Econometrics Journal (2018), volume **00**, pp. 1–32.
Article No. ectj??????

Testing for Moderate Explosiveness

Gangzheng Guo[†], Yixiao Sun[‡] and Shaoping Wang[†]

[†] *School of Economics, Huazhong University of Science and Technology, Wuhan, China.*
E-mail: gangzheng_guo@hust.edu.cn, wangshaoping@hust.edu.cn (correspondence)
[‡] *Department of Economics, University of California, San Diego, California, USA.*
E-mail: yisun@ucsd.edu
Received:

Summary This paper considers a moderately explosive AR(1) process where the autoregressive root approaches unity from the right at a certain rate. We first develop a test for the null of moderate explosiveness under independent and identically distributed errors. We show that the t statistic is asymptotically standard normal regardless of whether the true process is dominated by the stochastic moderately-explosive trend or the deterministic nonlinear drift trend. This result is in sharp contrast with the existing literature wherein nonstandard limiting distributions are obtained under different model assumptions. When the errors are weakly dependent, we show that the t statistic based on a heteroskedasticity and autocorrelation robust standard error follows Student's t distribution in large samples. Monte Carlo simulations show that our tests have satisfactory size and power performances in finite samples. Applying the asymptotic t test to ten major stock indexes in the pre-2008 financial exuberance period, we find that most indexes are only mildly explosive or not explosive at all, which implies that the bout of the irrational rise was not as serious as previously thought.

Keywords: *Heteroskedasticity and autocorrelation robust standard error, Irrational exuberance, Local to unity, Moderate explosiveness, Student's t distribution, Unit root.*

1. INTRODUCTION

Explosive processes have attracted much recent attention. Phillips and Magdalinos (2007a) consider moderately explosive (ME) processes where the autoregressive (AR) root is greater than unity but its deviation from unity decreases as the sample size increases. Such triangular array data processes have been shown to capture the ME behaviour in many economic and financial time series. The work of Phillips and Magdalinos (2007a, hereafter PM) has stimulated many subsequent studies including Phillips and Magdalinos (2007b), Magdalinos and Phillips (2009), Phillips et al. (2010), Phillips et al. (2011), Magdalinos (2012), Phillips et al. (2014, 2015a,b), and Arvanitis and Magdalinos (2018), among others.

Research on explosive processes can be traced back to White (1958) and Anderson (1959). For a simple Gaussian AR(1) process $y_t = \rho y_{t-1} + u_t$ $(t = 1, 2, \ldots, T)$ with fixed $\rho > 1$, $y_0 = 0$, and independent and identically distributed (i.i.d.) Gaussian errors $\{u_t\}$, White (1958) shows that $\rho^T(\hat{\rho} - \rho)/(\rho^2 - 1)$ converges to a standard Cauchy distribution, where $\hat{\rho}$ is the ordinary least-squares (OLS) estimator of ρ. Anderson (1959) points out that the normality of the error process is necessary for this result. This poses a challenge in the application of explosive processes, as we have to use different reference distributions for different distributions of the errors, which are often not known.

Phillips and Magdalinos (2007a) show that for an ME process wherein ρ is parametrized as $\rho := \rho_T = 1 + c/k_T$ for some $c > 0$ and $k_T = o(T) \to \infty$ as $T \to \infty$, the limiting behaviour of the OLS estimator of ρ is invariant to the distribution of the errors. More specifically, it is shown that the coefficient-based statistic $k_T \rho_T^T(\hat{\rho} - \rho_T)/(2c)$ converges weakly to the standard Cauchy distribution, even if the errors are not Gaussian. Inference can then be made without accounting for the exact distribution of the errors in large samples. More importantly, compared with the original explosive processes of White (1958) and Anderson (1959), ME processes are better able to capture the empirical regularities found in many economic and financial data, such as the Dow Jones Industrial Average.

In this paper, we generalize PM (2007a) to allow for an intercept in the AR(1) process and develop an asymptotically valid test for moderate explosiveness. The ME process under consideration, i.e., $y_t = \mu_T + \rho_T y_{t-1} + u_t$, has two components: the stochastic ME component and the deterministic drift trend component, both of which can render the process explosive. Generally, the deterministic trend component dominates the stochastic trend component, but when the drift μ_T decreases to zero at a certain rate with the sample size, e.g., $\mu_T \sqrt{k_T} \to 0$, the stochastic trend will become stronger in relation to the drift component. Regardless of which component dominates, this paper shows that under the null of moderate explosiveness, the asymptotic distributions of the OLS t statistic are the same, even though the asymptotic distributions of the underlying OLS estimator of ρ_T are different. In particular, in the presence of i.i.d. errors, the OLS t statistic is asymptotically standard normal regardless of whether the drift is large or small, or simply equal to zero. This invariance property extends the existing literature on how the drift specification affects the least squares limit theory of an explosive AR(1) model (see, e.g., Wang and Yu, 2015, Fei, 2018, and Liu and Peng, 2018), and releases us from having to choose a reference distribution in practice. Compared with the nonstandard test of Wang and Yu (2015), who also accommodate a drift but assume a fixed ρ greater than 1, our asymptotic normal test is much easier to use, as critical values are readily available.

Our invariance result is in sharp contrast with the unit-root case and the conventional local-to-unity case where the nonstandard limiting distribution of the t-type statistic is a functional of Brownian motions. The nonstandard distribution is different depending on whether an intercept is included in the regression or not. When an intercept is included, a demeaning effect will emerge and will be retained in the limiting distribution. For more detailed discussion, see Dickey and Fuller (1979, 1981) and MacKinnon (1996) in the unit root setting, Phillips (1987) and Phillips and Perron (1988) in the local-to-unity setting, and Phillips et al. (2014) and Phillips et al. (2015a,b) in the periodically collapsing explosive bubble setting.

Another contribution of this paper is that we extend our basic results to allow for weakly dependent errors. The limiting distribution of the OLS estimator of ρ_T is still normal or mixed normal, but it now depends on the long-run variance (LRV) of the error process. We employ the simple average of the first few periodograms to estimate the LRV and construct the heteroskedasticity and autocorrelation robust (HAR) standard error of the OLS estimator of ρ_T. Under the conventional asymptotics where the number of periodograms grows but at a slower rate than the sample size, the LRV estimator is consistent. In this case, the t statistic is still asymptotically standard normal, and our invariance result continues to hold.

Given that the normal approximation is not very accurate when the error process has high autocorrelation, we develop the t approximation theory using the fixed-smoothing asymptotics. Under this type of asymptotics where the number of periodograms used in the LRV estimation is held fixed, we show that the t statistic based on the HAR standard error follows Student's t distribution in large samples. This result holds regardless of whether a drift term is present or not. The asymptotic t test achieves double robustness: it is asymptotically valid no matter whether the errors are autocorrelated or not, and whether the drift is large, small, or is simply not present.

Monte Carlo (MC) simulations show that the asymptotic normal test under i.i.d. errors and the asymptotic t test under weakly dependent errors have accurate size and satisfactory power in finite samples. When it is not clear whether the errors are i.i.d., we recommend using the HAR t test with a data-driven smoothing parameter.

To identify the degree of the moderate explosiveness of a time series in practice, we propose a two-step empirical testing strategy that involves pretesting.[1] The pretesting aims at detecting whether the series is an explosive process. This is necessary, as the asymptotic t test is based on the primitive condition $\rho_T > 1$, and the asymptotic t theory does not hold when $\rho_T = 1$. For this reason, we have to exclude the nonexplosive root at first. After finding evidence on explosive behaviour, we proceed to employ our asymptotic t test to obtain a confidence interval for the explosiveness. The confidence interval consists of all permissible null values of ρ_T that are not rejected by our t test. Categorizing the seemingly severe or slight explosiveness according to ρ_T will be helpful in bubble identification, classification, and provision of warning. We apply our empirical testing strategy to ten major stock indexes in various countries/districts of the world in a period before the 2008 financial crisis. Interestingly, we find that most indexes are only mildly explosive, or not explosive at all. The pre-2008-financial-crisis bout of irrational rise did not seem so serious as previously thought. This is consistent with Greenspan (2008)'s perception that the financial bubble was not so large.

The rest of the paper is organized as follows. Section 2 establishes the limit theory for ME processes with a sample-size dependent drift. The drift is allowed to be large or small, or simply equal to zero. This section also compares our limit theory with the limit theory developed by Wang and Yu (2015) for severely explosive processes. Section 3 extends the results in Section 2 by allowing weakly dependent errors. Section 4 contains simulation evidence. Section 5 provides the empirical testing strategy and documents the empirical application. The last section concludes. Appendix A presents some technical lemmas that are used in the proofs of the key results, and Appendix B comprises the proofs of the key results. Proofs of the technical lemmas and some additional simulation results are relegated to the online supplement.

2. ASYMPTOTIC NORMAL TEST UNDER I.I.D. ERRORS

2.1. Preliminaries

Following PM (2007a), we consider an ME series $\{\xi_t\}$:

$$\begin{aligned} \xi_t &= \rho \xi_{t-1} + u_t, \\ \rho &= \rho_T = 1 + \frac{c}{k_T}, \ c > 0, \end{aligned} \tag{2.1}$$

for $t = 1, 2, \ldots, T$, where $\{u_t\}$ is a sequence of i.i.d. innovations with $Eu_t = 0$ and $Eu_t^2 = \sigma^2 < \infty$, and k_T increases with T but at a slower rate, i.e., $k_T \to \infty$ but

[1] There will be some size distortion from pretesting. In principle, a Bonferroni correction can be used to alleviate the problem.

$k_T/T \to 0$ as $T \to \infty$. Under the rate condition on k_T, we can show that, for any $a > 0$, ρ_T^{aT} grows at an exponential rate in T/k_T, which is faster than any polynomial rate in T/k_T; see Lemma A.1 in Appendix A.

We further assume that the initial value of the ME process, ξ_0, satisfies $\xi_0 = o_p(\sqrt{k_T})$ and that ξ_0 is independent of $\{u_t, t = 1, \ldots, T\}$. The triangular parametrization of ρ_T and the assumption on ξ_0 ensure that an invariance principle can be established for the ME process. If ρ is a fixed value greater than 1, the effects of a nonzero initial value would not disappear, even asymptotically. In this case, as shown in Anderson (1959), an invariance principle is not applicable.

Define

$$X_T := k_T^{-1/2} \sum_{t=1}^{T} \rho_T^{-(T-t)-1} u_t \text{ and } Y_T := k_T^{-1/2} \sum_{j=1}^{T} \rho_T^{-j} u_j. \tag{2.2}$$

Let X and Y be independent random variables, each distributed as $N(0, \sigma_c^2)$, where $\sigma_c^2 := \sigma^2/(2c)$. PM (2007a) show that

$$(X_T, Y_T)' \Rightarrow (X, Y)'. \tag{2.3}$$

The symbol "$\Rightarrow$" signifies the weak convergence. Moreover, they show that

$$\left(k_T \rho_T^T\right)^{-2} \sum_{t=1}^{T} \xi_{t-1}^2 = \frac{1}{2c} Y_T^2 + o_p(1), \tag{2.4}$$

$$\left(k_T \rho_T^T\right)^{-1} \sum_{t=1}^{T} \xi_{t-1} u_t = X_T Y_T + o_p(1), \tag{2.5}$$

and

$$k_T \rho_T^T (\hat{\rho}_{T,\xi} - \rho_T) \Rightarrow 2cX/Y, \tag{2.6}$$

where $\hat{\rho}_{T,\xi}$ is the OLS estimator of ρ_T and X/Y follows the standard Cauchy distribution. See PM (2007a, page 122) for more details.

Let

$$\hat{\sigma}_{\rho,\xi}^2 = s_{T,\xi}^2 \left(\sum_{t=1}^{n} \xi_{t-1}^2 \right)^{-1} \text{ and } s_{T,\xi}^2 = \frac{1}{T-1} \sum_{t=1}^{T} \left(\xi_t - \hat{\rho}_{T,\xi} \xi_{t-1}\right)^2.$$

Taking $\hat{\sigma}_{\rho,\xi}$ as an estimator of the standard error of $(\hat{\rho}_{T,\xi} - \rho_T)$, we construct the OLS t statistic as follows

$$t_{\mathrm{PM}} := \frac{\hat{\rho}_{T,\xi} - \rho_T}{\hat{\sigma}_{\rho,\xi}}. \tag{2.7}$$

Using (2.3)–(2.6), we can show that

$$t_{\mathrm{PM}} \Rightarrow \frac{2cX/Y}{\sigma/(Y/\sqrt{2c})} = \frac{X}{\sigma_c} \overset{d}{=} N(0, 1).$$

The symbol "$\overset{d}{=}$" signifies the equivalence in distribution.

2.2. Model and Test for ME Processes with Drift

We consider an ME process with drift (MED) defined by

$$y_t = \mu_T + \rho y_{t-1} + u_t,$$
$$\rho = \rho_T = 1 + \frac{c}{k_T}, \ c > 0. \tag{2.8}$$

We maintain the following assumption.

ASSUMPTION 2.1. *(a) $u_t \sim i.i.d.(0, \sigma^2)$; (b) $k_T \to \infty$ and $k_T/T \to 0$ as $T \to \infty$; (c) $\mu_T\sqrt{k_T} \to \nu \in [0, \infty]$ as $T \to \infty$; (d) y_0 is independent of $\{u_t, t = 1, \ldots, T\}$ and $y_0 = o_p(\sqrt{k_T})$.*

As discussed earlier, by imposing an upper bound on the rate of divergence of k_T, i.e., $k_T = o(T)$, we assume that the AR root deviates more from the unity than the usual local-to-unity specification under which $\rho_T = 1 + c/T$. The relatively larger deviation leads to explosive behaviour. On the other hand, the deviation is decaying in T so that the process is only mildly explosive.

The drift in our model can be both large or small. When $\mu_T\sqrt{k_T} \to \infty$, we say that the drift is large. When μ_T is a fixed constant, then $\mu_T\sqrt{k_T} \to \infty$, and we have a large drift. On the other hand, when $\mu_T\sqrt{k_T} \to \nu \in [0, \infty)$, we say that the drift is small. In this case, μ_T approaches zero at a certain rate with the sample size. Note that ν can be arbitrarily close to zero or just equal to zero. So our model allows for a small drift or no drift at all. In practice, we do not know the size of the true drift. To avoid model misspecification, it is advisable to include a drift in our model specification.

Expanding (2.8), we obtain

$$y_t = \rho_T^t y_0 + \sum_{j=1}^{t} \rho_T^{t-j} u_j + \mu_T \left(\rho_T^t - 1\right) / \left(\rho_T - 1\right)$$
$$= \xi_t + \mu_T \left(\rho_T^t - 1\right) k_T/c, \tag{2.9}$$

where

$$\xi_t = \rho_T^t \xi_0 + \sum_{j=1}^{t} \rho_T^{t-j} u_j \text{ for } \xi_0 = y_0.$$

$\{\xi_t\}$ satisfies model (2.1) and is an ME process without drift. So, the stochastic approximations in (2.4) and (2.5) in Section 2.1 hold. When $\mu_T \neq 0$, the process $\{y_t\}$ has two components: the stochastic ME component ξ_t and the deterministic nonlinear trend component $\mu_T(\rho_T^t - 1)k_T/c$, both of which can render the process explosive.

Based on (2.9), we obtain Theorem 2.1 which characterizes the limits of the main sample statistics of interest.

THEOREM 2.1. *Let Assumption 2.1 hold with $\nu \in (0, \infty]$. Define $1/\infty = 0$. Then the following convergence results hold jointly:*

(a) $\left(\mu_T^2 k_T^3 \rho_T^{2T}\right)^{-1} \sum_{t=1}^{T} y_{t-1}^2 \Rightarrow \frac{1}{2c}\left(\frac{Y}{\nu} + \frac{1}{c}\right)^2$;

(b) $\left(\mu_T k_T^2 \rho_T^T\right)^{-1} \sum_{t=1}^{T} y_{t-1} \Rightarrow \frac{1}{c}\left(\frac{Y}{\nu}+\frac{1}{c}\right)$;

(c) $(\mu_T k_T^{3/2} \rho_T^T)^{-1} \sum_{t=1}^{T} y_{t-1} u_t \Rightarrow X\left(\frac{Y}{\nu}+\frac{1}{c}\right)$.

When Assumption 2.1(c) holds with $\nu = \infty$, the convergence rates of the sample statistics in Theorem 2.1 are all higher than those obtained for the ME processes without drift. The faster rates of convergence when the drift satisfies $\mu_T\sqrt{k_T} \to \infty$ are due to the accumulation of the drift term. In large samples, $\{y_t\}$ behaves like a deterministic trending process with consequential effects on the asymptotic behaviour of the sample statistics. This also explains why μ_T appears as a normalization factor in Theorem 2.1.

We proceed to investigate the asymptotic distribution of the OLS estimator $\hat{\rho}_T$ of ρ_T. Define

$$Z_T := T^{-1/2} \sum_{t=1}^{T} u_t. \tag{2.10}$$

Using the Lindeberg-Feller central limit theorem, we can show that Z_T converges in distribution to Z, where $Z \sim N(0, \sigma^2)$. Moreover, the convergence holds jointly with the convergence in (2.3) with Z independent of (X, Y); see Wang and Yu (2015) or the proof of Theorem 3.3 in Appendix B.

To characterize the rate of convergence of the OLS estimator $\hat{\rho}_T$ of ρ_T, we let

$$\boldsymbol{D}_T = \begin{pmatrix} T^{1/2} & 0 \\ 0 & \mu_T k_T^{3/2} \rho_T^T \end{pmatrix}.$$

Then, for $\boldsymbol{x}_t = (\ 1, \quad y_{t-1}\ )'$, we have

$$\begin{aligned} \boldsymbol{D}_T^{-1}\left(\sum_{t=1}^{T} \boldsymbol{x}_t \boldsymbol{x}_t'\right)\boldsymbol{D}_T^{-1} &= \begin{pmatrix} \frac{1}{T}\sum_{t=1}^{T} 1 & \frac{1}{\mu_T\sqrt{T}k_T^{3/2}\rho_T^T}\sum_{t=1}^{T} y_{t-1} \\ \frac{1}{\mu_T\sqrt{T}k_T^{3/2}\rho_T^T}\sum_{t=1}^{T} y_{t-1} & \frac{1}{\mu_T^2 k_T^3 \rho_T^{2T}}\sum_{t=1}^{T} y_{t-1}^2 \end{pmatrix} \\ &= \begin{pmatrix} 1 & O_p\left(\sqrt{k_T/T}\right) \\ O_p\left(\sqrt{k_T/T}\right) & \frac{1}{\mu_T^2 k_T^3 \rho_T^{2T}}\sum_{t=1}^{T} y_{t-1}^2 \end{pmatrix} \Rightarrow \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{2c}\left(\frac{Y}{\nu}+\frac{1}{c}\right)^2 \end{pmatrix}, \end{aligned}$$

using Theorems 2.1(a) and (b). In addition, using Theorem 2.1(c), we have

$$\boldsymbol{D}_T^{-1}\sum_{t=1}^{T} \boldsymbol{x}_t u_t \Rightarrow \left(\ Z, \quad X\left(Y/\nu + 1/c\right)\ \right)'.$$

It then follows that

$$\begin{aligned} \mu_T k_T^{3/2} \rho_T^T \left(\hat{\rho}_T - \rho_T\right) &= e_2' \left[\boldsymbol{D}_T^{-1}\left(\sum_{t=1}^{T} \boldsymbol{x}_t \boldsymbol{x}_t'\right)\boldsymbol{D}_T^{-1}\right]^{-1}\left[\boldsymbol{D}_T^{-1}\sum_{t=1}^{T} \boldsymbol{x}_t u_t\right] \\ &\Rightarrow \frac{X\left(Y/\nu+1/c\right)}{\left(Y/\nu+1/c\right)^2/\left(2c\right)} = \frac{2cX}{Y/\nu+1/c}, \end{aligned} \tag{2.11}$$

where $e_2 = (\ 0, \quad 1\ )'$.

When $\nu = \infty$, we have $\mu_T k_T^{3/2} \rho_T^T (\hat{\rho}_T - \rho_T) \Rightarrow 2c^2 X$ and so $\hat{\rho}_T$ is asymptotically normal. The rate of convergence of $\hat{\rho}_T$ to ρ_T (i.e., $\mu_T k_T^{3/2} \rho_T^T$) is faster than the rate of $k_T \rho_T^T$ in PM (2007a). This result is consistent with Theorem 2.7(b) in Fei (2018) and Theorem 1(P7) in Liu and Peng (2018), both of which allow for a nonzero constant intercept in the MED model such that $\mu_T \sqrt{k_T} \to \infty$. On the other hand, when $\nu \in (0, \infty)$, the limit distribution is mixed normal. Like Theorem 4.3(b) in PM (2007a) and Theorem 1(P4) in Liu and Peng (2018), it is a ratio of two independent normal random variables, but it is not the Cauchy distribution. Depending on the value of ν, we obtain an asymptotically normal or mixed-normal distribution.

We now construct the t statistic as follows:

$$t_{\mathrm{MED}} := \frac{\hat{\rho}_T - \rho_T}{\hat{\sigma}_\rho},$$

where

$$\hat{\sigma}_\rho^2 = s_T^2 e_2' \left(\sum_{t=1}^{T} \boldsymbol{x}_t \boldsymbol{x}_t' \right)^{-1} e_2 \text{ and } s_T^2 = \frac{1}{T-2} \sum_{t=1}^{T} \left(y_t - \hat{\mu}_T - \hat{\rho}_T y_{t-1} \right)^2 .$$

Then we have

$$t_{\mathrm{MED}} \Rightarrow \frac{2cX}{Y/\nu + 1/c} \left(\frac{1}{\sigma} \frac{Y/\nu + 1/c}{\sqrt{2c}} \right) = \frac{X}{\sigma_c} \overset{d}{=} N(0,1) .$$

The limiting distribution of t_{MED} is the standard Gaussian distribution rather than some nonstandard distribution that involves functionals of Brownian motions. The main reason is that, after being normalized by the scaling matrix $\boldsymbol{D}_T$, the off-diagonal elements of $\boldsymbol{D}_T^{-1} (\sum_{t=1}^T \boldsymbol{x}_t \boldsymbol{x}_t') \boldsymbol{D}_T^{-1}$ vanish as $T \to \infty$. A key assumption behind this result is that $k_T = o(T)$. In contrast, these elements converge weakly to a nonzero constant or random variate in the conventional unit-root or local-to-unity framework.

The ME process can be regarded as an approximation to the unit root process from the explosive side. When k_T is of the same order as T, our parametrization resembles a near unit-root parametrization but on the explosive side. Note that when $k_T = T$, we have $\lim_{T\to\infty} \rho_T^T = \lim_{T\to\infty} (1 + c/T)^T = e^c$. So when k_T is of the same order as T and μ_T is a constant, the orders of $\sum_{t=1}^T y_{t-1}^2$, $\sum_{t=1}^T y_{t-1}$, and $\sum_{t=1}^T y_{t-1} u_t$ become close to T^3, T^2, and $T^{3/2}$, respectively. These convergence rates match those in the local-to-unity case.

To investigate the asymptotic properties of the t test when $\nu = 0$, we establish the theorem below, which is a modified version of Theorem 2.1. Given that the proof is essentially the same as that for Theorem 2.1 with only minor modifications, we omit it here.

THEOREM 2.2. *Let Assumption 2.1 hold with $\nu = 0$. Then the following convergence results hold jointly:*

(*a*) $\left(k_T^2 \rho_T^{2T} \right)^{-1} \sum_{t=1}^{T} y_{t-1}^2 \Rightarrow \frac{1}{2c} Y^2$;

(*b*) $\left(k_T^{3/2} \rho_T^T \right)^{-1} \sum_{t=1}^{T} y_{t-1} \Rightarrow \frac{1}{c} Y$;

(c) $\left(k_T \rho_T^T\right)^{-1} \sum_{t=1}^{T} y_{t-1} u_t \Rightarrow XY$.

The limiting behaviours of the sample statistics are the same as the case with no drift. Combining Theorem 2.2 with the argument for the asymptotic normal result for the case $\nu > 0$, we obtain

$$k_T \rho_T^T \left(\hat{\rho}_T - \rho_T\right) \Rightarrow 2cX/Y, \tag{2.12}$$

and

$$t_{\text{MED}} \Rightarrow X/\sigma_c \overset{d}{=} N\left(0,1\right).$$

We formalize our asymptotic standard normal limit theory in the theorem below.

THEOREM 2.3. *Let Assumption 2.1 hold. Then* $t_{MED} \Rightarrow N(0,1)$ *as* $T \to \infty$.

Regardless of the size of the drift, the t statistic is asymptotically standard normal. This is a very encouraging and convenient result. We obtain the same limiting distribution even though the asymptotic distribution of the coefficient estimator is different for different drift sizes. Note that the t test based on the PM regression (with no intercept included) is asymptotically normal only in the absence of a drift term. The asymptotic normal t-test based on the PM regression can have large size distortion if a drift is actually present. When the nature of the drift is not known, we recommend employing the t_{MED} test, which is asymptotically valid no matter whether the drift is large or small.

Wang and Yu (2015, hereafter WY) develop the limit theory for the model

$$y_t = \mu + \rho y_{t-1} + u_t,\ u_t \sim i.i.d.(0, \sigma^2),$$

where both μ and ρ are fixed and $\rho > 1$. Compared with a moderate deviation from unity, a fixed ρ value that is strictly greater than 1 can be viewed as a severely explosive (SE) parametrization.

The t statistic t_{WY} in WY (2015) is identical to t_{MED}. Let

$$\ddot{X}_T := \sum_{t=1}^{T} \rho^{-(T-t)} u_t \text{ and } \ddot{Y}_T := \rho \sum_{j=1}^{T-1} \rho^{-j} u_j + \rho y_0.$$

WY (2015) show that $(\ddot{X}_T, \ddot{Y}_T) \Rightarrow (\ddot{X}, \ddot{Y})$ and that

$$t_{\text{WY}} \Rightarrow t_{\text{WY},\infty}\left(y_0, \rho, \sigma^2, \mu\right) := \frac{\ddot{X}}{\ddot{Y} + \rho\mu/\left(\rho - 1\right)} \cdot \left|\ddot{Y} + \frac{\rho\mu}{\rho - 1}\right| \cdot \left(\frac{\rho^2 - 1}{\rho^2\sigma^2}\right)^{1/2}. \tag{2.13}$$

The limiting distribution is nonstandard. It is also not pivotal, as it depends on the unknown parameters ρ, μ, and σ, and the initial value y_0. This feature makes the limiting distribution less convenient to use in empirical applications.

If we replace ρ with $\rho_T = 1 + c/k_T$ and maintain the same initial condition that $y_0 = o_p(\sqrt{k_T})$, we will have $k_T^{-1/2}\ddot{X}_T = X_T(1 + o_p(1))$ and $k_T^{-1/2}\ddot{Y}_T = Y_T(1 + o_p(1))$.[2] Then the distribution of the random variable $t_{\mathrm{WY},\infty}(y_0, \rho, \sigma^2, \mu)$ will become asymptotically standard normal as $T \to \infty$, viz.

$$
\begin{aligned}
& t_{\mathrm{WY},\infty}\left(y_0, \rho, \sigma^2, \mu\right) \\
&= \frac{\sqrt{k_T}X}{\sqrt{k_T}Y + \mu k_T/c + \mu} \cdot \left|\sqrt{k_T}Y + \mu k_T/c + \mu\right| \cdot \left(\frac{2c/k_T}{\sigma^2}\right)^{1/2} \left(1 + o_p\left(1\right)\right) \\
&= \frac{X}{\sigma_c} \cdot \frac{\left|\sqrt{k_T}Y + \mu k_T/c + \mu\right|}{\sqrt{k_T}Y + \mu k_T/c + \mu} \left(1 + o_p\left(1\right)\right) = \pm\frac{X}{\sigma_c}\left(1 + o_p\left(1\right)\right) \Rightarrow N\left(0, 1\right),
\end{aligned}
$$

no matter whether $(\sqrt{k_T}Y)/(\mu k_T/c+\mu) = O_p(1)$ or $(\mu k_T/c+\mu)/(\sqrt{k_T}Y) = O_p(1)$. This is a type of informal sequential asymptotics. We first establish the limiting distribution of the t statistic for a fixed $\rho > 1$ and a given initial value y_0. We then investigate the behaviour of the limiting distribution when ρ approaches 1 from the right-hand side (i.e., $\rho = 1 + c/k_T$) and when the initial value becomes stochastically manageable (i.e., $y_0 = o_p(\sqrt{k_T})$). There is a smooth transition from the limiting distribution in the severely explosive case (i.e., ρ is fixed and greater than 1) to that in the moderately explosive case (i.e., $\rho = 1 + c/k_T$ for $c > 0$ and $k_T = o(T)$).

3. ASYMPTOTIC t TEST UNDER WEAKLY DEPENDENT ERRORS

The previous section has been confined to the case wherein the sequence of errors driving the model is independently and identically distributed. A natural extension is to develop a test for MED that does not rely on this strong assumption. Assumption 3.1 below allows the error process to have a general dependence structure.

Assumption 3.1. *(a)* $u_t = C(L)\varepsilon_t$ *with* $\varepsilon_t \sim i.i.d.(0, \sigma^2)$, $C(L) = \sum_{j=0}^{\infty} c_j L^j$, $c_0 = 1$, *and* L *is the lag operator; (b)* $C(1) \in (0, \infty)$ *and* $\sum_{j=0}^{\infty} j \cdot |c_j| < \infty$*; (c)* $E|\varepsilon_t|^l < \infty$ *for some* $l \geq 4$*; (d)* $k_T \to \infty$ *and* $k_T/T \to 0$ *as* $T \to \infty$*; (e)* $\mu_T\sqrt{k_T} \to \nu \in [0, \infty]$ *as* $T \to \infty$*; (f)* y_0 *is independent of* $\{u_t, t = 1, \ldots, T\}$ *and* $y_0 = o_p(\sqrt{k_T})$.

[2] The two results $k_T^{-1/2}\ddot{X}_T = X_T(1 + o_p(1))$ and $k_T^{-1/2}\ddot{Y}_T = Y_T(1 + o_p(1))$ hold because

$$
\begin{aligned}
k_T^{-1/2}\ddot{X}_T &= k_T^{-1/2}\sum_{t=1}^{T}\rho^{-(T-t)}u_t = k_T^{-1/2}\rho_T\sum_{t=1}^{T}\rho_T^{-(T-t)-1}u_t \\
&= \left[k_T^{-1/2}\left(1 + ck_T^{-1}\right)\right]\sum_{t=1}^{T}\rho_T^{-(T-t)-1}u_t = X_T\left(1 + o_p\left(1\right)\right),
\end{aligned}
$$

and

$$
\begin{aligned}
k_T^{-1/2}\ddot{Y}_T &= k_T^{-1/2}\left(\rho\sum_{j=1}^{T-1}\rho^{-j}u_j + \rho y_0\right) = k_T^{-1/2}\rho_T\left(\sum_{j=1}^{T}\rho_T^{-j}u_j - \rho_T^{-T}u_T\right) + k_T^{-1/2}\rho_T y_0 \\
&= \left[k_T^{-1/2}\left(1 + ck_T^{-1}\right)\right]\sum_{j=1}^{T}\rho_T^{-j}u_j + O_p\left(k_T^{-1/2}\rho_T^{-(T-1)}\right) + o_p\left(1 + ck_T^{-1}\right) \\
&= Y_T\left(1 + o_p\left(1\right)\right).
\end{aligned}
$$

Assumptions 3.1(a)–(c) are the same as those maintained in Phillips and Solo (1992) and Phillips and Magdalinos (2007b). Under these assumptions, $\{u_t\}$ is weakly stationary.[3] Assumption 3.1(b) ensures that $\{u_t\}$ has a martingale decomposition:

$$u_t = C(1)\varepsilon_t + \tilde{\varepsilon}_{t-1} - \tilde{\varepsilon}_t, \tag{3.1}$$

where $\tilde{\varepsilon}_t = \sum_{j=0}^{\infty} \tilde{c}_j \varepsilon_{t-j}$ and $\tilde{c}_j = \sum_{k=j+1}^{\infty} c_k$. In addition, $\sum_{j=0}^{\infty} \tilde{c}_j^2 < \infty$ and so $var(\tilde{\varepsilon}_t) < \infty$. For more details, see Phillips and Solo (1992, Theorem 2.5). Using the martingale decomposition, we have

$$T^{-1/2} \sum_{t=1}^{T} u_t \Rightarrow N\left(0, \lambda^2\right),$$

where λ^2 is the LRV of $\{u_t\}$ defined by

$$\lambda^2 := \lim_{T\to\infty} T^{-1} E\left(\sum_{t=1}^{T} u_t\right)^2 = \sigma^2 C(1)^2.$$

Define $\lambda_c^2 := \lambda^2/(2c)$. The above martingale decomposition also facilitates the proof of Lemma 3.1 below.

Lemma 3.1. *Let Assumption 3.1 hold. Then*

(*a*)

$$\tilde{X}_T := k_T^{-1/2} \sum_{t=1}^{T} \rho_T^{-(T-t)-1} u_t = C(1)\, k_T^{-1/2} \sum_{t=1}^{T} \rho_T^{-(T-t)-1} \varepsilon_t + o_p(1);$$

(*b*)

$$\tilde{Y}_T := k_T^{-1/2} \sum_{t=1}^{T} \rho_T^{-t} u_t = C(1)\, k_T^{-1/2} \sum_{t=1}^{T} \rho_T^{-t} \varepsilon_t + o_p(1);$$

(*c*) $(\tilde{X}_T, \tilde{Y}_T) \Rightarrow (\tilde{X}, \tilde{Y})$ *where* $\tilde{X}$ *and* $\tilde{Y}$ *are independent* $N(0, \lambda_c^2)$ *random variables.*

Lemma 3.1 shows that the effect of temporal dependence on the distribution of $(\tilde{X}_T, \tilde{Y}_T)$ is to re-scale the distribution under i.i.d. errors by a constant $C(1)$. As a result, the asymptotic distributions of the main sample statistics under $\nu \in (0, \infty]$ and under $\nu = 0$ follow in a direct way from the approach that we pursue in Section 2. The proof of Theorem 3.1 is given in Appendix B while the proof of Theorem 3.2 is similar and is therefore omitted.

Theorem 3.1. *Let Assumption 3.1 hold with* $\nu \in (0, \infty]$*. Define* $1/\infty = 0$*. Then the following convergence results hold jointly:*

(*a*) $\left(\mu_T^2 k_T^3 \rho_T^{2T}\right)^{-1} \sum_{t=1}^{T} y_{t-1}^2 \Rightarrow \frac{1}{2c}\left(\frac{\tilde{Y}}{\nu} + \frac{1}{c}\right)^2;$

(*b*) $\left(\mu_T k_T^2 \rho_T^{T}\right)^{-1} \sum_{t=1}^{T} y_{t-1} \Rightarrow \frac{1}{c}\left(\frac{\tilde{Y}}{\nu} + \frac{1}{c}\right);$

[3] We could also assume that $\{\varepsilon_t\}$ is a martingale difference sequence satisfying a L_1-mixingale condition so that only second-moment conditions of $\{u_t\}$ are required; see, e.g., Magdalinos (2012) and Arvanitis and Magdalinos (2018).

(c) $\left(\mu_T k_T^{3/2} \rho_T^T\right)^{-1} \sum_{t=1}^{T} y_{t-1} u_t \Rightarrow \tilde{X} \left(\frac{\tilde{Y}}{\nu} + \frac{1}{c}\right).$

THEOREM 3.2. *Let Assumption 3.1 hold with $\nu = 0$. Then the following convergence results hold jointly:*

(a) $\left(k_T^2 \rho_T^{2T}\right)^{-1} \sum_{t=1}^{T} y_{t-1}^2 \Rightarrow \frac{1}{2c}\tilde{Y}^2;$

(b) $\left(k_T^{3/2} \rho_T^T\right)^{-1} \sum_{t=1}^{T} y_{t-1} \Rightarrow \frac{1}{c}\tilde{Y};$

(c) $\left(k_T \rho_T^T\right)^{-1} \sum_{t=1}^{T} y_{t-1} u_t \Rightarrow \tilde{X}\tilde{Y}.$

Note that $\sum_{t=1}^T y_{t-1}^2$, $\sum_{t=1}^T y_{t-1}$, and $\sum_{t=1}^T y_{t-1}u_t$ have the same convergence rates as in the i.i.d. case. When $\nu \in (0, \infty]$, the OLS estimator $\hat{\rho}_T$ of ρ_T satisfies

$$\mu_T k_T^{3/2} \rho_T^T \left(\hat{\rho}_T - \rho_T\right) \Rightarrow \frac{2c\tilde{X}}{\tilde{Y}/\nu + 1/c}.$$

When $\nu = 0$, the coefficient estimator satisfies

$$k_T \rho_T^T \left(\hat{\rho}_T - \rho_T\right) \Rightarrow 2c\tilde{X}/\tilde{Y}.$$

These two results are analogous to (2.11) and (2.12), respectively.

To make an inference on ρ_T, we need to estimate the LRV λ^2 of $\{u_t\}$. Let

$$\hat{u}_t = y_t - \hat{\mu}_T - \hat{\rho}_T y_{t-1}$$

be the estimated residual. The commonly-used estimator of λ^2 takes the form

$$\hat{\lambda}_K^2 = \frac{1}{T} \sum_{t=1}^{T} \sum_{s=1}^{T} Q_K(t, s) \hat{u}_t \hat{u}_s,$$

where $Q_K(\cdot, \cdot)$ is a weighting function that depends on the smoothing parameter K. This includes the kernel LRV estimator if we let $Q_K(t, s) = \kappa((t - s)/(TK^{-1}))$ for a kernel function $\kappa(\cdot)$. In this paper, we take a simple average of the first few periodograms to construct $\hat{\lambda}_K^2$. More specifically, we let K be even and

$$Q_K(t, s) = \frac{1}{K} \sum_{\ell=1}^{K} \phi_\ell\left(\frac{t}{T}\right) \phi_\ell\left(\frac{s}{T}\right),$$

where $\phi_{2\ell}(x) = \sqrt{2}\sin(2\pi\ell x)$ and $\phi_{2\ell-1}(x) = \sqrt{2}\cos(2\pi\ell x)$ are the Fourier basis functions. With the above weighting function, $\hat{\lambda}_K^2$ takes the average form:

$$\hat{\lambda}_K^2 = \frac{1}{K} \sum_{\ell=1}^{K} \left[\frac{1}{\sqrt{T}} \sum_{t=1}^{T} \phi_\ell\left(\frac{t}{T}\right) \hat{u}_t\right]^2. \tag{3.2}$$

Other basis functions can be used, leading to a new class of orthonormal series LRV estimators. For theoretical developments of this type of LRV estimators and their advantages, see, e.g., Phillips (2005), Müller (2007), and Sun (2011, 2013, 2014). For simplicity, we opt for the Fourier basis functions here.

On the basis of $\hat{\lambda}_K^2$ in (3.2), we construct the t statistic as follows:

$$\tilde{t}_{\mathrm{MED}} := \frac{\hat{\rho}_T - \rho_T}{\tilde{\sigma}_{\rho,K}},$$

where

$$\tilde{\sigma}_{\rho,K}^2 = \hat{\lambda}_K^2 e_2' \left(\sum_{t=1}^{T} \boldsymbol{x}_t \boldsymbol{x}_t' \right)^{-1} e_2.$$

The limiting distribution of the $\tilde{t}_{\mathrm{MED}}$ statistic is given in the theorem below.

THEOREM 3.3. *Let Assumption 3.1 hold. Under the fixed-K asymptotics where $T \to \infty$ for a fixed K, the following convergence results hold jointly:*

(a) $\hat{\lambda}_K^2/\lambda^2 \Rightarrow \chi_K^2/K$ where χ_K^2 is a random variable following the chi-square distribution with K degrees of freedom;

(b) $\tilde{t}_{MED} \Rightarrow t_K$ where t_K is the Student's t distribution with K degrees of freedom.

Theorem 3.3 indicates that if $K \to \infty$, then $\hat{\lambda}_K^2$ will become consistent, and the asymptotic t_K distribution approaches the standard normal distribution. This is a type of sequential asymptotics. More rigorously, under the joint asymptotics under which $K \to \infty$ but $K/T \to 0$ as $T \to \infty$, we can establish that $\hat{\lambda}_K^2$ is consistent for λ^2 and $\tilde{t}_{\mathrm{MED}}$ is asymptotically standard normal. So, under the conventional asymptotics that ensures the consistency of the standard error estimator, the asymptotic normality of the t statistic holds for both i.i.d. errors and weakly dependent errors and for both small and large drifts.

To understand the invariance of the asymptotic distribution of the t statistic, we note that both Theorem 3.1 and Theorem 3.2 imply that $T^{-1}\sum_{t=1}^{T} y_{t-1}^2$ stochastically dominates $(T^{-1}\sum_{t=1}^{T} y_{t-1})^2$ and that $T^{-1}\sum_{t=1}^{T} y_{t-1}u_t$ stochastically dominates $(T^{-1}\sum_{t=1}^{T} y_{t-1})(T^{-1}\sum_{t=1}^{T} u_t)$. So, if $\hat{\lambda}_K \xrightarrow{p} \lambda$, then the t statistic $\tilde{t}_{\mathrm{MED}}$ satisfies

$$\begin{aligned} \tilde{t}_{\mathrm{MED}} &:= \frac{1}{\hat{\lambda}_K} \frac{\frac{1}{T}\sum_{t=1}^{T} y_{t-1}u_t - \left(\frac{1}{T}\sum_{t=1}^{T} y_{t-1}\right)\left(\frac{1}{T}\sum_{t=1}^{T} u_t\right)}{\left[\frac{1}{T}\sum_{t=1}^{T} y_{t-1}^2 - \left(\frac{1}{T}\sum_{t=1}^{T} y_{t-1}\right)^2\right]^{1/2}} \\ &= \frac{1}{\hat{\lambda}_K} \frac{\frac{1}{T}\sum_{t=1}^{T} y_{t-1}u_t}{\left(\frac{1}{T}\sum_{t=1}^{T} y_{t-1}^2\right)^{1/2}} (1 + o_p(1)) \Rightarrow \frac{\tilde{X}}{\lambda_c} \stackrel{d}{=} N(0,1) \end{aligned}$$

by Theorem 3.1 or 3.2.

Theorem 3.3(b) shows that, under the fixed-K asymptotics, the HAR t statistic is asymptotically t distributed. There is a growing literature showing that the fixed-K asymptotic approximation for the studentized test statistic is more accurate than the corresponding increasing-K asymptotic approximation. The reason is that the former captures the randomness in $\hat{\lambda}_K^2$ while the latter does not. Theorem 3.3 holds for $\nu \in [0, \infty]$. The asymptotic t approximation for the $\tilde{t}_{\mathrm{MED}}$ statistic is valid regardless of whether the drift is present or not. In this sense, the asymptotic t test achieves double robustness: it is asymptotically valid no matter whether the errors are autocorrelated or not, and whether the drift is large or small, or equal to zero.

To establish the asymptotic t theory in Theorem 3.3(b), we have to show that the estimator error in $\hat{\rho}_T$ is asymptotically independent of the LRV estimator $\hat{\lambda}_K^2$. The asymptotic independence is due to the explosive behaviour of the underlying time series. It is similar to the asymptotic independence of (X_T, Y_T) from Z_T, defined in (2.2) and (2.10), respectively. We also have to show that $\{T^{-1/2}\sum_{t=1}^{T}\phi_\ell(\frac{t}{T})\hat{u}_t\}$ forms an i.i.d. sequence in large samples. The key driving forces behind this result are the orthonormality of the basis functions $\{\phi_\ell\}$ on $L_2[0,1]$ and the "zero mean" condition, i.e., $\int_0^1 \phi_\ell(r)dr = 1$.

For the asymptotic t theory to hold, it is necessary to employ the orthonormal series LRV estimator. Using a kernel LRV estimator will not allow us to develop the convenient t approximation. Nevertheless, it will enable us to make asymptotically pivotal inferences — the limiting distribution of the associated t statistic will be a nonstandard mixed-normal distribution that is nuisance parameter free. It is not very convenient to use a nonstandard distribution, as critical values have to be simulated.

4. MONTE CARLO SIMULATION

4.1. Simulation Evidence Under i.i.d. Errors

In this subsection, we conduct MC simulations to evaluate the finite sample performance of our asymptotic normal test, the t_{MED} test, when the errors are independently and identically distributed.

The data generating process (DGP) is given by

$$y_t = \mu_T + \rho y_{t-1} + u_t,\ t = 1, 2, \ldots, T, \tag{4.1}$$

where $\rho = 1 + c/k_T$ with $c = 1$ and $k_T = T^\alpha$ for some $\alpha \in (0, 1)$.[4] The initial value is set to be $y_0 = \mu_T$. The intercept is set to be $\mu_T = 0, T^{-\alpha/2}, T^{-\alpha/4}, 1$. Such a setting is compatible with $y_0 = o_p(\sqrt{k_T})$. We conduct two groups of MC simulations. The first group employs i.i.d. Gaussian errors while the second group employs i.i.d. uniform errors. That is, $u_t \sim i.i.d.N(0,1)$ or $u_t \sim i.i.d.U(-\sqrt{3}, \sqrt{3})$.

We examine the empirical size of the t_{MED} test. For comparison we also examine the empirical size of the t_{PM} and t_{WY} tests. The PM test based on the statistic in (2.7) ignores the intercept, while the WY test assumes that ρ is fixed and strictly larger than 1. The null hypothesis of interest is $H_0 : \rho = 1 + 1/T^\alpha$ for different configurations of α and T, where $1/T^\alpha$ represents the moderate deviation from unity for a sample of size T. To save space, we discuss the case with $\alpha = 0.5$ and $T = 100$ in the main text. This case is representative of other configurations. More detailed simulation results are reported and discussed in the online supplement. For the t_{PM} test and t_{MED} test, we use critical values from the standard normal distribution. The t_{WY} test is similar to the t_{MED} test but uses critical values from the asymptotic distribution shown in (2.13), which is simulated using true parameter values. To a great extent, we give the t_{WY} test some edge, as some of the true parameter values are not known under the null. The nominal level is 5%, and the number of simulation replications is 5,000.

We also examine the empirical power of the three competing tests. The parameter configuration is the same as those for size calculations except the DGP is generated under the local-to-unity alternative $H_A : \rho = 1 + 1/T$. To avoid the size difference in the

[4]Other parameter configurations of c and k_T, e.g., $c = 0.1$ and $k_T = \log T$, have also been examined. The results are similar and are available upon request.

Table 1. Size and power under i.i.d. errors: $\alpha = 0.5$ and $T = 100$.

	Size ($\rho = 1 + 1/T^{\alpha}$)			Power ($\rho = 1 + 1/T$)	
	t_{PM}	t_{WY}	t_{MED}	t_{PM}	t_{MED}
			(a) i.i.d. Gaussian errors		
$\mu_T = 0$	0.050	0.053	0.056	1.000	1.000
$\mu_T = T^{-\alpha/2}$	0.277	0.058	0.055	1.000	1.000
$\mu_T = T^{-\alpha/4}$	0.641	0.056	0.055	1.000	1.000
$\mu_T = 1$	0.969	0.058	0.055	1.000	1.000
			(b) i.i.d. uniform errors		
$\mu_T = 0$	0.049	0.047	0.049	1.000	1.000
$\mu_T = T^{-\alpha/2}$	0.276	0.049	0.049	1.000	1.000
$\mu_T = T^{-\alpha/4}$	0.637	0.047	0.048	1.000	1.000
$\mu_T = 1$	0.973	0.047	0.048	1.000	1.000

Note: This table reports the empirical size and size-adjusted power of 5% tests with 5,000 Monte Carlo replications. The model used for the experiment is (4.1) with $u_t \sim i.i.d.N(0,1)$ in the i.i.d. Gaussian group and with $u_t \sim i.i.d.U(-\sqrt{3}, \sqrt{3})$ in the i.i.d. uniform group. Different parameter combinations are configured to conduct simulations for the null of moderate explosiveness $\rho = 1 + 1/T^{\alpha}$ against the alternative of local-to-unity $\rho = 1 + 1/T$.

power comparison, we simulate and compare the size-adjusted power using the empirical finite sample critical values obtained from the null distribution. Since the t_{WY} and t_{MED} tests are based on the same test statistic, the size-adjusted power of these two tests is identical. We report the power for the t_{MED} test only.

Table 1 reports the size and power results of the t_{PM}, t_{WY}, and t_{MED} tests under Gaussian errors and uniform errors, respectively. The two groups of results are qualitatively similar, providing further evidence that normality of the errors is not necessary for these tests. First, as can be seen from the table, both the t_{WY} and t_{MED} tests have quite accurate size in all drift cases. Note that we employ the true parameter values to simulate the asymptotic distribution of the t_{WY} statistic. In an absolute and overall sense, the standard normal distribution approximates the distribution of the t_{MED} statistic very well. Second, we observe that when $\mu_T = 0$, the size performance of the t_{MED} test is not worse than that of the t_{PM} test, while as μ_T departs farther away from zero, the t_{PM} test suffers from large size distortion but the t_{MED} test still enjoys a good size control. It is encouraging to see that the t_{MED} test dominates the t_{PM} test in terms of the size accuracy. Finally, both the t_{PM} and t_{MED} tests have satisfactory power performance. Simulation results in the online supplement show that the t_{MED} test is generally more powerful than the t_{PM} test. Given the simulation evidence, we can conclude that the t_{MED} test succeeds in controlling size without power loss.

4.2. Simulation Evidence Under Weakly Dependent Errors

Using the same DGP in (4.1), we examine the finite sample performance of the $\tilde{t}_{\text{MED}}$ test under two different experiment designs in this subsection: the AR design and the

Table 2. Size and power in the presence of autocorrelated errors: $\alpha = 0.5$ and $T = 100$.

	Size ($\rho = 1 + 1/T^{\alpha}$)		Power ($\rho = 1 + 1/T$)	
	t_{MED}	$\tilde{t}_{\text{MED}}$	t_{MED}	$\tilde{t}_{\text{MED}}$
		(a) AR design		
$\theta = 0.00$	0.053	0.057	1.000	1.000
$\theta = 0.25$	0.134	0.068	1.000	1.000
$\theta = 0.50$	0.242	0.070	1.000	0.998
$\theta = 0.75$	0.413	0.081	0.993	0.971
		(b) MA design		
$\theta = 0.00$	0.055	0.060	1.000	1.000
$\theta = 0.25$	0.114	0.060	1.000	1.000
$\theta = 0.50$	0.157	0.066	1.000	1.000
$\theta = 0.75$	0.169	0.066	1.000	1.000

Note: This table reports the empirical size and size-adjusted power of 5% tests with 5,000 Monte Carlo replications. The model used for the experiment is (4.1) with $u_t = \theta u_{t-1} + \sqrt{1-\theta^2} e_{1,t}$ under the AR design and with $u_t = \theta e_{2,t-1} + \sqrt{1-\theta^2} e_{2,t}$ under the MA design, where $e_{1,t} \sim i.i.d.N(0,1)$ and $e_{2,t} \sim i.i.d.N(0,1)$. Different parameter combinations are configured to conduct simulations for the null of moderate explosiveness $\rho = 1 + 1/T^{\alpha}$ against the alternative of local-to-unity $\rho = 1 + 1/T$.

moving average (MA) design. To save space, we only consider the case with $\mu_T = T^{-\alpha/4}$. In the AR design, u_t follows an AR(1) process $u_t = \theta u_{t-1} + \sqrt{1-\theta^2} e_{1,t}$, where $e_{1,t} \sim i.i.d.N(0,1)$. In the MA design, $u_t = \theta e_{2,t-1} + \sqrt{1-\theta^2} e_{2,t}$, where $e_{2,t} \sim i.i.d.N(0,1)$. By construction, the error has a unit variance in both designs. We take $\theta = 0.00, 0.25, 0.50,$ and 0.75. The $\tilde{t}_{\text{MED}}$ statistic is based on the LRV estimator in (3.2). Following Phillips (2005), we choose K based on the asymptotic mean squared error (AMSE) criterion implemented using the AR(1) plug-in procedure. We round the data-driven value of K to a closest even number between 4 and T. For both the AR and MA designs, we consider different combinations of α and T; see Table 2 for the case with $\alpha = 0.5$ and $T = 100$ and the online supplement for more detailed simulation evidence. For comparison, we also consider the t_{MED} test, which ignores the autocorrelation in $\{u_t\}$. The initial value is set to be $y_0 = \mu_T$ and the number of simulation replications is 5,000.

Table 2 reports the size and power results of the t_{MED} and $\tilde{t}_{\text{MED}}$ tests. The table shows that compared with the t_{MED} test, the $\tilde{t}_{\text{MED}}$ test achieves a satisfactory size-adjusted power performance with only relatively small size distortion in both the AR and MA designs. This result is consistent with our theoretical analysis. Ignoring the autocorrelation leads to an inaccurate test.

5. EMPIRICAL APPLICATION

5.1. Background and Data: Explosive Ups of the World Stock Indexes

Before the "Great Recession" of 2007–2009, led by the loose monetary policy and irrational real estate boom, the U.S. stock market experienced a spectacular rise (Allen et al., 2009; Taylor, 2009; Allen and Carletti, 2010; Stiglitz, 2010). The most impressive phenomenon is that the Dow Jones Industrial Average (DJI) reached its peak at 14,198.1 points on October 12, 2007, after witnessing continuous gain. Most regard this type of increase as an explosive process and as the first half of a financial bubble episode (Phillips et al., 2015a). Shiller (2008) argued that the irrational prosperity was the root cause of the subprime crisis, which was the crux of the financial crisis. Greenspan (1996) coined the phrase “irrational exuberance” in his remark on December 5, 1996 to describe the herd phenomenon in the stock market.

The global stock markets were also affected by such a rise. Different economies experienced different degrees of the boom during the exuberance period, largely owing to their corresponding global financial participation and dependence on the U.S. economy. China, for instance, held massive foreign exchange reserves, especially the U.S. treasury bonds, in the pre-2008 period (Woo et al., 2013). Along with the American economic prosperity and appreciation of the dollar, a great deal of capital entered into China’s foreign exchange market, stimulating the explosive growth of China’s major stock indexes.

Greenspan (2008) argued that not all of the increasingly growing processes should be characterized by irrational exuberance and that the bubble was not so large. We are sympathetic to this argument. Sometimes it may be better to describe a surge series as a mildly explosive process instead of a severe explosion. Furthermore, some series may have only a unit root or be trend-stationary and do not pertain to the so-called “explosive” process.

In this study, we examine ten major stock indexes listed in Table 3. These ten indexes are representatives of the world stock markets in different continents: Americas, Asia-Pacific, Europe, and Africa. We select the most representative stock index for each country/district and collect weekly observations. The data are taken from the *Wind Economic Database*. To investigate the dynamics in the exuberance episode, which is our focus here, we use a sample window that ends at the highest point of the exuberance episode. More specifically, we choose each stock index’s highest point in the pre-2008-financial-crisis period as the end point of the rise and take 100 periods before this highest point. For the purposes of comparison, we employ the same sample window width for different stock indexes. The window width $T = 100$ is roughly in line with Allen and Carletti (2010)’s argument that the Federal Reserve’s low interest rate policy in 2005 is the most immediate and important reason to cause prices to take off. Other window widths have also been examined, and the results are available upon request.

Figure 1 plots the ten stock indexes. All of the ten indexes experienced considerable rises, revealing the co-movement among the major stock markets in the world. On the one hand, several series display relatively pronounced explosive features, even though there are some random ups and downs around their explosion paths; see, e.g., DJI, CSI300, HSI, and CASE. On the other hand, some series, such as AS51 and ITLMS, are more like difference-stationary processes with stochastic trends or even trend-stationary processes with deterministic linear trends, rather than explosive processes. It is worth noting that the stock indexes in three Western European countries — France, Germany, and Italy — have similar growth patterns, as Figure 1 shows. However, further investigations are required to detect whether they are explosive processes and, if they are, to identify their degrees of explosiveness.

Table 3. Description of stock indexes.

Region	Country/District	Stock index	Peak date
Americas	U.S.	DJI	Oct 12, 2007
	Brazil	IBOVESPA	May 30, 2008
Asia-Pacific	China	CSI300	Oct 19, 2007
	Hong Kong	HSI	Nov 02, 2007
	Australia	AS51	Nov 02, 2007
Europe	France	FCHI	Jun 01, 2007
	Germany	GDAXI	Jul 13, 2007
	Italy	ITLMS	May 18, 2007
Africa	Egypt	CASE	Apr 24, 2008
	Nigeria	NGSEINDX	Mar 07, 2008

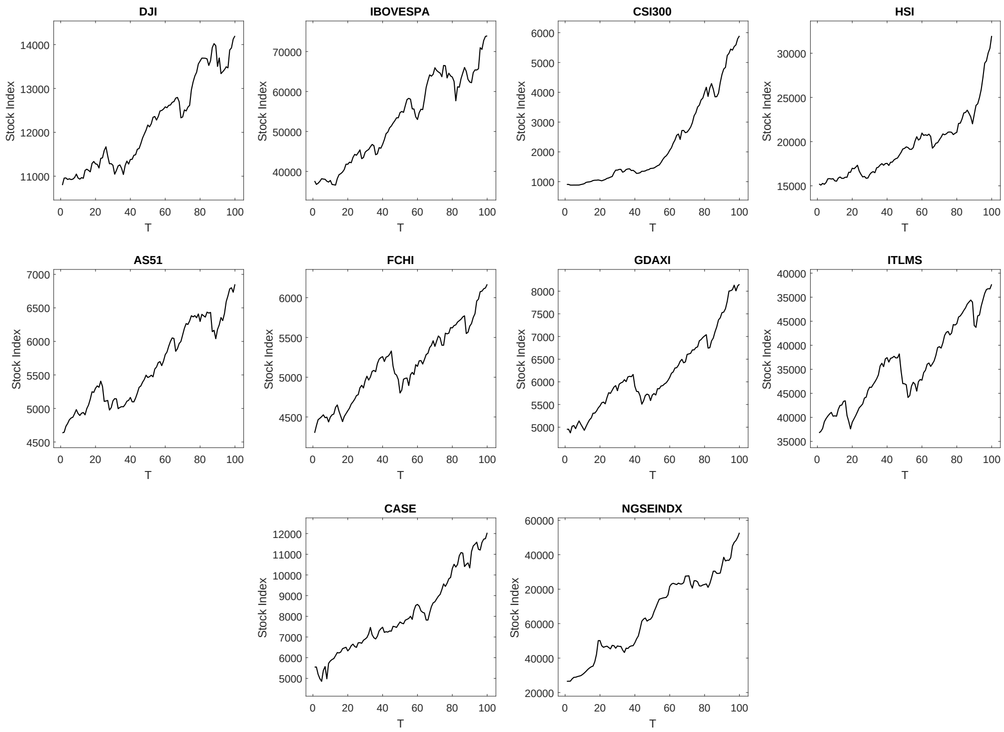

Figure 1. Time series plots of different stock indexes.

5.2. Empirical Testing Strategy

Our empirical study starts with a two-step empirical testing strategy. The first step is a pretest aimed at confirming whether each index is an explosive process. This is necessary, as the asymptotic t test is based on the primitive condition that the AR root is greater than 1, and we must exclude the nonexplosive root at first. We propose to use the right-tailed augmented Dickey-Fuller (RADF) method and the supremum augmented Dickey-Fuller (SADF) method, both of which are adopted in Phillips et al. (2011) and are designed to test the null hypothesis $\rho = 1$ against the alternative hypothesis $\rho > 1$. The RADF test is the conventional augmented Dickey-Fuller (ADF) unit root test but uses the right-tailed critical values. We use the RADF test in order to target at the explosive-root alternative. The SADF method employs a sequence of forward recursive RADF unit root tests, using subsets of the sample data increased by one observation at each pass until the full sample is used. The SADF statistic is then the sup value of the corresponding ADF statistic sequence, whose limiting distribution under the null $\rho = 1$ is obtained by Phillips et al. (2011, Section 2), viz.

$$\begin{aligned}\mathrm{SADF}\,(r_0) &= \sup_{r\in[r_0,1]} \mathrm{ADF}_r \\ &\Rightarrow t_{\mathrm{SADF},\infty} := \sup_{r\in[r_0,1]} \left[\left(\int_0^r \tilde{W}\,(s)\, dW\,(s) \right) \left(\int_0^r \tilde{W}^2\,(s)\, ds \right)^{-1/2} \right], \quad (5.1)\end{aligned}$$

where r_0 is the smallest window size and $W\,(s)$ and $\tilde{W}\,(s)$ are the standard Brownian motion and its demeaned version.

The second and main step of our empirical testing strategy is to perform the asymptotic t test, $\tilde{t}_{\mathrm{MED}}$, on the indexes that are regarded as explosive according to the first step. We invert the $\tilde{t}_{\mathrm{MED}}$ test and construct a confidence interval (set) for each AR parameter ρ. The confidence interval consists of all the values of ρ that are not rejected by our asymptotic t test. Theoretically speaking, we should test for each possible AR value in the region $\rho \in [1 + c_{\min}/k_T, 1 + c_{\max}/k_T]$ for some k_T and positive numbers $c_{\min}$ and $c_{\max}$. Here we consider the following grid

$$\{H_0 : \rho \in \{1.001, 1.002, \ldots, 1.500\}\}\,.$$

We can also consider a more refined grid if needed.[5] Conceptually, smaller values of ρ correspond to low deviations of the AR roots from the unity and mildly explosive behaviours. Larger values of ρ correspond to high deviations of the AR roots from the unity and highly explosive behaviours. Following Phillips et al. (2011, Section 3), we label the explosive AR roots not greater than 1.05 as mildly explosive. This informative label will be useful in conveying the severity of bubbles, if they exist, to policy makers.

[5]Our "nonrejection" confidence set is essentially the same as the more conventional confidence interval based on $\hat{\rho}_T$ because the asymptotic approximation is the same for all null values under consideration. In view that the region $\rho \leq 1$ would be excluded when the unit-root null is rejected in the pretest step, the confidence interval for each AR parameter can be, in practice, constructed as $\rho \in [\hat{\rho}_L, \hat{\rho}_U]$, where $\hat{\rho}_L = \max\{1.001, \hat{\rho}_T - t_{\alpha/2,K}\tilde{\sigma}_{\rho,K}\}$ and $\hat{\rho}_U = \hat{\rho}_T + t_{\alpha/2,K}\tilde{\sigma}_{\rho,K}$ and $t_{\alpha/2,K}$ is the $1-\alpha/2$ percentiles in the Student's t distribution with K degrees of freedom. The boundary 1.001 can be more refined if needed.

Table 4. Testing for moderately explosive behaviours.

Stock index	Step 1: Explosive behaviour test		Step 2: Moderate explosiveness test
	ADF	SADF	Confidence interval
DJI	0.201	1.506	$\rho \in [1.001, 1.028]$
IBOVESPA	0.078	3.592	$\rho \in [1.001, 1.027]$
CSI300	3.548	5.665	$\rho \in [1.011, 1.039]$
HSI	3.103	8.207	$\rho \in [1.023, 1.079]$
AS51	-0.012	0.557	Non-explosion
FCHI	-0.148	0.313	Non-explosion
GDAXI	1.201	2.635	$\rho \in [1.001, 1.032]$
ITLMS	-0.227	0.413	Non-explosion
CASE	1.050	1.638	$\rho \in [1.001, 1.031]$
NGSEINDX	0.210	7.544	$\rho \in [1.001, 1.022]$
	Critical values		
90%	-0.440	1.100	
95%	-0.080	1.370	
99%	0.600	1.880	

Note: This table reports the results of the RADF and SADF tests and the results of the asymptotic t test, $\tilde{t}_{\mathrm{MED}}$. The lag length for each regression in the RADF and SADF tests is selected by the Akaike information criterion, with the maximum lag set to 8. The critical values for the RADF and SADF tests are from Phillips et al. (2011) and Phillips et al. (2015a), respectively. For the asymptotic t test, we report the confidence intervals for the AR parameter ρ.

5.3. Empirical Results

Table 4 reports the results of the RADF and SADF tests in the first step and the asymptotic t test for those explosive stock indexes in the second step. In implementing the SADF test, we follow the empirical rule recommended by Phillips et al. (2015a) to set the user-chosen parameter, $r_0 = 0.01 + 1.8/\sqrt{T}$, and accordingly use the asymptotic critical values given in the same paper.[6] At the 5% significance level, the combination of RADF and SADF tests indicates that DJI, IBOVESPA, CSI300, HSI, GDAXI, CASE, and NGSEINDX follow the explosive processes in their respective sampling periods. However, the major stock indexes of some countries, such as Australia, France, and Italy, could not be described by explosive processes.

For the seven explosive stock indexes, the results of the asymptotic t test in Table 4 show that their explosiveness degrees largely fall in the range $\rho \in [1.001, 1.040]$. This indicates that most stock indexes during the pre-2008 exuberance period are only mildly

[6] In practice, we choose $r_0 = \lfloor (0.01 + 1.8/\sqrt{T})T \rfloor / T$ to ensure that $r_0 T$ is a positive integer. The right-tailed critical values for the ADF statistic are available from Phillips et al. (2011, Table 1).

explosive.[7] Take the DJI and CSI300 as examples. The DJI in the 100 booming weeks before October 12, 2007 could be described by an MED process with some AR parameter $\rho \in [1.001, 1.028]$. This signifies that the U.S. stock market witnessed an explosive process with a quite slow pace of explosion. For CSI300 — the main stock index in the largest developing country (China) — we fail to reject the null of moderate explosiveness for $\rho \in [1.011, 1.039]$. Again, while the process is explosive, it is only mildly explosive.

Similarly, for IBOVESPA and GDAXI, the confidence intervals of ρ are $[1.001, 1.027]$ and $[1.001, 1.032]$, respectively. These two stock markets responded closely to the "exuberance" in the US. For CASE and NGSEINDX, two representative indexes in the African stock markets, the degrees of explosiveness are also quite mild. African countries' thin market capitalization and shortage of liquidity led to the volatility and vulnerability of the stock markets (Allen et al., 2011), making them easily affected by the mild exuberance from external economies.

The HSI of Hong Kong is relatively special. We fail to reject the null for $\rho \in [1.023, 1.079]$. Thus, the Hong Kong market appeared to be more explosive. This could be due to the smaller scale of the market, which made an explosive outburst relatively easier.

Finally, the three series, AS51, FCHI, and ITLMS, are neither explosive processes nor MED processes. When the unit-root null against the explosive-root alternative is not rejected by either the RADF or SADF method, we can use the conventional unit root tests to examine these three indexes further. In this paper, we employ the ADF test and Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test, the critical values of which are obtained from MacKinnon (1996) and Kwiatkowski et al. (1992), respectively. Table 5 reports the unit root test results. According to the ADF results, all three time series have a unit root at the 5% level, but their differences have no unit root and appear to be stationary. The KPSS results also provide significant evidence that the three time series are difference-stationary instead of being trend-stationary. Thus, we may conclude that the AS51, FCHI, and ITLMS are all $I(1)$ processes during their respective sampling periods. These quantitative testing results lend some supplementary support to the conclusion that the rises in Australia, France, and Italy's stock markets were not explosive.

In summary, we find evidence that seven of the ten major stock indexes under our consideration are moderately explosive, while the remaining ones are nonexplosive and difference-stationary processes. However, for the former group of indexes, the degree of explosiveness is quite mild. This finding is consistent with the remark of Jagannathan et al. (2013): the 2008 financial crisis was more like a symptom than the disease. Despite the severity and ample effects (Martin and Ventura, 2012; Miao and Wang, 2015; Kunieda and Shibata, 2016), this financial crisis was similar to past crises (Allen and Carletti, 2010) that did not show an extremely serious irrational explosion.

6. CONCLUSION

This paper considers a moderately explosive process wherein the AR root is greater than one by a margin diminishing with the sample size. We allow for a drift in the model so that the true process is driven by both the stochastic moderately-explosive trend and the deterministic nonlinear drift trend. New asymptotic approximations are established to test for the degree of the moderate explosiveness under i.i.d. errors and under weakly

[7] When we use other window widths such as $T = 80$ and 120, we obtain the same qualitative conclusion. Details are not reported here to conserve space.

Table 5. Testing for unit root.

	ADF		KPSS
	Level	First difference	
AS51	-0.012	-9.163	3.225
FCHI	-0.148	-9.296	3.086
ITLMS	-0.227	-4.778	3.002
	Critical values		
1% (99%)	-3.498	-3.498	0.739
5% (95%)	-2.891	-2.891	0.463
10% (90%)	-2.583	-2.583	0.347

Note: This table reports the results of the ADF and KPSS tests. The lag length for the ADF test is selected by the Akaike information criterion, with the maximum lag set to 8. The critical values for the ADF (left-tailed) and KPSS (right-tailed) tests are from MacKinnon (1996) and Kwiatkowski et al. (1992), respectively.

dependent errors. When the errors are weakly dependent, we show that under the fixed-smoothing asymptotics, the HAR t statistic follows Student's t distribution in large samples. The asymptotic t test achieves double robustness: it is asymptotically valid no matter whether the errors are autocorrelated or not, and whether the drift is large or small, or simply equal to zero. Monte Carlo experiments lend some support to our asymptotic results.

The paper also proposes a two-step empirical testing strategy that involves first identifying whether a time series is explosive or not and then employing our asymptotic t test to measure the degree of moderate explosiveness if it is indeed explosive. We apply our empirical strategy to ten major stock indexes in the world during the pre-2008 financial exuberance period. The results show that seven of these indexes follow the MED processes with AR roots slightly larger than unity. In addition, the other three stock indexes are nonexplosive and difference-stationary processes. These results conform with Greenspan (2008)'s perception and imply that the stock market boom before the 2008 financial crisis is not as explosive as in the existing literature.

ACKNOWLEDGEMENTS

Correspondence to: Shaoping Wang, School of Economics, Huazhong University of Science and Technology, Wuhan, China, 430074. Guo and Wang acknowledge research support from the National Natural Science Foundation of China under Grant No. 71671070. Sun acknowledges partial research support from the NSF under Grant No. SES-1530592. For helpful comments, the authors gratefully thank Dennis Kristensen (the co-editor), two anonymous referees, Liangjun Su (SMU), Xiaohu Wang (CUHK), seminar participants at Huazhong University of Science and Technology, Shandong University, Sun Yat-Sen University, Xiamen University, the 2017 China Meeting of the Econometric Society at Wuhan University, and the 2017 Tsinghua International Conference on Econometrics at Tsinghua University. All errors are our own.

REFERENCES

Allen, F., A. Babus, and E. Carletti (2009). Financial crises: Theory and evidence. *Annual Review of Financial Economics 1*(1), 97–116.

Allen, F. and E. Carletti (2010). An overview of the crisis: Causes, consequences, and solutions. *International Review of Finance 10*(1), 1–26.

Allen, F., I. Otchere, and L. W. Senbet (2011). African financial systems: A review. *Review of Development Finance 1*(2), 79–113.

Anderson, T. W. (1959). On asymptotic distributions of estimates of parameters of stochastic difference equations. *Annals of Mathematical Statistics 30*(3), 676–687.

Arvanitis, S. and T. Magdalinos (2018). Mildly explosive autoregression under stationary conditional heteroskedasticity. *Journal of Time Series Analysis 39*(6), 892–908.

Dickey, D. A. and W. A. Fuller (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association 74*(366), 427–431.

Dickey, D. A. and W. A. Fuller (1981). Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica 49*(4), 1057–1072.

Fei, Y. (2018). Limit theory for mildly integrated process with intercept. *Economics Letters 163*, 98–101.

Greenspan, A. (1996). Minutes of the federal open market committee. Available from www.federal.reserve.gov/transcript/1996/19960924meeting.pdf.

Greenspan, A. (2008). *The age of turbulence: Adventures in a new world.* New York: Penguin Press.

Jagannathan, R., M. Kapoor, and E. Schaumburg (2013). Causes of the great recession of 2007–2009: The financial crisis was the symptom not the disease! *Journal of Financial Intermediation 22*(1), 4–29.

Kunieda, T. and A. Shibata (2016). Asset bubbles, economic growth, and a self-fulfilling financial crisis. *Journal of Monetary Economics 82*, 70–84.

Kwiatkowski, D., P. C. B. Phillips, P. Schmidt, and Y. Shin (1992). Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? *Journal of Econometrics 54*(1-3), 159–178.

Liu, X. and L. Peng (2018). Asymptotic theory and unified confidence region for an autoregressive model. *Journal of Time Series Analysis*.

MacKinnon, J. G. (1996). Numerical distribution functions for unit root and cointegration tests. *Journal of Applied Econometrics 11*(6), 601–618.

Magdalinos, T. (2012). Mildly explosive autoregression under weak and strong dependence. *Journal of Econometrics 169*(2), 179–187.

Magdalinos, T. and P. C. B. Phillips (2009). Limit theory for cointegrated systems with moderately integrated and moderately explosive regressors. *Econometric Theory 25*(2), 482–526.

Martin, A. and J. Ventura (2012). Economic growth with bubbles. *American Economic Review 102*(6), 3033–3058.

Miao, J. and P. Wang (2015). Banking bubbles and financial crises. *Journal of Economic Theory 157*, 763–792.

Müller, U. K. (2007). A theory of robust long-run variance estimation. *Journal of Econometrics 141*(2), 1331–1352.

Phillips, P. C. B. (1987). Time series regression with a unit root. *Econometrica 55*(2), 277–301.

Phillips, P. C. B. (2005). HAC estimation by automated regression. *Econometric Theory 21*(1), 116–142.

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

Phillips, P. C. B. and T. Magdalinos (2007b). *Limit theory for moderate deviations from a unit root under weak dependence*, pp. 123–162. Cambridge: Cambridge University Press.

Phillips, P. C. B., T. Magdalinos, and L. Giraitis (2010). Smoothing local-to-moderate unit root theory. *Journal of Econometrics 158*(2), 274–279.

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

Phillips, P. C. B., S. Shi, and J. Yu (2014). Specification sensitivity in right-tailed unit root testing for explosive behaviour. *Oxford Bulletin of Economics and Statistics 76*(3), 315–333.

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

Phillips, P. C. B., S. Shi, and J. Yu (2015b). Testing for multiple bubbles: Limit theory of real-time detectors. *International Economic Review 56*(4), 1079–1134.

Phillips, P. C. B. and V. Solo (1992). Asymptotics for linear processes. *Annals of Statistics 20*(2), 971–1001.

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*(1), 201–226.

Shiller, R. J. (2008). *The subprime solution: How today's global financial crisis happened, and what to do about it.* Princeton: Princeton University Press.

Stiglitz, J. E. (2010). Robust growth or anemic recovery in the US and the global economy. *Journal of Policy Modeling 32*(5), 632–636.

Sun, Y. (2011). Robust trend inference with series variance estimator and testing-optimal smoothing parameter. *Journal of Econometrics 164*(2), 345–366.

Sun, Y. (2013). A heteroskedasticity and autocorrelation robust F test using orthonormal series variance estimator. *Econometrics Journal 16*(1), 1–26.

Sun, Y. (2014). Fixed-smoothing asymptotics in a two-step generalized method of moments framework. *Econometrica 82*(6), 2327–2370.

Taylor, J. B. (2009). Economic policy and the financial crisis: An empirical analysis of what went wrong. *Critical Review 21*(2-3), 341–364.

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

White, J. S. (1958). The limiting distribution of the serial correlation coefficient in the explosive case. *Annals of Mathematical Statistics 29*(4), 1188–1197.

Woo, W. T., R. Garnaut, and L. Song (2013). *China's new place in a world in crisis: Economic, geopolitical and environmental dimensions.* Canberra: ANU Press.

APPENDIX

Appendix A presents some technical lemmas that are used in the proofs of the key results in Sections 2 and 3. The proofs of these lemmas are available from the online supplement. Appendix B presents the proofs of the key results of the paper.

Appendix A. Technical Lemmas

LEMMA A.1. *Let $\rho_T = 1 + c/k_T$ for some $c > 0$ and k_T satisfy $1/k_T + k_T/T = o(1)$. Then $\rho_T^{-aT} = o\left(k_T^b/T^b\right)$ for any a and $b \in \mathbb{R}^+$. In addition, if $k_T \to \infty$ and $k_T/T \to 0$ as $T \to \infty$, then $\rho_T^{-aT} = o\left(1/T^b\right)$ for any a and $b \in \mathbb{R}^+$.*

LEMMA A.2. *Let Assumption 2.1 hold. Then*

(a) $\left(k_T^{3/2}\rho_T^T\right)^{-1}\sum_{t=1}^{T}\sum_{j=t}^{T}\rho_T^{t-1-j}u_j = o_p(1)$;

(b) $\left(k_T^{3/2}\rho_T^{2T}\right)^{-1}\sum_{t=1}^{T}\sum_{j=t}^{T}\rho_T^{2(t-1)-j}u_j = o_p(1)$.

LEMMA A.3. *Let Assumption 2.1 hold. Then*

$$\left(k_T^{3/2}\rho_T^T\right)^{-1}\sum_{t=1}^{T}\xi_{t-1} = \frac{1}{c}Y_T + o_p(1).$$

LEMMA A.4. *Let Assumption 3.1 hold. Then*

(a) $\left(k_T\rho_T^T\right)^{-1}\sum_{t=1}^{T}\sum_{j=t}^{T}\rho_T^{t-1-j}u_ju_t = o_p(1)$;

(b) $\left(k_T^{3/2}\rho_T^T\right)^{-1}\sum_{t=1}^{T}\sum_{j=t}^{T}\rho_T^{t-1-j}u_j = o_p(1)$;

(c) $\left(k_T^{3/2}\rho_T^{2T}\right)^{-1}\sum_{t=1}^{T}\sum_{j=t}^{T}\rho_T^{2(t-1)-j}u_j = o_p(1)$.

LEMMA A.5. *Let Assumption 3.1 hold. Then*

(a) $\left(k_T\rho_T^T\right)^{-2}\sum_{t=1}^{T}\xi_{t-1}^2 = \frac{1}{2c}\tilde{Y}_T^2 + o_p(1)$;

(b) $\left(k_T^{3/2}\rho_T^T\right)^{-1}\sum_{t=1}^{T}\xi_{t-1} = \frac{1}{c}\tilde{Y}_T + o_p(1)$;

(c) $\left(k_T\rho_T^T\right)^{-1}\sum_{t=1}^{T}\xi_{t-1}u_t = \tilde{X}_T\tilde{Y}_T + o_p(1)$.

LEMMA A.6. *Let Assumption 3.1 hold. Then*

(a) $\sum_{t=1}^{T}\phi_\ell\left(\frac{t}{T}\right)\rho_T^t = O\left(k_T\rho_T^T\right)$;

(b) $\sum_{t=1}^{T}\phi_\ell\left(\frac{t}{T}\right)\sum_{j=t}^{T}\rho_T^{t-1-j}u_j = o_p\left(\sqrt{T}k_T\rho_T^T\right)$.

Appendix B: Proofs of the Key Results

Proof of Theorem 2.1: Part (a). Using (2.9), we obtain

$$\left(\mu_T^2 k_T^3 \rho_T^{2T}\right)^{-1} \sum_{t=1}^{T} y_{t-1}^2 = \left(\mu_T^2 k_T^3 \rho_T^{2T}\right)^{-1} \sum_{t=1}^{T} \left(\xi_{t-1} + \frac{1}{c}\mu_T k_T \rho_T^{t-1} - \frac{1}{c}\mu_T k_T\right)^2$$

$$= \left(\mu_T^2 k_T^3 \rho_T^{2T}\right)^{-1} \left(\sum_{t=1}^{T} \xi_{t-1}^2 + \frac{1}{c^2}\mu_T^2 k_T^2 \sum_{t=1}^{T} \rho_T^{2t-2} + \frac{2}{c}\mu_T k_T \sum_{t=1}^{T} \xi_{t-1}\rho_T^{t-1}\right.$$

$$\left. + \frac{1}{c^2} T \mu_T^2 k_T^2 - \frac{2}{c}\mu_T k_T \sum_{t=1}^{T} \xi_{t-1} - \frac{2}{c^2}\mu_T^2 k_T^2 \sum_{t=1}^{T} \rho_T^{t-1}\right)$$

$$= \left(\mu_T^2 k_T^3 \rho_T^{2T}\right)^{-1} \sum_{t=1}^{T} \xi_{t-1}^2 + \frac{1}{c^2}\left(k_T \rho_T^{2T}\right)^{-1} \frac{\rho_T^{2T} - 1}{\rho_T^2 - 1} + \frac{2}{c}\left(\mu_T k_T^2 \rho_T^{2T}\right)^{-1} \sum_{t=1}^{T} \xi_{t-1}\rho_T^{t-1}$$

$$- \frac{2}{c}\left(\mu_T k_T^2 \rho_T^{2T}\right)^{-1} \sum_{t=1}^{T} \xi_{t-1} + O\left(T k_T^{-1} \rho_T^{-2T} + \rho_T^{-T}\right)$$

$$= \frac{Y_T^2}{2c\mu_T^2 k_T} + \frac{1}{2c^3} + \frac{2}{c}\left(\mu_T k_T^2 \rho_T^{2T}\right)^{-1} \sum_{t=1}^{T} \xi_{t-1}\rho_T^{t-1} + o_p(1),$$

by (2.4) and Lemmas A.1 and A.3.

Note that

$$\frac{2}{c}\left(\mu_T k_T^2 \rho_T^{2T}\right)^{-1} \sum_{t=1}^{T} \xi_{t-1}\rho_T^{t-1} = \frac{2}{c}\left(\mu_T k_T^2 \rho_T^{2T}\right)^{-1} \sum_{t=1}^{T} \left(\rho_T^{t-1}\xi_0 + \sum_{j=1}^{t-1} \rho_T^{t-1-j} u_j\right)\rho_T^{t-1}$$

$$= \frac{2}{c}\xi_0 \left(\mu_T k_T^2 \rho_T^{2T}\right)^{-1} \sum_{t=1}^{T} \rho_T^{2t-2} + \frac{2}{c}\left(\mu_T k_T^2 \rho_T^{2T}\right)^{-1} \sum_{t=1}^{T}\sum_{j=1}^{t-1} \rho_T^{2(t-1)-j} u_j$$

$$= \frac{2}{c}\left(\mu_T k_T^2 \rho_T^{2T}\right)^{-1} \sum_{t=1}^{T}\sum_{j=1}^{T} \rho_T^{2(t-1)-j} u_j - \frac{2}{c}\left(\mu_T k_T^2 \rho_T^{2T}\right)^{-1} \sum_{t=1}^{T}\sum_{j=t}^{T} \rho_T^{2(t-1)-j} u_j + o_p\left(\mu_T^{-1} k_T^{-1/2}\right)$$

$$= \frac{2}{c}\left(\mu_T k_T^2 \rho_T^{2T}\right)^{-1} \sum_{t=1}^{T} \rho_T^{2(t-1)} \sum_{j=1}^{T} \rho_T^{-j} u_j + o_p\left(\mu_T^{-1} k_T^{-1/2}\right)$$

$$= \frac{2}{c}\left(\mu_T k_T^2 \rho_T^{2T}\right)^{-1} \frac{\rho_T^{2T} - 1}{\rho_T^2 - 1} k_T^{1/2} Y_T + o_p(1) = \frac{Y_T}{c^2 \mu_T k_T^{1/2}} + o_p(1),$$

by $\xi_0 = o_p\left(k_T^{1/2}\right)$ and Lemma A.2(b). The key assumption behind this result is that $\mu_T k_T^{1/2} \to \nu > 0$. Thus,

$$\left(\mu_T^2 k_T^3 \rho_T^{2T}\right)^{-1} \sum_{t=1}^{T} y_{t-1}^2 \Rightarrow \frac{Y^2}{2c\nu^2} + \frac{1}{2c^3} + \frac{Y}{c^2\nu} = \frac{1}{2c}\left(\frac{Y}{\nu} + \frac{1}{c}\right)^2.$$

Part (b). The normalized sample mean is

$$
\begin{aligned}
\left(\mu_T k_T^2 \rho_T^T\right)^{-1} \sum_{t=1}^{T} y_{t-1} &= \left(\mu_T k_T^2 \rho_T^T\right)^{-1} \sum_{t=1}^{T} \left(\xi_{t-1} + \frac{1}{c}\mu_T k_T \rho_T^{t-1} - \frac{1}{c}\mu_T k_T\right) \\
&= \left(\mu_T k_T^2 \rho_T^T\right)^{-1} \sum_{t=1}^{T} \xi_{t-1} + \frac{1}{c}\left(k_T \rho_T^T\right)^{-1} \frac{\rho_T^T - 1}{\rho_T - 1} - \frac{1}{c} T k_T^{-1} \rho_T^{-T} \\
&= \frac{Y_T}{c\mu_T k_T^{1/2}} + \frac{1}{c^2} + o_p(1) \Rightarrow \frac{1}{c}\left(\frac{Y}{\nu} + \frac{1}{c}\right),
\end{aligned}
$$

by (2.3) and Lemmas A.1 and A.3.

Part (c). The normalized sample covariance is

$$
\begin{aligned}
&\left(\mu_T k_T^{3/2} \rho_T^T\right)^{-1} \sum_{t=1}^{T} y_{t-1} u_t = \left(\mu_T k_T^{3/2} \rho_T^T\right)^{-1} \sum_{t=1}^{T} \left(\xi_{t-1} + \frac{1}{c}\mu_T k_T \rho_T^{t-1} - \frac{1}{c}\mu_T k_T\right) u_t \\
&= \left(\mu_T k_T^{3/2} \rho_T^T\right)^{-1} \sum_{t=1}^{T} \xi_{t-1} u_t + \frac{1}{c}\left(k_T^{1/2} \rho_T^T\right)^{-1} \sum_{t=1}^{T} \rho_T^{t-1} u_t - \frac{1}{c}\left(k_T^{1/2} \rho_T^T\right)^{-1} \sum_{t=1}^{T} u_t \\
&= \frac{X_T Y_T}{\mu_T k_T^{1/2}} + \frac{1}{c} k_T^{-1/2} \sum_{t=1}^{T} \rho_T^{-(T-t)-1} u_t + O_p\left(T^{1/2} k_T^{-1/2} \rho_T^{-T}\right) + o_p(1) \Rightarrow X\left(\frac{Y}{\nu} + \frac{1}{c}\right),
\end{aligned}
$$

by (2.3), (2.5), and Lemma A.1.

The joint convergence of (a), (b), and (c) follows from the Cramér-Wold theorem.□

Proof of Lemma 3.1: Parts (a) and (b). We prove (b) first. Using the decomposition in (3.1), we have

$$
\begin{aligned}
k_T^{-1/2} \sum_{t=1}^{T} \rho_T^{-t} u_t &= k_T^{-1/2} \sum_{t=1}^{T} \rho_T^{-t} C(1) \varepsilon_t + k_T^{-1/2} \sum_{t=1}^{T} \rho_T^{-t} \left(\tilde{\varepsilon}_{t-1} - \tilde{\varepsilon}_t\right) \\
&= C(1) k_T^{-1/2} \sum_{t=1}^{T} \rho_T^{-t} \varepsilon_t + k_T^{-1/2} \sum_{t=1}^{T} \rho_T^{-t} \left(\tilde{\varepsilon}_{t-1} - \tilde{\varepsilon}_t\right).
\end{aligned}
$$

But

$$
\begin{aligned}
\sum_{t=1}^{T} \rho_T^{-t} \left(\tilde{\varepsilon}_{t-1} - \tilde{\varepsilon}_t\right) &= \sum_{t=1}^{T} \rho_T^{-t} \tilde{\varepsilon}_{t-1} - \sum_{t=1}^{T} \rho_T^{-t} \tilde{\varepsilon}_t = \sum_{t=0}^{T-1} \rho_T^{-(t+1)} \tilde{\varepsilon}_t - \sum_{t=1}^{T} \rho_T^{-t} \tilde{\varepsilon}_t \\
&= \rho_T^{-1} \tilde{\varepsilon}_0 - \rho_T^{-T} \tilde{\varepsilon}_T + \sum_{t=1}^{T-1} \left(\rho_T^{-(t+1)} - \rho_T^{-t}\right) \tilde{\varepsilon}_t \\
&= \rho_T^{-1} \tilde{\varepsilon}_0 - \rho_T^{-T} \tilde{\varepsilon}_T - c k_T^{-1} \sum_{t=1}^{T-1} \rho_T^{-(t+1)} \tilde{\varepsilon}_t.
\end{aligned}
$$

Since $var\left(\tilde{\varepsilon}_t\right) < \infty$, we have

$$
k_T^{-1/2} \rho_T^{-1} \tilde{\varepsilon}_0 = o_p(1) \text{ and } k_T^{-1/2} \rho_T^{-T} \tilde{\varepsilon}_T = o_p(1).
$$

Now, using the Cauchy inequality, we obtain

$$var\left(\sum_{t=1}^{T-1}\rho_T^{-(t+1)}\tilde{\varepsilon}_t\right) = \left(\sum_{t=1}^{T-1}\rho_T^{-2(t+1)}\right)var(\tilde{\varepsilon}_t) + 2\sum_{t<s}^{T-1}\rho_T^{-(t+s+2)}cov\left(\tilde{\varepsilon}_t,\tilde{\varepsilon}_s\right)$$

$$= \left(\sum_{t=1}^{T-1}\rho_T^{-2(t+1)}\right)var(\tilde{\varepsilon}_t) + \left(2\sum_{t<s}^{T-1}\sum_{j=0}^{\infty}\tilde{c}_j\tilde{c}_{(s-t)+j}\rho_T^{-(t+s+2)}\right)var(\varepsilon_t)$$

$$= O\left(\frac{\rho_T^{-4}-\rho_T^{-2(T+1)}}{1-\rho_T^{-2}}\right) + O\left(\sum_{t<s}^{T-1}\left(\sum_{j=0}^{\infty}\tilde{c}_j^2\right)\rho_T^{-(t+s+2)}\right)$$

$$= O\left(k_T\right) + O\left(\left(\sum_{t=1}^{T-1}\rho_T^{-(t+1)}\right)^2\right) = O\left(k_T\right) + O\left(\left(\frac{1}{\rho_T^T}-\frac{1}{\rho_T}\right)^2\left(\rho_T-1\right)^{-2}\right)$$

$$= O\left(k_T^2\right).$$

Therefore,

$$k_T^{-1/2}\left(ck_T^{-1}\sum_{t=1}^{T-1}\rho_T^{-(t+1)}\tilde{\varepsilon}_t\right) = O_p\left(k_T^{-1/2}\right) = o_p\left(1\right).$$

Combining the above results yields

$$\tilde{Y}_T = k_T^{-1/2}\sum_{t=1}^{T}\rho_T^{-t}u_t = C\left(1\right)k_T^{-1/2}\sum_{t=1}^{T}\rho_T^{-t}\varepsilon_t + o_p\left(1\right).$$

To prove part (a), we use the same arguments, starting with

$$k_T^{-1/2}\sum_{t=1}^{T}\rho_T^{-(T-t)-1}u_t = C\left(1\right)k_T^{-1/2}\sum_{t=1}^{T}\rho_T^{-(T-t)-1}\varepsilon_t + k_T^{-1/2}\sum_{t=1}^{T}\rho_T^{-(T-t)-1}\left(\tilde{\varepsilon}_{t-1}-\tilde{\varepsilon}_t\right).$$

But

$$\sum_{t=1}^{T}\rho_T^{-(T-t)-1}\left(\tilde{\varepsilon}_{t-1}-\tilde{\varepsilon}_t\right) = \rho_T^{-T}\tilde{\varepsilon}_0 - \rho_T^{-1}\tilde{\varepsilon}_T + \left(\rho_T-1\right)\sum_{t=1}^{T-1}\rho_T^{-(T-t)-1}\tilde{\varepsilon}_t$$

$$= ck_T^{-1}\sum_{t=1}^{T-1}\rho_T^{-(T-t)-1}\tilde{\varepsilon}_t + O_p\left(1\right).$$

By similar calculations, we have

$$var\left(\sum_{t=1}^{T-1}\rho_T^{-(T-t)-1}\tilde{\varepsilon}_t\right)$$

$$= \left(\sum_{t=1}^{T-1}\rho_T^{-2(T-t+1)}\right)var(\tilde{\varepsilon}_t) + 2\sum_{t<s}^{T-1}\rho_T^{-(T-t)-1}\rho_T^{-(T-s)-1}cov\left(\tilde{\varepsilon}_t,\tilde{\varepsilon}_s\right)$$

$$= \left(\sum_{t=1}^{T-1}\rho_T^{-2(t+1)}\right)var(\tilde{\varepsilon}_t) + \left(2\sum_{t<s}^{T-1}\sum_{j=0}^{\infty}\tilde{c}_j\tilde{c}_{(s-t)+j}\rho_T^{-(T-t)-1}\rho_T^{-(T-s)-1}\right)var(\varepsilon_t)$$

$$= O(k_T) + O\left(\sum_{t<s}^{T-1}\left(\sum_{j=0}^{\infty}\tilde{c}_j^2\right)\rho_T^{-(T-t)-1}\rho_T^{-(T-s)-1}\right)$$

$$= O(k_T) + O\left(\left(\sum_{t=1}^{T-1}\rho_T^{-(T-t)-1}\right)^2\right) = O\left(k_T^2\right).$$

Therefore

$$k_T^{-1/2}\left(ck_T^{-1}\sum_{t=1}^{T-1}\rho_T^{-(T-t)-1}\tilde{\varepsilon}_t\right) = O_p\left(k_T^{-1/2}\right) = o_p(1).$$

Combining the above results yields

$$\tilde{X}_T = k_T^{-1/2}\sum_{t=1}^{T}\rho_T^{-(T-t)-1}u_t = C(1)\,k_T^{-1/2}\sum_{t=1}^{T}\rho_T^{-(T-t)-1}\varepsilon_t + o_p(1).$$

Part (c). This follows immediately from Parts (a) and (b) and equation (2.3).□

Proof of Theorem 3.1: The proof is similar to that of Theorem 2.1, but we employ Lemmas A.4(c) and A.5 which accommodate weak dependence in $\{u_t\}$. For completeness, we sketch the proof here.

Part (a). Using (2.9), we obtain

$$\left(\mu_T^2k_T^3\rho_T^{2T}\right)^{-1}\sum_{t=1}^{T}y_{t-1}^2 = \left(\mu_T^2k_T^3\rho_T^{2T}\right)^{-1}\sum_{t=1}^{T}\left(\xi_{t-1} + \frac{1}{c}\mu_Tk_T\rho_T^{t-1} - \frac{1}{c}\mu_Tk_T\right)^2$$

$$= \left(k_T^3\rho_T^{2T}\right)^{-1}\left(\mu_T^{-2}\sum_{t=1}^{T}\xi_{t-1}^2 + \frac{2}{c}\mu_T^{-1}k_T\sum_{t=1}^{T}\xi_{t-1}\rho_T^{t-1} - \frac{2}{c}\mu_T^{-1}k_T\sum_{t=1}^{T}\xi_{t-1} + \frac{1}{c^2}k_T^2\sum_{t=1}^{T}\rho_T^{2t-2}\right) + o(1)$$

$$= \left(k_T^3\rho_T^{2T}\right)^{-1}\left(\mu_T^{-2}\sum_{t=1}^{T}\xi_{t-1}^2 + \frac{2}{c}\mu_T^{-1}k_T\sum_{t=1}^{T}\xi_{t-1}\rho_T^{t-1} - \frac{2}{c}\mu_T^{-1}k_T\sum_{t=1}^{T}\xi_{t-1}\right) + \frac{1}{2c^3} + o(1).$$

It follows from Lemmas A.5(a&b) that

$$\left(k_T^3\rho_T^{2T}\right)^{-1}\left(\mu_T^{-2}\sum_{t=1}^{T}\xi_{t-1}^2\right) = \frac{\tilde{Y}_T^2}{2c\mu_T^2k_T} + o_p(1),$$

and

$$\left(k_T^3\rho_T^{2T}\right)^{-1}\left(\frac{2}{c}\mu_T^{-1}k_T\sum_{t=1}^{T}\xi_{t-1}\right) = O_p\left(\frac{\tilde{Y}_T}{\mu_Tk_T^{1/2}\rho_T^T}\right) = o_p(1).$$

Moreover, using $\xi_0 = o_p\left(k_T^{1/2}\right)$ and Lemma A.4(c), we obtain

$$\left(k_T^3\rho_T^{2T}\right)^{-1}\left(\frac{2}{c}\mu_T^{-1}k_T\sum_{t=1}^{T}\xi_{t-1}\rho_T^{t-1}\right) = \frac{2}{c}\left(\mu_Tk_T^2\rho_T^{2T}\right)^{-1}\sum_{t=1}^{T}\left(\rho_T^{t-1}\xi_0 + \sum_{j=1}^{t-1}\rho_T^{t-1-j}u_j\right)\rho_T^{t-1}$$

$$= \frac{2}{c}\xi_0\left(\mu_Tk_T^2\rho_T^{2T}\right)^{-1}\sum_{t=1}^{T}\rho_T^{2t-2} + \frac{2}{c}\left(\mu_Tk_T^2\rho_T^{2T}\right)^{-1}\sum_{t=1}^{T}\sum_{j=1}^{t-1}\rho_T^{2(t-1)-j}u_j$$

$$= \frac{2}{c}\left(\mu_T k_T^2 \rho_T^{2T}\right)^{-1} \sum_{t=1}^{T}\sum_{j=1}^{T} \rho_T^{2(t-1)-j} u_j - \frac{2}{c}\left(\mu_T k_T^2 \rho_T^{2T}\right)^{-1} \sum_{t=1}^{T}\sum_{j=t}^{T} \rho_T^{2(t-1)-j} u_j + o_p\left(\mu_T^{-1} k_T^{-1/2}\right)$$

$$= \frac{2}{c}\left(\mu_T k_T^2 \rho_T^{2T}\right)^{-1} \sum_{t=1}^{T} \rho_T^{2(t-1)} \sum_{j=1}^{T} \rho_T^{-j} u_j + o_p\left(\mu_T^{-1} k_T^{-1/2}\right)$$

$$= \frac{2}{c}\left(\mu_T k_T^2 \rho_T^{2T}\right)^{-1} \frac{\rho_T^{2T}-1}{\rho_T^2-1} k_T^{1/2}\tilde{Y}_T + o_p(1) = \frac{\tilde{Y}_T}{c^2 \mu_T k_T^{1/2}} + o_p(1),$$

when $\nu \in (0,\infty]$.

Combining the above results and Lemma 3.1(c) leads to

$$\left(\mu_T^2 k_T^3 \rho_T^{2T}\right)^{-1} \sum_{t=1}^{T} y_{t-1}^2 \Rightarrow \frac{\tilde{Y}^2}{2c\nu^2} + \frac{\tilde{Y}}{c^2\nu} + \frac{1}{2c^3} = \frac{1}{2c}\left(\frac{\tilde{Y}}{\nu} + \frac{1}{c}\right)^2.$$

Part (b). By Lemmas 3.1(c), A.1, and A.5(b), we have

$$\begin{aligned}\left(\mu_T k_T^2 \rho_T^{T}\right)^{-1} \sum_{t=1}^{T} y_{t-1} &= \left(\mu_T k_T^2 \rho_T^{T}\right)^{-1} \sum_{t=1}^{T} \left(\xi_{t-1} + \frac{1}{c}\mu_T k_T \rho_T^{t-1} - \frac{1}{c}\mu_T k_T\right) \\ &= \left(\mu_T k_T^2 \rho_T^{T}\right)^{-1} \sum_{t=1}^{T} \xi_{t-1} + \frac{1}{c}\left(k_T \rho_T^{T}\right)^{-1} \frac{\rho_T^{T}-1}{\rho_T - 1} - \frac{1}{c} T k_T^{-1} \rho_T^{-T} \\ &= \frac{\tilde{Y}_T}{c\mu_T k_T^{1/2}} + \frac{1}{c^2} + o_p(1) \Rightarrow \frac{1}{c}\left(\frac{\tilde{Y}}{\nu} + \frac{1}{c}\right).\end{aligned}$$

Part (c). It follows from Lemmas 3.1(c), A.1, and A.5(c) that

$$\left(\mu_T k_T^{3/2} \rho_T^{T}\right)^{-1} \sum_{t=1}^{T} y_{t-1} u_t = \left(\mu_T k_T^{3/2} \rho_T^{T}\right)^{-1} \sum_{t=1}^{T} \left(\xi_{t-1} + \frac{1}{c}\mu_T k_T \rho_T^{t-1} - \frac{1}{c}\mu_T k_T\right) u_t$$

$$= \left(\mu_T k_T^{3/2} \rho_T^{T}\right)^{-1} \sum_{t=1}^{T} \xi_{t-1} u_t + \frac{1}{c}\left(k_T^{1/2} \rho_T^{T}\right)^{-1} \sum_{t=1}^{T} \rho_T^{t-1} u_t + O_p\left(T^{1/2} k_T^{-1/2} \rho_T^{-T}\right)$$

$$= \frac{\tilde{X}_T \tilde{Y}_T}{\mu_T k_T^{1/2}} + \frac{\tilde{X}_T}{c} + o_p(1) \Rightarrow \tilde{X}\left(\frac{\tilde{Y}}{\nu} + \frac{1}{c}\right).$$

The joint convergence of the results in the theorem follows from the Cramér-Wold theorem.□

Proof of Theorem 3.3: We prove the case with $\nu \in (0,\infty]$ only. The proof for the case with $\nu = 0$ is essentially the same with only minor modifications. Detailed calculations for the latter case are available upon request.

Part (a). Note that

$$\hat{u}_t = y_t - \hat{\mu}_T - \hat{\rho}_T y_{t-1} = u_t - \begin{pmatrix} 1, & y_{t-1} \end{pmatrix} \left(\sum_{t=1}^{T} \boldsymbol{x}_t \boldsymbol{x}_t'\right)^{-1} \sum_{t=1}^{T} \boldsymbol{x}_t u_t.$$

So

$$
\begin{aligned}
&\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\phi_\ell\left(\frac{t}{T}\right)\hat{u}_t \\
&= \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\phi_\ell\left(\frac{t}{T}\right)u_t - \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\phi_\ell\left(\frac{t}{T}\right)\begin{pmatrix} 1, & y_{t-1}\end{pmatrix}\left(\sum_{t=1}^{T}\boldsymbol{x}_t\boldsymbol{x}_t'\right)^{-1}\sum_{t=1}^{T}\boldsymbol{x}_t u_t \\
&= \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\phi_\ell\left(\frac{t}{T}\right)u_t \\
&\quad - \left[\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\phi_\ell\left(\frac{t}{T}\right)\begin{pmatrix} 1, & y_{t-1}\end{pmatrix}\boldsymbol{D}_T^{-1}\right]\left[\boldsymbol{D}_T^{-1}\left(\sum_{t=1}^{T}\boldsymbol{x}_t\boldsymbol{x}_t'\right)\boldsymbol{D}_T^{-1}\right]^{-1}\boldsymbol{D}_T^{-1}\sum_{t=1}^{T}\boldsymbol{x}_t u_t,
\end{aligned}
$$

where

$$
\begin{aligned}
\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\phi_\ell\left(\frac{t}{T}\right)\begin{pmatrix} 1, & y_{t-1}\end{pmatrix}\boldsymbol{D}_T^{-1} &= \left(\frac{1}{T}\sum_{t=1}^{T}\phi_\ell\left(\frac{t}{T}\right), \frac{1}{\mu_T\sqrt{T}k_T^{3/2}\rho_T^T}\sum_{t=1}^{T}\phi_\ell\left(\frac{t}{T}\right)y_{t-1}\right) \\
&= \left(o\left(1\right), \frac{1}{\mu_T\sqrt{T}k_T^{3/2}\rho_T^T}\sum_{t=1}^{T}\phi_\ell\left(\frac{t}{T}\right)y_{t-1}\right).
\end{aligned}
$$

For the second element in the above vector, using Lemma A.6(a), we have

$$
\begin{aligned}
&\frac{1}{\mu_T\sqrt{T}k_T^{3/2}\rho_T^T}\sum_{t=1}^{T}\phi_\ell\left(\frac{t}{T}\right)y_{t-1} \\
&= \frac{1}{\mu_T\sqrt{T}k_T^{3/2}\rho_T^T}\sum_{t=1}^{T}\phi_\ell\left(\frac{t}{T}\right)\left[\xi_{t-1} + \mu_T\left(\rho_T^{t-1}-1\right)k_T/c\right] \\
&= \frac{1}{\mu_T\sqrt{T}k_T^{3/2}\rho_T^T}\sum_{t=1}^{T}\phi_\ell\left(\frac{t}{T}\right)\xi_{t-1} + \frac{1}{c\sqrt{T}k_T^{1/2}\rho_T^T}\sum_{t=1}^{T}\phi_\ell\left(\frac{t}{T}\right)\left(\rho_T^{t-1}-1\right) \\
&= \frac{1}{\mu_T\sqrt{T}k_T^{3/2}\rho_T^T}\sum_{t=1}^{T}\phi_\ell\left(\frac{t}{T}\right)\xi_{t-1} + o\left(1\right).
\end{aligned}
$$

By Lemmas 3.1(b&c) and A.6(a&b), we have, for $\nu > 0$,

$$
\begin{aligned}
&\frac{1}{\mu_T\sqrt{T}k_T^{3/2}\rho_T^T}\sum_{t=1}^{T}\phi_\ell\left(\frac{t}{T}\right)\xi_{t-1} = \frac{1}{\mu_T\sqrt{T}k_T^{3/2}\rho_T^T}\sum_{t=1}^{T}\phi_\ell\left(\frac{t}{T}\right)\left(\rho_T^{t-1}\xi_0 + \sum_{j=1}^{t-1}\rho_T^{t-1-j}u_j\right) \\
&= \frac{\xi_0}{\mu_T\sqrt{T}k_T^{3/2}\rho_T^T}\sum_{t=1}^{T}\phi_\ell\left(\frac{t}{T}\right)\rho_T^{t-1} + \frac{1}{\mu_T\sqrt{T}k_T^{3/2}\rho_T^T}\sum_{t=1}^{T}\phi_\ell\left(\frac{t}{T}\right)\sum_{j=1}^{t-1}\rho_T^{t-1-j}u_j \\
&= \frac{1}{\mu_T\sqrt{T}k_T^{3/2}\rho_T^T}\sum_{t=1}^{T}\phi_\ell\left(\frac{t}{T}\right)\sum_{j=1}^{t-1}\rho_T^{t-1-j}u_j + o_p\left(\frac{1}{\mu_T k_T^{1/2}\sqrt{T/k_T}}\right) \\
&= \frac{1}{\mu_T\sqrt{T}k_T^{3/2}\rho_T^T}\sum_{t=1}^{T}\phi_\ell\left(\frac{t}{T}\right)\sum_{j=1}^{T}\rho_T^{t-1-j}u_j - \frac{1}{\mu_T\sqrt{T}k_T^{3/2}\rho_T^T}\sum_{t=1}^{T}\phi_\ell\left(\frac{t}{T}\right)\sum_{j=t}^{T}\rho_T^{t-1-j}u_j + o_p\left(1\right)
\end{aligned}
$$

$$= \frac{1}{\mu_T \sqrt{T} k_T \rho_T^T} \left[\sum_{t=1}^{T} \phi_\ell \left(\frac{t}{T} \right) \rho_T^{t-1} \right] \tilde{Y}_T + o_p \left(\mu_T^{-1} k_T^{-1/2} \right) + o_p (1)$$

$$= O_p \left(\frac{1}{\mu_T k_T^{1/2} \sqrt{T/k_T}} \right) + o_p (1) = o_p (1) .$$

Therefore

$$\frac{1}{\mu_T \sqrt{T} k_T^{3/2} \rho_T^T} \sum_{t=1}^{T} \phi_\ell \left(\frac{t}{T} \right) y_{t-1} = o_p (1) ,$$

and

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{T} \phi_\ell \left(\frac{t}{T} \right) \hat{u}_t = \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \phi_\ell \left(\frac{t}{T} \right) u_t + o_p (1) = C(1) \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \phi_\ell \left(\frac{t}{T} \right) \varepsilon_t + o_p (1) .$$

Now under Assumption 3.1,

$$\frac{1}{\sqrt{T}} \sum_{j=1}^{[Tr]} u_t \Rightarrow \lambda W (r) .$$

Since $\phi_\ell (\cdot)$ is continuously differentiable, using summation by parts and the continuous mapping theorem, we have

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{T} \phi_\ell \left(\frac{t}{T} \right) \hat{u}_t \Rightarrow \lambda \eta_\ell \text{ for } \eta_\ell = \int_0^1 \phi_\ell (r) \, dW(r), \tag{A.1}$$

jointly over $\ell = 1, \ldots, K$. Since $\phi_\ell (\cdot)$ are orthonormal bases, we have $\eta_\ell \sim i.i.d.N(0, 1)$. It then follows that

$$\hat{\lambda}_K^2 / \lambda^2 \Rightarrow \frac{1}{K} \sum_{\ell=1}^{K} \eta_\ell^2 \overset{d}{=} \frac{1}{K} \chi_K^2 .$$

Part (b). Note that

$$\begin{aligned} &\mu_T k_T^{3/2} \rho_T^T (\hat{\rho}_T - \rho_T) \\ &= \left(\frac{1}{\mu_T^2 k_T^3 \rho_T^{2T}} \sum_{t=1}^{T} y_{t-1}^2 \right)^{-1} \frac{1}{\mu_T k_T^{3/2} \rho_T^T} \sum_{t=1}^{T} y_{t-1} u_t + o_p (1) \\ &\Rightarrow \frac{\tilde{X} \left(\tilde{Y}/\nu + 1/c \right)}{\left(\tilde{Y}/\nu + 1/c \right)^2 / (2c)} = \frac{2c\tilde{X}}{\tilde{Y}/\nu + 1/c} . \end{aligned}$$

It is easy to show that the above convergence holds jointly with (A.1) for $\ell = 1, \ldots, K$. Moreover, using Lemma A.6(a), we have

$$\begin{aligned} &\left| cov \left(\frac{1}{\sqrt{k_T}} \sum_{t=1}^{T} \rho_T^{-(T-t)-1} \varepsilon_t , \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \phi_\ell \left(\frac{t}{T} \right) \varepsilon_t \right) \right| \\ &= \left| \frac{\sigma^2}{\sqrt{T k_T}} \sum_{t=1}^{T} \phi_\ell \left(\frac{t}{T} \right) \rho_T^{-(T-t)-1} \right| = \left| \frac{\sigma^2}{\sqrt{T k_T} \rho_T^{T+1}} \sum_{t=1}^{T} \phi_\ell \left(\frac{t}{T} \right) \rho_T^t \right| \end{aligned}$$

$$= O\left(\frac{1}{\sqrt{T k_T}\rho_T^{T+1}} k_T \rho_T^T\right) = o\left(1\right).$$

This implies that $\tilde{X}$ is independent of $\{\eta_1, \ldots, \eta_K\}$.

Let $\eta_0 = \tilde{X}/\lambda_c$. Then, $\eta_0 \sim N(0,1)$, and η_0 is independent of $\{\eta_1, \ldots, \eta_K\}$. Now

$$\begin{aligned}\frac{\hat{\rho}_T - \rho_T}{\tilde{\sigma}_{\rho,K}} &= \frac{\mu_T k_T^{3/2} \rho_T^T \left(\hat{\rho}_T - \rho_T\right)}{\sqrt{\hat{\lambda}_K^2}\sqrt{e_2'\left[\boldsymbol{D}_T^{-1}\left(\sum_{t=1}^T \boldsymbol{x}_t \boldsymbol{x}_t'\right)\boldsymbol{D}_T^{-1}\right]^{-1} e_2}} \\ &\Rightarrow \frac{\frac{2c\tilde{X}}{\tilde{Y}/\nu + 1/c}}{\sqrt{\frac{\sum_{\ell=1}^K \eta_\ell^2}{K}\lambda^2}} \frac{\tilde{Y}/\nu + 1/c}{\sqrt{2c}} = \frac{\eta_0}{\sqrt{\frac{\sum_{\ell=1}^K \eta_\ell^2}{K}}} \overset{d}{=} t_K,\end{aligned}$$

as desired.$\square$