Discussion Papers
Department of Economics
University of Copenhagen

No. 19-03

A PRIMER ON BOOTSTRAP TESTING OF HYPOTHESES IN TIME SERIES MODELS: WITH AN APPLICATION TO DOUBLE AUTOREGRESSIVE MODELS

Giuseppe Cavaliere and Anders Rahbek

Øster Farimagsgade 5, Building 26, DK-1353 Copenhagen K., Denmark
Tel.: +45 35 32 30 01 – Fax: +45 35 32 30 00
http://www.econ.ku.dk

ISSN: 1601-2461 (E)

A PRIMER ON BOOTSTRAP TESTING OF HYPOTHESES IN TIME SERIES MODELS: WITH AN APPLICATION TO DOUBLE AUTOREGRESSIVE MODELS

GIUSEPPE CAVALIERE[*] AND ANDERS RAHBEK[‡]

ABSTRACT

In this paper we discuss the general application of the bootstrap as a tool for statistical inference in econometric time series models. We do this by considering the implementation of bootstrap inference in the class of double-autoregressive [DAR] models discussed in Ling (2004). DAR models are particularly interesting to illustrate implementation of the bootstrap to time series: first, standard asymptotic inference is usually difficult to implement due to the presence of nuisance parameters under the null hypothesis; second, inference involves testing whether one or more parameters are on the boundary of the parameter space; third, under the alternative hypothesis, fourth or even second order moments may not exist. In most of these cases, the bootstrap is not considered an appropriate tool for inference. Conversely, and taking testing (non-) stationarity to illustrate, we show that although a standard bootstrap based on unrestricted parameter estimation is invalid, a correct implementation of a bootstrap based on *restricted* parameter estimation (restricted bootstrap) is first-order valid; that is, it is able to replicate, under the null hypothesis, the correct limiting null distribution. Importantly, we also show that the behaviour of this bootstrap under the alternative hypothesis may be different because of possible lack of finite second-order moments of the bootstrap innovations. This features makes – for some parameter configurations – the restricted bootstrap unable to replicate the null asymptotic distribution when the null is false. We show that this drawback can be fixed by using a new 'hybrid' bootstrap, where the parameter estimates used to construct the bootstrap data are obtained with the null imposed, while the bootstrap innovations are sampled with replacement from the unrestricted residuals. We show that this bootstrap, novel in this framework, mimics the correct asymptotic null distribution, irrespetively of the null to be true or false. Throughout the paper, we use a number of examples from the bootstrap time series literature to illustrate the importance of properly defining and analyzing the bootstrap generating process and associated bootstrap statistics.

KEYWORDS: Bootstrap; Hypothesis testing; Double-Autoregressive models; Parameter on the boundary; Infinite Variance.

JEL CLASSIFICATION: C32.

---

[*]Department of Economics, University of Bologna, Italy, and Department of Economics, Exeter Business School, UK. [†]Department of Economics, University of Copenhagen, Denmark. Correspondence to: Anders Rahbek, Department of Economics, University of Copenhagen, Øster Farimagsgade 5, 1353 København, Denmark; email: anders.rahbek@econ.ku.dk. The paper is based on the Econometric Theory Lecture given by Anders Rahbek at the 8^{th} Italian Congress of Econometrics and Empirical Economics (ICEEE), Lecce, 23-25 January 2019. We thank conference participants and, importantly, James MacKinnon, Ye Lu and and David Harris for important feedbacks on previous versions of the paper. Part of the paper was written when Anders Rahbek visited the Department of Economics of the University of Bologna, whose hospitality is gratefully acknowledged. This research was supported by the Danish Council for Independent Research (DSF Grant 015-00028B) and by an ALMA IDEA grant from the University of Bologna.

1 Introduction

Outcomes of various bootstrap schemes applied to econometric time series models are routinely reported in the literature. This is generally done in cases where (i) the limiting distribution of the reference estimator or test statistics depends on a (possibly infinite-dimensional) vector of unknown nuisance parameters; (ii) critical values or standard errors can be obtained by simulations only; (iii) the asymptotic approximation to the distribution of the reference estimator or test statistics is poor. The increasing computational power available to researchers coupled with the fact that the implementation of bootstrap algorithms is typically straightforward, makes the bootstrap one of the most popular inference tools in the econometric analysis of time series data; see, *inter alia*, Davidson and MacKinnon (2006) and MacKinnon (2009).

Despite its many appealing features, the application of the bootstrap to time series models requires a detailed analysis of its asymptotic properties. This is necessary in order to establish asymptotic validity of the bootstrap, at least up to first order. Taking hypothesis testing to illustrate – as we do throughout this paper – a proper statistical analysis of any bootstrap test would necessarily involve two main, interconnected steps. First, it requires to determine whether, conditionally on the original data, the bootstrap correctly mimics the null asymptotic distribution of the reference test statistics under the null hypothesis. This step is generally more involved than the asymptotic analysis of the original test statistics, as the conditional distribution of the bootstrap statistic given the data is a random element in the space of distribution functions. Hence, specific probability tools are required. In general, further high level conditions over those required for asymptotic inference are necessary and, consequently, any application of the bootstrap which is not backed up by a proper analysis of these conditions must be taken with caution.

The second step, which is often neglected in applications of the bootstrap, is the statistical analysis of the properties of the test under the alternative hypothesis, *i.e.* consistency of the bootstrap test. This step is more involved than assessing bootstrap validity under the null. Essentially, difficulties may arise because it requires to analyze the asymptotic behaviour of the estimators used to generate the bootstrap data when the null is false: in particular when estimators restricted by the null hypothesis are considered. More specifically, under the alternative hypothesis it is important to check whether the estimators restricted by the (false) null hypothesis do not converge to pseudo-true values that ruin the large sample features of the bootstrap sample.

In this paper we aim at discussing the two aforementioned steps by considering a novel application of the bootstrap to econometric time series models. Specifically, we consider bootstrap inference in the class of double-autoregressive [DAR] models, see *e.g.* Borkovec and Klüppelberg (2001), Ling (2004, 2007*a*) and Chen, Li and Ling (2013). The DAR is a time series model with an autoregressive structure both in the conditional mean and in the conditional variance. The conditional mean has the classic autoregressive formulation, *i.e.* it is linear in the lagged level of the process. The conditional variance, in contrast to the classic ARCH-type or AR–ARCH type specifications (Ling and Li, 1998; Ling and MacAleer, 2003; Lange, Jensen and Rahbek, 2006; Ling, 2007*b*; Nielsen and Rahbek, 2014) where it is linear in the squared lagged innovations, is also linear in the lagged *level* of the process. In this sense, it allows the levels of the process to affect both the conditional mean and conditional variance, as expected e.g. in the econometric modelling of interest rates[1] as in Nielsen and Rahbek (2014). In the framework of the DAR model, we consider as leading testing application the bootstrap implementation of the likelihood-ratio [LR] test for reduction to random walk. In essence, this can be viewed as a (non-)stationarity test within the DAR model. Previous studies of this testing problem are given in Ling (2004), who focus on the score test, and in Kluppenberg *et al.* (2001), who consider a LR testing approach.

DAR models and the associated (non-)stationarity testing problem are particularly interesting to illustrate implementation of the bootstrap to time series, for several reasons. First, standard asymptotic inference is usually difficult to implement, due to the presence of nuisance parameters under the null hypothesis. The asymptotic distribution of the test statistics, for instance, depends on nuisance parameters (such as the kurtosis of the innovations) which makes it hard to construct tables of critical values. Second, the autoregressive parameter entering the conditional variance equation is – in order to guarantee non-negativity of the conditional variance – usually restricted to be non-negative. As a consequence, inference must deal with possible parameters on the boundary of the parameter space, a situation where the bootstrap is usually regarded as invalid (see, e.g. Andrews, 2000, Cavaliere, Nielsen and Rahbek, 2016). Third, under strict stationarity, second order moments may not exist. Hence, understanding the properties of the bootstrap under the alternative hypothesis, which would require re-sampling from an infinite variance process, may be cumbersome, if not even impossible (seminal results about the possible invalidity of the bootstrap when second order moments may not exist are given in Athreya, 1987, and Knight, 1988; for time series models see also Cavaliere, Nielsen and Rahbek, 2018, and the references therein).

In the following of the paper we show that, as expected in the aforementioned cases, classic bootstrap hypothesis testing, based on generating the bootstrap data using estimators (and residuals) obtained without imposing the null hypothesis (as suggested in Hall, 1992), is not valid. Despite this fact, we also show that the problem of (non-)stationarity testing in a DAR model can be successfully solved by a proper implementation of the bootstrap. More specifically, we initially show that the bootstrap based on *restricted* parameter estimation (the so-called 'restricted bootstrap') is first-order valid under the null hypothesis; that is, it is able to replicate the correct limiting null distribution when the null hypothesis is true. However, we also show that the behaviour of this bootstrap under the alternative hypothesis may be different because of possible lack of finite second-order moments of the bootstrap innovations. This features makes – for some parameter configurations – the restricted bootstrap unable to replicate the null asymptotic distribution when the null is false. This is a typical instance where validity of the bootstrap under the null does not imply its validity under the alternative.

We next show that this drawback can be fixed by using a new 'hybrid' bootstrap, where the parameter estimates used to construct the bootstrap data are obtained with the null imposed, while the bootstrap innovations are sampled with replacement from the unrestricted residuals. This simple modification of the bootstrap algorithm, which is novel in this framework, mimics the correct asymptotic null distribution also under the alternative.

[1] The Cox-Ingersoll-Ross (CIR) Model is a classic example of a level-dependent heteroskedasticity model.

We use a Monte Carlo experiment to analyze the finite sample properties of the different bootstrap algorithms. We show substantial gains in terms accuracy of the empirical rejection probabilities under the null hypothesis, while under the alternative we show that our bootstrap has power very close to the (pointwise) size-adjusted power of the (infeasible) asymptotic test.

Throughout the paper, we use a number of examples from the bootstrap (time series) literature to illustrate the importance of properly defining the bootstrap generating process and associated bootstrap statistic, as well as the need of looking at the appropriate bootstrap statistic on the base of a rigorous, case-by-case analysis of its theoretical properties, both under the null and under the alternative hypothesis.

1.1 Structure of the paper

The structure of the paper is the following. In section 2 we introduce the reference DAR model and the testing problem we consider throughout the paper. In section 3 we introduce the main bootstrap approaches and discuss their validity under the null hypothesis. Section 4 focuses on the behaviour of the bootstrap test under the alternative hypothesis. Here we also introduce and discusse the hybrid bootstrap scheme. Results from a small Monte Carlo study on the finite sample behaviour of the asymptotic and bootstrap tests are reported in section 5. We consider some extensions of the model and of the tests considered in section 6. Section 7 concludes. All mathematical proofs are reported in the appendix.

1.2 Notation

The following notation is used throughout. With $x := y$ ($y =: x$) we mean that x is defined by y (y defined by x). For any $q \in \mathbb{R}$ ($\mathbb{R}$ denoting the set of real numbers), $q^+ := \max\{0, q\}$ and $\lfloor q \rfloor$ denotes the integer part of q. The set of non-negative real numbers is denoted by $\mathbb{R}^+$. The space of $m \times 1$ vectors of càdlàg functions on the unit interval $[0, 1]$ is denoted by D^m. With $X_n \to_w X$ and $X = \text{wlim}\, X_n$ we mean that X_n converges weakly to X. Also, $\overset{d}{=}$ denotes equality in distribution. We use P^*, E^* and V^* respectively to denote probability, expectation and variance, conditional on the original sample. With $\overset{w^*}{\to}_p$ we denote weak convergence in probability; that is, $X_n^* \overset{w^*}{\to}_p X$ means that, as the sample size n diverges, the cdf G_n^* of X_n, conditional on the original data, converges in probability to the cdf G of X, at all continuity points of G. For a given sequence X_n^* computed from the bootstrap data, $X_n^* - X = o_p^*(1)$, in probability, or $X_n^* \overset{p^*}{\to}_p X$, means that for any $\epsilon > 0$, $P^*(||X_n^* - X|| > \epsilon) \to_p 0$, as $n \to \infty$. Similarly, $X_n^* = O_p^*(1)$, in probability, means that, for every $\epsilon > 0$, there exists a constant $M > 0$ such that, for all large n, $P(P^*(||X_n^*|| > M) < \epsilon)$ is arbitrarily close to one. Unless otherwise specified, integrals are between 0 and 1.

2 (Non-)stationarity in a DAR model

In this section we present the leading DAR model and the associated (non-)stationary testing problem which we discuss throughout the paper. We introduce the main assumptions in Section 2.1, discuss estimation in Section 2.2 and the key testing procedure in section 2.3. Bootstrap inference and hypothesis testing is be discussed in section 3.

2.1 Model and assumptions

Consider the double-autogressive [DAR] model (Ling, 2004), as defined through the recursion

$$\Delta x_t = \pi x_{t-1} + \varepsilon_t, \quad \varepsilon_t := \sigma_t z_t, \quad \sigma_t^2 := \omega + \alpha x_{t-1}^2 \tag{1}$$

where the z_t's is an i.i.d. random variables with zero mean and unit variance, and with a continuous, strictly positive density with respect to the Lebesgue measure[2]. The initial value, denoted by x_0, is independent of the future z_t's and will be considered fixed in the statistical analysis. As is customary for this class of models, it is also assumed that[3] $\xi := Ez_t^3 = 0$, and $\kappa := Ez_t^4 - 1 < \infty$. In this model, the mean of x_t conditional on the σ-field generated by $\{x_0, z_1, ..., z_{t-1}\}$, say $\mathcal{I}_{t-1}$, equals $(1+\pi)x_{t-1}$ while the conditional variance is given by $\sigma_t^2 := \omega + \alpha x_{t-1}^2$ and hence is level-dependent. In this respect, it differs from the standard AR-ARCH model (see *e.g.* Lange, Rahbek and Jensen, 2011), where the conditional variance σ_t^2 depends on ε_{t-1}^2 rather than on x_{t-1}^2 (see also Nielsen and Rahbek, 2014, for a discussion of the multivariate DAR). Clearly, the model reduces to a standard autoregression with i.i.d. innovation when $\alpha = 0$, and to the ARCH model when $\pi = -1$, which implies $x_t = (\omega + \alpha x_{t-1}^2)^{1/2} z_t$. In the DAR model, a sufficient condition for σ_t^2 to be positive a.s. is given by the usual non-negativity constraint $\alpha \geq 0$, which we assume to hold throughout. A necessary and sufficient condition for $Ex_t^2 < \infty$ is $(1+\pi)^2 + \alpha < 1$; moreover, provided $E \log |1 + \pi + \alpha^{1/2} z_t| < 0$, the process can be given an initial distribution such that it is strictly stationary and geometrically ergodic if some mild regularity conditions on the density function of z_t also hold. A key feature of the model is that the classical autoregressive unit root condition, $\pi = 0$, does not imply that the process is non-stationary. More specifically, $\pi = 0$ implies non-stationarity only if $\alpha = 0$; see Figure 1 in Ling (2004). We discuss the issue of testing for non-stationarity in Section 2.3 below.

In the following we assume that the parameter space for the true value, denoted as θ_0, is given by $\Theta_0 := \Theta_{\mathcal{S}} \cup \Theta_{\mathcal{N}}$, where $\Theta_{\mathcal{S}} := \{\theta := (\pi, \alpha, \omega)' : E \log |1 + \pi + \alpha^{1/2} z_t| < 0$ with $\alpha \geq 0$ and $\omega > 0\}$ and $\Theta_{\mathcal{N}} := \{\theta := (0, 0, \omega)' : \omega > 0\}$. That is, we assume that either the process is strictly stationary (the true parameter is in $\Theta_{\mathcal{S}}$), or that the process is non-stationary and, specifically, reduces to a standard random walk with i.i.d. increments (the true parameter is in $\Theta_{\mathcal{N}}$). In both cases it is assumed that the innnovations z_t have finite fourth order moments.

2.2 Estimation

As in Ling (2004) and in Klüppenberg *et al.* (2002), we consider quasi maximum likelihood [QML] estimation based on the auxiliary assumption of Gaussian innovations. The results given here are employed in sections 3 and 4 in order to establish the properties of the bootstrap test. We further assume that the user-chosen optimization set employed for maximization of the likelihood function is given by

[2] The assumption of a continuous and positive density with respect to the Lebesgue measure can be relaxed.

[3] In Section 6, we consider $\zeta \neq 0$.

$\mathcal{T} := \{\theta := (\pi, \alpha, \omega)' : -\pi_L \leq \pi \leq \pi_U,\ 0 \leq \alpha \leq \alpha_U,\ \omega_L \leq \omega \leq \omega_U\}$, with $\pi_L, \pi_U, \alpha_U, \omega_L$ and ω_U positive constants and $\omega_L < \omega_U$. In practice, estimation is performed imposing the non-negativity restriction $\alpha \geq 0$ while leaving π unrestricted (and ω positive).

For a time series $\{x_1, ..., x_n\}$, and with x_0 fixed in the statistical analysis, the Gaussian QMLE is given by

$$\hat{\theta}_n := \arg\max_{\theta \in \mathcal{T}} L_n(\theta),\ L_n(\theta) := \sum_{t=1}^{n} l_t(\theta)$$

where

$$l_t(\theta) := -\frac{1}{2}\log\sigma_t^2(\theta) - \frac{1}{2}\left(\frac{\Delta x_t - \pi x_{t-1}}{\sigma_t(\theta)}\right)^2,\ \sigma_t^2(\theta) := \omega + \alpha x_{t-1}^2,\ t = 1, ..., n.$$

Theory for the QMLE under the strict stationarity assumption, *i.e.* when the true parameter θ_0 is in $\Theta_{\mathcal{S}}$, is provided in Ling (2004) under the assumption that α_0 is *not* on the boundary (specifically, it is required that $\alpha_0 \in [\alpha_L, \alpha_U]$ with $\alpha_L > 0$), hence not covering the case where α_0 may be zero, that is, on the boundary. By employing non-standard arguments as e.g. in Andrews (1999, 2001), see also Cavaliere, Nielsen and Rahbek (2017), we generalize Ling (2004, Theorem 1) as follows:

Theorem 1 *Suppose that $\{x_t\}$ is generated as in (1) with $\xi = 0$ and $\kappa < \infty$, and that the true parameter vector $\theta_0 \in \Theta_{\mathcal{S}}$. Then, as $n \to \infty$, $\hat{\theta}_n = (\hat{\pi}_n, \hat{\alpha}_n, \hat{\omega}_n)'$ is consistent,* i.e. *$\hat{\theta}_n \to_p \theta_0 = (\pi_0, \alpha_0, \omega_0)'$. The asymptotic distribution of $\hat{\theta}_n$ is given by*

$$n^{1/2}(\hat{\theta}_n - \theta_0) \to_w \zeta = (\zeta_\pi, \zeta_\gamma)',$$

with $\zeta_\pi \overset{d}{=} N(0, \sigma_\pi^2)$, $\sigma_\pi^2 := 1/E(x_{t-1}^2/\sigma_t^2)$. Moreover, ζ_π is independent of the bivariate random vector $\zeta_\gamma := (\zeta_\alpha, \zeta_\omega)'$, where:
(i) for $\alpha_0 > 0$, $\zeta_\gamma \overset{d}{=} N(0, \Omega_{\gamma\gamma})$ with $\Omega_{\gamma\gamma}$ given in the Appendix, eq.(A.22);
(ii) for $\alpha_0 = 0$, then, with $\varrho := Ex_t^2$,

$$\zeta_\alpha = \max(0, \zeta_\alpha^0),\ \textit{and}\ \zeta_\omega = \zeta_\omega^0 - \varrho\max(0, \zeta_\alpha^0),$$

where $\zeta_\alpha^0 \overset{d}{=} N(0, \sigma_\alpha^2)$ and $\zeta_\omega^0 \overset{d}{=} N(0, \sigma_\omega^2)$ are independent, $\sigma_\alpha^2 = \sigma_\omega^2/\delta$, $\sigma_\omega = \sqrt{\kappa}\omega_0$ and $\delta = E(x_t^4) - (E(x_t^2))^2$.

With respect to Ling (2004), the asymptotic distribution is no longer Gaussian when $\alpha_0 = 0$ due to the restriction that $\alpha \geq 0$. As a result, the asymptotic distribution of ($n^{1/2}$ times) $\hat{\alpha}_n$ is 'half-normal', i.e. of the form $\zeta^+ := \max(0, \zeta)$ with ζ Gaussian. For the case of $\alpha_0 > 0$, the asymptotic distribution of ζ is as in Ling (2004, Theorem 1). Note that asymptotic normality and consistency at the $n^{1/2}$-rate is established even in cases where $E(\Delta x_t)^2 = +\infty$, due to the structure of the score of the likelihood function, see Appendix A.2 (and Jensen and Rahbek (2004) for similar arguments in the pure (G)ARCH case).

Remark 2.1 Note that the results in Theorem 1 can be generalized to the case of $\xi \neq 0$. In this case however, see Appendix Appendix A.2, ζ_π and ζ_γ are dependent with covariance matrix $\mathrm{Cov}(\zeta_\pi, \zeta_\gamma) = \xi\Omega_{\pi\gamma} \neq 0$, with $\Omega_{\pi\gamma}$ given in Appendix A.2, eq.(A.16).

As it will appear clear from the discussions about the bootstrap tests, we also need to understand the properties of the estimator under non-stationarity. These are provided in the following theorem.

THEOREM 2 *Suppose that $\{x_t\}$ is generated as in (1) with $\xi = 0$ and $\kappa < \infty$ and that the true parameter vector $\theta_0 \in \Theta_{\mathcal{N}}$, i.e. $\pi_0 = 0$ and $\alpha_0 = 0$. Then, as $n \to \infty$, $\hat{\theta}_n \to_p \theta_0$. Moreover,*

$$\text{diag}(n, n^{3/2}n^{1/2})(\hat{\theta}_n - \theta_0) \to_w \lambda = (\lambda_\pi, \lambda_\alpha, \lambda_\omega)',$$

where, with B and W independent standard Brownian motions,

$$\lambda_\pi := \int B\mathsf{d}B/\left(\int B_u^2\mathsf{d}u\right), \; \lambda_\alpha := (\lambda_\alpha^0)^+ = \max\left(0, \lambda_\alpha^0\right),$$

for

$$\lambda_\alpha^0 := \sqrt{\kappa}\left(\int B^2\mathsf{d}W - \int B_u^2\mathsf{d}uW_1\right)/\left(\int B_u^4\mathsf{d}u - (\int B_u^2\mathsf{d}u)^2\right).$$

Moreover, $\lambda_\omega = \lambda_\omega^0 - (\int B^2\mathsf{d}u)\lambda_\alpha$, where $\lambda_\omega^0 \overset{d}{=} \sigma_\omega W_1$ and $\sigma_\omega := \sqrt{\kappa}\omega_0$.

REMARK 2.2 With respect to the (strict) stationary case, we observe that the rate of convergence of the estimator varies across parameters. In particular, $\hat{\pi}_n$ converges at the rate of n, similar to the standard autoregressive case with a unit roots, while the volatility parameter, $\hat{\alpha}_n$, converges at the faster rate of $n^{3/2}$. The estimator of the intercept term in the variance equation has the usual stationary, $n^{1/2}$, rate.

REMARK 2.3 While λ in Theorem 2 clearly is non-Gaussian, and thus different from the stationary case with $\alpha_0 = 0$ in Theorem 1, one can immediately observe some similarities: *(i)* in the expression for λ_π, the term $(\int B_u^2\mathsf{d}u)^{-1}$ corresponds to the variance σ_π^2 of ζ_π; *(ii)* in λ_α^0, the term $\sqrt{\kappa}(\int B_u^4\mathsf{d}u - (\int B_u^2\mathsf{d}u)^2)^{-1}$ corresponds to $\sigma_a^2 = \sigma_\omega^2/\delta$ in ζ_α^0; *(iii)* finally, in the expression for λ_ω, while $\lambda_\omega^0 \overset{d}{=} \zeta_\omega^0$, the loading $\int B^2\mathsf{d}u$ corresponds to the ϱ term in ξ_ω.

REMARK 2.4 Similar to the case of Theorem 1, also Theorem 2 can be modified to the asymmetric case of $\xi \neq 0$, see the discussion in Section 6.

2.3 TESTING (NON)STATIONARITY

Suppose that the econometrician is interested in testing whether $\{x_t\}$ is non-stationary, against the alternative of (strict) stationarity. In a pure AR–ARCH framework, the (unit root) null hypothesis corresponds to $\pi = 0$ in eq. (1). However, the DAR process can be strictly stationary even if $\pi = 0$, provided $\alpha > 0$ and $E \log|1 + \alpha^{1/2}z_t| < 0$; hence, testing nullity of π is not alone sufficient to assess the (non-stationarity) of x_t. Rather, as discussed in Ling (2004), one may test the pure random walk hypothesis, as given by $\mathsf{H}_0 : \pi = 0$, $\alpha = 0$, against the alternative $\mathsf{H}_1 : \pi \neq 0$, $\alpha \geq 0$. The likelihood ratio test can easily be computed in the usual way as

$$LR_n := -2(L_n(\tilde{\theta}_n) - L_n(\hat{\theta}_n)) \tag{2}$$

where $\tilde{\theta}_n := (0, 0, \tilde{\omega}_n)'$, $\tilde{\omega}_n := n^{-1} \sum_{t=1}^{n} (\Delta x_t)^2$, denotes the restricted estimator of θ, i.e. $\tilde{\theta}_n := \arg\max_{\theta \in \mathcal{T}_0} L_n(\theta)$ where $\mathcal{T}_0 := \{\theta := (0, 0, \omega)' : \omega_L \leq \omega \leq \omega_U\}$. Now, the asymptotics in the previous Theorem 1 obviously breaks down when $\theta_0 \in \Theta_{\mathcal{N}}$, see Theorem 2. In this case, Klüppelberg *et al.* (2002) establish the following result for the LR test statistic in (2).

THEOREM 3 *Suppose that $\{x_t\}$ is generated as in (1) with $Ez_t^4 < \infty$ and that the true parameter vector $\theta_0 \in \Theta_{\mathcal{N}}$, i.e. $\pi_0 = 0$ and $\alpha_0 = 0$. Then, as $n \to \infty$, $LR_n \to_w \mathcal{LR}_\infty(\kappa)$, where*

$$\mathcal{LR}_\infty(\kappa) = \frac{\kappa}{2}\left(\max\left(0, \frac{\int B_u^2 \mathsf{d}u W_1 - \int B_u^2 \mathsf{d}W_u}{(\int B_u^4 \mathsf{d}u - (\int B_u^2 \mathsf{d}u)^2)^{1/2}}\right)\right)^2 + \frac{(\int B_u \mathsf{d}B_u)^2}{\int B_u^2 \mathsf{d}u} \tag{3}$$

where B and W are as in Theorem 2.

Some remarks follow.

REMARK 2.5 Notice that since B and W are independent, conditionally on B, we have in particular that

$$\frac{W_1 \int B_u^2 \mathsf{d}u - \int B_u^2 \mathsf{d}W_u}{(\int B_u^4 \mathsf{d}u - (\int B_u^2 \mathsf{d}u)^2)^{1/2}} \stackrel{d}{=} \frac{N\left(0, \int (B_u^2 - (\int B_u^2 \mathsf{d}u))^2 \mathsf{d}u\right)}{(\int (B_u^2 - (\int B_u^2 \mathsf{d}u))^2 \mathsf{d}u)^{1/2}} \stackrel{d}{=} N(0, 1).$$

This implies that the first term in (3) is distributed as $\frac{\kappa}{2}(\max(0, N(0,1)))^2$, *i.e.* as the half-$\chi_1^2$ distribution. Moreover, it is independent of the second term , $(\int B^2 \mathsf{d}u)^{-1}(\int B \mathsf{d}B)^2$, which is a squared Dickey-Fuller distribution. Should the condition $\xi = 0$ fail to hold, both the half χ_1^2 property and the independence of the two terms in (3) would no longer hold true; see also Section 6.

REMARK 2.6 The distribution in (3) is non-pivotal, since it depends on κ. A consistent estimator of this quantity can be constructed by using the unrestricted residuals, as $\hat{\kappa}_n := n^{-1} \sum_{t=1}^{m} (1 - \hat{z}_t^2)^2$, where $\hat{z}_t := \hat{\varepsilon}_t / \hat{\sigma}_t$ for $\hat{\varepsilon}_t := \Delta x_t - \hat{\pi}_n x_{t-1}$, $\hat{\sigma}_t^2 := \hat{\omega}_n + \hat{\alpha}_n x_{t-1}^2$. An estimator $\tilde{\kappa}_n$ which imposes the null hypothesis may be constructed using the restricted residuals, $\tilde{z}_t := \tilde{\omega}_n^{-1/2} \Delta x_t$. However, this estimator overestimates κ when the null hypothesis does not hold, hence reducing the power of an asymptotic test based on $\mathcal{LR}_\infty(\tilde{\kappa}_n)$. □

3 BOOTSTRAPPING THE ASYMPTOTIC DISTRIBUTION UNDER THE NULL HYPOTHESIS

3.1 PRELIMINARIES AND BOOTSTRAP ALGORITHMS

The classical requirement of any bootstrap implementation is consistent estimation of the asymptotic null distribution of the reference test statistic when the null hypothesis is true. Specifically, and taking the LR_n test statistic to illustrate, consider a bootstrap analog, say LR_n^*, which is a function of the original sample and of a vector of bootstrap innovations, say $\eta_1^*, ...\eta_m^*$, defined jointly with the original data on a possibly expanded probability space. With $G_n^*(\cdot) := P^*(LR_n^* \leq x) := P(LR_n^* \leq x|\{x_t\})$ denoting the conditional distribution of LR_n^* given the original data, this requires that, *under the null hypothesis*, $G_n^*(\cdot) \to_p G_\infty(\cdot)$, where G_∞ denotes the cdf of $\mathcal{LR}_\infty(\kappa)$, the asymptotic distribution of LR_n under the null; see eq. (3). This is the well-known concept of 'weak convergence, in probability', denoted as $LR_n^* \xrightarrow{w^*}_p \mathcal{LR}_\infty$ (or, $\mathcal{LR}_\infty(\kappa)$) as used repeatedly to emphasize the dependence on the nuisance parameter κ). If, additionally, $G_\infty(\cdot)$ is continuous, then by Polya's theorem proximity of $G_n^*(\cdot)$ to $G_\infty(\cdot)$ holds in the sup norm,

$$\sup_{x\in\mathbb{R}} |G_n^*(x) - G_\infty(x)| \to_p 0,$$

and the bootstrap p-value, given by

$$p_n^* := 1 - G_n^*(LR_n),$$

is asymptotically uniformly distributed, i.e. $p_n^* \to_w U[0,1]$. This allows to construct a bootstrap test with the correct asymptotic size at any nominal significance level.

Two main approaches can be given in order to define the bootstrap statistic LR_n^*. The first, the 'restricted bootstrap', is based on estimation of the original model with the null hypothesis imposed; i.e. with $\pi = \alpha = 0$. In this case, the bootstrap statistic mimics the original test statistic and tests the restriction $\pi = \alpha = 0$ on the bootstrap data. The second, the 'unrestricted bootstrap', uses the unrestricted parameter estimates $\hat{\pi}_n, \hat{\alpha}_n$ to generate the bootstrap data and the bootstrap statistic is based on testing $\pi = \hat{\pi}_n$ and $\alpha = \hat{\alpha}_n$ on the bootstrap data; see e.g. Hall (1992). We introduce the restricted bootstrap first.

Restricted (i.i.d.) Bootstrap:

(i) Estimate model (1) using Gaussian QML under the null hypothesis, yielding the estimates $\tilde{\theta}_n := (0, 0, \tilde{\omega}_n)'$, together with the corresponding restricted QML residuals, $\tilde{\varepsilon}_t := \Delta x_t$ and $\tilde{z}_t := \tilde{\omega}_n^{-1/2}\tilde{\varepsilon}_t$, as defined above;

(ii) Standardize the residuals as

$$\tilde{z}_{s,t} := \frac{\tilde{z}_t - n^{-1}\sum_{t=1}^n \tilde{z}_t}{(n^{-1}\sum_{t=1}^n (\tilde{z}_t - n^{-1}\sum_{t=1}^n \tilde{z}_t)^2)^{1/2}}$$

and construct the bootstrap innovations using the i.i.d. bootstrap re-sampling scheme; i.e., $z_t^* := \tilde{z}_{s,\eta_t^*}$, where η_t^*, $t = 1, ..., n$ is an i.i.d. sequence of discrete uniform distributions on $\{1, 2, ..., n\}$;

(iii) Construct the bootstrap sample $\{x_t^*\}$ from the recursion

$$\Delta x_t^* = \varepsilon_t^*, \quad \varepsilon_t^* := \sigma_t^* z_t^*, \quad \sigma_t^{*2} = \tilde{\omega}_n, \, t = 1, \ldots, n, \tag{4}$$

with the n bootstrap errors z_t^* generated in Step (ii) and with initial values $x_0^* = x_0$.

(iv) Using the bootstrap sample, $\{x_t^*\}$, compute the bootstrap test statistic LR_n^*. Define the corresponding p-value as $p_n^* := 1 - G_n^*(LR_n)$ with $G_n^*(\cdot)$ denoting the conditional (on the original data) cumulative density function (cdf) of LR_n^*.

(v) The restricted bootstrap test of H_0 at level ζ rejects if $p_n^* \leq \zeta$.

There are many variants of the restricted bootstrap, as exemplified in the following remarks.

Remark 3.1 In the definition above, the length of the bootstrap sample equals the length of the original sample, n. A different sample size, say $m < n$, could be used in order to form the bootstrap sample. This is the so-called 'm out of n' bootstrap, which (under proper conditions on m as n increases, such as $m^{-1}+mn^{-1} \to 0$) has been proved to be asymptotically valid in certain cases where bootstraps based on n observations fail; see Politis, Romano and Wolf (1999) and the references therein. However, for the 'm out of n' bootstrap, while mathematically appealing in the derivations of the asymptotic theory, the choice of m is 'delicate' (see Davison, Hinkley and Young, 2003), and, moreover, in general it does not deliver satisfactory finite sample results.

Remark 3.2 The bootstrap shocks in Step 2 are based on i.i.d. re-sampling (*i.e.*, with replacement) from the standardized residuals. Different bootstrap schemes could in principle be used. For instance, the so-called wild bootstrap (Wu, 1986; Liu, 1988; Mammen, 1993) generates the bootstrap innovations as the (conditionally) independent sequence $z_t^* := \tilde{z}_{s,t} w_t^*$ where w_t^* is i.i.d.(0,1) with bounded fourth order moments. Alternatively, re-sampling without replacement of the $\tilde{z}_{s,t}$'s could be employed, leading to the permuted bootstrap sample $z_t^* = \tilde{z}_{s,\pi^*(t)}$, $t = 1, ..., n$, where $\{\pi^*(1), ..., \pi^*(n)\}$ is a (uniformly distributed) random permutation of $\{1, ..., n\}$ (Cavaliere, Georgiev and Taylor, 2016; Cavaliere, Nielsen and Rahbek, 2018). Finally, a fully parametric bootstrap could be obtained by generating z_t^* as i.i.d. from any pre-specified zero mean, unit variance, distribution.

Remark 3.3 In practice, the cdf G_n^* required in Step (iv) of Algorithm 1 can only be approximated through numerical simulation. As is standard, this requires generating B (conditionally) independent bootstrap statistics, $LR_{n:b}^*$, $b = 1, \ldots, B$, computed as above. The approximated bootstrap p-value for LR_n, is then computed as $\tilde{p}_n^* := B^{-1} \sum_{b=1}^{B} \mathbb{I}(LR_{n:b}^* > LR_n)$, and is such that $\tilde{p}_n^* \overset{a.s.}{\to} p_n^*$ as $B \to \infty$. For the choice of B, see, *inter alia*, Andrews and Buchinsky (2000) and Davidson and MacKinnon (2000). □

The key feature of the restricted bootstrap is that the parameter estimates used in constructing the bootstrap sample data are obtained under the restriction of the null hypothesis, H_0. As discussed for instance in Hall (1992), in the statistics literature it is often the case that in bootstrap implementations parameters are estimated without imposing the null hypothesis, and to subsequently calculate a bootstrap test statistic for the hypothesis $\theta = \hat{\theta}_n$, that is, the hypothesis that θ equals the unrestricted estimate. Formally, this corresponds to the unrestricted bootstrap, as defined through the following steps.

Unrestricted (i.i.d.) Bootstrap:

(i) Estimate model (1) using Gaussian QML without imposing the null hypothesis, yielding the estimates $\hat{\theta}_n := (\hat{\pi}_n, \hat{\alpha}_n, \hat{\omega}_n)'$, together with the corresponding unrestricted QML residuals, $\hat{\varepsilon}_t := \Delta x_t - \hat{\pi}_n x_{t-1}$ and $\hat{z}_t := (\hat{\omega}_n + \hat{\alpha}_n x_{t-1}^2)^{-1/2} \hat{\varepsilon}_t$, as defined above;

(ii) Standardize the residuals as

$$\hat{z}_{s,t} := \frac{\hat{z}_t - n^{-1}\sum_{t=1}^{n}\hat{z}_t}{(n^{-1}\sum_{t=1}^{n}(\hat{z}_t - n^{-1}\sum_{t=1}^{n}\hat{z}_t)^2)^{1/2}}$$

and construct the bootstrap innovations using the i.i.d. bootstrap re-sampling scheme; i.e., $z_t^* := \hat{z}_{s,\eta_t^*}$, where η_t^*, $t = 1, ..., n$ is an i.i.d. sequence of discrete uniform distributions on $\{1, 2, ..., n\}$;

(iii) Construct the bootstrap sample $\{x_t^*\}$ from the recursion

$$\Delta x_t^* = \hat{\pi}_n x_{t-1}^* + \varepsilon_t^*, \quad \varepsilon_t^* := \sigma_t^* z_t^*, \quad \sigma_t^{*2} = \hat{\omega}_n + \hat{\alpha}_n (x_{t-1}^*)^2,\ t = 1, \ldots, n,$$

with the n bootstrap errors z_t^* generated in step (ii) and with initial values $x_0^* = x_0$.

(iv) Using the bootstrap sample, $\{x_t^*\}$, compute the bootstrap test statistic LR_n^* for the (auxiliary) null hypothesis $\pi = \hat{\pi}_n$, $\alpha = \hat{\alpha}_n$. Define the corresponding p-value as $p_n^* := 1 - G_n^*(LR_n)$ with $G_n^*(\cdot)$ denoting the conditional (on the original data) cumulative distribution function (cdf) of LR_n^*.

(v) The unrestricted bootstrap test of H_0 at level ζ rejects if $p_n^* \leq \zeta$.

The logic behind the unrestricted bootstrap is to avoid potential power losses that the restricted bootstrap test may experience because of incorrectly imposing a false null hypothesis when the null does not hold. There are, however, many cases where the unrestricted bootstrap fails to mimic the asymptotic distribution, whereas the restricted bootstrap does not. Among those, two cases are extremely relevant for the testing problem considered here. The first is the case of bootstrapping when data have unit roots – as it happens in the DAR model when $\pi = 0$. The second is the case where a parameter lies on the boundary of the parameter space – which again appears in our testing problem as $\alpha = 0$ is a boundary point under the maintained hypothesis that $\alpha \geq 0$. We briefly discuss these two examples in the following.

Example 1 (Unit roots and unrestricted bootstrap) *As in Basawa et al. (1991), consider the first order autoregression with a unit root,*

$$\Delta x_t = \pi x_{t-1} + \varepsilon_t,\ \pi = 0,$$

ε_t *i.i.d.*$N(0, \omega)$, $x_0 = 1$ *and* $t = 1, ..., n$. *Let* J_c *denote an Ornstein-Uhlenbeck process with mean reversion parameter* c *(such that* $c = 0$ *corresponds to a standard Brownian Motion) and set* $\tau(c) := \int J_c \mathsf{d}J_c / \int J_c^2 \mathsf{d}u$. *The QMLE of* π *is the least squares estimator,* $\hat{\pi}_n = \sum_{t=1}^{n} \Delta x_t x_{t-1} / \sum_{t=1}^{n} x_{t-1}^2$, *which satisfies*

$$\tau_n := n\hat{\pi}_n \to_w \tau_\infty := \tau(0) \tag{5}$$

see Phillips (1987) and the references therein. Now, consider a (fully parametric) unrestricted bootstrap, based on the recursion

$$\Delta x_t^* = \hat{\pi}_n x_{t-1}^* + \varepsilon_t^*, \tag{6}$$

initialized at $x_0^* = x_0$, *and* ε_t^* *i.i.d.* $N(0,1)$ $(t = 1, ..., n)$ *With* $\hat{\pi}_n^*$ *the bootstrap (least squares) estimator,* $\hat{\pi}_n^* = \sum_{t=1}^n \Delta x_t^* x_{t-1}^* / \sum_{t=1}^n (x_{t-1}^*)^2$, *the bootstrap analog of* τ_n *is defined as* $\tau_n^* := n\hat{\pi}_n$. *Unfortunately, despite* $\hat{\pi}_n$ *is superconsistent,* τ_n^* *fails to mimic the asymptotic distribution in (5). Essentially, because* $n\hat{\pi}_n = O_p(1)$ *rather than* $o_p(1)$, *the bootstrap sample (normalized by the usual rate* $n^{-1/2}$*) behaves, in large samples, as an Ornstein-Uhlenbeck process with* random *drift parameter, rather than as a Brownian motion. To see why, replace* $\hat{\pi}_n$ *in (6) by a sequence* π_n *such that* $n \cdot \pi_n \to v$. *Then, by Chan and Wei (1987, Theorem 1) we have that (conditionally on the original data),* $\tau_n^* := n(\hat{\pi}_n^* - \hat{\pi}_n)$ *is asymptotically distributed as* $\tau(v)$ *(see Basawa* et al.*, 1991). In our case,* $\tau_n := n\hat{\pi}_n \to_w \tau_\infty$ *and, as a result, the bootstrap statistic has a random distribution function, even for* $n \to \infty$, *given by* $\tau(\tau_\infty)$. *More specifically, it can be proved that*

$$\begin{aligned} P^*(\tau_n^* \le x) &= P(\tau_n^* \le x | \tau_n) = P(n(\hat{\pi}_n^* - \hat{\pi}_n) \le x | \tau_n) \\ &\to_w P\left(\int J_{\tau_\infty} \mathrm{d} J_{\tau_\infty} / \int J_{\tau_\infty}^2 \mathrm{d}u \le x \middle| \tau_\infty \right). \end{aligned}$$

That is, the limiting distribution can be written in terms of an Ornstein-Uhlenbeck process with a random drift, distributed as τ_∞, *i.e. as a Dickey-Fuller distribution. Similar arguments are applied in Cavaliere, Nielsen and Rahbek (2015), see also the next section, and in terms of aymptotically random bootstrap measures in Cavaliere and Georgiev (2018).*

Example 2 (Unit roots and the restricted bootstrap) *Despite the unrestricted bootstrap fails to mimic the unit root distribution, the restricted bootstrap does not; see Cavaliere and Taylor (2008, 2009*a*) and Cavaliere, Rahbek and Taylor (2012) for the multivariate case. Specifically, by imposing the unit root on the bootstrap sample, i.e. by setting*

$$\Delta x_t^* = \varepsilon_t^*,$$

where ε_t^* *are i.i.d.* $N(0,1)$ *and* $t = 1, ..., n$, *it is guaranteed that* $\tau_n^* := n\hat{\pi}_n^* \overset{w^*}{\to}_p \tau(0)$, *in probability.*

Alternatively, it follows by standard arguments, that one may use an 'm out of n' version of the unrestricted bootstrap which, by considering samples of size $m = o(n)$ *ensures that* $m\hat{\pi}_n = o_p(1)$ *as* $m \to \infty$, *which is sufficient for* $\tau_m^* := m\hat{\pi}_m^* \overset{w^*}{\to}_p \tau(0)$, *in probability. However, as already emphasized, while the asymptotic arguments are mathematically appealing, in practice the 'm out of n' bootstrap in this case does not have adequate finite sample properties.*

Example 3 (Boundary problems and the unrestricted bootstrap) *The standard unrestricted bootstrap is also known to fail when (some of) the parameters lie on the boundary of the parameter space. Consider, as in Cavaliere, Nielsen and Rahbek (2017), see also Andrews (2000), the Gaussian ARCH model,*

$$x_t = \sqrt{\omega + \alpha x_{t-1}^2} z_t,$$

with z_t *i.i.d. N(0,1). Moreover, the optimization set is given by* $\mathcal{T}_\alpha = \{\alpha : \alpha \in [0, \alpha_U]\}$, $\alpha_0 \in \Theta_{\mathcal{S}_\alpha}$, *with* $\Theta_{\mathcal{S}_\alpha} = \{\alpha : E \log (\alpha z_t^2) < 0\}$, *while* ω *is kept fixed for simplicity here.*

We consider here testing $\alpha_0 = 0$ by the likelihood ratio statistic, LR_n. As in Theorem 3 for the DAR, the MLE $\hat{\alpha}_n$ satisfies for $\alpha_0 > 0$,

$$\sqrt{n}(\hat{\alpha}_n - \alpha_0) \to_w \tfrac{\kappa}{\delta}\zeta,\ \zeta \sim N(0,1),$$

$\delta = V\left(x_t^2\right)$, and the associated LR statistic for $\alpha = \alpha_0$ is asymptotically χ_1^2 (times $\frac{\kappa}{2}$).

In constrast, if $\alpha_0 = 0$,

$$\sqrt{n}(\hat{\alpha}_n - \alpha_0) \to_w \alpha_\infty = \max\{0, \zeta\},$$

and the associated LR statistic for $\alpha = 0$ has asymptotic distribution given by,

$$LR_n \to_w \tfrac{\kappa}{2}\zeta^2 1(\zeta \geq 0) = \tfrac{\kappa}{2}\max\{0,\zeta\}^2.$$

Now, consider instead the (parametric) unrestricted bootstrap sample, as given by $x_t^* = \sqrt{\omega + \hat{\alpha}_n x_{t-1}^{*2}} z_t^*$, with z_t^* i.i.d. $N(0,1)$ (independent of the original data), and the associated bootstrap statistic, LR_n^*, for the (bootstrap) hypothesis that α equals the bootstrap true value, $\hat{\alpha}_n$. With $\zeta^* \sim N(0,1)$ and independent of ζ, we conjecture from the theory in Cavaliere, Nielsen, Pedersen and Rahbek (2018) that, conditionally on the original data, the asymptotic distribution of the LR_n^* statistic has a random limit,

$$\tfrac{\kappa}{2}(\zeta + \alpha^*)^2 1(\zeta + \alpha^* \geq 0)\Big|\ \alpha^*,$$

where α^* is a function of α_∞ given above. Thus, as expected the unrestricted bootstrap fails to mimic the null asymptotic distribution. □

3.2 Bootstrap validity of the DAR model

Testing the pure random walk hypothesis in the DAR framework features the complications discussed in the previous Examples 1 and 3. First, since the null hypothesis implies a unit root in the data, a bootstrap which does not impose the unit root on the bootstrap sample is likely to fail to be first-order valid. Second, since the null hypothesis implies a parameter (α) on the boundary of the parameter space, a bootstrap which does not account for this feature may display a random limiting distribution. The unrestricted bootstrap is neither imposing the unit root nor restricting α to be on the boundary of the parameter space; hence, it fails to be first-order valid. Conversely, under mild conditions the restricted bootstrap is able to replicate the correct null limiting distribution of the LR test when the null hypothesis holds true. This is proved in the next theorem.

Theorem 4 *Under the conditions of Theorem 3, provided $Ez_t^8 < \infty$, as $n \to \infty$ the restricted bootstrap LR statistic satisfies:*

$$LR_n^* \overset{w^*}{\to}_p \mathcal{LR}_\infty(\kappa).$$

The logic behind the proof of bootstrap validity under the null hypothesis is the following. When the restricted bootstrap is employed, the sample bootstrap is generated as

$$\Delta x_t^* = \varepsilon_t^* = \hat{\omega}_n z_t^*.$$

Conditionally on the original data, the bootstrap score, see Appendix A, depends on the vector $\left(z_i^*, z_t^{*2} - 1\right)$, which needs satisfying a (bootstrap) functional central limit theorem of the form,

$$\mathcal{Z}_n^* (\cdot) := n^{-1/2} \sum\nolimits_{i=1}^{\lfloor n\cdot \rfloor} \left(z_i^*, z_t^{*2} - 1\right) \xrightarrow{w^*}_p \left(B^*, \sqrt{\kappa} W^*\right) \tag{7}$$

with B^* and W^* two independent standard Brownian motions. It is therefore crucial to control what conditions are needed for (7) to hold, given that z_t^* is a zero mean (conditionally) i.i.d. sample from the centered standardized residuals, $\tilde{z}_{s,t}$. This requires checking whether the (conditional) variance of $\mathcal{Z}_n^* (\cdot)$ converges to $\text{diag}(1, \kappa)$ and whether the Lindeberg condition holds. As shown in the Appendix, these requirements hold provided z_t has bounded eight order moments.

It is worth emphasizing that for the DAR model, the limiting distribution of the LR_n test statistic for reduction to pure random walk feaures a nuisance parameter, namely the constant κ. This makes the testing problem based on asymptotic inference convoluted, since the practitioner needs first to estimate κ using a proper (consistent) estimator, say $\hat{\kappa}$, and then using Monte Carlo methods to simulate the quantiles of limiting distribution $\mathcal{LR}_\infty (\hat{\kappa})$. The bootstrap allows to circumvent this problem, as it replicate the correct limiting distribution without the need of plug-in methods. This is an example of a classic application of the bootstrap to time series data, where it is used to retrieve quantiles from an asymptotic distribution which depends on a (possibly infinite dimensional) vector of nuisance parameters, see the following example.

Example 4 (Non-stationary volatility) *A classic instance of a limiting distribution depending on nuisance parameter is the case of 'non-stationary' volatility. In this case, in the simplest form the innovations of an econometric model can be represented as $\varepsilon_t = \sigma_t z_t$, where z_t is an i.i.d. finite variance sequence and $\sigma_t = h(t/n)$, where h is a bounded function satisfying some regularity conditions (e.g., it is càdlàg; see Cavaliere, 2004, Cavaliere, Rahbek and Taylor, 2014, and Boswijk* et al.*, 2017). In this case, the partial sum process associated to ε_t delivers the following result*

$$S_n(\cdot) := \frac{1}{n^{1/2}} \sum_{t=1}^{\lfloor n\cdot \rfloor} \varepsilon_t \xrightarrow{w} M(\cdot) := \int_0^{\cdot} h(u) \, \mathsf{d}B(s),$$

where B is a Brownian motion. In this specific case, M is a continuous-time martingale with covariance kernel given by $Cov\left(M(s), M(s')\right) = \int_0^{\min\{s,s'\}} h(u)^2 \, \mathsf{d}u$. Limit distributions of estimators and test statistics usually depend on such covariance kernel, which is unknown in practice. Although consistent estimators could be constructed (see e.g. Cavaliere and Taylor, 2007), the bootstrap can in general automatically replicate the limiting functional M. That is, consider a vector of residuals $\hat{\varepsilon}_t$ satisfying $n^{-1} \sum\nolimits_{t=1}^{n} (\hat{\varepsilon}_t^2 - \varepsilon_t^2) = o_p(1)$, and construct the bootstrap errors using the 'wild' bootstrap as

$$\varepsilon_t^* := \hat{\varepsilon}_t w_t^*, \; t = 1, ..., n,$$

where the w_t's are i.i.d. $N(0, 1)$. Then, it holds, as $n \to \infty$, see Boswijk et al. *(2017) and the references therein,*

$$S_n^*(\cdot) := \frac{1}{n^{1/2}} \sum_{t=1}^{\lfloor n\cdot \rfloor} \varepsilon_t^* \xrightarrow{w^*}_p M(\cdot) \tag{8}$$

and hence the wild bootstrap replicates the same limiting distribution of the original functional S_n. □

4 The behaviour of the bootstrap under the alternative hypothesis

4.1 Preliminaries and bootstrap consistency

The analysis of the large sample properties of the bootstrap test statistic under the alternative hypothesis is a key requirement for a correct implementation of the bootstrap. Unfortunately, as it will be exemplified later in this section, this step is in general more involved than just proving validity under the null hypothesis.

Ideally, one would aim that, under the alternative hypothesis, LR_n^* is (asymptotically) distributed as the LR_n limit under the null. This would require that, as $n \to \infty$,

$$LR_n^* \xrightarrow{w^*}_p \mathcal{LR}_\infty(\kappa) \tag{9}$$

also when H_0 does not hold.

This immediately implies that the (bootstrap) test is consistent: if LR_n diverges to $+\infty$ under the alternative hypothesis then, with G_n^* denoting the cdf of LR_n^* conditional on the original data, it holds that the bootstrap p-value satisfies $p_n^* := 1 - G_n^*(LR_n) \to_p 0$. Moreover, in large samples a test based on the (conditional) quantiles of LR_n^* would have power approximately equal to the size-adjusted power of the (asymptotic) test based on the quantiles of $\mathcal{LR}_\infty$.

In fact, a weaker result that implies bootstrap consistency can be used in case (9) does not hold. Specifically, a sufficient condition for the bootstrap p-value to shrink to zero under the alternative is (again, provided $LR_n \to \infty$ under the alternative hypothesis)

$$LR_n^* = O_p^*(1), \text{ in probability,} \tag{10}$$

or the even weaker result that

$$LR_n^* = o_p^*(LR_n), \text{ in probability.} \tag{11}$$

In the first case, the bootstrap test statistic is bounded in probability, which implies consistency of the test at the usual rate. In the second case, both the bootstrap and the original test statistics diverge to $+\infty$; however, the fact that the conditional quantiles of LR_n^* diverge at a slower rate implies bootstrap consistency.

Two simple examples are now given.

Example 5 (ARCH | Boundary and restricted bootstrap) *In Example 3, unrestricted bootstrap based testing for $\mathsf{H}_0 : \alpha = 0$ was discussed in the ARCH model given by,*

$$x_t = \sigma_t z_t, \quad \sigma_t^2 = \omega + \alpha x_{t-1}^2, \quad \theta = (\alpha, \omega)'.$$

Recall furthermore that the likelihood ratio statistic LR_n has the asymptotic limiting distribution as given by,

$$\mathcal{LR}_\infty(\kappa) = \frac{\kappa}{2}(\zeta^+)^2 = \frac{\kappa}{2}\zeta^2 \max\{0, \zeta\},$$

with ζ a $N(0,1)$ random variable. Consider here the restricted bootstrap based on i.i.d. resampling of the (standardized) restricted residuals proposed in Cavaliere, Nielsen and Rahbek (2017), hereafter. With $\tilde{\theta}_n := (\tilde{\omega}_n, 0)'$ denoting the restricted (QML) estimator, the bootstrap data are given by

$$x_t^* := \sqrt{\tilde{\omega}_n} z_t^*, \tag{12}$$

with z_t^* sampled with replacement from the standardized residuals from restricted estimation, given by $\tilde{z}_t^s := (\tilde{z}_t - \bar{\tilde{z}}_n)/(n^{-1}\sum_{t=1}^n (\tilde{z}_t - \bar{\tilde{z}}_n)^2)^{1/2}$, $\bar{\tilde{z}}_n := n^{-1}\sum_{t=1}^n \tilde{z}_t$, with $\tilde{z}_t := x_t/\sqrt{\tilde{\omega}_n}$. The bootstrap shocks $\{z_t^* : t \leq n\}$ are an i.i.d. sample from $\tilde{z}_t^s$, $t = 1, ..., n$, such that, conditionally on the original data, $E^*(z_t^*) = 0$ and $V^*(z_t^*) = 1$. Cavaliere, Nielsen and Rahbek (2017, Theorem 1) show that under the null hypothesis, the bootstrap QLR statistic, say LR_n^*, satisfies

$$LR_n^* \xrightarrow{w^*}_p \mathcal{LR}_\infty(\kappa), \tag{13}$$

hence mimicking the correct asymptotic null distribution. However, if the null hypothesis does not hold, result (13) may no longer hold. Essentially, the reason is that the unrestricted estimator $\tilde{\omega}_n$ equals $n^{-1}\sum_{t=1}^n x_t^2$, which may even diverge under the stated assumptions. For instance, while under the null hypothesis $x_t = \omega^{1/2} z_t$, which implies that also $\{x_t : t \geq 1\}$ has finite fourth order moments, under the alternative hypothesis x_t may have infinite fourth order moments. If, additionally, it is assumed that x_t has finite fourth order moments, such that $\kappa^\dagger := E(x_t^4)/(E(x_t^2))^2 - 1 < \infty$, by Theorem 1 in Cavaliere, Nielsen and Rahbek (2017) it follows that under the alternative,

$$LR_n^* \xrightarrow{w^*}_p \mathcal{LR}_\infty(\kappa^\dagger),$$

such that $LR_n^* = O_p^*(1)$, in probability. Hence, while as shown in Example 3 the unrestricted bootstrap is invalid, the restricted is.

Finally, note that when $\alpha_0 \neq 0$ the constant $\kappa^\dagger > \kappa$, hence implying a potential power loss of the bootstrap test with respect to the asymptotic test.

Example 6 (Hypothesis testing on the cointegrating vectors) *Consider a p-dimensional VAR process with r co-integrating relations, as given by*

$$\Delta x_t = \pi x_{t-1} + \varepsilon_t, \quad \pi = \alpha\beta' \quad (t = 1, ..., n), \tag{14}$$

with $\{\varepsilon_t\}$ independent and identically distributed (i.i.d.) with mean zero and full-rank variance matrix Ω, and where the initial values $x_{1-k}, ..., x_0$ are fixed in the statistical analysis. Furthermore, assume that the so-called '$I(1, r)$ conditions' holds; that is, (a) the characteristic polynomial associated with (14) has $p - r$ roots equal to 1 and all other roots outside the unit circle, and (b) α and β have full column rank r. Under these conditions x_t is $I(1)$ with co-integration rank r, such that the co-integrating relations $\beta' x_t$ are stationary. We want to test the null hypothesis $\mathsf{H}_0 : \beta = \tau$, where τ a known $p \times r$ matrix of full column rank r. To this aim, it is customary to consider the LR test of Johansen (1996), which rejects H_0 when the associated LR statistic LR_n is large. Under the null, LR_n is asymptotically $\chi^2_{p(p-r)}$, see Johansen (1996), while under the alternative LR_n diverges, see Cavaliere, Nielsen and Rahbek (2015, Remark 3.4). Hence, the asymptotic LR test is consistent. Now, consider a restricted bootstrap for H_0, as initially proposed in Fachin (2000), Gredenhoff and Jacobson (2001) and later *discussed in Fachin and Omzigt (2006). This bootstrap requires estimation of (14) under* H_0 *and then use the corresponding (restricted) estimates* $\tilde{\alpha}_n$ *and* τ *to generate the bootstrap sample as*

$$\Delta x_t^* = \tilde{\alpha}_n \tau' x_{t-1}^* + \varepsilon_t^*, \tag{15}$$

where the bootstrap shocks ε_t^* *are obtained by re-sampling (after re-centering) from the corresponding restricted residuals,* $\tilde{\varepsilon}_t := \Delta x_t - \tilde{\alpha}_n \tau' x_{t-1}$. *Under* H_0*, consistency of* $\tilde{\alpha}_n$ *implies, along with a bootstrap (functional) CLT for* $\{\varepsilon_t^*\}$*, that the bootstrap LR statistic, say* LR_n^**, satisfies*

$$LR_n^* \xrightarrow{w^*}_p \chi^2_{p(p-r)}.$$

Hence, the bootstrap mimics the correct asymptotic distribution under the null. However, as proved in Cavaliere et al. (2015), the same result does not *hold when* H_0 *is false. Intuitively, this is the case because when* H_0 *is false,* $\tau' X_{t-1}$ *is no longer stationary, and hence the restricted estimators* $\tilde{\alpha}_n, \tau, \tilde{\Gamma}_i$ *are based on the unbalanced regression of* Δx_t *(stationary) on* $\tau' x_{t-1}$ *(non-stationary in* $p - r^*$ *directions , with* $r^* < r$*) and lags of* Δx_t *(stationary). This implies that* $\tilde{\alpha}_n \tau'$*, properly normalized, does not converge to a constant but, rather, to a* stochastic *matrix of reduced rank* r^* *(see Cavaliere, Nielsen and Rahbek, 2015, Proposition 1). As a consequence, the bootstrap estimator of* β *is no longer mixed Gaussian (as it is under the null hypothesis) and the statistic* LR_n^* *has a random limiting distribution which differs from the target* χ^2 *distribution. However, it still holds that* $LR_n^* = O_p^*(1)$*, in probability, as in (10), hence implying that the bootstrap test is consistent.*

Example 7 (Bootstrap financial bubbles) *Phillips, Wu and Yu (2011) consider testing for an explosive bubble regime, based on the supremum of a set of recursive right-tailed DF test statistics,* τ_n. *While Harvey, Leybourne, Sollis, and Taylor (2016) show that the restricted (Wild) bootstrap statistic* τ_n^* *mimics the right limiting distribution under the null hypothesis, this result does not hold under the alternative; neither does it hold that* $\tau_n^* = O_p^*(1)$*, in probability. Rather, Harvey* et al. *(2016) show that* $\tau_n^* = O_p^*\left(n^{1/2}\right)$*, in probability and hence both the original and the bootstrap statistics diverge to* $+\infty$. *But since the bootstrap statistic diverges at a polynomial rate* $n^{1/2}$ *while the original statistic diverges at the exponential rate* $n^{1/2}(1+\delta_1)^{n(\tau_2-\tau_1)}$*, see Theorem 3 in Harvey* et al. *(2016), the bound in (11) applies and the bootstrap test rejects with probability tending to one as* n *diverges.* □

4.2 On consistency of the boostrap for the DAR model

Despite the restricted bootstrap correctly estimates the null asymptotic distribution under the null hypothesis, its performance under the alternative is not at all straightforward to establish. This is because, under the alternative hypothesis of strict stationarity, the retricted residuals $\tilde{z}_t$ are no longer close enough to the true innovations, z_t, and do not share the same properties in terms of moments. Consequently, the bootstrap score and information may have different asymptotic properties with respect to their sample analogs. Intuitively, this happens because while under the null hypothesis, $\tilde{z}_t \approx z_t$, under the alternative hypothesis $\tilde{z}_t = \tilde{\omega}_n^{-1/2} \Delta x_t$, where x_t may not possess finite fourth order moments (take, for instance, the case where $\alpha + (1+\pi)^2 = 1$ with $\pi \neq 0$, such that x_t is strictly stationary and ergodic but $E x_t^2 = +\infty$).

More precisely, recall that a first requirement for the asymptotic result in Theorem 4 is to assess whether the bootstrap functional CLT [FCLT] in (7) holds, with z_t^* (conditionally on the original data) i.i.d. from the centered standardized residuals, $\tilde{z}_{s,t} := (n^{-1}\sum_{t=1}^{n}(\tilde{z}_t - n^{-1}\sum_{t=1}^{n}\tilde{z}_t)^2)^{-1/2}(\tilde{z}_t - n^{-1}\sum_{t=1}^{n}\tilde{z}_t)$. In term of z_t^*, conditions for

$$n^{-1/2}\sum_{i=1}^{[n\cdot]} z_i^* \xrightarrow{w^*}_p B^*(\cdot)$$

with B^* a standard Brownian motion, are: (i) $E^* z_i^* = 0$, (ii) $E^*(z_i^*)^2 = \frac{1}{n}\sum_{t=1}^{n}(\tilde{z}_t^s)^2 \to_p 1$, and a Lindeberg condition (similar results are required for $n^{-1/2}\sum_{i=1}^{[n\cdot]}(z_t^{*2}-1)$). While (i) and (ii) are trivially satisfied, from e.g. Lemma B.1 in Cavaliere *et al.* (2016) it follows that the Lindeberg condition holds provided Δx_t has bounded kurtosis. Specifically, if $E(\Delta x_t)^4 < \infty$, then

$$\hat{\kappa}_{z^s} := E^*((z_t^*)^2 - 1)^2 \to \kappa^\dagger := \frac{E(\Delta x_t)^4}{(E(\Delta x_t)^2)^2} - 1 < \infty \tag{16}$$

and the Lindeberg condition holds. A sufficient condition for this is

$$(1+\pi)^4 + (\kappa+1)\,\alpha^2 + 6\,(1+\pi)^2\,\alpha^2 < 1 \tag{17}$$

such that, in the case special case where $\pi = -1$ (pure ARCH(1) process) and under Gaussianity, we get the well-known condition $\alpha < \frac{1}{\sqrt{\kappa+1}} = \frac{1}{\sqrt{3}}$. Hence, the following Theorem can be established.

Theorem 5 *Let the conditions of Theorem 3 hold, and consider the retricted bootstrap test statistic, LR_n^*. Then, under H_1, if additionally (17) holds, then as $n \to \infty$:*

$$LR_n^* \xrightarrow{w^*}_p \mathcal{LR}_\infty(\kappa^\dagger)$$

where $\kappa^\dagger > \kappa$, with $\kappa^\dagger$ defined in (16).

This theorem proves that even in the case of bounded fourth order moments of Δx_t, under the alternative hypothesis the bootstrap does not mimic the asymptotic distribution given in Theorem 3. Rather, it is shifted to the right: as n get large, the bootstrap critical values are therefore expected to be shifted to the right with respect to the critical values from the true null distribution $\mathcal{LR}_\infty(\kappa^\dagger)$. However, since LR_n^* remains of order $O_p^*(1)$, in probability, the bootstrap test is consistent.

We now turn to the case where the moment condition on Δx_t fails. Establishing the limiting ditribution in this case is extremely complicated, in particular because under lack of moments (in particular, second order moments), the bootstrap CLT no longer holds. Specifically, it is well known from Athreya (1987) and Knight (1989) that in this case the bootstrap delivers a random limiting distribution, as reported in the following example.

Example 8 (Bootstrap of the sample mean under infinite variance) *Suppose that the x_t's form an i.i.d. sequence in the domain of attraction of a Stable law with tail index denoted by $\nu \in (0,2)$. In this case it is well known that there are sequences a_n and b_n such that $S_n := a_n^{-1}\sum_{t=1}^{n}(x_t - b_n) \to_w S(\nu)$, a Stable random variable with tail index ν. Its i.i.d. bootstrap analog is given by $S_n^* := a_n^{-1}\sum_{t=1}^{n}(x_t^* - E^* x_t^*)$, where the*

x_t^*'s are (conditionally on the original data) i.i.d. from $\{x_1, ..., x_n\}$. Bootstrap validity would require that, in probability, $S_n^* \overset{w^*}{\to} S(\nu)$. However, this is not the case. As shown by Knight (1989), because of the lack of finite second order moments the large extremes in the original sample does not 'wash away' and, consequently, the cdf of the bootstrap statistic depends on the original data also asymptotically. Specifically, the following representation can be given to the cdf of S_n^* conditionally on the original data (see Knight, 1989)

$$P^*(S_n^* \leq x) \to_w P\left(\sum\nolimits_{i=1}^{\infty} \delta_i Z_i (M_i^* - 1) \leq x \middle| \delta_1, \delta_2, ..., Z_1, Z_2, ...\right) \tag{18}$$

where the δ_i's are i.i.d. random signs (with $P(\delta_i = 1) = p \in (0,1)$); for all $i = 1, 2, ...$, $Z_i := (e_1 + ... + e_i)^{1/\nu}$ where the e_i's form a sequence of i.i.d. exponential random variables (independent of the δ_i's) with mean 1; finally the M_i^*'s are i.i.d. Poisson with mean 1. Due to its dependence on the δ_i's and the Z_i's, the distribution on the right hand side of (18) is a random cdf, which is clearly different from the cdf of a $S(\nu)$ random variable. Moreover, should the bootstrap being based on the standardized residuals, i.e. based on $\tilde{S}_n^* := n^{-1/2} s_n^{-1/2} \sum_{t=1}^n (x_t^* - E^* x_t^*)$, with $s_n := E^*(x_t^* - E^* x_t^*)^2 = n^{-1} \sum_{t=1}^n (x_t - \bar{x}_n)^2$, it is straightforward that Knight's result extend to the following

$$P^*(\tilde{S}_n^* \leq x) \to_w P\left(\sum\nolimits_{i=1}^{\infty} \delta_i Z_i (M_i^* - 1) \leq x (\sum\nolimits_{i=1}^{\infty} Z_i^2) \middle| \delta_1, \delta_2, ..., Z_1, Z_2, ...\right)$$

which implies that $\tilde{S}_n = O_p^*(1)$, in probability. That is, althought the CLT does not hold on $\tilde{S}_n$, it is still of order $O_p^*(1)$ in probability. Extensions to other bootstraps and to (stationary and non-stationary) time series models are provided in Cavaliere, Georgiev and Taylor (2013, 2016, 2018) and extended to non-causal time series in Cavaliere, Nielsen and Rahbek (2018). □

In particular, as in the previous example it is reasonable to conjecture that the term $n^{-1/2} \sum_{t=1}^n z_t^*$ is of order $O_p^*(1)$, in probability. Hence, the central limit theorem does not hold on z_t^*; however, its sum is still of order $n^{1/2}$. This would suggest that the bootstrap LR statistic may have a random limiting distribution which, however, is bounded in probability.

4.3 A hybrid bootstrap

We here propose a bootstrap method which is able to mimic the null asymptotic distribution even if the null is false. This is simply a hybrid bootstrap, where we combine the use of the restricted parameter estimators (typically employed for the restricted bootstrap) with the use of the unrestricted residuals (typically employed for the unrestricted bootstrap). The hybrid bootstrap test statistic is defined through the following steps.

Hybrid (i.i.d.) Bootstrap:

(i) Estimate model (1) using Gaussian QML under the null hypothesis, yielding the estimates $\tilde{\theta}_n := (0, 0, \tilde{\omega}_n)'$; similarly, also estimate model (1) using Gaussian QML *without* imposing the null hypothesis, yielding the estimates $\hat{\theta}_n := (\hat{\pi}_n, \hat{\alpha}_n, \hat{\omega}_n)'$, together with the corresponding unrestricted QML residuals, $\hat{\varepsilon}_t := \Delta x_t - \hat{\pi}_n x_{t-1}$ and $\hat{z}_t := (\hat{\omega}_n + \hat{\alpha}_n x_{t-1}^2)^{-1/2} \hat{\varepsilon}_t$, as defined above;

(ii) Standardize the unrestricted residuals as

$$\hat{z}_{s,t} := \frac{\hat{z}_t - n^{-1}\sum_{t=1}^{n}\hat{z}_t}{\left(n^{-1}\sum_{t=1}^{n}(\hat{z}_t - n^{-1}\sum_{t=1}^{n}\hat{z}_t)^2\right)^{1/2}}$$

and construct the bootstrap innovations using the i.i.d. bootstrap re-sampling scheme; i.e., $z_t^* := \hat{z}_{s,\eta_t^*}$, where η_t^*, $t = 1, ..., n$ is an i.i.d. sequence of discrete uniform distributions on $\{1, 2, ..., n\}$;

(iii)-(v) As Steps (iii)-(v) of the restricted bootstrap.

This bootstrap is simple to implement and – with respect to the standard restricted bootstrap – it only requires unrestricted estimation of the model on the original data. Since this step is done one time only, implementation of this bootstrap is not more time consuming than the two bootstrap described earlier.

The crucial features of this bootstrap is that, due to the use of the unrestricted residuals, a bootstrap invariance principle for $(z_t^*, z_t^{*2} - 1)$ holds irrespectively of the null hypothesis to be true or not. Hence, the issue of possible lack of (fourth order) moments for z_t^* described in the previous Section 4.2 does not arise when this bootstrap is implemented. Moreover, the use of the restricted parameter estimates in the construction of the bootstrap sample allows to avoid possible randomness of the limiting bootstrap measures due to unit roots and a parameter on the boundary under the null hypothesis. We have the following theorem.

Theorem 6 *Let the conditions of Theorem 3 holds, and consider the hybrid bootstrap test statistic, LR_n^*. Then, both under H_0 and H_1, as $n \to \infty$:*

$$LR_n^* \xrightarrow{w^*}_p \mathcal{LR}_\infty(\kappa)$$

Remark 4.1 In principle, under the null hypothesis it is well expected that the restricted bootstrap delivers better size control than the hybrid version discussed here. This is a well-known property of bootstrap tests for a unit root; see e.g. Cavaliere and Taylor (2008, 2009*b*) and Palm et al. (2008) and the references therein. The amount of size accuracy which is lost by bootstrapping unrestricted residuals instead of plain restricted residuals is usually negligible. However, how the DAR structure affects these finite-sample properties of these two bootstrap schemes cannot be inferred from the proofs of (first-order) bootstrap validity. In the next section we aim at sheding some light on this issue by means of Monte Carlo simulation.

5 Simulations

In this section we compare the finite sample properties of the LR test for the pure random walk null hypothesis with its (asymptotically valid) bootstrap analogs: the restricted bootstrap LR test and the hybrid bootstrap test of section 4.3. By considering a detailed simulation study based on the DAR model, see (1), we aim at analyzing the finite-sample performance our the various bootstrap schemes across different choices of the bootstrap true values and different distributions of the innovations, both under the null and under the alternative hypothesis of (strict) stationarity.

The section is organized as follows. First, in Section 5.1 we describe (i) the model; (ii) the null hypothesis; (iii) the reference LR test and associated bootstrap test statistics. Finally, we describe the design of the Monte Carlo experiment. The empirical rejection probabilities [ERP] of the tests under the null hypothesis are investigated in Section 5.2. Section 5.3 is devoted to the analysis of the behaviour of the test when the null hypothesis is false. Here we investigate both raw and (pointwise) size-adjusted ERPs under the alternative hypothesis.

5.1 Monte Carlo design

We consider the DAR data generating process of (1),

$$\Delta x_t = \pi x_{t-1} + \varepsilon_t, \quad \varepsilon_t = \sigma_t z_t, \quad \sigma_t^2 = \omega + \alpha x_{t-1}^2,\ z_t \sim \text{i.i.d.}(0,1),\ (t = 1, ..., n) \tag{19}$$

for different choices of the distributions of the innovations z_t (in all cases, we consider distributions satisfying $\kappa < \infty$. The sample is initialized at $x_0 = 0$. The parameters are collected in the vector $\theta := (\pi, \alpha, \omega)'$ and, when necessary, the true values are denoted by the subscript '0'.

We consider three cases for the distribution of the innovations:

($\mathcal{E}_1$) z_t is a zero mean, unit variance Gaussian random variable;

($\mathcal{E}_2$) z_t is a standardized Student t random variable with ν degrees of freedom, i.e. z_t is distributed as $t(\nu)/\sqrt{\nu(\nu-2)}$, where $t(\nu)$ denoting a t random variable with $\nu \in \mathbb{R}^+$ degrees of freedom;

($\mathcal{E}_3$) z_t is a symmetric, standardized $\chi^2(1)$ random variable, i.e. z_t is distributed as $S(\chi_k^2 - k)/\sqrt{2k}$, with χ_k^2 denoting a χ^2 random variable with $k \in \mathbb{N}$ degrees of freedom and S is a Rademacher random variable (*i.e.*, a two-point distribution with $P(S = 1) = P(S = -1) = 1/2$).

Notice that for the (unimodal) distribution in $\mathcal{E}_2$, the moment of order m exists provideed $\nu > m$; moreover, for $\nu > 4$ the fourth-order moment (which appears in the asymptotic distribution of the LR test of Section 2.3) is given by $\frac{3(\nu-2)}{\nu-4}$. For $k \geq 2$, under $\mathcal{E}_3$ the distributions of the innovations is bimodal with modes at $\pm(k-2)^+$; moreover, all moments exist and in particular the fourth-order moment is given by $12/k + 3$. In the simulations, we force the t and the symmetric χ^2 distributions to have the same fourth-order moments, which requires setting $k = 2(\nu - 4)$; specifically, we set $\nu = 5.5$ and $k = 3$, which corresponds to $\kappa = 6$. In all cases, $\xi = 0$.

The null hypothesis is the pure random walk hypothesis $\mathsf{H}_0 : \pi = \alpha = 0$, see Section 2.1. We focus in particular on alternatives of the form $\pi < 0$ and $\alpha = 0$ (no unit root in the mean equation and no conditional heteroskedasticity) and on alternatives of the form $\pi = 0$ and $\alpha > 0$ (conditionally heteroskedastic strictly stationary with a unit root in the mean equation). In order to investigate local power, we consider these alternatives under Pitman drifts. Precisely, we first consider the sequence of (near unit root) local (pitman) alternatives

$$\mathsf{H}_1^{(\pi)} : \pi = c_\pi n^{-1},\ \alpha = 0 \tag{20}$$

with $c_\pi < 0$ fixed. For n fixed, this alternative lies in the region of the parameter space where the process is strictly stationary, conditionally homoskedastic and with finite fourth order moments. Moreover, we also consider the sequence of local (pitman) alternatives

$$\mathsf{H}_1^{(\alpha)} : \pi = 0,\ \alpha = c_\alpha n^{-3/2} \tag{21}$$

with $c_\alpha < 0$ fixed. For n fixed, even if $\pi = 0$ (such that the mean equation has a unit root, i.e. $\Delta x_t = \varepsilon_t$) this alternative lies in the region of the parameter space where the process displays (volatility-induced) strictly stationary, is conditionally heteroskedastic, but does not possess finite second order moments. The rate in (21) implies that the (normalized) score associated to the DAR Gaussian likelihood does not diverges under the alternative, and still makes x_t of order $n^{1/2}$.

Restricted and unrestricted estimation and associated LR tests are based on the Gaussian likelihood associated to (19), with x_0 considered fixed in the statistical analysis. Maximization of the likelihood function imposes the non-negativity constraints $\alpha \geq 0$ and $\omega > 0$.[4] The (asymptotic) LR test is based on asymptotic critical values obtained numerically by discretizing the distribution in (3) over 100,000 steps and using 100,000 Monte Carlo repetitions (these do not substantially differ from those reported in Table 1 of Kluppelberg *et al.*, 2002) under the assumption that $\kappa = 2$ and $\xi = 0$; hence, the asymptotic case is not expected to be correctly sized, even in large samples, when the actual distribution of z_t departs from the Gaussian distribution.

We consider the two (asymptotically valid) bootstrap schemes introduced earlier in the paper. First, the plain restricted bootstrap of section 3.1, which is based on resampling the residuals from restricted estimation and impose the null hypothesis on the bootstrap generating process. Second, the hybrid bootstrap scheme, which employs the residuals from unrestricted parameter estimation but still imposes the null on the bootstrap sample.

Throughout, we use $10,000$ Monte Carlo replications and use $B = 399$ bootstrap repetitions. Sample of size $n \in \{50, 100, 200, 500\}$ are considered throughout. All tests are run at the nominal 1%, 5% and 10% significance levels.

5.2 Empirical rejection probabilities under the null

Table 1 reports the empirical rejection probabilities (as estimated on the $10,000$ Monte Carlo replications) under the null hypothesis, $\mathsf{H}_0 : \alpha = \pi = 0$, for the three distributions for the innovations.

[Table 1 about here]

The following points can be made out of the analysis.

For the leading case of Gaussian errors, the asymptotic LR test tends to be undersized for samples of size $n \in \{50, 100, 200\}$. For $n = 500$, the ERPs are closer to the nominal level, although at the 10% level the test appears to be slightly oversized. In contrast, both the restricted bootstrap and the hybrid bootstrap tests show excellent size control for samples of $n \in \{50, 100, 200\}$, with ERPs very close to the corresponding

[4] All computations are performed in Matlab R2018*b* using the 'fmincon' constrained optimization routine. Code is available upon request.

nominal levels. The bootstrap tests do not seem to dominate the asymptotic test in terms of size when $n = 500$.

For t-distributed errors, the asymptotic LR test is significantly oversized. This is expected, since this test is implicitly based on the (false) assumption that the errors are Gaussian. The bootstrap tests show very good size control, with the restricted bootstrap being slightly more accurate than the hybrid bootstrap, as expected; see the discussion in section 3.1.

For the bimodal χ^2-type errors, the asymptotic tests is again substantially unreliable. For instance, when the nominal level is 1% and $n = 500$, the ERP equals 5.5%. The bootstrap seems to fix this problem very well, again with ERPs very close to the corresponding nominal levels at all the sample sizes considered. Again, the restricted bootstrap seems to marginally outperform the hybrid bootstrap.

In summary, the performance of the bootstrap tests is largely satisfactory. Not only the bootstrap allows to circumvent the non-pivotality of the asymptotic test, whose distribution depends on the unknown parameter κ, but it also delivers an excellent control of the ERP when the null hypothesis holds true.

5.3 Empirical rejection probabilities under local alternatives

We now turn to the inspection of the ERPs of the (asymptotic and bootstrap) tests when the null hypothesis does not hold. To this aim, we first consider alternatives satisfying $\pi < 0$ and $\alpha = 0$ (pure homoskedastic autoregressions) in Section 5.3.1. Alternatives satisfying $\pi = 0$ and $\alpha > 0$ (heteroskedastic processes with a unit root in the mean equation) are considered in Section 5.3.2. Throughout this section we present both raw ERPs and (pointwise) size-adjusted rejection probabilities. To compute the latter, as suggested in Cavaliere *et al.* (2015) for each given point in the parameter space, we first perform the simulation under the null and record the nominal level that would have given an ERP equal to the desired significance level. Next, we use this adjusted nominal level in the simulations under the alternative hypothesis. Let, for instance, p_n^* denote the p-value of the bootstrap test, and let $p_0(\eta) := P(p_n^* \leq \eta | \mathsf{H}_0)$, with η denoting the choosen significance level. Then, the size-adjusted bootstrap test at the $100\eta\%$ level corresponds to rejecting H_0 when $p_n^* \leq \tilde{\eta}$, where $\tilde{\eta}$ is such that $p(\tilde{\eta}) := P(p_n^* \leq \tilde{\eta} | \mathsf{H}_0) = \eta$.

5.3.1 Pure autoregressive alternatives

In Figure 1 we report the size-adjusted power of the asymptotic test as well as of the two bootstrap tests, under the assumption of Gaussian shocks. We set $-\pi \in [0, 0.2]$ and $\alpha = 0$ (with $\pi = 0$ clearly corresponding to the null hypothesis), for samples of sizes $n = 50$ and $n = 100$.

[Figure 1]

From the inspection of Figure 1 it is clear that the bootstrap tests are not outperformed, in terms of size-adjusted ERPs, by the power function of the LR test. In this respect, implementation of the bootstrap does not entail a power loss relatively to the asymptotic tests, as predicted by the theory. If anything, the restricted bootstrap, at the smallest nominal levels, seems to be slightly more powerful than the asymptotic test. While this fact may appear surprising and may depend on the chosen Monte Carlo design, similar evidence has already been documented in the literature (see, Davidson and MacKinnon, 2002, Figure 14). Moreover, in terms of theory there is no result that prevents this from happening (see, e.g, Davidson and Mackinnon, 2006). In general, there are no significant power gains from implementing the hybrid bootstrap over the plain restricted bootstrap – however, in the next section we provide a case where this actually is not the case.

We now turn to the analysis of the local power of the tests under this alternative. Here we consider samples of size $n \in \{50, 100, 200, 500\}$ and all the three distributions described earlier. The local power functions for the alternative $\pi = cn^{-1}$ and $\alpha = 0$ are reported in Table 2 for $c = -10$. For completeness, we also report the raw ERPs in Table 3. Obviously, these ERPs are affected by the deviations of the actual size of the tests from the corresponding nominal levels, see Table 1.

[Tables 2 and 3 about here]

For Gaussian errors, the results are in line with those represented in Figure 1. The two bootstrap tests are not outperformed by the (size-adjusted) asymptotic LR test. Interestingly, at the smallest sample sizes the restricted bootstrap experiences some power gains over the other two tests. For samples of size 500, there are no discernible differences between the three tests.

For t-distributed errors, at nominal significance levels of 10% the restricted bootstrap test behaves very similarly to the asymptotic test. At smaller nominal levels, however, there seems to be some power gains stemming from the implementation of the restricted bootstrap test. The hybrid bootstrap seems somehow less powerful than the restricted bootstrap. A similar pattern can be observed for the case of symmetrically χ^2-distributed errors. Here, again, the restricted bootstrap test seems to be preferrable.

Interestingly, although the LR test is asymptotically valid for all the alternatives considered here, its (local) power function seem to be affected by deviations of the errors' distribution from the Gaussian distribution used to construct the (pseudo) likelihood function.

In summary, the restricted bootstrap tests display power which is not inferior (sometimes even superior) to the power of the corresponding asymptotic test. Moreover, implementation of the hybrid bootstrap does not seem to provide power gains.

5.3.2 Heteroskedastic, unit root alternatives

Results for alternatives satisfying $\pi = 0$ and $\alpha > 0$ (strictly stationary heteroskedastic DAR with a unit root in the mean equation) are reported in Figure 2. Here we let $\alpha \in [0, 0.2]$ (with $\alpha = 0$ corresponding to the null hypothesis) and, again, we consider samples of size $n = 50$ and $n = 100$. As for the previous case, the bootstrap tests do not display substantial differences in terms of power from the asymptotic test. The only exception is for the 1% nominal level and $n = 50$: here, while the hybrid bootstrap has the same power function of the asymptotic test, the restricted bootstrap shows a power loss with respect to the asymptotic test. This is somehow expected from the discussion in Section 4.2: under the alternative considered here, which implies that the data are generated by a strictly stationary, *infinite variance* process, the restricted bootstrap is based on resampling residuals which, even in large samples, do not possess finite second order moment (and therefore the bootstrap CLT may not apply). Hence, while the bootstrap statistic may in fact be of $O_p^*(1)$, in probability, it does not replicate the same distribution of the asymptotic test. From the results in Figure 2 it seems that this is indeed the case – the distribution of the restricted bootstrap statistic is shifted to the left (at least in terms of the first percentile) with respect to the distribution of the hybrid bootstrap test. For larger sample sizes or more liberal nominal levels, however, this does not seem to be an issue, as the power functions of the two bootstrap tests and of the asymptotic tests do not significantly differ.

[Figure 2 about here]

We now turn to the local power analysis. As for the previous alternative, we consider samples of size $n \in \{50, 100, 200, 500\}$ and all the three distributions described earlier. The local power functions for the alternative $\pi = 0$ and $\alpha = cn^{-3/2}$ are reported for $c = 10$ in Table 4; the corresponding raw ERPs are reported in 5.

[Tables 4 and 5 about here]

In terms of this local power analysis, the hybrid bootstrap seems to be slightly preferrable to the restricted bootstrap in terms of power for the smallest sample sizes, especially for the case of Gaussian errors. For the cases of t-distributed errors and of symmetric χ^2 distributed errors, the difference between the two bootstrap seems to be marginal.

In summary, the bootstrap tests display power which is generally not inferior (and sometimes superior) to the power of the corresponding asymptotic test. The implementation of the restricted bootstrap seems to provide the best performance not only in terms of size, but also in terms of size adjusted power. The hybrid bootstrap may provide some power gains when the first differenced process do not possess finite variance, but such power gains do not make this bootstrap preferrable to the plain, restricted bootstrap.

6 Extension to asymmetric innovations

One of the assumptions in Ling (2004) and that we have assumed so far is that the third order moment of the innovations, $\xi = Ez_t^3$, equals zero. This condition ensures that the two Brownian motions characterizing the asymptotic distribution in (3) are independent.

If this moment condition fails to hold, the limiting distribution of LR_n can no longer be expressed as the (weighted) sum of a squared Dickey-Fuller and a half-χ_1^2

independent random variables, see Remark 2.6. More precisely in Theorem 3, as shown in Klüppelberg et al (2002, Theorem 3.1), the second term in the expression for the $\mathcal{LR}_\infty(\kappa)$ in (3) for general ξ is given by:

$$\frac{1}{2}\max\left(0,\left[\xi\int B^2\mathsf{d}B+\sqrt{\kappa-\xi^2}\int B^2\mathsf{d}W-\int B^2\mathsf{d}u\left\{\xi B_1+\sqrt{\kappa-\xi^2}W_1\right\}\right]\right)^2 \quad (22)$$
$$\times\left[\int B_u^4\mathsf{d}u-\left(\int B_u^2\mathsf{d}u\right)^2\right]^{-1}$$

where, as before, B and W are independent standard Brownian motions.

Interestingly, the bootstrap may take care of this non-pivotality and we can establish the following result.

Theorem 7 *The results of Theorem 4 and Theorem 6 hold independently of whether* $\xi = 0$ *or not.*

For the restricted bootstrap, where the z_t^*'s are based on the restricted (standardized) residuals $\tilde{z}_{s,t}$, a key insight is the following. It holds that the bootstrap (conditional) third order moment ξ_n^*, is given by,

$$\xi_n^* = E^*\left(z_t^{*3}\right) = \frac{1}{n}\sum_{t=1}^{n}\tilde{z}_t^3,$$

such that, under suitable moment restrictions on the $\{z_t\}$ sequence, $\xi_n^* \rightarrow_p \xi$. This implies, under some additional algebra, that $\mathcal{Z}_n^*(\cdot)$ of (7) satisfies in this more general setting,

$$\mathcal{Z}_n^*(\cdot) := n^{-1/2}\textstyle\sum_{i=1}^{[n\cdot]}\left(z_t^*, z_t^{*2}-1\right)' \xrightarrow{w^*}_p \begin{pmatrix} 1 & 0 \\ \xi & \sqrt{\kappa-\xi^2}\end{pmatrix}\begin{pmatrix} B_\cdot^* \\ W_\cdot{}^* \end{pmatrix},$$

see Appendix A.5.1, eq.(A.40). Hence, the bootstrap mimics the asymptotic distributional properties of the original statistics even if $\xi \neq 0$.

7 Conclusions

In this paper we have discussed several issues which may arise in the implementation of the bootstrap hypothesis testing to time series econometric models. Essentially, these are related to the assessment of bootstrap validity under the null hypothesis (*i.e.*, establishing that the bootstrap mimics the correct limiting distribution of the original test statistic under the null hypothesis) as well as to the behaviour of the bootstrap statistic under the alternative hypothesis.

Our discussion has focused on the double-autoregressive, or DAR, model, where the time series properties of the data – such as strict strationarity or the existence of moments – are determined through a very delicate balance between the parameters of conditional mean and the conditional variance equations.

Focusing on tests of the null hypothesis of non-stationarity, *i.e.* reduction to the pure random walk, we have initially shown that – due to the possible presence of unit roots and of parameters on the boundary of the parameter space a classic – unrestricted bootstrap fails to mimic the null distribution under the null. Conversely, the restricted bootstrap works, irrespectively of a parameter of the conditional variance equation being on the boundary of the parameter space under the null hypothesis.

Next, we have discussed the possible issues which may arise under the alternative. Here, the crucial issue is that, under the alternative, the data may have have infinite variance. Hence, the restricted bootstrap, based on re-sampling the residuals with the null imposed, may in fact be based on re-sampling an infinite variance sequence. As a consequence, the bootstrap statistic may have a *random* limiting distribution which may lead to a lack of power over the infeasible size-adjusted asymptotic test. This observation is the basis of our next suggestion, which is a *hybrid* implementation of the bootstrap where the parameters used to generate the bootstrap sample are based on restricted estimation while the residuals used to construct the bootstrap shocks are based on unrestricted estimation.

Although most of our analysis is based on the DAR model, most of these issues are common to the great majority of econometric models. Hence, a thorough investigation of the properties of the bootstrap under the null and under the alternative are always required before its practical implementation.

There are further issues which have not been touched in this paper but may as well be important to establish bootstrap validity. For instance, in our testing example the parameters of the model are (up to an intercept) all restricted by the null hypothesis. In most cases, however, where the null hypothesis restricts only a subset of the parameters. For instance, when testing hypothesis of the cointegration rank in a VAR model, estimation with the wrong rank imposed may lead to a bootstrap sample which is not I(1), see the discussion in Cavaliere *et al.* (2012). Another case is testing if a parameter is on the boundary of the parameter space when the remaining parameters might be on the boundary as well (see Cavaliere, Nielsen, Pedersen and Rahbek, 2018). In this case the limiting distribution of the bootstrap statistic depends on the asymptotic properties of the estimators used to generate the bootstrap data. In both examples, establishing bootstrap validity requires to determine to what pseudo-true values such estimators converge to, at what speed, and what are the implications on the properties of the bootstrap sample.

References

Andrews, D.W.K. (1999): "Estimation when a Parameter is on a Boundary", *Econometrica*, 67(6), 1341–1383.

—— (2000): "Inconsistency of the Bootstrap when a Parameter is on the Boundary of the Parameter Space", *Econometrica*, 68(2), 399–405.

—— (2001): "Testing when a Parameter is on the Boundary of the Maintained Hypothesis", *Econometrica*, 69(3), 683–734.

Andrews, D.W.K. and M. Buchinsky (2000), "A three-step method for choosing the number of bootstrap repetitions", *Econometrica* 68, 23–51.

Athreya K.B. (1987): "Bootstrap of the Mean in the Infinite Variance Case", *Annals of Statistics*, 15, 724–731.

Basawa, I.V., A.K. Mallik, W.P. McCormick, J.H. Reeves, and R.L. Taylor (1991): "Bootstrapping Unstable First-Order Autoregressive Processes", *Annals of Statistics*, 19, 1098–1101.

Borkovec M. and C. Klüppelberg (2001): "The tail of the stationary distribution of an autoregressive process with ARCH(1) errors", *Annals of Applied Probability*, 11, 1220–1241.

Brown, B.M. (1971): Martingale central limit theorems, *Annals of Mathematical* Statistics 42, 59–66.

Cavaliere G., H.B. Nielsen and A. Rahbek (2015): "Bootstrap Testing of Hypotheses on Co-Integration Relations in Vector Autoregressive Models", *Econometrica,* 83(2), 813–831.

—— (2017): "On the Consistency of Bootstrap Testing for a Parameter on the Boundary of the Parameter Space", *Journal of Time Series Analysis*, 38(4), 513–534.

Cavaliere G., H.B. Nielsen, R.S. Pedersen and A. Rahbek (2018): "Bootstrap inference on the boundary of the parameter space with application to conditional volatility models", Discussion Papers 18–10, University of Copenhagen, Department of Economics.

Cavaliere, G., A. Rahbek and A.M.R. Taylor (2010a): "Co-integration rank testing under conditional heteroskedasticity", *Econometric Theory* 26, 1719–1760.

—— (2010b): "Testing for co-integration in vector autoregressions with non-stationary volatility", *Journal of Econometrics* 158, 7–24.

—— (2012): "Bootstrap determination of the co-integration rank in vector autoregressive models", *Econometrica* 80, 1721–1740.

—— (2014): "Bootstrap determination of the co-integration rank in heteroskedastic VAR models", *Econometric Reviews*, 33, 606–650

Cavaliere G. and I. Georgiev (2018): "Inference under random limit bootstrap measures", mimeo.

Cavaliere G. and A.M.R. Taylor (2008): "Bootstrap unit root tests for time series with non-stationary volatility", *Econometric Theory* 24, 43-71.

—— (2009a), "Bootstrap M Unit Root Tests", *Econometric Reviews*, 28, 393–421.

—— (2008): "Heteroskedastic time series with a unit root", *Econometric Theory,* 25, 1228–1276.

—— (2007): "Testing for unit roots in time series models with non-stationary volatility", *Journal of Econometrics* 140, 919–947.

Chan N.H. and C.Z. Wei (1987): "Asymptotic inference for nearly non-stationary AR(1) processes", *Annals of Statistics*, 15, 1050–1063.

Chen M., D. Li and S. Ling (2013): "Non-stationarity and quasi-maximum likelihood estimation on a double autoregressive model", *Journal of Time Series Analysis*, 35, 189–202.

Davidson, R. and J.G. MacKinnon (2000): "Bootstrap tests: How many bootstraps?", *Econometric Reviews* 19, 55–68.

—— (2006): "Bootstrap methods in econometrics", Chapter 23 in *Palgrave Handbooks of Econometrics*: Volume 1 Econometric Theory, ed. T. C. Mills and K. D. Patterson, Basingstoke, Palgrave Macmillan, 812–838.

—— (2002): "Bootstrap J tests of nonnested linear regression models", *Journal of Econometrics* 109, 167–193.

—— (2002): "The power of bootstrap and asymptotic tests", *Journal of Econometrics* 133, 421–441.

Davison, A. C., D. V. Hinkley and G.A. Young (2003): "Recent developments in bootstrap methodology", *Statistical Science* 18, 2, 141–157.

Deo, R.S. (2000): "Spectral tests of the martingale hypothesis under conditional heteroskedasticity", *Journal of Econometrics* 99, 291–315.

Engle, R.F. and C.W.J. Granger (1987): "Co-integration and error correction: Representation, estimation and testing", *Econometrica* 55, 251-276.

Fachin, S. (2000): "Bootstrap and Asymptotic Tests of Long-Run Relationships in Cointegrated Systems", *Oxford Bulletin of Economics and Statistics*, 62, 543–551.

Fachin S. and P. Omtzigt (2006): "Boostrapping and Bartlett correction in the cointegrated VAR model", *Econometric Reviews* 25, 41–60.

Giné, E. and J. Zinn (1990): "Bootstrapping general empirical measures", *Annals of Probability* 18, 851–869.

Gonçalves, S. and L. Kilian (2004): "Bootstrapping autoregressions with conditional heteroskedasticity of unknown form", *Journal of Econometrics* 123, 89–120.

Gredenhoff, M. and T. Jacobson (2001): "Bootstrap Testing Linear Restrictions on Cointegrating Vectors", *Journal of Business & Economic Statistics*, 19, 63–72.

Hall, P. and C.C. Heyde (1980): *Martingale Limit Theory and its Application.* Academic Press, New York.

Hall P. (1992), *The bootstrap and Edgeworth expansions*, New York: Springer-Verlag.

Harvey D., S.J. Leybourne, R.Sollis, A.M.R. Taylor (2016): "Tests for explosive financial bubbles in the presence of non-stationary volatility", *Journal of Empirical Finance* 38(B), 548–574

Jensen, S.T. and A. Rahbek (2004): "Asymptotic normality of the QMLE estimator of ARCH in the nonstationary case", *Econometrica*, 72, 641–646.

Johansen, S. (1996), *Likelihood-based inference in cointegrated vector autoregressive models*, Oxford: Oxford University Press.

Knight K. (1989): "On the bootstrap of the mean in the infinite variance case", *Annals of Statistics*, 17, 1168–1175.

Klüppelberg C., R.A. Maller, M. Van de Vyver and D. Wee (2002): "Testing for reduction to random walk in autoregressive conditional heteroskedasticity models", *Econometrics Journal*, 5, 387–416.

Kristensen, D. and A. Rahbek (2013): "Testing and inference in nonlinear cointegrating vector error-correction models", *Econometric Theory*, 29, 1238–1288

Lange, T., A. Rahbek and S.T.Jensen (2011): "Estimation and asymptotic inference in the AR-ARCH model", *Econometric Reviews*, 30, 129–153.

Ling S. (2004): "Estimation and testing stationarity for double-autoregressive models", *Journal of the Royal Statistical Society B*, 66, 63–78.

—— (2007a): "A double AR(p) model: Structure and estimation", *Statistica Sinica* 17, 161–175.

—— (2007b): "Self-weighted and local quasi-maximum likelihood estimators for ARMA-GARCH/IGARCH models", *Journal of Econometrics* 140, 849–873.

Ling, S. and M. McAleer (2003), "Asymptotic theory for a vector ARMA-GARCH model", *Econometric Theory* 19, 280–310.

Ling, S. and W.K. Li (1998): "Limiting distributions of maximum likelihood estimators for unstable ARMA models with GARCH errors", *Annals of Statistics* 26, 84–12.

Liu, R.Y. (1988): "Bootstrap procedures under some non i.i.d. models", *Annals of Statistics* 16, 1696–1708.

Mammen, E. (1993): "Bootstrap and wild bootstrap for high dimensional linear models", *Annals of Statistics* 21, 255–285.

MacKinnon, J.G. (2009): "Bootstrap hypothesis testing", in *Handbook of Computational Econometrics*, ed. David A. Belsley and John Kontoghiorghes, Chichester, Wiley, 183–213.

Nielsen, H.B. and A. Rahbek (2014): "Unit root vector autoregression with volatility induced stationarity", *Journal of Empirical Finance*, 29, 144–167.

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

Phillips, P.C.B., Y. Wu, Y. and J. Yu (2011): "Explosive behavior in the 1990s Nasdaq: when did exuberance escalate stock values?" *International Economic Review* 52, 201–226.

Politis, D., J. Romano and M. Wolf (1999): *Subsampling*, Springer, Berlin.

Swensen, A.R. (2003): "Bootstrapping unit root tests for integrated processes", *Journal of Time Series Analysis, 24*, 99–126.

Swensen, A.R. (2006): "Bootstrap algorithms for testing and determining the cointegration rank in VAR models", *Econometrica* 74, 1699–1714.

Wu, C.F.J. (1986): "Jackknife, bootstrap, and other resampling methods", *Annals of Statistics* 14, 1261–1295.

A Mathematical Appendix

A.1 Introduction

This appendix contains the proofs of the theory for the bootstrap implementation in the DAR model for testing the null of non-stationarity, that is $\mathsf{H}_0 : \pi = \alpha = 0$.

In Appendix A.2 and A.3, we first establish new asymptotic (non-bootstrap) results for the QMLE $\hat{\theta}_n := (\hat{\pi}_n, \hat{\alpha}_n, \hat{\omega}_n)'$ under both stationarity as well as under the null H_0 of non-stationarity, see Theorem 1 and Theorem 2, respectively. Appendix A.3 additionally provides asymptotic theory for the LR_n statistic under H_0, see Theorem 3. The asymptotic results for the QMLE, as well as LR_n, are then applied in Appendix A.5, where asymptotic results for the (restricted and hybrid) bootstrap variants LR_n^* of LR_n are derived.

As to the general (nonstandard) likelihood theory, recall that the parameter (or, optimization) set for the DAR model is given by

$$\mathcal{T} := \{\theta = (\pi, \alpha, \omega)' : \pi \in [-\pi_L, \pi_U], \alpha \in [0, \alpha_U] \text{ and } \omega \in [\omega_L, \omega_U]\},$$

and, for estimation with the null hypothesis imposed, by $\mathcal{T}_0 := \{\theta = (\pi, \alpha, \omega)' : \pi = \alpha = 0 \text{ and } \omega \in [\omega_L, \omega_U]\}$. As $\alpha \geq 0$, inference and testing is nonstandard and we apply theory from Andrews (1999, 2001) which treat estimation and testing under inequality constraints (and more general boundary issues), see also Vu and Zhou (1997), Klüppelberg et al. (2002) and Cavaliere, Nielsen and Rahbek (2017). Thus, the asymptotic distributions of the QMLE $\hat{\theta}_n$ and the associated LR_n statistic follow by verifying regularity conditions for *(i)* the parameter spaces $\mathcal{T}$ and $\mathcal{T}_0$; *(ii)* consistency of $\hat{\theta}_n$; and, *(iii)* convergence of the score, information and third derivative of the log-likelihood function. For the bootstrap asymptotic theory, we verify the analogous regularity conditions for the bootstrap log-likelihood quantities, applying convergence (weakly, and in probability) conditional on the data, see *e.g.* Cavaliere, Nielsen and Rahbek (2015, 2017) and Cavaliere, Rahbek and Taylor (2012).

As to *(i)*, consider first the stationary case, where the true parameter $\theta_0 \in \Theta_{\mathcal{S}}$, with $\theta_0 = (\pi_0, \alpha_0, \omega_0)'$. In this case, $\mathcal{T} - \theta_0$, in the sense of Andrews (1999, 2001), is locally equal to the cone(s),

$$\Lambda(A) = \mathbb{R} \times A \times \mathbb{R}, \tag{A.1}$$

where $A = \mathbb{R}$ if $\alpha_0 > 0$, and $A = \mathbb{R}^+$ if $\alpha_0 = 0$, such that Assumption $5^{2*(b)}$ in Andrews (1999) holds with $B_T = n^{1/2}$. For the non-stationary case, where $\theta_0 \in \Theta_{\mathcal{N}}$, then $\mathcal{T} - \theta_0$ and $\mathcal{T}_0 - \theta_0$ are locally equal to the cones $\Lambda := \Lambda\left(\mathbb{R}^+\right)$ and

$$\Lambda_0 := \{0\} \times \{0\} \times \mathbb{R}, \tag{A.2}$$

respectively. That is, with $B_T := G_n := \text{diag}\left(n, n^{3/2}, n^{1/2}\right)$ in the non-stationary case, Assumption $5^{2*(b)}$ in Andrews (1999) holds.

With respect to *(ii)*, the regularity conditions verified under *(iii)* imply, with probability tending to one, that $\hat{\theta}_n \to_p \theta_0$. As to *(iii)*, note that we verify suitable bounds on the third-order log-likelihood derivative(s), rather than, as is standard, establish uniform convergence of the information (that is, the second order log-likelihood derivative); see Jensen and Rahbek (2004, Lemma 1) and Kristensen and Rahbek (2010, Lemmas 11 and 12) for general asymptotic likelihood theory in the stationary and non-stationary cases respectively.

Finally note that while the results quoted in Theorems 1 and 2, and Theorem 3, are for the case of the nuisance (asymmetry) parameter $\xi = 0$, the results are derived in the next under the general assumption of $\xi \neq 0$ as needed for the discussion in Section 6 where we extend the asymptotic (and bootstrap) theory to address also the nuisance parameter ξ (in addition to κ).

A.2 QMLE under stationarity – Proof of Theorem 1

In this section we derive the asymptotic theory for the QMLE $\hat{\theta}_n = (\hat{\pi}_n, \hat{\alpha}_n, \hat{\omega}_n)'$ in Theorem 1 for the stationary case where $\theta_0 \in \Theta_{\mathcal{S}}$. We verify conditions (A.1)–(A.3) in Jensen and Rahbek (2004, Lemma 1) [JR hereafter] which imply, with probability tending to one, that $\hat{\theta}_n \to_p \theta_0$. Conditions (A.1) and (A.2), that is convergence of the score and information, are detailed below, while condition (A.3) for the third order derivative follows as for the proof of establishing condition (C.ii) in Section A.6 for the non-stationary case.

A.2.1 Score and observed information

In terms of the log-likelihood function $L_n(\theta) = \sum_{t=1}^n l_t(\theta)$, define the score quantities,

$$S_n(\theta) = \sum_{t=1}^n s_t = \sum_{t=1}^n \partial l_t(\theta)/\partial\theta \text{ and } S_n = S_n(\theta)|_{\theta=\theta_0}. \tag{A.3}$$

Likewise, the observed information is given by

$$I_n(\theta) = \sum_{t=1}^n i_t = \sum_{t=1}^n (-\partial l_t(\theta)/\partial\theta\partial\theta') \text{ and } I_n = I_n(\theta)|_{\theta=\theta_0}. \tag{A.4}$$

The terms in score $S_n(\theta)$ are given by

$$\begin{aligned} s_t' &= \left(s_t^\pi, s_t^\alpha, s_t^\omega\right) \\ &= \left(\varepsilon_t x_{t-1}/\sigma_t^2, \tfrac{1}{2}\left(\varepsilon_t^2/\sigma_t^2 - 1\right)x_{t-1}^2/\sigma_t^2, \tfrac{1}{2}\left(\varepsilon_t^2/\sigma_t^2 - 1\right)/\sigma_t^2\right). \end{aligned} \tag{A.5}$$

At the the true value θ_0, the score is (the sum of) a martingale differences (MGD) sequence,

$$s'_{t,0} := s'_t\big|_{\theta=\theta_0} = \left(z_t v_{t-1}, \tfrac{1}{2}\left(z_t^2-1\right)v_{t-1}^2, \tfrac{1}{2}\left(z_t^2-1\right)/\sigma_t^2\right), \tag{A.6}$$

with $v_{t-1} := x_{t-1}/\sigma_t$.

The terms i_t of the observed information are given by

$$\begin{aligned} i_t &= \begin{pmatrix} i_t^{\pi\pi} & i_t^{\pi\alpha} & i_t^{\pi\omega} \\ i_t^{\alpha\pi} & i_t^{\alpha\alpha} & i_t^{\alpha\omega} \\ i_t^{\pi\omega} & i_t^{\alpha\omega} & i_t^{\omega\omega} \end{pmatrix} \qquad\qquad (\text{A.7}) \\ &= \begin{pmatrix} v_{t-1}^2 & \left(\varepsilon_t/\sigma_t\right)v_{t-1}^3 & \varepsilon_t v_{t-1}/\sigma_t^3 \\ \varepsilon_t/\sigma_t^2 v_{t-1}^3 & \left(\varepsilon_t^2/\sigma_t^2-1/2\right)v_{t-1}^4 & \left(\varepsilon_t^2/\sigma_t^2-1/2\right)v_{t-1}^2/\sigma_t^2 \\ \varepsilon_t v_{t-1}/\sigma_t^3 & \left(\varepsilon_t^2/\sigma_t^2-1/2\right)v_{t-1}^2/\sigma_t^2 & \left(\varepsilon_t^2/\sigma_t^2-1/2\right)/\sigma_t^4 \end{pmatrix} \end{aligned}$$

which at the true value θ_0 reduces to

$$i_{t,0} := i_t\big|_{\theta=\theta_0} = \begin{pmatrix} v_{t-1}^2 & z_t v_{t-1}^3 & z_t v_{t-1}/\sigma_t^2 \\ z_t v_{t-1}^3 & \left(z_t^2-1/2\right)v_{t-1}^4 & \left(z_t^2-1/2\right)v_{t-1}^2/\sigma_t^2 \\ z_t v_{t-1}/\sigma_t^2 & \left(z_t^2-1/2\right)v_{t-1}^2/\sigma_t^2 & \left(z_t^2-1/2\right)/\sigma_t^4 \end{pmatrix}. \tag{A.8}$$

A.2.2 Asymptotics for the Score and the Hessian – Proofs of Conditions (A.1) and (A.2) in JR

Note initially that, by $\kappa < \infty$, standard application of central limit theory for i.i.d. variables gives

$$n^{-1/2}\sum_{t=1}^{n}\left(z_t, z_t^2-1\right) \overset{w}{\to} V, \qquad \mathrm{Var}\left(V\right) = \begin{pmatrix} 1 & \xi \\ \xi & \kappa \end{pmatrix}. \tag{A.9}$$

Next, the MGD representation of the score $s_{t,0}$,

$$s'_{t,0} = \left(z_t, z_t^2-1\right)\begin{pmatrix} v_{t-1} & 0 & 0 \\ 0 & \frac{1}{2}v_{t-1}^2 & \frac{1}{2\sigma_t^2} \end{pmatrix},$$

together with (A.9), implies by standard arguments that condition (A.1) holds, *i.e.*:

$$n^{-1/2}S_n\left(\theta_0\right) = n^{-1/2}\sum_{t=1}^{n} s_{t,0} \overset{w}{\to} \mathcal{S}_\infty := \left(\mathcal{S}_\pi, \mathcal{S}'_\gamma\right)', \quad \mathcal{S}_\gamma := \left(\mathcal{S}_\alpha, \mathcal{S}_\omega\right)'.$$

Here $\mathcal{S}_\infty$ is Gaussian with covariance matrix

$$\Omega_{\mathcal{S}} := \begin{pmatrix} \Omega_{\mathcal{S},\pi\pi} & \Omega'_{\mathcal{S},\gamma\pi} \\ \Omega_{\mathcal{S},\gamma\pi} & \Omega_{\mathcal{S},\gamma\gamma} \end{pmatrix}, \tag{A.10}$$

where

$$\Omega_{\mathcal{S},\pi\pi} = E\left(x_{t-1}^2/\sigma_t^2\right), \qquad \Omega'_{\mathcal{S},\gamma\pi} = \begin{pmatrix} \frac{\xi}{2}E\left(x_{t-1}^3/\sigma_t^3\right) & \frac{\xi}{2}E\left(x_{t-1}/\sigma_t^3\right) \end{pmatrix}$$

and

$$\Omega_{\mathcal{S},\gamma\gamma} = \frac{\kappa}{4}\begin{pmatrix} E\left(x_{t-1}^4/\sigma_t^4\right) & E\left(x_{t-1}^2/\sigma_t^4\right) \\ E\left(x_{t-1}^2/\sigma_t^4\right) & E\left(1/\sigma_t^4\right) \end{pmatrix}.$$

Note that on the one hand for $\alpha_0 > 0$, it follows that $E\left(x_{t-1}^4/\sigma_t^4\right) < \infty$ under stationarity of x_t as in JR. If, on the other hand, $\alpha_0 = 0$, then $Ex_t^4 < \infty$ is implied by $\kappa < \infty$. Moreover, for $\alpha_0 = 0$ and denoting $\Omega_{\mathcal{S}}$ under $\alpha_0 = 0$ by $\Omega_{\mathcal{S}}^0$, the covariance $\Omega_{\mathcal{S}}$ simplifies to the following

$$\begin{aligned}\Omega^0_{\mathcal{S},\pi\pi} &= \frac{1}{\omega_0}E\left(x_{t-1}^2\right), \ \Omega^{0\prime}_{\mathcal{S},\gamma\pi} = \left(\tfrac{\xi}{2\omega_0^{3/2}}E\left(x_{t-1}^3\right)\ , 0\right) \ \text{ and} \\ \Omega^0_{\mathcal{S},\gamma\gamma} &= \frac{\kappa}{4\omega_0^2}\begin{pmatrix} E\left(x_{t-1}^4\right) & E\left(x_{t-1}^2\right) \\ E\left(x_{t-1}^2\right) & 1\end{pmatrix}.\end{aligned} \tag{A.11}$$

As to condition (A.2) for the observed information, it follows by the same arguments used for the score, that by standard application of the law of large numbers,

$$n^{-1}I_n\left(\theta_0\right) = n^{-1}\sum_{t=1}^n i_{t,0} \xrightarrow{p} \mathcal{I}_\infty = \begin{pmatrix} \mathcal{I}_{\pi\pi} & 0 \\ 0 & \mathcal{I}_{\gamma\gamma}\end{pmatrix} = \begin{pmatrix} \Omega_{S,\pi\pi} & 0 \\ 0 & \frac{2}{\kappa}\Omega_{S,\gamma\gamma}\end{pmatrix}. \tag{A.12}$$

A.2.3 Asymptotics for the QMLE

Define first the tri-variate Gaussian variable, $\mathcal{Z} := \left(\mathcal{Z}_\pi, \mathcal{Z}_\gamma'\right)'$ with $\mathcal{Z}_\gamma := \left(\mathcal{Z}_a, \mathcal{Z}_\omega\right)'$ and

$$\mathcal{Z} := \mathcal{I}_\infty^{-1}\mathcal{S}_\infty \overset{d}{=} N\left(0, \Omega_{\mathcal{Z}}\right), \ \text{ where } \Omega_{\mathcal{Z}} = \mathcal{I}_\infty^{-1}\Omega_{\mathcal{S}}\mathcal{I}_\infty^{-1}. \tag{A.13}$$

For $\alpha_0 \geq 0$, $\Omega_{\mathcal{S}}$ is given by (A.10), while from (A.12) it follows that

$$\mathcal{I}_\infty^{-1} = \begin{pmatrix} \mathcal{I}_{\pi\pi}^{-1} & 0 \\ 0 & \mathcal{I}_{\gamma\gamma}^{-1}\end{pmatrix} = \begin{pmatrix} \Omega_{S,\pi\pi}^{-1} & 0 \\ 0 & \frac{\kappa}{2}\Omega_{S,\gamma\gamma}^{-1}\end{pmatrix}. \tag{A.14}$$

Hence,

$$\Omega_{\mathcal{Z}} = \mathcal{I}_\infty^{-1}\Omega_S\mathcal{I}_\infty^{-1} = \begin{pmatrix} \Omega_{\mathcal{Z},\pi\pi} & \Omega'_{\mathcal{Z},\gamma\pi} \\ \Omega_{\mathcal{Z},\gamma\pi} & \Omega_{\mathcal{Z},\gamma\gamma}\end{pmatrix}, \tag{A.15}$$

where

$$\Omega_{\mathcal{Z},\pi\pi} = \Omega_{S,\pi\pi}^{-1}, \quad \Omega_{\mathcal{Z},\gamma\gamma} = \tfrac{\kappa}{2}\Omega_{S,\gamma\gamma}^{-1}$$

and

$$\Omega_{\mathcal{Z},\gamma\pi} = \tfrac{\xi}{\delta}\Omega_{\mathcal{Z},\pi\pi}\begin{pmatrix} E(\frac{1}{\sigma_t^4})E(\frac{x_{t-1}^3}{\sigma_t^3}) - E(\frac{x_{t-1}^2}{\sigma_t^4})E(\frac{x_{t-1}}{\sigma_t^3}) \\ E(\frac{x_{t-1}^4}{\sigma_t^4})E(\frac{x_{t-1}}{\sigma_t^3}) - E(\frac{x_{t-1}^2}{\sigma_t^4})E(\frac{x_{t-1}^3}{\sigma_t^3})\end{pmatrix}, \tag{A.16}$$

with $\delta = E(\frac{x_{t-1}^4}{\sigma_t^4})E(\frac{1}{\sigma_t^3}) - (E(\frac{x_{t-1}^2}{\sigma_t^4}))^2$.

Now, with $\hat{\theta}_n := \arg\max_{\theta\in\mathcal{T}}\sum l_t$ by Andrews (1999, Theorem 3),

$$n^{1/2}(\hat{\theta}_n - \theta_0) \xrightarrow{w} \arg\inf_{\lambda\in\Lambda(A)}\|\lambda - \mathcal{Z}\|^2_{\mathcal{I}_\infty}, \tag{A.17}$$

where $A = \mathbb{R}$ if $\alpha_0 > 0$ and $A = \mathbb{R}^+$ if $\alpha_0 = 0$, see (A.1).

In the case of $\alpha_0 > 0$, it follows that for $\Lambda\left(\mathbb{R}\right)$, with $\mathcal{Z}$ defined in (A.13),

$$n^{1/2}(\hat{\theta}_n - \theta_0) \xrightarrow{w} \mathcal{Z} = \left(\mathcal{Z}_\pi, \mathcal{Z}_\gamma'\right)'.$$

Consider now the the case of $\alpha_0 = 0$. For $\Lambda(\mathbb{R}^+)$, use the block-diagonality of $\mathcal{I}_\infty$ to rewrite the quadratic form on the right hand side of (A.17) as

$$\arg \inf_{\lambda \in \Lambda(\mathbb{R}^+)} \|\lambda - \mathcal{Z}\|^2_{\mathcal{I}_\infty} = (\mathcal{Z}_\pi, (\arg \inf_{\lambda \in \mathbb{R}^+ \times \mathbb{R}} \|\lambda - \mathcal{Z}_\gamma\|^2_{\mathcal{I}_{\gamma\gamma}})')'$$

where $\mathcal{Z}_\gamma = \mathcal{I}^{-1}_{\gamma\gamma}\mathcal{S}_\gamma$ has covariance $\Omega_{\mathcal{Z},\gamma\gamma} = \mathcal{I}^{-1}_{\gamma\gamma} = \frac{\kappa}{2}\Omega^{-1}_{\mathcal{S},\gamma\gamma}$, see (A.15).

Next, diagonalization of $\mathcal{I}_{\gamma\gamma}$ is obtained by using the matrix M,

$$M := \begin{pmatrix} 1 & -\varrho \\ 0 & 1 \end{pmatrix}, \quad \varrho := Ex_t^2,$$

such that $\mathcal{I}_{\gamma\gamma}$ is diagonalized by post- (and pre-multiplying) with M (M'). That is,

$$I_{\gamma\gamma} = M\mathcal{I}_{\gamma\gamma}M' = \frac{1}{2\omega_0^2}\begin{pmatrix} \delta_0 & 0 \\ 0 & 1 \end{pmatrix}, \tag{A.18}$$

with $\delta_0 := E\left(x^4_{t-1}\right) - \left(E\left(x^2_{t-1}\right)\right)^2$. Define next,

$$Z_\gamma := (Z_\alpha, Z_\omega)' := \left(M'\right)^{-1}\mathcal{Z}_\gamma = (\mathcal{I}_{\gamma\gamma}M)^{-1}\mathcal{S}_\gamma,$$

which by definition, using the identity $\Omega_{S,\gamma\gamma} = \frac{\kappa}{2}\mathcal{I}_{\gamma\gamma}$, has covariance

$$\frac{\kappa}{2}\left(M\mathcal{I}_{\gamma\gamma}M'\right)^{-1} = \begin{pmatrix} \kappa\omega_0^2/\delta & 0 \\ 0 & \kappa\omega_0^2 \end{pmatrix}.$$

Finally, note that $\Lambda(\mathbb{R}^+)$ is invariant under transformation with the transpose of M^{-1}. That is, for any $(x, y)' \in \Lambda(\mathbb{R}^+)$,

$$\left(M'\right)^{-1}(x, y)' = (x, y - \varrho x) \in \Lambda\left(\mathbb{R}^+\right).$$

Collecting terms,

$$\begin{aligned} \inf_{\lambda \in \mathbb{R}^+ \times \mathbb{R}} \|\lambda - \mathcal{Z}_\gamma\|^2_{\mathcal{I}_{\gamma\gamma}} &= \inf_{\eta = (\eta_\alpha, \eta_\omega)' \in \mathbb{R}^+ \times \mathbb{R}} \|\eta - Z_\gamma\|^2_{I_{\gamma\gamma}} \\ &= \frac{1}{2\omega_0^2} \inf_{\eta \in \mathbb{R}^+ \times \mathbb{R}} \{(\eta_\alpha - Z_\alpha)^2 \delta_0 + (\eta_\omega - Z_\omega)^2\}. \end{aligned} \tag{A.19}$$

It follows that

$$\arg \inf_{\eta \in \mathbb{R}^+ \times \mathbb{R}} \{(\eta_\alpha - Z_\alpha)^2 \delta_0 + (\eta_\omega - Z_\omega)^2\} = (\max(0, Z_\alpha), Z_\omega)',$$

such that by (A.19), and using that by definition, $\lambda = M'\eta$,

$$\begin{aligned} \arg \inf_{\lambda \in \mathbb{R}^+ \times \mathbb{R}} \|\lambda - \mathcal{Z}_\gamma\|^2_{\mathcal{I}_{\gamma\gamma}} &= M'(\max(0, Z_\alpha), Z_\omega)' \\ &= (\max(0, Z_\alpha), Z_\omega - \varrho\max(0, Z_\alpha)). \end{aligned} \tag{A.20}$$

Here, Z_α and Z_ω are independent Gaussian distributed with

$$Z_\alpha \overset{d}{=} N\left(0, \kappa\omega_0^2/\delta\right) \quad \text{and} \quad Z_\omega \overset{d}{=} N\left(0, \kappa\omega_0^2\right) \tag{A.21}$$

This establishes Theorem 1.

Remark A.1 Note that if $\xi = 0$, the covariance of $\mathcal{Z}$, see (A.15) becomes block-diagonal,

$$\Omega_{\mathcal{Z}} = \begin{pmatrix} \Omega_{\mathcal{Z},\pi\pi} & 0 \\ 0 & \Omega_{\mathcal{Z},\gamma\gamma} \end{pmatrix}. \tag{A.22}$$

Remark A.2 The above also reduces to Ling (2004, Theorem 1) for the case of $\alpha_0 > 0$ and $\xi = 0$.

A.3 QMLE and LR test under non-stationarity – Theorem 2 and Theorem 3

We proceed in the following by establishing regularity conditions under which the asymptotic distribution of the QMLE and the likelihood ratio test can be derived for the non-stationary case where $\theta_0 \in \Theta_{\mathcal{N}}$. Specifically, we verify the following regularity conditions (C.i)-(C.ii) in terms of the log-likelihood function, $L_n(\theta)$, and its derivatives.

Condition (C.*i*). *With* $G_n = diag(g_{n,i})_{i=1,2,3}$, *where*

$$(g_{1,n}, g_{2,n}, g_{3,n}) = (n, n^{3/2}, n^{1/2}),$$

it holds that

$$\left(G_n^{-1} S_n(\theta_0), G_n^{-1} I_n(\theta_0) G_n^{-1}\right) \xrightarrow{w} (\mathcal{S}_\infty, \mathcal{I}_\infty). \tag{A.23}$$

Condition (C.*ii*). *With* $\theta := (\theta_1, \theta_2, \theta_3)' = (\pi, \alpha, \omega)'$, *and* $i, j, k = 1, 2, 3$,

$$\sup_{\theta \in N_n(\theta_0)} \left\| n^{1/2} \left(\partial^3 L_n(\theta) / \partial\theta_i \partial\theta_j \partial\theta_k\right) / (g_{i,n} g_{j,n} g_{k,n}) \right\| = O_p(1) \tag{A.24}$$

where the supremum is over a the sequence of neigborhoods given by,

$$N_n(\theta_0) = \left\{\theta : g_{1,n}^2 \pi^2 + g_{2,n}^2 \alpha^2 + g_{3,n}^2 (\omega - \omega_0)^2 < \varepsilon/n\right\}.$$

Conditions (C.i) and (C.ii) are from Kristensen and Rahbek (2010, Lemma 11 and Lemma 12) where general asymptotic theory is presented for (non-)stationary variables. With the parameter spaces $\mathcal{T}$ and $\mathcal{T}_0$ satisfying *(i)*, that is, shifted they are locally equal to Λ and Λ_0, it follows as in Klüppelberg *et al.* (2002, Lemma B.1), see also Vu and Zhou (1997) and Andrews (2001), that with

$$\mathcal{Z} := \mathcal{I}_\infty^{-1} \mathcal{S}_\infty,$$

then the LR_n statistic converges in distribution:

$$LR_n \to_w \mathcal{LR}_\infty(\kappa) = \inf_{\lambda \in \Lambda_0} \|\lambda - \mathcal{Z}\|_{\mathcal{I}_\infty}^2 - \inf_{\lambda \in \Lambda} \|\lambda - \mathcal{Z}\|_{\mathcal{I}_\infty}^2 . \tag{A.25}$$

Likewise, as in Andrews (1999, Theorem 3), under (C.i)-(C.ii) it follows that

$$G_n(\hat{\theta}_n - \theta_0) \to_w \arg\inf_{\lambda \in \Lambda} \|\lambda - \mathcal{Z}\|_{\mathcal{I}_\infty}^2 . \tag{A.26}$$

A.3.1 Preliminaries

Note initially, that under the null hypothesis H_0, $S_n(\theta_0) = \sum_{t=1}^{n} s_{t,0}$, see (A.3), where with $\theta_0 \in \Theta_{\mathcal{N}}$,

$$s_{t,0} = (v_{t-1}z_t, \tfrac{1}{2}v_{t-1}^2\left(z_t^2-1\right), \tfrac{1}{2\omega_0}\left(z_t^2-1\right))', \quad \text{with } v_t := \sum_{i=1}^{t} z_i.$$

Standard application of the invariance principle implies convergence to the Brownian motion V_u, $u \in (0,1)$:

$$n^{-1/2}\sum_{t=1}^{\lfloor n\cdot\rfloor}\left(z_t, z_t^2-1\right)' \xrightarrow{w} V := (V_1, V_2)', \ E(V_1V_1') = \begin{pmatrix} 1 & \xi \\ \xi & \kappa \end{pmatrix}. \tag{A.27}$$

Define next the matrix

$$Q = \begin{pmatrix} 1 & 0 \\ -\xi/\sqrt{\kappa-\xi^2} & 1/\sqrt{\kappa-\xi^2} \end{pmatrix}, \text{ with } Q^{-1} = \begin{pmatrix} 1 & 0 \\ \xi & \sqrt{\kappa-\xi^2} \end{pmatrix}, \tag{A.28}$$

and use it to define the bivariate standard Brownian motion $(B, W)'$:

$$(B, W)' := QV = (V_1, (V_2 - \xi V_1)/\sqrt{\kappa-\xi^2})'. \tag{A.29}$$

It then follows that

$$n^{-1/2}\sum_{t=1}^{\lfloor n\cdot\rfloor} Q\left(z_t, z_t^2-1\right)' \xrightarrow{w} (B_\cdot, W_\cdot)'. \tag{A.30}$$

A.3.2 Score – Condition (C.i)

Consider next the score $S_n(\theta_0)$, normalized by G_n where

$$G_n^{-1} = \operatorname{diag}\left(n^{-1}, n^{-3/2}, n^{-1/2}\right). \tag{A.31}$$

It follows, with

$$G_n^{-1}S_n(\theta_0) = \frac{1}{n^{1/2}}\sum^{n}\left(\begin{pmatrix} n^{-1/2}\sum_{i=1}^{t-1} z_t & 0 \\ 0 & \left(n^{-1/2}\sum_{i=1}^{t-1} z_t\right)^2/2 \\ 0 & \frac{1}{2\omega_0} \end{pmatrix}\right)\left(z_t, z_t^2-1\right)',$$

that

$$\begin{aligned} G_n^{-1}S_n(\theta_0) &\xrightarrow{w} \mathcal{S}_\infty(\xi) = \left(\mathcal{S}_\pi, \mathcal{S}_\gamma'\right)' = (\mathcal{S}_\alpha, \mathcal{S}_\alpha, \mathcal{S}_\omega)' \qquad \text{(A.32)} \\ &= \left(\int B\mathsf{d}B, \tfrac{\xi}{2}\int B^2\mathsf{d}B + \tfrac{\sqrt{\kappa-\xi^2}}{2}\int B^2\mathsf{d}W, \tfrac{\xi}{2\omega_0}B_1 + \tfrac{\sqrt{\kappa-\xi^2}}{2\omega_0}W_1\right)'. \end{aligned}$$

A.3.3 Information – Condition (C.i)

Under H_0, it follows by standard arguments that (jointly with the score) the information $I_n(\theta_0) = \sum_1^n i_{t,0}$ converges weakly

$$G_n^{-1} I_n(\theta_0) G_n^{-1} \xrightarrow{w} \mathcal{I}_\infty = \begin{pmatrix} \mathcal{I}_{\pi\pi} & 0 \\ 0 & \mathcal{I}_{\gamma\gamma} \end{pmatrix}, \tag{A.33}$$

with

$$\mathcal{I}_{\gamma\gamma} = \frac{1}{2} \begin{pmatrix} \int B^4 \mathsf{d}u & \frac{1}{\omega_0} \int B^2 \mathsf{d}u \\ \frac{1}{\omega_0} \int B^2 \mathsf{d}u & \frac{1}{\omega_0^2} \end{pmatrix}, \quad \text{and} \quad \mathcal{I}_{\pi\pi} = \int B^2 \mathsf{d}u.$$

Also observe that by, definition,

$$\mathcal{I}_\infty^{-1} = \begin{pmatrix} \mathcal{I}_{\pi\pi}^{-1} & 0 \\ 0 & \mathcal{I}_{\gamma\gamma}^{-1} \end{pmatrix}, \quad \mathcal{I}_{\gamma\gamma}^{-1} = \frac{2}{\delta} \begin{pmatrix} 1 & -\omega_0 \int B^2 \mathsf{d}u \\ -\omega_0 \int B^2 \mathsf{d}u & \omega_0^2 \int B^4 \mathsf{d}u \end{pmatrix} \tag{A.34}$$

with $\delta = \int B^4 \mathsf{d}u - \left(\int B^2 \mathsf{d}u\right)^2$.

A.3.4 Third order derivatives – Condition (C.ii)

It follows that (C.ii) holds by the considerations in Appendix A.6 below.

A.4 QMLE – Proof of Theorem 2

By (A.26) we have,

$$G_n(\hat{\theta}_n - \theta_0) \xrightarrow{w} \arg \inf_{\lambda \in \Lambda} \|\lambda - \mathcal{Z}\|_{\mathcal{I}_\infty}^2 =: (\lambda_\pi, \lambda_\alpha, \lambda_\omega)',$$

with $\mathcal{Z}$ given by (A.35). As before, by block-diagonality of $\mathcal{I}_\infty$ in (A.34) and the definition of $\mathcal{Z}$ in (A.35),

$$\lambda_\pi = \mathcal{Z}_\pi = \int B \mathsf{d}B \Big/ \int B^2 \mathsf{d}u.$$

For $\lambda_\gamma = (\lambda_\alpha, \lambda_\omega)'$ use that by definition of $Z_\gamma := (Z_\alpha, Z_\omega)'$ defined in (A.37), we have

$$\inf_{\lambda_\gamma \in \mathbb{R}^+ \times \mathbb{R}} \|\lambda_\gamma - \mathcal{Z}_\gamma\|_{\mathcal{I}_{\gamma\gamma}}^2 = \inf_{\eta \in \mathbb{R}^+ \times \mathbb{R}} \|\eta - Z_\gamma\|_{I_{\gamma\gamma}}^2.$$

In terms of $\eta = (\eta_\alpha, \eta_\omega)'$ we find

$$\begin{aligned} \arg \inf_{\eta \in \mathbb{R}^+ \times \mathbb{R}} \|\eta - Z_\gamma\|_{I_{\gamma\gamma}}^2 &= \arg \inf_{\eta = (\eta_\alpha, \eta_\omega)' \in \mathbb{R}^+ \times \mathbb{R}} \left((\eta_\alpha - Z_\alpha)^2 + (\eta_\omega - Z_\omega)^2\right) \\ &= (\max(0, Z_\alpha), Z_\omega)'. \end{aligned}$$

Finally, use the identity $\lambda_\gamma = (\lambda_\alpha, \lambda_\omega)' = M'\eta$ to see that

$$\lambda_\gamma = M' (\max(0, Z_\alpha), Z_\omega)' = (\max(0, Z_\alpha), Z_\omega - (\omega_0 \int B^2 \mathsf{d}u) \max(0, Z_\alpha))'.$$

Collecting terms, and setting $\xi = 0$, ends the proof of Theorem 2.

A.4.1 LR_n convergence – Proof of Theorem 3

From (A.25),

$$LR_n \xrightarrow{w} \mathcal{LR}_\infty(\kappa) = \inf_{\lambda \in \Lambda_0} \|\lambda - \mathcal{Z}\|^2_{\mathcal{I}_\infty} - \inf_{\lambda \in \Lambda} \|\lambda - \mathcal{Z}\|^2_{\mathcal{I}_\infty},$$

where $\mathcal{Z} := (\mathcal{Z}_\pi, \mathcal{Z}_\alpha, \mathcal{Z}_\omega)' = \left(\mathcal{Z}_\pi, \mathcal{Z}_\gamma'\right)' = \mathcal{I}_\infty^{-1}\mathcal{S}_\infty$ satisfies $\mathcal{Z}_\pi = \int B \mathsf{d}B / \int B^2 \mathsf{d}u$,

$$\mathcal{Z}_\alpha = \tfrac{1}{\delta}((\xi \int B^2 \mathsf{d}B + \sqrt{\kappa - \xi^2} \int B^2 \mathsf{d}W) - \int B^2 \mathsf{d}u(\xi B_1 + \sqrt{\kappa - \xi^2} W_1)), \quad \text{(A.35)}$$

and

$$\mathcal{Z}_\omega = \tfrac{\omega_0}{\delta}(\int B^4 \mathsf{d}u(\xi B_1 + \sqrt{\kappa - \xi^2} W_1) - \int B^2 \mathsf{d}u(\xi \int B^2 \mathsf{d}B + \sqrt{\kappa - \xi^2} \int B^2 \mathsf{d}W)).$$

By the block-diagonality of $\mathcal{I}_\infty$ in (A.34), we may write $\mathcal{LR}_\infty(\kappa)$ as

$$\mathcal{LR}_\infty(\kappa) = \mathcal{Z}_\pi^2 \mathcal{I}_{\pi\pi} + \inf_{\lambda \in \{0\} \times \mathbb{R}} \|\lambda - \mathcal{Z}_\gamma\|^2_{\mathcal{I}_{\gamma\gamma}} - \inf_{\lambda \in \mathbb{R}^+ \times \mathbb{R}} \|\lambda - \mathcal{Z}_\gamma\|^2_{\mathcal{I}_{\gamma\gamma}}$$

Diagonalization of $\mathcal{I}_{\gamma\gamma}$ can next be obtained by using the matrix M defined as

$$M := \begin{pmatrix} 1 & -\omega_0 \int B^2 \mathsf{d}u \\ 0 & 1 \end{pmatrix},$$

such that

$$I_{\gamma\gamma} := M \mathcal{I}_{\gamma\gamma} M' = \tfrac{1}{2} \begin{pmatrix} \delta & 0 \\ 0 & \frac{1}{\omega_0^2} \end{pmatrix}. \quad \text{(A.36)}$$

Next, note that $\mathcal{Z}_\gamma = \mathcal{I}_{\gamma\gamma}^{-1} \mathcal{S}_\gamma$, and hence we can define $Z_\gamma := (Z_\alpha, Z_\omega)'$, where

$$Z_\gamma = \left(M'\right)^{-1} \mathcal{Z}_\gamma = \left(\mathcal{I}_{\gamma\gamma} M'\right)^{-1} \mathcal{S}_\gamma. \quad \text{(A.37)}$$

By definition,

$$\mathcal{I}_{\gamma\gamma} M' = \tfrac{1}{2} \begin{pmatrix} \delta & \frac{1}{\omega_0} \int B^2 \mathsf{d}u \\ 0 & \frac{1}{\omega_0^2} \end{pmatrix}$$

and hence

$$Z_\gamma = \begin{pmatrix} \delta^{-1}((\xi \int B^2 \mathsf{d}B + \sqrt{\kappa - \xi^2} \int B^2 \mathsf{d}W) - \int B^2(\{\mathsf{d}u \xi B_1 + \sqrt{\kappa - \xi^2} W_1)) \\ \omega_0(\xi B_1 + \sqrt{\kappa - \xi} W_1) \end{pmatrix}$$

Finally, the cones $\mathbb{R}^+ \times \mathbb{R}$ and $\{0\} \times \mathbb{R}$ are invariant to multiplication by $(M')^{-1}$, such that we get, using the identity (A.36),

$$\begin{aligned} \inf_{\lambda \in \mathbb{R}^+ \times \mathbb{R}} \|\lambda - \mathcal{Z}_\gamma\|^2_{\mathcal{I}_{\gamma\gamma}} &= \inf_{\lambda \in \mathbb{R}^+ \times \mathbb{R}} (\lambda - \mathcal{Z}_\gamma)' \mathcal{I}_{\gamma\gamma} (\lambda - \mathcal{Z}_\gamma) \\ &= \inf_{\lambda \in \mathbb{R}^+ \times \mathbb{R}} (\lambda - Z_\gamma)' \left(M \mathcal{I}_{\gamma\gamma} M'\right) (\lambda - Z_\gamma) \\ &= \tfrac{\delta}{2} \inf_{\lambda \in \mathbb{R}^+} (\lambda - Z_\alpha)^2 + \inf_{\lambda \in \mathbb{R}} (\lambda - Z_\omega)^2 / \left(2\omega_0^2\right) \\ &= \tfrac{\delta}{2} Z_\alpha^2 \mathbb{I}(Z_\alpha < 0). \end{aligned} \quad \text{(A.38)}$$

Here, by definition,

$$Z_\alpha = \delta^{-1}((\xi \int B^2 \mathsf{d}B + \sqrt{\kappa - \xi^2} \int B^2 \mathsf{d}W) - \int B^2 \mathsf{d}u(\xi B_1 + \sqrt{\kappa - \xi^2} W_1)). \quad \text{(A.39)}$$

Collecting terms we find

$$\begin{aligned}\mathcal{LR}_\infty(\kappa) &= \mathcal{Z}_\pi^2 \mathcal{I}_{\pi\pi} + \frac{\delta}{2} Z_\alpha^2 1(Z_\alpha < 0) \\ &= (\int B \mathsf{d}B)^2 / \int B^2 \mathsf{d}u + \frac{\delta}{2} Z_\alpha^2 1(Z_\alpha < 0),\end{aligned}$$

and setting $\xi = 0$ ends the proof of Theorem 3.

A.5 Bootstrap – Proof of Theorem 4

We verify here the equivalent of the conditions (C.i) and (C.ii) for the bootstrap from which the bootstrap results are derived.

A.5.1 Bootstrap score and information

It follows that the bootstrap score is given by

$$s_{t,0}^* = \left(v_{t-1}^* z_t^*, \tfrac{1}{2} v_{t-1}^{*2} \left(z_t^{*2} - 1 \right), \tfrac{1}{2\tilde{\omega}_n} \left(z_t^{*2} - 1 \right) \right)', \quad \text{with } v_t^* = \sum_{i=1}^t z_i^*.$$

The bootstrap invariance principle (cf. Cavaliere, Rahbek and Taylor, 2012) implies the main result of convergence to the Brownian motion V^*, as stated in the following Lemma.

Lemma A.1 *Assume that $E(z_t^8) < \infty$. Then, as $n \to \infty$,*

$$n^{-1/2} \sum_{t=1}^{\lfloor n \cdot \rfloor} \left(z_t^*, z_t^{*2} - 1 \right)' \xrightarrow{w^*}_p V_\cdot^* = (V_{1\cdot}^*, V_{2\cdot}^*)', \; E(V_1^* V_1^{*\prime}) = \begin{pmatrix} 1 & \xi \\ \xi & \kappa \end{pmatrix}.$$

Proof. By definition, z_t^* is re-sampled with replacement from $\tilde{z}_{s,t}$,

$$\tilde{z}_{s,t} = \frac{\tilde{z}_t - n^{-1} \sum_{t=1}^n \tilde{z}_t}{(n^{-1} \sum_{t=1}^n (\tilde{z}_t - n^{-1} \sum_{t=1}^n \tilde{z}_t)^2)^{1/2}},$$

where, under H_0,

$$\tilde{z}_t = \tilde{\omega}_n^{-1/2} \Delta x_t = \tilde{\omega}_n^{-1/2} \omega_0^{1/2} z_t.$$

With $m_t^* := \left(z_t^*, z_t^{*2} - 1 \right)'$ consider, for any $\lambda \in \mathbb{R}^2$, $\lambda \neq 0$,

$$\lambda' m_t^* = \lambda_1 z_t^* + \lambda_2 \left(z_t^{*2} - 1 \right).$$

Again, conditional on data, $\lambda' m_t^*$ is i.i.d., and hence as in Swensen (2003, eq. (10), proof of Theorem 1) it suffices to establish

$$E^* \left(\lambda' m_t^* \right)^2 \xrightarrow{p} E \left(\lambda' m_t \right)^2, \quad \text{and} \quad E^* \left(\lambda' m_t^* \right)^4 \xrightarrow{p} E \left(\lambda' m_t \right)^4$$

where $m_t = \left(z_t, z_t^2 - 1\right)'$, which by standard arguments holds if $Ez_t^8 < \infty$. This ends the proof of Lemma A.1. $\square$

Next, with $Q^* = Q$ in (A.28), construct the bivariate standard Brownian motion $(B^*, W^*)'$ as

$$(B^*, W^*)' = Q^* V^* = (V_1^*, (V_2^* - \xi V_1^*) / \sqrt{\kappa - \xi^2})',$$

such that

$$n^{-1/2} \sum_{t=1}^{\lfloor n\cdot \rfloor} Q^* \left(z_t^*, z_t^{*2} - 1\right)' \xrightarrow{w^*}_p (B_\cdot^*, W_\cdot^*)'. \tag{A.40}$$

Then, the following lemma follows.

Lemma A.2 *If $Ez_t^8 < \infty$, and with G_n defined in (A.31), then the bootstrap score sasfisties,*

$$G_n^{-1} S_n^* \xrightarrow{w^*}_p \mathcal{S}_\infty^*,$$

where $\mathcal{S}_\infty^* = (\mathcal{S}_\pi^*, \mathcal{S}_\alpha^*, \mathcal{S}_\omega^*)'$ with,

$$\mathcal{S}_\infty^{*\prime} = (\int B^* \mathsf{d} B^*, \tfrac{\xi}{2} \int B^{*2} \mathsf{d} B^* - \tfrac{\sqrt{\kappa - \xi^2}}{2} \int B^{*2} \mathsf{d} W^*, \tfrac{\xi}{2\omega_0} B_1^* - \tfrac{\sqrt{\kappa - \xi^2}}{2\omega_0} W_1^*).$$

We also have the following result on the information.

Lemma A.3 *Under the conditions of Lemma A.2 it follows that the bootstrap information converges jointly with the score as follows:*

$$G_n^{-1} \left(\sum_1^n i_{t,0}^*\right) G_n^{-1} \xrightarrow{w^*}_p \mathcal{I}_\infty^* = \begin{pmatrix} \mathcal{I}_{\pi\pi}^* & 0 \\ 0 & \mathcal{I}_{\gamma\gamma}^* \end{pmatrix}, \text{ with}$$

$$\mathcal{I}_{\gamma\gamma}^* = \frac{1}{2} \begin{pmatrix} \int B^{*4} \mathsf{d} u & \frac{1}{\omega_0} \int B^{*2} \mathsf{d} u \\ \frac{1}{\omega_0} \int B^{*2} \mathsf{d} u & \frac{1}{\omega_0^2} \end{pmatrix},$$

and $\mathcal{I}_{\pi\pi}^* = \int B^{*2} \mathsf{d} u$.

Finally, condition (C.ii) is shown in Appendix A.6 to hold also for the bootstrap case.

A.5.2 Bootstrap LR_n^* statistic

Observe that by definition

$$\mathcal{I}_\infty^{*-1} = \begin{pmatrix} \mathcal{I}_{\pi\pi}^{*-1} & 0 \\ 0 & \mathcal{I}_{\gamma\gamma}^{*-1} \end{pmatrix}, \quad \mathcal{I}_{\gamma\gamma}^{*-1} = \frac{2}{\delta^*} \begin{pmatrix} 1 & -\omega_0 \int B^{*2} \mathsf{d} u \\ -\omega_0 \int B^{*2} \mathsf{d} u & \omega_0^2 \int B^{*4} \mathsf{d} u \end{pmatrix},$$

with $\delta^* = \left[\int B^{*4} \mathsf{d} u - \left(\int B^{*2}\right)^2\right]$. We define $\mathcal{Z}^* = (\mathcal{Z}_\pi^*, \mathcal{Z}_\alpha^*, \mathcal{Z}_\omega^*)' = \mathcal{I}_\infty^{*-1} \mathcal{S}_\infty^* (\xi)$, where

$$\mathcal{Z}_\pi^* = \int B^* \mathsf{d} B^* / \int B^{*2} \mathsf{d} u \tag{A.41}$$

$$\mathcal{Z}_\pi^* = \tfrac{1}{\delta^*} (\{\xi \int B^{*2} \mathsf{d} B^* + \sqrt{\kappa - \xi^2} \int B^{*2} \mathsf{d} W\} - \int B^{*2} \mathsf{d} u \{\xi B_1^* + \sqrt{\kappa - \xi^2} W_1\})$$

$$\mathcal{Z}^*_\omega = -\tfrac{\omega_0}{\delta^*}(\int B^{*2}\mathsf{d}u\{\xi\int B^{*2}\mathsf{d}B^* + \sqrt{\kappa-\xi^2}\int B^{*2}\mathsf{d}W\}$$
$$+\int B^{*4}\mathsf{d}u\{\xi B^*_1 + \sqrt{\kappa-\xi^2}W_1\})$$

It follows that, as for the LR_n statistic under H_0, $LR^*_n \xrightarrow{w^*}_p \mathcal{LR}^*_\infty(\kappa)$, where,

$$\mathcal{LR}^*_\infty(\kappa) = (\mathcal{Z}^*_\pi)^2\mathcal{I}^*_{\pi\pi} + \frac{\delta^*}{2}Z^{*2}_\alpha 1(Z^*_\alpha > 0)$$
$$= (\int B^*\mathsf{d}B^*)^2/\int B^{*2}\mathsf{d}u + \frac{\delta^*}{2}Z^{*2}_\alpha 1(Z^*_\alpha > 0),$$

with

$$Z^*_\alpha = \delta^{*-1}(\{\xi\int B^{*2}\mathsf{d}B^* + \sqrt{\kappa-\xi^2}\int B^{*2}\mathsf{d}W^*\} \tag{A.42}$$
$$-\int B^{*2}\mathsf{d}u\{\xi B^*_1 + \sqrt{\kappa-\xi^2}W^*_1\}).$$

which ends the proof of Theorem 4 using $\mathcal{LR}^*_\infty(\kappa) \overset{d}{=} \mathcal{LR}_\infty(\kappa)$.

A.5.3 Bootstrap – Proof of Theorem 6 and Theorem 7

The proof of Theorem 6 follows by replicating the proof of Theorem 4, as Lemma A.1 also applies to the case where the bootstrap innovations z^*_t are resampled from

$$\hat{z}_{s,t} = \frac{\hat{z}_t - n^{-1}\sum_{t=1}^n \hat{z}_t}{(n^{-1}\sum_{t=1}^n(\hat{z}_t - n^{-1}\sum_{t=1}^n \hat{z}_t)^2)^{1/2}}, \tag{A.43}$$

where the unrestricted residuals are given by

$$\hat{z}_t = (\Delta x_t - \hat{\pi}_n x_{t-1}) / \left(\hat{\omega}_n + \hat{\alpha}_n x^2_{t-1}\right)^{1/2}.$$

The proof of Theorem 7 holds trivially as all arguments used to establish Theorems 4 and 6 allow $\xi \neq 0$.

A.6 On the third order derivatives – Condition (C.ii)

A.6.1 Non-bootstrap case

With c and $(c_i)_{i=1}^3$ generic constants, it follows that (C.ii) holds as follows:

$$\partial^3 L_n(\theta)/\partial\pi^3 = 0$$
$$g^{-3}_{1,n}\partial^3 L_n(\theta)/\partial\pi^2\partial\alpha = n^{-3}\sum_{t=1}^n \frac{x^4_{t-1}}{\sigma^4_t} \leq cn^{-3}\sum_{t=1}^n x^4_{t-1} = O_p(1)$$
$$n^{1/2}g^{-2}_{1,n}g^{-1}_{3,n}\partial^3 L_n(\theta)/\partial\pi^2\partial\omega = n^{-2}\sum_{t=1}^n \frac{x^2_{t-1}}{\sigma^4_t} \leq cn^{-2}\sum_{t=1}^n x^2_{t-1} = O_p(1)$$

$$\left|n^{1/2}g^{-3}_{2,n}\partial^3 L_n(\theta)/\partial\alpha^3\right| = \left|n^{-4}\sum_{t=1}^n[3\frac{\varepsilon^2_t}{\sigma^2_t} - 1]\left(\frac{x^6_{t-1}}{\sigma^6_t}\right)\right|$$

$$\leq c_1 n^{-4}\sum_{t=1}^{n} x_{t-1}^6\left(z_t^2-1\right)+c_2 n^{-4}\sum_{t=1}^{n} x_{t-1}^6 = O_p\left(1\right)$$

$$\left|n^{1/2}g_{2,n}^{-2}g_{3,n}^{-1}\partial^3 L_n\left(\theta\right)/\partial\alpha^2\partial\omega\right| = \left|n^{-3}\sum_{t=1}^{n}[3\frac{\varepsilon_t^2}{\sigma_t^2}-1]\left(\frac{x_{t-1}^4}{\sigma_t^6}\right)\right|$$

$$\leq c_1 n^{-3}\sum_{t=1}^{n} x_{t-1}^4\left(z_t^2-1\right)+c_2 n^{-3}\sum_{t=1}^{n} x_{t-1}^4 = O_p\left(1\right)$$

$$\left|n^{1/2}g_{2,n}^{-2}g_{1,n}^{-1}\partial^3 L_n\left(\theta\right)/\partial\alpha^2\partial\pi\right| = \left|n^{-7/2}\sum_{t=1}^{n}2\frac{\varepsilon_t x_{t-1}^5}{\sigma_t^6}\right|$$

$$\leq c_1 n^{-7/2}\sum_{t=1}^{n}\left(|z_t|-E\left|z_t\right|\right)\left|x_{t-1}^5\right|+c_2 n^{-7/2}\sum_{t=1}^{n}\left|x_{t-1}^5\right| = O_p\left(1\right)$$

$$\left|n^{1/2}g_{3,n}^{-3}\partial^3 L_n\left(\theta\right)/\partial\omega^3\right| = \left|n^{-1}\sum_{t=1}^{n}[3\frac{\varepsilon_t^2}{\sigma_t^2}-1]\left(\frac{1}{\sigma_t^6}\right)\right|$$

$$\leq c_1 n^{-1}\sum_{t=1}^{n}[z_t^2-1]+c_2 = O_p\left(1\right)$$

$$\left|n^{1/2}g_{1,n}^{-1}g_{3,n}^{-2}\partial^3 L_n\left(\theta\right)/\partial\omega^2\partial\pi\right| = \left|n^{-3/2}\sum_{t=1}^{n}2\frac{\varepsilon_t x_{t-1}}{\sigma_t^6}\right|$$

$$\leq c_1 n^{-3/2}\sum_{t=1}^{n}\left(|z_t|-E\left|z_t\right|\right)\left|x_{t-1}\right|+c_2 n^{-3/2}\sum_{t=1}^{n}\left|x_{t-1}\right| = O_p\left(1\right)$$

$$\left|n^{1/2}g_{3,n}^{-2}g_{2,n}^{-1}\partial^3 L_n\left(\theta\right)/\partial\omega^2\partial\alpha\right| = \left|n^{-2}\sum_{t=1}^{n}[3\frac{\varepsilon_t^2 x_{t-1}^2}{\sigma_t^2}-1]\left(\frac{1}{\sigma_t^6}\right)\right|$$

$$\leq c_1 n^{-2}\sum_{t=1}^{n}[z_t^2-1]x_{t-1}^2+c_2 n^{-2}\sum_{t=1}^{n}\left(x_{t-1}^2+1\right) = O_p\left(1\right).$$

$$\left|n^{1/2}g_{3,n}^{-1}g_{2,n}^{-1}g_{1,n}^{-1}\partial^3 L_n\left(\theta\right)/\partial\pi\partial\omega\partial\alpha\right| = \left|n^{-5/2}\sum_{t=1}^{n}2\frac{\varepsilon_t x_{t-1}^3}{\sigma_t^6}\right|$$

$$\leq c_1 n^{-5/2}\sum_{t=1}^{n}\left(|z_t|-E\left|z_t\right|\right)\left|x_{t-1}^3\right|+c_2 n^{-5/2}\sum_{t=1}^{n}\left|x_{t-1}^3\right| = O_p\left(1\right)$$

Remark A.3 Note that we here used that an invariance principle applies to the term, $\sum_{t=1}^{[n\cdot]}\left(|z_t|-E\left|z_t\right|\right)$ normalized by $n^{-1/2}$.

A.6.2 Bootstrap case

With c and $(c_i)_{i=1}^3$ generic constants, it follows that (C.ii) holds for the bootstrap by replicating the arguments in Appendix A.6. That is, we have:

$$\partial^3 L_n^*\left(\theta\right)/\partial\pi^3 = 0$$

$$g_{1,n}^{-3}\partial^3 L_n^*(\theta)/\partial\pi^2\partial\alpha = n^{-3}\sum_{t=1}^{n}\frac{x_{t-1}^{*4}}{\sigma_t^{*4}} \le cn^{-3}\sum_{t=1}^{n}x_{t-1}^{*4} = O_p^*(1)$$

$$n^{1/2}g_{1,n}^{-2}g_{3,n}^{-1}\partial^3 L_n^*(\theta)/\partial\pi^2\partial\omega = n^{-2}\sum_{t=1}^{n}\frac{x_{t-1}^{*2}}{\sigma_t^{*4}} \le cn^{-2}\sum_{t=1}^{n}x_{t-1}^{*2} = O_p^*(1)$$

$$\left|n^{1/2}g_{2,n}^{-3}\partial^3 L_n^*(\theta)/\partial\alpha^3\right| = \left|n^{-4}\sum_{t=1}^{n}[3\frac{\varepsilon_t^{*2}}{\sigma_t^{*2}}-1]\left(\frac{x_{t-1}^{*6}}{\sigma_t^{*6}}\right)\right|$$
$$\le c_1 n^{-4}\sum_{t=1}^{n}x_{t-1}^{*6}\left(z_t^{*2}-1\right) + c_2 n^{-4}\sum_{t=1}^{n}x_{t-1}^{*6} = O_p^*(1)$$

$$\left|n^{1/2}g_{2,n}^{-2}g_{3,n}^{-1}\partial^3 L_n^*(\theta)/\partial\alpha^2\partial\omega\right| = \left|n^{-3}\sum_{t=1}^{n}[3\frac{\varepsilon_t^{*2}}{\sigma_t^{*2}}-1]\left(\frac{x_{t-1}^{*4}}{\sigma_t^{6}}\right)\right|$$
$$\le c_1 n^{-3}\sum_{t=1}^{n}x_{t-1}^{*4}\left(z_t^{*2}-1\right) + c_2 n^{-3}\sum_{t=1}^{n}x_{t-1}^{*4} = O_p^*(1)$$

$$\left|n^{1/2}g_{2,n}^{-2}g_{1,n}^{-1}\partial^3 L_n^*(\theta)/\partial\alpha^2\partial\pi\right| = \left|n^{-7/2}\sum_{t=1}^{n}2\frac{\varepsilon_t^{*}x_{t-1}^{*5}}{\sigma_t^{*6}}\right| \le$$
$$\le c_1 n^{-7/2}\sum_{t=1}^{n}\left(|z_t^*| - E^*|z_t^*|\right)\left|x_{t-1}^{*5}\right| + c_2 n^{-7/2}\sum_{t=1}^{n}\left|x_{t-1}^{*5}\right| = O_p^*(1)$$

$$\left|n^{1/2}g_{3,n}^{-3}\partial^3 L_n^*(\theta)/\partial\omega^3\right| = \left|n^{-1}\sum_{t=1}^{n}[3\frac{\varepsilon_t^{*2}}{\sigma_t^{*2}}-1]\left(\frac{1}{\sigma_t^{*6}}\right)\right|$$
$$\le c_1 n^{-1}\sum_{t=1}^{n}[z_t^{*2}-1] + c_2 = O_p^*(1)$$

$$\left|n^{1/2}g_{1,n}^{-1}g_{3,n}^{-2}\partial^3 L_n^*(\theta)/\partial\omega^2\partial\pi\right| = \left|n^{-3/2}\sum_{t=1}^{n}2\frac{\varepsilon_t^{*}x_{t-1}^{*}}{\sigma_t^{*6}}\right|$$
$$\le c_1 n^{-3/2}\sum_{t=1}^{n}\left(|z_t^*| - E^*|z_t^*|\right)\left|x_{t-1}^{*}\right| + c_2 n^{-3/2}\sum_{t=1}^{n}\left|x_{t-1}^{*}\right| = O_p^*(1)$$

$$\left|n^{1/2}g_{3,n}^{-2}g_{2,n}^{-1}\partial^3 L_n^*(\theta)/\partial\omega^2\partial\alpha\right| = \left|n^{-2}\sum_{t=1}^{n}[3\frac{\varepsilon_t^{*2}x_{t-1}^{*2}}{\sigma_t^{*2}}-1]\left(\frac{1}{\sigma_t^{*6}}\right)\right|$$
$$\le c_1 n^{-2}\sum_{t=1}^{n}[z_t^{*2}-1]x_{t-1}^{*2} + c_2 n^{-2}\sum_{t=1}^{n}x_{t-1}^{*2} + c_3 = O_p^*(1).$$

$$\left|n^{1/2}g_{3,n}^{-1}g_{2,n}^{-1}g_{1,n}^{-1}\partial^3 L_n^*(\theta)/\partial\pi\partial\omega\partial\alpha\right| = \left|n^{-5/2}\sum_{t=1}^{n}2\frac{\varepsilon_t^{*}x_{t-1}^{*3}}{\sigma_t^{*6}}\right|$$
$$\le c_1 n^{-5/2}\sum_{t=1}^{n}\left(|z_t^*| - E^*|z_t^*|\right)\left|x_{t-1}^{*}\right|^3 + c_2 n^{-5/2}\sum_{t=1}^{n}\left|x_{t-1}^{*}\right|^3 = O_p^*(1)$$

Remark A.4 We have here used that a Bootstrap invariance principle holds for the term, $n^{-1/2}\sum_1^{[nu]}\left(|z_t^*|-E^*|z_t^*|\right)$, under the conditions in Lemma A.1.

Table 1: Size of the asymptotic and bootstrap tests

		Asymp. Test			Restr. BS			Hybrid BS		
	n	1%	5%	10%	1%	5%	10%	1%	5%	10%
$z_t \sim N$	50	0.6	4.0	9.0	1.0	4.7	9.8	1.0	5.0	10.1
	100	0.5	3.9	8.8	0.9	4.7	9.6	1.0	4.9	10.0
	200	0.6	4.2	9.2	1.0	4.8	9.8	1.0	5.1	9.9
	500	0.8	4.9	10.5	1.1	5.7	10.8	1.1	5.8	10.9
$z_t \sim t$	50	2.1	6.5	11.7	0.8	4.4	9.7	1.8	5.9	10.7
	100	2.8	7.8	13.6	0.9	5.0	10.3	2.0	6.3	11.2
	200	3.2	8.1	14.1	0.9	4.9	9.8	1.8	5.8	10.5
	500	3.9	9.9	16.4	0.9	5.2	10.5	1.4	5.9	10.9
$z_t \sim \chi^2$	50	2.7	7.6	13.2	0.9	4.7	9.8	2.4	6.8	11.8
	100	3.8	9.8	15.3	0.9	5.0	10.8	2.3	7.4	11.9
	200	4.6	11.2	16.9	1.0	5.2	10.8	2.1	6.8	11.7
	500	5.5	12.5	19.2	1.2	5.6	11.0	1.8	6.5	11.5

Notes: The parameter setting under the null is $\pi = 0, \alpha = 0$ and $\omega = 1$. The innovation process (z_t) is drawn, respectively, from standard normal distribution, standardized t distribution with degree of freedom 5.5, and standardized symmetric χ^2 distribution with degree of freedom 3. The results are obtained from 10000 Monte Carlo simulation iterations each of which evaluated using 399 bootstrap samples.

Table 2: Size-adjusted power of the asymptotic and bootstrap tests under the local alternative $\pi = cn^{-1}, \alpha = 0$.

		Asymp. Test			Restr. BS			Hybrid BS		
	n	1%	5%	10%	1%	5%	10%	1%	5%	10%
$z_t \sim N$	50	22.0	61.6	81.2	24.9	63.1	82.1	21.9	61.3	80.5
	100	21.9	60.3	81.0	24.2	62.3	81.0	24.6	60.8	80.3
	200	21.6	58.8	78.8	23.7	59.5	78.9	20.3	58.0	79.0
	500	20.0	55.0	77.2	20.3	54.8	76.3	20.4	54.4	76.1
$z_t \sim t$	50	7.3	45.0	72.9	17.5	53.7	72.8	9.2	41.5	67.9
	100	4.5	35.7	66.0	14.6	44.0	65.5	7.8	34.8	61.5
	200	2.9	33.0	63.7	11.3	38.6	62.3	6.1	33.1	59.9
	500	2.6	26.7	58.2	8.3	33.4	57.4	5.5	28.9	55.2
$z_t \sim \chi^2$	50	4.6	36.2	66.5	15.5	44.6	66.9	5.5	32.0	57.4
	100	4.0	26.2	55.9	11.5	36.5	57.2	7.0	25.9	50.3
	200	2.5	21.8	49.6	7.9	30.6	53.0	4.2	22.6	46.6
	500	1.8	18.5	44.7	4.4	22.7	46.8	2.7	20.4	43.2

Notes: The parameter setting is $c = -10$, and $\omega = 1$. See also notes to Table 1.

Table 3: Raw power of the asymptotic and bootstrap tests under the local alternative $\pi = cn^{-1}$, $\alpha = 0$.

		Asymp. Test			Restr. BS			Hybrid BS		
	n	1%	5%	10%	1%	5%	10%	1%	5%	10%
$z_t \sim N$	50	16.2	54.5	78.2	21.2	60.0	80.7	21.9	59.9	80.5
	100	15.4	53.4	77.6	20.5	59.4	79.7	21.1	59.2	79.5
	200	15.7	53.6	76.8	20.3	57.9	78.2	20.3	58.0	78.2
	500	16.4	54.7	78.5	20.3	57.7	78.5	20.4	57.8	78.5
$z_t \sim t$	50	17.9	55.2	78.3	14.4	49.4	71.3	15.5	47.6	69.6
	100	18.2	54.5	78.4	11.9	42.7	66.3	12.9	42.3	65.2
	200	18.6	54.3	77.7	9.3	37.3	61.5	10.1	37.1	60.9
	500	20.0	57.7	79.7	6.5	33.4	58.2	7.4	33.5	58.0
$z_t \sim \chi^2$	50	17.4	54.0	78.3	12.8	43.4	66.0	14.8	42.0	63.4
	100	18.6	54.8	78.5	9.5	36.5	60.4	11.4	35.8	59.1
	200	19.7	56.0	79.8	6.2	30.6	55.3	7.8	30.8	54.9
	500	21.8	57.4	79.6	4.4	25.2	50.1	5.3	25.4	49.7

Notes: The parameter setting is $c = -10$ and $\omega = 1$. See also notes to Table 1.

Table 4: Size-adjusted power of the asymptotic and bootstrap tests under the local alternative $\pi = 0, \alpha = cn^{-3/2}$.

		Asymp. Test			Restr. BS			Hybrid BS		
	n	1%	5%	10%	1%	5%	10%	1%	5%	10%
$z_t \sim N$	50	24.3	39.6	50.3	21.5	38.2	50.0	24.5	40.2	50.6
	100	29.1	43.7	54.6	27.8	43.6	54.6	30.3	44.4	55.0
	200	32.4	46.9	56.9	31.9	46.6	57.0	31.4	46.6	57.4
	500	34.8	48.6	59.1	34.3	48.3	58.6	35.0	48.8	58.7
$z_t \sim t$	50	17.0	32.3	43.2	16.3	33.0	43.5	17.9	32.1	43.5
	100	18.8	34.7	45.1	20.6	35.4	45.4	20.1	34.0	45.0
	200	20.3	37.7	48.2	23.9	38.7	48.2	21.7	37.3	48.0
	500	22.7	38.2	48.7	26.1	39.7	49.0	24.7	38.5	48.5
$z_t \sim \chi^2$	50	15.8	30.6	40.5	15.6	30.8	41.1	14.6	29.9	40.2
	100	18.4	31.9	41.4	18.8	33.2	42.2	19.4	31.6	41.3
	200	19.0	34.1	43.4	20.8	35.2	44.3	19.8	33.1	42.8
	500	21.1	35.9	45.9	22.3	36.3	46.2	20.5	36.1	45.4

Notes: The parameter setting is $c = 10$ and $\omega = 1$. See also notes to Table 1.

Table 5: Raw power of the asymptotic and bootstrap tests under the local alternative $\pi = 0, \alpha = cn^{-3/2}$.

		Asymp. Test			Restr. BS			Hybrid BS		
	n	1%	5%	10%	1%	5%	10%	1%	5%	10%
$z_t \sim N$	50	21.7	36.4	48.4	19.6	36.8	49.1	24.5	39.6	50.6
	100	25.9	41.2	52.6	25.9	42.4	53.7	28.9	43.8	54.6
	200	29.4	44.6	55.9	30.1	45.8	56.5	31.4	46.6	56.8
	500	33.4	48.5	59.8	34.3	49.8	59.8	35.0	50.1	60.0
$z_t \sim t$	50	23.0	35.9	46.2	14.7	30.8	42.5	21.6	34.7	44.2
	100	27.6	40.8	50.3	18.8	34.8	45.8	24.1	37.1	46.7
	200	31.6	44.9	54.3	22.2	38.1	47.9	25.6	39.1	48.4
	500	35.6	48.5	57.8	24.4	39.7	49.4	26.6	40.6	49.7
$z_t \sim \chi^2$	50	23.5	36.1	45.2	14.1	30.3	40.6	21.9	34.0	42.3
	100	28.8	41.2	50.1	17.1	33.2	43.5	22.9	35.8	44.5
	200	33.1	45.1	53.3	19.2	35.2	45.1	23.5	36.8	45.8
	500	37.3	49.9	58.0	22.3	37.7	47.5	24.8	38.5	48.1

Notes: The parameter setting is $c = 10$ and $\omega = 1$. See also notes to Table 1.

Figure 1: Size-adjusted power of the asymptotic and bootstrap LR tests under the alternative $\pi < 0, \alpha = 0$ for different values of $-\pi$. Upper panel: $n = 50$; lower panel: $n = 100$. Nominal levels 1% (left panels), 5% (central panels) and 10% (right panels)

Figure 2: Size-adjusted power of the asymptotic and bootstrap LR tests under the alternative $\pi = 0, \alpha > 0$ for different values of α. Upper panel: $n = 50$; lower panel: $n = 100$. Nominal levels 1% (left panels), 5% (central panels) and 10% (right panels)