Random Coefficient Continuous Systems: Testing for Extreme Sample Path Behavior*

Yubo Tao[1], Peter C.B. Phillips[1,2,3,4], and Jun Yu[1]

[1]Singapore Management University
[2]Yale University
[3]University of Auckland
[4]University of York

October 17, 2018

Abstract

This paper studies a continuous time dynamic system with a random persistence parameter. The exact discrete time representation is obtained and related to several discrete time random coefficient models currently in the literature. The model distinguishes various forms of unstable and explosive behavior according to specific regions of the parameter space that open up the potential for testing these forms of extreme behavior. A two-stage approach that employs realized volatility is proposed for the continuous system estimation, asymptotic theory is developed, and test statistics to identify the different forms of extreme sample path behavior are proposed. Simulations show that the proposed estimators work well in empirically realistic settings and that the tests have good size and power properties in discriminating characteristics in the data that differ from typical unit root behavior. The theory is extended to cover models where the random persistence parameter is endogenously determined. An empirical application based on daily real S&P 500 index data over 1928-2018 reveals strong evidence against parameter constancy over the whole sample period leading to a long duration of what the model characterizes as extreme behavior in real stock prices.

JEL Classification: C13, C22, G13.

Keywords: Continuous time models; Explosive path; Extreme behavior; Random coefficient autoregression; Infill asymptotics; Bubble testing.

*We thank the Coeditor, Associate Editor and three referees for helpful comments on the original version. Yubo Tao, School of Economics, Singapore Management University, 90 Stamford Road, Singapore 178903. Email: yubo.tao.2014@phdecons.smu.edu.sg. Peter C.B. Phillips, Cowles Foundation for Research in Economics, Yale University, Box 208281, Yale Station, New Haven, Connecticut 06520-8281. Email: peter.phillips@yale.edu. Jun Yu, School of Economics and Lee Kong Chian School of Business, Singapore Management University, 90 Stamford Road, Singapore 178903. Email: yujun@smu.edu.sg.

1 Introduction

Many macroeconomic and financial time series are well described by autoregressive processes with roots that are close to unity but not necessarily constant over time. Motivated by this empirical characteristic, various strands of the literature have sought to extend pure unit root models to more flexible dynamic systems. One approach allows for structural breaks in which the autoregressive coefficient takes a constant value in each regime but changes value in different regimes (e.g. Chong, 2001; Pang et al., 2014; Jiang et al., 2017). Another assumes that the autoregressive coefficient is a continuous random variable or evolves according to a stochastic process (e.g. Granger and Swanson, 1997; Lieberman and Phillips, 2014, 2017b). Yet another allows for a time varying autoregressive parameter to capture evolution in the stochastic process, introduce flexibility, and enhance forecasting capability (Bykhovskaya and Phillips, 2018, 2019; Giraitis et al., 2014; Kristensen, 2012).

Complementary to this literature on autoregressive specification is a growing interest in modelling explosive behavior and collapse, particularly since the events leading up to and following the global financial crisis, where strong upward movements and subsequent major downturns in asset prices have occurred in various markets (Phillips and Yu, 2011). Empirical methods used to model these events have made extensive use of the concepts of mildly explosive and mildly integrated autoregressive processes (see Phillips and Magdalinos, 2007). Thus, Phillips et al. (2011, PWY hereafter), Phillips and Yu (2011), Phillips et al. (2015a,b, PSY hereafter) assume data are generated according to unit root processes in one regime and as mildly explosive processes in another regime; and methods of date-stamping such regime changes have been developed (Phillips et al., 2011, 2015a,b) stimulating new empirical research and improvements in test methodology (e.g. Cavaliere et al., 2016; Phillips and Shi, 2017). Developments in random autoregressive coefficient approaches have also been pursued, with work by Aue (2008), who analyzed a near-integrated random coefficient autoregressive model, and by Banerjee et al. (2017) who studied a near-explosive random coefficient autoregressive model.[1]

The present paper contributes to this literature by working with a continuous time model in which the parameter that measures persistence is randomized. A novel advantage arising from this formulation is that extreme sample path behavior can be classified into distinct scenarios that represent various forms of instability and explosiveness. These scenarios are distinguished parametrically and corresponding hypotheses are formulated to facilitate empirical testing. Continuous time specification also enables the localizing coefficients that appear

[1] Considerable work has been done on discrete time random coefficient autoregressive models in the literature, including Aue et al. (2006), Berkes et al. (2009), Aue and Horváth (2011), and Horváth and Trapani (2016) among many others.

in mildly integrated and mildly explosive processes to be represented in terms of sampling frequency, which facilitates econometric estimation. These parameters are of great importance empirically because they control distance from martingale and unit root behavior in discrete time models (Banerjee et al. (2017)). This advantage of continuous systems has been used in other recent work by Chen et al. (2017) and Wang and Yu (2016) in developing the discrete time methodology of Phillips and Magdalinos (2007).

Continuous system formulation and high frequency data open up the opportunity to employ methods such as realized volatility in estimating parameters that are identified in the quadratic variation process using in-fill asymptotic methods. The two-stage realized volatility approach employed here naturally accommodates heteroskedasticity in the process and allows for consistent estimation of the parameters in the diffusion function under both stationarity and explosiveness. The approach therefore offers potential for a unified in-fill limit theory of consistent parameter estimation in random coefficient autoregression.

A further well-known feature of continuous system formulations is that the effects of initial conditions are naturally incorporated by in-fill asymptotics (as in Phillips, 1987) without having to specify orders of magnitude or use distant past representations (as in Phillips and Magdalinos, 2009) which involve additional unknown parameters. Moreover, continuous systems readily accommodate endogeneity by allowing for dependence between the random coefficient elements and system shocks. In this respect the present research relates to recent work on generalized random coefficient autoregressive models (Hwang and Basawa, 1998) and localized endogenous stochastic unit root models in (Lieberman and Phillips, 2017a). Initial condition effects appear directly in the asymptotic theory and, as is shown in the paper, the endogeneity parameter can be consistently estimated using realized volatility.

The continuous time model used in the present study is a special case of a financial market model developed in Föllmer and Schweizer (1993) obtained by applying an invariance principle to a discrete time market equilibrium model derived from first principles. In particular, Föllmer and Schweizer (1993) developed a microeconomic model of rational expectations equilibrium for a market that involves both information traders and noise traders. They showed that when the proportions of different types of traders fluctuates randomly the equilibrium outcome is a discrete time model with a random coefficient. The mapping from the theory model implies that noise traders contribute positively to the random coefficient whereas information traders contribute negatively. Correspondingly, the ratio of trader types affects the recurrence or transient properties of the resulting price process, thereby impacting the nature of the resulting price trajectories. Thus, the extent of randomness in the coefficient reflects underlying market composition characteristics and is informative about the respective trader proportions. Econometric estimation of such models can therefore help to

shed light on some of these properties and possibly also the changing nature of the market trader composition.

The remainder of the paper is organized as follows. Section 2 introduces a continuous system with randomized persistence and relates this system to several discrete time models already used in the literature. The multiple forms of behavior induced by this system are described and characterized parametrically. Section 3 proposes a novel two-stage approach to parameter estimation using realized volatility. Asymptotic theory is developed and test statistics for distinguishing different forms of explosive behavior are proposed in Section 4. Section 5 extends the methodology to the case of endogenous persistence. Section 6 gives the results of Monte Carlo simulations that explore the finite sample performance of the estimators and test statistics. Empirical applications of the model are reported in Section 7 using daily real S&P 500 index data from December 1927 to June 2018. Some empirical applications of the extended model using 5-minute real S&P 500 index data over the period from November 1, 1997 to October 31, 2013 are also discussed. Section 8 concludes. Proofs and other technical material are given in the Appendix. Additional simulation and empirical results can be found in an online supplement.

2 The Model

The model used here is a modified version of the Ornstein-Uhlenbeck process

$$dy(t) = y(t)\widetilde{\mu}dt + \sigma dB_{\varepsilon}(t), y(0) = y_0. \tag{2.1}$$

where B_{ε} is a standard Brownian motion and the sign of the drift parameter $\widetilde{\mu}$ determines stationary (< 0), nonstationary ($= 0$), and explosive (> 0) behavior in $y(t)$, the latter corresponding to a discrete time autoregression with a root that exceeds unity and whose variance grows exponentially with t. In (2.1), the drift parameter $\widetilde{\mu}$ is taken as constant, an assumption that may not be well supported by data over extended periods of time.

The model considered in the present paper extends (2.1) by introducing random shocks to the drift component of (2.1) so that

$$dy(t) = y(t)\left[\widetilde{\mu}dt + \widetilde{\sigma}dB_u(t)\right] + \sigma dB_{\varepsilon}(t), y(0) = y_0, \tag{2.2}$$

where $B_u(t)$ and $B_{\varepsilon}(t)$ are both standard Brownian motions, and y_0 is independent of $B_u(t)$ and $B_{\varepsilon}(t)$. When $\widetilde{\sigma}^2 \neq 0$, model (2.2) may be viewed as an Ornstein-Uhlenbeck process with randomized drift or persistence. Initially, we focus on the case of independent noise processes $B_u(t)$ and $B_{\varepsilon}(t)$, and later consider the endogenous case where these processes are dependent.

Model (2.2) is a special case of a general model introduced by Föllmer and Schweizer (1993),

$$dy(t) = y(t)\left[\widetilde{\mu}(t)dt + \widetilde{\sigma}(t)dB_u(t)\right] + \mu(t)dt + \sigma(t)dB_{\varepsilon}(t), y(0) = y_0, \tag{2.3}$$

called an Ornstein-Uhlenbeck process in a random environment. Föllmer and Schweizer (1993) developed a discrete time version of this process in a market equilibrium setting that involved both information traders and noise traders and then derived its continuous-time limit given by the process in (2.3). Persistence in the dynamic model is determined by the relative proportions of the two types of traders, so random proportions lead to a randomized degree of persistence in the solution. Information traders contribute negatively to persistence while noise traders contribute positively.[2]

Föllmer and Schweizer (1993) derived the strong solution of (2.2) which takes the explicit form

$$y(t) = \exp\left(\widetilde{\sigma}B_u(t) + \left(\widetilde{\mu} - \frac{1}{2}\widetilde{\sigma}^2\right)t\right)y(0) + K(t), \tag{2.4}$$

where

$$\begin{aligned} K(t) &= \sigma\int_0^t \exp\left(\widetilde{\sigma}\left(B_u(t) - B_u(s)\right) + \left(\widetilde{\mu} - \frac{1}{2}\widetilde{\sigma}^2\right)(t-s)\right)dB_\varepsilon(s) \\ &\sim \mathcal{MN}\left(0, \sigma^2\int_0^t e^{2\widetilde{\sigma}(B_u(t)-B_u(s))+2\left(\widetilde{\mu}-\frac{1}{2}\widetilde{\sigma}^2\right)(t-s)}ds\right) \end{aligned} \tag{2.5}$$

under independence of B_u and B_ε and with

$$E\left\{K(t)^2\right\} = \sigma^2 E\left\{\int_0^t e^{2\widetilde{\sigma}(B_u(t)-B_u(s))+2\left(\widetilde{\mu}-\frac{1}{2}\widetilde{\sigma}^2\right)(t-s)}ds\right\} = \frac{\sigma^2}{2}\frac{e^{2\left(\widetilde{\mu}+\frac{1}{2}\widetilde{\sigma}^2\right)t} - 1}{\left(\widetilde{\mu}+\frac{1}{2}\widetilde{\sigma}^2\right)}. \tag{2.6}$$

Notably, $E\left\{K(t)^2\right\}$ diverges exponentially when $\widetilde{\mu} + \frac{1}{2}\widetilde{\sigma}^2 > 0$.

The exact discrete time model corresponding to (2.2) at the sampling interval Δ follows directly from the strong solution and has the explicit form

$$\begin{aligned} y_{t\Delta} = &\exp\left\{\left(\widetilde{\mu} - \frac{1}{2}\widetilde{\sigma}^2\right)\Delta + \widetilde{\sigma}\left[B_{u,t\Delta} - B_{u,(t-1)\Delta}\right]\right\}y_{(t-1)\Delta} \\ &+ \sigma\int_{(t-1)\Delta}^{t\Delta}\exp\left\{\left(\widetilde{\mu} - \frac{1}{2}\widetilde{\sigma}^2\right)(t\Delta - s) + \widetilde{\sigma}\left[B_{u,t\Delta} - B_u(s)\right]\right\}dB_\varepsilon(s), \end{aligned} \tag{2.7}$$

where $t = 1, ..., T/\Delta$, T is the time span, and we write discrete time data in subscripted form. This model is a random coefficient autoregression (RCAR) of the type considered by Nicholls and Quinn (1980) in which the autoregressive (AR) coefficient is

$$\rho_{t\Delta} = \exp\left\{\left(\widetilde{\mu} - \frac{1}{2}\widetilde{\sigma}^2\right)\Delta + \widetilde{\sigma}\left[B_{u,t\Delta} - B_{u,(t-1)\Delta}\right]\right\}, \tag{2.8}$$

and is random when $\widetilde{\sigma}^2 > 0$.

[2]Granger (1980) showed how simple cross section aggregation of random coefficient AR processes can – under certain conditions – generate long memory in the aggregated series. In Föllmer and Schweizer (1993) the aggregation is more complex with time varying weights; see Equation (3.6) in Föllmer and Schweizer (1993).

For the ensuing development it will be helpful to fix the following simpler notations for the discrete system

$$\phi := \widetilde{\mu} - \frac{1}{2}\widetilde{\sigma}^2, \quad \kappa := \widetilde{\mu} + \frac{1}{2}\widetilde{\sigma}^2, \quad u_{t\Delta} := \frac{B_{u,t\Delta} - B_{u,(t-1)\Delta}}{\sqrt{\Delta}} \sim \mathcal{N}(0,1),$$

$$\rho_{t\Delta} := \exp\left\{\left(\widetilde{\mu} - \frac{1}{2}\widetilde{\sigma}^2\right)\Delta + \widetilde{\sigma}\left[B_{u,t\Delta} - B_{u,(t-1)\Delta}\right]\right\} = \exp\left\{\phi\Delta + \widetilde{\sigma}\sqrt{\Delta}u_{t\Delta}\right\},$$

$$\eta_{t\Delta} := \int_{(t-1)\Delta}^{t\Delta} \exp\left\{\left(\widetilde{\mu} - \frac{1}{2}\widetilde{\sigma}^2\right)(t\Delta - s) + \widetilde{\sigma}\left[B_{u,t\Delta} - B_{u,s}\right]\right\} dB_{\varepsilon,s} \sim \mathcal{N}\left(0, \gamma_\Delta^2\right),$$

where $\gamma_\Delta = \sqrt{\left(e^{2\kappa\Delta} - 1\right)/2\kappa}$. Model (2.7) is then

$$y_{t\Delta} = \exp\left\{\phi\Delta + \widetilde{\sigma}\sqrt{\Delta}u_{t\Delta}\right\} y_{(t-1)\Delta} + \sigma\eta_{t\Delta} = \rho_{t\Delta}y_{(t-1)\Delta} + \sigma\eta_{t\Delta}, \tag{2.9}$$

where y_t is initiated at y_0.

Importantly, when the driver Wiener processes B_u and B_ε are independent, data generated from (2.2) is observationally equivalent to data from the continuous system

$$dy(t) = y(t)\widetilde{\mu}dt + \sqrt{\widetilde{\sigma}^2 y^2(t) + \sigma^2}dB_v(t), \quad y(0) = y_0, \tag{2.10}$$

where $B_v(t)$ is another standard Brownian motion. In the same way, model (2.7), is observationally equivalent to the (approximate) discrete system

$$y_{t\Delta} = \exp\left\{\widetilde{\mu}\Delta\right\} y_{(t-1)\Delta} + \sqrt{(\widetilde{\sigma}^2 y_{(t-1)\Delta}^2 + \sigma^2)\Delta} \cdot v_{t\Delta}, \tag{2.11}$$

where $v_{t\Delta} \sim \mathcal{N}(0,1)$ and $y_{t\Delta}$ exhibits conditional heteroskedasticity. Notably, by standardization of $1/\sqrt{\Delta}$, the conditional variance of the process is $\widetilde{\sigma}^2 y_{(t-1)\Delta}^2 + \sigma^2$, so that large realizations of the process magnify its variability. This dependence has a substantial bearing on the properties of $y_{t\Delta}$ and the form of its trajectories. Moreover, $y_{t\Delta}$ has a submartingale property when $e^{\widetilde{\mu}\Delta} > 1$ and given $y_{(t-1)\Delta} > 0$ because in that case $E_{(t-1)\Delta}(y_{t\Delta}) = e^{\widetilde{\mu}\Delta}y_{(t-1)\Delta} > y_{(t-1)\Delta}$.

Assuming $\widetilde{\sigma}^2 > 0$, models (2.9) and (2.2) have the following properties: (1) $E(\rho_{t\Delta}) = e^{\widetilde{\mu}\Delta}$, which is unity if and only if $\widetilde{\mu} = 0$ and exceeds unity if and only if $\widetilde{\mu} > 0$; (2) $E(\rho_{t\Delta}^2) = \exp\left(2\widetilde{\mu}\Delta + \widetilde{\sigma}^2\Delta\right) = \exp(2\kappa\Delta)$, which exceeds unity if and only if $\kappa > 0$; (3) $Var(\rho_{t\Delta}) = e^{2\widetilde{\mu}\Delta}\left(e^{\widetilde{\sigma}^2\Delta} - 1\right) > 0$; (4) $E(\rho_{t\Delta}^k) = \exp\left(k\Delta\left[\widetilde{\mu} + \frac{1}{2}(k-1)\widetilde{\sigma}^2\right]\right) \to \infty$ when $k \to \infty$; (5) As shown in Föllmer and Schweizer (1993), when $\phi = \widetilde{\mu} - \frac{1}{2}\widetilde{\sigma}^2 < 0$, the process is asymptotically stationary but may not have finite second moments.[3] To ensure the existence of second moments, we should impose a stronger condition that $\kappa = \widetilde{\mu} + \frac{1}{2}\widetilde{\sigma}^2 < 0$. From (2.5), when $\kappa < 0$, it is apparent that the variance of $K(t)$ exists and converges to $-0.5\sigma^2/\kappa < \infty$ as

[3]When y_0 is fixed, the process is not stationary for finite t but is asymptotically stationary.

$t \to \infty$. It then follows that (2.2) is asymptotically covariance stationary; (6) If $\kappa = 0$, the variance of $K(t)$ equals to $\sigma^2 t$ that diverges as $t \to \infty$, which means (2.2) is not asymptotically covariance stationary. Since $\kappa = 0$ implies $\widetilde{\mu} < 0$ and $\phi < 0$, (2.2) is asymptotic stationarity; (7) If $\phi = \widetilde{\mu} - \frac{1}{2}\widetilde{\sigma}^2 \geq 0$, $y(t)$ is no longer asymptotically stationary as shown in Föllmer and Schweizer (1993).

Table 1: Properties of Proposed Model Under Different Scenarios

Scenario	Asymptotically Stationary	Asym. Covariance Stationary	$E(\rho_{t\Delta})$	$E(\rho_{t\Delta}^2)$
$\widetilde{\mu} + \widetilde{\sigma}^2/2 < 0$	Yes	Yes	< 1	< 1
$\widetilde{\mu} + \widetilde{\sigma}^2/2 = 0$	Yes	No	< 1	$= 1$
$\widetilde{\mu} + \widetilde{\sigma}^2/2 > 0$ & $\widetilde{\mu} < 0$	Yes	No	< 1	> 1
$\widetilde{\mu} = 0$	Yes	No	$= 1$	> 1
$\widetilde{\mu} > 0$ & $\widetilde{\mu} - \widetilde{\sigma}^2/2 < 0$	Yes	No	> 1	> 1
$\widetilde{\mu} - \widetilde{\sigma}^2/2 \geq 0$	No	No	> 1	> 1

Table 1 summarizes the stationarity properties mentioned above and the respective values of $E(\rho_{t\Delta})$, and $E(\rho_{t\Delta}^2)$ under different regions of the parameter space depending on the values of $\widetilde{\mu}$ and $\widetilde{\sigma}^2$. When $\widetilde{\mu} + \widetilde{\sigma}^2/2 < 0$, the model is asymptotically covariance stationary with both $E(\rho_{t\Delta}) < 1$ and $E(\rho_{t\Delta}^2) < 1$. Figure 1(a) plots a simulated time series in this case with $\widetilde{\mu} = -5, \widetilde{\sigma}^2 = 0.5$ and $\widetilde{\mu} + \widetilde{\sigma}^2/2 = -4.75$ where stationary behavioral features of the data are apparent. When $\widetilde{\mu} + \widetilde{\sigma}^2/2 = 0$, the model retains asymptotic stationarity but is no longer covariance stationary with $E(\rho_{t\Delta}) < 1$ and $E(\rho_{t\Delta}^2) = 1$. Figure 1(b) plots a simulated time series in this case with $\widetilde{\mu} = -2, \widetilde{\sigma}^2 = 4$ and $\widetilde{\mu} + \widetilde{\sigma}^2/2 = 0$ where stationarity is again apparent but with more evidence of persistence in the trajectory than in Figure 1(a). It was suggested in Granger and Swanson (1997) that the unit root hypothesis in a STUR random environment might be represented by the expectation $E(\rho_{t\Delta}^2) = 1$. However, the stationary properties of the time series in this case suggest stable and mean recursive trajectories that have greater persistence than when $E(\rho_{t\Delta}^2) < 1$.

When $\widetilde{\mu} + \widetilde{\sigma}^2/2 > 0$ and $\widetilde{\mu} < 0$, the model is asymptotically stationary but is not covariance stationary and $E(\rho_{t\Delta}^2) > 1$. Figure 1(c) plots a simulated time series in this case with $\widetilde{\mu} = -1, \widetilde{\sigma}^2 = 3.5$ and $\widetilde{\mu} + \widetilde{\sigma}^2/2 = 0.75$. Whereas the expectation $E(\rho_{t\Delta})$ is still less than unity, unstable behavior is evident in the simulated time series. In particular, the unstable subperiod of growth and collapse in the trajectory mimics bubble phenomena that are observed in actual data, such as that in Figure 5 in the empirical section of the present paper and in Figure 1 of PWY (2011). If $\widetilde{\mu} = 0$, the model continues to be asymptotically stationary but is not covariance stationary and $E(\rho_{t\Delta}) = 1$, so the model reduces to the

stochastic unit root (STUR) model of Granger and Swanson (1997). Figure 1(d) plots a simulated time series in this case with $\widetilde{\mu} = 0, \widetilde{\sigma}^2 = 2$ and $\widetilde{\mu} + \widetilde{\sigma}^2/2 = 1$. Compared to the traditional (nonstochastic) unit root model, unstable behavior with bubble-like phenomenon in a subperiod of the simulated trajectory is now more evident.

When $\widetilde{\mu} > 0$, $E(\rho_{t\Delta}) > 1$ and $\Pr(\rho_{t\Delta} > 1) > 0.5$, giving greater probability to the realization of an explosive root than a unit or stationary root. However, unlike the traditional (nonstochastic) explosive AR(1) model which is nonstationary, this model is still asymptotically stationary although not covariance stationary. Figure 1(e) plots a simulated time series in this case with $\widetilde{\mu} = 0.5, \widetilde{\sigma}^2 = 2$, $\widetilde{\mu} + \widetilde{\sigma}^2/2 = 1.5$, $\widetilde{\mu} - \widetilde{\sigma}^2/2 = -0.5$. Although the trajectory in Figure 1(e) appears similar to those of Figure 1(c) and Figure 1(d), the process exhibits larger variation, as is apparent from the vertical scale of the figure. When $\widetilde{\mu} - \widetilde{\sigma}^2/2 > 0$, the model is asymptotically nonstationary and both moments $E(\rho_{t\Delta})$ and $E(\rho_{t\Delta}^2)$ exceed unity. Figure 1(f) plots a simulated time series in this case with $\widetilde{\mu} = 1, \widetilde{\sigma}^2 = 0.5$ and $\widetilde{\mu} - \widetilde{\sigma}^2/2 = 1.25$. The explosive growth behavior is clearly evident in the plotted trajectory.

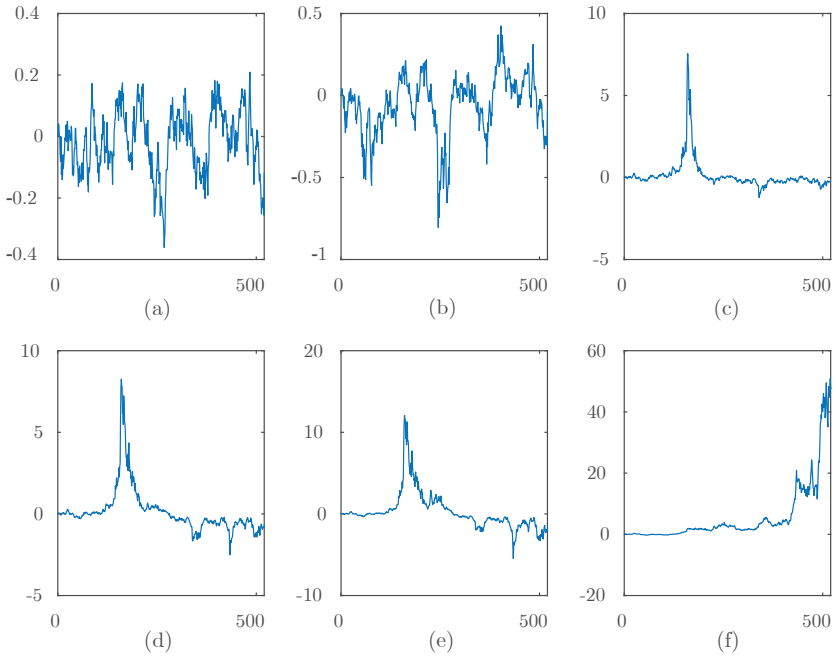


Figure 1: Simulated paths from the proposed model (2.2) when $\widetilde{\mu}$ and $\widetilde{\sigma}^2$ are in different regions.

The exact discrete time representation of our model is closely related to the near explosive random coefficient (NERC) model proposed recently in Banerjee et al. (2017) and to the multivariate local STUR model that is studied in Lieberman and Phillips (2017c) which combines deterministic local unit root (LUR) and random STUR component departures

from unity. In particular, if Δ is chosen as $1/T^{\alpha}$ and $y_0 = 0$, then model (2.9) is the same as model (1) in Banerjee et al. (2017); and if Δ is chosen as $1/T$ and $y_0 = 0$, then our (2.9) has the same form as equation (4) in Lieberman and Phillips (2017c). As discussed in Phillips and Magdalinos (2007), the power rate α in the fraction $1/T^{\alpha}$ controls the degree of mild deviation from a unit root and is typically assumed to lie strictly between zero and unity, which assures that such deviations are localized to unity and exceed the usual local to unity departure of $O\left(T^{-1}\right)$.

In the standard discrete time modeling framework, the localizing rate parameter α is difficult to estimate, although it is possible to do so at a slow rate (Phillips , 2012). Following the argument used in Wang and Yu (2016) with double asymptotics (i.e., both large span and infill schemes), the discrete time model (2.9), or equivalently (2.11), implies mild deviations from a unit root in which the localizing rate is determined by the sampling frequency Δ, and so there is no need to estimate a separate parameter α. This distinction implies an important advantage of the underlying continuous system framework when it is appropriate in practical work to employ this model using discrete time observations. A further useful difference is that the continuous system allows for flexible initial condition assumptions.

The model reduces to a simple autoregression with a time-invariant coefficient when $\widetilde{\sigma}^2 = 0$, in which case $\kappa = \phi = \widetilde{\mu}$ and then explosive behavior applies when $\phi > 0$. Conventional tests for a unit versus an explosive root therefore reduce to testing $\phi = 0$ against $\phi > 0$. This formulation explains the focus on right-tailed unit root testing (Diba and Grossman, 1988), including the recursive methodology used in PWY (2011), Phillips and Yu (2011), PSY (2015a, b) and related work.

In the extended model (2.9), a wider set of dynamic patterns are possible for studying various types of extreme behavior in realized sample trajectories. More specifically, we consider three cases distinguished by the following typology.

1. Unstable trajectory: $\kappa = \widetilde{\mu} + \widetilde{\sigma}^2/2 > 0$ which is equivalent to $E(\rho_{t\Delta}^2) > 1$. In this case, the model is covariance nonstationary asymptotically and is capable of generating trajectories with explosive and collapse behavior;

2. Locally Explosive trajectory: $\widetilde{\mu} > 0$ which is equivalent to $E(\rho_{t\Delta}) > 1$. In this case, there is greater probability for an explosive root to be realized in the sample than a unit or stationary root and the model is covariance nonstationary asymptotically. The model is capable of generating both explosive and collapsing behavior;

3. Explosive trajectory: $\phi = \widetilde{\mu} - \widetilde{\sigma}^2/2 > 0$. Here the model is nonstationary asymptotically and generates explosive behavior.

According to this terminology *explosiveness* implies *local explosiveness* which implies *instability*. We characterize all of these cases as various forms of extreme behavior. Figure 2 shows regions of the parameter space $\left(\widetilde{\mu}, \widetilde{\sigma}^2\right)$ that accord with these classifications of sample behavior.

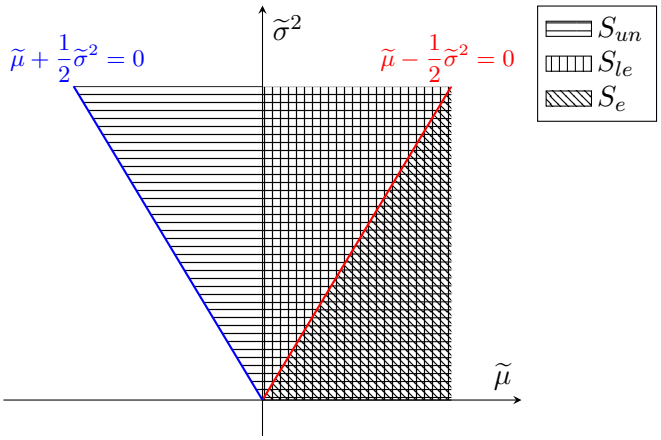


Figure 2: Various explosive regions of $\{y_t\}$ characterized by different parameter combinations of $\left(\widetilde{\mu}, \widetilde{\sigma}^2\right) \in \mathbb{R} \times \mathbb{R}_+$. S_{un} is the region for instability. S_{le} is the region for local explosiveness. S_e is the region for explosiveness.

It is helpful to link the above concepts of instability, local explosiveness and explosiveness to some well-known concepts in the stochastic process literature and to those used recently in Kim and Park (2016). Note first that the observational equivalent model (2.11) is a special case of generalized Höpfner and Kutoyants (GHK) diffusion (Höpfner and Kutoyants, 2003):[4]

$$dX_t = \frac{\widetilde{\mu} X_t}{\left(\widetilde{\sigma}^2 X_t^2 + \sigma^2\right)^{1-d}} dt + \left(\widetilde{\sigma}^2 X_t^2 + \sigma^2\right)^{d/2} dW_t$$

with $d = 1$. In this case, we can easily calculate the scale density $(s'(x))$ and the speed density $(m(x))$ of the model (2.11) as follows:

$$s'(x) = (\sigma^2 + \widetilde{\sigma}^2 x^2)^{-\widetilde{\mu}/\widetilde{\sigma}^2} \quad \text{and} \quad m(x) = (\sigma^2 + \widetilde{\sigma}^2 x^2)^{(\widetilde{\mu}/\widetilde{\sigma}^2 - 1)}. \tag{2.12}$$

Thus, the model (2.11) is recurrent if $\widetilde{\mu}/\widetilde{\sigma}^2 \leq 1/2$, i.e., $\phi \leq 0$. It is positive recurrent (PR) if $\widetilde{\mu}/\widetilde{\sigma}^2 < 1/2$, i.e., $\phi < 0$. Thus, it is null recurrent (NR) when $\phi = 0$ and transient (TR) when $\phi > 0$. Therefore, our definition of explosiveness corresponds to the transient property, which typically applies to processes that trend upwards or downwards and may be rendered recurrent after suitable detrending techniques as discussed by Kim and Park (2016)

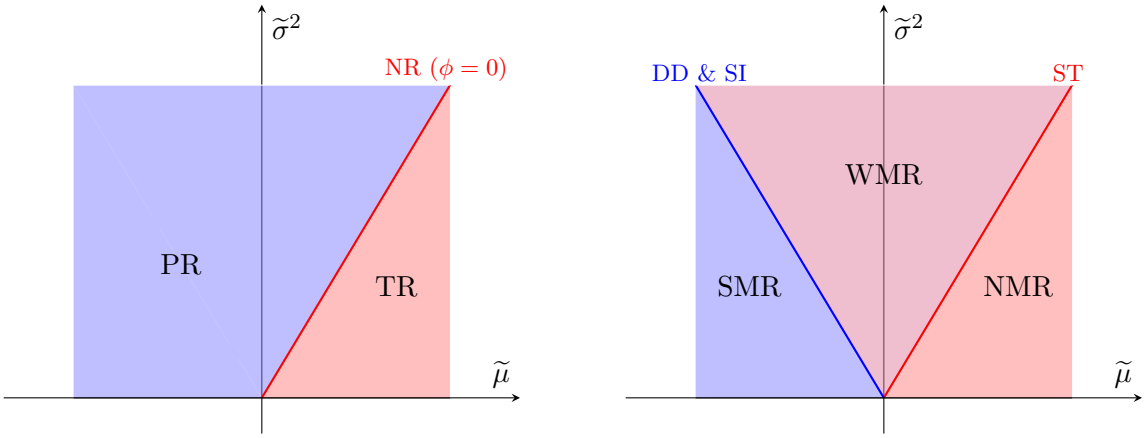
[4]The diffusion process studied here is a generalization of Example 2.1 in Kim and Park (2016) by adding a coefficient in front of X_t^2.

who considered various notions of mean reversion for financial time series. These authors related the mean-reversion property to the following three conditions:

(ST): the speed measure m is either integrable or barely nonintegrable[5];

(DD): The inverse of the scale density $1/s'$ is either integrable or barely nonintegrable;

(SI): square of identity function, ι^2, is either m-integrable[6] or m-barely nonintegrable.



(a) Recurrence Properties (b) Mean Reversion Properties

Figure 3: Subfigures (a) and (b) characterize the recurrence properties and the mean reversion properties of $\{y_t\}$ under different combinations of $(\widetilde{\mu}, \widetilde{\sigma}^2) \in R \times R_+$. PR=positive recurrent, NR=null recurrent, TR=transient; SMR=strong mean reversion, WMR=weak mean reversion, NMR=no mean reversion.

Kim and Park (2016) showed that when both ST and DD hold, the process has strong mean reversion (SMR) and if only one of ST and DD holds the process has weak mean reversion (WMR). By checking these conditions, we find that model (2.11) satisfies: DD if and only if $\widetilde{\mu}/\widetilde{\sigma}^2 \leq -1/2$, i.e., $\kappa \leq 0$; ST if and only if $\widetilde{\mu}/\widetilde{\sigma}^2 \leq 1/2$, i.e., $\phi \leq 0$; and SI if $\widetilde{\mu}/\widetilde{\sigma}^2 \leq -1/2$, i.e., $\kappa \leq 0$. So in our model the condition that ensures ST is the same as that which ensures SI, and is stronger than that which ensures DD. Thus, if $\kappa \leq 0$, our model has strong mean reversion; if $\phi \leq 0$ but $\kappa > 0$, our model has weak mean reversion; and if $\phi > 0$, our model does not imply mean reversion. Hence, our definition of explosiveness is the same as no mean reversion in Kim and Park (2016). Figure 3 summarizes the mean reversion properties of the process, viz., strong mean reversion (SMR), weak mean reversion

[5]A function m is defined to be *barely nonintegrable* if there exists some slowly varying function ℓ such that $m\ell$ is integrable.

[6]The square of the identity function ι^2 is defined by $\iota^2(x) = x^2$; and a function f is defined to be *m-integrable* if mf is integrable.

(WMR), and no mean reversion (NMR) of the diffusion process (2.11) in different regions of the respective parameter spaces.

3 Model Estimation using Realized Volatility

To estimate the continuous-time model (2.2) based on discretely sampled data, we employ the two-stage estimation procedure proposed by Phillips and Yu (2009). In the first stage we make use of the feasible central limit theory for realized volatility to set up a regression model for estimating $\widetilde{\sigma}^2$ and σ^2. In the second stage the in-fill log-likelihood function is maximized to estimate $\widetilde{\mu}$. Consistency and asymptotic distribution theory are established for all estimates.

To explain the estimation method and to establish the large sample theory of the estimators, we assume the time interval $[0, T]$ with span length T can be split into N equispaced blocks. The time span of each block is $h := T/N$ and we assume there are M observations of y_t within each block. So in total $M \times N$ observations on y_t are available over $[0, T]$ and $M \times N = T/\Delta$. Further assume that as $\Delta \to 0$, $M \to \infty$ and $M \times N \to \infty$. Figure 4 illustrates this notation and the sampling scheme. For example, if ten years of daily observa-

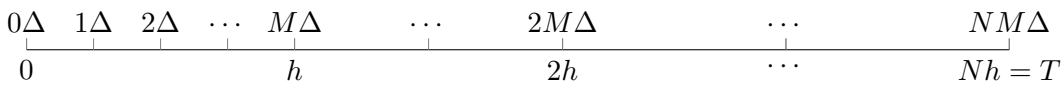


Figure 4: Notational schematic for individual observations, block divisions, and full sample span

tions are available and we split the data into ten blocks, then $T = 10$, $\Delta = 1/252$, $M = 252$, $h = 1$, and $N = 10$. The total number of observations is 2520 and the number of observations contained in each block is 252.

The quadratic variation process $[y]_t$ of $y(t)$ in (2.2) satisfies $d[y]_t = \left(\widetilde{\sigma}^2 y_t^2 + \sigma^2\right) dt$, giving

$$[y]_t = \int_0^t \left(\widetilde{\sigma}^2 y_s^2 + \sigma^2\right) ds. \tag{3.1}$$

Barndorff-Nielsen and Shephard (2002) showed that quadratic variation may be consistently estimated using realized variance (RV) when $\Delta \to 0$. Realized variance and realized quarticity (RQ) are computed using increments $y_{(n-1)h+i\Delta} - y_{(n-1)h+(i-1)\Delta}$ in the observed process by means of the following formulae calculated over the nth block

$$RV_n = \sum_{i=1}^{M} \left[y_{(n-1)h+i\Delta} - y_{(n-1)h+(i-1)\Delta}\right]^2, \quad n = 1, 2, \cdots, N,$$

$$RQ_n = \frac{1}{3\Delta} \sum_{i=1}^{M} \left[y_{(n-1)h+i\Delta} - y_{(n-1)h+(i-1)\Delta}\right]^4, \quad n = 1, 2, \cdots, N.$$

From Barndorff-Nielsen (2002) realized variance has the following asymptotic distribution for large M within each block

$$\sqrt{M}\left([y]_{(n-1)h}^{nh} - RV_n\right) \xrightarrow{L} \mathcal{MN}\left(0, 2h\int_{(n-1)h}^{nh}\left(\widetilde{\sigma}^2 y_s^2 + \sigma^2\right)^2 ds\right), \tag{3.2}$$

where $\mathcal{MN}$ signifies mixed normal and $[y]_{(n-1)h}^{nh} = \int_{(n-1)h}^{nh}\left(\widetilde{\sigma}^2 y_s^2 + \sigma^2\right) ds$.

Following the algorithm of Phillips and Yu (2009), the first-stage estimation step aims to estimate $\boldsymbol{\theta} := (\widetilde{\sigma}^2, \sigma^2)'$ by least squares using the criterion

$$\widehat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} Q_\Delta(\boldsymbol{\theta}), \tag{3.3}$$

where

$$Q_\Delta(\boldsymbol{\theta}) = \Delta\sum_{n=1}^{N}\frac{\left(\log RV_n - \log[y]_{(n-1)h}^{nh} + \frac{1}{2}s_n^2\right)^2}{s_n^2},$$

with

$$s_n = \max\left\{\sqrt{2\Delta\frac{RQ_n}{RV_n^2}}, \sqrt{\frac{2}{M}}\right\},$$

and where Θ is a compact subset of R_+^2 containing the true value $\theta_0 = \left(\widetilde{\sigma}_0^2, \sigma_0^2\right)'$ as an interior point. The term $s_n^2/2$ in the numerator of $Q_\Delta(\boldsymbol{\theta})$ is a finite sample correction on the asymptotic theory. In practice, the quadratic variation element $[y]_{(n-1)h}^{nh}$ in $Q_\Delta(\boldsymbol{\theta})$ can be approximated by Riemann sums as follows

$$[y]_{(n-1)h}^{nh} = \int_{(n-1)h}^{nh}(\widetilde{\sigma}^2 y_s^2 + \sigma^2)ds = \Delta\sum_{t=1}^{M}\left\{\widetilde{\sigma}^2 y_{(n-1)h+t\Delta}^2 + \sigma^2\right\} + O(\sqrt{\Delta}).$$

In the second stage, $\widetilde{\mu}$ is estimated by maximizing the approximate log-likelihood function, viz.,

$$\widehat{\widetilde{\mu}} = \arg\max_{\widetilde{\mu}}\frac{1}{MN}\log\ell_{ALF}(\widetilde{\mu}), \tag{3.4}$$

$$\ell_{ALF}(\widetilde{\mu}) = \sum_{t=1}^{M\times N}\frac{\widetilde{\mu}y_{(t-1)\Delta}}{\widehat{\widetilde{\sigma}^2}y_{(t-1)\Delta}^2 + \widehat{\sigma^2}}\left(y_{t\Delta} - y_{(t-1)\Delta}\right) - \frac{\Delta}{2}\sum_{t=1}^{M\times N}\frac{\widetilde{\mu}^2 y_{(t-1)\Delta}^2}{\widehat{\widetilde{\sigma}^2}y_{(t-1)\Delta}^2 + \widehat{\sigma^2}}., \tag{3.5}$$

giving

$$\widehat{\widetilde{\mu}} = \Delta^{-1}\frac{\sum_{t=1}^{M\times N}\frac{y_{(t-1)\Delta}\left(y_{t\Delta} - y_{(t-1)\Delta}\right)}{\widehat{\widetilde{\sigma}^2}y_{(t-1)\Delta}^2 + \widehat{\sigma^2}}}{\sum_{t=1}^{M\times N}\frac{y_{(t-1)\Delta}^2}{\widehat{\widetilde{\sigma}^2}y_{(t-1)\Delta}^2 + \widehat{\sigma^2}}} =: \Delta^{-1}\frac{\widehat{A}_N}{\widehat{B}_N}. \tag{3.6}$$

This estimator of $\widetilde{\mu}$ has the same form as the weighted least squares estimator used by Hwang and Basawa (2005) in the context of a discrete time RCAR.

4 Asymptotic Theory

This section develops asymptotic theory for the diffusion parameter and drift parameter estimates. Limit theory for estimates of the diffusion parameters is first obtained by using infill asymptotic methods where $\Delta \to 0$. We then investigate the asymptotic theory of the drift parameter by resorting to double asymptotic methods in which both $\Delta \to 0$ and $T \to \infty$.

4.1 Diffusion Parameters

This section derives asymptotic theory for the estimates $\widehat{\widetilde{\sigma}^2}$ and $\widehat{\sigma}^2$ by assuming $\Delta \to 0$ (with a fixed T) in an infill asymptotic scheme. Since $M = h/\Delta$ it follows that $\Delta \to 0$ implies $M \to \infty$. Realized variance and realized quarticity are therefore consistent estimates of integrated variance and integrated quarticity. Since T is fixed, $N = T/h$ is also fixed. Let $\{y_t\}_{t=\Delta}^{MN\Delta}$ be a discrete sample generated from (2.2) where the true parameter values for $\widetilde{\mu}, \widetilde{\sigma}^2, \sigma^2$ are denoted $\widetilde{\mu}_0, \widetilde{\sigma}_0^2, \sigma_0^2$. Assume that $\boldsymbol{\theta_0} = \left(\widetilde{\sigma}_0^2, \sigma_0^2\right)' \in \text{Int}(\boldsymbol{\Theta})$ where $\boldsymbol{\Theta}$ is a compact set in R_+^2. Let $\rho_0 = \exp\left(\widetilde{\mu}_0 \Delta\right) = E\left(\rho_{t\Delta}\right)$, and $\widehat{\rho} = \exp\left(\widehat{\widetilde{\mu}}\Delta\right)$. The following result provides within block infill asymptotics as $\Delta \to 0$.

Theorem 4.1. *If $\boldsymbol{\theta_0} \in Int(\boldsymbol{\Theta})$ and $\Delta \to 0$,*

$$\frac{1}{\sqrt{\Delta}}\left(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta_0}\right) \xrightarrow{L} \left(\sum_{n=1}^{N} \frac{\int_{(n-1)h}^{nh} \frac{\partial \breve{\sigma}^2(y_s;\boldsymbol{\theta_0})}{\partial \boldsymbol{\theta}} \cdot \frac{\partial \breve{\sigma}^2(y_s;\boldsymbol{\theta_0})}{\partial \boldsymbol{\theta}'} ds}{\int_{(n-1)h}^{nh} \breve{\sigma}^4(y_s;\boldsymbol{\theta_0}) ds}\right)^{-1} \left(\sum_{n=1}^{N} \frac{\sqrt{2}\int_{(n-1)h}^{nh} \frac{\partial \breve{\sigma}^2(y_s;\boldsymbol{\theta_0})}{\partial \boldsymbol{\theta}} \breve{\sigma}^2(y_s;\boldsymbol{\theta_0}) dB_s}{\int_{(n-1)h}^{nh} \breve{\sigma}^4(y_s;\boldsymbol{\theta_0}) ds}\right),$$

where $\breve{\sigma}^2(y_t;\boldsymbol{\theta_0}) = \widetilde{\sigma}_0^2 y_t^2 + \sigma_0^2$ is the spot variance of $y(t)$.

Remark 4.1. *In discrete time modeling, it is common for the parameters $\widetilde{\sigma}^2$ and σ^2 to be estimated by MLE or QMLE by imposing ARCH-type innovations, see for example Jensen and Rahbek (2004); Ling and Li (2008); Francq and Zakoïan (2012); Chen et al. (2014). This approach provides consistent estimates and associated asymptotics for $\widetilde{\sigma}^2$ rather than σ^2 when y_t is nonstationary. The explanation is that as $T \to \infty$, the log-likelihood function becomes flat because of the dominating scale effects of y_T that occur in the direction where $\widetilde{\sigma}^2$ is fixed and σ^2 varies. Unlike previous work, our approach applies an infill asymptotic scheme which fixes the time span (T) and shrinks the sampling interval (Δ) to 0. These asymptotics ensure that y_T is measurable and finite, so that σ^2 continues to play a role in the limit as $\Delta \to 0$. With this approach it is possible to consistently estimate both variance parameters and establish their asymptotic properties as in Theorem 4.1.*

Remark 4.2. *As found in Phillips and Yu (2009), there is a trade-off between large N and small N. When N is too small, the variance of the estimator is large. However, when N is sufficiently large, further increases in N fail to improve the variance. The choice of the optimal N is therefore of interest but is beyond the scope of the present paper.*

It is interesting in practical applications to test the null hypothesis $\widetilde{\sigma}^2 = 0$, which corresponds to the special case of no randomness in the persistence properties of $y(t)$. To test this boundary condition hypothesis we apply a modified version of the locally best invariant test (LBI-test) by Lee (1998) for $\widetilde{\sigma}^2 = 0$, viz.,

$$\widetilde{Z}_N := \frac{\sum_{t=1}^{M\times N}\left(\widetilde{\varepsilon}_{t\Delta}^2 - \left(\frac{1}{MN}\sum_{t=1}^{M\times N}\widetilde{\varepsilon}_{t\Delta}^2\right)\right)\widetilde{y}_{(t-1)\Delta}^2}{\sqrt{\frac{1}{MN}\sum_{t=1}^{M\times N}\widetilde{\varepsilon}_{t\Delta}^4 - \left(\frac{1}{MN}\sum_{t=1}^{M\times N}\widetilde{\varepsilon}_{t\Delta}^2\right)^2}\sqrt{\frac{1}{MN}\sum_{t=1}^{M\times N}\widetilde{y}_{(t-1)\Delta}^4 - \left(\frac{1}{MN}\sum_{t=1}^{M\times N}\widetilde{y}_{(t-1)\Delta}^2\right)^2}}$$

where $\widetilde{y}_{t\Delta} = \dfrac{y_{t\Delta}}{\sqrt{1+y_{t\Delta}^2}}$, $\widetilde{\varepsilon}_{t\Delta} = y_{t\Delta} - \widetilde{\rho}y_{(t-1)\Delta}$ and $\widetilde{\rho} = \left(\sum_{t=1}^{M\times N}\widetilde{y}_{(t-1)\Delta}y_{(t-1)\Delta}\right)^{-1}\sum_{i=1}^{M\times N}\widetilde{y}_{(t-1)\Delta}y_{t\Delta}$.
Then, as $N \to \infty$,

$$(MN)^{-1/2}\widetilde{Z}_N \xrightarrow{L} \mathcal{N}(0,1), \text{ under } H_0 : \widetilde{\sigma}^2 = 0,$$

and

$$|(MN)^{-1/2}\widetilde{Z}_N| \xrightarrow{p} \infty, \quad \text{under } H_1 : \widetilde{\sigma}^2 > 0.$$

Remark 4.3. *Note first that we use the self-normalized variable $\widetilde{y}_{t\Delta}$ for constructing the test statistic. This is because the normalization ensures that $\widetilde{y}_{t\Delta}$ is stationary when $y_{t\Delta}$ is nonstationary, which is crucial for $\widetilde{Z}_N$ to converge under the null hypothesis (Lee, 1998; Nagakura, 2009). In fact, the weighting function $1+y_{t\Delta}^2$ can be replaced by any function $g(x)$ where $g : [0,\infty) \to (0,\infty)$ is a Borel function satisfying $x^2/g(x) \to 1$ as $|x| \to \infty$. In practice, we follow the usual convention by setting the weighting function to be $1+y_{t\Delta}^2$ as in Hill and Peng (2014) and Horváth and Trapani (2016).*

Remark 4.4. *A second point worthy of note concerns the use of the IV estimate $\widetilde{\rho}$ here. Following Chan et al. (2012), the IV estimate*

$$\widetilde{\rho} = \sum_{t=1}^{M\times N}\frac{y_{(t-1)\Delta}y_{t\Delta}}{\sqrt{\delta + y_{(t-1)\Delta}^2}}\Bigg/\sum_{t=1}^{M\times N}\frac{y_{(t-1)\Delta}^2}{\sqrt{\delta + y_{(t-1)\Delta}^2}}$$

is uniformly asymptotically normally distributed for both stationary and nonstationary $y_{t\Delta}$. Further, the IV estimate $\widetilde{\rho}$ includes the Cauchy estimator (So and Shin, 1999) as a special case ($\delta = 0$), which is known to be asymptotically median-unbiased. This helps improve the finite sample performance of the test statistic which depends explicitly on the residuals.

Remark 4.5. *The above test for coefficient constancy remains valid in the presence of correlation between the random coefficient and innovations. When the random coefficients are endogenous the quadratic covariation $\langle B_u, B_\varepsilon\rangle_t = \int_0^t \gamma_s ds$ and the conditional variance of $\widetilde{\varepsilon}_{t\Delta}$ under the null is $Var(\widetilde{\varepsilon}_{t\Delta}|y_{(t-1)\Delta}) = \sigma^2$, whereas under the alternative $Var(\widetilde{\varepsilon}_{t\Delta}|y_{(t-1)\Delta}) =$*

$\widetilde{\sigma}^2 y^2_{(t-1)\Delta} + 2\gamma_t \widetilde{\sigma}\sigma y_{(t-1)\Delta} + \sigma^2$. *The test may therefore be interpreted as examining evidence for the presence of a relationship between* $\widetilde{\varepsilon}^2_{t\Delta}$ *and* $y^2_{(t-1)\Delta}$ *and* $y_{(t-1)\Delta}$ *– in other words, a test for conditional heteroskedasticity.*

4.2 Drift Parameter

This section first derives limit theory for the estimated drift parameter under the double asymptotics scheme with $\Delta \to 0$ and $T \to \infty$. In this framework both $M \to \infty$ and $N \to \infty$. Then we derive limit theory for the expectation of the random autoregressive coefficient $\rho_{t\Delta}$ by requiring only that $\Delta \to 0$. Based on the limit theory for these estimated parameters we construct test statistics for assessing the nature of extreme sample path behavior in data.

Theorem 4.2. *In model (2.2), assume* $\widetilde{\sigma}_0^2 > 0$. *When* $T \to \infty$ *and* $\Delta \to 0$, $\widetilde{\mu} \xrightarrow{p} \widetilde{\mu}_0$. *Additionally, if* $T\Delta \to 0$, *the asymptotic distribution of* $\widehat{\widetilde{\mu}}$ *is given by*

$$\sqrt{T}\left(\widehat{\widetilde{\mu}} - \widetilde{\mu}_0\right) \xrightarrow{L} \mathcal{N}\left(0, V_1^{-1}\right),$$

where

$$V_1 = \begin{cases} E\left(\dfrac{y_t^2}{\widetilde{\sigma}_0^2 y_t^2 + \sigma_0^2}\right), & \text{if } \kappa = \widetilde{\mu}_0 + \dfrac{1}{2}\widetilde{\sigma}_0^2 < 0; \\ \widetilde{\sigma}_0^{-2}, & \text{if } \kappa = \widetilde{\mu}_0 + \dfrac{1}{2}\widetilde{\sigma}_0^2 \geq 0. \end{cases} . \tag{4.1}$$

Remark 4.6. *The asymptotics (4.2) hold regardless of the value of* $\widetilde{\mu}_0 + \widetilde{\sigma}_0^2/2$, *which may be less than zero, equal zero, or greater than zero. By contrast, it is well-known that in the case of the pure AR(1) model, the asymptotic theory for the least squares estimator of the autoregressive coefficient depends critically on the true value of the coefficient. However, in the RCAR model asymptotic normality may hold in both the stationary and explosive cases under certain conditions, as discussed in Hwang and Basawa (2005). The above result reinforces this finding and extends applicability to the continuous-time random coefficient model examined here.*

The asymptotic theory given in (4.2) suggests that consistent estimation of $\widetilde{\mu}$ requires $T \to \infty$. In practical work, however, the time span is often short making large span asymptotics less relevant. The following theorem provides infill asymptotics for estimating $\rho = \exp\{\widetilde{\mu}\Delta\}$, which is useful for testing nonstationarity in a finite time span setting.

Theorem 4.3. *In model (2.2), assume* $\widetilde{\sigma}_0^2 > 0$. *When* T *is fixed and* $\Delta \to 0$, $\widehat{\rho} \xrightarrow{p} \rho_0 = 1$ *and the asymptotic distribution of* $\widehat{\rho}$ *is given by*

$$\frac{1}{\Delta}\left(\widehat{\rho} - 1\right) \xrightarrow{L} \mathcal{N}\left(0, (TV_1)^{-1}\right). \tag{4.2}$$

where

$$\widehat{\rho} = 1 + \frac{\widehat{A}_N}{\widehat{B}_N}, \tag{4.3}$$

and V_1 is given in (4.1).

Remark 4.7. *Although the above result does not deliver a consistent estimate of $\widetilde{\mu}$ with a finite T, the asymptotic theory in (4.2) shows that consistent estimation of ρ is possible when $\Delta \to 0$. This result motivates estimation of $\beta_\kappa := \exp\left\{\left(\widetilde{\mu} + \widetilde{\sigma}^2/2\right)\Delta\right\}$ and $\beta_\phi := \exp\left\{\left(\widetilde{\mu} - \widetilde{\sigma}^2/2\right)\Delta\right\}$ instead of the continuous time parameters κ and ϕ when the time span of the data is short.*

Proposition 4.1. *For model (2.2) with T fixed and $\widetilde{\sigma}_0^2 > 0$, as $\Delta \to 0$,*

$$\frac{1}{\Delta}\left(\widehat{\beta}_\kappa - \beta_\kappa^0\right) \xrightarrow{L} \mathcal{N}\left(0, (TV_1)^{-1}\right), \quad \frac{1}{\Delta}\left(\widehat{\beta}_\phi - \beta_\phi^0\right) \xrightarrow{L} \mathcal{N}\left(0, (TV_1)^{-1}\right).$$

where

$$\widehat{\beta}_\kappa = \exp\left\{\left(\widehat{\widetilde{\mu}} + \frac{1}{2}\widehat{\widetilde{\sigma}^2}\right)\Delta\right\} \text{ and } \widehat{\beta}_\phi = \exp\left\{\left(\widehat{\widetilde{\mu}} - \frac{1}{2}\widehat{\widetilde{\sigma}^2}\right)\Delta\right\}, \tag{4.4}$$

$$\beta_\kappa^0 = \exp\left\{\left(\widetilde{\mu}_0 + \widetilde{\sigma}_0^2/2\right)\Delta\right\} \text{ and } \beta_\phi^0 = \exp\left\{\left(\widetilde{\mu}_0 - \widetilde{\sigma}_0^2/2\right)\Delta\right\}.$$

Remark 4.8. *To test different forms of unstable/explosive behavior, we need to test whether $\kappa = \widetilde{\mu} + \widetilde{\sigma}^2/2 = 0$, or $\widetilde{\mu} = 0$, or $\phi = \widetilde{\mu} - \widetilde{\sigma}^2/2 = 0$. Testing these restrictions corresponds to testing the hypotheses $\beta_\kappa = 1$, or $\rho = 1$, or $\beta_\phi = 1$. In the spirit of Theorem 4.3 and Proposition 4.1, as $\Delta \to 0$, we can construct the following test statistics and derive their asymptotic distributions as detailed below:*

$$t_\kappa = \left(\frac{1}{\Delta}\sum_{t=1}^{M\times N} \frac{y_{(t-1)\Delta}^2}{\widehat{\widetilde{\sigma}^2} y_{(t-1)\Delta}^2 + \widehat{\sigma^2}}\right)^{1/2} \left(\widehat{\beta}_\kappa - \beta_\kappa^0\right) \xrightarrow{L} \mathcal{N}(0,1), \tag{4.5}$$

$$t_{\widetilde{\mu}} = \left(\frac{1}{\Delta}\sum_{t=1}^{M\times N} \frac{y_{(t-1)\Delta}^2}{\widehat{\widetilde{\sigma}^2} y_{(t-1)\Delta}^2 + \widehat{\sigma^2}}\right)^{1/2} (\widehat{\rho} - \rho_0) \xrightarrow{L} \mathcal{N}(0,1), \tag{4.6}$$

$$t_\phi = \left(\frac{1}{\Delta}\sum_{t=1}^{M\times N} \frac{y_{(t-1)\Delta}^2}{\widehat{\widetilde{\sigma}^2} y_{(t-1)\Delta}^2 + \widehat{\sigma^2}}\right)^{1/2} \left(\widehat{\beta}_\phi - \beta_\phi^0\right) \xrightarrow{L} \mathcal{N}(0,1). \tag{4.7}$$

These three t-test statistics can be calculated sequentially and compared with the right-tailed critical value of the asymptotic distributions, giving a real-time testing strategy of empirical evidence of instability/explosiveness in the data. Accordingly, the origination and termination dates of different types of extreme behavior may be estimated in the same fashion

as Phillips et al. (2015a). More specifically, date estimates can be determined from first crossing times as follows

$$\hat{r}_{un}^{ie} = \inf_{s\in[\hat{r}_{un}^{(i-1)f},1]} \{s : t_\kappa(s) > Z_{0.95}\} \quad \text{and} \quad \hat{r}_{un}^{if} = \inf_{s\in[\hat{r}_{un}^{ie},1]} \{s : t_\kappa(s) < Z_{0.95}\},$$
$$\hat{r}_{le}^{ie} = \inf_{s\in[\hat{r}_{le}^{(i-1)f},1]} \left\{s : t_{\widetilde{\mu}}(s) > Z_{0.95}\right\} \quad \text{and} \quad \hat{r}_{le}^{if} = \inf_{s\in[\hat{r}_{le}^{ie},1]} \left\{s : t_{\widetilde{\mu}}(s) < Z_{0.95}\right\},$$
$$\hat{r}_{e}^{ie} = \inf_{s\in[\hat{r}_{e}^{(i-1)f},1]} \{s : t_\phi(s) > Z_{0.95}\} \quad \text{and} \quad \hat{r}_{e}^{if} = \inf_{s\in[\hat{r}_{e}^{ie},1]} \{s : t_\phi(s) < Z_{0.95}\},$$

where: $Z_{0.95} = 1.645$ is the 95% critical value of the standard normal distribution; $\hat{r}_{un}^{ie}/\hat{r}_{le}^{ie}/\hat{r}_{e}^{ie}$ represent estimates of the origination date of the ith explosive period; and $\hat{r}_{un}^{if}/\hat{r}_{le}^{if}/\hat{r}_{e}^{if}$ represent estimates of the termination date of the ith explosive period. To identify the first unstable/explosive period in the sample, a minimum window is needed to start the recursion. The time-stamping strategy used here is based on the standard normal distribution whereas the PWY and PSY algorithms rely on non-standard unit root and sup unit root distributions.

5 The Model with Endogeneity

This section extends the base model (2.2) by allowing for endogeneity, quantified by the correlation between the random coefficient and the equation innovation. In the discrete time literature Hwang and Basawa (1997, 1998) described this framework as a generalized random coefficient autoregressive model. With stationarity imposed they studied the local asymptotic normality of the maximum likelihood estimator and the weighted least squares estimator of the autoregressive coefficient. Zhao and Wang (2012) considered empirical likelihood estimation of the stationary model and proposed a likelihood ratio test for testing stationary/ergodicity. Lieberman and Phillips (2017b) studied the effects of endogeneity in a multivariate context and derived the asymptotic distribution for the non-linear least squares (NLLS) estimator for the autoregressive coefficient, showing that NLLS is inconsistent for the autoregressive coefficient under endogeneity. To address inconsistency of NLLS, Lieberman and Phillips (2017a) proposed a non-linear instrumental variable technique and a GMM approach, establishing consistency and deriving the asymptotic distribution for the IV estimator of the autoregressive coefficient.

To incorporate endogeneity in a continuous time random coefficient setting, we rewrite the model (2.2) as the following continuous time system

$$\begin{aligned} dy(t) &= y(t)d\widetilde{Z}(t) + dZ(t), \quad y(0) = y_0, \\ d\widetilde{Z}(t) &= \widetilde{\mu}dt + \widetilde{\sigma}dB_u(t), \\ dZ(t) &= \sigma dB_\varepsilon(t), \end{aligned} \tag{5.1}$$

where (B_u, B_ε) is a two-dimensional Brownian motion with covariance parameter γ so that the quadratic covariation process satisfies $d\langle B_u, B_\varepsilon\rangle_t = \gamma dt$. Then, $d\langle \widetilde{Z}, Z\rangle_t = \gamma\widetilde{\sigma}\sigma dt := \omega dt$, where $\omega = \gamma\widetilde{\sigma}\sigma$ is the covariance parameter of $(\widetilde{Z}, Z)$. According to Föllmer et al. (1994), the strong solution to this continuous system is

$$y(t) = \exp\left(\left(\widetilde{\mu} - \frac{1}{2}\widetilde{\sigma}^2\right)t + \widetilde{\sigma}B_u(t)\right)y(0) + J(t), \tag{5.2}$$

where

$$\begin{aligned} J(t) &= \sigma\int_0^t \exp\left\{\left(\widetilde{\mu} - \frac{1}{2}\widetilde{\sigma}^2\right)(t-s) + \widetilde{\sigma}\left(B_u(t) - B_u(s)\right)\right\}dB_\varepsilon(s) \\ &\quad - \omega\int_0^t \exp\left\{\left(\widetilde{\mu} - \frac{1}{2}\widetilde{\sigma}^2\right)(t-s) + \widetilde{\sigma}\left(B_u(t) - B_u(s)\right)\right\}ds \\ &= K(t) - L(t). \end{aligned}$$

Compared to the model without endogeneity in (2.4), the dynamics of the process are now driven by $J(t)$ instead of $K(t)$. $J(t)$ has two components, one being $K(t)$ and the other depending on the covariance of the random coefficient and the innovation, ω. The model specified in the system (5.1) is the continuous time limit of the endogenous stochastic unit root (STUR) model of Lieberman and Phillips (2017b) and ω corresponds to the one-sided long-run covariance in the STUR model.

The following proposition shows that the given characterization of instability/explosiveness in the model without endogeneity remains valid for the model with endogeneity.

Proposition 5.1. *The sample path characteristics of the process (5.2) may be classified into the following three types,*

1. *unstable:* $\kappa = \widetilde{\mu} + \frac{1}{2}\widetilde{\sigma}^2 > 0$;
2. *locally explosive:* $\widetilde{\mu} > 0$;
3. *explosive:* $\phi = \widetilde{\mu} - \frac{1}{2}\widetilde{\sigma}^2 > 0$.

The fact that sample path characteristics of (5.2) are unaffected by endogeneity may be explained intuitively by noting that the model (5.1) is observationally equivalent to the following continuous system

$$dy_t = \widetilde{\mu}y_t dt + \sqrt{\widetilde{\sigma}^2 y_t^2 + 2\omega y_t + \sigma^2}dB_v(t), \tag{5.3}$$

where $B_v(t)$ is another standard Brownian motion in an expanded probability space. Note that when the variance of y_t goes to infinity as t increases, the dominant term in the diffusion function $\widetilde{\sigma}^2 y_t^2 + 2\omega y_t + \sigma^2$ is $\widetilde{\sigma}^2 y_t^2$, which explains why $\widetilde{\sigma}^2$ is the key parameter in determining long-run volatility.

Remark 5.1. *From the perspective of diffusion process asymptotics, the recurrence and mean reversion characterizations given in Figure 3 also remain valid. This robustness is evident by checking the limit of the scale index function:*

$$p = \lim_{y\to\infty} v(y) = \lim_{y\to\infty} \frac{-2\widetilde{\mu}y^2}{\sigma^2 + 2\omega y + \widetilde{\sigma}^2 y^2} = -\frac{2\widetilde{\mu}}{\widetilde{\sigma}^2},$$

which is apparently unaffected by endogeneity in the limit.

We can rewrite the discrete time model in AR(1) format as

$$y_{t\Delta} = \exp\left(\left(\widetilde{\mu} - \frac{1}{2}\widetilde{\sigma}^2\right)\Delta + \widetilde{\sigma}\sqrt{\Delta}u_t\right) y_{(t-1)\Delta} + J_\Delta(t) = \rho_{t\Delta} y_{(t-1)\Delta} + J_{t\Delta},$$

where $u_t \overset{i.i.d}{\sim} \mathcal{N}(0,1)$, and

$$\begin{aligned} J_{t\Delta} = \sigma \int_{(t-1)\Delta}^{t\Delta} \exp\left\{\left(\widetilde{\mu} - \frac{1}{2}\widetilde{\sigma}^2\right)(t\Delta - s) + \widetilde{\sigma}\left(B_{u,t\Delta} - B_{u,s}\right)\right\} dB_{\varepsilon,s} \\ - \omega \int_{(t-1)\Delta}^{t\Delta} \exp\left\{\left(\widetilde{\mu} - \frac{1}{2}\widetilde{\sigma}^2\right)(t\Delta - s) + \widetilde{\sigma}\left(B_{u,t\Delta} - B_{u,s}\right)\right\} ds. \end{aligned}$$

From earlier derivations we know that

$$E\left(J_{t\Delta}\right) = \frac{\omega}{\widetilde{\mu}}(1 - \exp(\widetilde{\mu}\Delta)) = -\omega\Delta + O(\Delta^2), \tag{5.4}$$

$$Var(J_{t\Delta}) = O(\Delta). \tag{5.5}$$

Therefore, when standardizing the model by the factor $1/\sqrt{\Delta}$, the expectation of the correspondingly standardized error process $J_{t\Delta}/\sqrt{\Delta}$ in (5.4) has order $O(\sqrt{\Delta})$ as $\Delta \to 0$. This means that under infill asymptotics we can consistently estimate the expectation of the random coefficient, $\rho_0 = E\rho_{t\Delta} = \exp\left(\widetilde{\mu}\Delta\right)$. This result is naturally achieved in the continuous time setup with infill asymptotics and contrasts with inconsistency of least squares estimation in discrete time models with endogeneity (Lieberman and Phillips, 2017b).

As before, we continue to apply the two stage estimation procedure of Phillips and Yu (2009) to estimate the model under endogeneity. Note that the quadratic variation of y_t now satisfies

$$d[y]_t = (\widetilde{\sigma}^2 y_t^2 + 2\omega y_t + \sigma^2)dt. \tag{5.6}$$

In light of the argument of Remark 4.1 we cannot consistently estimate ω and σ^2 in explosive cases under long-span sampling because the signal of y_t^2 is so strong that it drowns information in the linear and constant terms (i.e., $2\omega y_t$ and σ^2). However, infill asymptotics for $\widehat{\boldsymbol{\theta}^*} := \left(\widehat{\widetilde{\sigma}^2}, \widehat{\gamma}, \widehat{\sigma^2}\right)'$ can be developed in the same way as before and the results are summarized in the following theorem.

Theorem 5.1. *Assume $\boldsymbol{\theta}_{\mathbf{0}}^{*} \in Int(\boldsymbol{\Theta}^{*})$ where $\boldsymbol{\Theta}^{*}$ is a compact set in $R_{+} \times [-1,1] \times R_{+}$. As T is fixed and $\Delta \to 0$, we have*

$$\frac{1}{\sqrt{\Delta}}\left(\widehat{\boldsymbol{\theta}^{*}} - \boldsymbol{\theta}_{\mathbf{0}}^{*}\right) \xrightarrow{L} \left(\sum_{n=1}^{N} \frac{\int_{(n-1)h}^{nh} \frac{\partial \breve{\sigma}^{2}(y_s;\boldsymbol{\theta}_{\mathbf{0}}^{*})}{\partial \boldsymbol{\theta}^{*}} \cdot \frac{\partial \breve{\sigma}^{2}(y_s;\boldsymbol{\theta}_{\mathbf{0}}^{*})}{\partial \boldsymbol{\theta}^{*\prime}} ds}{\int_{(n-1)h}^{nh} \breve{\sigma}^{4}(y_s;\boldsymbol{\theta}_{\mathbf{0}}^{*}) ds}\right)^{-1} \left(\sum_{n=1}^{N} \frac{\sqrt{2}\int_{(n-1)h}^{nh} \frac{\partial \breve{\sigma}^{2}(y_s;\boldsymbol{\theta}_{\mathbf{0}}^{*})}{\partial \boldsymbol{\theta}^{*}} \breve{\sigma}^{2}(y_s;\boldsymbol{\theta}_{\mathbf{0}}^{*}) dB_s}{\int_{(n-1)h}^{nh} \breve{\sigma}^{4}(y_s;\boldsymbol{\theta}_{\mathbf{0}}^{*}) ds}\right),$$

where $\breve{\sigma}^{2}(y_t;\boldsymbol{\theta}_{\mathbf{0}}^{*}) = \widetilde{\sigma}_{0}^{2} y_t^{2} + 2\omega_0 y_t + \sigma_0^{2}$ is the spot variance of $y(t)$.

Remark 5.2. *In principle at least, this limit theory enables us to construct a test for endogeneity based on the asymptotic distribution of $\widehat{\gamma}$. However, the limit theory above is hard to implement as this distribution is non-standard and non-pivotal and $\widehat{\gamma}$ is biased when the frequency is low. Instead, to test the most relevant hypothesis of interest $H_0 : \gamma_0 = 0$ we propose the likelihood ratio test based on the objective function $Q_{\Delta}(\boldsymbol{\theta}^{*})$:*

$$LR = \Delta^{-1}\left(Q_{\Delta}^{r} - Q_{\Delta}^{ur}\right) \sim \chi^{2}(1), \text{ under } H_0 : \gamma_0 = 0. \tag{5.7}$$

For consistent estimation of $\widetilde{\mu}$, as in the base model, we maximize the following approximated likelihood

$$\ell_{ALF}(\widetilde{\mu}) = \sum_{t=1}^{M\times N} \frac{\widetilde{\mu} y_{(t-1)\Delta}\left(y_{t\Delta} - y_{(t-1)\Delta}\right)}{\widehat{\widetilde{\sigma}^{2}} y_{(t-1)\Delta}^{2} + 2\widehat{\omega} y_{(t-1)\Delta} + \widehat{\sigma^{2}}} - \frac{\Delta}{2} \sum_{t=1}^{M\times N} \frac{\widetilde{\mu}^{2} y_{(t-1)\Delta}^{2}}{\widehat{\widetilde{\sigma}^{2}} y_{(t-1)\Delta}^{2} + 2\widehat{\omega} y_{(t-1)\Delta} + \widehat{\sigma^{2}}}, \tag{5.8}$$

where $\widehat{\omega} = \widehat{\gamma}\sqrt{\widehat{\widetilde{\sigma}^{2}}\widehat{\sigma^{2}}}$, which gives

$$\widehat{\widetilde{\mu}} = \Delta^{-1} \frac{\sum_{t=1}^{M\times N} \frac{y_{(t-1)\Delta}\left(y_{t\Delta} - y_{(t-1)\Delta}\right)}{\widehat{\widetilde{\sigma}^{2}} y_{(t-1)\Delta}^{2} + 2\widehat{\omega} y_{(t-1)\Delta} + \widehat{\sigma^{2}}}}{\sum_{t=1}^{M\times N} \frac{y_{(t-1)\Delta}^{2}}{\widehat{\widetilde{\sigma}^{2}} y_{(t-1)\Delta}^{2} + 2\widehat{\omega} y_{(t-1)\Delta} + \widehat{\sigma^{2}}}} =: \Delta^{-1} \frac{\widehat{A}_{N}^{*}}{\widehat{B}_{N}^{*}}. \tag{5.9}$$

The following theorem provides asymptotic theory for $\widehat{\widetilde{\mu}}$ and $\widehat{\rho} := \exp\left(\widehat{\widetilde{\mu}}\Delta\right)$.

Theorem 5.2. *In model (5.1) assume $\widetilde{\sigma}_{0}^{2} > 0$. When $T \to \infty$ and $\Delta \to 0$, we have $\widetilde{\mu} \xrightarrow{p} \widetilde{\mu}_0$. Additionally, if $T\Delta \to 0$, we have,*

$$\sqrt{T}\left(\widehat{\widetilde{\mu}} - \widetilde{\mu}_0\right) \xrightarrow{L} \mathcal{N}\left(0, V_2^{-1}\right), \tag{5.10}$$

where

$$V_2 = \begin{cases} E\left(\dfrac{y_t^{2}}{\widetilde{\sigma}_{0}^{2} y_t^{2} + 2\omega_0 y_t + \sigma_0^{2}}\right) & \text{if } \kappa = \widetilde{\mu} + \dfrac{1}{2}\widetilde{\sigma}^{2} < 0 \\ \widetilde{\sigma}_{0}^{-2} & \text{if } \kappa = \widetilde{\mu} + \dfrac{1}{2}\widetilde{\sigma}^{2} \geq 0 \end{cases}. \tag{5.11}$$

Theorem 5.3. *In model (5.1), assume $\widetilde{\sigma}_0^2 > 0$. When T is fixed and $\Delta \to 0$, we have $\widehat{\rho} \xrightarrow{p} \rho_0$ and its asymptotic distribution is*

$$\frac{1}{\Delta}\left(\widehat{\rho} - \rho_0\right) \xrightarrow{L} \mathcal{N}\left(0, (TV_2)^{-1}\right), \tag{5.12}$$

where

$$\widehat{\rho} = \frac{\sum_{t=1}^{M\times N} \dfrac{y_{(t-1)\Delta} y_{t\Delta}}{\widehat{\widetilde{\sigma}^2} y_{(t-1)\Delta}^2 + 2\widehat{\omega} y_{(t-1)\Delta} + \widehat{\sigma^2}}}{\sum_{t=1}^{M\times N} \dfrac{y_{(t-1)\Delta}^2}{\widehat{\widetilde{\sigma}^2} y_{(t-1)\Delta}^2 + 2\widehat{\omega} y_{(t-1)\Delta} + \widehat{\sigma^2}}} =: 1 + \frac{\widehat{A}_N^*}{\widehat{B}_N^*}.$$

According to Theorems 5.2 and 5.3 the estimates $\widehat{\widetilde{\mu}}$ and $\widehat{\rho}$ continue to have asymptotic normal distributions under infill asymptotics. This convenient feature allows us to apply the testing procedures proposed in the previous section after making a minor change in the variance of the limit distribution to accommodate endogeneity.

6 Simulations

This section reports the results of Monte Carlo simulations designed to evaluate the performance of the two-stage estimator. We also examine the finite sample adequacy of the asymptotic theory for the test statistics developed in Sections 4 and 5. To save space, we only report the results for model (2.2) under explosiveness. More tables reporting bias and standard errors under stationary, unstable, and locally explosive cases are given in the online supplement to this paper (Tao, Phillips and Yu (2018)).

The simulations involved 10,000 replications of sample paths generated from model (2.2) under explosiveness with parameter values $\widetilde{\mu} = 1$, $\widetilde{\sigma} = 1$, $\sigma = 1$, and with initial condition $y_0 = 10$.[7] Since $\phi > 0$ this generating process leads to explosiveness. In the first experiment, we set the time span $T = 5$, but varied Δ from 1/252 to 1/19656 and varied M from 21, 63 to 252. $\Delta = 1/252$ corresponds to daily observations whereas $\Delta = 1/19656$ corresponds to 5-minute (high frequency) observations. When $\Delta = 1/252$, $M = 21$, 63 and 252 implies a corresponding block size that is monthly, quarterly, and annual, respectively. When $\Delta = 1/19656$, we report the estimation bias and standard errors by holding the number of observations for calculating the realized volatilities (M) constant as in a daily frequency. In panel A of Table 2, we report the bias and the standard errors of the two-stage estimates when there is no endogeneity in the model, i.e. when $\gamma = 0$, and in panel B, we report the corresponding results for the model with endogeneity, specifically with $\gamma = 0.8$. The bias and the standard errors are computed using 5,000 replications.

[7] We also report bias and standard errors under stationary, unstable, and locally explosive cases in the online appendix.

Table 2: Bias and standard errors of the two-stage estimates for different Δ and M and a fixed $T(=5)$. The parameter values are $\widetilde{\mu}=1$, $\widetilde{\sigma}^2=1$, $\sigma^2=1$. $y_0=10$.

	$\Delta=1/252$						$\Delta=1/19656$					
Params	$M=21$		$M=63$		$M=252$		$M=21$		$M=63$		$M=252$	
	Bias	S.E.	Bias	S.E.	Bias	S.E.	Bias	S.E.	Bias	S.E.	Bias	S.E.
Panel A: $\gamma=0$												
$\widetilde{\mu}$	-0.0328	0.4942	-0.0347	0.4946	-0.0455	0.4938	-0.0279	0.5173	-0.0279	0.5173	-0.0279	0.5173
$\widetilde{\sigma}^2$	-0.0093	0.0471	-0.0135	0.0493	-0.0190	0.0611	0.0014	0.0056	4.7e-04	0.0055	-8.3e-05	0.0055
σ^2	4.6415	13.0303	6.1285	18.5101	28.7040	138.7416	0.2518	1.3496	0.2385	1.3241	0.2414	1.3443
κ	-0.0375	0.4952	-0.0414	0.4958	-0.0549	0.4962	-0.0272	0.5172	-0.0276	0.5172	-0.0279	0.5172
ρ	-1.3e-04	0.0020	-1.4e-04	0.0020	-1.8e-04	0.0020	-1.4e-06	2.6e-05	-1.4e-06	2.6e-05	-1.4e-06	2.6e-05
ϕ	-0.0282	0.4943	-0.0279	0.4945	-0.0360	0.4933	-0.0285	0.5174	-0.0281	0.5174	-0.0278	0.5174
Panel B: $\gamma=0.8$												
$\widetilde{\mu}$	-0.0487	0.5213	-0.0511	0.5214	0.0208	1.5405	-0.0368	0.5194	-0.0368	0.5194	-0.0368	0.5194
$\widetilde{\sigma}^2$	0.0326	0.0974	0.0350	0.1119	0.0776	0.2239	0.0037	0.0105	0.0028	0.0119	0.0028	0.0104
σ^2	19.2645	45.3508	28.3990	76.9436	193.6521	1.1e+03	1.4455	3.7584	1.4215	3.6797	1.4215	3.7170
γ	-0.4179	0.6293	-0.4519	0.6521	-0.6296	0.7499	-0.0884	0.1778	-0.0874	0.1750	-0.0874	0.1762
κ	-0.0324	0.5208	-0.0336	0.5199	0.0596	1.5431	-0.0349	0.5193	-0.0354	0.5193	-0.0354	0.5193
ρ	-1.9e-04	0.0021	-2.0e-04	0.0021	9.9e-05	0.0064	-4.9e-05	1.1e-05	-4.1e-05	2.4e-05	-4.1e-05	2.4e-05
ϕ	-0.0650	0.5262	-0.0686	0.5289	-0.0180	1.5461	-0.0386	0.5196	-0.0382	0.5196	-0.0382	0.5196

First, from Table 2 it is apparent that when the sampling frequency increases the parameters $\widetilde{\sigma}^2$, γ and σ^2 are all better estimated in terms of bias and standard error. On the other hand, there is little improvement in the estimation of $\widetilde{\mu}$ because the time span does not change. This finding corroborates the asymptotic theory for $\widetilde{\mu}$ given in Theorem 4.2 and also supports results found in Yu (2012). Furthermore, due to the difference in the convergence rates shown in Theorems 4.1 and 4.2, the bias and the standard errors of $\widehat{\kappa}$ and $\widehat{\phi}$ are mainly determined by those of $\widehat{\widetilde{\mu}}$, which explains why estimation performance of $\widehat{\kappa}$ and $\widehat{\phi}$ does not improve as sampling frequency increases. Finally, bias and standard errors both appear reasonably robust across different values of M.

In the second experiment, we fix $\Delta=1/252$, but vary T from 30 to 60 and M from 21, 63 to 252. In Panel A of Table 3, we report the bias and the standard errors of the two-stage estimators across 5,000 simulated samples for the model without endogeneity. The same experiment is repeated for the model with endogeneity and the results are reported in Panel B. Several findings are evident from Table 3. First, as the time span enlarges, sharp reductions occur in the bias and standard error of $\widehat{\widetilde{\mu}}$. Combined with the results of Table 2, this finding suggests that time span, not sampling frequency, is the primary influence on performance of $\widehat{\widetilde{\mu}}$. Second, the bias and standard errors of $\widehat{\widetilde{\sigma}^2}$, $\widehat{\gamma}$ and $\widehat{\sigma^2}$ do not change significantly as T increases. Finally, both bias and standard errors are again robust with respect to M.

Table 3: Bias and standard error of the two-stage estimates for different T and M and fixed $\Delta(= 1/252)$. The parameter values are $\widetilde{\mu} = 1$, $\widetilde{\sigma}^2 = 1$, $\sigma^2 = 1$. $y_0 = 10$.

	$T = 30$						$T = 60$					
Params	$M = 21$		$M = 63$		$M = 252$		$M = 21$		$M = 63$		$M = 252$	
	Bias	S.E.	Bias	S.E.	Bias	S.E.	Bias	S.E.	Bias	S.E.	Bias	S.E.
Panel A: $\gamma = 0$												
$\widetilde{\mu}$	-0.0058	0.1861	-0.0060	0.1861	-0.0075	0.1860	-0.0014	0.1301	-0.0015	0.1301	-0.0023	0.1302
$\widetilde{\sigma}^2$	-0.0031	0.0175	-0.0066	0.0179	-0.0077	0.0208	-0.0023	0.0123	-0.0058	0.0126	-0.0068	0.0147
σ^2	3.8582	11.6119	5.1172	16.6821	24.0530	136.4710	3.8333	11.6501	5.0750	16.5899	24.4001	137.0690
κ	-0.0073	0.1864	-0.0093	0.1862	-0.0113	0.1860	-0.0026	0.1304	-0.0044	0.1303	-0.0057	0.1303
ρ	-2.3e-05	7.4e-04	-2.4e-05	7.4e-04	-2.9e-05	7.4e-04	-5.6e-06	5.2e-04	-6.0e-06	5.2e-04	-9.0e-06	5.2e-04
ϕ	-0.0042	0.1862	-0.0027	0.1864	-0.0036	0.1866	-2.7e-04	0.1302	0.0013	0.1303	0.0011	0.1304
Panel B: $\gamma = 0.8$												
$\widetilde{\mu}$	-9.7e-04	0.1796	-0.0013	0.1796	0.0045	0.6583	0.0016	0.1277	0.0014	0.1277	0.0127	0.5476
$\widetilde{\sigma}^2$	-1.6e-04	0.0188	-0.0035	0.0193	-0.0034	0.0225	-0.0010	0.0126	-0.0043	0.0130	-0.0048	0.0153
σ^2	12.4166	32.3889	17.3784	53.4253	81.2221	396.9127	12.1161	31.6064	16.9538	52.3937	80.3542	393.2385
γ	-0.2879	0.5323	-0.3136	0.5496	-0.3924	0.6223	-0.2768	-0.5228	-0.3011	0.5417	-0.3848	0.6138
κ	-0.0010	0.1799	-0.0030	0.1798	0.0028	0.6584	0.0011	0.1279	-7.8e-04	0.1278	0.0103	0.5475
ρ	-3.6e-06	7.2e-04	5.0e-06	7.2e-04	2.2e-05	0.0027	6.4e-06	5.1e-04	5.7e-06	5.1e-04	5.3e-05	0.0022
ϕ	-8.9e-04	0.1797	4.2e-04	0.1800	0.0063	0.6584	0.0021	0.1278	0.0036	0.1279	0.0151	0.5477

From Table 2 and Table 3, it is evident that the proposed two-stage method is effective in estimating $\widetilde{\mu}$, γ, $\widetilde{\sigma}^2$, κ, ρ, ϕ even in the presence of endogeneity. While the estimate of σ^2 is less satisfactory,[8] this outcome is unsurprising because when $\kappa > 0$, y_t^2 grows exponentially with t. Hence, estimates of γ and σ^2 are dominated by the component $\widetilde{\sigma}^2 y_t^2$ in $\widetilde{\sigma}^2 y_t^2 + 2\omega y_t + \sigma^2$ when t is large. More importantly, the three forms of explosive behavior do not depend on γ and σ^2 in that case. Hence, it is expected that the performance of $\widehat{\gamma}$ and $\widehat{\sigma^2}$ will have little impact on the performance of the proposed t-tests and the time-stamping strategy.

The third experiment is designed to evaluate the performance of the test statistics introduced in Remark 4.8. To do so, we simulate 5,000 sample paths from model (2.2) and from model (5.1) with $\gamma = 0.8$, and calculate the power and size of the three tests. In the first case, we set the nominal size to $5\%, M = 21$ and $\Delta = 1/252$, but vary the time span T. In the second case, we set the nominal size to $5\%, M = 21$ and $T = 10$, but vary the sampling interval Δ. The values of Δ correspond to 1-day (1/252), 1-hour (1/1638), and 30-minute (1/3276), respectively. Results for the power and size are reported in Table 4 and 5.

The simulation results show that size distortion of the proposed tests for different types of explosive behavior becomes smaller when data is sampled more frequently. In addition, the power of the tests rises rapidly as the sample size increases (by either increasing the time span

[8] In both Table 2 and Table 3, the bias and S.E. of $\widetilde{\sigma}^2$ and σ^2 evidently increase with M. The explanation is that, given Δ and T, as M increases the effective sample size in the first stage estimation (N) decreases, which harms performance of $\widehat{\widetilde{\sigma}^2}$ and $\widehat{\sigma^2}$.

Table 4: Power and size of the t tests under different forms of unstable/explosive behavior with fixed Δ $(= 1/252)$ and growing time span T.

T	t_κ				$t_{\widetilde{\mu}}$				t_ϕ			
$\gamma = 0$	Size	$\kappa = 0.5$	$\kappa = 1$	$\kappa = 2$	Size	$\widetilde{\mu} = 0.5$	$\widetilde{\mu} = 1$	$\widetilde{\mu} = 2$	Size	$\phi = 0.5$	$\phi = 1$	$\phi = 2$
5	0.0388	0.2564	0.6580	0.9948	0.0436	0.2812	0.7062	0.9974	0.0472	0.2924	0.7170	0.9976
10	0.0342	0.3304	0.8066	1.0000	0.0394	0.4406	0.9116	1.0000	0.0470	0.4680	0.9328	1.0000
15	0.0366	0.4094	0.8946	1.0000	0.0376	0.5472	0.9664	1.0000	0.0472	0.6002	0.9836	1.0000
30	0.0384	0.5804	0.9840	1.0000	0.0380	0.7468	0.9990	1.0000	0.0484	0.8528	1.0000	1.0000
T	t_κ				$t_{\widetilde{\mu}}$				t_ϕ			
$\gamma = 0.8$	Size	$\kappa = 0.5$	$\kappa = 1$	$\kappa = 2$	Size	$\widetilde{\mu} = 0.5$	$\widetilde{\mu} = 1$	$\widetilde{\mu} = 2$	Size	$\phi = 0.5$	$\phi = 1$	$\phi = 2$
5	0.0448	0.2608	0.6340	0.9860	0.0452	0.2732	0.6690	0.9942	0.0436	0.2700	0.6816	0.9964
10	0.0410	0.4384	0.8898	0.9998	0.0464	0.4818	0.9154	1.0000	0.0436	0.4532	0.9176	1.0000
15	0.0414	0.5830	0.9750	1.0000	0.0538	0.6650	0.9814	1.0000	0.0458	0.5928	0.9810	1.0000
30	0.0404	0.8382	0.9994	1.0000	0.0532	0.9260	1.0000	1.0000	0.0466	0.8614	1.0000	1.0000

Table 5: Power and size of the t tests under different forms of unstable/explosive behavior with fixed T $(= 10)$ and shrinking sampling interval Δ.

Δ	t_κ				$t_{\widetilde{\mu}}$				t_ϕ			
$\gamma = 0$	Size	$\kappa = 0.5$	$\kappa = 1$	$\kappa = 2$	Size	$\widetilde{\mu} = 0.5$	$\widetilde{\mu} = 1$	$\widetilde{\mu} = 2$	Size	$\phi = 0.5$	$\phi = 1$	$\phi = 2$
1/252	0.0342	0.3304	0.8066	1.0000	0.0394	0.4406	0.9116	1.0000	0.0470	0.4680	0.9328	1.0000
1/1638	0.0354	0.3260	0.8126	0.9992	0.0442	0.4412	0.9158	1.0000	0.0476	0.4734	0.9302	1.0000
1/3276	0.0358	0.3340	0.8156	0.9998	0.0476	0.4462	0.9102	1.0000	0.0504	0.4740	0.9270	1.0000
Δ	t_κ				$t_{\widetilde{\mu}}$				t_ϕ			
$\gamma = 0.8$	Size	$\kappa = 0.5$	$\kappa = 1$	$\kappa = 2$	Size	$\widetilde{\mu} = 0.5$	$\widetilde{\mu} = 1$	$\widetilde{\mu} = 2$	Size	$\phi = 0.5$	$\phi = 1$	$\phi = 2$
1/252	0.0410	0.4384	0.8898	0.9998	0.0464	0.4818	0.9154	1.0000	0.0436	0.4532	0.9176	1.0000
1/1638	0.0410	0.4274	0.8868	0.9998	0.0520	0.4734	0.9146	1.0000	0.0474	0.4586	0.9214	1.0000
1/3276	0.0426	0.4418	0.8852	0.9988	0.0538	0.4830	0.9114	1.0000	0.0510	0.4596	0.9164	1.0000

T or shrinking the sampling interval Δ) and for greater departures of the true parameters from the null. The reason for smaller size distortions as Δ shrinks is that the approximation errors of the distributions of $\Delta^{-1}(\widehat{\beta}_\kappa - \beta_\kappa^0)$ and $\Delta^{-1}(\widehat{\beta}_\phi - \beta_\phi^0)$ to the limiting distributions are of order $O(\sqrt{\Delta})$; see (A.9).

Next, we check size and power of the tests under endogeneity. Simulations are generated by setting $\widetilde{\sigma}^2 = 1$ and $\sigma^2 = 100$ with $y_0 = 10$ under the sampling scheme $\Delta = 1/252$, $M = 21$ and $T = \{5, 10, 15, 30\}$. Sample paths are generated in 1000 replications for parameter values $\widetilde{\mu} = \{-1, -0.5, 0.5, 1\}$ and for correlation coefficients $\gamma = \{0, 0.04, 0.08, 0.4, 0.8\}$. The results in Table 6 show that size distortion is very small under all parameter scenarios and that test power grows more slowly as the process becomes more unstable. This phenomenon is due to the structure of the quadratic variation $\widetilde{\sigma}^2 y_t^2 + 2\omega y_t + \sigma^2$ under endogeneity. When the process y_t is more unstable, the signal from y_t^2 is stronger and a much larger value of ω is needed for the component $2\omega y_t$ in the quadratic variation to enhance the probability of rejecting the null. Also, as expected, the power of the test increases with the increase in sample size.

Table 6: Power and size of the LR test for endogeneity.

T	Stationary ($\widetilde{\mu} = -1$)					Unstable ($\widetilde{\mu} = -0.5$)				
	$\gamma = 0$	0.04	0.08	0.4	0.8	$\gamma = 0$	0.04	0.08	0.4	0.8
5	0.0470	0.0780	0.1950	0.9510	0.9940	0.0500	0.0810	0.1920	0.9210	0.9850
10	0.0440	0.1290	0.4110	0.9990	1.0000	0.0490	0.1200	0.4030	0.9970	1.0000
15	0.0540	0.1960	0.6000	1.0000	1.0000	0.0560	0.2160	0.5970	1.0000	1.0000
30	0.0490	0.3820	0.8790	1.0000	1.0000	0.0540	0.3990	0.8920	1.0000	1.0000
T	Locally Explosive ($\widetilde{\mu} = 0.5$)					Explosive ($\widetilde{\mu} = 1$)				
	$\gamma = 0$	0.04	0.08	0.4	0.8	$\gamma = 0$	0.04	0.08	0.4	0.8
5	0.0600	0.0880	0.1580	0.7260	0.8710	0.0460	0.0580	0.1070	0.5940	0.7850
10	0.0460	0.1120	0.2770	0.8650	0.9490	0.0530	0.0690	0.1640	0.7120	0.8800
15	0.0540	0.1210	0.3680	0.9090	0.9690	0.0540	0.0870	0.1780	0.7310	0.9020
30	0.0520	0.1910	0.5270	0.9480	0.9860	0.0540	0.0810	0.1900	0.7530	0.9130

The final experiment checks performance of the proposed tests of coefficient constancy, i.e. $\mathcal{H}_0 : \widetilde{\sigma}^2 = 0$. To do so, we simulate 10,000 sample paths from model (5.1) with different parameter values to cover the various explosive scenarios. Both size and power are calculated. More specifically, we vary $\widetilde{\mu}$ from -0.1 to 0.1, $\widetilde{\sigma}$ from 0 to 0.2 and $\gamma \in \{0, 0.8\}$ holding $\sigma = 1$, which covers all explosive scenarios. In these experiments, we set nominal size to 5%, $M = 21$ and $\Delta = 1/252$, but vary the time span T to control for sample sizes. Test size and power are reported in Table 7.

Table 7: Power and size of the modified LBI-test for different null models.

T	$\widetilde{\mu}=-0.1$				$\widetilde{\mu}=0$				$\widetilde{\mu}=0.1$			
	$\widetilde{\sigma}=0$	0.04	0.10	0.20	$\widetilde{\sigma}=0$	0.04	0.10	0.20	$\widetilde{\sigma}=0$	0.04	0.10	0.20
Panel A: $\gamma=0$												
5	0.0490	0.1727	0.8675	0.9977	0.0468	0.1551	0.7868	0.9977	0.0501	0.4940	0.9251	0.9977
10	0.0472	0.2845	0.9929	1.0000	0.0495	0.3377	0.9838	1.0000	0.0497	0.9659	0.9985	1.0000
15	0.0489	0.3449	0.9980	1.0000	0.0458	0.5043	0.9982	1.0000	0.0457	0.9980	0.9998	1.0000
30	0.0467	0.3908	0.9998	1.0000	0.0472	0.7936	1.0000	1.0000	0.0534	1.0000	1.0000	1.0000
Panel B: $\gamma=0.8$												
5	0.0490	0.5755	0.9435	0.9989	0.0468	0.4338	0.9134	0.9990	0.0501	0.7190	0.9499	0.9989
10	0.0472	0.9262	0.9993	1.0000	0.0495	0.7996	0.9980	1.0000	0.0497	0.9819	0.9995	1.0000
15	0.0489	0.9861	1.0000	1.0000	0.0458	0.9438	0.9998	1.0000	0.0457	0.9993	1.0000	1.0000
30	0.0467	0.9895	0.9982	0.9990	0.0472	0.9968	0.9973	0.9975	0.0534	1.0000	0.9996	0.9996

7 Empirical Studies

7.1 Daily data

As a practical illustration of our methods with real data, we analyze daily S&P 500 real prices over the period from December 31, 1927 to June 29, 2018 with $T = 91.5$. For the first stage estimation we use daily prices to calculate monthly realized variance and monthly realized quarticity within each month. Thus 1098 blocks are chosen with each block containing daily observations within each calendar month, i.e., $N = 1098$, $h = 1/12$. While on average there are 21 trading days within each month, the actual number of days within a month is year-dependent as well as month-dependent. Similarly, while on average there are 252 trading days within each year, the actual number of days within a year is year-dependent. If there are M days within a month, we set $\Delta = h/M$ for that month.

We first apply our estimation, testing, and time-stamping strategies to these daily price data based on the model with no endogeneity, that is, model (2.2).[9] Following PWY (2011), the initial window is taken as the first 5-year of the full sample. For comparison purposes, we also implement the BADF test of PWY and the BSADF test of PSY. The empirical results are shown in Figure 5, where we plot the test statistic sequences under the three forms of explosiveness and the test statistic sequences under the assumption of time-invariant coefficients. We also plot the 95% critical values and the data in each panel.

The last panel in Figure 5 plots the recursive test statistic sequence for testing a time-invariant autoregressive coefficient. The test results suggest that throughout the whole sample period the data are well described by a model with time varying coefficients as the test statistic

[9]We used model (2.2) largely because the bias in estimation of ω is relatively large in long-span, low-frequency samples and the bias becomes severe when the process is explosive (c.f., Lieberman and Phillips (2017c)). The methods of the present paper are more relevant in models without endogeneity when high-frequency data are unavailable.

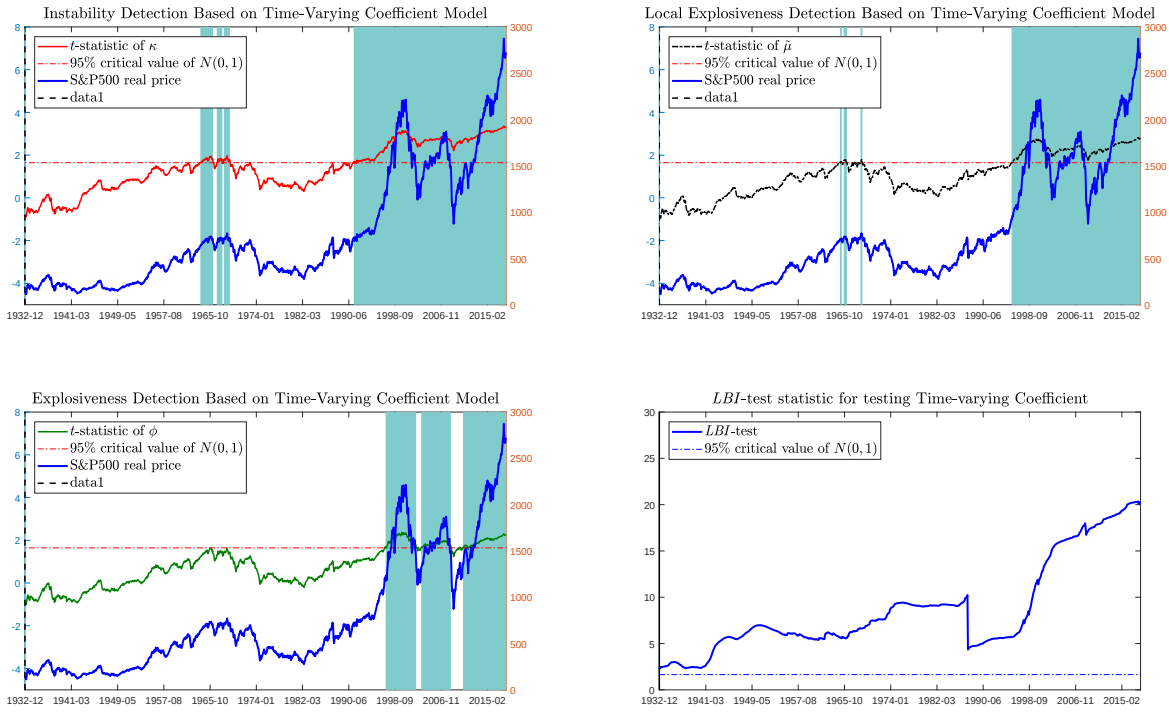


Figure 5: Date-stamping Explosive Periods in the S&P500 Real Price Index with Time-Varying Coefficient Model without Endogeneity.

for a time varying coefficient always exceeds the 95% critical value. In fact, the test statistic generally increases as the time period expands. Although the value of the test statistic drops in October 1987, indicative of a possible change in market behavior at that point, it is still larger than the critical value. Evidence for time variation becomes much stronger from July 1995 onwards. This dating coincides well with the estimated origination dates of the three forms of extreme behavior indicated by the other three panels in Figure 5. For example, the first panel in Figure 5 indicates that real stock prices have experienced two major unstable periods. The unstable behavior is first detected at February 1964 and is interrupted twice over the succeeding period to May 1969. Afterwards real stock prices are not unstable or explosive until May 1991, from which point the instability continues until the end of the sample.

The second panel in Figure 5 indicates that real prices are not locally explosive between January 1928 to January 1965, at which point locally explosive (submartingale) behavior is detected. This behavior is interrupted twice over the succeeding period to January 1969. The locally explosive behavior is detected again in July 1995 which continues to the end of the sample. Since local explosiveness is more extreme behavior than instability, its duration tends to be much shorter than that of instability.

The third panel in Figure 5 indicates that the real price index does not experience explosive behavior between January 1928 and July 1997. An episode of explosive behavior is detected in July 1997, lasting for a few years and ending in June 2002. The end date corresponds to the termination of the well-known internet bubble period. Then, the test detects a second explosive period starting from May 2003 to August 2008, which corresponds to the pre global financial crisis period. The panel also suggests a further explosive episode starting in October 2010 and continuing to the end of the sample period. These three periods of major price escalation in the sample (namely, the second half of 1990s, the pre global financial crisis period, and the recovery from the global financial crisis) are all deemed to have explosiveness which seems to be a reasonable empirical finding. Time horizons for the three forms of extreme behavior detected by our method are summarized in a table in the online supplement (Tao et al. (2018)).

For comparison purposes, Figure 6 plots the recursive BADF statistic (used in PWY), the recursive BSADF statistic (used in PSY), and corresponding 95% critical values together with the sample data in each panel. It is clear that both PWY and PSY tests identify explosive behavior in the second half of the 1990s earlier than the method proposed in the present paper. This early origination date identification is achieved by using a more restrictive reduced form autoregressive model. Interestingly, the PSY test recursion indicates two similar pronounced periods of explosive behavior, one in the second half of the 1990s and the other at the end of the sample, both matching those identified by methods of the present paper using a more complex modeling framework. Time horizons for the explosive behavior detected by PWY and PSY are summarized in a table in Tao, Phillips and Yu (2018).

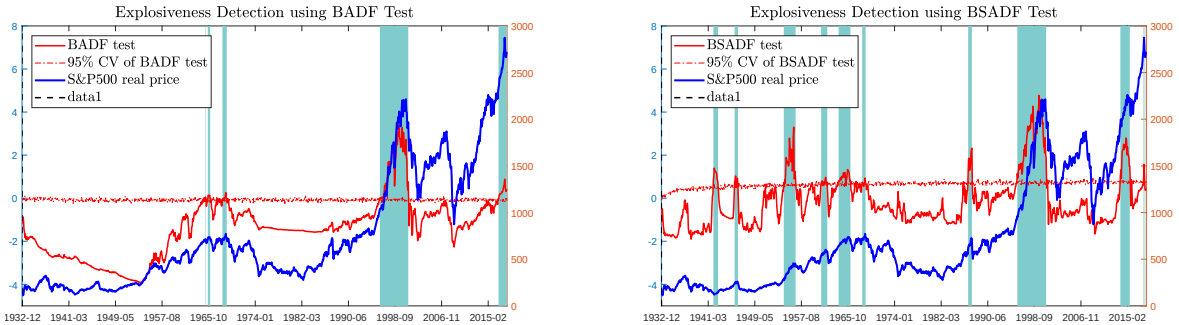


Figure 6: Date-stamping Explosive Periods in the S&P500 Real Price Index with Fixed Coefficient Model.

To address the possible presence of endogeneity we estimate the more general model in which endogeneity effects are permitted, i.e., model (5.1). The results are summarized by the recursions plotted in Figure 7. First, from the test for endogeneity it is apparent that the null of exogeneity is rejected almost everywhere throughout the entire sample, confirming

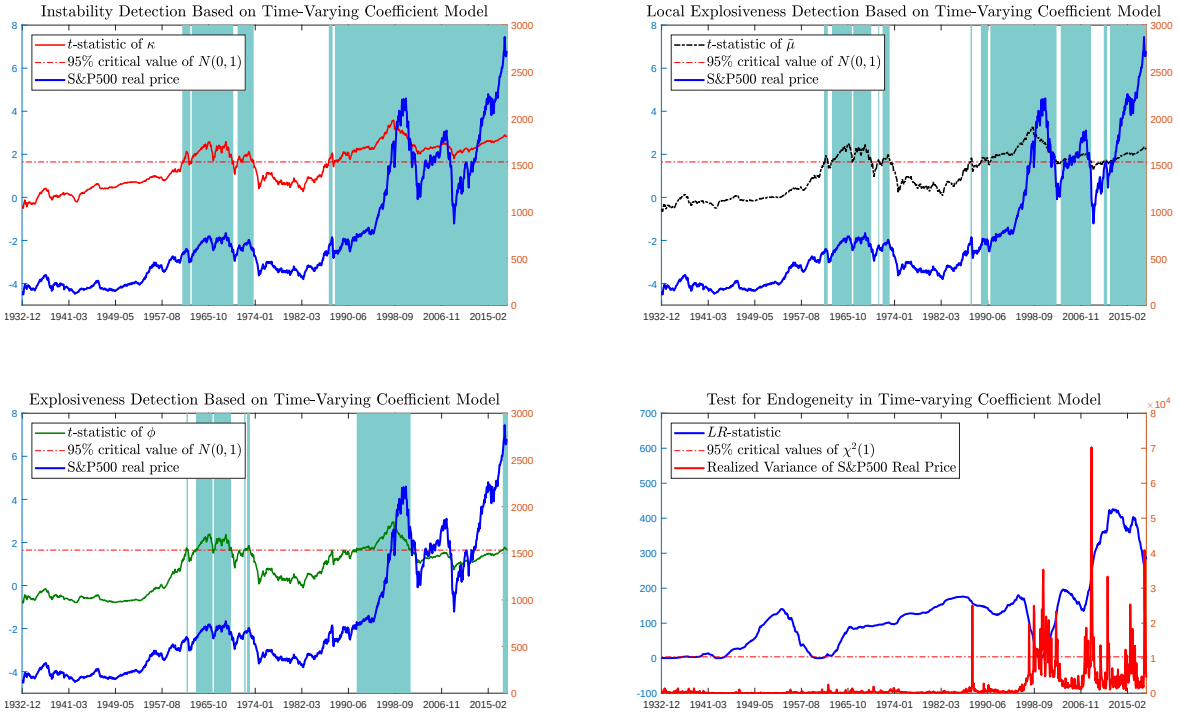


Figure 7: Date-stamping Explosive Periods in the S&P500 Real Price Index with Time-varying Coefficient Model with Endogeneity.

that endogeneity is important in the generating mechanism for these data. From the plotted realized variance graphic in Figure 7 it is further evident that the rejection of exogeneity is closely associated with the behavior of the quadratic variation of the process. This is explained by the fact that the likelihood ratio statistic is based on an objective function that is constructed using a central limit theorem (CLT) for the realized variance time series. The test statistic for endogeneity therefore captures differences in the realized variance estimates using different models, as is shown in Figure 8. Further, more nonstationary horizons are detected in the data when using the model with endogeneity. In particular, extreme behavior is detected in the form of both unstable and locally explosive behavior during price spike in 1987. Note that this explosive behavior is also reported in the PSY test. However, previous empirical evidence for explosiveness before global financial crisis (namely, from May 2003 to August 2008) disappears in the fitted model where endogeneity effects are incorporated in the autoregressive response mechanism. These findings indicate that endogeneity feedbacks in the random coefficient autoregressive model framework can play an important role in assessing evidence for various types of instability and explosiveness in the data. These time horizons for different form of extreme behavior are summarized in a table in Tao et al. (2018).

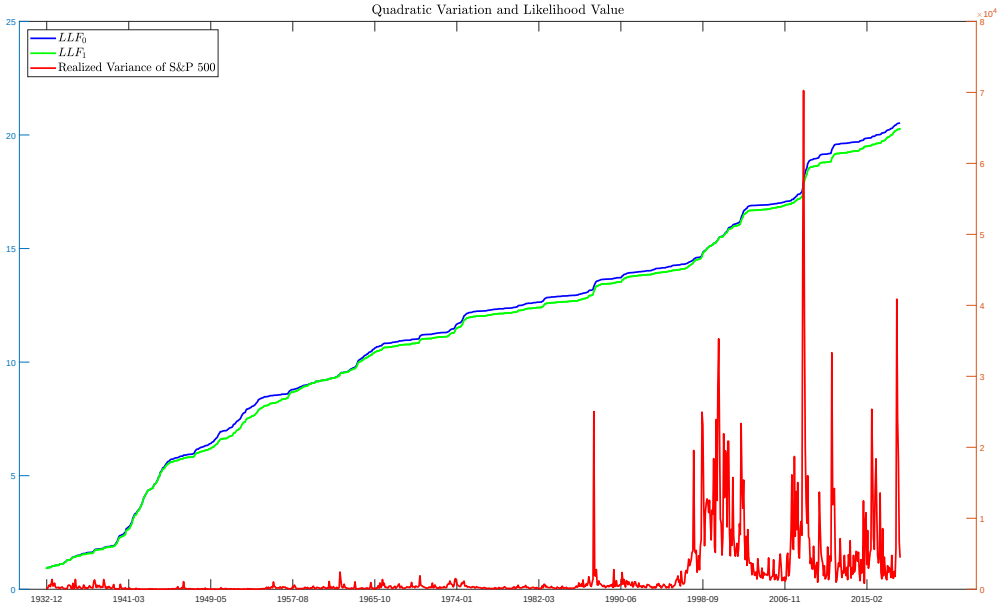


Figure 8: Realized Variance and Likelihood Values for Random Coefficient Autoregression.

7.2 Intra-day data

To further assess evidence for endogeneity and to reduce bias in the estimation of γ, we estimate the same model (5.1) using 5-minute high-frequency data for the real prices of the E-mini S&P 500 futures contract over the period from the closing on October 31, 1997 to the closing on October 31, 2013. Use of this high frequency intra-day data leads to a substantial increase in sample size, accords more closely with infill asymptotic theory, but has the limitation that the model itself abstracts from possible intra-day effects that are known to be present in ultra high frequency data. On the other hand, use of 5-minute data (rather than even higher frequency observations) helps to mitigate some of these intra-day effects and gives the benefit of bias reduction in estimation of the correlation between the equation errors and the random autoregressive coefficient, thereby improving estimation of the degree of endogeneity in the random coefficient driver variables.

We collect 16 years of 5-minute high-frequency S&P 500 real prices with $T = 16$. That is, in each trading day we have 79 real prices, sampled at 9:30am, 9:35am, ..., 3:55pm, 4:00pm. The overnight price movement is captured by the difference between the price observed at 4:00pm on day t and that at 9:30am on day $t + 1$. By doing so, we treat overnight price movements in the same way as intra-day price movements over 5-minute intervals.

For the first stage estimation we use price changes over 5-minute intervals and overnight to calculate monthly realized variance and monthly realized quarticity within each month. That is, 192 blocks are chosen with each block containing high-frequency data within each

calendar month, i.e., $N = 192$, $h = 1/12$. If there are M days within a month we set $\Delta = h/(78 \times M)$ for that month. Model (5.1) is fitted recursively with an initial window size of 5 years. The empirical results are summarized in Figure 9 on monthly basis.

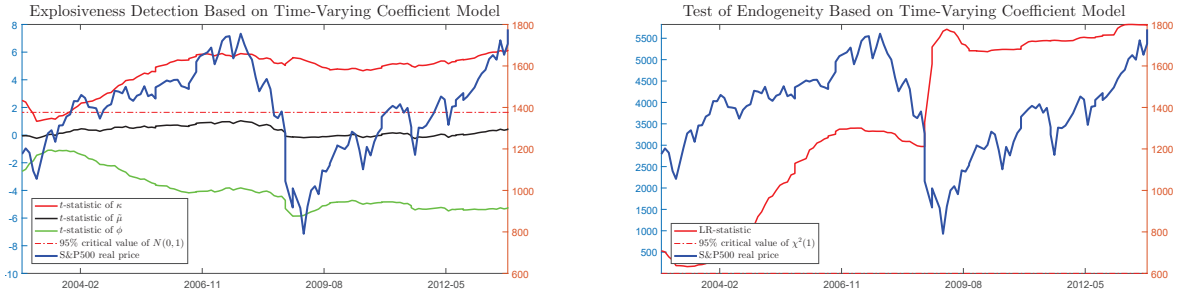


Figure 9: Testing Explosiveness and Endogeneity in the 5-minute S&P 500 Real Price Index with Time-varying Coefficient Model.

The recursive test statistic graphics in Figure 9 indicate that over this sample period and allowing for high frequency fluctuations the data are unstable but not locally explosive or explosive. Based on the simulation findings in the previous section, estimates of the endogeneity parameter γ can be expected to have reasonably small bias at this frequency and the t-tests to have good size and power. From the second panel in Figure 9, the LR-test for endogeneity always exceeds the 5% critical value of the χ^2_1 distribution, which reinforces from the 5-minute high-frequency data the evidence in support of endogenous effects on the autoregressive coefficient found in the daily-frequency sample.

8 Summary and Conclusions

This paper introduces a continuous time model for financial data where the persistence parameter is allowed to be random and time varying. The model has an analytical solution and an exact discrete time representation which make analysis convenient for studying the properties of the system that are associated with extreme sample path behavior. The discrete time model relates to some models already in the literature, including the stochastic unit root model (Granger and Swanson (1997); Lieberman and Phillips (2014); Lieberman and Phillips (2017b)) and the near-explosive random coefficient model of Banerjee et al. (2017). The statistical properties of our model reveal three different forms of potential extreme behavior in generated sample paths: instability, local explosiveness, and explosiveness. These forms of extreme behavior depend directly on the values of model parameters, including the possible presence of endogeneity in the random autoregressive coefficient.

A novel two-stage estimation method that relies on empirical quadratic variation is developed to estimate the model parameters. Limit theory is developed using an infill asymptotic

scheme that provides a convenient basis for testing parameter constancy and the various forms of extreme sample path behavior. The test statistics all have asymptotically pivotal standard normal distributions which makes implementation of the tests straightforward in practical work. Similar to other recent work in the literature on bubbles, a time-stamping strategy is proposed to detect origination and termination dates of extreme behavior.

In an empirical application to daily S&P 500 real prices between December 31, 1927 and June 29, 2018. Strong evidence against parameter constancy is found in the whole sample period and this evidence strengthens after July 1997, leading to a finding of long durations of parameter instability in the model. Three periods of explosive instability in the data match well with observed periods of major price escalation in the data and these largely overlap with the periods of price exuberance identified in earlier work. Tests for endogeneity in these data provide strong evidence in support of endogenous feedbacks in the random coefficient model framework that materially influence quadratic variation and hence recursive estimates of realized variation in the data. The empirical findings of extreme sample path behavior in real S&P 500 stock prices are broadly in line with the conclusions of other recent work on stock price exuberance but now provide new evidence against parameter constancy and in support of the role of endogenous feedbacks that influence autoregressive behavior and the time forms of extreme sample paths.

List of Symbols and Notations

The symbols and notations used in the present paper are summarized in the following table.

Symbol	Meaning	Definition
$\widetilde{\mu}$	Drift parameter of the random coefficient	Equation (2.2)
$\widetilde{\sigma}$	Diffusion parameter of the random coefficient	Equation (2.2)
σ	Diffusion parameter of the Ornstein-Uhlenbeck process in a random environment	Equation (2.2)
B_u, B_ε	Standard Brownian motion that generates randomness in the drift component/the diffusion component	Equation (2.2)
γ	Covariance between the Brownian motions B_u and B_ε	$\gamma = \frac{d\langle B_u, B_\varepsilon \rangle_t}{dt}$
ω	Covariance between the drift process and diffusion process	$\omega = \gamma\widetilde{\sigma}\sigma$
Δ, h	Sampling interval/block	
M	Number of observations within each sampling block	$M = h/\Delta$
T	Time span	
N	Number of sampling blocks	$N = T/h$
$\rho_{t\Delta}$	Random autoregressive coefficient	Equation (2.8)
$\eta_{t\Delta}$	Innovation of the random coefficient autoregression	Equation (2.9)
ρ	Expectation of the random autoregressive coefficient	$\rho = E(\rho_{t\Delta}) = e^{\widetilde{\mu}\Delta}$
κ	Key parameter controls the unstable behavior	$\kappa = \widetilde{\mu} + \widetilde{\sigma}^2/2$
ϕ	Key parameter controls the explosive behavior	$\phi = \widetilde{\mu} - \widetilde{\sigma}^2/2$
β_κ, β_ϕ	Exponential transformation of κ and ϕ	$\beta_\kappa = e^{\kappa\Delta}$, $\beta_\phi = e^{\phi\Delta}$
$[y]_a^b$	Quadratic variation process of y_t for $t \in [a, b]$	Equation (3.1)
$Q_\Delta(\cdot)$	Objective function of the first-stage estimation	Equation (3.3)
$\ell_{ALF}(\cdot)$	Approximate log-likelihood function in the second stage	Equation (3.5)

A Appendix

The proofs of Theorem 4.1 and 5.1 follow directly from Phillips and Yu (2009) and are omitted.

A.1 Proof of Theorem 4.2

To show consistency of $\widehat{\widetilde{\mu}}$, by the Taylor expansion of $\exp\{\widetilde{\mu}_0\Delta\}$ in the observationally equivalent model (2.11), we have

$$\begin{aligned} y_{t\Delta} - y_{(t-1)\Delta} &= \exp\{\widetilde{\mu}_0\Delta\}\, y_{(t-1)\Delta} - y_{(t-1)\Delta} + \sqrt{(\widetilde{\sigma}_0^2 y_{(t-1)\Delta}^2 + \sigma_0^2)\Delta}\cdot v_{t\Delta} \\ &= (\widetilde{\mu}_0\Delta + o(\Delta))y_{(t-1)\Delta} + \sqrt{(\widetilde{\sigma}_0^2 y_{(t-1)\Delta}^2 + \sigma_0^2)\Delta}\cdot v_{t\Delta}, \end{aligned} \tag{A.1}$$

where $v_{t\Delta} \overset{i.i.d}{\sim} \mathcal{N}(0,1)$. Let $\zeta_{(t-1)\Delta} := \dfrac{1}{\widehat{\widetilde{\sigma}^2} y_{(t-1)\Delta}^2 + \widehat{\sigma^2}} - \dfrac{1}{\widetilde{\sigma}_0^2 y_{(t-1)\Delta}^2 + \sigma_0^2}$, then according to (3.6), we have

$$\begin{aligned} \widehat{A}_N &= \sum_{t=1}^{M\times N} \frac{y_{(t-1)\Delta}\left(y_{t\Delta} - y_{(t-1)\Delta}\right)}{\widehat{\widetilde{\sigma}^2} y_{(t-1)\Delta}^2 + \widehat{\sigma^2}} \\ &= \sum_{t=1}^{M\times N} \frac{y_{(t-1)\Delta}\left(y_{t\Delta} - y_{(t-1)\Delta}\right)}{\widetilde{\sigma}_0^2 y_{(t-1)\Delta}^2 + \sigma_0^2} + \sum_{t=1}^{M\times N} y_{(t-1)\Delta}\left(y_{t\Delta} - y_{(t-1)\Delta}\right)\zeta_{(t-1)\Delta} \\ &:= A_N + R_A, \end{aligned}$$

and

$$\widehat{B}_N = \sum_{t=1}^{M\times N} \frac{y_{(t-1)\Delta}^2}{\widehat{\widetilde{\sigma}^2} y_{(t-1)\Delta}^2 + \widehat{\sigma^2}} = \sum_{t=1}^{M\times N} \frac{y_{(t-1)\Delta}^2}{\widetilde{\sigma}_0^2 y_{(t-1)\Delta}^2 + \sigma_0^2} + \sum_{t=1}^{M\times N} y_{(t-1)\Delta}^2\zeta_{(t-1)\Delta} := B_N + R_B.$$

Note that, by Theorem 4.1, we have

$$\zeta_{(t-1)\Delta} = \frac{(\widetilde{\sigma}_0^2 - \widehat{\widetilde{\sigma}^2})y_{(t-1)\Delta}^2 + (\sigma_0^2 - \widehat{\sigma^2})}{(\widehat{\widetilde{\sigma}^2} y_{(t-1)\Delta}^2 + \widehat{\sigma^2})(\widetilde{\sigma}_0^2 y_{(t-1)\Delta}^2 + \sigma_0^2)} = O_p\left(\frac{\sqrt{\Delta}}{y_{(t-1)\Delta}^2}\right).$$

This implies that

$$\begin{aligned} R_A &= \widetilde{\mu}_0\Delta \sum_{t=1}^{M\times N} y_{(t-1)\Delta}^2\zeta_{(t-1)\Delta} + \sqrt{\Delta}\sum_{t=1}^{M\times N} y_{(t-1)\Delta} v_{t\Delta}\sqrt{\widetilde{\sigma}_0^2 y_{(t-1)\Delta}^2 + \sigma_0^2}\zeta_{(t-1)\Delta} \\ &= O_p(MN\Delta^{3/2}) + O_p(\sqrt{MN}\Delta) = O_p(T\sqrt{\Delta}) + O_p(\sqrt{T\Delta}) = O_p(T\sqrt{\Delta}), \end{aligned}$$

and

$$R_B = \sum_{t=1}^{M\times N} y_{(t-1)\Delta}^2\zeta_{(t-1)\Delta} = O_p(MN\sqrt{\Delta}) = O_p(T/\sqrt{\Delta}).$$

Now, first solve

$$A_N = \widetilde{\mu}_0 \sum_{t=1}^{M\times N} \frac{y_{(t-1)\Delta}^2 \cdot \Delta}{\widetilde{\sigma}_0^2 y_{(t-1)\Delta}^2 + \sigma_0^2} + \sum_{t=1}^{M\times N} \frac{y_{(t-1)\Delta} v_{t\Delta} \cdot \sqrt{\Delta}}{\sqrt{\widetilde{\sigma}_0^2 y_{(t-1)\Delta}^2 + \sigma_0^2}} + o_p(\Delta B_N)$$
$$= \widetilde{\mu}_0 \Delta B_N + \sqrt{\Delta} C_N + o_p(\Delta B_N),$$

where

$$C_N := \sum_{t=1}^{M\times N} \frac{y_{(t-1)\Delta} v_{t\Delta}}{\sqrt{\widetilde{\sigma}_0^2 y_{(t-1)\Delta}^2 + \sigma_0^2}}.$$

This leads to

$$\breve{\widetilde{\mu}} := \Delta^{-1} \frac{A_N}{B_N} = \widetilde{\mu}_0 + \frac{C_N}{\sqrt{\Delta} B_N} + o_p(1). \tag{A.2}$$

and hence,

$$\sqrt{T}\left(\breve{\widetilde{\mu}} - \widetilde{\mu}_0\right) = \frac{\frac{1}{\sqrt{MN}} \sum_{t=1}^{M\times N} \frac{y_{(t-1)\Delta} v_{t\Delta}}{\sqrt{\widetilde{\sigma}_0^2 y_{(t-1)\Delta}^2 + \sigma_0^2}}}{\frac{1}{MN} \sum_{t=1}^{M\times N} \frac{y_{(t-1)\Delta}^2}{\widetilde{\sigma}_0^2 y_{(t-1)\Delta}^2 + \sigma_0^2}} + o_p(1). \tag{A.3}$$

Denote $U_{t\Delta} := \frac{y_{(t-1)\Delta}^2}{\widetilde{\sigma}_0^2 y_{(t-1)\Delta}^2 + \sigma_0^2}$. Clearly $\{U_{t\Delta}\}$ is bounded for all t. Let us look at the asymptotic behavior of the denominator. First consider the case $\phi < 0$ where $\{y_{t\Delta}\}$ is asymptotically stationary and ergodic as shown in Föllmer and Schweizer (1993). In this case $\{U_{t\Delta}\}$ is also asymptotically stationary and ergodic. By the ergodic theorem we then have,

$$\frac{1}{MN} \sum_{t=1}^{M\times N} \frac{y_{(t-1)\Delta}^2}{\widetilde{\sigma}_0^2 y_{(t-1)\Delta}^2 + \sigma_0^2} \xrightarrow{a.s.} E\left(\frac{y_t^2}{\widetilde{\sigma}_0^2 y_t^2 + \sigma_0^2}\right).$$

In the case $\phi \geq 0$ where $\{y_t\}$ is nonstationary, by equation (2.9),

$$y_{(t-1)\Delta} = \rho_{1\Delta} y_0 + \sigma_0 \left(\eta_{(t-1)\Delta} + \rho_{(t-1)\Delta}\eta_{(t-2)\Delta} + \cdots + \rho_{(t-1)\Delta}\rho_{(t-2)\Delta}\cdots\rho_{2\Delta}\eta_{1\Delta}\right).$$

Hence, we have

$$E\left(\left|U_{t\Delta} - \widetilde{\sigma}_0^{-2}\right| | \rho_{(t-1)\Delta}, \cdots, \rho_{2\Delta}\right) = E\left[\left|\frac{\sigma_0^2}{\widetilde{\sigma}_0^2 y_{(t-1)\Delta}^2 + \sigma_0^2}\right| | \rho_{(t-1)\Delta}, \cdots, \rho_{2\Delta}\right]$$
$$= \int_0^\infty \frac{\sigma_0^2}{\widetilde{\sigma}_0^2 y_0^2 + \widetilde{\sigma}_0^2 \lambda^2 x^2 + \sigma_0^2} d\Phi(x),$$

where $\Phi(\cdot)$ is the cdf of $\eta_{t\Delta} \sim N(0, \gamma_\Delta^2)$ and $\lambda^2 = 1 + \rho_{(t-1)\Delta}^2 + \cdots + \rho_{(t-1)\Delta}^2 \rho_{(t-2)\Delta}^2 \cdots \rho_{2\Delta}^2 \xrightarrow{p} \infty$ as $t \to \infty$. Thus, $E\left(\left|U_{t\Delta} - \widetilde{\sigma}_0^{-2}\right| | \rho_{(t-1)\Delta}, \cdots, \rho_{2\Delta}\right) \xrightarrow{p} 0$. Since $\{U_{t\Delta}\}$ is bounded,

$E\left|U_{t\Delta} - \widetilde{\sigma}_0^{-2}\right| \xrightarrow{p} 0$ as $t \to \infty$. Consequently,

$$\frac{1}{MN}\sum_{t=1}^{M\times N} \frac{y_{(t-1)\Delta}^2}{\widetilde{\sigma}_0^2 y_{(t-1)\Delta}^2 + \sigma_0^2} \xrightarrow{p} \widetilde{\sigma}_0^{-2}.$$

To unify the results for both stationary and nonstationary cases, we write

$$\frac{1}{MN}\sum_{t=1}^{M\times N} \frac{y_{(t-1)\Delta}^2}{\widetilde{\sigma}_0^2 y_{(t-1)\Delta}^2 + \sigma_0^2} \xrightarrow{p} V_1, \tag{A.4}$$

where V_1 is defined in (4.1).

To examine the asymptotic behavior of the numerator, denote

$$\xi_t := \frac{y_{(t-1)\Delta} v_{t\Delta}}{\sqrt{\widetilde{\sigma}_0^2 y_{(t-1)\Delta}^2 + \sigma_0^2}},$$

and note that ξ_t is a martingale difference sequence with respect to the filtration $\mathcal{F}_t := \sigma(B_v(s) : 0 \le s \le t)$ because

$$E\left(\xi_t|\mathcal{F}_{t-1}\right) = E\left(\left.\frac{y_{(t-1)\Delta} v_{t\Delta}}{\sqrt{\widetilde{\sigma}_0^2 y_{(t-1)\Delta}^2 + \sigma_0^2}}\right| \mathcal{F}_{t-1}\right) = \frac{y_{(t-1)\Delta}}{\sqrt{\widetilde{\sigma}_0^2 y_{(t-1)\Delta}^2 + \sigma_0^2}} E\left(\left. v_{t\Delta}\right| \mathcal{F}_{t-1}\right) = 0.$$

To apply the martingale CLT to $\{\xi_t\}$, we need to check the stability and Lindeberg conditions. For the stability condition, the conditional variance of the standardized martingale is

$$\left\langle \frac{1}{\sqrt{MN}}\sum_{t=1}^{M\times N} \xi_t \right\rangle = \frac{1}{MN}\sum_{t=1}^{M\times N} E\left(\xi_t^2|\mathcal{F}_{t-1}\right) = \frac{1}{MN}\sum_{t=1}^{M\times N} \frac{y_{(t-1)\Delta}^2}{\widetilde{\sigma}_0^2 y_{(t-1)\Delta}^2 + \sigma_0^2} \xrightarrow{a.s.} V_1.$$

For the Lindeberg condition, we have for any $\delta > 0$,

$$\begin{aligned} &E\left[\frac{1}{MN}\sum_{t=1}^{M\times N} E\left\{\xi_t^2 \mathbf{1}\left(|\xi_t| > \sqrt{MN}\delta\right)\middle|\, \mathcal{F}_{t-1}\right\}\right] \\ &\le \sup_t E\left\{\xi_t^2 \mathbf{1}\left\{\frac{y_{(t-1)\Delta}^2 v_{1\Delta}^2}{\widetilde{\sigma}_0^2 y_{(t-1)\Delta}^2 + \sigma_0^2} > MN\delta^2\right\}\right\} \to 0, \end{aligned}$$

since ξ_t^2 is uniformly integrable. By the martingale CLT, as $T \to \infty$, we deduce that

$$\frac{1}{\sqrt{MN}}\sum_{t=1}^{M\times N} \xi_t \xrightarrow{L} \mathcal{N}\left(0, V_1\right).$$

It follows from this result, (A.3) and (A.4) that

$$\sqrt{T}\left(\breve{\widetilde{\mu}} - \widetilde{\mu}_0\right) \xrightarrow{L} \mathcal{N}\left(0, V_1^{-1}\right).$$

Then, by the fact that $A_N = O_p(\Delta B_N)$, $R_A = O_p(\Delta R_B)$ and $B_N + R_B = O_p(T/\Delta)$, we know

$$\begin{aligned}\widehat{\widetilde{\mu}} &= \Delta^{-1}\frac{\widehat{A}_N}{\widehat{B}_N} = \Delta^{-1}\frac{A_N + R_A}{B_N + R_B} = \Delta^{-1}\frac{A_N}{B_N} + \Delta^{-1}\left(\frac{A_N + R_A}{B_N + R_B} - \frac{A_N}{B_N}\right) \\ &= \breve{\widetilde{\mu}} + \Delta^{-1}\frac{R_A B_N - R_B A_N}{B_N(B_N + R_B)} = \breve{\widetilde{\mu}} + O_p\left(\frac{R_B}{B_N + R_B}\right) = \breve{\widetilde{\mu}} + O_p(\sqrt{\Delta}).\end{aligned}$$

Therefore, by assuming $T\Delta \to 0$, we have

$$\sqrt{T}\left(\widehat{\widetilde{\mu}} - \widetilde{\mu}_0\right) = \sqrt{T}\left(\breve{\widetilde{\mu}} - \widetilde{\mu}_0\right) + \sqrt{T}\left(\widehat{\widetilde{\mu}} - \breve{\widetilde{\mu}}\right) \xrightarrow{L} \mathcal{N}\left(0, V_1^{-1}\right). \tag{A.5}$$

Consistency of $\widehat{\widetilde{\mu}}$ follows naturally from the results above.

A.2 Proof of Theorem 4.3

Similar to the previous proof, by equation (2.11) and consistency of $\widehat{\boldsymbol{\theta}}$, we have

$$\begin{aligned}\widehat{\rho} - \rho_0 &= \frac{\sum_{t=1}^{T/\Delta} \frac{y_{(t-1)\Delta}(y_{t\Delta} - \rho_0 y_{(t-1)\Delta})}{\widehat{\widetilde{\sigma}^2} y_{(t-1)\Delta}^2 + \widehat{\sigma^2}}}{\sum_{t=1}^{T/\Delta} \frac{y_{(t-1)\Delta}^2}{\widehat{\widetilde{\sigma}^2} y_{(t-1)\Delta}^2 + \widehat{\sigma^2}}} \\ &= \frac{\sqrt{\Delta}\sum_{t=1}^{T/\Delta} \frac{y_{(t-1)\Delta} v_{t\Delta}}{\sqrt{\widetilde{\sigma}_0^2 y_{(t-1)\Delta}^2 + \sigma_0^2}} + \sqrt{\Delta}\sum_{t=1}^{T/\Delta} y_{(t-1)\Delta} v_{t\Delta}\sqrt{\widetilde{\sigma}_0^2 y_{(t-1)\Delta}^2 + \sigma_0^2}\zeta_{(t-1)\Delta}}{\sum_{t=1}^{T/\Delta} \frac{y_{(t-1)\Delta}^2}{\widetilde{\sigma}_0^2 y_{(t-1)\Delta}^2 + \sigma_0^2} + R_B} \\ &= \frac{\sqrt{\Delta}C_N + R_C}{B_N + R_B},\end{aligned}$$

which leads to the decomposition

$$\frac{1}{\Delta}(\widehat{\rho} - \rho_0) = \frac{\sqrt{\Delta}C_N + R_C}{\Delta B_N + \Delta R_B},$$

where B_N and C_N are defined in the proof of Theorem 4.2, and

$$R_C := \sqrt{\Delta}\sum_{t=1}^{T/\Delta} y_{(t-1)\Delta} v_{t\Delta}\sqrt{\widetilde{\sigma}_0^2 y_{(t-1)\Delta}^2 + \sigma_0^2}\zeta_{(t-1)\Delta} = O_p(\sqrt{\Delta}). \tag{A.6}$$

when T is finite.

In the previous proof, we have shown that $C_N = O_p(\sqrt{\Delta}B_N)$ and one can easily check that $R_C = O_p(\Delta R_B)$ given the order of R_B in the previous proof. Therefore, we have

$$\frac{1}{\Delta}(\widehat{\rho} - \rho_0) = \frac{\sqrt{\Delta}C_N + R_C}{\Delta B_N + \Delta R_B} = \frac{\sqrt{\Delta}C_N}{\Delta B_N} + O_p\left(\frac{R_B}{B_N + R_B}\right) = \frac{\sqrt{\Delta}C_N}{\Delta B_N} + O_p(\sqrt{\Delta}).$$

By the similar arguments to those in the proof of Theorem 4.2,

$$\frac{\Delta}{T}B_N = \frac{\Delta}{T}\sum_{t=1}^{T/\Delta}\frac{y_{(t-1)\Delta}^2}{\widetilde{\sigma}_0^2 y_{(t-1)\Delta}^2 + \sigma_0^2} \xrightarrow{p} V_1, \text{ i.e. } \Delta B_N \xrightarrow{p} TV_1.$$

Further, by the martingale CLT,

$$\sqrt{\frac{\Delta}{T}}C_N = \sqrt{\frac{\Delta}{T}}\sum_{t=1}^{T/\Delta}\frac{y_{(t-1)\Delta}v_{t\Delta}}{\sqrt{\widetilde{\sigma}_0^2 y_{(t-1)\Delta}^2 + \sigma_0^2}} \xrightarrow{L} \mathcal{N}(0, V_1),$$

when $\Delta \to 0$. This is equivalent to $\sqrt{\Delta}C_N \xrightarrow{L} \mathcal{N}(0, TV_1)$. Combining these results gives

$$\frac{1}{\Delta}(\widehat{\rho} - \rho_0) = \frac{\sqrt{\Delta}\cdot C_N}{\Delta \cdot B_N} \xrightarrow{L} \mathcal{N}(0, (TV_1)^{-1}).$$

A.3 Proof of Proposition 4.1

Under the assumption that $\Delta \to 0$ with fixed T, we have

$$\begin{aligned}
\widehat{\beta}_\kappa - \beta_\kappa^0 &= \exp\left(\widehat{\widetilde{\mu}}\Delta + \frac{\widehat{\widetilde{\sigma}^2}\Delta}{2}\right) - \exp\left(\widetilde{\mu}_0\Delta + \frac{\widetilde{\sigma}_0^2\Delta}{2}\right) \\
&= \left(\widehat{\widetilde{\mu}}\Delta - \widetilde{\mu}_0\Delta\right) + \frac{1}{2}\left(\widehat{\widetilde{\sigma}^2}\Delta - \widetilde{\sigma}_0^2\Delta\right) + O_p(\Delta^2) \\
&= \left(\widehat{\widetilde{\mu}}\Delta - \widetilde{\mu}_0\Delta\right) + \frac{\Delta^{3/2}}{2}\left\{\frac{1}{\sqrt{\Delta}}\left(\widehat{\widetilde{\sigma}^2} - \widetilde{\sigma}_0^2\right)\right\} + O_p(\Delta^2) \\
&= \left(\widehat{\widetilde{\mu}}\Delta - \widetilde{\mu}_0\Delta\right) + O_p(\Delta^{3/2}).
\end{aligned}$$

By Theorem 4.3,

$$\frac{1}{\Delta}(\widehat{\rho} - \rho_0) \xrightarrow{L} \mathcal{N}(0, (TV_1)^{-1}). \tag{A.7}$$

Then, by the Taylor expansion, we obtain

$$\widehat{\rho} - \rho_0 = \widehat{\widetilde{\mu}}\Delta - \widetilde{\mu}_0\Delta + O_p(\Delta^2). \tag{A.8}$$

Therefore, by Theorem 4.1 and 4.3 we have

$$\frac{1}{\Delta}\left(\widehat{\beta}_\kappa - \beta_\kappa^0\right) = \frac{1}{\Delta}(\widehat{\rho} - \rho_0) + O_p(\sqrt{\Delta}) \xrightarrow{L} \mathcal{N}(0, (TV_1)^{-1}). \tag{A.9}$$

The same argument yields the asymptotic result for $\widehat{\beta}_\phi$. Details of the proof are omitted.

A.4 Proof of Modified LBI Test Statistic $\widetilde{Z}_N$

Under the null, by Chan et al. (2012), we have the following asymptotic distribution result for $\widetilde{\rho}$,

$$\left(\sum_{t=1}^{M\times N}\frac{y_{(t-1)\Delta}^2}{\delta+y_{(t-1)\Delta}^2}\right)^{-1/2}\left(\sum_{t=1}^{M\times N}\frac{y_{(t-1)\Delta}^2}{(\delta+y_{(t-1)\Delta}^2)^{1/2}}\right)\times(\widetilde{\rho}-\rho_0)\xrightarrow{L}\mathcal{N}\left(0,\,Var(\varepsilon_{t\Delta})\right). \quad \text{(A.10)}$$

Then, we know for $y_{(t-1)\Delta}$, no matter it is stationary or nonstationary, we have

$$\widetilde{\varepsilon}_{t\Delta}-\varepsilon_{t\Delta}=(\widetilde{\rho}-\rho_0)y_{(t-1)\Delta}=o_p(1).$$

Note $\widetilde{y}_{t\Delta}$ is always stationary, by WLLN and the ergodic theorem, we can easily show that, for any $p\in Z_+$ such that $p\leq 4$,

$$\frac{1}{MN}\sum_{t=1}^{M\times N}\widetilde{\varepsilon}_{t\Delta}^p\xrightarrow{p}E(\varepsilon_{t\Delta}^p),\quad \frac{1}{MN}\sum_{t=1}^{M\times N}\widetilde{y}_{t\Delta}^p\xrightarrow{a.s.}E(\widetilde{y}_{t\Delta}^p).$$

Therefore, for the denominator, we have

$$\sqrt{\frac{1}{MN}\sum_{t=1}^{M\times N}\widetilde{\varepsilon}_{t\Delta}^4-\left(\frac{1}{MN}\sum_{t=1}^{M\times N}\widetilde{\varepsilon}_{t\Delta}^2\right)^2}\xrightarrow{p}\sqrt{E(\varepsilon_{t\Delta}^4)-E(\varepsilon_{t\Delta}^2)^2}=Std(\varepsilon_{t\Delta}^2),$$

$$\sqrt{\frac{1}{MN}\sum_{t=1}^{M\times N}\widetilde{y}_{t\Delta}^4-\left(\frac{1}{MN}\sum_{t=1}^{M\times N}\widetilde{y}_{t\Delta}^2\right)^2}\xrightarrow{a.s.}\sqrt{E(\widetilde{y}_{t\Delta}^4)-E(\widetilde{y}_{t\Delta}^2)^2}=Std(\widetilde{y}_{t\Delta}^2).$$

For the numerator, denote $\tilde{\xi}_{t\Delta}=\widetilde{\varepsilon}_{t\Delta}^2-\left(\frac{1}{MN}\sum_{t=1}^{M\times N}\widetilde{\varepsilon}_{t\Delta}^2\right)$ and $\xi_{t\Delta}=\varepsilon_{t\Delta}^2-\left(\frac{1}{MN}\sum_{t=1}^{M\times N}\varepsilon_{t\Delta}^2\right)$. We know $E(\tilde{\xi}_{t\Delta})=0=E(\xi_{t\Delta})$ and $Var(\tilde{\xi}_{t\Delta})=Var(\widetilde{\varepsilon}_{t\Delta}^2)\xrightarrow{p}Var(\varepsilon_{t\Delta}^2)$.

$$\frac{1}{\sqrt{MN}}\sum_{t=1}^{M\times N}\widetilde{y}_{(t-1)\Delta}^2\tilde{\xi}_{t\Delta}=\frac{1}{\sqrt{MN}}\sum_{t=1}^{M\times N}\widetilde{y}_{(t-1)\Delta}^2(\tilde{\xi}_{t\Delta}-\xi_{t\Delta})+\frac{1}{\sqrt{MN}}\sum_{t=1}^{M\times N}\widetilde{y}_{(t-1)\Delta}^2\xi_{t\Delta}.$$

By equation (3.3) in Lee (1998), one can easily show

$$\frac{1}{\sqrt{MN}}\sum_{t=1}^{M\times N}\widetilde{y}_{(t-1)\Delta}^2(\tilde{\xi}_{t\Delta}-\xi_{t\Delta})=o_p(1),$$

and by applying the martingale CLT (cf. Hall and Heyde, 1980), we have

$$\frac{1}{\sqrt{MN}}\sum_{t=1}^{M\times N}\widetilde{y}_{(t-1)\Delta}^2\xi_{t\Delta}\xrightarrow{L}\mathcal{N}\left(0,\,Var(\widetilde{y}_{t\Delta}^2)\,Var(\varepsilon_{t\Delta}^2)\right).$$

Then, by combining the results above, we can derive the asymptotic distribution of $\widetilde{Z}_N$ under $H_0:\widetilde{\sigma}_0^2=0$, i.e.,

$$\frac{1}{\sqrt{MN}}\widetilde{Z}_N\xrightarrow{L}\mathcal{N}(0,1).$$

Lastly, under the alternative, one just need to note that $Cov(\varepsilon_{t\Delta}^2,y_{(t-1)\Delta}^2)$ diverges when $\widetilde{\sigma}_0^2\neq 0$, and this leads to the divergence of $\widetilde{Z}_N$.

A.5 Proof of Proposition 5.1

It has been proved in Föllmer et al. (1994) that y_t is strictly stationary and ergodic when $\widetilde{\mu} - \frac{1}{2}\widetilde{\sigma}^2 < 0$. This means that we can still characterize explosiveness using $\phi := \widetilde{\mu} - \frac{1}{2}\widetilde{\sigma}^2$. However, for characterizing instability, we need to calculate the second moment of y_t.

The expectation of $J(t)$ is

$$\begin{aligned} EJ(t) &= -\omega \int_0^t E\left(\exp\left\{\left(\widetilde{\mu} - \frac{1}{2}\widetilde{\sigma}^2\right)(t-s) + \widetilde{\sigma}\left(B_u(t) - B_u(s)\right)\right\}\right) ds \\ &= -\omega \int_0^t \exp\left\{\widetilde{\mu}(t-s)\right\} ds = \frac{\omega}{\widetilde{\mu}}(1 - \exp(\widetilde{\mu}t)), \end{aligned}$$

and so $EJ(t)$ is finite as $t \to \infty$ if and only if $\widetilde{\mu} < 0$. Further, to figure out the order of $Var(J(t))$, we apply the Cauchy-Schwartz inequality to $EJ(t)^2$, giving

$$EJ(t)^2 = E[K(t) - L(t)]^2 \leq 2EK(t)^2 + 2EL(t)^2$$

This inequality indicate that we only need to calculate $EK(t)^2$ and $EL(t)^2$ to evaluate the asymptotic order of $EJ(t)^2$. By Itō's isometry

$$EK(t)^2 = \sigma^2 \int_0^t E\left(\exp\left\{\left(2\widetilde{\mu} - \widetilde{\sigma}^2\right)(t-s) + 2\widetilde{\sigma}\left(B_u(t) - B_u(s)\right)\right\}\right) ds = \sigma^2 \frac{e^{2\kappa t} - 1}{2\kappa},$$

$$\begin{aligned} EL(t)^2 &= \omega^2 E\left(\int_0^t \exp\left\{\left(\widetilde{\mu} - \frac{1}{2}\widetilde{\sigma}^2\right)(t-s) + \widetilde{\sigma}\left(B_u(t) - B_u(s)\right)\right\} ds\right)^2 \\ &= \omega^2 E\left(\int_0^t \int_0^t \exp\left\{\left(\widetilde{\mu} - \frac{1}{2}\widetilde{\sigma}^2\right)(2t - s - r) + \widetilde{\sigma}\left(2B_u(t) - B_u(s) - B_u(r)\right)\right\} dsdr\right) \\ &= \omega^2 \int_0^t \int_0^t \exp\left\{\left(\widetilde{\mu} - \frac{1}{2}\widetilde{\sigma}^2\right)(2t - s - r) + \frac{1}{2}\widetilde{\sigma}^2\left(2t - s - r + 2\min\{t-s, t-r\}\right)\right\} dsdr \\ &= \omega^2 \int_0^t \int_0^r \exp\left\{\left(\widetilde{\mu} - \frac{1}{2}\widetilde{\sigma}^2\right)(2t - s - r) + \frac{1}{2}\widetilde{\sigma}^2\left(4t - s - 3r\right)\right\} dsdr \\ &\quad + \omega^2 \int_0^t \int_r^t \exp\left\{\left(\widetilde{\mu} - \frac{1}{2}\widetilde{\sigma}^2\right)(2t - s - r) + \frac{1}{2}\widetilde{\sigma}^2\left(4t - 3s - r\right)\right\} dsdr \\ &= \omega^2 \left(\int_0^t \frac{(e^{\widetilde{\mu}r} - 1)e^{2\kappa(t-r)}}{\widetilde{\mu}} dr + \int_0^t \frac{e^{\widetilde{\mu}(t-r)}(e^{(\widetilde{\mu}+\widetilde{\sigma}^2)(t-r)} - 1)}{\widetilde{\mu} + \widetilde{\sigma}^2} dr\right) \\ &= \omega^2 \left(\frac{\widetilde{\mu}(e^{2\kappa t} - 2e^{\widetilde{\mu}t} + 1) - \widetilde{\sigma}^2(e^{\widetilde{\mu}t} - 1)}{2\widetilde{\mu}\kappa(\widetilde{\mu} + \widetilde{\sigma}^2)} + \frac{1 - e^{\widetilde{\mu}t}}{\widetilde{\mu}(\widetilde{\mu} + \widetilde{\sigma}^2)} + \frac{e^{2\kappa t} - 1}{2\kappa(\widetilde{\mu} + \widetilde{\sigma}^2)}\right) \\ &= \omega^2 \left(\frac{\widetilde{\mu}e^{2\kappa t} - 2\kappa e^{\widetilde{\mu}t} + \widetilde{\mu} + \widetilde{\sigma}^2}{\widetilde{\mu}\kappa(\widetilde{\mu} + \widetilde{\sigma}^2)}\right) \end{aligned}$$

Note that for $\kappa < 0$

$$\lim_{t\to\infty} EJ(t)^2 \leq \lim_{t\to\infty} 2EK(t)^2 + 2EL(t)^2 = \frac{2\omega^2}{\widetilde{\mu}\kappa} - \frac{\sigma^2}{\kappa} < \infty, \tag{A.11}$$

showing that, when $\kappa < 0$, $J(t)$ has finite second-order moments as $t \to \infty$. Further, for $\kappa \to 0$, by L'Hôpital's rule, we have

$$\lim_{\kappa \to 0} EK(t)^2 = \lim_{\kappa \to 0} \sigma^2 \frac{2te^{2\kappa t}}{2} = \sigma^2 t, \tag{A.12}$$

$$\begin{aligned} \lim_{\kappa \to 0} EL(t)^2 &= \lim_{\widetilde{\mu} \to -\widetilde{\sigma}^2/2} \omega^2 \left(\frac{\widetilde{\mu} e^{2\kappa t} - 2\kappa e^{\widetilde{\mu} t} + \widetilde{\mu} + \widetilde{\sigma}^2}{\widetilde{\mu}\kappa(\widetilde{\mu} + \widetilde{\sigma}^2)} \right) \\ &= 4\gamma^2 \sigma^2 t + \frac{2(e^{-\frac{1}{2}\widetilde{\sigma}^2 t} - 1)}{\widetilde{\sigma}^2}. \end{aligned} \tag{A.13}$$

Combining results (A.12) and (A.13), we obtain $EJ(t)^2 = O_p(t)$, as $t \to \infty$.

Lastly, for $\kappa > 0$, note that $\kappa = \widetilde{\mu} + \frac{1}{2}\widetilde{\sigma}^2 \geq \widetilde{\mu}$, we then have

$$EK(t)^2 = \sigma^2 \frac{e^{2\kappa t} - 1}{2\kappa} = O_p(e^{2\kappa t}) \tag{A.14}$$

$$EL(t)^2 = \omega^2 \left(\frac{\widetilde{\mu} e^{2\kappa t} - 2\kappa e^{\widetilde{\mu} t} + \widetilde{\mu} + \widetilde{\sigma}^2}{\widetilde{\mu}\kappa(\widetilde{\mu} + \widetilde{\sigma}^2)} \right) = O_p(e^{2\kappa t}) \tag{A.15}$$

which leads to $EJ(t)^2 = O_p(e^{2\kappa t})$.

From the results above we know that $EJ(t)$ is finite if and only if $\widetilde{\mu} < 0$, and $EJ(t)^2$ is finite if and only if $\kappa < 0$. Since $\kappa < 0$ implies $\widetilde{\mu} < 0$, $Var(J(t)) < \infty$ if and only if $\kappa < 0$. We can now work out the first two moments of y_t,

$$Ey_t = E\left[\exp\left(\widetilde{\sigma} B_u(t) + \left(\widetilde{\mu} - \frac{1}{2}\widetilde{\sigma}^2 \right) t \right) \right] y_0 + EJ(t) = e^{\widetilde{\mu} t} y_0 + \frac{\omega}{\widetilde{\mu}} (1 - \exp(\widetilde{\mu} t)),$$

$$\begin{aligned} Ey_t^2 &= E\left[\exp\left(2\widetilde{\sigma} B_u(t) + \left(2\widetilde{\mu} - \widetilde{\sigma}^2 \right) t \right) \right] Ey_0^2 \\ &\quad + 2Ey_0 E\left[\exp\left(\widetilde{\sigma} B_u(t) + \left(\widetilde{\mu} - \frac{1}{2}\widetilde{\sigma}^2 \right) t \right) J(t) \right] + EJ(t)^2 \\ &= e^{2\kappa t} Ey_0^2 - 2Ey_0 E\left(\omega \int_0^t \exp\left\{ \left(\widetilde{\mu} - \frac{1}{2}\widetilde{\sigma}^2 \right)(2t - s) + \widetilde{\sigma}\left(2B_u(t) - B_u(s) \right) \right\} ds \right) + EJ(t)^2 \\ &= e^{2\kappa t} Ey_0^2 - 2\omega \int_0^t E\left(\exp\left\{ \left(\widetilde{\mu} - \frac{1}{2}\widetilde{\sigma}^2 \right)(2t - s) + \widetilde{\sigma}\left(2B_u(t) - B_u(s) \right) \right\} \right) ds Ey_0 + EJ(t)^2 \\ &= e^{2\kappa t} y_0^2 - 2\omega \int_0^t \exp\left(2\kappa t - \left(\widetilde{\mu} + \widetilde{\sigma}^2 \right) s \right) ds Ey_0 + EJ(t)^2 \\ &= e^{2\kappa t} Ey_0^2 - 2\omega \frac{e^{2\kappa t} - e^{\widetilde{\mu} t}}{\widetilde{\mu} + \widetilde{\sigma}^2} Ey_0 + EJ(t)^2. \end{aligned}$$

Evidently from these expressions Ey_t is asymptotically finite if and only if $\widetilde{\mu} < 0$, and Ey_t^2 is asymptotically finite if and only if $\kappa < 0$. This indicates that $Var(y_t) < \infty$ if and only if $\kappa < 0$. Therefore, we can still characterize instability with $\kappa \geq 0$ and locally explosiveness with $\widetilde{\mu} \geq 0$.

A.6 Proof of Remark 5.2

Denote $X_n = \dfrac{\log RV_n - \log[y]^{nh}_{(n-1)h} + \frac{1}{2}s_n^2}{s_n}$, where $s_n = \max\left\{\sqrt{2\Delta\dfrac{RQ_n}{RV_n^2}}, \sqrt{\dfrac{2}{M}}\right\}$. According to Barndorff-Nielsen and Shephard (2005), $\{X_n\}_{n=1}^N \xrightarrow{L} \mathcal{N}(0,1)$ as $\Delta \to 0$. Note $N = \frac{T}{M\Delta}$, so when $\Delta \to 0$ with T, M being finite, we have $N \to \infty$. Therefore, the log-likelihood function for $\theta = (\widetilde{\sigma}^2, \gamma, \sigma^2)$ is given by

$$\ell_{ur}(\theta) = -\frac{N}{2}\log 2\pi - \frac{1}{2}\sum_{n=1}^{N} X_n(\theta)^2 + O(\Delta). \tag{A.16}$$

As $\ell_{ur}(\theta)$ is based on the standard normal distribution, Wilks's theorem applies in this case, i.e. under $\mathcal{H}_0 : \gamma_0 = 0$, as $N \to \infty$,

$$\begin{aligned} LR &= -2\left(\ell_r - \ell_{ur}\right) = \sum_{n=1}^{N} X_n(\theta_0)^2 - \sum_{n=1}^{N} X_n(\theta)^2 + o_p(1) \\ &= \Delta^{-1}\left(Q_\Delta(\theta_0) - Q_\Delta(\theta)\right) + o_p(1) \xrightarrow{L} \chi^2(1), \end{aligned}$$

where $\theta_0 = \left(\widetilde{\sigma}_0^2, \gamma_0, \sigma_0^2\right)$.

A.7 Proof of Theorem 5.2

The dependence of B_u and B_ε leads to a complex relationship among $y_{(t-1)\Delta}$, $\rho_{t\Delta}$ and $J_{t\Delta}$ in model (5.4). Without loss of generality, we know that y_t can also be viewed as generated from model (5.3) by virtue of the observational equivalence of these mechanisms. Note that the approximate representation of the exact discretized model of (5.3) is

$$y_{t\Delta} = \exp\{\widetilde{\mu}_0\Delta\}y_{(t-1)\Delta} + \sqrt{(\widetilde{\sigma}_0^2 y_{(t-1)\Delta}^2 + 2\omega_0 y_{(t-1)\Delta} + \sigma_0^2)\Delta}\cdot v_{t\Delta}, \tag{A.17}$$

and by the Taylor expansion of $\exp\{\widetilde{\mu}_0\Delta\}$, we know

$$y_{t\Delta} - y_{(t-1)\Delta} = (\widetilde{\mu}_0\Delta + o(\Delta))y_{(t-1)\Delta}\cdot + \sqrt{(\widetilde{\sigma}_0^2 y_{(t-1)\Delta}^2 + 2\omega_0 y_{(t-1)\Delta} + \sigma_0^2)\Delta}\cdot v_{t\Delta}, \tag{A.18}$$

where $v_{t\Delta} \overset{i.i.d}{\sim} \mathcal{N}(0,1)$.

Similar to the proof of Theorem 4.2, the limiting distribution of the feasible estimator will converge to that of the infeasible estimator given $T\Delta \to 0$. Therefore, we only need to figure out the distribution of the infeasible estimator. According to (5.9), we have

$$\begin{aligned} A_N^* &= \sum_{t=1}^{M\times N} \frac{y_{(t-1)\Delta}\left(y_{t\Delta} - y_{(t-1)\Delta}\right)}{\widetilde{\sigma}_0^2 y_{(t-1)\Delta}^2 + 2\omega_0 y_{(t-1)\Delta} + \sigma_0^2} \\ &= \widetilde{\mu}_0\Delta \sum_{t=1}^{M\times N} \frac{y_{(t-1)\Delta}^2}{\widetilde{\sigma}_0^2 y_{(t-1)\Delta}^2 + 2\omega_0 y_{(t-1)} + \sigma_0^2} + \sqrt{\Delta}\sum_{t=1}^{M\times N} \frac{y_{(t-1)\Delta} v_{t\Delta}}{\sqrt{\widetilde{\sigma}_0^2 y_{(t-1)\Delta}^2 + 2\omega_0 y_{(t-1)\Delta} + \sigma_0^2}} + o_p(\Delta B_N^*) \\ &= \widetilde{\mu}_0\Delta B_N^* + \sqrt{\Delta}C_N^* + o_p(\Delta B_N^*), \end{aligned}$$

where $C_N^* = \sum_{t=1}^{M\times N} \frac{y_{(t-1)\Delta} v_{t\Delta}}{\sqrt{\widetilde{\sigma}_0^2 y_{(t-1)\Delta}^2 + 2\omega_0 y_{(t-1)\Delta} + \sigma_0^2}}$. This leads to

$$\breve{\widetilde{\mu}} = \widetilde{\mu}_0 + \frac{C_N^*}{\sqrt{\Delta} B_N^*} + o_p(1), \tag{A.19}$$

and hence,

$$\sqrt{T}\left(\breve{\widetilde{\mu}} - \widetilde{\mu}_0\right) = \frac{\frac{1}{\sqrt{MN}} \sum_{t=1}^{M\times N} \frac{y_{(t-1)\Delta} v_{t\Delta}}{\sqrt{\widetilde{\sigma}_0^2 y_{(t-1)\Delta}^2 + 2\omega_0 y_{(t-1)\Delta} + \sigma_0^2}}}{\frac{1}{MN} \sum_{t=1}^{M\times N} \frac{y_{(t-1)\Delta}^2}{\widetilde{\sigma}_0^2 y_{(t-1)\Delta}^2 + 2\omega_0 y_{(t-1)\Delta} + \sigma_0^2}} + o_p(1). \tag{A.20}$$

Note that $\frac{y_{(t-1)\Delta}^2}{\widetilde{\sigma}_0^2 y_{(t-1)\Delta}^2 + 2\omega_0 y_{(t-1)\Delta} + \sigma_0^2}$ is bounded above by $\widetilde{\sigma}_0^{-2}$. Similarly, by the same argument as in the proof of Theorem 4.2, we have

$$\frac{1}{MN} \sum_{t=1}^{M\times N} \frac{y_{(t-1)\Delta}^2}{\widetilde{\sigma}_0^2 y_{(t-1)\Delta}^2 + 2\omega_0 y_{(t-1)\Delta} + \sigma_0^2} \xrightarrow{p} V_2,$$

where V_2 is defined in (5.11).

Further, denote

$$\xi_t := \frac{y_{(t-1)\Delta} v_{t\Delta}}{\sqrt{\widetilde{\sigma}_0^2 y_{(t-1)\Delta}^2 + 2\omega_0 y_{(t-1)\Delta} + \sigma_0^2}}$$

and observe that ξ_t is a martingale difference sequence with respect to the filtration $\mathcal{F}_t := \sigma(B_v(s) : 0 \le s \le t)$ as

$$\begin{aligned} E\left(\xi_t | \mathcal{F}_{t-1}\right) &= E\left(\frac{y_{(t-1)\Delta} v_{t\Delta}}{\sqrt{\widetilde{\sigma}_0^2 y_{(t-1)\Delta}^2 + 2\omega_0 y_{(t-1)\Delta} + \sigma_0^2}} \middle| \mathcal{F}_{t-1} \right) \\ &= \frac{y_{(t-1)\Delta}}{\sqrt{\widetilde{\sigma}_0^2 y_{(t-1)\Delta}^2 + 2\omega_0 y_{(t-1)\Delta} + \sigma_0^2}} E\left(v_{t\Delta} | \mathcal{F}_{t-1}\right) = 0. \end{aligned}$$

To apply the martingale CLT, we check the stability condition and the Lindeberg condition. For the stability condition, we have

$$\left\langle \frac{1}{\sqrt{MN}} \sum_{t=1}^{M\times N} \xi_t \right\rangle = \frac{1}{MN} \sum_{t=1}^{M\times N} E\left(\frac{y_{(t-1)\Delta}^2}{\widetilde{\sigma}_0^2 y_{(t-1)\Delta}^2 + 2\omega_0 y_{(t-1)\Delta} + \sigma_0^2} \middle| \mathcal{F}_{t-1} \right) \xrightarrow{a.s.} V_2.$$

For the Lindeberg condition, we have for any $\delta > 0$

$$\begin{aligned} &\frac{1}{MN} \sum_{t=1}^{M\times N} E\left\{ \xi_t^2 \mathbf{1}\left(|\xi_t| > \sqrt{MN}\delta \right) \middle| \mathcal{F}_{t-1} \right\} \\ &\le \sup_t E\left\{ \xi_t^2 \mathbf{1}\left\{ \frac{y_{(t-1)\Delta}^2 v_{1\Delta}^2}{\widetilde{\sigma}_0^2 y_{(t-1)\Delta}^2 + 2\omega_0 y_{(t-1)\Delta} + \sigma_0^2} > MN\delta^2 \right\} \middle| \mathcal{F}_{t-1} \right\} \to 0, \end{aligned}$$

since ξ_t^2 is uniformly integrable and $MN \to \infty$. From the martingale CLT, as $T \to \infty$,

$$\sqrt{T}\left(\breve{\widetilde{\mu}} - \widetilde{\mu}_0\right) \xrightarrow{L} \mathcal{N}\left(0, V_2^{-1}\right),$$

and hence,

$$\sqrt{T}\left(\widehat{\widetilde{\mu}} - \widetilde{\mu}_0\right) \xrightarrow{L} \mathcal{N}\left(0, V_2^{-1}\right).$$

Consistency of $\widehat{\widetilde{\mu}}$ follows naturally from the results above.

A.8 Proof of Theorem 5.3

Similar to the proof of Theorem 4.3 by substituting equation (A.18) into $\widehat{\rho}$, we obtain

$$\widehat{\rho} - \rho_0 = \widehat{\widetilde{\mu}}\Delta - \widetilde{\mu}_0\Delta + o_p(\Delta) = \frac{A_N^*}{B_N^*} - \widetilde{\mu}_0\Delta + o_p(\Delta) = \frac{\sqrt{\Delta}C_N^*}{B_N^*} + o_p(\Delta).$$

Then, by the same arguments as those in the proof of Theorem 4.2, we have

$$\frac{\Delta}{T}B_N^* = \frac{\Delta}{T}\sum_{t=1}^{T/\Delta} \frac{y_{(t-1)\Delta}^2}{\widetilde{\sigma}_0^2 y_{(t-1)\Delta}^2 + 2\omega_0 y_{(t-1)\Delta} + \sigma_0^2} \xrightarrow{p} V_2, \text{ i.e. } \Delta B_N^* \xrightarrow{p} TV_2,$$

Further, as proved in the previous section, $\sqrt{\Delta/T}C_N^* \xrightarrow{L} \mathcal{N}(0, V_2)$ by the martingale CLT, which gives $\sqrt{\Delta}C_N^* \xrightarrow{L} \mathcal{N}(0, TV_2)$. Combining these results gives

$$\frac{1}{\Delta}\left(\widehat{\rho} - \rho_0\right) = \frac{\sqrt{\Delta}C_N^*}{\Delta B_N^*} + o_p(1) \xrightarrow{L} \mathcal{N}(0, (TV_2)^{-1}).$$

References

Aue, A. (2008): "Near-integrated Random Coefficient Autoregressive Time Series," *Econometric Theory*, 24, 1343–1372.

Aue, A. and L. Horváth (2011): "Quasi-likelihood estimation in stationary and nonstationary autoregressive models with random coefficients," *Statistica Sinica*, 21, 973–.

Aue, A., L. Horváth, and J. Steinebach (2006): "Estimation in random coefficient autoregressive models," *Journal of Time Series Analysis*, 27, 61–76.

Banerjee, A. N., G. Chevillon, and M. Kratz (2014): "Detecting and Forecasting Bubbles in a Near-Explosive Random Coefficient Model," *Working Paper, ESSEC Business School*.

Banerjee, A. N., G. Chevillon, and M. Kratz (2017): "Probabilistic Forecasting of Bubbles and Crashes," *Working Paper, ESSEC Business School*.

Barndorff-Nielsen, O. E. (2002): "Econometric Analysis of Realized Volatility and Its Use in Estimating Stochastic Volatility Models," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 64, 253–280.

Barndorff-Nielsen, O. E. and N. Shephard (2002): "Estimating Quadratic Variation Using Realized Variance," *Journal of Applied econometrics*, 17, 457–477.

Barndorff-Nielsen, O. E. and N. Shephard (2005): "How Accurate is the Asymptotic Approximation to the Distribution of Realized Variance," *Identification and inference for econometric models. A Festschrift in honour of TJ Rothenberg*, 306–311.

Berkes, I., L. Horváth, and S. Ling (2009): "Estimation in nonstationary random coefficient autoregressive models," *Journal of Time Series Analysis*, 30, 395–416.

Bykhovskaya, A. and P. C. B. Phillips (2018): "Boundary Limit Theory for Functional Local to Unity Regression," *Journal of Time Series Analysis*, 39, 523–562.

Bykhovskaya, A. and P. C. B. Phillips (2019): "Point Optimal Testing with Roots that are Functionally Local to Unity," *Journal of Econometrics*, forthcoming.

Cavaliere, G., I. Georgiev, and A. M. R. Taylor (2016): "Sieve-based Inference for Infinite-variance Linear Processes," *Annals of Statistics*, 44(4), 1467–1494.

Chan, N. H., D. Li, L. Peng (2012): "Toward a Unified Interval Estimation of Autoregressions," *Econometric Theory*, 28, 705–717.

Chen, M., D. Li, and S. Ling (2014): "Non-Stationarity and Quasi-Maximum Likelihood Estimation on a Double Autoregressive Model," *Journal of Time Series Analysis*, 35, 189–202.

Chen, Y., P. C. B. Phillips, and J. Yu (2017): "Inference in Continuous Systems with Mildly Explosive Regressors," *Journal of Econometrics*, 201, 400–416.

Chong, T. T.-L. (2001): "Structural Change in AR (1) Models," *Econometric Theory*, 87–155.

Diba, B. T. and H. I. Grossman (1988): "Explosive Rational Bubbles in Stock Prices?" *The American Economic Review*, 78, 520–530.

Föllmer, H., W. Cheung, and M. A. H. Dempster (1994): "Stock Price Fluctuation as a Diffusion in a Random Environment [and discussion]," *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 347, 471–483.

Föllmer, H. and M. Schweizer (1993): "A Microeconomic Approach to Diffusion Models for Stock Prices," *Mathematical Finance*, Vol. 3, 1–23.

Francq, C. and J.-M. Zakoïan (2012): "Strict Stationarity Testing and Estimation of Explosive and Stationary Generalized Autoregressive Conditional Heteroscedasticity Models," *Econometrica*, 80, 821–861.

Giraitis, L., G. Kapetanios, and T. Yates (2014): "Inference on Stochastic Time Varying Coefficient Models," *Journal of Econometrics*, 179, 46–65.

Granger, C. W. J. (1980): "Long Memory Relationships and the Aggregation of Dynamic Models," *Journal of Econometrics*, 14, 227–238.

Granger, C. W. J. and N. R. Swanson (1997): "An Introduction to Stochastic Unit-root Processes," *Journal of Econometrics*, 80, 35–62.

Hill, J. and Liang Peng (2014): "Unified Interval Estimation for Random Coefficient Autoregressive Models," *Journal of Time Series Analysis*, 35, 282-297.

Horváth, Lajos and Lorenso Trapani (2016): "Statistical Inference in Random Coefficient Panel Model," *Journal of Econometrics*, 193, 54-75.

Höpfner, Reinhard, and Yu Kutoyants (2003): "On a Problem of Statistical Inference in Null Recurrent Diffusions," *Statistical Inference for Stochastic Processes*, 6.1, 25-42.

Hwang, S. Y. and I. V. Basawa (1997): "The Local Asymptotic Normality of a Class of Generalized Random Coefficient Autoregressive Processes," *Statistics & Probability Letters*, 34, 165–170.

Hwang, S. Y. and I. V. Basawa (1998): "Parameter Estimation for Generalized Random Coefficient Autoregressive Processes," *Journal of Statistical Planning and Inference*, 68, 323–337.

Hwang, S. Y. and I. V. Basawa (2005): "Explosive Random-Coefficient AR (1) Processes and Related Asymptotics for Least-Squares Estimation," *Journal of Time Series Analysis*, 26, 807–824.

Jensen, S. T. and A. Rahbek (2004): "Asymptotic Normality of the QMLE Estimator of ARCH in the Nonstationary Case," *Econometrica*, 72, 641–646.

Jiang, L., X. Wang, and J. Yu (2017): "In-fill Asymptotic Theory for Structural Break Point in Autoregression: A Unified Theory," *Working Paper, Singapore Management University.*

Kim, J. and J. Park (2016): "Mean Reversion and Stationarity in Financial Time Series Generated from Diffusion Models," *Working Paper, Indiana University.*

Kristensen, D. (2012): "Nonparametric Detection and Estimation of Structural Change," *Econometrics Journal*, 15, 420–460.

Lee, S. (1998): "Coefficient Constancy Test in a Random Coefficient Autoregressive Model," *Journal of Statistical Planning and Inference*, 74(1), 93-101.

Lieberman, O. and P. C. B. Phillips (2014): "Norming Rates and Limit Theory for Some Time-Varying Coefficient Autoregressions," *Journal of Time Series Analysis*, 35, 592–623.

Lieberman, O. and P. C. B. Phillips (2017a): "IV and GMM Estimation and Testing of Multivariate Stochastic Unit Root Models." *Econometric Theory*, forthcoming.

Lieberman, O. and P. C. B. Phillips (2017b): "A Multivariate Stochastic Unit Root Model with an Application to Derivative Pricing," *Journal of Econometrics*, 196, 99–110.

Lieberman, O. and P. C. B. Phillips (2017c): "Hybrid Stochastic Local Unit Roots," *Working Paper, Yale University.*

Ling, S. and D. Li (2008): "Asymptotic Inference for a Nonstationary Double AR(1) Model," *Biometrika*, 95, 257–263.

Nagakura, D.(2009): "Asymptotic Theory for Explosive Random Coefficient Autoregressive Models and Inconsistency of a Unit Root Test against a Stochastic Unit Root Process," *Statistics and Probability Letters*, 76, 2476-2483.

Nicholls, D. and B. Quinn (1980): "The Estimation of Random Coefficient Autoregressive Models. I," *Journal of Time Series Analysis*, 1, 37–46.

Pang, T., D. Zhang, and T. T.-L. Chong (2014): "Asymptotic Inferences for an AR (1) Model with a Change Point: Stationary and Nearly Non-Stationary Cases," *Journal of Time Series Analysis*, 35, 133–150.

Phillips, P. C. B. (1987): "Towards a Unified Asymptotic Theory for Autoregression," *Biometrika*, 74(3), 535–547.

Phillips, P. C. B. (2012): "Estimation of the Localizing Rate for Mildly Integrated and Mildly Explosive Processes," *Working Paper, Yale University*.

Phillips, P. C. B. and T. Magdalinos (2007): "Limit Theory for Moderate Deviations from a Unit Root," *Journal of Econometrics*, 136, 115–130.

Phillips, P. C. B. and T. Magdalinos (2009): "Unit Root and Cointegrating Limit Theory When Initialization is in the Infinite Past," *Econometrics Theory*, 25(6), 1682-1715.

Phillips, P. C. B. and S.-P. Shi (2017): "Financial Bubble Implosion and Reverse Regression," *Econometric Theory*, 2017, 1-49.

Phillips, P. C. B., S.-P. Shi, and J. Yu (2015a): "Testing for Multiple Bubbles: Limit Theory of Real Time Detectors," *International Economic Review*, 56, 1079-1134.

Phillips, P. C. B., S.-P. Shi, and J. Yu (2015b): "Testing for Multiple Bubbles: Historical Episodes of Exuberance and Collapse in the S&P 500," *International Economic Review*, 56, 1043-1078.

Phillips, P. C. B., Y. Wu, and J. Yu (2011): "Explosive Behavior in the 1990s NASDAQ: When did Exuberance Escalate Asset Values?" *International economic review*, 52, 201–226.

Phillips, P. C. B. and J. Yu (2009): "A Two-stage Realized Volatility Approach to Estimation of Diffusion Processes with Discrete Data," *Journal of Econometrics*, 150, 139–150.

Phillips, P. C. B. and J. Yu (2011): "Dating the Timeline of Financial Bubbles during the Subprime Crisis," *Quantitative Economics*, 2, 455–491.

So, W. S. and D. W. Shin (1999): "Cauchy Estimators for Autoregressive Processes with Applications to Unit Root Tests and Confident Intervals," *Econometric Theory*, 15, 165–176.

Tao, Y., P. C. B. Phillips and J. Yu (2018): "Online Supplement to: 'Random Coefficient Continuous Systems: Testing for Extreme Sample Behavior,"' *Working Paper Supplement*.

Wang, X. and J. Yu (2016): "Double Asymptotics for Explosive Continuous Time Models," *Journal of Econometrics*, 193, 35–53.

Yu, J. (2012): "Bias in the Estimation of the Mean Reversion Parameter in Continuous Time Models," *Journal of Econometrics*, 169, 114–122.

Zhao, Z. and D. Wang (2012): "Statistical Inference for Generalized Random Coefficient Autoregressive Model," *Mathematical and Computer Modelling*, 56, 152–166.