \sum_{t=m_N+1}^{N} R_N^{N^a} u_{0t}, \frac{1}{\sqrt{N^a}} \sum_{j=1}^{m_N} R_N^{N^a} u_{0j} \right) \]
\[ = o_p(1) \]

just as in PM and MP, using the fact that the limit variates \( (U_0, U_k) \) are independent because

\[ \frac{1}{\sqrt{N^a}} \sigma_{0e} \to 0. \]

Hence

\[ \left( \frac{1}{\sqrt{N^a}} \sum_{t=1}^{N} R_N^{N^a} u_{0t}, \frac{1}{\sqrt{N^a}} \sum_{j=1}^{m_N} R_N^{N^a} u_{0j} \right) \]
\[ = o_p(1) \]

Combining the above results and using (A.4) we obtain

\[ \left( R_N^{N^a} \right)^{-1} \sum_{t=1}^{N} x_t u_{0t} \]

Please cite this article in press as: Chen Y., et al., Inference in continuous systems with mildly explosive regressors. Journal of Econometrics (2017), http://dx.doi.org/10.1016/j.jeconom.2017.08.016.
\[\begin{align*}
\frac{(R_N^N)^{-1}}{R_N^{N_N^u}} & = \sum_{t=1}^{N} R_N^N u_{0t} \left( x_0 - \frac{\mu}{1 - R_N} \right) \\
& + \sum_{t=1}^{N} \frac{\mu}{1 - R_N} \\
& + \frac{R_N}{\sqrt{N}} \sum_{t=m+1}^{N} u_{0t} \left( \frac{1}{\sqrt{N}} \sum_{j=1}^{t} R_N^{-1} u_{0j} \right) + o_p(1) \\
& = \frac{(R_N^N)^{-1}}{R_N^{N_N^u}} \sum_{t=1}^{N} R_N^N u_{0t} \left( x_0 - \frac{\mu}{1 - R_N} \right) \\
& + \frac{R_N}{\sqrt{N}} \sum_{t=m+1}^{N} u_{0t} \left( \frac{1}{\sqrt{N}} \sum_{j=1}^{t} R_N^{-1} u_{0j} \right) + o_p(1) \\
& \Rightarrow \frac{\sigma_{00}}{(2c)^{1/2}} U_0 + \frac{\sigma_{00} \sigma_{xx}}{2c} U_0, \\
& \Rightarrow \frac{\sigma_{00}}{(2c)^{1/2}} U_0 \left( D + \frac{\sigma_{00}}{(2c)^{1/2}} U_0 \right), \quad \text{(A.8)}
\end{align*}\]

giving the limit of the numerator.

(ii) From the identity
\[
\chi_t^2 = R_N^N x_t^2 + \mu^2 + u_t^2 + 2 R_N \mu x_t - 2 R_N x_t u_t + 2 \mu u_t,
\]
we have
\[
\sum_{t=1}^{N} R_N x_t^2 = R_N^N \sum_{t=1}^{N} x_t^2 - 2 R_N \mu \sum_{t=1}^{N} x_t - 2 R_N x_t u_t - 2 \mu u_t - 2 R_N \mu x_t + 2 \mu u_t.
\]

We show in the following that each of the following standardized terms
\[
\frac{R_N^N x_t^2}{R_N^{N_N^u}} = \frac{N \mu^2}{R_N^{N_N^u}}, \quad \frac{2 \mu \sum_{t=1}^{N} x_t u_t}{R_N^{N_N^u}}, \quad \frac{N \sum_{t=1}^{N} x_t u_t}{R_N^{N_N^u}}, \quad \frac{R_N^N \sum_{t=1}^{N} x_t u_t}{R_N^{N_N^u}}
\]
are asymptotically negligible. In particular, since the standardized initial condition and drift satisfy \(\tilde{x}_0 = x_0/N^{u/2} \Rightarrow X^* \) and \(\tilde{\mu} = N^{u/2}/\mu \Rightarrow \mu^*\) we find that
\[
\frac{R_N^N x_t^2}{R_N^{N_N^u}} = O_p \left( \left( x_0/N^{u/2} \right)^2 \left( 1/\sqrt{N} \right) \right) = o_p(1),
\]
\[
N \mu^2/R_N^{N_N^u} = O_p \left( (1/N^{u/2}) \right) = o_p(1),
\]
\[
2 \mu \sum_{t=1}^{N} x_t u_t/R_N^{N_N^u} = \left( 1/R_N^{N_N^u} \right) \left( 2 \mu \sqrt{N} \right) \left( 1/\sqrt{N} \right) \sum_{t=1}^{N} u_{xt} \Rightarrow O_p \left( (1/R_N^{N_N^u}) \times O_p(1) \right),
\]
\[
N \sum_{t=1}^{N} x_t u_t/R_N^{N_N^u} = \left( 1/R_N^{N_N^u} \right) \sum_{t=1}^{N} u_{xt} = O_p \left( N^{u/2-3/2} \right) \times O_p(1) = o_p(1),
\]
\[
R_N^N \sum_{t=1}^{N} x_t u_t/R_N^{N_N^u} = O_p \left( \frac{\sum_{t=1}^{N} x_t u_t}{R_N^{N_N^u}} \right) \times O_p \left( (1/R_N^{N_N^u}) \right) = o_p(1),
\]
since \((R_N^N)^{-1}\sum_{t=1}^{N} x_t u_t = O_p(1)\) just as in the analysis of \((R_N^N)^{-1}\sum_{t=1}^{N} x_t u_t\) in part (i); and finally
\[
R_N \mu \sum_{t=1}^{N} x_t/R_N^{N_N^u} = \frac{\mu}{N^{u/2} R_N^{N_N^u}} \sum_{t=1}^{N} x_t R_N^{-1} u_N \Rightarrow o_p \left( (1/N^{u/2}) \right) \times O_p \left( N \right) = o_p(1).
\]

Hence, from (A.9) and (A.2) we deduce that
\[
\frac{(R_N^N)^{-1}}{R_N^{N_N^u}} \sum_{t=1}^{N} x_t^2 \Rightarrow \frac{(X_N/R_N^{N_N^u})^2}{(1 + o_p(1))} \Rightarrow \left( x_0/R_N^{N_N^u} \right)^2 + \frac{1}{(1 + o_p(1))}
\]
\[
= \left( x_0/R_N^{N_N^u} + 1/N^{u/2} \sum_{t=1}^{N} R_N^N u_{0t} + N^{u/2}/c \mu \right)^2 \times \left( 1 + o_p(1) \right)
\]
\[
\Rightarrow \left( \frac{\sigma_{xx}}{(2c)^{1/2}} U_0 + \frac{\sigma_{00} \sigma_{xx}}{2c} U_0 \right)
\]
\[
\Rightarrow \left( \frac{\sigma_{xx}}{(2c)^{1/2}} U_0 + \frac{\sigma_{00} \sigma_{xx}}{(2c)^{1/2}} U_0 \right).
\]

(iii) Combining the results (A.8) and (A.10), we have
\[
\frac{(R_N^N)^{-1}}{R_N^{N_N^u}} \sum_{t=1}^{N} x_t u_{0t} \Rightarrow \frac{(R_N^N)^{-1}}{R_N^{N_N^u}} \sum_{t=1}^{N} x_t u_{0t} \Rightarrow \frac{2c - \sigma_{00} \sigma_{xx}}{\sigma_{xx} U_0 + (2c)^{1/2} D_0}.
\]

Therefore,
\[
R_N^{N_N^u} \left( \hat{A} - A \right) \Rightarrow \left( R_N^N \right)^{-1} \sum_{t=1}^{N} x_t u_{0t} \Rightarrow (2c - \sigma_{00} \sigma_{xx}) U_0 \Rightarrow \sigma_{00} U_0 \Rightarrow \sigma_{xx} U_0 + (2c)^{1/2} D_0.
\]
giving the stated result. ■

A.2. Proof of Corollary 2.2

Proof. The proof follows from Theorem 2.1 by noting the mappings
\[
\sigma_{00} \quad \mapsto \quad \frac{\sigma_{00}}{\sigma_{xx}},
\]
\[
\mu \quad \mapsto \quad \frac{\mu}{\sigma_{xx}} \cdot \frac{h^{1/2}}{\sigma_{xx}},
\]
\[
D_N \quad \mapsto \quad D_{N} \Rightarrow \frac{h^{1/2}}{\sigma_{xx}},
\]
\[
D_N \Rightarrow \frac{h^{1/2}}{\sigma_{xx}},
\]
\[
\sigma_{00} \quad \mapsto \quad \frac{\sigma_{00}}{\sigma_{xx}} \cdot \frac{h^{1/2}}{\sigma_{xx}},
\]
with \(h = 1/N^{u/2}\). It follows that
\[
\frac{\sigma_{00} h^{1/2}}{-2 \kappa} \Rightarrow \frac{\sigma_{00}}{-2 \kappa} U_0 \Rightarrow \left( \frac{\sigma_{00}}{-2 \kappa} U_0 \right) \Rightarrow \frac{\sigma_{00}}{-2 \kappa} U_0 \Rightarrow \left( \frac{\sigma_{00}}{-2 \kappa} U_0 \right).
\]

Please cite this article in press as: Chen Y., et al., Inference in continuous systems with mildly explosive regressors. Journal of Econometrics (2017), http://dx.doi.org/10.1016/j.jeconom.2017.08.016.
and hence
\[
\frac{\sigma_h^2}{\sqrt{h}} (\hat{\beta} - \beta) \Rightarrow (-2\kappa) \frac{\sigma_{00} U_0}{\sigma_{00} U_x + (-2\kappa)^{1/2} D^*} = (-2\kappa) \frac{\sigma_{00} U_0}{\sigma_{00} U_x + (-2\kappa)^{1/2} (x_0 - \mu)}. \]

Appendix B. Supplementary data

Supplementary material related to this article can be found online at http://dx.doi.org/10.1016/j.jeconom.2017.08.016.

References