AN INTRODUCTION TO BOOTSTRAP THEORY IN TIME SERIES ECONOMETRICS

Giuseppe Cavaliere*, Heino Bohn Nielsen†, and Anders Rahbek†

Abstract: This article provides an introduction to methods and challenges underlying application of the bootstrap in econometric modelling of economic and financial time series. Validity, or asymptotic validity, of the bootstrap is discussed as this is a key element in deciding whether the bootstrap is applicable in empirical contexts. That is, as detailed here, bootstrap validity relies on regularity conditions, which need to be verified on a case-by-case basis. To fix ideas, asymptotic validity is discussed in terms of the leading example of bootstrap-based hypothesis testing in the well-known first order autoregressive model. In particular, bootstrap versions of classic convergence in probability and distribution, and hence of laws of large numbers and central limit theorems, are discussed as crucial ingredients to establish bootstrap validity. Regularity conditions and their implications for possible improvements in terms of (empirical) size and power for bootstrap-based testing, when compared to asymptotic testing, are illustrated by simulations. Following this, an overview of selected recent advances in the application of bootstrap methods in econometrics is also given.

1 Introduction

Bootstrap in econometrics is frequently applied in the context of estimationa and testing, see e.g. Berkowitz and Kilian (2000), Cavaliere and Rahbek (2020), Davidson and MacKinnon (2006), Horowitz (2001, 2003) and MacKinnon (2009). As an example, consider the case where some test statistic, τ_n say, is of interest given a sample of (time-series) data $x_1, ..., x_n$ with initial values $x_0, x_{-1}, .., x_{-p}$ for $p \geq 0$, $\{x_t\}_{t=-p}^{n}$. Under suitable regularity conditions, including typically *(i)* stationarity and ergodicity of the x_t process, and, *(ii)*

*Department of Economics, University of Bologna, Italy and University of Exeter, United Kingdom.
†Department of Economics, University of Copenhagen, Denmark. This research was supported by Danish Council for Independent Research (DSF Grant 7015-00028B). Correspondence to: Anders Rahbek, Department of Economics, University of Copenhagen, Øster Farimagsgade 5, 1353 Copenhagen K, Denmark; email: anders.rahbek@econ.ku.dk.

finite moments conditions on the form $E|x_t^2|^k < \infty$ for some $k \geq 1$, it holds that under the null hypothesis of interest, $\mathcal{H}_0$ say,

$$\tau_n \xrightarrow{d} \chi_q^2 \text{ as } n \to \infty,$$

where q denotes the degrees of freedom and "$\xrightarrow{d}$" denotes convergence in distribution. Moreover, under the alternative, or when $\mathcal{H}_0$ is not true, the statistic τ_n diverges. In standard asymptotic testing, the χ_q^2 distributional approximation of the test statistic τ_n is frequently applied using the $1 - \alpha$ quantile of the χ_q^2 distribution or, equivalently, by calculating the p-value at the nominal level α.

A bootstrap-based test is often motivated by noting that the asymptotic approximation may not be good in finite samples, which may lead, for a finite number of observations n, to an actual size larger, or smaller, than the nominal level α. This is often corrected by applying even simple bootstrap schemes. An additional motivation for the bootstrap is that the underlying model used for estimation – and on which the derivation of τ_n is based – may be misspecified. A typical example is the assumption of homoskedasticity of model innovations, which often in practice is challenged in the modelling of macro and financial data, where (conditional and unconditional) heteroskedasticity is typically present. In this case, the so-called *wild bootstrap* may correct for such heteroskedasticity when doing inference which would otherwise be affected. Likewise, for other types of misspecification and model departures, different bootstrap schemes may be applied depending on which type of misspecification is of concern. In addition, in some cases the limiting distribution of τ_n cannot be tabulated, e.g. because it is a function of unknown nuisance parameters. Furthermore, the limiting distribution may even not exist in certain non-standard testing problems. The bootstrap may resolve such issues. However, it should be emphasized that *any* bootstrap based test is – as for the original asymptotic test – only valid under certain regularity conditions. These regularity conditions are important to check and understand as in many cases application of what may be thought of as a "standard bootstrap" may be invalid and misleading, despite its popularity in many empirical applications. Typical situations when a standard bootstrap does not work can be found in the context of non-stationary variables and when data exhibit heavy tailed distributions, as typical in financial data. When application of existing, or standard, bootstrap schemes fails to work, these may be corrected by more elaborate bootstrap schemes, as is richly documented in recent research in econometrics, see *inter alia* Cavaliere and Rahbek (2020) and the references therein.

The bootstrap is a simulation-based approach which is typically simple to apply in the context of inference and testing. Based on some bootstrap scheme, the bootstrap is based on generating new bootstrap samples of data and the test statistic of interest, denoted by $\{x_t^*\}_{t=-p}^n$ and τ_n^* respectively. The (re-)generation of such bootstrap data is based on keeping the original sample $\{x_t\}_{t=-p}^n$ fixed.

If the limiting distribution (in the bootstrap sense) of the bootstrap test statistic τ_n^* is identical to the limiting distribution of the original statistic τ_n under the null of $\mathcal{H}_0$, such as the χ^2_q distribution above, this implies validity under the null of the bootstrap. If, in addition, under the alternative when $\mathcal{H}_0$ is not true, the bootstrap statistic τ^*_n does not diverge as the sample size n increases (or, at most it diverges at a slower rate than the original statistic), the bootstrap is asymptotically valid. This is in particular the case, when under the alternative, the bootstrap limiting distribution of τ^*_n is identical to the limiting distribution of the original statistic, or when τ^*_n is bounded in (bootstrap) probability.

As a leading example throughout, and to fix ideas, we consider the first-order autoregressive (AR) model. Section 2 provides a summary of standard econometric *non-bootstrap* analysis of the AR model before turning to the discussion of the implementation and (asymptotic) theory of the bootstrap in Section 3. Next, a Monte Carlo study is used to discuss finite sample behavior of different bootstrap schemes in Section 4, and Section 5 contains selected recent advances in econometric time series bootstrap analysis. Section 6 provides an overview of some important further topics and approaches to bootstrap inference. The appendix contains some technical details.

2 The autoregressive model

In order to present bootstrap theory and arguments, we present here a brief summary of results and arguments for the (non-bootstrap) AR model of order one. The results and asymptotic arguments presented here are well-known, and details can be found several places, see *e.g.* Hamilton (1994) and Hayashi (2000).

Thus, consider the AR model of order one

$$x_t = \rho x_{t-1} + \varepsilon_t, \quad t = 1, 2, ..., n, \tag{2.1}$$

with the initial value x_0 fixed in the statistical analysis and ε_t i.i.d. $\mathrm{N}(0, \sigma^2)$, see also Remark 2.1 regarding relaxing the assumption of Gaussianity. The parameter of the model is $\theta = (\rho, \sigma^2)'$ with $\theta_0 = (\rho_0, \sigma_0^2)'$ denoting the true value. For estimation, the parameter space is given by

$$\theta \in \Theta = \mathbb{R} \times (0, \infty).$$

For the true parameter θ_0 we assume $\theta_0 \in \Theta_0 \subset \Theta$, with $\Theta_0 = \{\theta \in \Theta \mid |\rho_0| < 1\}$, such that x_t in (2.1) for $\theta_0 \in \Theta_0$ has a stationary and geometrically ergodic solution given by

$$x_t = \sum_{i=0}^{\infty} \rho_0^i \varepsilon_{t-i}.$$

With $\bar{\rho}$ some fixed value, consider testing the hypothesis $\mathcal{H}_0$ given by

$$\mathcal{H}_0 : \rho = \bar{\rho}.$$

The test is based on the likelihood ratio test statistic, τ_n, defined in terms of the Gaussian log-likelihood function

$$\ell_n(\theta) = -\frac{n}{2}(\log \sigma^2 + n^{-1} \sum_{t=1}^{n} \varepsilon_t^2(\rho) / \sigma^2), \tag{2.2}$$

with x_0 fixed, and $\varepsilon_t(\rho) = x_t - \rho x_{t-1}$. Standard optimization gives the unrestricted maximum likelihood estimator (MLE), $\hat{\theta}_n = (\hat{\rho}_n, \hat{\sigma}_n^2)'$, where

$$\hat{\rho}_n = n^{-1}\sum_{t=1}^{n} x_t x_{t-1}(n^{-1}\sum_{t=1}^{n} x_{t-1}^2)^{-1} \quad \text{and} \quad \hat{\sigma}_n^2 = n^{-1}\sum_{t=1}^{n}\varepsilon_t^2(\hat{\rho}_n). \tag{2.3}$$

The restricted estimator $\tilde{\theta}_n = (\tilde{\rho}_n, \tilde{\sigma}_n^2)'$ is obtained by maximization under $\mathcal{H}_0$, and is given by

$$\tilde{\rho}_n = \bar{\rho} \quad \text{and} \quad \tilde{\sigma}_n^2 = n^{-1}\sum_{t=1}^{n}\varepsilon_t^2(\bar{\rho}).$$

It follows that the likelihood ratio (LR) test statistic is given by

$$\tau_n = \text{LR}_n(\rho = \bar{\rho}) = 2(\ell_n(\hat{\theta}_n) - \ell_n(\tilde{\theta}_n)) = n\log\left(\tilde{\sigma}_n^2/\hat{\sigma}_n^2\right). \tag{2.4}$$

Moreover, under $\mathcal{H}_0$ and regularity conditions detailed in Section 2.1, with $\bar{\rho} = \rho_0$ and $\theta_0 \in \Theta_0$, it holds that $\tau_n \xrightarrow{d} \chi_1^2$.

Note in this respect that, as long as the true ρ_0 is not "too large" the asymptotic χ_1^2 approximation is a good approximation even for small samples. In contrast, and as exemplified in Table 1 in Section 4, if either ε_t is not i.i.d. $\text{N}(0, \sigma^2)$, or if $\rho_0 = 0.9$, the χ_1^2 distribution is not a good approximation of the actual distribution of τ_n for small, or even moderate, finite samples of size n.

Remark 2.1 (Quasi likelihood) *Often the assumption that the ε_t sequence is assumed to be i.i.d. Gaussian is relaxed to i.i.d. $(0, \sigma^2)$ for some unknown distribution. When the Gaussian log-likelihood function in (2.2) in this case is used to obtain estimators and test statistics, these are referred to as (Gaussian) quasi MLE and quasi LR statistics respectively. This is also the case when ε_t is assumed to be mean zero and (conditionally or unconditionally) heteroskedastic, as for example with some autoregressive conditional heteroskedastic (ARCH) specification or, as in Section 4, with a structural break in the variance.*

2.1 Asymptotics for the autoregressive model

To ease the presentation of the key arguments for the standard non-bootstrap asymptotic analysis, assume here without loss of generality that σ^2 is fixed at the true value, that is, $\sigma^2 = \sigma_0^2$. Accordingly, the parameter is $\theta = \rho$ and the likelihood function in (2.2) simplifies as

$$\ell_n(\theta) = \ell_n(\rho) = -\frac{n}{2}(\log\sigma_0^2 + n^{-1}\sum_{t=1}^{n}\varepsilon_t^2(\rho)/\sigma_0^2).$$

The MLE $\hat{\rho}_n$ is given by (2.3), while the $LR_n(\rho = \bar{\rho})$ statistic in (2.4) in this case is identical to the Wald statistic (W_n)

$$\tau_n = W_n(\rho = \bar{\rho}) = (\hat{\rho}_n - \bar{\rho})^2\sum_{t=1}^{n} x_{t-1}^2/\sigma_0^2. \tag{2.5}$$

By construction, τ_n is a simple function of $n^{1/2}(\hat{\rho}_n - \bar{\rho})$ and $n^{-1}\sum_{t=1}^{n} x_{t-1}^2$, and the limiting distribution is found by applying a law of large numbers (LLN) to the average $n^{-1}\sum_{t=1}^{n} x_{t-1}^2$, as well as a central limit theorem (CLT) to $n^{1/2}(\hat{\rho}_n - \bar{\rho})$.

With $\bar{\rho} = \rho_0 \in \Theta_0$, the AR process x_t is stationary and ergodic, and has finite variance, $\mathrm{V}(x_t) = \omega_0 = \sigma_0^2/(1-\rho_0^2)$. In particular, it follows by the LLN for stationary and ergodic processes that

$$n^{-1}\sum_{t=1}^{n} x_{t-1}^2 \xrightarrow{p} \mathrm{V}(x_t) = \omega_0.$$

Next, by definition of the MLE,

$$n^{1/2}(\hat{\rho}_n - \rho_0) = n^{-1/2}\sum_{t=1}^{n} \varepsilon_t x_{t-1}(n^{-1}\sum_{t=1}^{n} x_{t-1}^2)^{-1}.$$

Here $m_t = \varepsilon_t x_{t-1}$ is a martingale difference sequence (mds) with respect to the filtration $\mathcal{F}_t$, where $\mathcal{F}_t = \sigma(x_t, x_{t-1}...)$, as $E|m_t| < \infty$, and $E(m_t|\mathcal{F}_{t-1}) = 0$. Moreover, the conditional second order moment converges in probability,

$$n^{-1}\sum_{t=1}^{n} \mathrm{E}\left(m_t^2|\mathcal{F}_{t-1}\right) = \sigma_0^2 n^{-1}\sum_{t=1}^{n} x_{t-1}^2 \xrightarrow{p} \sigma_0^2\omega_0.$$

This implies by the CLT for mds that $n^{-1/2}\sum_{t=1}^{n} m_t \xrightarrow{d} \mathrm{N}(0, \sigma_0^2\omega_0)$, see e.g. Hamilton (1994) and Hall and Heyde (1980), such that

$$n^{1/2}(\hat{\rho}_n - \rho_0) \xrightarrow{d} \mathrm{N}\left(0, \sigma_0^2/\omega_0\right) \overset{d}{=} \mathrm{N}\left(0, 1-\rho_0^2\right).$$

Remark 2.2 (Lindeberg condition) *Note that for a CLT for mds in general to hold, see e.g. Hall and Heyde (1980), also a Lindeberg-type condition of the form,* $\gamma_n = n^{-1}\sum_{t=1}^{n} E\left(m_t^2 \mathbb{I}\left(|m_t| > \delta n^{1/2}\right)\right) \to 0$ *for any* $\delta > 0$*, must hold, where* $\mathbb{I}(\cdot)$ *is the indicator function. This holds here as, by stationarity,*

$$\gamma_n = n^{-1}\sum_{t=1}^{n} \mathrm{E}\left(m_t^2 \mathbb{I}\left(|m_t| > \delta n^{1/2}\right)\right) = \mathrm{E}\left(m_t^2 \mathbb{I}\left(|m_t| > \delta n^{1/2}\right)\right),$$

which tends to zero as $n \to \infty$ *by dominated convergence under the moment condition* $\mathrm{E}(m_t^2) < \infty$ *as implied by* $\mathrm{E}(\varepsilon_t^2) < \infty$ *and independence of* ε_t *and* x_{t-1}*.*

Collecting terms it follows that we have the following result.

Lemma 2.1 *For the AR(1) model in (2.1) with* $\theta_0 \in \Theta_0$ *and* ε_t *i.i.d.* $(0, \sigma_0^2)$*, it follows that with* τ_n *given by (2.5), under the null* $\tau_n \xrightarrow{d} \chi_1^2$ *as* $n \to \infty$*.*

Note that in Lemma 2.1, it is not assumed that the ε_t sequence is Gaussian, but instead the lemma is formulated in terms of the milder sufficient condition that ε_t is an i.i.d. $(0, \sigma^2)$ sequence, see also Remark 2.1. Thus the essential regularity conditions for the lemma to hold are that $|\rho_0| < 1$ and ε_t i.i.d. $(0, \sigma^2)$.

Remark 2.3 (Heteroskedasticity) *Note that if* $\varepsilon_t = \sigma_t z_t$, *with* z_t *i.i.d.* $\mathrm{N}(0,1)$ *and* σ_t^2 *given by an ARCH process with* $\mathrm{E}(\sigma_t^2) = \sigma_\varepsilon^2 < \infty$, *it follows that* $\tau_n \xrightarrow{d} c\chi_1^2$, *with* $c = \sigma_\varepsilon^2/\sigma_0^2$. *This reflects in general that if* ε_t *are (conditionally, or unconditionally) heteroskedastic, the limiting distribution of the test statistic* τ_n *is not* χ_1^2. *In the context of regression models this is well-known, and corrections of the test statistic, see e.g. White (1980), are typically applied to ensure valid asymptotic inference; the bootstrap, or more precisely, the wild bootstrap discussed later, may correct for such misspecification without the need to correct the original test statistic, cf. Gonçalves and Kilian (2004).*

Remark 2.4 (Test statistic) *For* σ^2 *an unknown parameter to be estimated, such that* $\theta = (\sigma^2, \rho)'$, *note that the LR statistic in (2.4) can be written as*

$$\tau_n = LR_n(\rho = \bar{\rho}) = -n \log(1 - n^{-1} W_n(\rho = \bar{\rho}) \frac{\sigma_0^2}{\tilde{\sigma}_n^2}),$$

with $\tilde{\sigma}_n^2 = n^{-1} \sum_{t=1}^n \varepsilon_t^2(\bar{\rho})$. *By Lemma 2.1,* $w_n = n^{-1} W_n(\rho = \bar{\rho}) \xrightarrow{p} 0$, *and by a Taylor expansion,* $-\log(1-w) = w + o(w)$, *with* $o(w)$ *a term which tends to zero as* $w \to 0$. *Hence,*

$$\tau_n = W_n(\rho = \bar{\rho}) + o_p(1),$$

where $o_p(1)$ *denotes a term which converges to zero in probability as* $n \to \infty$*; see* e.g. *van der Vaart (2000, Lemma 2.12) for further details on* $o_p(\cdot)$ *(and* $O_p(\cdot)$*) notation, and the stochastic (Taylor) expansion as applied here.*

3 Bootstrap in the autoregressive model

The general idea behind a bootstrap algorithm as implemented in the context of likelihood-based testing in the first order AR model can be summarized as follows:

Step $\mathcal{A}$ With the original data $\{x_t\}_{t=0}^n$ fixed, generate a sample of bootstrap data $\{x_t^*\}_{t=0}^n$ using some *bootstrap scheme* with bootstrap true parameter, $\theta_n^* = (\rho_n^*, \sigma_n^{*2})'$. Specifically, for the AR model, set $x_0^* = x_0$ and generate x_t^* recursively by

$$x_t^* = \rho_n^* z_t + \varepsilon_t^*, \tag{3.1}$$

with $z_t = x_{t-1}^*$ for a *recursive* bootstrap scheme, while $z_t = x_{t-1}$ for a *fixed design* bootstrap scheme. As to the choice of bootstrap innovations ε_t^* in terms of the original data $\{x_t\}_{t=0}^n$ and a bootstrap sampling distribution, this is detailed below.

Step $\mathcal{B}$ Compute the bootstrap (quasi) MLE $\hat{\rho}_n^*$ and the bootstrap LR statistic $\tau_n^* = LR_n^*(\rho = \rho_n^*)$, where the bootstrap log-likelihood function $\ell_n^*(\theta)$ as a function of θ is given by

$$\ell_n^*(\theta) = -\frac{n}{2}(\log \sigma^2 + n^{-1} \sum_{t=1}^n \varepsilon_t^{*2}(\rho)/\sigma^2), \quad \text{with } \varepsilon_t^*(\rho) = x_t^* - \rho z_t. \tag{3.2}$$

Step $\mathcal{C}$ Generate $\left\{x_{t,b}^{*}\right\}_{t=0}^{n}$, $\hat{\theta}_{n,b}^{*}$ and $\tau_{n,b}^{*}$ for $b = 1, 2, ..., B$ by repeating Steps $\mathcal{A}$ and $\mathcal{B}$, and use the empirical distribution of $\left\{\tau_{n,b}^{*}\right\}_{b=1}^{B}$ for testing based on the original statistic $\tau_n = LR_n(\rho = \rho_0)$. Precisely, with the bootstrap p-value set to $p_{n,B}^{*} = B^{-1}\sum_{b=1}^{B}\mathbb{I}(\tau_n \geq \tau_{n,b}^{*})$, $\mathcal{H}_0$ is rejected when $p_{n,B}^{*} < \alpha$, with α the nominal level. For the choice of bootstrap repetitions B, see Remark 3.3.

As to the definition in Step $\mathcal{A}$ of the bootstrap innovations ε_t^{*}, usually these are obtained by i.i.d. draws with replacement from re-centered estimated model residuals; henceforth referred to as *iid bootstrap*, or iid resampling. The model residuals can be either estimated under the null ($\mathcal{H}_0$ imposed), or without imposing the null. That is, the ε_t^{*} are in the first case resampled from centered residuals, $\{\tilde{\varepsilon}_t^{c}\}_{t=1}^{n}$, where

$$\tilde{\varepsilon}_t^{c} = \varepsilon_t(\bar{\rho}) - n^{-1}\sum_{t=1}^{n}\varepsilon_t(\bar{\rho}), \quad \text{with } \varepsilon_t(\bar{\rho}) = x_t - \bar{\rho}x_{t-1}. \tag{3.3}$$

Using unrestricted residuals, the bootstrap ε_t^{*} innovations are resampled from $\{\hat{\varepsilon}_t^{c}\}_{t=1}^{n}$, where

$$\hat{\varepsilon}_t^{c} = \varepsilon_t(\hat{\rho}_n) - n^{-1}\sum_{t=1t}^{n}\varepsilon_t(\hat{\rho}_n), \quad \text{with } \varepsilon_t(\hat{\rho}_n) = x_t - \hat{\rho}_n x_{t-1}. \tag{3.4}$$

Centering is required because by doing so, both of the residual series $\hat{\varepsilon}_t^{c}$ and $\tilde{\varepsilon}_t^{c}$ have empirical mean zero, $n^{-1}\sum_{t=1}^{n}\hat{\varepsilon}_t^{c} = 0$ and $n^{-1}\sum_{t=1}^{n}\tilde{\varepsilon}_t^{c} = 0$, and ideally "mimic" the true ε_t. In particular, under $\mathcal{H}_0$, $\varepsilon_t(\bar{\rho}) = \varepsilon_t(\rho_0) = \varepsilon_t$, and, moreover, $\hat{\rho}_n \xrightarrow{p} \rho_0$. In contrast, under the alternative, when $\mathcal{H}_0$ does not hold, while $\varepsilon_t(\bar{\rho}) \neq \varepsilon_t$, it still holds that $\hat{\rho}_n \xrightarrow{p} \rho_0$. Hence, while one may expect that the unrestricted residuals $\hat{\varepsilon}_t^{c}$ perform better under the alternative, at least asymptotically, in practice little difference is found between the two.

An alternative to iid resampling is the so-called *wild bootstrap*, where either $\varepsilon_t^{*} = \tilde{\varepsilon}_t^{c}w_t^{*}$, or $\varepsilon_t^{*} = \hat{\varepsilon}_t^{c}w_t^{*}$, with w_t^{*} an auxiliary i.i.d. sequence, independent of the original data, with $\mathrm{E}(w_t^{*}) = 0$ and $\mathrm{V}(w_t^{*}) = 1$. A simple example is w_t^{*} i.i.d. $\mathrm{N}(0,1)$, see Remark 3.4 below for alternative specifications. The wild bootstrap is typically motivated by the its potential ability to allow for possible model misspecification in the sense that it allows for (conditional and unconditional) heteroskedasticity in the innovations ε_t. To see this for e.g. $\varepsilon_t^{*} = \tilde{\varepsilon}_t^{c}w_t^{*}$, it follows that (conditionally on the data $\{x_t\}_{t=0}^{n}$), the ε_t^{*} are i.i.d. distributed with mean zero, and with time-varying variance,

$$\mathrm{V}\left(\varepsilon_t^{*} \mid \{x_t\}_{t=0}^{n}\right) = \left(\tilde{\varepsilon}_t^{c}\right)^2.$$

This way, the variation in ε_t^{*} from the wild bootstrap "*in essence reflects the heteroskedasticity of the original data*" (Liu, 1988, p.1704) as is also illustrated in the Monte Carlo simulations in Section 4. This differs from the iid bootstrap, where the bootstrap innovations ε_t^{*} (conditionally on the data) are i.i.d., with a discrete distribution given by $P\left(\varepsilon_t^{*} = \tilde{\varepsilon}_j^{c} \mid \{x_t\}_{t=0}^{n}\right) = n^{-1}$, for $j = 1, 2, ..., n$. Specifically, as mentioned earlier, the (empirical, or conditional) mean is zero, $\mathrm{E}\left(\varepsilon_t^{*} \mid \{x_t\}_{t=0}^{n}\right) = 0$, and the variance is given by

$$\mathrm{V}\left(\varepsilon_t^{*} \mid \{x_t\}_{t=0}^{n}\right) = n^{-1}\sum_{t=1}^{n}\left(\tilde{\varepsilon}_t^{c}\right)^2.$$

That is, for fixed n, the variance for the iid bootstrap innovations ε_t^* is constant (and, equal to the empirical variance of the estimated residuals), while depending on time t for the wild bootstrap.

In short, the bootstrap scheme in (3.1), $x_t^* = \rho_n^* z_t + \varepsilon_t^*$, with $z_t = x_{t-1}$ or $z_t = x_{t-1}^*$, depends on two types of randomness: *(i)* the variation of the original data $\{x_t\}_{t=0}^n$; and, *(ii)* the bootstrap re-sampling for $\{\varepsilon_t^*\}_{t=1}^n$ (wild or iid resampling). This is important for the application of the LLN and the CLT to establish bootstrap validity as demonstrated in the next.

Some further remarks are in order.

Remark 3.1 ((Un-)restricted bootstrap) *As to the choice in Step $\mathcal{A}$ of the bootstrap true value $\theta_n^* = (\rho_n^*, \sigma_n^{*2})'$ (or simply, ρ_n^* for the AR model), one may set θ_n^* to the value of the unrestricted estimator $\hat{\theta}_n$, $\theta_n^* = \hat{\theta}_n^*$, which is referred to as unrestricted bootstrap. If $\theta_n^* = \tilde{\theta}_n$, the restricted estimator, the bootstrap is referred to as restricted. It should be emphasized that for the unrestricted bootstrap, the bootstrap likelihood ratio statistic τ_n^* is derived for the hypothesis $\rho = \hat{\rho}_n$, while for the restricted bootstrap, the original hypothesis $\mathcal{H}_0 : \rho = \bar{\rho}$ is considered. While both choices are widely applied in existing literature, the restricted bootstrap in the context of testing is more popular in econometrics, see for example Davidson and MacKinnon (2000).*

Remark 3.2 (Recursive and fixed design bootstrap) *In Step $\mathcal{B}$ of the algorithm in equation (3.1), with $z_t = x_{t-1}^*$, this is an example of a recursive bootstrap. That is, the original autoregressive structure for x_t is replicated for the bootstrap process x_t^*. On the other hand, with $z_t = x_{t-1}$, the original data x_{t-1} are used as lagged value of x_t^*, such that x_t^* is not an autoregressive process even conditionally on the data. The fixed design bootstrap typically simplifies (some of) the asymptotic arguments, and is often found to behave as well as the recursive bootstrap, see for example Gonçalves and Kilian (2004, 2007) for general AR models, and Cavaliere, Pedersen and Rahbek (2018) for ARCH models. Section 3.1 below provides a detailed analysis of the recursive bootstrap.*

Remark 3.3 (Bootstrap p-value) *In Step $\mathcal{C}$, the bootstrap p-value $p_{n,B}^*$ is defined as $p_{n,B}^* = B^{-1}\sum_{b=1}^{B}\mathbb{I}(\tau_n \geq \tau_{n,b}^*)$ with B the number of bootstrap repetitions, see also Remark 3.8. Typical choices are $B = 199$, 399 or 999, see also Andrews and Buchinsky (2000) and Davidson and MacKinnon (2000) for details on the choice of B.*

Remark 3.4 (Choice of w_t^*) *With respect to the choice of distribution of the i.i.d. sequence w_t^* for the wild bootstrap, Liu (1988) provides a detailed discussion of various choices based on so-called Edgeworth expansions (see, e.g. Hall, 1992, and van der Vaart, 2000, for an introduction) of test statistics similar to τ_n. In particular, Liu (1988), with $\xi_k = \mathrm{E}(w_t^k)$, $k \geq 1$, emphasizes $\xi_1 = 0$, $\xi_2 = 1$ as well as $\xi_3 = 1$ as important. For the case of $\xi_3 = 1$, emphasis is on possible skewness of the test statistic, while $\xi_3 = 0$ works well in the case of symmetry. In applications, standard choices for w_t^*, all with $\xi_1 = 0$*

and $\xi_2 = 1$, include the Gaussian, Rademacher and Mammen distributions.[1] *It follows that $\xi_3 = 0$ for the first two (with $\xi_4 = 3$ and 1, respectively), while $\xi_3 = 1$ (and $\xi_4 = 2$) for the Mammen distribution.*

Remark 3.5 (Parametric bootstrap) *In Step $\mathcal{A}$ of the bootstrap scheme, one may also use a so-called parametric bootstrap, where the bootstrap innovations ε_t^* are generated as i.i.d. $\mathrm{N}(0, \sigma_n^{*2})$. While this parametric bootstrap performs well in the case where the true innovations ε_t are Gaussian, this may not be the case when the distribution of ε_t is non-Gaussian, see also Horowitz (2001).*

3.1 Asymptotic theory for the recursive bootstrap

In order to discuss regularity conditions under which bootstrap-based testing holds, we consider here the details of verification of the asymptotic validity of the recursive bootstrap for the AR model.

Thus, with the AR model given in (2.1), consider here the recursive restricted bootstrap scheme as defined by setting $z_t = x_{t-1}^*$ and $\rho_n^* = \bar{\rho}$ in Step $\mathcal{A}$ of the bootstrap algorithm. In short, the x_t^* bootstrap sequence is here generated as

$$x_t^* = \bar{\rho} x_{t-1}^* + \varepsilon_t^*, \tag{3.5}$$

with $x_0^* = x_0$ and $\bar{\rho} = \rho^*$ as the bootstrap true value. Moreover, we consider the classic case of iid resampling from the autoregressive residuals obtained under $\mathcal{H}_0$, that is, from $\{\tilde{\varepsilon}_t^c\}_{t=1}^n$ defined in (3.3). The statistic of interest is $\tau_n = W_n(\rho = \bar{\rho})$ in (2.5) which is computed using the bootstrap sample $\{x_t^*\}_{t=0}^n$ as

$$\tau_n^* = W_n^*(\rho = \bar{\rho}) = (\hat{\rho}_n^* - \bar{\rho})^2 \sum_{t=1}^n \left(x_{t-1}^*\right)^2 / \sigma_0^2 \quad \text{and} \quad \hat{\rho}_n^* = \sum_{t=1}^n x_t^* x_{t-1}^* (\sum_{t=1}^n x_{t-1}^{*2})^{-1}. \tag{3.6}$$

While x_t^* clearly has some features similar to x_t, one cannot apply standard concepts such as stationarity and ergodicity when analyzing the asymptotic behavior of τ_n^*, as we as have two types of randomness: the bootstrap resampling distribution, and the distribution of the original data, $\{x_t\}_{t=0}^n$. We therefore introduce the bootstrap equivalent concepts of convergence in probability and distribution, which reflects the fact that inference is based on conditioning on the original data, which are themselves random.

Bootstrap probability, expectation and convergence

With $\mathrm{P}^*(\cdot)$ denoting the bootstrap probability, that is, the probability conditional on the data, the iid bootstrap innovations ε_t^* are by definition i.i.d. distributed with

$$\mathrm{P}^*(\varepsilon_t^* = \tilde{\varepsilon}_j^c) = \mathrm{P}(\varepsilon_t^* = \tilde{\varepsilon}_j^c | \{x_t\}_{t=0}^n) = n^{-1} \quad \text{for} \quad j = 1, 2, ..., n.$$

[1] The Rademacher distribution is a two-point distribution on ± 1, each with probability a half, while the Mammen distribution is a two-point distribution on $\frac{1+\sqrt{5}}{2}$ and $\frac{1-\sqrt{5}}{2}$, with probabilities given by $\frac{\sqrt{5}-1}{2\sqrt{5}}$ and $\frac{\sqrt{5}+1}{2\sqrt{5}}$ respectively.

Similarly, the expectation $\mathrm{E}^*(\cdot)$ is defined by $\mathrm{E}^*(\cdot) = \mathrm{E}(\cdot|\{x_t\}_{t=0}^n)$. As an example, consider the expectation of ε_t^* conditionally on the data. It follows that, as already discussed, $\mathrm{E}^*(\varepsilon_t^*) = 0$ as

$$\mathrm{E}^*(\varepsilon_t^*) = \sum_{j=1}^n \mathrm{P}^*(\varepsilon_t^* = \tilde{\varepsilon}_j^c)\tilde{\varepsilon}_j^c = n^{-1}\sum_{j=1}^n \tilde{\varepsilon}_j^c = 0,$$

by the definition in (3.3). Next, consider the variance of ε_t^* conditionally on the data, $\mathrm{V}^*(\varepsilon_t^*)$. Again, by definition, $\mathrm{V}^*(\varepsilon_t^*) = \mathrm{E}^*(\varepsilon_t^{*2}) - (\mathrm{E}^*(\varepsilon_t^*))^2$ and hence as $\mathrm{E}^*(\varepsilon_t^*) = 0$, it follows that

$$\mathrm{V}^*(\varepsilon_t^*) = \sum_{j=1}^n \mathrm{P}^*(\varepsilon_t^* = \tilde{\varepsilon}_j^c)\left(\tilde{\varepsilon}_j^c\right)^2 = n^{-1}\sum_{j=1}^n \left(\tilde{\varepsilon}_j^c\right)^2.$$

That is, the variance conditional on the data is equal to the sample variance of the original estimated residuals under $\mathcal{H}_0$. In particular, $\mathrm{V}^*(\varepsilon_t^*)$ is a random variable (in terms of the original probability measure) and moreover, by the LLN for i.i.d. variables, under $\mathcal{H}_0$ with $\bar{\rho} = \rho_0$,

$$\mathrm{V}^*(\varepsilon_t^*) = \mathrm{E}^*\left(\varepsilon_t^{*2}\right) \xrightarrow{p} \mathrm{V}(\varepsilon_t) = \sigma_0^2. \tag{3.7}$$

Note that for the wild bootstrap, $\mathrm{V}^*(\varepsilon_t^*) = \mathrm{V}^*(\tilde{\varepsilon}_t^c w_t^*) = (\tilde{\varepsilon}_t^c)^2$, emphasizing that the wild bootstrap indeed "mimics" heteroskedasticity, while the iid bootstrap does not.

Similar to the definition of $\mathrm{P}^*(\cdot)$ and $\mathrm{E}^*(\cdot)$ this motivates the definition of the bootstrap equivalent of convergence in probability. Formally, a sequence of stochastic variables X_n^* is said to converge in probability conditional on the data (or, to converge in P^*-probability, in probability) to c (possibly random), if $\mathrm{P}^*(|X_n^* - c| > \delta)$ converge in probability to zero. This can be stated as

$$X_n^* - c \xrightarrow{p^*}_p 0 \quad \text{if} \quad \mathrm{P}^*(|X_n^* - c| > \delta) \xrightarrow{p} 0 \quad \text{for any} \quad \delta > 0.$$

As an example it follows that $X_n^* = n^{-1}\sum_{t=1}^n \varepsilon_t^* \xrightarrow{p^*}_p 0$ as by the bootstrap equivalent of Markov's inequality[2] one has,

$$\mathrm{P}^*(|X_n^*| > \delta) \leq \mathrm{E}^*(X_n^*)^2/\delta^2.$$

By definition,

$$\mathrm{E}^*(X_n^*)^2 = n^{-2}\,\mathrm{E}^*(\sum_{t=1}^n \varepsilon_t^{*2} + 2\sum_{t=1}^n\sum_{s=t+1}^n \varepsilon_t^*\varepsilon_s^*),$$

with $\mathrm{E}^*(\varepsilon_t^{*2}) = \mathrm{V}^*(\varepsilon_t^*)$, and for $s \neq t$,

$$\mathrm{E}^*(\varepsilon_s^*\varepsilon_t^*) = \mathrm{E}^*(\varepsilon_s^*)\,\mathrm{E}^*(\varepsilon_t^*) = 0.$$

Hence, since $\mathrm{E}^*(X_n^*)^2 = n^{-1}\,\mathrm{V}^*(\varepsilon_t^*)$, we conclude that

$$X_n^* = n^{-1}\sum_{t=1}^n \varepsilon_t^* \xrightarrow{p^*}_p 0. \tag{3.8}$$

[2] Markov's inequality: $P(|X| > \delta) \leq E|X|^k/\delta^k$ for any $\delta > 0$ and $\kappa \geq 1$.

A key ingredient in the asymptotic analysis of the non-bootstrap AR model is the CLT, and we need a bootstrap equivalent of the CLT, and a bootstrap equivalent of convergence in distribution. By definition, $X_n \xrightarrow{d} X$ if $F_{X_n}(x) = P(X_n \leq x) \to F_X(x) = P(X \leq x)$, at all continuity points of $F_X(\cdot)$. Likewise, X_n^* converge in distribution to X conditional on the data (or, as sometimes used, X_n^* converges "weakly in probability"), that is, $X_n^* \xrightarrow{d^*}_p X$, if the bootstrap cumulative distribution function converges in probability. Specifically, $X_n^* \xrightarrow{d^*}_p X$ if

$$F_{X_n^*}^*(x) = P^*(X_n^* \leq x) \xrightarrow{p} F_X(x),$$

at all continuity points of $F_X(\cdot)$. The next lemma illustrates that, as one might expect, the sum of bootstrap innovations ε_t^* is asymptotically Gaussian (in probability).

Lemma 3.1 (van der Vaart, 2000, Theorem 23.4) *With* $X_n^* = n^{-1/2}\sum_{t=1}^n \varepsilon_t^*$, *then*

$$X_n^* \xrightarrow{d^*}_p X \overset{d}{=} \mathrm{N}\left(0, \sigma_0^2\right). \tag{3.9}$$

The proof, as for most bootstrap CLTs, is based on applying a CLT for *triangular arrays*, as $\{\varepsilon_t^*\}_{t=1}^n$ are sampled from $\tilde{\varepsilon}_t^c$, which depends on n.

To give an idea of the underlying theory, consider here verifying (3.9) using a classic approach based on the *characteristic function*, see also Durret (2019, proof of Theorem 3.4.10). For a random variable X, the characteristic function defines uniquely the distribution of X and is defined by $\phi(s) = \mathrm{E}(\exp(isX))$. Here i is the complex (unit imaginary) number which satisfies $i^2 = -1$ and $s \in \mathbb{R}$, and for $X \overset{d}{=} \mathrm{N}(0, \sigma_0^2)$ it holds that $\phi(s) = \exp(-\frac{s^2}{2}\sigma_0^2)$.

With the bootstrap characteristic function of X_n^* defined by $\phi_n^*(s) = \mathrm{E}^*(\exp(isX_n^*))$, it follows that (3.9) holds if

$$\phi_n^*(s) \xrightarrow{p} \exp(-\tfrac{s^2}{2}\sigma_0^2) = \phi(s).$$

Note first, as ε_t^* are i.i.d. conditionally on the data,

$$\begin{aligned}\mathrm{E}^*(\exp(isX_n^*)) &= \mathrm{E}^*(\prod_{t=1}^n \exp(isX_n^*)) = \prod_{t=1}^n \mathrm{E}^*\left(\exp\left(isn^{-1/2}\varepsilon_t^*\right)\right) \\ &= \left(\mathrm{E}^*\left(\exp\left(isn^{-1/2}\varepsilon_t^*\right)\right)\right)^n.\end{aligned}$$

Next, see Durret (2019, Lemma 3.3.19), a Taylor expansion of $\exp(\cdot)$ at $s = 0$ gives

$$\begin{aligned}\mathrm{E}^*\left(\exp\left(isn^{-1/2}\varepsilon_t^*\right)\right) &= 1 + isn^{-1/2}\,\mathrm{E}^*(\varepsilon_t^*) - \tfrac{1}{2}s^2 n^{-1}\,\mathrm{E}^*\left(\varepsilon_t^{*2}\right) + o_p\left(n^{-1}\right) \\ &= 1 - n^{-1}(\tfrac{1}{2}s^2\sigma_0^2) + o_p\left(n^{-1}\right),\end{aligned}$$

using $\mathrm{E}^*(\varepsilon_t^*) = 0$ and $\mathrm{E}^*(\varepsilon_t^{*2}) \xrightarrow{p} \sigma_0^2$, see (3.7). It therefore follows as desired that

$$\mathrm{E}^*(\exp(isX_n^*)) = \left(1 - n^{-1}(\tfrac{1}{2}s^2\sigma_0^2) + o_p\left(n^{-1}\right)\right)^n \xrightarrow{p} \phi(s),$$

as for any sequence c_n, with $c_n \xrightarrow{p} c \in \mathbb{C}$ as $n \to \infty$, then similar to Durret (2019, Theorem 3.4.2), $(1 - n^{-1}c_n)^n \xrightarrow{p} \exp(-c)$.

Remark 3.6 (Lindeberg condition) *The CLTs in Durret (2019, Theorem 3.4.10) and van der Vaart (2000, Theorem 23.4) for triangular arrays follow by verifying*

$$\mathrm{E}^*\left(X_n^{*2}\right) = n^{-1}\sum_{t=1}^{n}\mathrm{E}^*\left(\varepsilon_t^{*2}\right) \xrightarrow{p} \sigma_0^2,$$

in addition to the bootstrap Lindeberg condition,

$$\gamma_n^* = n^{-1}\sum_{t=1}^{n}\mathrm{E}^*\left(\varepsilon_t^{*2}\mathbb{I}\left(|\varepsilon_t^*| > \delta n^{1/2}\right)\right) = n^{-1}\sum_{t=1}^{n}\left(\tilde{\varepsilon}_t^c\right)^2\mathbb{I}\left(|\tilde{\varepsilon}_t^c| > \delta n^{1/2}\right) \xrightarrow{p} 0.$$

A simple way to see that $\gamma_n^* \xrightarrow{p} 0$ is for example to note that if (the rather strong moment condition) $\mathrm{E}(\varepsilon_t^4) < \infty$ holds, the LLN applies to $n^{-1}\sum_{t=1}^{n}\left(\tilde{\varepsilon}_t^c\right)^4$ and hence,

$$\gamma_n^* = n^{-1}\sum_{t=1}^{n}\left(\tilde{\varepsilon}_t^c\right)^2\mathbb{I}\left(|\tilde{\varepsilon}_t^c| > \delta n^{1/2}\right) \leq \frac{1}{n\delta^2}n^{-1}\sum_{t=1}^{n}\left(\tilde{\varepsilon}_t^c\right)^4 \xrightarrow{p} 0.$$

3.2 Bootstrap validity under $\mathcal{H}_0$

As briefly mentioned in the introduction, it is important for the application of the bootstrap that the limiting distribution (in probability) of the bootstrap test statistic has the same limiting distribution as the original test statistic when the null is true. Stated differently, we wish here to establish that under $\mathcal{H}_0$ with $\bar{\rho} = \rho_0$,

$$\tau_n^* = W_n^*\left(\rho = \rho_0\right) = \left(\hat{\rho}_n^* - \rho_0\right)^2\sum_{t=1}^{n}x_{t-1}^{*2}/\sigma_0^2 \xrightarrow{d^*}_p \chi_1^2.$$

By definition, the bootstrap estimator $\hat{\rho}_n^*$ is given by

$$\hat{\rho}_n^* = n^{-1}\sum_{t=1}^{n}x_t^*x_{t-1}^*(n^{-1}\sum_{t=1}^{n}x_{t-1}^{*2})^{-1},$$

such that by the bootstrap scheme employed, that is $x_t^* = \rho_0 x_{t-1}^* + \varepsilon_t^*$, it follows that

$$n^{1/2}\left(\hat{\rho}_n^* - \rho_0\right) = \underbrace{n^{-1/2}\sum_{t=1}^{n}\varepsilon_t^*x_{t-1}^*}_{(i)}\underbrace{(n^{-1}\sum_{t=1}^{n}x_{t-1}^{*2})^{-1}}_{(ii)}.$$

Here a bootstrap CLT should be used for the first term *(i)*, and a bootstrap LLN for the second term *(ii)* in order to find the limiting behavior of the bootstrap estimator, and hence of the test statistic τ_n^*. Consider first *(ii)*, which is an average of lagged x_t^* squared, with

$$x_t^* = \rho_0 x_{t-1}^* + \varepsilon_t^* = \sum_{i=0}^{t-1}\rho_0^i\varepsilon_{t-i}^* + \rho_0^t x_0. \tag{3.10}$$

As ε_t^* depends on n, and the data $\{x_t\}_{t=0}^n$, the concepts of stationarity and ergodicity – while applying to x_t – do not to apply for the x_t^*. However, the following lemma holds which establishes that the LLN holds for the average of x_t^* and x_t^{*2}:

Lemma 3.2 *Suppose that $\{x_t\}_{t=0}^n$ is given by (2.1) with $|\rho_0| < 1$ and ε_t i.i.d. $(0, \sigma_0^2)$. Assume furthermore, with ε_t^* are iid sampled with replacement from $\{\tilde{\varepsilon}_t^c\}_{t=1}^n$ and x_t^* given by (3.5). Then, as $n \to \infty$,*

$$n^{-1} \sum_{t=1}^{n} x_{t-1}^* \xrightarrow{p^*}_p 0 \quad \textit{and} \quad n^{-1} \sum_{t=1}^{n} x_{t-1}^{*2} \xrightarrow{p^*}_p \omega_0 = \sigma_0^2 \left(1 - \rho_0^2\right)^{-1}.$$

The proof of Lemma 3.2 is given in the appendix. Note that for the case of $\rho_0 = 0$ the arguments are similar to the arguments used to establish $n^{-1} \sum_{t=1}^n \varepsilon_t^* \xrightarrow{p^*}_p 0$ in the previous section.

Remark 3.7 (LLN triangular arrays) *Lemma 3.2 is the bootstrap equivalent of the weak law of large numbers for triangular arrays, see also Durret (2019, Theorem 2.2.6).*

Next, consider the CLT candidate term *(ii)*,

$$n^{-1/2} \sum_{t=1}^{n} \varepsilon_t^* x_{t-1}^*.$$

As for the non-bootstrap case, with $\mathcal{F}_t^* = \sigma\left(x_t^*, x_{t-1}^*, ..., x_0^*\right)$, then (conditionally on the data) a bootstrap CLT for martingale difference *arrays* (mda) can be applied. In particular, $\mathrm{E}^*\left(\varepsilon_t^* x_{t-1}^* | \mathcal{F}_{t-1}^*\right) = x_{t-1}^* \mathrm{E}^*\left(\varepsilon_t^*\right) = 0$, while for the conditional second order moment (conditional on the data), it follows by application of Lemma 3.2 that

$$n^{-1} \sum_{t=1}^{n} \mathrm{E}^*(\left(\varepsilon_t^* x_{t-1}^*\right)^2 | \mathcal{F}_{t-1}^*) = n^{-1} \sum_{t=1}^{n} x_{t-1}^{*2} \mathrm{E}^*\left(\varepsilon_t^{*2}\right) \xrightarrow{p^*}_p \sigma_0^2 \omega_0.$$

It remains to establish the bootstrap Lindeberg condition, $\gamma_n^* \xrightarrow{p^*}_p 0$, where

$$\gamma_n^* = n^{-1} \sum_{t=1}^{n} \mathrm{E}^*\left(\left(\varepsilon_t^* x_{t-1}^*\right)^2 \mathbb{I}\left(|\varepsilon_t^* x_{t-1}^*| > \delta\sqrt{n}\right) | \mathcal{F}_{t-1}^*\right). \tag{3.11}$$

Similar to Remark 3.6, this follows by using that for some (arbitrarily small) $\eta > 0$,

$$\begin{aligned}
\gamma_n^* &\leq \frac{1}{n^{1+\eta/2}\delta^\eta} \sum_{t=1}^{n} \mathrm{E}^*(|\varepsilon_t^* x_{t-1}^*|^{2+\eta} | \mathcal{F}_{t-1}^*) \\
&= \frac{1}{n^{\eta/2}\delta^\eta} n^{-1} \sum_{t=1}^{n} |x_{t-1}^*|^{2+\eta} \mathrm{E}^* |\varepsilon_t^*|^{2+\eta} \xrightarrow{p^*}_p 0,
\end{aligned}$$

which holds provided $\mathrm{E}\,|\varepsilon_t|^{2+\eta} < \infty$, using arguments as in Lemma 3.2. Note that in Remark 3.6 the same argument is used for $\eta = 2$.

Hence, with ε_t^* iid resampled from $\{\tilde{\varepsilon}_t^c\}_{t=1}^n$, ε_t i.i.d. $(0, \sigma_0^2)$ with $\mathrm{E}\,|\varepsilon_t|^{2+\eta} < \infty$, it follows that the limiting distribution (in probability) of the bootstrap MLE is given by

$$n^{1/2}\left(\hat{\rho}_n^* - \rho_0\right) \xrightarrow{d^*}_p \mathrm{N}\left(0, \sigma_0^2 \omega_0^{-1}\right). \tag{3.12}$$

This immediately leads to the desired result:

Theorem 3.1 *Under* $\mathcal{H}_0$ *with* $|\rho_0| < 1$, *and with* ε_t^* *iid resampled from* $\{\tilde{\varepsilon}_t^c\}_{t=1}^n$, *with* ε_t *being i.i.d.* $(0, \sigma_0^2)$ *with* $\mathrm{E}\,|\varepsilon_t|^{2+\eta} < \infty$ *for some* $\eta > 0$, *it holds that*

$$\tau_n^* = W_n^*\,(\rho = \rho_0) \xrightarrow{d^*}_p \chi_1^2. \tag{3.13}$$

Remark 3.8 (Bootstrap p-value) *The bootstrap p-value* $p_{n,B}^*$ *in Step* $\mathcal{C}$ *of the bootstrap algorithm is an approximation to the "true" bootstrap p-value* p_n^*, *where* $p_n^* = \mathrm{P}^*\,(\tau_n^* > \tau_n)$, *in the sense that* $p_{n,B}^* \xrightarrow{p} p_n^*$ *as* B *tends to infinity. See e.g. Cavaliere, Nielsen and Rahbek (2015, Rem.2) for details in terms of the stronger concept of "almost sure" convergence.*

Remark 3.9 (Validity under $\mathcal{H}_0$) *Note that by Cavaliere, Nielsen and Rahbek (2015, Corollary 1) for the bootstrap p-value* p_n^* *(see Remark 3.8), it follows that as the limiting* χ_1^2 *distribution has a continuous distribution function, under the conditions of Theorem 3.1,* $p_n^* \xrightarrow{d} \mathrm{U}$, *with* U *uniformly distributed on* $[0, 1]$, *see e.g. Hansen (1996, 2000).*

Remark 3.10 (Moments of ε_t) *Note that* $\mathrm{E}\,|\varepsilon_t|^{2+\eta} < \infty$ *for some* η, *or* $\mathrm{E}\,(\varepsilon_t^4) < \infty$ *as is often used, is required. This reflects the fact, as typically found for the bootstrap, that to prove bootstrap validity further moment restrictions are used when compared to the non-bootstrap case. This is, as illustrated above, due to the complexities arising when applying bootstrap LLNs and arguments in connection to establishing bootstrap Lindeberg-type conditions, see also Cavaliere and Rahbek (2020). However, notably, as also discussed in the Monte Carlo Section 4, while the higher order moment conditions are sufficient for the mathematical arguments, their necessity is often not reflected in bootstrap simulations.*

Remark 3.11 (Introducing σ^2) *As for the non-bootstrap case, the result in Theorem 3.1 also holds for the case where* σ^2 *is treated as a parameter.*

3.3 Bootstrap validity under the alternative

We consider here the convergence of the bootstrap statistic τ_n^* in (3.6) when the alternative holds. That is, assume here that the original data are generated with true value $\theta_0 = (\rho_0, \sigma_0^2)'$, but the hypothesis tested is as before $\mathcal{H}_0 : \rho = \bar{\rho}$ with $\bar{\rho} \neq \rho_0$. As argued below it holds that Theorem 3.1 hold under the alternative as well, such that

$$\tau_n^* \xrightarrow{d^*}_p \chi_1^2. \tag{3.14}$$

For the application of bootstrap-based testing, this implies that under the alternative, as $W_n\,(\rho = \bar{\rho})$ diverges, while $W_n^*\,(\rho = \bar{\rho})$ converges in distribution, the bootstrap-based test will reject with probability tending to one. That is, asymptotic bootstrap validity holds since by Cavaliere, Nielsen and Rahbek (2015, Corollary 1) the bootstrap p-value p_n^*, defined in Remark 3.9, tends to zero in probability under the alternative, $p_n^* \xrightarrow{p} 0$.

A key argument for (3.14) to hold is to note that the identity,

$$\varepsilon_t^* = x_t^* - \bar{\rho} x_{t-1}^*,$$

holds independently of whether $\bar{\rho}$ is the data true value ρ_0 or not. That is, $\bar{\rho}$ is by construction *the bootstrap true value*, such that under the null and also under the alternative, the bootstrap estimator can be rewritten as

$$n^{1/2}\left(\hat{\rho}_n^* - \bar{\rho}\right) = n^{-1/2}\sum_{t=1}^{n}\varepsilon_t^* x_{t-1}^* (n^{-1}\sum_{t=1}^{n} x_{t-1}^{*2})^{-1}.$$

What differs is that ε_t^* under the alternative is resampled from recentered residuals $\tilde{\varepsilon}_t^c$ with

$$\tilde{\varepsilon}_t = \varepsilon_t\left(\bar{\rho}\right) = x_t - \bar{\rho}x_{t-1} = \varepsilon_t + \left(\bar{\rho} - \rho_0\right)x_{t-1} \neq \varepsilon_t.$$

That is, while under the null hypothesis when $\bar{\rho} = \rho_0$, the identity $\tilde{\varepsilon}_t = \varepsilon_t$ holds, this is not the case under the alternative. Hence to establish (3.14) a repeated application of the bootstrap LLN (applied to $\sum_{t=1}^{n} x_{t-1}^{*2}$) and CLT (applied to $\sum_{t=1}^{n}\varepsilon_t^* x_{t-1}^*$) are needed under the alternative. For the AR process of order one considered here, the arguments are based on simple modifications of the theory under $\mathcal{H}_0$.

Remark 3.12 (Theory for the wild bootstrap) *The same results can be shown to apply for the wild bootstrap in the case of conditional heteroskedasticity, see e.g. Gonçalves and Kilian (2004, 2007).*

4 Finite sample behavior

Throughout, the focus has been on establishing asymptotic validity. This was done by verifying that the bootstrap statistic τ_n^* has the same limiting distribution (in probability) as the original statistic τ_n under the null hypothesis. Moreover, the same was argued to hold under the alternative. To illustrate the finite sample performance of the iid and wild bootstraps for the AR model, we highlight in this section some selected typical findings for the bootstrap based on a small and simple (to replicate) Monte Carlo study. Thus the Monte Carlo study here is not meant to be elaborate; exhaustive and detailed bootstrap Monte Carlo-based investigations are given in several papers, see for example Gonçalves and Kilian (2004) with special attention higher order AR models, as well as the references in Cavaliere and Rahbek (2020).

The Monte Carlo results reported here highlight the importance of the assumptions for the established validity of the bootstrap based test of $\mathcal{H}_0 : \rho = \bar{\rho}$ in the autoregressive model. Specifically, it was argued that the true value of the autoregressive root ρ_0 for x_t should satisfy $|\rho_0| < 1$, and it was emphasized that ε_t is an i.i.d. $(0, \sigma_0^2)$ sequence, such that $\mathrm{E}\,|\varepsilon_t|^{2+\eta} < \infty$, or rather, $\mathrm{E}\left(\varepsilon_t^4\right) < \infty$.

With details of the Monte Carlo designs and consideration given below, we initially mention the following findings for bootstrap simulations in the AR model with a constant term. The findings are typical for existing applications of the bootstrap and are standard in time series contexts.

(i) With ε_t i.i.d. $\mathrm{N}(0, \sigma_0^2)$ the χ_1^2-based asymptotic test performs well for even small samples of size n in terms of empirical rejection frequencies, or empirical size, for $\rho_0 = 0.5$, while for $\rho_0 = 0.9$ the asymptotic test fails as its empirical size is not close to the nominal level α. In comparison, the iid (and wild) bootstrap-based test has empirical size close the nominal level in both cases, see Table 1, columns (A) and (B) for $\rho_0 \in \{0.5, 0.9\}$, $\sigma_0^2 = 1$ and $n \in \{15, 25, ..., 1000\}$ with $\alpha = 0.05$.

(ii) With ε_t independently $\mathrm{N}(0, \sigma_t^2)$ distributed with σ_t^2 time-varying (heteroskedasticity), neither the χ_1^2-based asymptotic nor the iid bootstrap-based tests have empirical size close to the nominal level α, which contrasts the wild bootstrap-based test. This is illustrated in Table 1, columns (C)-(F) for a time-changed volatility, $\sigma_t^2 = \sigma_{0,1}^2 + \sigma_{0,2}^2 \mathbb{I}(t \geq [n/2])$, with $\sigma_{0,1}^2 = 1 < \sigma_{0,2}^2 = 15$, and as before $\rho_0 = \{0.5, 0.9\}$ and $n \in \{15, 25, ..., 1000\}$ with $\alpha = 0.05$.

(iii) In terms of empirical rejection frequencies under the alternative, bootstrap based tests and the asymptotic test are comparable. This is illustrated in Table 2, with ε_t i.i.d. $\mathrm{N}(0, \sigma_0^2)$ for $n = 250$ and $\rho_0 = 0.9$, and the test of $\mathcal{H}_0 : \rho = \bar{\rho}$ is considered for values of $\bar{\rho}$ ranging from 0.70 to 0.875 (with $\bar{\rho} = \rho_0$ included as a benchmark).

(iv) As mentioned in the discussion of the wild bootstrap in Section 3, see also Remark 3.4, the wild bootstrap is often motivated by its ability to replicate underlying heteroskedasticity. This is illustrated in Figure 1, where panel (A) shows the empirical residuals $\tilde{\varepsilon}_t$ from one of the draws in Table 1 with $\sigma_t^2 = \sigma_{0,1}^2 + \sigma_{0,2}^2 \mathbb{I}(t \geq [n/2])$, see above. Figure 1, in panels (B), (C) and (D), illustrates that, while ε_t^* replicates the heteroskedasticity for the wild bootstrap, this is not the case for the iid bootstrap.

(v) In terms of requirements for finite moments of the i.i.d. sequence ε_t, Section 3.2 discussed sufficiency and necessity of the condition $\mathrm{E}|\varepsilon_t|^{2+\eta} < \infty$ for some $\eta > 0$, and it was conjectured that $\mathrm{E}(\varepsilon_t^2) < \infty$ was sufficient. This is illustrated in Table 2, which shows that when ε_t does not have a finite variance, then the asymptotic test, as well as the wild and iid bootstrap based tests, fail to have correct empirical size. On the other hand, when ε_t has a finite variance, while the asymptotic test has empirical size far from the nominal, the bootstraps work despite the lack of e.g. fourth order moments. Also Table 2 shows results from a so-called permutation bootstrap, see Section 4.5 below.

4.1 Asymptotic test

We consider x_t as given by the AR model of order one, with a constant term δ included,

$$x_t = \delta + \rho x_{t-1} + \varepsilon_t, \quad t = 1, 2, ..., n, \tag{4.1}$$

with ε_t i.i.d. $\mathrm{N}(0, \sigma^2)$ and x_0 fixed. The parameters are given by $\theta = (\delta, \rho, \sigma^2)' \in \Theta = \mathbb{R}^2 \times (0, \infty)$, with $\theta_0 = (\delta_0, \rho_0, \sigma_0^2)'$ the true value, where $\theta_0 \in \Theta_0 = \{\theta \in \Theta \mid |\rho_0| < 1\}$, and the hypothesis of interest is given by $\mathcal{H}_0 : \rho = \bar{\rho}$.

The *unrestricted* and *restricted* (Gaussian likelihood based) estimators which maximizes

$$\ell_n(\theta) = -\frac{n}{2}(\log \sigma^2 + n^{-1}\sum_{t=1}^{n} \varepsilon_t^2(\rho,\delta)/\sigma^2),$$

with x_0 fixed, and $\varepsilon_t(\rho,\delta) = x_t - \rho x_{t-1} - \delta$, are given by

$$\hat{\theta}_n = (\hat{\delta}_n, \hat{\rho}_n, \hat{\sigma}_n^2)' \quad \text{and} \quad \tilde{\theta}_n = (\tilde{\delta}_n, \rho_0, \tilde{\sigma}_n^2)',$$

respectively. The theory from the case of no constant term immediately carries over, such that for $\bar{\rho} = \rho_0$, $|\rho_0| < 1$ and ε_t i.i.d. $(0, \sigma_0^2)$, as $n \to \infty$,

$$\tau_n = 2(\ell_n(\hat{\theta}_n) - \ell_n(\tilde{\theta}_n)) \xrightarrow{d} \chi_1^2.$$

In the implementations of the asymptotic test, we use the p-value, p_n, calculated as the tail probability of τ_n in the limiting χ_1^2−distribution, and reject if p_n is smaller than the nominal level α.

Column (A) in Table 1 illustrates this for x_t in (4.1) generated with $\delta_0 = 0$, $\rho_0 = \{0.5, 0.9\}$, and $\sigma_0^2 = 1$. Moreover, ε_t is simulated as an i.i.d. $\mathrm{N}(0, \sigma_0^2)$ sequence, and $x_0 = 0$. The empirical rejection frequencies are reported based on $N = 10,000$ repetitions, with nominal level, $\alpha = 0.05$. Results for $n \in \{15, 25, ..., 1000\}$ are given in column (A) of Table 1. Observe, as noted, that quite a large sample is required for the limiting χ_1^2 distribution to be a good approximation, in particular with $\rho_0 = 0.9$, see e.g. Duffee and Stanton (2008) and references therein.

Remark 4.1 (Empirical rejection probabilities) *At the chosen (nominal) level α, with q_α the corresponding $1-\alpha$ quantile of the limiting distribution, the true rejection probability at sample length n is $\alpha_n = P(\tau_n > q_\alpha)$. The Monte Carlo estimator is the empirical rejection frequency computed as $\alpha_{n,N} = \frac{1}{N}\sum_{i=1}^{N} \mathbb{I}(p_{n,i} < \alpha)$, where $p_{n,i}$ is the p-value in Monte Carlo replication i, $i = 1, 2, ..., N$. It follows that the simulation uncertainty of $\alpha_{n,N}$ is given by*

$$\mathrm{V}(\alpha_{n,N}) = \alpha_n(1-\alpha_n)/N,$$

see e.g. Hendry (1984), and for a correctly sized test, with $\alpha_n = 0.05$, and $N = 10^4$, the 95% confidence bound for $\alpha_{n,N}$ is $[0.0456, 0.0544]$. Similar considerations hold for the bootstrap simulations of the test, with $p_{n,i}$ replaced by $p^*_{n,B,i}$, see also Remark 3.8.

[Table 1 around here]

4.2 The iid bootstrap test

To illustrate bootstrap based testing, we apply a restricted recursive bootstrap in terms of residuals estimated under $\mathcal{H}_0$, see Section 3.2.

Specifically, for Steps $\mathcal{A}$ and $\mathcal{B}$ in Section 3, the bootstrap samples $\{x_t^*\}_{t=0}^n$ are sampled from

$$x_t^* = \tilde{\delta}_n + \bar{\rho}x_{t-1}^* + \varepsilon_t^*, \quad t = 1, 2, ..., n, \tag{4.2}$$

with $x_0^* = x_0$ and ε_t^* drawn with replacement (for wild, see below) from $\{\tilde{\varepsilon}_t^c\}_{t=1}^n$, where $\tilde{\varepsilon}_t^c$ are defined as in[3] (3.3) in terms of,

$$\tilde{\varepsilon}_t = \varepsilon_t(\bar{\rho}, \tilde{\delta}_n) = x_t - \bar{\rho}x_{t-1} - \tilde{\delta}_n. \tag{4.3}$$

For the bootstrap sample, $\{x_t^*\}_{t=0}^n$, we estimate the unrestricted and restricted models and calculate the bootstrap statistic τ_n^*, given by

$$\tau_n^* = 2(\ell_n^*(\hat{\theta}_n^*) - \ell_n^*(\tilde{\theta}_n^*)).$$

Here $\hat{\theta}_n^* = (\hat{\delta}_n^*, \hat{\rho}_n^*, \hat{\sigma}_n^{*2})'$ and $\tilde{\theta}_n^* = (\tilde{\delta}_n^*, \bar{\rho}, \tilde{\sigma}_n^{*2})'$ denote the unrestricted and restricted bootstrap estimators, respectively, in terms of the bootstrap log-likelihood function,

$$\ell_n^*(\theta) = -\frac{n}{2}(\log \sigma^2 + n^{-1}\sum_{t=1}^{n} \varepsilon_t^{*2}(\rho, \delta)/\sigma^2), \quad \varepsilon_t^*(\rho, \delta) = x_t^* - \rho x_{t-1}^* - \delta.$$

For Step $\mathcal{C}$, the bootstrap test is based on replicating the above to obtain $\left\{\tau_{n,b}^*\right\}_{b=1}^B$ with B denoting the number of bootstrap repetitions. As discussed in Remark 3.3, the empirical bootstrap p-value is computed as the tail probability,

$$p_{n,B}^* = \frac{1}{B}\sum_{b=1}^{B} \mathbb{I}\left(\tau_{n,b}^* \geq \tau_n\right). \tag{4.4}$$

With $B = 399$ bootstrap repetitions, the empirical rejection frequencies for the $N = 10,000$ Monte Carlo repetitions are presented in column (B) in Table 1.

4.3 The wild bootstrap test

The wild bootstrap design is as in Section 4.2, except that ε_t^* for the wild bootstrap is resampled by,

$$\varepsilon_t^* = w_t^*\tilde{\varepsilon}_t,$$

with w_t^* i.i.d. $(0, 1)$ distributed and $\tilde{\varepsilon}_t$ defined in (4.3). In the simulations w_t^* is chosen as $\mathrm{N}(0, 1)$ and Rademacher distributed, respectively, see also Remark 3.4.

For the simulations reported in Table 1, ε_t are assumed not to be i.i.d. $\mathrm{N}(0, \sigma_0^2)$ distributed in order to illustrate the impacts of heteroskedasticity. Specifically, we set $\varepsilon_t \overset{d}{=} \mathrm{N}(0, \sigma_t^2)$ with $\sigma_t^2 = 1$ for $t = 1, 2, ..., [n/2]$ and $\sigma_t^2 = 15$ for $t = [n/2] + 1, ..., n$. That is,

$$\sigma_t^2 = \sigma_{0,1}^2 + \sigma_{0,2}^2\mathbb{I}(t > [n/2]), \tag{4.5}$$

[3] Note that as the model here includes a constant term, strictly speaking the re-centering of $\tilde{\varepsilon}_t$ is not needed.

with $\sigma_{0,1}^2 = 1$ and $\sigma_{0,2}^2 = 15$. In this case the asymptotic test is *not* consistent as demonstrated in column (A) of Table 1, which reports the empirical rejection probabilities for the asymptotic test. The asymptotic test is severely over-sized – and even for $n = 500$ the empirical size is not close to $\alpha = 0.05$. Also the iid bootstrap is *not* asymptotically valid. With the iid bootstrap as in Section 4.2, the results are reported in column (D), and we observe that the results are similar to the asymptotic test. Intuitively, this reflects that the bootstrap series, $\{x_t^*\}_{t=0}^n$, does not mimic the properties of the original data series $\{x_t\}_{t=0}^n$. This can be illustrated by Figure 1, where panel (A) shows pronounced heteroskedasticity of the estimated residuals, $\tilde{\varepsilon}_t$, for one sample. Panel (B) shows a single iid resampled sample $\{\varepsilon_t^*\}_{t=1}^n$ from $\{\tilde{\varepsilon}_t\}_{t=1}^n$, and as, by definition, in particular the ordering change, the ε_t^* series does not mimic the heteroskedasticity as seen in the estimated residuals in panel (A).

For the wild bootstrap, columns (E)-(F) in Table 1 report the empirical rejection frequencies for the wild bootstrap test with, as mentioned, w_t^* $\mathrm{N}(0,1)$ and Rademacher distributed respectively. The empirical size for the wild bootstrap is quite close to the nominal level, with the Rademacher distribution performing slightly better. Likewise, panels (C) and (D) in Figure 1, illustrate that the wild bootstrap ε_t^* series more closely mimic the properties of the original $\tilde{\varepsilon}_t$ series.

[Figure 1 around here]

4.4 Bootstrap under the alternative

To establish asymptotic validity, it was shown that τ_n^* is also asymptotically χ^2-distributed under the alternative, leading to a consistent bootstrap test. To illustrate this, we consider the empirical probability of rejecting a false hypothesis.

Specifically for the iid bootstrap test in Section 4.2, let the data x_t be generated with true value $\theta_0 = (\rho_0, \sigma_0^2, \delta_0)'$ as before. The hypothesis of interest is $\mathcal{H}_0 : \rho = \bar{\rho}$ and we let $\bar{\rho} \in \{0.875, ..., 0.70\}$ to illustrate the empirical power of the bootstrap test and asymptotic test. From Table 2 it follows that the bootstrap is comparable to the asymptotic test in terms of empirical power.

[Table 2 around here]

4.5 Moment condition, $\mathrm{E}\,|\varepsilon_t|^k < \infty$ for $k \geq 2$

When establishing asymptotic validity of the bootstrap the moment condition, $\mathrm{E}(\varepsilon_t^4) < \infty$, was discussed and it was mentioned that while it is a sufficient condition, it may not be necessary. On the other hand $\mathrm{E}(\varepsilon_t^2) < \infty$ seems necessary, unless the bootstrap algorithm is based on permutation, see below.

To illustrate this, we consider here the data x_t as generated with true value $\theta_0 = (\rho_0, \sigma_0^2, \delta_0)'$ as before for samples of size n, $n \in \{15, ..., 1000\}$. The i.i.d. innovations ε_t are simulated from the Student's t_v-distribution, where the degrees of freedom v,

$v \in \{3/2, 3, 5\}$. Specifically, for $v = 3/2$, $\mathrm{E}\,|\varepsilon_t|^k$ is finite, only for $k < 3/2$, thus allowing first order, but not second order, moments of ε_t. For $v = 3$ and $v = 5$, the second and finite fourth order moments are finite respectively. Also note that the Gaussian case is included as a reference. Table 3 shows that the asymptotic test (based on the χ_1^2 approximation) for all the four cases has empirical rejection rates far from the nominal level of $\alpha = 0.05$, even for large samples n, where $n \in \{15, ..., 1000\}$.

As to the bootstrap design, Table 3 reports bootstrap simulations based on iid sampling and the wild (Rademacher). For $v = 3/2$, as expected, neither the wild nor the iid bootstraps have empirical rejection rates close the nominal level. For $v \geq 3$, both the wild and iid bootstrap works surprisingly well, even when $v = 3$.

Additionally, Table 3 reports bootstrap based testing where ε_t^* are iid sampled, but without replacement. This, which is referred to as the permutation bootstrap, works well in terms of empirical rejection frequencies even for $v = 3/2$. In general, the permutation bootstrap works well in the context of heavy-tailed i.i.d. ε_t (see Cavaliere, Nielsen and Rahbek, 2020, and also the discussion in Section 5.4 for AR models with heavy-tailed innovations).

[Table 3 around here]

5 A selection of further topics

In the previous sections the simple AR model of order one was used to introduce key ideas and challenges of the iid and wild bootstrap schemes when applied to testing the hypothesis, $\mathcal{H}_0 : \rho = \bar{\rho}$. In this section, we provide an overview of recent selected results for the bootstrap when applied to different testing problems in econometric time series models. The overview is *not meant to be exhaustive*, see *e.g.* Cavaliere and Rahbek (2020) for a review of the bootstrap with more technical details, as well the references therein.

5.1 Non-stationary (vector) autoregressive models

5.1.1 The iid bootstrap

Consider initially the AR model in (2.1) again. The hypothesis of non-stationarity is given by $\mathcal{H}_0 : \rho = 1$, or equivalently, with $\pi = \rho - 1$, $\mathcal{H}_0 : \pi = 0$, in the AR model restated as

$$\Delta x_t = \pi x_{t-1} + \varepsilon_t, \tag{5.1}$$

where $\Delta x_t = x_t - x_{t-1}$. It follows that the test-statistic τ_n in (2.5) can be written as

$$\tau_n = W_n\,(\pi = 0) = \hat{\pi}_n^2 \sum_{t=1}^{n} x_{t-1}^2 / \sigma_0^2,$$

and under $\mathcal{H}_0$,

$$\tau_n \xrightarrow{d} \tau = (\int_0^1 B\,(u)\,dB\,(u))^2 / \int_0^1 B^2\,(u)\,du, \tag{5.2}$$

where $B(u)$ is a standard Brownian motion, $u \in (0,1)$, and τ is the (squared) Dickey-Fuller distribution, see Hamilton (1994). In this case, in terms of the bootstrap in Section 3, the restricted recursive bootstrap is given by $\Delta x_t^* = \varepsilon_t^*$, with ε_t^* iid resampled, is asymptotically valid. This holds as in this case, $\tau_n^* \xrightarrow{d^*}_p \tau$, both under $\mathcal{H}_0$ and the alternative. In contrast, and as discussed in Basawa *et al.* (1991), the unrestricted recursive bootstrap based on the recursion $\Delta x_t^* = \hat{\pi}_n x_{t-1}^* + \varepsilon_t^*$, with ε_t^* iid resampled, is invalid. This follows as the corresponding bootstrap statistic, τ_n^*, under $\mathcal{H}_0$ converges in distribution to $\tilde{\tau}$, $\tilde{\tau} \neq \tau$. Precisely, the bootstrap conditional distribution function converges weakly rather than in probability; see Basawa *et al.* (1991), Cavaliere and Georgiev (2019) and Cavaliere, Nielsen and Rahbek (2015).

The univariate case of testing for non-stationarity is a special case of the more general hypothesis of non-stationarity in vector AR models for $X_t \in \mathbb{R}^p$ with general lag-structure, as given by

$$\Delta X_t = \pi X_{t-1} + \sum_{i=1}^{k} \gamma_i \Delta X_{t-i} + \varepsilon_t, \tag{5.3}$$

where ε_t are i.i.d. $\mathrm{N}_p(0, \Omega)$ distributed and the initial values $(X_0, \Delta X_0, ..., \Delta X_{1-k})$ are fixed in the statistical analysis. Moreover, π and $(\gamma_i)_{i=1}^k$ are $p \times p$ matrices. The hypothesis of non-stationarity of X_t is given by the hypothesis of reduced rank r, $0 \leq r < p$, of π, see Johansen (1996). Specifically, with $\mathcal{H}_r : \mathrm{rank}(\pi) \leq r$, it follows that this may be written as

$$\mathcal{H}_r : \pi = \alpha \beta',$$

where α and β are $(p \times r)$ dimensional matrices. Under the non-stationarity conditions in Johansen (1996, Theorem 4.2), it follows that X_t is a non-stationary process, with r stationary, or co-integrating, relations $\beta' X_t$, and $(p - r)$ common trends given by $\delta' \sum_{i=1}^t \varepsilon_i$, with δ $(p \times (p - r))$ dimensional of full rank, and such that $\delta' \alpha = 0$. The likelihood-ratio statistic $\tau_n(r)$ for co-integration rank r satisfies, under $\mathcal{H}_r$ and the mentioned non-stationarity conditions, that

$$\tau_n(r) \xrightarrow{d} \tau(r) = \mathrm{tr}\{(\int_0^1 B(u)\, dB'(u))'(\int_0^1 B(u)\, B(u)'\, du)^{-1}(\int_0^1 B(u)\, dB'(u))\},$$

which is a multivariate version of (5.2) in terms of the $(p - r)$-dimensional standard Brownian motion, $B(\cdot)$. Cavaliere, Rahbek and Taylor (2012) consider the recursive restricted bootstrap based on

$$\Delta X_t^* = \pi_n^* X_{t-1}^* + \sum_{i=1}^{k} \gamma_{n,i}^* \Delta X_{t-i}^* + \varepsilon_t^*,$$

with $\pi_n^* = \tilde{\alpha}_n \tilde{\beta}_n'$and $\gamma_{n,i}^* = \tilde{\gamma}_{n,i}$; that is, the bootstrap true values are given by the estimators under $\mathcal{H}_r$. Asymptotic validity of the iid bootstrap is established in Cavaliere, Rahbek and Taylor (2012), by showing that $\tau_n^*(r) \xrightarrow{d^*}_p \tau(r)$ under $\mathcal{H}_r$ and the alternative.

Cavaliere, Nielsen and Rahbek (2015) extend the analysis to hypothesis testing on the co-integration (matrix) parameter β. Specifically, Cavaliere *et al.* (2015, Proposition 1 and Theorem 1), establish that under the hypothesis $\mathcal{H}_{\bar{r}} : \beta = \bar{\beta}$, the bootstrap likelihood ratio statistic, τ_n^* satisfies $\tau_n^* \xrightarrow{d^*}_p \chi^2_{(p-r)r}$. Importantly, it is also established that under the alternative τ_n^* has a limiting distribution (in distribution) in terms of a diffusion process with a stochastic diffusion coefficient, and hence it is bounded in probability such that the bootstrap based test is asymptotically valid.

5.1.2 The wild bootstrap

In order to allow for possible heteroskedasticity in the ε_t sequence in (5.3), also the application of the wild bootstrap has been studied. Results for application of the wild bootstrap in general lag univariate AR models in Gonçalves and Kilian (2004), with ε_t allowed to have general time-varying volatility structures, such as ARCH and stochastic volatility, have been generalized to the testing the hypothesis of co-integration $\mathcal{H}_r$ in Cavaliere, Rahbek and Taylor (2010a, 2010b, 2014) and Boswijk *et al.* (2016).

Moreover, Boswijk, Cavaliere, Rahbek and Taylor (2016) and Boswijk, Cavaliere, Georgiev and Rahbek (2020) consider general hypothesis testing on the co-integration parameters α and β, with $\pi = \alpha\beta'$ in (5.3). They consider the case of stochastic volatility, where $\varepsilon_t = \Omega_t^{1/2} z_t$, with the p-dimensional z_t i.i.d. $(0, 1)$ and the time-varying $(p \times p)$-dimensional $\Omega_t = \Omega\,(t/n)$. Moreover, with "$\xrightarrow{w}$" denoting weak convergence, it is assumed that for $u \in (0, 1)$,

$$n^{-1/2} \sum_{t=1}^{[nu]} \varepsilon_t \xrightarrow{w} \int_0^u \Omega^{1/2}(s)\, dB(s)\,, \tag{5.4}$$

where B is a p-dimensional standard Brownian motion, which generalizes the i.i.d. $(0, \Omega)$ assumption, where $n^{-1/2} \sum_{t=1}^{[nu]} \varepsilon_t \xrightarrow{w} \Omega^{1/2} B(u)$. Specifically, the limiting process in (5.4) is a continuous-time martingale, with in general an unknown covariance (kernel). This implies that the limiting distribution of the test statistic(s) τ_n, for example for the mentioned hypotheses $\mathcal{H}_r$ and $\mathcal{H}_{\bar{r}}$, will depend on unknown nuisance parameters, which again means asymptotic inference is infeasible in practice. In contrast, for the wild bootstrap, it is established that $n^{-1/2} \sum_{t=1}^{[nu]} \varepsilon_t^*$ has the same limiting distribution (in probability), and, as a result, the wild bootstrap is asymptotically valid as shown in Boswijk *et al.* (2016, 2020), under some additional regularity conditions to be verified.

5.2 Time-varying conditional volatility models

As discussed in Andrews (2000), applying bootstrap based testing in ARCH models is in general difficult, and may be invalid in certain cases, due to general problems arising when testing hypotheses in time series models, when one or more parameters under the null may be "on the boundary of the parameter space".

To illustrate, consider here x_t given by a linear ARCH model of order q

$$x_t = \sigma_t(\theta)\, z_t, \;\; t = 1, ..., n,$$

with z_t i.i.d. $(0, 1)$, and

$$\sigma_t^2(\theta) = \omega + \sum_{i=1}^{q} \alpha_i x_{t-i}^2.$$

In the statistical analysis, the initial values $(x_0, ..., x_{-q+1})$ are fixed, and the parameter $\theta = (\omega, \alpha_1, ..., \alpha_q)' \in \Theta$, where

$$\Theta = \{\theta \in \mathbb{R}^{q+1} : \ \omega^2 > 0, \text{ and } \alpha_i \geq 0 \text{ for } i = 1, ..., q\}.$$

Thus by definition of the parameter space Θ, if for the true value $\theta_0 = (\omega_0, \alpha_{0,1}, ..., \alpha_{0,q})'$, it holds that $\alpha_{0,j} = 0$ for some j, the true value θ_0 is on the boundary of Θ.

The fact that it is unknown a priory which of the ARCH coefficients may, or may not be zero, leads to non-pivotal limiting distributions of test statistics and estimators. Consider here the likelihood ratio statistic τ_n for the nullity of the q-th order ARCH coefficient, that is the hypothesis $\mathcal{H}_q : \alpha_q = 0$. With the Gaussian likelihood function given by

$$\ell_n(\theta) = -\frac{1}{2} \sum_{t=1}^{n} \left(\log \sigma_t^2(\theta) + x_t^2 / \sigma_t^2(\theta)\right),$$

by definition, $\tau_n = 2(\ell_n(\hat{\theta}_n) - \ell_n(\tilde{\theta}_n))$, where the unrestricted[4] Gaussian MLE is given by $\hat{\theta}_n = \arg\max_{\theta \in \Theta} \ell_n(\theta)$, while $\tilde{\theta}_n$ is the Gaussian MLE under $\mathcal{H}_q$. By Andrews (1999, 2001), it follows that τ_n has a limiting distribution which is non-standard. In addition, the limiting distribution of τ_n is non-pivotal as, crucially, it depends on whether $\alpha_{0,i} > 0$, or $\alpha_{0,i} = 0$ for $i = 1, ..., q-1$ under $\mathcal{H}_q$.

While this implies that the unrestricted bootstrap is invalid, see Andrews (2000), it follows by Cavaliere, Nielsen and Rahbek (2017) that the iid restricted bootstrap is asymptotically valid under mild conditions for the simple case of the first order ARCH with $q = 1$. Moreover, Cavaliere, Nielsen, Pedersen and Rahbek (2020) demonstrate validity of a modified restricted bootstrap, which can be applied for general testing problems in parametric models with parameters on the boundary under the null. Specifically, for the case of ARCH of order q, consider the bootstrap process

$$x_t^* = \sigma_t^*(\theta_n^*) z_t^*,$$

with z_t^* iid resampled from $\hat{z}_t = x_t / \sigma_t(\hat{\theta}_n)$, after recentering and rescaling these. The bootstrap conditional volatility process $\sigma_t^{*2}(\theta_n^*)$ is given by

$$\sigma_t^{*2}(\theta_n^*) = \omega_n^* + \sum_{i=1}^{n} \alpha_{n,i}^* x_{t-1}^2, \tag{5.5}$$

with $\omega_n^* = \tilde{\omega}_n$ and $\alpha_{n,i}^* = \tilde{\alpha}_{n,i} \mathbb{I}(\tilde{\alpha}_{n,i} > c_n)$, with c_n a deterministic sequence which satisfies *(i)* $c_n \to 0$, and *(ii)* $n^{1/2} c_n \to \infty$, as $n \to \infty$. The bootstrap scheme is referred to as

[4] A complete discussion of specifications and properties of the parameter space Θ is given in Cavaliere, Nielsen, Pedersen and Rahbek (2020).

modified since in (5.5) the bootstrap true values $\alpha^*_{n,i}$ – by "shrinking" – are set to zero for $i = 1, ..., q-1$, provided $\tilde{\alpha}_{n,i}$ is small relative to c_n. Note that with the time-varying bootstrap volatility defined by (5.5), this is a case of a fixed design (or, rather fixed volatility) bootstrap, see Step $\mathcal{A}$ in Section 3. The general fixed volatility bootstrap for ARCH models is considered in Cavaliere, Pedersen and Rahbek (2018), and the modified, by shrinking, fixed volatility bootstrap is shown in Cavaliere, Nielsen, Pedersen and Rahbek (2020, Proposition 1) to be asymptotically valid. Simulations there show that both the fixed volatility, and the recursive with $x^2_{t-i} = x^{*2}_{t-i}$ in (5.5) bootstrap based tests have empirical rejection frequencis which are close to the nominal level for small, and moderate, sample sizes n. Moreover, as for the discussion of the moment requirements for the AR bootstrap, simulations indicate that while sufficient, the moment constraints on the original ARCH process imposed to establish validity are also not necessary.

5.3 Double autoregressive models

The double autoregressive (DAR) model combines the AR and ARCH models, as both the conditional mean and conditional variance depend on lagged levels of the process, see Ling (2004, 2007), and for a multivariate "co-integrated" version, Nielsen and Rahbek (2014).

Consider here the first order DAR model as given by

$$\Delta x_t = \pi x_{t-1} + \sigma_t(\theta) z_t, \quad \sigma^2_t(\theta) = \sigma^2 + \alpha x^2_{t-1}, \quad t = 1, 2, ..., n, \tag{5.6}$$

with z_t i.i.d. $\mathrm{N}(0,1)$, x_0 is fixed in the statistical analysis, and the parameter given by $\theta = (\pi, \sigma^2, \alpha)' \in \Theta$, with $\Theta = \{\theta \in \mathbb{R}^3 : \sigma^2 > 0 \text{ and } \alpha \geq 0\}$.

A notable special feature of the DAR process is that for $\pi = 0$, the process is strictly stationary for any $0 < \alpha < 2.42$, see Borkovec and Klüppelberg (2001) and Ling (2004, 2007), while the process is non-stationary when $\pi = \alpha = 0$. With $\pi = 0$, and $\alpha \in (0, 2.42)$, while being strictly stationary, the DAR process x_t has infinite variance (and only finite fractional moments).

From the specification of the parameter space Θ evidently for $\alpha_0 = 0$, the true value θ_0 is on the boundary, raising the issues discussed in Section 5.2 in relation to the ARCH model. Asymptotic theory for the Gaussian likelihood-based MLE with $\alpha_0 > 0$ is given by Ling (2004), while Cavaliere and Rahbek (2020) extend the results to allow for the boundary case. Klüppelberg, Maller, van de Vyver and Wee (2002) derive the asymptotic distribution of the likelihood ratio statistic τ_n for the hypothesis of non-stationarity as given by $\mathcal{H}_0 : \pi = \alpha = 0$, see also Chen, Li and Ling (2013). Different versions of a bootstrap based test are discussed in Cavaliere and Rahbek (2020). In particular, validity is established for a restricted bootstrap given by, $\Delta x^*_t = \tilde{\sigma}_n z^*_t$, where $\tilde{\sigma}^2_n$ is the MLE of σ^2 under $\mathcal{H}_0$, while z^*_t is obtained by iid resampling of *unrestricted* residuals, $\hat{z}_t$ (recentered and rescaled). That is, the unrestricted residuals $\hat{z}_t$ are given by

$$\hat{z}_t = (\Delta x_t - \hat{\pi}_n x_{t-1}) / \sigma_t(\hat{\theta}_n),$$

with $\hat{\theta}_n = (\hat{\pi}_n, \hat{\sigma}_n, \hat{\alpha}_n)'$ the unrestricted MLE. In line with the discussion in Section 3 regarding restricted and unrestricted residuals, this choice ensures that $\hat{z}_t$ for large n, is "close" to the true z_t, irrespective of whether the null $\mathcal{H}_0$ is true or not. While the validity result is shown for this choice, simulations indicate that in practice the difference between choosing to resample from $\hat{z}_t$, or from $\tilde{z}_t = \Delta x_t / \tilde{\sigma}_n$, is negligible. See Cavaliere and Rahbek (2020) for a detailed discussion of this as well as asymptotic theory for different bootstraps.

5.4 Heavy-tailed autoregressive models

So far results for the (vector) AR models have been derived under the assumption that the innovations ε_t have mean zero, and a finite variance σ^2, or some time-varying, possibly conditional, variance σ_t^2 when discussing heteroskedasticity. To allow for more extreme events, and phenomena such as "bubble" periods with local explosive behavior, this assumption was relaxed in Davis and Resnick (1985a, 1985b, 1986) and Davis and Song (2020) where the i.i.d. innovations ε_t are allowed to have infinite variance. Specifically, they consider the case of ε_t i.i.d with a stable distribution such as the Cauchy; that is, "heavy-tailed" as the tails of the distribution of ε_t are assumed to decay at a rate which is slower than the Gaussian (exponential) rate.

Two key examples are given by the classic AR model and the so-called non-causal AR model of order one in terms of i.i.d. stable distributed ε_t,

$$\text{AR} : x_t = \rho x_{t-1} + \varepsilon_t, \quad \text{and} \quad \text{AR}^+ : x_t = \rho^+ x_{t+1} + \varepsilon_t. \tag{5.7}$$

For the standard, and hence causal, AR recall that with $t = 1, ..., n$, x_0 is the initial value which is fixed in the statistical analysis, while for the non-causal AR^+, x_n is the "initial value" due to the forward recursion. Recently non-causal AR^+ type models have become popular as they seem to capture well the dynamics of phenomena such as bubbles where, after period of exponential type growth, the process "collapses". Interestingly, and linking the heavy-tail AR^+ models with the DAR model in Section 5.3, the AR^+ process in (5.7) can be shown to have a causal "semi-strong" representation as the DAR process in (5.6), see Gourieroux and Zakoïan (2017).

Consider testing the hypothesis $\mathcal{H}_0^+ : \rho^+ = \bar{\rho}$ using the Gaussian likelihood based statistic τ_n^+, given by

$$\tau_n^+ = (\hat{\rho}_n^+ - \bar{\rho})(\sum_{t=1}^{n-1} x_{t+1})^{1/2} / \hat{\sigma}_n,$$

where $\hat{\rho}_n^+ = \sum_{t=1}^{n-1} x_t x_{t+1} / \sum_{t=1}^{n} x_{t+1}^2$, and $\hat{\sigma}_n^2 = n^{-1} \sum_{t=1}^{n-1} (x_t - \hat{\rho}_n^+ x_{t-1})^2$. While the test statistic τ_n^+ is analogous to the (square root of the) previously studied statistic τ_n in (2.5), the limiting distribution is non-standard, as the ε_t are assumed to be i.i.d. stable distributed. For example, with ε_t Cauchy distributed, $(n/\log n)(\hat{\rho}_n^+ - \bar{\rho})$ is asymptotically distributed as $(1 + \bar{\rho})\mathcal{C}\chi_1^2$, where $\mathcal{C}$ is standard Cauchy distributed, and

$\tau_n^+ = O_P\left(n^{-1/2}\log n\right)$. For general stable distributions, see Cavaliere, Nielsen and Rahbek (2020), asymptotic testing is infeasible as the limiting distributions depend on the "tail index" (which is one for the Cauchy) of ε_t, and moreover the normalization (which is $n^{-1/2}\log n$ for the Cauchy case) depend on further, and in practice, unknown quantities.

Cavaliere, Nielsen and Rahbek (2020) discuss validity of the recursive bootstrap scheme similar to (3.5),

$$x_t^* = \bar{\rho} x_{t+1}^* + \varepsilon_t^*,$$

initialized with $x_n^* = x_n$, and with ε_t^* resampled from the restricted residuals $\left\{\tilde{\varepsilon}_t^+\right\}_{t=1}^{n-1}$, where $\tilde{\varepsilon}_t^+ = x_t - \bar{\rho}x_{t+1}$. Crucially the ε_t^* are not sampled by iid resampling with replacement, as this would lead to an invalid bootstrap test, see Athreya (1987) and Knight (1989). Instead, ε_t^* are resampled without replacement, that is, by permuting $\left\{\tilde{\varepsilon}_t^+\right\}_{t=1}^{n-1}$, or in combination with the wild, by permuting $\left\{\tilde{\varepsilon}_t^+ w_t^*\right\}_{t=1}^{n-1}$, where w_t^* are i.i.d. Rademacher distributed, see Remark 3.4. With τ_n^{+*} the bootstrap statistic based on the permutation, or the combined permutation-wild bootstrap, Theorem 1 in Cavaliere, Nielsen and Rahbek (2020) establish validity of the bootstrap based test under the null hypothesis for general AR$^+$ models.

Similarly, Cavaliere, Georgiev and Taylor (2016) establish bootstrap validity for the so-called sieve bootstrap in Bühlmann (1997) for general causal AR models with heavy tails. Also note that in terms of testing for the presence of bubbles based on one-sided testing using the supremum of recursively computed Dickey-Fuller type statistics as in Phillips, Wu, and Yu (2011), Harvey, Leybourne, Sollis and Taylor (2016) establish validity of a wild bootstrap based test to allow for heteroskedasticity.

6 Conclusions and further readings

In this article we have provided an introduction to key steps required for a successful implementation of bootstrap hypothesis testing to time series models. In the framework of a simple autoregressive model, we have discussed the (large-sample) validity of recursive bootstrap algorithms, where the bootstrap sample is constructed by mimicking the dependence structure of the original data, as implied by the econometric model at hand. We have discussed the main requirements for bootstrap inference to be valid; that is, to mimic the asymptotic distribution of the original test statistic under the null hypothesis, and to be bounded (in probability) when the null hypothesis is false. To do so, we have introduced the needed probability tools (which involve dealing with the conditional – hence random – nature of the bootstrap measure) and shown how to apply them in order to assess the large sample behavior of the bootstrap estimators and test statistics.

Needless to say, this article is not meant to be exhaustive or to provide a comprehensive treatment of bootstrap inference. We have focused on bootstrap hypothesis testing which, although being one of the most common applications of the bootstrap, is not the only one. Further common implementations of the bootstrap involve computation of standard errors, the construction of confidence intervals for the parameters of interest and bias correction of the estimators; see *inter alia*, Horowitz (2001) or MacKinnon (2006) for reviews.

This article is also not exhaustive in terms of bootstrap algorithms. For instance we have not discussed bootstrap algorithms based on re-sampling block of observations (or residuals) rather than re-sampling single observations (or residuals). Block bootstrap methods, as introduced by Künsch (1989), see also Politis and Romano (1994), Lahiri (2003) and Shao (2010) *inter alia*, are quite flexible and powerful in cases where one does not have strong a priori beliefs or knowledge about the dependence structure of the data. As for subsampling methods (Politis, Romano and Wolf, 1999), they also represent a valid alternative (or complement) to recursive, model based bootstrap. We have focused on a case (the first order AR model) where the data generating process is described by a finite dimensional vector of parameters. Extensions to infinite-dimensional parameter spaces, as for the case of general linear processes, are available in the literature; a classic example is the AR(∞) sieve bootstrap of Kreiss (1988, 1992), Bühlmann (1997) and Gonçalves and Kilian (2007).

It is important also to emphasize that there is a rich literature discussing finite-sample improvements of the bootstrap in various models, including time series models, based on Edgeworth expansions of the τ_n and τ_n^* statistics, as briefly mentioned also in Remark 3.4. An introduction to Edgeworth expansions as applied for the bootstrap can be found in van der Vaart (2000, Chapter 23.3), see also Hall (1992) and Horowitz (2001).

Finally, it is worth noticing that throughout this article bootstrap validity under the null is defined as the fact that under the null the bootstrap test statistic converges to the asymptotic (null) distribution of the original statistic. This implies that, under rather mild regularity conditions (such as continuity of the limiting distribution), if the null hypothesis holds then the bootstrap p-value p_n^* is asymptotically uniformly distributed, see Remark 3.9. Combined with establishing that the bootstrap statistic is bounded in probability under the alternative such that $p_n^* \xrightarrow{p} 0$, see Section 3.3, this was used to establish validity of the bootstrap. Cavaliere and Rahbek (2020) discuss in detail different definitions of asymptotic validity and their verification under different assumptions. An example is Cavaliere and Georgiev (2019), where validity of the bootstrap test is defined directly in terms of asymptotic uniformity of the bootstrap p-values, rather than the properties of (consistent) estimation of the limiting null distribution of the original test statistic. By doing so, the bootstrap can be employed also in cases where the limiting distribution of the original statistic may not be well defined, or in cases where the boostrap distribution does not converge in probability but, rather, converges in distribution, see e.g. Boswijk, Cavaliere, Georgiev and Rahbek (2020) and the discussions in Sections 5.1 and 5.4 where this was applied.

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.

Berkowitz, J. and L. Kilian (2000): "Recent developments in bootstrapping time series", *Econometric Reviews*, 19(1), 1-48.

Borkovec, M. and C. Klüppelberg (2001), "The Tail of the Stationary Distribution of an Autoregressive Process with ARCH(1) Errors", *The Annals of Applied Probability*, 11(4), 1220-1241.

Boswijk, H.P., G. Cavaliere, A. Rahbek and A.M.R. Taylor (2016): "Inference on Co-Integration Parameters in Heteroskedastic Vector Autoregressions", *Journal of Econometrics*, 192, 64–85.

Boswijk, H.P., G. Cavaliere, I. Georgiev and A. Rahbek (2019): "Bootstrapping Non-Stationary Stochastic Volatility", *Working paper, Tinbergen Institute,* TI 2019-083/III.

Bühlmann P. (1997): "Sieve Bootstrap for Time Series", *Bernoulli* 3, 123–148.

Cavaliere G., and I. Georgiev (2019): "Inference under Random Limit Bootstrap Measures", arXiv:2930679.

Cavaliere, G., I. Georgiev, and A.M.R. Taylor (2016). "Sieve-based inference for infinite-variance linear processes". *The Annals of Statistics*, 44(4), 1467–1494.

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.

—— (2020): "Bootstrapping Noncausal Autoregressions: With Applications to Explosive Bubble Modeling", *Journal of Business & Economic Statistics*, 38(1), 55-67.

Cavaliere G., H.B. Nielsen, R.S. Pedersen and A. Rahbek (2020): "Bootstrap Inference on the Boundary of the Parameter Space with Application to Conditional Volatility Models", *Journal of Econometrics*, forthcoming.

Cavaliere, G, R.S. Pedersen, and A. Rahbek (2018): "The Fixed Volatility Bootstrap for a Class of ARCH(q) Models", *Journal of Time Series Analysis*, 39(6), 920-941.

Cavaliere, G., and A. Rahbek (2020): "A Primer on Bootstrap Testing of Hypotheses in Time Series Models: with an Application to Double Autoregressive Models", *Econometric Theory*, 1–48.

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 Vector Autoregressive Heteroskedastic VAR Models", *Econometric Reviews*, 33, 606–650.

Cavaliere G., and I. Georgiev (2019): "Inference under Random Limit Bootstrap Measures", arXiv:2930679.

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.

Davis, R., and S. Resnick (1985a): "Limit Theory for Moving Averages of Random Variables with Regularly Varying Tail Probabilities," *The Annals of Probability*, 13, 179–195.

——— (1985b), "More Limit Theory for the Sample Correlation Function of Moving Averages," *Stochastic Processes and their Applications*, 20, 257–279.

——— (1986), "Limit Theory for the Sample Covariance and Correlation Functions of Moving Averages," *The Annals of Statistics*, 14, 533–558.

Davis, R., and L. Song (2020), "Noncausal vector AR processes with application to economic time series", *Journal of Econometrics,* 216(1), 246-267.

Duffee, G.R., and R.H. Stanton (2008), "Evidence on Simulation Inference for Near Unit-Root Processes with Implications for Term Structure Estimation", *Journal of Financial Econometrics*, 6(1), 108–142.

Durret, R. (2019), *Probability: Theory and Examples*, Fifth edition, Cambridge University Press.

Gonçalves, S. and L. Kilian (2004): "Bootstrapping Autoregressions with Conditional Heteroskedasticity of Unknown Form", *Journal of Econometrics* 123, 89–120.

—— (2007): "Asymptotic and Bootstrap Inference for AR(∞) Processes with Conditional Heteroskedasticity," *Econometric Reviews*, 26(6), 609-641.

Gourieroux, C., and Zakoïan, J.-M. (2017): "Local Explosion Modelling by Non-causal Process," *Journal of the Royal Statistical Society, Series B*, 79, 737–756.

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.

Hamilton, J.D. (1994): *Time Series Analysis*, Princeton University Press.

Hansen, B. (1996): "Inference when a Nuisance Parameter is not Identified under the Null Hupothesis", *Econometrica* 64(2), 413–430.

—— (2000): "Testing for Structural Change in Conditional Models", *Journal of Econometrics* 97, 93–115.

Harvey D., S.J. Leybourne, R.Sollis, and 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.

Hayashi, F. (2000): *Econometrics*, Princeton University Press.

Hendry, D. F. (1984): "Monte Carlo Experimentation in Econometrics". In Z. Grilishes & M. D. Intriligator (Eds.), *Handbook of Econometrics, Vol. 2*, 937–976.

Horowitz, J.L. (2001): "The Bootstrap". In Z. Grilishes & M. D. Intriligator (Eds.), *Handbook of Econometrics, Vol. 5*, 3159-3228.

—— (2003): "The Bootstrap in Econometrics", *Statistical Science,* 18(2), 211–218.

Johansen, S. (1996): *Likelihood-based Inference in Cointegrated Vector Autoregressive Models*, Oxford: Oxford University Press.

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.

Knight K. (1989): "On the Bootstrap of the Mean in the Infinite Variance Case", *Annals of Statistics*, 17, 1168–1175.

Kreiss, J. P. (1988): Asymptotical inference for a class of stochastic processes. *Habilitationsschrift,* Univ. Hamburg.

—— (1992): Bootstrap procedures for AR(∞)-processes. In *Bootstrapping and Related Techniques (Trier, 1990), Lecture Notes in Econom.and Math.Systems*, 107–113, Springer, Berlin.

Künsch, H. (1989): "The jackknife and the bootstrap for general stationary observations", *Annals of Statistics*, 17, 1217–1241.

Lahiri, S.N. (2003). *Resampling Methods for Dependent Data*, New York, Springer Verlag.

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.

Liu, R.Y. (1988): "Bootstrap Procedures under Some Non i.i.d. Models", *Annals of Statistics* 16, 1696–1708.

MacKinnon, J.G. (2006): "Bootstrap methods in econometrics", *The Economic Record*, 82, Special Issue, 2–18.

—— (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., Y. Wu, and J. Yu (2011): "Explosive Behavior in the 1990s Nasdaq: when did Exuberance Escalate Stock Values?" *International Economic Review* 52, 201–226.

Politis, D., and J. Romano, (1994): "The stationary bootstrap", *Journal of the American Statistical Association*, 89(428), 1303–1313.

Shao, X. (2010): "The Dependent Wild Bootstrap", *Journal of the American Statistical Association*, 105, 218–235.

van der Vaart, A.W. (2000): *Asymptotic Statistics*, Cambridge University Press.

White, H. (1980): "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity", *Econometrica*, 48, 817–838.

Appendix

A Proof of Lemma 3.2

Consider the bootstrap process $\{x_t^*\}_{t=0}^n$ as given by (3.5). Without loss of generality set $x_0^* = 0$, such that by simple recursion,

$$x_t^* = \sum_{i=0}^{t-1} \rho_0^i \varepsilon_{t-i}^* = \sum_{i=0}^{\infty} \rho_0^i \varepsilon_{t-i}^* = \rho(L)\,\varepsilon_t^*,$$

with $\varepsilon_j^* = 0$ for $j \leq 0$, L the lag operator, $L\varepsilon_t^* = \varepsilon_{t-1}^*$, and $\rho(z) = \sum_{i=0}^{\infty} \rho_0^i z^i$, for $z \in \mathbb{C}$. The summation to infinity means one can use standard manipulations of the lag polynomial $\rho(\cdot)$ as well known from time series literature, see e.g. Johansen (1996), Hamilton (1994) and references therein. Note first that as $|\rho_0| < 1$, the coefficients ρ_0^i in $\rho(z)$ are exponentially decreasing and $\rho(z)$ is convergent for $|z| < 1 + \delta$ for some $\delta > 0$ and $\rho(1) = (1-\rho_0)^{-1}$.

When considering the average of x_t^*, $n^{-1}\sum_{t=1}^n x_t^*$ one may for example use an expansion of $\rho(z)$ (see Remark A.1 below) of $\rho(z)$ around $z = 1$, such that

$$x_t^* = \rho(L)\,\varepsilon_{t-i}^* = \rho(1)\,\varepsilon_t^* + \rho^*(L)\,\Delta\varepsilon_t^*,$$

with $\rho^*(z) = \sum_{i=1}^{\infty} \theta_i z^i$ with $\theta_i = -i\rho_0^i$ exponentially decreasing. This immediately gives

$$n^{-1}\sum_{t=1}^n x_t^* = (1-\rho_0)^{-1} n^{-1} \sum_{t=1}^n \varepsilon_t^* + n^{-1}\rho^*(L)\,\varepsilon_n^*.$$

The first term tends to zero by (3.8). Next, for the second term use that ε_t^* are i.i.d. (conditionally on the data) such that

$$\mathrm{E}^* \left| n^{-1} \rho^* (L) \varepsilon_n^* \right| = n^{-1} \mathrm{E}^* \left| \sum_{i=1}^{\infty} i \rho_0^i \varepsilon_{n-i}^* \right| \leq n^{-1} \sum_{i=1}^{\infty} i |\rho_0|^i \mathrm{E}^* \left| \varepsilon_{n-i}^* \right| \ \leq c n^{-2} \sum_{t=1}^{n} |\tilde{\varepsilon}_t^c| \xrightarrow{p} 0,$$

with $c = |\rho_0| (1 - |\rho_0|)^{-2}$ and where $n^{-1} \sum_{t=1}^n |\tilde{\varepsilon}_t^c| \xrightarrow{p} \kappa < \infty$ for κ a constant by standard application of the LLN.

Turning to the average of x_t^{*2}, note that

$$n^{-1} \sum_{t=1}^{n} x_t^{*2} = n^{-1} \sum_{t=1}^{n} (\sum_{i=0}^{\infty} \rho_0^i \varepsilon_{t-i}^*)^2 = n^{-1} \sum_{t=1}^{n} \sum_{i=0}^{\infty} \rho_0^{2i} \varepsilon_{t-i}^{*2} + \delta_n^*$$

where

$$\delta_n^* = n^{-1} \sum_{t=1}^{n} \sum_{i \neq j=1}^{n} \rho_0^{i+j} \varepsilon_{t-i}^* \varepsilon_{t-j}^*.$$

With $\rho_2(z) = \sum_{i=0}^\infty \rho_0^{2i} z^i$ and $\omega_0 = \rho_2(1) \sigma_0^2 = \sigma_0^2 / (1 - \rho_0^2)$ the first term can be written as

$$n^{-1} \sum_{t=1}^{n} \sum_{i=0}^{\infty} \rho_0^{2i} \varepsilon_{t-i}^{*2} = n^{-1} \sum_{t=1}^{n} \rho_2(L) \eta_t^* + \omega_0,$$

where $\eta_t^* = \varepsilon_t^{*2} - \sigma_0^2$. Thus using arguments identical to above, $n^{-1} \sum_{t=1}^n \rho_2(L) \eta_t^* \xrightarrow{p^*}_p 0$.

It remains to show that $\delta_n^* \xrightarrow{p^*}_p 0$. First, observe that by definition of the double summation,

$$\delta_n^* = 2n^{-1} \sum_{t=1}^{n} (\sum_{i=0}^{\infty} \rho_0^i \varepsilon_{t-i}^* \sum_{j=1}^{\infty} \rho_0^{i+j} \varepsilon_{t-i-j}^*) = 2 \sum_{i=0}^{\infty} \rho_0^{2i} (n^{-1} \sum_{t=1}^{n} \varepsilon_{t-i}^* \sum_{j=1}^{\infty} \rho_0^j \varepsilon_{t-i-j}^*).$$

As $\mathrm{E}^* \left(\varepsilon_{t-i}^* \varepsilon_{t-j}^* \right) = 0$ for $j \neq i$, it follows that

$$\mathrm{E}^* \left(\delta_n^{*2} \right) = 4 \sum_{i=0}^{\infty} \rho_0^{4i} \mathrm{E}^* (n^{-1} \sum_{t=1}^{n} \varepsilon_{t-i}^* \sum_{j=1}^{\infty} \rho_0^j \varepsilon_{t-i-j}^*)^2 = 4n^{-2} \sum_{i=0}^{\infty} \rho_0^{4i} \sum_{t=1}^{n} \mathrm{E}^* (\varepsilon_{t-i}^* \sum_{j=1}^{\infty} \rho_0^j \varepsilon_{t-i-j}^*)^2.$$

Here

$$\begin{aligned} \mathrm{E}^* (\varepsilon_{t-i}^* \sum_{j=1}^{\infty} \rho_0^j \varepsilon_{t-i-j}^*)^2 &= \sum_{j=1}^{\infty} \sum_{m=1}^{\infty} \rho_0^{2j} \mathrm{E}^* \left(\varepsilon_{t-i}^{*2} \varepsilon_{t-i-j}^* \varepsilon_{t-i-m}^* \right) = \sum_{j=1}^{\infty} \rho_0^{2j} \mathrm{E}^* \left(\varepsilon_{t-i}^{*2} \varepsilon_{t-i-j}^{*2} \right) \\ &= \sum_{j=1}^{\infty} \rho_0^{2j} \left(\mathrm{E}^* \left(\varepsilon_t^{*2} \right) \right)^2 = c (n^{-1} \sum_{t=1}^{n} \tilde{\varepsilon}_t^2)^2 \xrightarrow{p} \tilde{c}, \end{aligned}$$

with $\tilde{c} = \rho_0^2 \sigma_0^4 / (1 - \rho_0^2)$.

Remark A.1 *The equality*

$$\rho(z) = \rho(1) + \rho^*(z)(1 - z),$$

with $\rho^*(z) = \sum_{i=1}^{\infty} \theta_i z^{i-1}$ *and* $\theta_i = -i\rho_0^i$ *follows by the identity*

$$\rho^*(z) = \tfrac{\rho(z)-\rho(1)}{1-z} = -\sum_{i=1}^{\infty} i\rho_0^i z^{i-1}.$$

Also note that the polynomial $\rho^*(z)$ *is convergent for* $|z| < 1+\delta$ *as the coefficients* $\theta_i = -i\rho_0^i$ *of* $\rho^*(z)$ *are exponentially decreasing.*

Figures and Tables

TABLE 1: Empirical rejection frequencies of asymptotic and bootstrap tests in the homoskedastic and heteroskedastic case.

	Homoskedastic case		Heteroskedastic case			
	(A)	(B)	(C)	(D)	(E)	(F)
n	Asymp.	iid boot.	Asymp.	iid boot.	Wild boot. N(0,1)	Wild boot. Rademacher
$\rho_0 = \bar{\rho} = 0.5$						
15	0.0825	0.0512	0.1319	0.1015	0.0720	0.0462
25	0.0733	0.0541	0.1329	0.1121	0.0685	0.0496
50	0.0623	0.0519	0.1370	0.1265	0.0625	0.0486
100	0.0538	0.0511	0.1425	0.1383	0.0586	0.0488
250	0.0521	0.0508	0.1428	0.1419	0.0593	0.0557
500	0.0498	0.0478	0.1398	0.1375	0.0526	0.0491
1000	0.0495	0.0493	0.1434	0.1463	0.0515	0.0502
$\rho_0 = \bar{\rho} = 0.9$						
15	0.1971	0.0558	0.1797	0.0597	0.0688	0.0534
25	0.1651	0.0552	0.1727	0.0688	0.0670	0.0563
50	0.1218	0.0526	0.1603	0.0875	0.0628	0.0530
100	0.0866	0.0507	0.1518	0.1016	0.0565	0.0501
250	0.0664	0.0511	0.1444	0.1226	0.0538	0.0494
500	0.0559	0.0480	0.1452	0.1323	0.0500	0.0474
1000	0.0548	0.0516	0.1463	0.1376	0.0515	0.0503

Notes: The data generating process is given by (4.1) with $\rho_0 \in \{0.5, 0.9\}$ *and* $\delta_0 = 0$ *and the bootstrap process defined in (4.2). In panels (A) and (B), the innovations* ε_t *are* $i.i.d.N(0,\sigma_0^2)$ *distributed with* $\sigma_0^2 = 1$*, while in panels (C)–(F)* ε_t *are independently* $N(0,\sigma_t^2)$ *distributed, with* σ_t^2 *given in (4.5).*

Table 2: Empirical rejection frequencies of asymptotic and bootstrap tests under the alternative.

$\bar{\rho}$	Asymptotic	iid bootstrap
0.9	0.0651	0.0485
0.875	0.0980	0.0810
0.85	0.2911	0.2658
0.825	0.5533	0.5283
0.8	0.7730	0.7567
0.75	0.9617	0.9578
0.7	0.9958	0.9954

Notes: The data generating process is given by (4.1) with $n = 250$ *and* $\rho_0 = 0.9$*, such that* $\rho_0 \neq \bar{\rho}$ *except for the first row entry.*

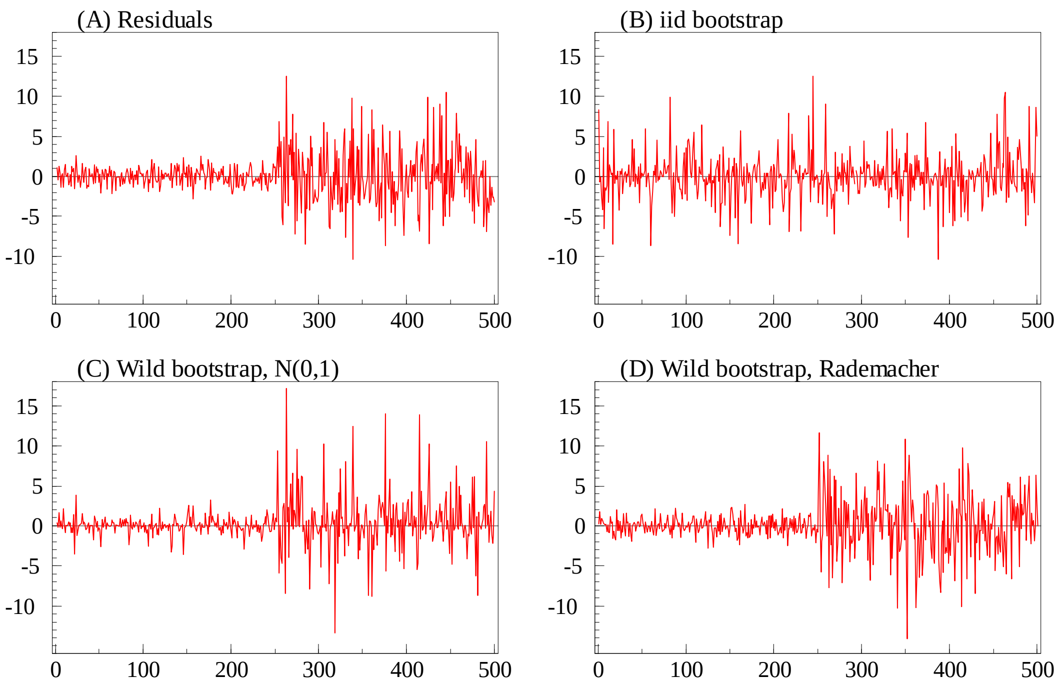

Figure 1: Residuals from a typical simulation with sample length $n = 500$. Panel (A) shows a sample of restricted residuals $\tilde{\varepsilon}_t^c$, while panels (B)-(D) show corresponding bootstrap innovations $\{\varepsilon_t^*\}_{t=1}^n$ for the iid bootstrap and the wild bootstrap, with the auxiliary w_t^* Gaussian and Rademacher distributed respectively.

TABLE 3: Empirical rejection frequencies with heavy tailed innovations.

	Asymptotic	iid	Wild	Permutation
n		Student's t_ν, $\nu = 3/2$		
15	0.1434	0.0395	0.0248	0.0499
25	0.0959	0.0382	0.0256	0.0520
50	0.0563	0.0350	0.0268	0.0488
100	0.0363	0.0340	0.0262	0.0443
250	0.0290	0.0350	0.0326	0.0480
500	0.0249	0.0333	0.0322	0.0465
1000	0.0278	0.0351	0.0348	0.0513
n		Student's t_ν, $\nu = 3$		
15	0.1809	0.0508	0.0494	0.0508
25	0.1512	0.0491	0.0475	0.0522
50	0.1129	0.0494	0.0496	0.0521
100	0.0806	0.0491	0.0486	0.0505
250	0.0596	0.0487	0.0514	0.0490
500	0.0578	0.0528	0.0515	0.0534
1000	0.0499	0.0465	0.0486	0.0481
n		Student's t_ν, $\nu = 5$		
15	0.1894	0.0519	0.0531	0.0511
25	0.1571	0.0486	0.0515	0.0499
50	0.1119	0.0523	0.0495	0.0507
100	0.0807	0.0503	0.0497	0.0503
250	0.0639	0.0486	0.0477	0.0486
500	0.0542	0.0484	0.0478	0.0485
1000	0.0540	0.0498	0.0501	0.0503

Notes: Empirical rejection frequencies of asymptotic and bootstrap tests with ε_t i.i.d. Student's t_ν for $\nu \in \{1, 3, 5\}$. Results are reported for the iid bootstrap, the wild boostrap with w_t Rademacher distributed and the permutation bootstrap. Simulations are reported for sample lengths $n \in \{15, ..., 1000\}$. The number of bootstrap replications is $B = 399$ and $N = 10,000$ replications.