Mildly explosive autoregression under stationary conditional heteroskedasticity*

Stelios Arvanitis
Athens University of Economics and Business

Tassos Magdalinos
University of Southampton

18 January 2018

Abstract

A limit theory is developed for mildly explosive autoregressions under stationary (weakly or strongly dependent) conditionally heteroskedastic errors. The conditional variance process is allowed to be stationary, integrable and mixingale, thus encompassing general classes of GARCH type or stochastic volatility models. No mixing conditions nor moments of higher order than 2 are assumed for the innovation process. As in Magdalinos (2012), we find that the asymptotic behaviour of the sample moments is affected by the memory of the innovation process both in the form of the limiting distribution and, in the case of long range dependence, the rate of convergence, while conditional heteroskedasticity affects only the asymptotic variance. These effects are cancelled out in least squares regression theory and thereby the Cauchy limit theory of Phillips and Magdalinos (2007a) remains invariant to a wide class of stationary conditionally heteroskedastic innovations processes.

Keywords: Central limit theory, Explosive autoregression, Long Memory, Conditional heteroskedasticity, GARCH, mixingale, Cauchy distribution.

*Magdalinos acknowledges research support by the British Academy under grant SG151171. Arvanitis acknowledges research support by the Research Centre of AUEB, EP-2215-01/00-01.

1 Introduction

First order autoregressive processes with an explosive root, i.i.d. Gaussian innovations and zero initial condition were first analysed by White (1958), who using a moment generating function technique, derived a Cauchy limit theory for the OLS/ML estimator. Using martingale methods, Anderson (1959) arrived to the same conclusion and showed that the Cauchy limit theory is not invariant to deviations from Gaussianity and that, in general, the limit distribution of the OLS/ML estimator depends on the distribution of the (i.i.d.) innovations.

Invariance of the Cauchy least squares regression limit theory to the distribution of the innovations can be recovered when the explosive root approaches unity as the sample size n tends to infinity at sufficiently slow rate. Phillips and Magdalinos (2007a, hereafter PM_a) considered mildly explosive processes of the form

$$X_t = \rho_n X_{t-1} + u_t, \quad \rho_n = 1 + \frac{c}{n^\alpha}, \ \alpha \in (0,1), \ c > 0. \tag{1}$$

When the innovation process $(u_t)_{t\in\mathbb{Z}}$ is i.i.d. and square integrable, PM_a establish central limit theorems for sample moments generated by mildly explosive processes and obtain the following least squares regression theory:

$$\frac{1}{2c} n^\alpha \rho_n^n (\hat{\rho}_n - \rho_n) \Rightarrow \mathcal{C} \quad \text{as } n \to \infty. \tag{2}$$

This Cauchy limit theory is invariant to both the distribution of the innovations and to the initialization of the mildly explosive process.

The results of PM_a were generalised by Phillips and Magdalinos (2007b, hereafter PM_b) to include a class of weakly dependent innovations. Aue and Horvath (2007) relaxed the moment conditions on the innovations by considering an i.i.d. innovation sequence that belongs to the domain of attraction of a stable law. The limiting distribution is represented by a ratio of two independent and identically distributed stable random variables and reduces to a Cauchy distribution when the innovations have finite variance. Multivariate extensions are included in Magdalinos and Phillips (2008).

Magdalinos (2012, hereafter M_a) considered mildly explosive autoregressions generated by a correlated innovation sequence that may exhibit long range dependence. The asymptotic behaviour of the sample moments that appear in the ratio of the centred least squares estimator $\hat{\rho}_n - \rho_n$ was found to be affected by long range dependence both in the rate of convergence and in the form of the limiting distribution, crucially, in the same way and by the same amount for both components of the ratio. Hence, there is an asymptotic cancellation and, unlike its constituent components, the ratio $\hat{\rho}_n - \rho_n$ is not affected by the memory of the innovation sequence and continues to be asymptotically Cauchy with the rate of convergence of (2). The limit theory of PM_a was thus generalised and found invariant to the dependence structure of the innovation sequence even in the long memory case.

Phillips, Wu and Yu (2009) and Phillips and Yu (2011) employ the limit theory of PM_a to construct inferential procedures for the detection and dating of financial bubbles. Since the empirical stylized facts of financial asset returns are consistent with conditional heteroskedasticity, see for example Ghysels et al. (1996), it is natural to ask whether these Cauchy based confidence intervals remain valid in the presence of time varying conditional second moments. Lee (2017) and Oh, Lee and Chan (2017) confirm the Cauchy limit theory of PM_a for conditionally heteroskedastic innovations under restrictive assumptions on the innovation sequence u_t that include strong mixing with exponentially decaying coefficients, the existence of fourth moments and, in the case of Lee (2017), restrictions on the distribution of the conditional variance in some neighborhood of the origin. In it well known that finite four moments impose severe restrictions on the parameter space of GARCH type models. Also, the fact that the Cauchy limit theory (2) is directly generalisable to long memory innovations that violate the strong mixing condition, suggests that strong mixing may not be an appropriate medium of testing the invariance of the Cauchy limit theory (2) to the dependence and distributional properties of u_t.

In this paper we extend the homoskedastic framework of M_a and that of Lee (2017) and Oh et.al. (2017) by allowing the sequence u_t in (1) to be a stationary (possibly long memory) linear process with to be constructed upon a stationary square integrable conditionally heteroskedastic process where the conditional variance is a mixingale. The innovation sequence u_t is not assumed to be strong mixing nor to have finite moments of higher order than 2. We provide detailed examples of general classes of conditionally heteroskedastic models that satisfy our framework, including stationary $\mathrm{ARCH}(\infty)$ processes, asymmetric GARCH type models and log-linear stochastic volatility models in Examples 1-3 in the next section. Our asymptotic development is based on the establishment of a new law of large numbers for weakly dependent heterogeneous triangular arrays (Lemma 1) below, which constitutes a partial generalisation of the L_1-mixingale law of large numbers in Andrews (1988). Employing this law of large numbers and following the martingale approximation approach of M_a, we establish the invariance of the Cauchy limit theory (2) under this extended dependence and conditional heteroskedasticity framework. In doing so, we confirm the robustness of the Phillips, Wu and Yu (2009) and Phillips and Yu (2011) procedures in environments consistent with the empirical properties of financial asset returns.

2 Main Results

Consider the mildly explosive process in (1) with innovations $(u_t)_{t\in\mathbb{N}}$ that take the form of a covariance stationary linear process with possible long memory, as in M_a (see Assumption LP below). We propose a framework for the introduction of conditional heteroskedasticity to the innovations of (1) that: (i) maintains the potential for strong dependence in the innovation sequence (by avoiding to impose mixing conditions on $(u_t)_{t\in\mathbb{N}}$); (ii) does not require the existence of moments of higher order than 2 for $(u_t)_{t\in\mathbb{N}}$, thus giving rise to GARCH-type models with sufficiently general parameter spaces. We impose this framework on the primitive innovations $(\varepsilon_t)_{t\in\mathbb{Z}}$ of the linear process u_t in (1) by Assumption CH bellow. We denote conditional expectation by $\mathbb{E}_{\mathcal{F}}(\cdot)$ and the L_p norm by $\|\cdot\|_p$.

Assumption CH. *The process $(\varepsilon_t)_{t\in\mathbb{Z}}$ satisfies $\varepsilon_t = z_t\sqrt{h_t}$ a.s. for all $t\in\mathbb{Z}$, where the sequence $(z_t)_{t\in\mathbb{Z}}$ is i.i.d. with $\mathbb{E}(z_1)=0$ and $\mathbb{E}(z_1^2)=1$. Given a filtration $(\mathcal{F}_t)_{t\in\mathbb{Z}}$, h_t is $\mathcal{F}_{t-1}$-adapted and z_t is independent of $\mathcal{F}_{t-1}$ for all $t\in\mathbb{Z}$. The process $(z_t,h_t)_{t\in\mathbb{Z}}$ is strictly stationary with $h_t>0$ a.s. and $\sigma^2 := \mathbb{E}(h_1)\in(0,\infty)$. Finally, there exist real positive sequences $(\zeta_t)_{t\in\mathbb{Z}}$ and $(\psi_m)_{m\in\mathbb{N}}$ satisfying $\sup_{t\in\mathbb{Z}}\zeta_t<\infty$, $\psi_m\to 0$ as $m\to\infty$, and*

$$\left\|\mathbb{E}_{\mathcal{F}_{t-1-m}}\left(h_t-\sigma^2\right)\right\|_1 \le \zeta_t\psi_m \quad \textit{for any} \quad t,m\ge 0. \tag{3}$$

Under Assumption CH, $(\varepsilon_t)_{t\in\mathbb{Z}}$ is conditionally heteroskedastic w.r.t. $(\mathcal{F}_t)_{t\in\mathbb{Z}}$. Typically, $\mathcal{F}_t$ represents the informational content of the history, up to time t, of the i.i.d. process $(z_t)_{t\in\mathbb{Z}}$ or any other process upon which h_t is formed. Furthermore, $(\varepsilon_t)_{t\in\mathbb{Z}}$ is a stationary white noise with variance equal to σ^2, and additionally a martingale difference when z_t is $\mathcal{F}_t$-adapted. The conditional variance process $(h_t)_{t\in\mathbb{Z}}$ is strictly stationary and integrable. The adaptation property of h_t to $\mathcal{F}_{t-1}$ combined with (3) implies that its demeaned version $h_t - \sigma^2$ is an L_1 mixingale (see for example Andrews (1988)).

Assumption CH encompasses several classes of conditionally heteroskedastic processes. For example, for (strong) GARCH-type models we have that $\mathcal{F}_t := \sigma(z_{t-i}, i \geq 0)$. Then, stationarity typically follows by a representation of h_t as a measurable function of $\mathcal{F}_{t-1}$. Positivity and uniform integrability are usually ensured via properties of the aforementioned representation possibly combined with parameter restrictions. The mixingale property is readily verifiable for a large class of frequently used conditionally heteroskedastic models, such as ARCH(∞) processes which include finite order covariance stationary GARCH processes; see Example 1 below. In more complicated cases, such as Examples 2 and 3 below, the mixingale property can be established by stricter integrability conditions on the primitive innovations z_t of h_t (only the first moment of h_t is assumed to exist) along with strong mixing properties due to the relation between the mixingale and the strong mixing properties implied by relevant mixing inequalities, see for example McLeish (1975). Notice, however, that these sufficient conditions are not necessary: the conditional variance process h_t of the ARCH(∞) model of Example 1 satisfies the mixingale property without higher order moments nor mixing conditions. Note also that the strong mixing property of h_t does not impose weak dependence on the innovation sequence u_t in (1): the latter may have long memory (see Assumption LP(ii) below). In what follows, we provide details on certain general classes of models that satisfy Assumption CH.

Example 1 (ARCH(∞) process). For $\omega > 0$, some non negative real sequence $(\alpha_i)_{i\in\mathbb{N}}$ and $\mathcal{F}_t := \sigma(z_{t-i}, i \geq 0)$ consider the infinite order recursion defining the ARCH(∞) model:

$$h_t = \omega + \sum_{i=1}^{\infty} \alpha_i z_{t-i}^2 h_{t-i} = \omega + \sum_{i=1}^{\infty} \alpha_i \varepsilon_{t-i}^2. \tag{4}$$

A sufficient condition for the existence of a unique stationary causal solution to the above is that $\sum_{i=1}^{\infty} \alpha_i < 1$ in which case the latter admits a Volterra expansion and $\sigma^2 = \omega \left(1 - \sum_{i=1}^{\infty} \alpha_i\right)^{-1}$ (see Theorem 2.1 of Giraitis et al. (2000)). Furthermore, by Theorem 4.1 of Giraitis et al. (2000), for $(\tilde{\alpha}_i)_{\,i\in\mathbb{N}}$ defined by $\tilde{\alpha}(z) = \sum_{i=0}^{\infty} \tilde{\alpha}_i z^i = 1/\alpha(z)$ with $\alpha(z) = \sum_{i=0}^{\infty} \alpha_i z^i$ and $|z| \leq 1$, we have that

$$\varepsilon_t^2 - \sigma^2 = \sum_{i=0}^{\infty} \tilde{\alpha}_i v_{t-i},$$

with $\left((v_t)_{t\in\mathbb{Z}}, (\mathcal{G}_t)_{t\in\mathbb{Z}}\right)$ a stationary martingale difference defined by $v_t := (z_t^2 - 1)\, h_t$, for all $t \in \mathbb{Z}$. Using the above, we have that for any t and any $m > 0$,

$$\left\|\mathbb{E}_{\mathcal{F}_{t-1-m}}\left(h_t - \sigma^2\right)\right\|_1 = \left\|\mathbb{E}_{\mathcal{F}_{t-1-m}}\left(\varepsilon_t^2 - \sigma^2\right)\right\|_1 \leq 2\sigma^2 \sum_{i=m}^{\infty} |\tilde{\alpha}_i|\,,$$

hence (3) holds with $\zeta_t = 2\sigma^2$ and $\psi_m = \sum_{i=m}^{\infty} |\tilde{\alpha}_i|$, since $\sum_{i=0}^{\infty} |\tilde{\alpha}_i| < \infty$. Hence, the previous assertions hold also for any GARCH(p,q) model under the $\sum_{i=1}^{\infty} \alpha_i < 1$ restriction when applied to its ARCH(∞) representation.

Example 2 (Asymmetric GARCH type Models). Similarly to Carrasco and Chen (2002) consider the stochastic recursion

$$\Lambda\left(h_t\right) = c\left(m_t\right)\Lambda\left(h_{t-1}\right) + g\left(m_t\right),$$

with Λ increasing and continuous on $\mathbb{R}_+$, m_t a measurable function of z_t, and c, g polynomials. This formulation encompasses several GARCH(1,1)-type models as for example the (1,1) versions of the LGARCH, VGARCH, EGARCH, MGARCH, GJR, and TGARCH models (for definitions and references see Carrasco and Chen (2002)). Their properties are, among others and in some varying extend, in accordance with the empirical stylized fact of dynamic asymmetry in financial time series (for the so-called asymmetric leverage effects, see for example Bollerslev et al. (2011)). Suppose now that the distribution of m_t is absolutely continuous w.r.t. the Lebesgue measure on $\mathbb{R}$ with support that contains zero, and that for some $s \geq 1$, $|c\,(0)| < 1, \mathbb{E}\left(c^s\left(m_1\right)\right) < 1$, and $\mathbb{E}\left(g^s\left(m_1\right)\right) < 1$. Then, by Proposition 5 of Carrasco and Chen (2002) the adaptation and stationarity parts of Assumption CH follow, and furthermore the conditional variance process h_t is strongly mixing with exponentially decaying mixing coefficients. If furthermore $\mathbb{E}\left(\left(\Lambda^{-1}\left(c\left(m_t\right)\Lambda\left(h_{t-1}\right) + g\left(m_t\right)\right)\right)^s\right) < +\infty$, for some $s > 1$, then $\sigma^2 = \mathbb{E}\left(h_1\right)$ exists and by the mixing inequality of Lemma 2 of McLeish (1975) (3) holds.

In a similar manner consider the Power GARCH(p,q) recursion for $\omega > 0$,

$\alpha_i \geq 0, i = 1, \ldots, p$, $\beta_i \geq 0, i = 1, \ldots, q$, $\delta > 0$,

$$h_t^\delta = \omega + \sum_{i=1}^{p} \alpha_i h_{t-1}^\delta z_{t-i}^{2\delta} + \sum_{i=1}^{p} \beta_i h_{t-1}^\delta,$$

which is essentially a Box-Cox transformation for the conditional variance. If the distribution of z_t is absolutely continuous w.r.t. the Lebesgue measure on $\mathbb{R}$ with support equal to the real line, for some $s > \frac{1}{\delta}$, $\mathbb{E}\left(z_1^{2s\delta}\right) < \infty$, and $\mathbb{E}\left(\lambda_{\max}^s (z_1)\right) < 1$ where $\lambda_{\max}$ is the largest root of $\sum_{i=1}^{\max(p,q)} \left(\alpha_i z_1^{2\delta} + \beta_i\right) / \lambda^i$ (where a parameter is considered equal to zero if its index exceeds the relevant order), then all the assertions of Assumption CH follow by Proposition 13 of Carrasco and Chen (2002) and Lemma 2 of McLeish (1975).

Example 3 (Log-Linear Stochastic Volatility). Suppose now that $(\eta_t)_{t\in\mathbb{Z}}$ is another i.i.d. sequence such that $(\eta_t, z_t)_{t\in\mathbb{Z}}$ is stationary, $\mathcal{F}_t := \sigma\left(u_{t-i}, z_{t-i}, i \geq 0\right)$, z_t is independent of $\mathcal{F}_{t-1}$, and $\omega \in \mathbb{R}$ while $(\beta_i)_{i\in\mathbb{N}}$ is a real sequence. Consider the process

$$\ln h_t = \omega + \sum_{i=0}^{\infty} \beta_i \eta_{t-i-1}. \tag{5}$$

The conditional variance is defined as an (exogenous) log-linear process w.r.t. $(\eta_t)_{t\in\mathbb{Z}}$, and thereby the previous specify a stochastic volatility model (see for example Straumann (2004)). The possibility of contemporaneous dependence between η_t and z_t is also related to the empirical dynamic asymmetry in financial data (see above). It is easy to see that if the distribution of η_1 has a well-defined moment generating function, say M_η, on the range of $(\beta_i)_{i\in\mathbb{N}}$ and $\sum_{i=0}^{\infty} \ln M_\eta(\beta_i)$ converges, then all the assertions of Assumption CH except for (3) hold with $\sigma^2 = \exp\left(\omega + \sum_{i=0}^{\infty} \ln M_\eta(\beta_i)\right)$. For example when $\eta_1 \sim N(0,1)$ then square summability for the $(\beta_i)_{i\in\mathbb{N}}$ suffices for the above and $\sigma^2 = \exp\left(\omega + \frac{1}{2}\sum_{i=0}^{\infty} \beta_i^2\right)$. For (3) notice that any set of conditions for strong mixing of linear processes like (5), see for example Theorem 14.9 of Davidson (1994), along with the convergence of $\sum_{i=0}^{\infty} \ln M_\eta((1+\varepsilon)\beta_i)$ for some $\varepsilon > 0$ would suffice due to Theorem 14.1 of Davidson (1994), and Lemma 2 of McLeish (1975). In the standard normal case those are reduced to the absolute summability of $(\beta_i)_{i\in\mathbb{N}}$ due to Theorem 13.3.3. in Ibragimov and Linnik (1971).

The mixingale property of Assumption CH facilitates the validity of a law of large numbers for weakly dependent heterogeneous triangular arrays, Lemma 1(ii) below, which constitutes a partial generalisation of the L_1-mixingale law of large numbers in Andrews (1988). This is a key result for the asymptotic development of the paper, as it characterises the asymptotic behaviour of the conditional variance of martingale transforms that arise in mildly explosive least squares theory and allows the application of a central limit theorem to these martingale transforms to establish their asymptotic normality (see Lemma 2). Part (i) of the lemma is an auxiliary result of sums of martingale differences weighted by triangular arrays of constants, leading to the main result of part (ii), where the martingale difference is generalised to a L_1 mixingale processes without the imposition of rates to the original mixingale numbers. We denote by $(\mathcal{G}_t)_{t\in\mathbb{Z}}$ a generic filtration that need not coincide with that of Assumption CH.

Lemma 1. *For an integer valued sequence $(k_n)_{n\in\mathbb{N}}$ with $k_n \to \infty$, consider an array of real numbers $\{\alpha_{n,t} : 1 \le t \le k_n\}$ satisfying*

$$\sup_{n\in\mathbb{N}} \sum_{t=1}^{k_n} |a_{n,t}| < \infty \quad \text{and} \quad \sum_{t=1}^{k_n} a_{n,t}^2 \to 0 \text{ as } n \to \infty. \tag{6}$$

(i) *If $\left((\epsilon_t)_{t\in\mathbb{Z}}, (\mathcal{G}_t)_{t\in\mathbb{Z}}\right)$ is a uniformly integrable martingale difference process, then $\left\|\sum_{t=1}^{k_n} a_{n,t}\epsilon_t\right\|_1 \to 0$.*

(ii) *Let $\left((y_t)_{t\in\mathbb{Z}}, (\mathcal{G}_t)_{t\in\mathbb{Z}}\right)$ be a uniformly integrable adapted process with zero-mean satisfying*

$$\left\|\mathbb{E}_{\mathcal{G}_{t-m}}(y_t)\right\|_1 \le \zeta_t \psi_m \text{ for each } t, m \ge 0 \tag{7}$$

for real positive sequences $(\zeta_t)_{t\in\mathbb{Z}}$ and $(\psi_m)_{m\in\mathbb{N}}$ with $\sup_{t\in\mathbb{Z}} \zeta_t < \infty$ and $\psi_m \to 0$ as $m \to \infty$. Then $\left\|\sum_{t=1}^{k_n} a_{n,t} y_t\right\|_1 \to 0$.

The adaptation property along with (7) imply that $(y_t)_{t\in\mathbb{Z}}$ is an L_1 mixingale (see for example Andrews (1988)).

Having introduced a convenient conditional heteroskedasticity framework, we proceed to defining the linear relationship between the innovations of the mildly explosive autoregression (1) with the process $(\varepsilon_t)_{t\in\mathbb{Z}}$ of Assumption CH.

Assumption LP. *For each $t \in \mathbb{N}$, u_t has Wold representation $u_t = \sum_{j=0}^{\infty} c_j \varepsilon_{t-j}$ where $(\varepsilon_t)_{t\in\mathbb{Z}}$ is as in Assumption CH and $(c_j)_{j\geq 0}$ is a sequence of constants satisfying $c_0 = 1$ and one of the following conditions:*

(i) $\sum_{j=0}^{\infty} |c_j| < \infty$ *and* $\sum_{j=0}^{\infty} c_j \neq 0$.

(ii) *For each $j \in N$, $c_j = L(j) j^{-\kappa}$, for some $\kappa \in (1/2, 1)$, where L is a slowly varying function at infinity such that $\varphi(t) := L(t) t^{-\kappa}$ is eventually non-increasing and $\sup_{t\in[0,B]} t^{\delta} L(t) < \infty$ for any $\delta, B > 0$.*

(iii) *$c_j = \theta j^{-1}$, $j \in N$, for some $\theta \neq 0$.*

Assumption IC. *X_0 can be any fixed constant or a random process $X_0(n)$ satisfying $X_0(n) = o_p\left(n^{\alpha/2}\right)$ under LP(i), $X_0(n) = o_p\left(n^{(3/2-\kappa)\alpha} L(n^{\alpha})\right)$ under LP(ii) and $X_0(n) = o_p\left(n^{\alpha/2} \log n\right)$ under LP(iii).*

Assumptions LP and IC are identical to the assumption framework in M_a. Under the first, $(u_t)_{t\in\mathbb{N}}$ is a covariance and strictly stationary linear process, since $(c_j)_{j\geq 0}$ is square summable and $(\varepsilon_t)_{t\in\mathbb{Z}}$ is a stationary conditionally heteroskedastic white noise.

LP(i) ensures absolute summability of the autocovariance function of $(u_t)_{t\in\mathbb{Z}}$ thereby giving rise to a weakly dependent stationary process. LP(ii) induces long memory to $(u_t)_{t\in\mathbb{Z}}$. Recall that a function L is slowly varying at infinity if and only if $L(ut)/L(t) \to 1$ for any $u > 0$; see Bingham, Goldie and Teugels (1987), abbreviated hereafter as BGT. The parametrisation $c_j = L(j) j^{-\kappa}$ is standard for stationary linear processes that exhibit long memory, see e.g. Giraitis, Koul and Surgailis (1996) and Wu and Min (2005), including stationary AFRIMA processes, for $\kappa = 1 - d$, $d \in (0, 1/2)$ in the relevant notation. The boundary $\kappa = 1$ between weak and strong dependence in the memory of the innovation sequence is investigated via the harmonic coefficients of Assumption LP(iii).

The property of $\varphi(t)$ being eventually non-increasing facilitates the computation of asymptotic variances by means of Euler summation in M_a. This property is for instance satisfied by the Zygmund class of differentiable slowly varying functions (see BGT, Theorem 1.5.5). Boundedness of $t^{\delta} L(t)$ in a neighbourhood of the origin is a standard requirement for the validity of Abelian theorems for integrals involving regularly varying functions, see BGT, Proposition 4.1.2(a). Both conditions hold trivially for the stationary AFRIMA processes with fractional parameter as above, see Samorodnitsky (2006).

Under Assumptions LP-IC, and conditional homoskedasticity for $(\varepsilon_t)_{t\in\mathbb{Z}}$, M_a establishes the invariance of the Cauchy regression theory of PM_a. This invariance holds despite the different rates of convergence and limit distributions (that arise as a result of the memory properties of u_t) satisfied by the sample moments that enter least squares regression, as those are asymptotically cancelled out between the components of the OLS estimator. The question here is whether an analogous result holds under conditional heteroskedasticity. Once the appropriate asymptotic framework has been set by Assumption CH and the mixingale law of large numbers of Lemma 1, it turns out that not only the above invariance remains true, but also that every intermediate result of M_a continues to hold.

Specifically, as in the previous analyses of (mildly) explosive autoregression by Anderson (1959), PM_a and M_a, the limit theory of the OLSE for ρ_n depends on properties of the stochastic sequences

$$Y_n(\kappa) = \frac{1}{n^{\left(\frac{3}{2}-\kappa\right)\alpha}} \sum_{t=1}^{\tau_n(\beta)} \rho_n^{-t} u_{n+1-t} \quad \text{and} \quad Z_n(\kappa) = \frac{1}{n^{\left(\frac{3}{2}-\kappa\right)\alpha}} \sum_{t=1}^{\tau_n(\beta)} \rho_n^{-t} u_t \tag{8}$$

with $\rho_n = 1 + c/n^\alpha$ as defined in (1) and

$$\tau_n(\beta) = \left\lfloor \frac{n^\beta}{2} \right\rfloor \quad \text{for some } \beta \in \left(\alpha, \min\left\{\frac{3\alpha}{2}, 1\right\}\right). \tag{9}$$

For notational convenience, following M_a, we employ the notation $Y_n(1)$ and $Z_n(1)$ for the sequences in (8) under both Assumptions LP(i) and LP(iii). This is consistent with the $n^{\alpha/2}$ normalisation that applies under weak dependence.

By covariance stationarity of $(u_t)_{t\in\mathbb{Z}}$, $Y_n(\kappa)$ and $Z_n(\kappa)$ have equal variance; their asymptotic variance is computed in Lemma 1 of M_a for any white noise process $(\varepsilon_t)_{t\in\mathbb{Z}}$ with variance equal to σ^2 and is given by

$$V_\kappa := \sigma^2 c^{2\kappa-3} \frac{\Gamma(1-\kappa)^2}{2\cos\{\pi(1-\kappa)\}}, \quad \kappa \in (1/2, 1) \tag{10}$$

where $\Gamma(x) = \int_0^\infty u^{x-1} e^{-u} du$ is the gamma function. Under Assumption CH, the above expression for V_κ continues to apply with $\sigma^2 = \mathbb{E}(h_1)$. Lemma 2 below provides a limit theory $Y_n(\kappa)$ and $Z_n(\kappa)$ under conditional heteroskedasticity and both short and long memory in the innovations. Both the rates of convergence and the limit distributions depend crucially on the linear dependence properties of $(u_t)_{t\in\mathbb{Z}}$ via the memory parameter κ. Under CH the conditional variance process $(h_t)_{t\in\mathbb{Z}}$ does not affect the convergence rates, but affects the limit distributions via $\sigma^2 = \mathbb{E}(h_1)$. For this value of σ^2, denote $\omega^2 := \sigma^2 \left(\sum_{j=0}^{\infty} c_j\right)^2$.

Lemma 2. *Under Assumptions CH and LP, the sequences $Z_n(\kappa)$ and $Y_n(\kappa)$ in (8) have the following joint asymptotic behaviour as $n \to \infty$:*

(i) *Under LP(i), $[Y_n(1), Z_n(1)] \Rightarrow [Y_1, Z_1]$, where Y_1 and Z_1 are independent $N(0, \omega^2/2c)$ random variables.*

(ii) *Under LP(ii), $L(n^\alpha)^{-1}[Y_n(\kappa), Z_n(\kappa)] \Rightarrow [Y_\kappa, Z_\kappa]$, where where Y_κ and Z_κ are independent $N(0, V_\kappa)$ random variables and V_κ is given by (10).*

(iii) *Under LP(iii), $(\log n^\alpha)^{-1}[Y_n(1), Z_n(1)] \Rightarrow [Y_1', Z_1']$, where Y_1' and Z_1' are independent $N\left(0, \sigma^2\theta^2/2c\right)$ random variables.*

Lemma 2 shows that Lemmata 2-4 of M_a continue to hold under conditional heteroskedasticity with σ^2 arising as the expectation of the conditional variance process h_t of Assumption CH (instead of the conditional homoskedasticity assumption $h_t = \sigma^2$ for all t *a.s.* maintained in M_a). The key insight is the application of the mixingale law of large numbers of Lemma 1(ii) which ensures the validity of a standard martingale central limit theorem in each of the cases (i)-(iii) above. The joint asymptotic behaviour of $Y_n(\kappa)$ and $Z_n(\kappa)$ completely determines the limit theory of the sample moments of X_t and of the normalised and centred OLS estimator in the context of (1), as long as the standard approximation argument of Anderson (1959) continues to apply under Assumptions CH, LP, IC. The validity of this approximation argument is established in the following lemma.

Lemma 3. *Let L denote an arbitrary slowly varying function at infinity. Then, under Assumptions CH, LP and IC,*

$$\frac{\rho_n^{-2n}}{n^\alpha n^{(3-2\kappa)\alpha} L(n^\alpha)^2} \sum_{t=1}^{n} X_{t-1}^2 = \frac{1}{2c}\left[\frac{1}{L(n^\alpha)} Z_n(\kappa)\right]^2 + o_p(1)$$

$$\frac{\rho_n^{-n}}{n^{(3-2\kappa)\alpha} L(n^\alpha)^2} \sum_{t=1}^{n} X_{t-1} u_t = \frac{Y_n(\kappa)}{L(n^\alpha)} \frac{Z_n(\kappa)}{L(n^\alpha)} + o_p(1)$$

as $n \to \infty$ where: (a) Under LP(i), $\kappa = 1$ and $L(x) = 1$. (b) Under LP(ii), $\kappa \in (1/2, 1)$ and L satisfies LP(ii). (c) Under LP(iii), $\kappa = 1$ and $L(x) = \log x$.

Lemma 3 simply asserts that Lemma 5 of M_a continues to apply when the conditional homoskedasticity assumption of M_a is replaced by Assumption CH and confirms the validity of the standard approximation argument pertaining to (mildly) explosive sample moments. Combing Lemmata 2 and 3, we deduce that, under appropriate normalisation, joint convergence in distribution of $\left(\sum_{t=1}^{n} X_{t-1}u_t, \sum_{t=1}^{n} X_{t-1}^2\right)$ applies in all cases LP(i)-LP(iii). Moreover, the same normalisation applies to the centred OLS estimator $\hat{\rho}_n - \rho_n$ irrespective of the dependence properties of u_t. The resulting Cauchy limit distribution for the normalised and centred OLS estimator is a simple corollary of the continuous mapping theorem and the fact that the limiting random vectors (Y_1, Z_1), (Y_κ, Z_κ) and (Y_1', Z_1') of Lemma 2 consist of independent components.

Theorem 1. *For the mildly explosive process generated by (1) under Assumptions CH, LP and IC, the following limit theory applies as $n \to \infty$:*

$$\frac{1}{2c} n^{\alpha} \rho_n^n \left(\hat{\rho}_n - \rho_n\right) \Rightarrow \mathcal{C} \quad \text{as } n \to \infty,$$

where $\mathcal{C}$ denotes a standard Cauchy random variable.

Remarks.

1. Theorem 1 shows that standard Cauchy mildly explosive regression theory continues to hold under stationarity, weak or strong linear dependence, mixingale conditional variance and second order integrability for the innovation process. Even in this general framework, the limit theory depends only on the parameters c and α that determine the degree of mild explosion, i.e. the neighbourhood of unity that contains the mildly explosive root ρ_n. As remarked in M_a and also holds true in the current conditional heteroskedasticity context, this invariance of least squares limit theory to the memory properties of the innovation sequence is due to the strength of the (mildly) explosive regression signal. Exponential signal strength gives rise to a fundamental property of explosive and mildly explosive autoregression, established in our context by Lemma 3, that the asymptotic behaviour of the normalised and

centred least squares estimator is completely characterised by the ratio $Y_n(\kappa)/Z_n(\kappa)$ in which the numerator and the denominator have identical rates and limiting distributions (by Lemma 2). Hence any idiosyncratic characteristic of the limit theory of the individual components $Y_n(\kappa)$ and $Z_n(\kappa)$ is essentially canceled out in the ratio. Apart from strict stationarity of u_t (which is inherent in GARCH-type processes), Theorem 1 constitutes a generalisation of the corresponding theorem of M_a.

2. Mildly explosive autoregression with conditionally heteroskedastic innovations has recently been investigated by Lee (2017) and Oh et al. (2017) under a more restrictive framework. In particular, the innovation sequence $(u_t)_{t\in\mathbb{Z}}$ is assumed to be strong mixing with exponentially decaying coefficients and finite fourth moments (equivalently finite second moments for the conditional variance process h_t). It is well known that higher order moment assumptions severely restrict the parameter space of GARCH type models. We avoid this problem since Assumption CH does not require the existence of second moments (or inverse second moments) for the conditional variance process. Also, Assumption CH does not require the innovation sequence $(u_t)_{t\in\mathbb{Z}}$ to be strong mixing. Example 1 shows that Assumption CH is satisfied by ARCH(∞) processes (and hence all GARCH(p,q) processes) satisfying the standard stability condition, irrespective of whether h_t is strong mixing. Even when the mixingale condition (3) is verified by the strong mixing property of h_t (Examples 2 and 3), $(u_t)_{t\in\mathbb{Z}}$ will not be strong mixing under the strongly dependent correlation schemes of Assumption LP(ii) and LP(iii). To our knowledge, Assumption CH provides the most general framework of conditional heteroskedasticity in the literature of (mildly) explosive autoregressions. Further generalisation may be possible, in the direction of non-integrability of the conditional variance process, with the truncated first moment of h_1 being slowly varying at infinity, see Goldie (1991); this slow variation is likely to appear in the rates in Lemma 2, yet the Cauchy limit theory of Theorem 1 should remain unaffected. We leave such considerations for further research.

3. Theorem 1 provides a limit distribution that can be used for interval estimation. Phillips, Wu and Yu (2009) and Phillips and Yu (2011) apply the construction of Cauchy confidence intervals for the detection

of financial bubbles. Given the fact that financial asset returns in relevant frequencies exhibit stylized facts consistent with several patterns of conditional heteroskedasticity, see for example Ghysels et al. (1996), Theorem 1 above ensures the robustness of those procedures to general form of conditional heteroskedasticity in the innovations.

3 Proofs

This section contains the proofs of mathematical statements in the paper. We employ a similar approach to M_a and Magdalinos (2009, hereafter M_b) with the important addition of the mixingale law of large numbers of Lemma 1 permits the use of the martingale CLT in Corollary 3.1 of Hall and Heyde (1980) in the present framework of conditional heteroskedasticity.

3.1 Proof of Lemma 1.

Define $\hat{\epsilon}_{n,t} := \epsilon_t \mathbf{1}\{|\epsilon_t| \leq \Delta_n\}$ and $\tilde{\epsilon}_{n,t} := \epsilon_t \mathbf{1}\{|\epsilon_t| > \Delta_n\}$ for a sequence $(\Delta_n)_{n\in\mathbb{N}}$ satisfying

$$\Delta_n \to \infty \text{ and } \Delta_n^2 \sum_{t=1}^{k_n} a_{n,t}^2 \to 0 \text{ as } n \to \infty. \tag{11}$$

The martingale difference property implies that $\epsilon_t = \hat{\epsilon}_{n,t} - \mathbb{E}_{\mathcal{G}_{t-1}}(\hat{\epsilon}_{n,t}) + \tilde{\epsilon}_{n,t} - \mathbb{E}_{\mathcal{G}_{t-1}}(\tilde{\epsilon}_{n,t})$ so

$$\begin{aligned}\left\|\textstyle\sum_{t=1}^{k_n} a_{n,t} y_t\right\|_1 &\leq \left\|\textstyle\sum_{t=1}^{k_n} a_{n,t}\left(\hat{\epsilon}_{n,t} - \mathbb{E}_{\mathcal{G}_{t-1}}(\hat{\epsilon}_{n,t})\right)\right\|_1 + \left\|\textstyle\sum_{t=1}^{k_n} a_{n,t}\left(\tilde{\epsilon}_{n,t} - \mathbb{E}_{\mathcal{G}_{t-1}}(\tilde{\epsilon}_{n,t})\right)\right\|_1 \\ &\leq \left\|\textstyle\sum_{t=1}^{k_n} a_{n,t}\left(\hat{\epsilon}_{n,t} - \mathbb{E}_{\mathcal{G}_{t-1}}(\hat{\epsilon}_{n,t})\right)\right\|_2 + 2\left(\sup_{n\in\mathbb{N}} \textstyle\sum_{t=1}^{k_n} |a_{n,t}|\right) \sup_{t\in\mathbb{N}} \|\tilde{\epsilon}_{n,t}\|_1\end{aligned}$$

by the Lyapounov inequality and the Jensen inequality for conditional expectations. By (11) and uniform integrability of (ϵ_t), $\sup_{t\in\mathbb{N}} \|\tilde{\epsilon}_{n,t}\|_1 \to 0$ so the second term on the right tends to 0. For the first term, orthogonality of $\left\{\hat{\epsilon}_{n,t} - \mathbb{E}_{\mathcal{G}_{t-1}}(\hat{\epsilon}_{n,t}) : t \geq 1\right\}$ yields

$$\left\|\sum_{t=1}^{k_n} a_{n,t}\left(\epsilon_{1t} - \mathbb{E}_{\mathcal{G}_{t-1}}(\epsilon_{1t})\right)\right\|_2^2 = \sum_{t=1}^{k_n} a_{n,t}^2 \mathbb{E}\left[\left(\epsilon_{1t} - \mathbb{E}_{\mathcal{G}_{t-1}}(\epsilon_{1t})\right)^2\right] \leq \Delta_n^2 \sum_{t=1}^{k_n} a_{n,t}^2 \to 0$$

by the choice of $(\Delta_n)_{n\in\mathbb{N}}$ in (11). This proves part (i).

For part (ii), the fact that $(y_t)_{t\in\mathbb{Z}}$ is $\mathcal{G}_t$-adapted implies that, for any fixed integer $M > 0$

$$\sum_{t=1}^{k_n} a_{n,t} y_t = \sum_{m=0}^{M-1} \sum_{t=1}^{k_n} a_{n,t} \left(\mathbb{E}_{\mathcal{G}_{t-m}}(y_t) - \mathbb{E}_{\mathcal{G}_{t-m-1}}(y_t) \right) + \sum_{t=1}^{k_n} a_{n,t} \mathbb{E}_{\mathcal{G}_{t-M}}(y_t) \quad (12)$$

as in equation (6) of Andrews (1988). For each m,

$$\epsilon_t^{(m)} := \mathbb{E}_{\mathcal{G}_{t-m}}(y_t) - \mathbb{E}_{\mathcal{G}_{t-m-1}}(y_t)$$

is a $\mathcal{G}_{t-m}$-martingale difference process that inherits the uniform integrability property from y_t and so $\epsilon_t^{(m)}$ satisfies the conclusion of part (i). Applying the triangle inequality and (7) to (12) we obtain

$$\begin{aligned} \left\| \sum_{t=1}^{k_n} a_{n,t} y_t \right\|_1 &\leq \sum_{m=0}^{M-1} \left\| \sum_{t=1}^{k_n} a_{n,t} \epsilon_t^{(m)} \right\|_1 + \sum_{t=1}^{k_n} |a_{n,t}| \left\| \mathbb{E}_{\mathcal{G}_{t-M}}(y_t) \right\|_1 \\ &\leq M \max_{0\leq m<M} \left\| \sum_{t=1}^{k_n} a_{n,t} \epsilon_t^{(m)} \right\|_1 + \sup_{n\in\mathbb{N}} \sum_{t=1}^{k_n} |a_{n,t}| \sup_{t\in\mathbb{Z}} \zeta_t \psi_M . (13) \end{aligned}$$

Let $\delta > 0$ be arbitrary. Since $C := \sup_{n\in\mathbb{N}} \sum_{t=1}^{k_n} |a_{n,t}| \sup_{t\in\mathbb{Z}} \zeta_t < \infty$ and $\psi_M \to 0$, there exists $M_0(\delta) \in \mathbb{N}$ such that $\psi_{M_0(\delta)} \leq \delta/(2C)$. Choosing $M = M_0(\delta)$ in (13), we obtain

$$\begin{aligned} \left\| \sum_{t=1}^{k_n} a_{n,t} y_t \right\|_1 &\leq M_0(\delta) \max_{0\leq m<M_0(\delta)} \left\| \sum_{t=1}^{k_n} a_{n,t} \epsilon_t^{(m)} \right\|_1 + \frac{\delta}{2} \\ &\leq M_0(\delta) \frac{\delta}{2M_0(\delta)} + \frac{\delta}{2} = \delta \end{aligned}$$

where the second inequality applies for all but finitely many n since $\max_{1\leq m<M_0(\delta)} \left\| \sum_{t=1}^{k_n} a_{n,t} \epsilon_t^{(m)} \right\|_1 \to 0$ by part (i).

3.2 Proof of Lemma 2.

Let us first establish some useful notation. As in M_a, using the linear process representation of u_t, $Y_n(\kappa)$ and $Z_n(\kappa)$ of (8) are factored as the sum of pairs of uncorrelated components: $Z_n(\kappa) = Z_n^{(1)}(\kappa) + Z_n^{(2)}(\kappa)$ and $Y_n(\kappa) = Y_n^{(1)}(\kappa) + Y_n^{(2)}(\kappa)$ with

$$Z_n^{(1)}(\kappa) = \frac{1}{n^{\left(\frac{3}{2}-\kappa\right)\alpha}} \sum_{t=1}^{\tau_n(\beta)} \rho_n^{-t} \sum_{j=0}^{t} c_j \varepsilon_{t-j}, \quad Z_n^{(2)}(\kappa) = \sum_{j=1}^{\infty} B_{nj}(\kappa) \varepsilon_{-j} \tag{14}$$

$$Y_n^{(1)}(\kappa) = \sum_{j=1}^{\tau_n(\beta)} C_{nj}(\kappa) \varepsilon_{n+1-j}, \quad Y_n^{(2)}(\kappa) = \sum_{k>\tau_n(\beta)} \sum_{t=1}^{\tau_n(\beta)} \frac{\rho_n^{-t} c_{k-t}}{n^{\left(\frac{3}{2}-\kappa\right)\alpha}} \varepsilon_{n+1-k} \tag{15}$$

where $B_{nj}(\kappa)$ and $C_{nj}(\kappa)$ are arrays of real numbers defined by

$$B_{nj}(\kappa) = n^{-\left(\frac{3}{2}-\kappa\right)\alpha} \sum_{t=1}^{\tau_n(\beta)} \rho_n^{-t} c_{t+j}, \quad C_{nj}(\kappa) = n^{-\left(\frac{3}{2}-\kappa\right)\alpha} \sum_{t=1}^{j} \rho_n^{-t} c_{j-t} \tag{16}$$

and $\tau_n(\beta)$ is the sequence defined in (9). Denote the array

$$A_{nj}(\kappa) = n^{-\left(\frac{3}{2}-\kappa\right)\alpha} \rho_n^{-j} \sum_{i=0}^{\tau_n(\beta)} c_i \rho_n^{-i}, \tag{17}$$

related to $Z_n^{(1)}(\kappa)$. When $\kappa = 1$, we write

$$A_{nj} = A_{nj}(1), \; B_{nj} = B_{nj}(1) \; \text{ and } C_{nj} = C_{nj}(1). \tag{18}$$

As in M_a, the asymptotic behaviour of (Z_n, Y_n) is determined by $Z_n^{(1)}$ and $Y_n^{(1)}$ under Assumptions LP(i) and LP(iii), and by $Z_n^{(1)}, Z_n^{(2)}$ and $Y_n^{(1)}$ under Assumption LP(ii). $Y_n^{(2)}$ is asymptotically negligible in all cases. Finally, we say an array (ψ_{nk}) of random vectors satisfies the Lindeberg condition if

$$\sum_{k=0}^{n} \mathbb{E}_{\mathcal{F}_{k-1}} \left(\|\psi_{nk}\|^2 \mathbf{1}\{\|\psi_{nk}\| > \delta\} \right) \to_p 0 \quad \forall \delta > 0. \tag{19}$$

Proof of Lemma 2(i). Under Assumption CH, $(\varepsilon_t)_{t\in\mathbb{Z}}$ is a white noise process with variance equal to σ^2, so Lemmata B1 and B2 of M_b continue to apply. We have that

$$\begin{aligned}
\left[Y_n(1), Z_n(1)\right]' &= \left[Y_n^{(1)}(1), Z_n^{(1)}(1)\right]' + o_p(1) \\
&= \left[\sum_{j=1}^{\tau_n(\beta)} C_{nj}\varepsilon_{n+1-j}, \sum_{j=0}^{\tau_n(\beta)} A_{nj}\varepsilon_j\right]' + o_p(1) \\
&= \sum_{j=0}^{n} \zeta_{nj} + o_p(1), \qquad (20)
\end{aligned}$$

where ζ_{nj} denotes the $\mathcal{F}_j$-martingale difference array

$$\zeta_{nj} = \left[\bar{C}_{nj}, \bar{A}_{nj}\right]' \varepsilon_j$$

with

$$\bar{A}_{nj} = A_{nj}\mathbf{1}\{j \leq \tau_n(\beta)\} \quad \bar{C}_{nj} = C_{n,n+1-j}\mathbf{1}\{j > n - \tau_n(\beta)\}$$

and A_{nj} and C_{nj} defined in (17), (16) and (18). By (9), $n - \tau_n(\beta) > \tau_n(\beta)$ so $\bar{U}_{nk}\bar{C}_{nk} = 0$ for all k. By Assumption CH the conditional variance of $\sum_{k=0}^{n} \zeta_{nk}$ is given by

$$U_n := \sum_{j=0}^{n} \mathbb{E}_{\mathcal{F}_{j-1}} \zeta_{nj}\zeta_{nj}' = \operatorname{diag}\left(\sum_{j=1}^{\tau_n(\beta)} C_{nj}^2 h_j, \sum_{j=0}^{\tau_n(\beta)} A_{nj}^2 h_j\right)$$

By Lemma B2(ii) in M_b, $U_n \to_p \frac{\omega^2}{2c} I_2$ provided that

$$\left\|\sum_{j=1}^{\tau_n(\beta)} C_{nj}^2\left(h_j - \sigma^2\right)\right\|_1 + \left\|\sum_{j=0}^{\tau_n(\beta)} A_{nj}^2\left(h_j - \sigma^2\right)\right\|_1 \to 0. \qquad (21)$$

To prove (21), we employ Lemma 1(ii) with the identifications $y_t := h_t - \sigma^2$, $k_n := \tau_n(\beta)$, $\mathcal{G}_t := \mathcal{F}_{t-1}$, and $a_{n,t} \in \{C_{nt}^2, A_{nt}^2\}$. Since h_t is $\mathcal{F}_{t-1}$-adapted, (3) of Assumption CH implies that $y_t = h_t - \sigma^2$ satisfies (7) of Lemma 1. The sequences $\sum_{t=0}^{\tau_n(\beta)} A_{nt}^2$ and $\sum_{t=0}^{\tau_n(\beta)} C_{nt}^2$ are convergent by Lemma B2(ii) in M_b so the first condition of (6) is satisfied. For the second condition of (6), since $C := \sum_{i=0}^{\infty} |c_i| < \infty$ we obtain

$$\sum_{k=1}^{\tau_n(\beta)} C_{nk}^4 = \frac{1}{n^{2\alpha}} \sum_{k=1}^{\tau_n(\beta)} \left(\sum_{t=1}^{k} \rho_n^{-t} c_{k-t}\right)^4 \leq C^4 \frac{\tau_n(\beta)}{n^{2\alpha}} = o\left(\frac{1}{n^{\alpha/2}}\right)$$

by (9) and, similarly,

$$\sum_{k=1}^{\tau_n(\beta)} U_{nk}^4 = \frac{1}{n^{2\alpha}} \sum_{k=1}^{\tau_n(\beta)} \rho_n^{-4k} \left(\sum_{j=0}^{\tau_n(\beta)} c_j \rho_n^{-j} \right)^4 \leq \frac{C^4}{n^{2\alpha}} \sum_{k=0}^{\infty} \rho_n^{-4k} = O\left(\frac{1}{n^\alpha}\right).$$

This proves (21) and the required convergence $U_n \to_p \frac{\omega^2}{2c} I_2$ for the conditional variance. By the proof of equation (11) in M_b, it is clear that the array (ζ_{nj}) in (20) satisfies the Lindeberg condition (19) provided that $(\varepsilon_t^2)_{t\in\mathbb{Z}}$ is a uniformly integrable sequence. The latter is guaranteed by Assumption CH because of strict stationarity and integrability of $(\varepsilon_t^2)_{t\in\mathbb{Z}}$. The lemma now follows by applying the martingale CLT in Corollary 3.1 of Hall and Heyde (1980) to the martingale difference array (ζ_{nj}) in (20).

Proof of Lemma 2(iii). From the proof of Lemma 4 of M_b (which only employs unconditional second moment bounds) we obtain

$$\begin{aligned} \frac{1}{\log n^\alpha} \left[Y_n(1), Z_n(1)\right] &= \frac{1}{\log n^\alpha} \left[Y_n^{(1)}(1), Z_n^{(1)}(1)\right] + o_p(1) \\ &= \frac{1}{\log n^\alpha} \sum_{j=0}^{n} \zeta_{nj} + o_p(1) \end{aligned}$$

with ζ_{nj} defined as in (20). By the argument of part (i), it is sufficient to verify (6) of Lemma 1 with the same identifications as in part (i), apart from $a_{n,t} \in \left\{(C_{nt}/\log n^\alpha)^2, (A_{nt}/\log n^\alpha)^2\right\}$. The first part of (6) follows since $(\log n^\alpha)^{-2} \sum_{j=0}^{\tau_n(\beta)} C_{nt}^2$ and $(\log n^\alpha)^{-2} \sum_{j=0}^{\tau_n(\beta)} A_{nt}^2$ both converge to $\theta^2/2c$ (equations (16) and (17) of M_b). For the second part of (6), since $\sum_{i=0}^{n} |c_i| = O(\log n)$,

$$\begin{aligned} \frac{1}{(\log n^\alpha)^4} \sum_{k=1}^{\tau_n(\beta)} C_{nk}^4 &= \frac{1}{n^{2\alpha} (\log n^\alpha)^4} \sum_{k=1}^{\tau_n(\beta)} \left(\sum_{t=1}^{k} \rho_n^{-t} c_{k-t} \right)^4 \\ &\leq \frac{\left(\sum_{i=0}^{n} |c_i|\right)^4}{(\log n^\alpha)^4} \frac{\tau_n(\beta)}{n^{2\alpha}} = o\left(\frac{1}{n^{\alpha/2}}\right) \end{aligned}$$

$$
\begin{aligned}
\frac{1}{(\log n^{\alpha})^4}\sum_{k=1}^{\tau_n(\beta)} A_{nk}^4 &= \frac{1}{n^{2\alpha}(\log n^{\alpha})^4}\sum_{k=1}^{\tau_n(\beta)} \rho_n^{-4k}\left(\sum_{j=0}^{\tau_n(\beta)} c_j\rho_n^{-j}\right)^4 \\
&\leq \frac{\left(\sum_{i=0}^{n}|c_i|\right)^4}{(\log n^{\alpha})^4}\frac{1}{n^{2\alpha}}\sum_{k=0}^{\infty}\rho_n^{-4k} = O\left(\frac{1}{n^{\alpha}}\right).
\end{aligned}
$$

Proof of Lemma 2(ii). First note that the definition of $A_{nj}(\kappa)$ in (17) differs from A_{nj} in M_a by a slowly varying factor $L(n^{\alpha})$. By Propositions 3.2.1-3.2.3 of M_a,

$$
\begin{aligned}
L(n^{\alpha})^{-1}\left[Z_n(\kappa), Y_n(\kappa)\right]' &= L(n^{\alpha})^{-1}\left[Z_n^{(1)}(\kappa), Z_n^{(2)}(\kappa), Y_n^{(1)}(\kappa)\right] + o_p(1) \\
&= \sum_{j=-\tau_n(\beta)}^{n} \xi_{nj} + o_p(1) \qquad (22)
\end{aligned}
$$

where $\xi_{nk} := \left[\tilde{A}_{nj}, \tilde{B}_{nj}, \tilde{C}_{nj}\right]'\varepsilon_j$ is a $\mathcal{F}_j$-martingale difference array in $\mathbb{R}^3$ with components given by

$$
\begin{aligned}
\tilde{A}_{nj} &= L(n^{\alpha})^{-1} A_{nj}(\kappa)\,\mathbf{1}\{0 \leq j \leq \tau_n(\beta)\}, \quad \tilde{B}_{nj} = L(n^{\alpha})^{-1} B_{n,-j}(\kappa)\,\mathbf{1}\{j < 0\}, \\
\tilde{C}_{nj} &= L(n^{\alpha})^{-1} C_{n,n+1-j}(\kappa)\,\mathbf{1}\{j > n - \tau_n(\beta)\}
\end{aligned}
$$

(see equation (23) of M_a), with $A_{nj}(\kappa)$, $B_{nj}(\kappa)$ and $C_{nj}(\kappa)$ defined in defined in (17) and (16). By (9), $n - \tau_n(\beta) > \tau_n(\beta)$ so $\tilde{A}_{nk}\tilde{B}_{nk} = \tilde{A}_{nk}\tilde{C}_{nk} = \tilde{B}_{nk}\tilde{C}_{nk} = 0$ for all k, so the conditional variance of the martingale array in (22) is given by

$$
\begin{aligned}
U_n &= \sum_{j=-\tau_n(\beta)}^{n} \mathbb{E}_{\mathcal{F}_{j-1}}\xi_{nj}\xi_{nj}' = \sum_{j=-\tau_n(\beta)}^{n} \mathrm{diag}\left[\tilde{A}_{nj}^2, \tilde{B}_{nj}^2, \tilde{C}_{nj}^2\right] h_j \\
&= \frac{1}{L(n^{\alpha})^2}\mathrm{diag}\left[\sum_{j=0}^{\tau_n(\beta)} A_{nj}^2(\kappa) h_j, \sum_{j=1}^{\tau_n(\beta)} B_{nj}^2(\kappa) h_j, \sum_{j=1}^{\tau_n(\beta)} C_{nj}^2(\kappa) h_j\right].
\end{aligned}
$$

Denoting $y_j = h_j - \sigma^2$,

$$
\begin{aligned}
\tilde{U}_n &= \frac{1}{L(n^{\alpha})^2}\mathrm{diag}\left[\sum_{j=0}^{\tau_n(\beta)} A_{nj}^2(\kappa) y_j, \sum_{j=1}^{\tau_n(\beta)} B_{nj}^2(\kappa) y_j, \sum_{k=1}^{\tau_n(\beta)} C_{nj}^2(\kappa) y_j\right] \qquad (23) \\
U &= \frac{\sigma^2 c^{2\kappa-3}\Gamma(1-\kappa)^2}{2} diag\left[1, \frac{1}{\cos\pi(1-\kappa)} - 1, \frac{1}{\cos\pi(1-\kappa)}\right],
\end{aligned}
$$

Propositions 3.2.1-3.2.3 of M_a, imply that

$$U_n = U + \tilde{U}_n + o(1).$$

Therefore, $\tilde{U}_n \to_p 0$ is sufficient for $U_n \to_p U$ as in Proposition 3.2.4 of M_a. The proof of equation (24) of M_a, shows that uniform integrability of $(\varepsilon_t^2)_{t\in\mathbb{Z}}$ (which is guaranteed by Assumption CH) is sufficient for the array (ξ_{nj}) in (22) to satisfy the Lindeberg condition (19). Thus, if $\tilde{U}_n \to_p 0$ holds, the martingale CLT (Corollary 3.1 of Hall and Heyde (1980)) applied to (22) yields

$$L(n^\alpha)^{-1}\left[Z_n^{(1)}(\kappa), Z_n^{(2)}(\kappa), Y_n^{(1)}(\kappa)\right] \Rightarrow \left[Z^{(1)}(\kappa), Z^{(2)}(\kappa), Y(\kappa)\right] \qquad (24)$$

for each $\kappa \in (1/2, 1)$, as in Proposition 3.2.4 in M_a where $Z^{(1)}(\kappa), Z^{(2)}(\kappa)$ and $Y(\kappa)$ are independent zero mean Gaussian variates with variances $V_\kappa^{(1)}$, $V_\kappa - V_\kappa^{(1)}$ and V_κ respectively, where $V_\kappa^{(1)} = \sigma^2 c^{2\kappa-3}\Gamma(1-\kappa)^2/2$ and V_κ is defined in (10). This completes the proof of the Lemma 2, provided that $\tilde{U}_n \to_p 0$. To prove the latter, we employ Lemma 1(ii) to each term of (10), by taking $a_{n,j} \in \left\{\frac{A_{nj}^2(\kappa)}{L(n^\alpha)^2}, \frac{B_{nj}^2(\kappa)}{L(n^\alpha)^2}, \frac{C_{nj}^2(\kappa)}{L(n^\alpha)^2}\right\}$. The first part of (6) is satisfied since $\sum_{j=0}^{\tau_n(\beta)}|a_{n,j}|$ converges by Propositions 3.2.1-3.2.3 of M_a. For the second part of (6), $\sum_{j=0}^{\tau_n(\beta)} a_{n,j}^2$ is bounded by:

$$\begin{aligned}
\frac{1}{L(n^\alpha)^4}\sum_{j=0}^{\tau_n(\beta)} A_{nj}^4(\kappa) &= \frac{1}{L(n^\alpha)^4 n^{(6-4\kappa)\alpha}}\left(\sum_{i=0}^{\tau_n(\beta)} c_i\rho_n^{-i}\right)^4 \sum_{j=0}^{\tau_n(\beta)}\rho_n^{-4j} \\
&= \frac{1}{L(n^\alpha)^4 n^{(6-4\kappa)\alpha}} O\left(n^{4(1-\kappa)\alpha} L(n^\alpha)^4\right)\sum_{j=0}^{\tau_n(\beta)}\rho_n^{-4j} \\
&= O(1)\frac{1}{n^{2\alpha}}\sum_{j=0}^{\tau_n(\beta)}\rho_n^{-4j} = O\left(\frac{1}{n^\alpha}\right)
\end{aligned}$$

by Lemma A2(ii) of M_a; since $t \mapsto L(t)\, t^{-\kappa}$ is non-increasing on $[t_0, +\infty)$,

$$\begin{aligned}
\frac{1}{L(n^\alpha)^4} \sum_{j=1}^{\tau_n(\beta)} B_{nj}^4(\kappa) &= \frac{1}{L(n^\alpha)^4 n^{(6-4\kappa)\alpha}} \sum_{j=1}^{\tau_n(\beta)} \left(\sum_{t=1}^{\tau_n(\beta)} \rho_n^{-t} c_{t+j} \right)^4 \\
&= \frac{1}{L(n^\alpha)^4 n^{(6-4\kappa)\alpha}} \sum_{j=1}^{\tau_n(\beta)} \left[\left(\sum_{t=\lfloor t_0 \rfloor+1}^{\tau_n(\beta)} \rho_n^{-t} L(t+j)(t+j)^{-\kappa} \right) + O(1) \right]^4 \\
&\leq \frac{\tau_n(\beta)}{L(n^\alpha)^4 n^{(6-4\kappa)\alpha}} \left[\left(\sum_{t=\lfloor t_0 \rfloor+1}^{\tau_n(\beta)} \rho_n^{-t} L(t)\, t^{-\kappa} \right) + O(1) \right]^4 \\
&= \frac{\tau_n(\beta)}{L(n^\alpha)^4 n^{(6-4\kappa)\alpha}} O\left(n^{4(1-\kappa)\alpha} L(n^\alpha)^4 \right) \\
&= O\left(\frac{\tau_n(\beta)}{n^{2\alpha}} \right) = o\left(\frac{1}{n^{\alpha/2}} \right)
\end{aligned}$$

by Lemma A2(ii) of M_a; by using the C_r-inequality with $r = 4$ we obtain

$$\begin{aligned}
\frac{1}{L(n^\alpha)^4} \sum_{j=1}^{\tau_n(\beta)} C_{nj}^4(\kappa) &= \frac{1}{L(n^\alpha)^4 n^{(6-4\kappa)\alpha}} \sum_{j=1}^{\tau_n(\beta)} \left(\sum_{t=1}^{j} \rho_n^{-t} c_{j-t} \right)^4 \\
&= \frac{1}{L(n^\alpha)^4 n^{(6-4\kappa)\alpha}} \sum_{j=1}^{\tau_n(\beta)} \left(\sum_{t=0}^{j-1} \rho_n^{-(j-t)} c_t \right)^4 \\
&\leq \frac{1}{L(n^\alpha)^4 n^{(6-4\kappa)\alpha}} (r_{1n} + 8 r_{2n} + 8 r_{3n}) \qquad (25)
\end{aligned}$$

where

$$\begin{aligned}
r_{1n} &= \sum_{j=1}^{\lfloor n^\alpha \rfloor+1} \left(\sum_{t=0}^{j-1} \rho_n^{-(j-t)} c_t \right)^4 \leq n^\alpha \left(\sum_{t=0}^{\lfloor n^\alpha \rfloor} c_t \right)^4 = O\left(n^\alpha n^{4(1-\kappa)\alpha} L(n^\alpha)^4 \right) \\
r_{2n} &= \sum_{j=\lfloor n^\alpha \rfloor+2}^{\tau_n(\beta)} \left(\sum_{t=0}^{\lfloor n^\alpha \rfloor} \rho_n^{-(j-t)} c_t \right)^4 \leq \tau_n(\beta) \left(\sum_{t=0}^{\lfloor n^\alpha \rfloor} c_t \right)^4 = o\left(n^{3\alpha/2} n^{4(1-\kappa)\alpha} L(n^\alpha)^4 \right)
\end{aligned}$$

by (9) and

$$r_{3n} = \sum_{j=\lfloor n^\alpha \rfloor+2}^{\tau_n(\beta)} \left(\sum_{t=\lfloor n^\alpha \rfloor+1}^{j-1} \rho_n^{-(j-t)} c_t \right)^4 .$$

It clear that from the above bounds that the terms of (9) corresponding to r_{1n} and r_{2n} are of order $O\left(n^{-\alpha}\right)$ and $o\left(n^{-\alpha/2}\right)$ respectively. It remains to estimate r_{3n}. By Lemma A2(i) of M$_a$, $\rho_n^{-(j-t)} = e^{-\frac{c}{n^\alpha}(j-t)} + O\left(n^{-\alpha/2}\right)$ uniformly in t and j

$$\begin{aligned} r_{3n} &= \sum_{j=\lfloor n^\alpha\rfloor+2}^{\tau_n(\beta)} \left(\sum_{t=\lfloor n^\alpha\rfloor+1}^{j-1} \left[e^{-\frac{c}{n^\alpha}(j-t)} + O\left(n^{-\alpha/2}\right)\right] c_t \right)^4 \\ &\leq 8 \sum_{j=\lfloor n^\alpha\rfloor+2}^{\tau_n(\beta)} \left(\sum_{t=\lfloor n^\alpha\rfloor+1}^{j-1} e^{-\frac{c}{n^\alpha}(j-t)} c_t \right)^4 + 8O\left(n^{-2\alpha}\right) \sum_{j=\lfloor n^\alpha\rfloor+2}^{\tau_n(\beta)} \left(\sum_{t=\lfloor n^\alpha\rfloor+1}^{j-1} c_t \right)^4 \\ &= 8r'_{3n} + O\left(n^{-2\alpha}\right) O\left(n^\beta n^{4(1-\kappa)\beta} L\left(n^\beta\right)^4\right) \\ &= 8r'_{3n} + o\left(n^{(6-4\kappa)\alpha}\right) \end{aligned}$$

because (9) implies that $4(1-\kappa)(\beta-\alpha) \leq 2(1-\kappa)\alpha$, so

$$\begin{aligned} n^{-(6-4\kappa)\alpha} n^{-2\alpha} n^\beta n^{4(1-\kappa)\beta} L\left(n^\beta\right)^4 &= o\left(n^{-\alpha/2}\right) n^{-2\alpha} n^{4(1-\kappa)(\beta-\alpha)} L\left(n^\beta\right)^4 \\ &\leq o\left(n^{-\alpha/2} n^{-\kappa\alpha} L\left(n^\beta\right)^4\right). \end{aligned}$$

To deal with the final remainder term r'_{3n}, we employ Euler summation:

$$\begin{aligned} r'_{3n} &= \sum_{j=\lfloor n^\alpha\rfloor+2}^{\tau_n(\beta)} \left(\sum_{t=\lfloor n^\alpha\rfloor+1}^{j-1} e^{-\frac{c}{n^\alpha}(j-t)} t^{-\kappa} L(t) \right)^4 \\ &= \sum_{j=\lfloor n^\alpha\rfloor+2}^{\tau_n(\beta)} \left(\int_{\lfloor n^\alpha\rfloor+1}^{j-1} e^{-\frac{c}{n^\alpha}(j-\lfloor t\rfloor)} \lfloor t\rfloor^{-\kappa} L(\lfloor t\rfloor)\, dt \right)^4 \\ &= n^{4\alpha} \sum_{j=\lfloor n^\alpha\rfloor+2}^{\tau_n(\beta)} \left(\int_{(\lfloor n^\alpha\rfloor+1)/n^\alpha}^{(j-1)/n^\alpha} e^{-\frac{c}{n^\alpha}(j-\lfloor n^\alpha t\rfloor)} \lfloor n^\alpha t\rfloor^{-\kappa} L(\lfloor n^\alpha t\rfloor)\, dt \right)^4 \\ &= n^{4\alpha} \int_{\lfloor n^\alpha\rfloor+2}^{\tau_n(\beta)} \left(\int_{(\lfloor n^\alpha\rfloor+1)/n^\alpha}^{(\lfloor s\rfloor-1)/n^\alpha} e^{-\frac{c}{n^\alpha}(\lfloor s\rfloor-\lfloor n^\alpha t\rfloor)} \lfloor n^\alpha t\rfloor^{-\kappa} L(\lfloor n^\alpha t\rfloor)\, dt \right)^4 ds \\ &= n^{5\alpha} n^{-4\kappa\alpha} L(n^\alpha)^4 \int_{(\lfloor n^\alpha\rfloor+2)/n^\alpha}^{\tau_n(\beta)/n^\alpha} \left(\int_{(\lfloor n^\alpha\rfloor+1)/n^\alpha}^{(\lfloor n^\alpha s\rfloor-1)/n^\alpha} g_n(s,t)\, dt \right)^4 ds \\ &\leq n^\alpha n^{4(1-\kappa)\alpha} L(n^\alpha)^4 \int_1^\infty \left(\int_1^s g_n(s,t)\, dt \right)^4 ds \qquad (26) \end{aligned}$$

where, for all $t \geq 1$

$$
\begin{aligned}
g_n(s,t) &= e^{-\frac{c}{n^\alpha}(\lfloor n^\alpha s\rfloor - \lfloor n^\alpha t\rfloor)} \left(\frac{\lfloor n^\alpha t\rfloor}{n^\alpha}\right)^{-\kappa} \frac{L(\lfloor n^\alpha t\rfloor)}{L(n^\alpha)} \\
&= e^{-c\left(\frac{\lfloor n^\alpha s\rfloor}{n^\alpha s} - \frac{\lfloor n^\alpha t\rfloor}{n^\alpha t}\right)} \frac{L(n^\alpha)}{L(\lfloor n^\alpha\rfloor)} \left(\frac{\lfloor n^\alpha\rfloor}{n^\alpha}\right)^{-\kappa} e^{-c(s-t)} \left(\frac{\lfloor n^\alpha t\rfloor}{\lfloor n^\alpha\rfloor}\right)^{-\kappa} \frac{L(\lfloor n^\alpha t\rfloor)}{L(\lfloor n^\alpha\rfloor)} \\
&\leq C_1 e^{-c(s-t)} \left(\frac{\lfloor n^\alpha t\rfloor}{\lfloor n^\alpha\rfloor}\right)^{-\kappa} \frac{L(\lfloor n^\alpha t\rfloor)}{L(\lfloor n^\alpha\rfloor)} \\
&\leq C_2 e^{-c(s-t)} \left(\frac{\lfloor n^\alpha t\rfloor}{\lfloor n^\alpha\rfloor}\right)^{-\kappa+\delta} \leq C e^{-c(s-t)} t^{-\kappa+\delta}
\end{aligned}
$$

eventually for arbitrary $\delta > 0$ and C_1, C_2, C uniform constants by Potter's Theorem (see BGT, Theorem 1.5.6.(i)). Using the above bound, the integral in (26) satisfies

$$\int_1^\infty \left(\int_1^s g_n(s,t)\,dt\right)^4 ds \leq C^4 \int_1^\infty e^{-4cs}\left(\int_1^s e^{ct} t^{-(\kappa-\delta)} dt\right)^4 ds < \infty$$

by choosing $\delta \in (0, \kappa - 1/2)$, because $I = \int_1^\infty e^{-4cs}\left(\int_1^s e^{ct} t^{-\lambda} dt\right)^4 ds < \infty$ for any $\lambda > 1/2$. To see this, note that $\lim_{s\to\infty} e^{-cs}\int_1^s e^{ct} t^{-\lambda} dt = 0$ by L'Hospital's rule; applying integration by parts twice we obtain

$$
\begin{aligned}
I &= \frac{1}{c}\int_1^\infty e^{-3cs} s^{-\lambda}\left(\int_1^s e^{ct} t^{-\lambda} dt\right)^3 ds \\
&= \frac{1}{3c^2}\int_1^\infty e^{-3cs}\left\{3s^{-2\lambda} e^{ct}\left(\int_1^s e^{ct} t^{-\lambda} dt\right)^2 - \lambda s^{-\lambda-1}\left(\int_1^s e^{ct} t^{-\lambda} dt\right)^3\right\} ds \\
&= \frac{1}{c^2}\int_1^\infty s^{-2\lambda}\left(\int_1^s e^{-c(s-t)} t^{-\lambda} dt\right)^2 ds - \frac{\lambda}{3c^2}\int_1^\infty s^{-\lambda-1}\left(\int_1^s e^{-c(s-t)} t^{-\lambda} dt\right)^3 ds \\
&\leq \frac{1}{c^2}\int_1^\infty s^{-2\lambda} ds \left(\int_0^\infty e^{-cu} du\right)^2 < \infty.
\end{aligned}
$$

By (26) we conclude that $r'_{3n} = O\left(n^\alpha n^{4(1-\kappa)\alpha} L(n^\alpha)^4\right)$ and standardising by the normalisation of (25), $L(n^\alpha)^{-4} n^{-(6-4\kappa)\alpha} r'_{3n} = O(n^{-\alpha})$. This shows that the right side of (25) tends to 0 as $n \to \infty$ and completes the proof of the lemma.

Proof of Lemma 3. Having established the joint asymptotic behaviour of $L(n^{\alpha})^{-1} Y_n(\kappa)$ and $L(n^{\alpha})^{-1} Z_n(\kappa)$ under Assumption CH in Lemma 2, the proof of Lemma 3 follows the same steps as the proof of Lemma 5 of M_a: see page 185 of M_a.

References.

Aue, A. and L. Horváth (2007). A limit theorem for mildly explosive autoregression with stable errors. Econometric Theory 23, 201-220.

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

Andrews, D. W. (1988). Laws of large numbers for dependent non-identically distributed random variables. Econometric Theory, 4(3), 458-467.

Bingham, N.H., Goldie, C.M. and J.L. Teugels (1987). Regular Variation. CUP.

Bollerslev, T., Sizova, N., & Tauchen, G. (2011). Volatility in equilibrium: Asymmetries and dynamic dependencies. Review of Finance, 16(1), 31-80.

Carrasco, M., & Chen, X. (2002). Mixing and moment properties of various GARCH and stochastic volatility models. Econometric Theory, 18(1), 17-39.

Davidson, J. (1994). Stochastic limit theory: An introduction for econometricians. OUP Oxford.

Ghysels, E., Harvey, A. C., & Renault, E. (1996). Stochastic volatility. Handbook of statistics, 14, 119-191.

Giraitis, L., Koul, H.L., D. Surgailis (1996). Asymptotic normality of regression estimators with long memory errors. Statistics and Prob. Letters, 29, 317-335.

Giraitis, L., Kokoszka, P., & Leipus, R. (2000). Stationary ARCH models: dependence structure and central limit theorem. Econometric Theory, 16(1), 3-22.

Giraitis, L. and P. C. B. Phillips (2006). Uniform Limit Theory for Stationary Autoregression. Journal of Time Series Analysis, 27, 51-60.

Goldie, C. M. (1991). Implicit renewal theory and tails of solutions of random equations. The Annals of Applied Probability, 126-166.

Hall, P. and C.C. Heyde (1980). Martingale Limit Theory and its Application. AP.

Ibragimov, I. A. and Linnik, I. U. V. (1971). Independent and stationary sequences of random variables. Groningen: Wolters-Noordhoff.

Lee, J. H. (2017). Limit theory for explosive autoregression under conditional heteroskedasticity. Journal of Statistical Planning and Inference, Forthcoming.

Magdalinos, T. (2009). Appendix to "Mildly explosive autoregression under weak and strong dependence".

Magdalinos, T. and P. C. B. Phillips (2008). Limit Theory for Cointegrated Systems with Moderately Integrated and Moderately Explosive Regressors. Econometric Theory, 25, 482-526.

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

McLeish, D. L. (1975). A maximal inequality and dependent strong laws. The Annals of Probability, 3(5), 829-839.

Oh, H., Lee, S., & Chan, N. H. (2017). Mildly explosive autoregression with mixing innovations. Journal of the Korean Statistical Society, Forthcoming.

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

Phillips, P. C. B. and T. Magdalinos (2007b). Limit theory for Moderate deviations from a unit root under weak dependence. G. D. A. Phillips and E. Tzavalis (Eds.) The Refinement of Econometric Estimation and Test Procedures. CUP.

Phillips, P.C.B., Wu, Y. and J. Yu (2009). Explosive behavior in the 1990s NASDAC: When did exuberance escalate asset values? Cowles discussion paper 1699.

Phillips, P. C., & Yu, J. (2011). Dating the timeline of financial bubbles during the subprime crisis. Quantitative Economics, 2(3), 455-491.

Samorodnitsky, G. (2006). Long Range Dependence. Now Publishers.

Straumann, D. (2004). Estimation in Conditionally Heteroscedastic Time Series Models. Springer.

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

Wu, W.B. and W. Min (2005). On linear processes with dependent innovations. Stochastic Processes and their Applications 115, 939-958.