TI 2019-083/III
Tinbergen Institute Discussion Paper

Bootstrapping Non-Stationary Stochastic Volatility

H. Peter Boswijk[1]
Giuseppe Cavaliere[2,3]
Iliyan Georgiev[3]
Anders Rahbek[4]

[1] Amsterdam School of Economics and Tinbergen Institute, University of Amsterdam
[2] University of Bologna
[3] Exeter Business School
[4] University of Copenhagen

BOOTSTRAPPING NON-STATIONARY STOCHASTIC VOLATILITY

H. Peter Boswijk[a], Giuseppe Cavaliere[b,c], Iliyan Georgiev[b] and Anders Rahbek[d]

Abstract

To what extent can the bootstrap be applied to conditional mean models — such as regression or time series models — when the volatility of the innovations is random and possibly non-stationary? In fact, the volatility of many economic and financial time series displays persistent changes and possible non-stationarity. However, the theory of the bootstrap for such models has focused on deterministic changes of the unconditional variance and little is known about the performance and the validity of the bootstrap when the volatility is driven by a non-stationary stochastic process. This includes near-integrated exogenous volatility processes as well as near-integrated GARCH processes, where the conditional variance has a diffusion limit; a further important example is the case where volatility exhibits infrequent jumps. This paper fills this gap in the literature by developing conditions for bootstrap validity in time series and regression models with non-stationary, stochastic volatility. We show that in such cases the distribution of bootstrap statistics (conditional on the data) is random in the limit. Consequently, the conventional approaches to proofs of bootstrap consistency, based on the notion of weak convergence in probability of the bootstrap statistic, fail to deliver the required validity results. Instead, we use the concept of 'weak convergence in distribution' to develop and establish novel conditions for validity of the wild bootstrap, conditional on the volatility process. We apply our results to several testing problems in the presence of non-stationary stochastic volatility, including testing in a location model, testing for structural change using CUSUM-type functionals, and testing for a unit root in autoregressive models. Importantly, we show that sufficient conditions for conditional wild bootstrap validity include the absence of statistical leverage effects, i.e., correlation between the error process and its future conditional variance. The results of the paper are illustrated using Monte Carlo simulations, which indicate that a wild bootstrap approach leads to size control even in small samples.

Keywords: Bootstrap; Non-stationary stochastic volatility; Random limit measures; Weak convergence in Distribution.

JEL Classification: C32.

---

[a] Amsterdam School of Economics and Tinbergen Institute, University of Amsterdam, The Netherlands.
[b] Department of Economics, University of Bologna, Italy
[c] Department of Economics, Exeter Business School, UK.
[d] Department of Economics, University of Copenhagen, Denmark
Correspondence to: Giuseppe Cavaliere, Department of Economics, University of Bologna, Piazza Scaravilli 2, I-40126 Bologna, Italy. email: giuseppe.cavaliere@unibo.it. We thank participants to the conference in honour of Pierre Perron (Boston University, 14-15 March 2019) for useful feedbacks on an earlier version of the paper. This research was supported by the Danish Council for Independent Research (DSF Grant 015-00028B) and by the University of Bologna (ALMA IDEA 2017 Grant).

1 Introduction

In this paper we consider bootstrap and asymptotic inference on the conditional mean in econometric time series models when the (conditional) volatility is allowed to show a large degree of persistence due to possible permanent and stochastic changes, reflecting the well established fact that volatility in many economic and financial time series displays high persistence, and non-covariance stationarity.

Earlier references in macroeconomics include Kim and Nelson (1999) and McConnell and Perez-Quiroz (2000), who find evidence of an (unanticipated) structural change in the volatility of US GDP growth rates. Evidence of changes in the unconditional volatility appear in many key time series, such as aggregate consumption and income, in interest rate data and in nominal and real price variables; see Sensier and van Dijk (2004). Evidence on changes in the long-run component of volatility in stock and currency markets are initially reported in Loretan and Phillips (1994) and Hansen (1995), who show that when stochastic volatility [SV] models are taken to the data, the largest autoregressive root in the SV process is so close to one that the assumption of stationary volatility seems to be at odds with the data. Similarly, it is a well-known stylized fact that GARCH models fit to stock market returns display parameter estimates which reflect high persistence as they (nearly) violate covariance stationarity conditions (often referred to as "near-integrated GARCH"), and that such parameters are smaller when a slowly-varying long run component is accounted for in the model, see Engle and Rangel (2008). Harvey *et al.* (2016) list a number of empirical studies that have found strong evidence of structural breaks in the unconditional variance of asset returns, with break dates driven by major financial and macroeconomic crises. Such (possibly random) volatility shifts are known to affect the asymptotic properties of estimators of the parameters of models for the conditional mean; see Cavaliere and Taylor (2007), Xu and Phillips (2008) and, for multivariate models, Cavaliere *et al.* (2010a,b) and Boswijk *et al.* (2016).

In the framework of a conditional mean, or a general (stationary, or non-stationary) regression type model, the *wild bootstrap* is an important tool to deliver consistent estimation of the asymptotic distributions of test statistics or parameter estimators. The wild bootstrap allows in particular to track changes in the quadratic variation of an econometric model by simply mimicking the (unknown) volatility dynamics through the squared model residuals, see Gonçalves and Kilian (2004, 2007) for applications to stationary time series models and Cavaliere *et al.* (2010a,b) for non-stationary multivariate models.

Consider the simple case where the volatility, say σ_t, can be approximated by a *non-stochastic* element of the space $\mathscr{D}[0,1]$ of càdlàg functions on $[0,1]$, such that $\sigma_t = \sigma(t/n)$ ($t = 1,\dots,n$, n denoting the sample size) with $\sigma \in \mathscr{D}[0,1]$. Simple special cases are a single volatility break at time $\lfloor n\tau \rfloor$ (with $\lfloor\cdot\rfloor$ denoting the (floor) integer value), for some $\tau \in (0,1)$, as given by (with $\mathbb{I}_A(\cdot)$ denoting the indicator function of the set A)

$$\sigma(u) := \sigma_A + (\sigma_B - \sigma_A)\mathbb{I}_{[\tau,1]}(u),$$

or the case of trending volatility,

$$\sigma(u) := \sigma_A + (\sigma_B - \sigma_A)u^{\delta},$$

where $\delta > 0$. A classic wild bootstrap, based on the resampling scheme $\varepsilon_t^* = \hat{\varepsilon}_t w_t^*$, where the $\hat{\varepsilon}_t$'s are the estimated residuals from the regression model, and the w_t^*'s are i.i.d. $(0,1)$ bootstrap shocks, independent of the original sample, is in general able to track the volatility path (in terms of quadratic variation) of the original data, as loosely speaking under standard assumptions,

$$n^{-1}\sum_{t=1}^{\lfloor nu \rfloor}(\varepsilon_t^*)^2 = n^{-1}\sum_{t=1}^{\lfloor nu \rfloor}\hat{\varepsilon}_t^2(w_t^*)^2 = n^{-1}\sum_{t=1}^{\lfloor nu \rfloor}\varepsilon_t^2 + o_p(1) = \int_0^u \sigma(u)^2 \mathsf{d}u + o_p(1).$$

Existing theory of the bootstrap mainly focuses on such deterministic changes of the unconditional variance and little is known about the performance and the validity[1] of the bootstrap when the volatility is driven by a high-persistent, or (second order) non-stationary stochastic process. This includes leading key cases such as near-integrated exogenous volatility processes (as analyzed by Hansen, 1995), as well as near-integrated GARCH processes, where the conditional variance has a diffusion limit (Nelson, 1990).

This paper fills this gap in the literature by developing conditions for bootstrap validity and consistency of the associated bootstrap tests in regression and time series models with persistent stochastic volatility. That is, we replace the deterministic volatility assumption by allowing that volatility is the realization of a (non-stationary) stochastic process σ_t; specifically, we derive results under the general assumption that, for the càdlàg version of the volatility, it holds that

$$\sigma_{\lfloor \cdot n \rfloor + 1} \overset{w}{\to} \sigma(\cdot), \tag{1}$$

where σ is some random element in $\mathscr{D}[0,1]$.

The analysis of the bootstrap under a weak convergence assumption like (1) is not straightforward. As we show in the paper, a key fact under non-stationary stochastic volatility is that the distribution of bootstrap statistics (conditional on the data), rather than converging to the unconditional distribution of the statistic of interest, converges weakly to a *random* limit. By this we mean that the distribution function of the bootstrap statistic (conditional on the data) is stochastic not only for finite sample size n, but also in the limit as $n \to \infty$. Consequently, the conventional approach, based on the notion of weak convergence in probability of the bootstrap statistic to the limiting distribution of the original statistic (which is obviously non-stochastic), fails to deliver the required result of validity of the bootstrap. This problem is not new in the bootstrap literature, as it appears in various areas of application of the bootstrap; for example, in models with infinite variance innovations (Knight, 1989) and in autoregressive models with unit roots (Basawa et al., 1991; Cavaliere *et al.*, 2015).

Specifically, in this paper we analyze the wild bootstrap under (non-stationary) stochastic volatility by adopting a new approach to assess bootstrap validity under random limit bootstrap measures. Thus, rather than focusing on the usual weak convergence in probability of the bootstrap conditional distribution, we make use of the concept of weak convergence in distribution (see Cavaliere and Georgiev, 2019, for a general introduction) to develop novel conditions for validity of the wild bootstrap,

[1] Throughout the paper, with 'validity' of the bootstrap we mean that the associated bootstrap tests control size asymptotically. With 'consistency' of the bootstrap test we mean that the (bootstrap) test rejects with probability tending to one under the alternative.

conditional on the volatility process. This allows us to establish that, although the presence of a *random* limiting distribution for the bootstrap statistic makes the bootstrap unable to estimate the unconditional distribution of the statistic of interest, the bootstrap can still deliver hypothesis tests with the desired size. In particular, we do this by establishing that the high-level conditions for bootstrap validity in Cavaliere and Georgiev (2019) can be shown to hold for a large class of models with stochastic volatility, including the aforementioned near-integrated GARCH model and the non-stationary stochastic volatility model. We do so by showing new weak convergence results conditional on volatility paths.

To illustrate our new approach and its applicability, we apply our results to three leading testing problems in the presence of non-stationary stochastic volatility, including testing a hypothesis on the location of a time series, testing for a unit root and testing for stability of the conditional mean using CUSUM-type statistics. These illustrative examples can easily be extended to cover more general cases, such as cointegration (as in Cavaliere *et al.*, 2010a, Cavaliere *et al.*, 2015 and Boswijk *et al.*, 2016) with multivariate stochastic volatility, or multivariate stability tests (see Perron, 2006 and Casini and Perron, 2019). Importantly, for all examples we show that conditions for conditional wild bootstrap validity include the absence of statistical leverage effects, i.e. correlation between the error process and its future conditional variance. The results of the paper are illustrated using Monte Carlo simulations, which indicate that under the conditions developed in our paper, the wild bootstrap leads to excellent size control even in small samples.

Structure of the paper

The structure of the paper is the following. In Section 2 we introduce the reference data generating process and our main assumptions, in particular on the volatility. Here we also introduce three examples which are used throughout the paper to illustrate the main results. We also derive the reference limit distribution for non-bootstrap statistics under non-stationary volatility. In Section 3 we introduce the main (wild) bootstrap algorithm. We show that when volatility is non-stationary, the bootstrap fails to mimic the asymptotic distribution of the corresponding statistics, and hence it is not valid in the usual sense. In Section 4 we discuss the wild bootstrap and prove, under proper assumptions, validity conditionally on the volatility path, as well as consistency of the bootstrap tests under the alternative. We first introduce in Section 4.1 the concept of weak convergence in distribution and discuss how to prove validity of the bootstrap in the presence of random limit bootstrap distributions, as it happens here under non-stationary volatility. Then in Section 4.2 we provide our main results under the required additional conditions on the original data. In Section 4.3 we apply our results about validity of the bootstrap in our three applications. Finally, in Section 4.4 we discuss consistency of the bootstrap tests under the alternative hypothesis. Results from a Monte Carlo study on the finite sample behavior of the bootstrap tests are reported in Section 5. Section 6 concludes. All proofs are reported in the Appendix A.

Notation

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

2 Set-up and preliminaries

In this section we introduce our reference class of model for the conditional mean under stochastic volatility as well as the (test) statistics of interest. In Section 2.1 we focus on statistics which can be expressed (at least when the associated null hypothesis holds true) as functionals of the partial sum of the innovations and of the partial sum of the squared innovations. To illustrate ideas, we focus on three simple univariate cases (which can easily be extended to multivariate cases) throughout: (i) testing a hypothesis on the mean in a simple location model; (ii) CUSUM testing for parameter constancy in a location model; (iii) testing for an autoregressive unit root in an AR(1) model. The main assumption on the volatility — which, *inter alia*, allows for non-stationary stochastic volatility or near-integrated GARCH dynamics — is discussed next in Section 2.2. Under the assumptions in Sections 2.1 and 2.2, the asymptotic (null) distributions can be derived. We do this in Section 2.3, where we show that the limiting distribution can be expressed in terms of a continuous martingale and its quadratic variation process. The implications of these results on bootstrap inference and hypothesis testing are the focus of the main Sections 3 and 4.

2.1 Model and hypotheses of interest

We are concerned with inference and hypothesis testing on the regression parameters of a heteroskedastic time series regression model in a triangular array form:

$$y_{n,t} = \beta' x_{n,t} + \varepsilon_{n,t}, \qquad t = 1, \ldots, n;\ n = 1, 2, \ldots \tag{2}$$

where $\varepsilon_{n,t}$ is a martingale difference sequence (mds) relative to a suitable filtration $\mathcal{F}_{n,t}$, with conditional variance $\sigma_{n,t}^2 = E(\varepsilon_{n,t}^2 | \mathcal{F}_{n,t-1})$. To simplify notation, unless strictly required we simply write (2) as $y_t = \beta' x_t + \varepsilon_t$, with $\sigma_t^2 := E(\varepsilon_t^2 | \mathcal{F}_{t-1})$.

Inference focuses on test statistics, which we assume can be expressed (at least under the null hypothesis) as functionals of the partial sum processes

$$(M_n(u), U_n(u)) := \left(n^{-1/2} \sum\nolimits_{t=1}^{\lfloor nu \rfloor} \varepsilon_t, n^{-1} \sum\nolimits_{t=1}^{\lfloor nu \rfloor} \varepsilon_t^2\right), \qquad u \in [0,1]. \tag{3}$$

as is the case for many testing problems, see also the discussion below.

Defining $\mathcal{F}_n(u) := \mathcal{F}_{\lfloor nu \rfloor}$, the mds assumption implies that $\{M_n(u), \mathcal{F}_n(u)\}_{u\in[0,1]}$ is a martingale for all n, and $U_n(u)$ is its quadratic variation process, i.e.,

$$U_n(u) := [M_n](u) = \sum_{t=1}^{\lfloor nu \rfloor} \left(M_n\left(\frac{t}{n}\right) - M_n\left(\frac{t-1}{n}\right)\right)^2, \qquad u \in [0,1]. \tag{4}$$

Throughout it will also be useful to define the predictable quadratic variation or angle bracket process (see Jacod and Shiryaev, 2003):

$$V_n(u) := \langle M_n \rangle(u) = n^{-1} \sum\nolimits_{t=1}^{\lfloor nu \rfloor} \sigma_t^2, \qquad u \in [0,1], \tag{5}$$

with the defining property that $\left\{M_n^2(u) - \langle M_n \rangle(u), \mathcal{F}_n(u)\right\}_{u\in[0,1]}$ is a martingale.

The following three testing problems are discussed in the paper. These are all special cases of (2) where the statistic of interest is indeed a functional of $(M_n(\cdot), U_n(\cdot))$.

Example 1 (testing in a location model). Consider the location model $y_t = \theta + \varepsilon_t$, which is trivially obtained from (2) by setting $\beta = \theta$ and $x_t = 1$. The true location parameter is denoted by θ_0. Suppose that interest is in testing the simple null hypothesis $\theta = \bar{\theta}$. Then, one can consider the test statistic $S_n := \sqrt{n}(\overline{y}_n - \bar{\theta})$, where $\overline{y}_n = n^{-1} \sum_{t=1}^n y_t$, or, alternatively, its studentized version $T_n := \sqrt{n}(\overline{y}_n - \bar{\theta})/s_n$, with $s_n^2 = n^{-1} \sum_{t=1}^n (y_t - \bar{y}_n)^2$. It is not difficult to see that, under the null hypothesis, it holds that S_n and T_n can be expressed in terms of M_n and U_n as

$$S_n = \sqrt{n}(\overline{y}_n - \bar{\theta}) = M_n(1), \qquad T_n = \sqrt{n}\frac{(\overline{y}_n - \bar{\theta})}{s_n} = \frac{M_n(1)}{\sqrt{U_n(1) - n^{-1}M_n(1)^2}}.$$

If s_n is constructed with the null imposed, i.e. $s_n^2 = n^{-1} \sum_{t=1}^n (y_t - \bar{\theta})^2$, then T_n simplifies to

$$T_n = \frac{M_n(1)}{\sqrt{U_n(1)}},$$

provided the null hypothesis holds true.

Example 2 (CUSUM test in a location model). Consider the time-varying location model $y_t = \theta_t + \varepsilon_t$, and suppose that interest is in testing the null hypothesis of a constant location parameter, i.e. $\mathsf{H}_0 : \theta_t = \theta_1$, $t = 2, \ldots, n$. A standard CUSUM test can be constructed by considering the statistic (see e.g. Deng and Perron, 2008, and the references therein)

$$CS_n := \frac{1}{n^{1/2}} \max_{t=1,\ldots,n} \left| \sum_{i=1}^{t} (y_i - \bar{y}_n) \right|,$$

or its studentized version,

$$CT_n := \frac{1}{s_n n^{1/2}} \max_{t=1,\ldots,n} \left| \sum_{i=1}^{t} (y_i - \bar{y}_n) \right|,$$

which, as in Example 1, reduce to

$$CS_n = \sup_{u\in[0,1]} |M_n(u) - uM_n(1)|, \qquad CT_n = \frac{\sup_{u\in[0,1]} |M_n(u) - uM_n(1)|}{\sqrt{U_n(1) - n^{-1}M_n^2(1)}}$$

under H_0.

Example 3 (Testing for a unit root) Consider the first-order autoregression $y_t = (1+\theta)y_{t-1} + \varepsilon_t$, with $y_0 = 0$ (which again follows from (2) by setting $\beta = 1+\theta$ and $x_t = y_{t-1}$). A test of the unit root hypothesis $\theta = 0$ can be based on the Dickey-Fuller 'coefficient' statistic $R_n := n\widehat{\theta}_n$, where $\widehat{\theta}_n = \sum_{t=1}^n y_{t-1}\Delta y_t / \sum_{t=1}^n y_{t-1}^2$ is the least-squares estimator from the regression of Δy_t on y_{t-1}. Under the null hypothesis, $\widehat{\theta}_n = \sum_{t=1}^n \varepsilon_t (\sum_{i=1}^{t-1} \varepsilon_i)/\sum_{t=1}^n (\sum_{i=1}^{t-1}\varepsilon_i)^2$ and the test statistic may be expressed as

$$R_n = \frac{\int_0^1 M_n(u)\mathsf{d}M_n(u)}{\int_0^1 M_n^2(u)\mathsf{d}u} = \frac{\frac{1}{2}\left(M_n^2(1) - U_n(1)\right)}{\int_0^1 M_n^2(u)\mathsf{d}u}.$$

If the test is based on the Dickey-Fuller 'ratio' statistic $W_n := \widehat{\theta}_n (s_n/(\sum_{t=1}^n y_{t-1}^2))^{-1/2}$, where $s_n^2 := n^{-1}\sum_{t=1}^n (\Delta y_t - \hat{\theta} y_{t-1})^2$, then

$$W_n = \frac{\int_0^1 M_n(u)\mathsf{d}M_n(u)}{\sqrt{\int_0^1 M_n^2(u)\mathsf{d}u}} \frac{1}{\sqrt{U_n(1) - n^{-1}(\int_0^1 M_n(u)\mathsf{d}M_n(u))^2/\int_0^1 M_n^2(u)\mathsf{d}u}}$$

under the null hypothesis. □

Some remarks are in order.

Remark 2.1 Although for fixed n, $U_n(\cdot)$ can be determined from $M_n(\cdot)$ as seen in (4), it does not define a continuous function $h : \mathscr{D}[0,1] \to \mathscr{D}[0,1]$. Therefore, limit results for $U_n(\cdot)$ cannot be obtained from weak convergence of $M_n(\cdot)$ together with the continuous mapping theorem [CMT]. Joint weak convergence of $(M_n(\cdot), U_n(\cdot))$ is required to obtain the asymptotic null distribution of the statistics T_n, CT_n, R_n and W_n. This is related to the well-known fact that weak convergence of $\int_0^1 M_n \mathsf{d}M_n$ to the stochastic integral $\int_0^1 M\mathsf{d}M$ does not follow from $M_n(\cdot) \stackrel{w}{\to} M(\cdot)$ and the CMT; see e.g. Chan and Wei (1988).

Remark 2.2 The above examples involve single-parameter models. In more general testing situations, such as testing for a unit root in higher-order autoregressive models, the statistic of interest may be written as a functional of $(M_n(\cdot), U_n(\cdot))$ plus an asymptotically negligible term. The theory developed in this paper can be extended to cover such cases. □

2.2 Non-stationary stochastic volatility

We now introduce our basic hypotheses on the dynamic behavior of the conditional volatility σ_t^2 of the shocks ε_t. More specifically, we will allow volatility to be a persistent stochastic process, with a stochastic volatility weak limit, as formulated in the next two assumptions. These are in the spirit of the seminal paper by Hansen (1995), who considers conditional variances driven by nearly-integrated autoregressive shocks, although we do not constrain the behavior of the conditional variance to be of the autoregressive type.

Assumption 1 *In (2), we have $\varepsilon_t = \sigma_t z_t$, where z_t is an martingale difference sequence relative to $\mathcal{F}_t = \sigma(\{z_i\}_{i=1}^{t}, \{\sigma_i\}_{i=1}^{t+1})$, satisfying $E(z_t^2|\mathcal{F}_{t-1}) = 1$.*

Define now the $\mathscr{D}[0,1]$ version of the partial sum of the z_t's as $B_{z,n}(u) := n^{-1/2}\sum_{t=1}^{\lfloor nu \rfloor} z_t$, $u \in [0,1]$, and the $\mathscr{D}[0,1]$ version of σ_t as:

$$\sigma_n(u) := \sigma_{\lfloor nu \rfloor + 1}, \text{ for } u \in [0,1), \tag{6}$$

with $\sigma_n(1) := \sigma_n$.

Assumption 2 *As $n \to \infty$, $(\sigma_n(\cdot), B_{z,n}(\cdot)) \stackrel{w}{\to} (\sigma(\cdot), B_z(\cdot))$, where $\sigma(\cdot)$ is a stochastic process in $\mathscr{D}[0,1]$ satisfying $\inf_{u\in[0,1]}\sigma(u) > 0$ a.s., and $B_z(\cdot)$ is a standard Brownian motion on $[0,1]$.*

While the convergence of the partial sum $B_{z,n}(\cdot)$ is standard, the requirement on $\sigma_n(\cdot)$ is not. More specifically, this assumption requires the conditional variance process σ_t to be persistent enough such that its behavior can be approximated by an element of the space of càdlàg functions $\mathscr{D}[0,1]$. Some examples of processes satisfying Assumptions 1 and 2 are presented next. These will be analyzed in detail throughout the paper; for additional cases and discussions see e.g. Cavaliere and Taylor (2009).

Example V.1 (Stochastic volatility) *Let σ_t^2 be generated by*

$$\log \sigma_t^2 = \phi_n \log \sigma_{t-1}^2 + (1-\phi_n)\log \bar{\sigma}^2 + n^{-1/2}\eta_{t-1}, \qquad t = 1, 2, \ldots$$

where $\sigma_0^2 = \bar{\sigma}^2$ for some $\bar{\sigma} > 0$, where $\phi_n = e^{-\kappa/n}$ for some $\kappa \geq 0$, and where $\eta_t \sim$ i.i.d. $N(0, \sigma_\eta^2)$, independent of $z_t \sim$ i.i.d. $N(0,1)$. Then

$$\sigma_n^2(\cdot) = \bar{\sigma}^2 \exp\left(n^{-1/2}\sum_{t=1}^{\lfloor n\cdot \rfloor} \phi_n^{\lfloor n\cdot \rfloor - t}\eta_{t-1}\right) \stackrel{w}{\to} \bar{\sigma}^2 \exp\left(\sigma_\eta e^{-\kappa\cdot}\int_0^{\cdot} e^{\kappa u} \mathsf{d}B_\eta(u)\right) =: \sigma^2(\cdot),$$

independently of and hence jointly with $B_{z,n}(\cdot) \stackrel{w}{\to} B_z(\cdot)$; where (B_η, B_z) is a bivariate standard Brownian motion.

Example V.2 (Near-integrated GARCH) *Consider the case where σ_t^2 is generated by the standard GARCH recursion*

$$\sigma_t^2 = \omega_n + \alpha_n \varepsilon_{t-1}^2 + \beta_n \sigma_{t-1}^2 = \omega_n + \alpha_n \sigma_{t-1}^2 z_{t-1}^2 + \beta_n \sigma_{t-1}^2, \qquad t = 1, 2, \ldots$$

where $\sigma_0^2 = \bar{\sigma}^2$ *for some* $\bar{\sigma} > 0$*, where* $\alpha_n + \beta_n = 1 - n^{-1}\kappa$ *for some* $\kappa \geq 0$*, where* $\omega_n = n^{-1}\bar{\sigma}^2\kappa$ *and* $\alpha_n = (2n)^{-1/2}\sigma_\eta$ *for some* $\sigma_\eta > 0$*, and where* $z_t \sim$ i.i.d. $N(0,1)$. *Then it follows from Nelson (1990) that* $(\sigma_n(\cdot), B_{z,n}(\cdot)) \xrightarrow{w} (\sigma(\cdot), B_z(\cdot))$*, where*

$$\mathsf{d}\sigma^2(u) = \kappa(\sigma^2(u) - \bar{\sigma}^2)\mathsf{d}u + \sigma_\eta\sigma^2(u)\mathsf{d}B_\eta(u), \qquad u \in [0,1],$$

with $\sigma^2(0) = \bar{\sigma}^2$*, and where* (B_η, B_z) *is a bivariate standard Brownian motion.* □

Remark 2.3 In both examples, the process generating σ_t^2 depends on the sample size n, so that $\{\sigma_{nt}^2\}_{1\leq t\leq n; n\geq 1}$ is actually a triangular array. We will not make this explicit in the notation in this section.

Remark 2.4 The main difference between the examples is that in Example V.1, the volatility shocks $\{\eta_t\}_{t\geq 1}$ are independent of $\{z_t\}_{t\geq 1}$, whereas in Example V.2, the variance is driven by $\eta_t = (z_t^2 - 1)/\sqrt{2}$, which is fully determined by (although uncorrelated with) z_t. In both examples, however, the limiting volatility process σ is independent of the Brownian motion B_z generated by $\{z_t\}_{t\geq 1}$. If we think of ε_t as the deviation of a financial return from its conditional expectation, and σ_t as its conditional volatility, then this rules out so-called leverage effects, i.e., asymmetric effects of positive and negative return shocks ε_t on future volatility σ_{t+h}, $h > 0$. Although the results given in this section also apply to processes with leverage, we will assume (asymptotic) independence in order to establish bootstrap validity.

Remark 2.5 In Example V.1, the log-volatility follows a near-integrated first-order autoregression, converging weakly to an Ornstein-Uhlenbeck [OU] process. Similarly, in Hansen (1995) σ_t satisfies Assumption 2 with $\sigma(\cdot)$ a (possibly nonlinear) transformation of an OU process (or Brownian motion). Cases where the volatility is allowed to jump at a countable number of times (while being constant between these jump times) are also allowed by our assumption. For instance, let $\sigma_t = \exp(\omega_0 + \omega_1 J_t)$, $J_t := \sum_{i=1}^t \delta_i\eta_i$, $J_0 = 0$ a.s., where for all t, δ_t is a Bernoulli random variable which equals one if and only if a volatility jump occurs at time t. If the η_t's (which denote the random jump sizes) are i.i.d., independent of δ_t's, and if $P(\delta_t = 1) = \lambda n^{-1}$, then (see e.g. Georgiev, 2008) $J_n(\cdot) := J_{\lfloor n\cdot\rfloor}$ converges weakly to the compound Poisson process $C_\lambda(\cdot) := \sum_{i=1}^{N(\cdot)} \eta_i$, where $N(\cdot)$ is a Poisson process in $\mathscr{D}[0,1]$ with intensity parameter λ. As expected, the limiting volatility process is $\sigma(\cdot) = \exp(\omega_0 + \omega_1 C_\lambda(\cdot))$, a piecewise constant process with number of discontinuities given by $N(1)$. □

2.3 Standard asymptotics under non-stationary stochastic volatility

Assumptions 1 and 2 allow to analyze the asymptotic behavior of the functional (M_n, U_n), as is done in the following Lemma.

Lemma 1 *Under Assumptions 1 and 2, we have as* $n \to \infty$*,*

$$(M_n(\cdot), U_n(\cdot)) \xrightarrow{w} (M(\cdot), V(\cdot)) := \left(\int_0^{\cdot} \sigma(u)\mathsf{d}B_z(u), \int_0^{\cdot} \sigma^2(u)\mathsf{d}u\right), \tag{7}$$

with $V(\cdot) = \langle M \rangle (\cdot)$. *Furthermore,*

$$\sup_{u \in [0,1]} |U_n(u) - V_n(u)| \xrightarrow{p} 0. \tag{8}$$

The implications for the testing problems in Section 2.1 are given next.

Example 1 (cont'd). Consider the location model of Example 1. A straightforward application of Lemma 1 along with the CMT yields that, under H_0, $S_n = M_n(1) \xrightarrow{w} M(1)$, which corresponds to the mixed normal distribution $N(0, \int_0^1 \sigma^2(u) \mathsf{d}u)$ and hence is non-pivotal. For T_n it holds that

$$T_n = \frac{M_n(1)}{\sqrt{U_n(1)}} + o_p(1) \xrightarrow{w} \frac{M(1)}{\sqrt{V(1)}} = \frac{\int_0^1 \sigma(u) \mathsf{d}B_z(u)}{\sqrt{\int_0^1 \sigma^2(u) \mathsf{d}u}}. \tag{9}$$

Notice that the limit distribution in (9) is non-pivotal in cases where $\sigma(\cdot)$ and $B_z(\cdot)$ are not stochastically independent. In contrast, should independence hold, then (9) corresponds to a standard Gaussian distribution.

Example 2 (cont'd). For the CUSUM test statistics of Example 2 it holds that, again by Lemma 1 and the CMT, that under H_0

$$\begin{aligned} CS_n = \sup_{u \in [0,1]} |M_n(u) - uM_n(1)| &\xrightarrow{w} \sup_{u \in [0,1]} |M(u) - uM(1)|, \\ CT_n = \frac{\sup_{u \in [0,1]} |M_n(u) - uM_n(1)|}{\sqrt{U_n(1) - n^{-1} M_n^2(1)}} &\xrightarrow{w} \frac{\sup_{u \in [0,1]} |M(u) - uM(1)|}{\sqrt{V(1)}}. \end{aligned}$$

Both statistics have a non-pivotal asymptotic null distribution, even in cases where the limit stochastic volatility process and the limit Brownian motions are stochastically independent.

Example 3 (cont'd). Finally, as shown in Cavaliere and Taylor (2009), for the unit root testing problem the presence of non-stationary volatility renders the null distribution of the Dickey-Fuller coefficient and t-statistics non-pivotal. More specifically, under the unit root null hypothesis it holds that

$$R_n = \frac{\frac{1}{2}\left(M_n^2(1) - U_n(1)\right)}{\int_0^1 M_n^2(u) \mathsf{d}u} \xrightarrow{w} \frac{\frac{1}{2}(M(1)^2 - V(1))}{\int_0^1 M^2(u) \mathsf{d}u} = \frac{\int_0^1 M(u) \mathsf{d}M(u)}{\int_0^1 M^2(u) \mathsf{d}u},$$

and

$$W_n = \frac{\int_0^1 M_n(u) \mathsf{d}M_n(u)}{\sqrt{\int_0^1 M_n^2(u) \mathsf{d}u}} \frac{1}{\sqrt{U_n(1) + o_p(1)}} \xrightarrow{w} \frac{\int_0^1 M(u) \mathsf{d}M(u)}{\sqrt{V(1) \int_0^1 M^2(u) \mathsf{d}u}}$$

as the sample size diverges. □

3 Bootstrap under non-stationary stochastic volatility

Consider the standardized sample mean statistic S_n for the location model, see Example 1. Lemma 1 implies that under the null hypothesis $S_n \xrightarrow{w} M(1) = \int_0^1 \sigma \mathsf{d}B_z$, see (7). The distribution of $M(1)$ depends on the volatility process $\sigma(\cdot)$, implying that critical values cannot be tabulated without providing a complete specification of this process. When the Brownian motion B_z and the limit volatility process σ are independent, then in the location model of Example 1 this problem can be avoided by considering the studentized test statistic T_n, which has a standard normal limit distribution under the null. However, in other testing problems (such as those in Examples 2 and 3) it is generally not possible to find such asymptotically pivotal statistics.

This motivates the development of bootstrap tests. Following much of the literature (e.g. Cavaliere and Taylor, 2008, 2009), we consider the wild bootstrap, which replicates the volatility patterns in the original data. Let w_t^* be an i.i.d. sequence with mean zero and variance[2] 1, independent of $\{\sigma_t, z_t\}_{t\geq 1}$, and define the bootstrap shocks as

$$\varepsilon_t^* = \varepsilon_t w_t^*, \qquad t = 1, 2, \ldots, n.$$

Accordingly, we can define the bootstrap partial sum and the bootstrap partial sum of squares,

$$(M_n^*(\cdot), U_n^*(\cdot)) = \left(n^{-1/2} \sum\nolimits_{t=1}^{\lfloor n\cdot \rfloor} \varepsilon_t^*, n^{-1} \sum\nolimits_{t=1}^{\lfloor n\cdot \rfloor} (\varepsilon_t^*)^2 \right).$$

These processes are the bootstrap analogs of the processes M_n and U_n of Section 2. Notice that this implementation of the bootstrap assumes that ε_t is observed under the null hypothesis, which is the case in Examples 1 and 3 (location and unit root test). In more general testing problems, including Example 2 (CUSUM test), ε_t will be replaced by some residuals $\hat{\varepsilon}_t$ (either restricted by the null hypothesis or unrestricted).

3.1 Failure of classic bootstrap validity

Classic validity of the bootstrap (usually denoted as 'bootstrap consistency') is usually understood as the convergence in probability (or almost surely) of the conditional (on the original data) cdf of the bootstrap statistic to the limit cdf of the original statistic. We show here that, in the presence of stochastic volatility as in the previous section, in general classic validity of the bootstrap fails. This is essentially because the conditional cdf of the bootstrap statistic remains random in the limit. In this section we discuss this fact and its implications on bootstrap inference using, as the reference bootstrap algorithm, a wild bootstrap scheme as is typically applied when the data are heteroskedastic.

Focusing again on the location statistic S_n, its bootstrap counterpart is $S_n^* = n^{-1/2} \sum_{t=1}^{n} \varepsilon_t^* = M_n^*(1)$ where M_n^* is as previously defined. Define P^* as the bootstrap measure conditional on the original data, D_n. The bootstrap (conditional) cdf is

$$F_n^*(x) := P^*(S_n^* \leq x) = P(S_n^* \leq x | D_n).$$

[2]Further conditions on the moments of w_t^* may be required in some specific applications.

The classical condition for bootstrap validity is that, as $n \to \infty$, $S_n^* \xrightarrow{w^*}_p S := M(1)$. If the limit cdf $F(x) := P(S \le x)$ is continuous, then this weak convergence in probability corresponds to the property

$$\sup_{x \in \mathbb{R}} |F_n^*(x) - F(x)| \xrightarrow{p} 0, \tag{10}$$

where F denotes the cdf of the asymptotic distribution of S_n:

$$F(x) := P(N(0, V(1)) \le x) = \int \Phi(V(1)^{-1/2}x) \mathsf{d}P(V(1)),$$

where $\Phi(\cdot)$ is the standard normal cdf. Because $F(x)$ is the marginal cdf of $S = M(1)$, under independence of the processes B_z and σ it corresponds to the cdf of the mixed normal random variable $S = V(1)^{1/2}Z$, where $Z \sim N(0,1)$, independent of $V(1)$.

However, under Assumption 2, condition (10) fails to hold, which is seen as follows. Choosing $w_t^* \sim$ i.i.d. $N(0,1)$ for convenience, it is seen that

$$S_n^*|D_n \sim N(0, U_n(1))|U_n(1). \tag{11}$$

This follows because, conditional on the original data,

$$S_n^* = n^{-1/2} \textstyle\sum_{t=1}^n \varepsilon_t w_t^* \sim N(0, n^{-1} \sum_{t=1}^n \varepsilon_t^2) \sim N(0, U_n(1)).$$

In terms of the conditional distribution F_n^* of S_n^* given the data D_n, (11) corresponds to

$$\begin{aligned} F_n^*(x) := P^*(S_n^* \le x) &= P(N(0, U_n(1)) \le x|U_n(1)) \\ &= P(N(0,1) \le U_n(1)^{-1/2}x|U_n(1)) = \Phi(U_n(1)^{-1/2}x). \end{aligned}$$

Letting $n \to \infty$, the limit distribution of the bootstrap statistic given the data follows from Lemma 1 and the CMT. Specifically, we have that

$$F_n^*(x) \xrightarrow{w} \Phi(V(1)^{-1/2}x) \tag{12}$$

for all $x \in \mathbb{R}$; eq. (12) implies that the limit distribution of the bootstrap cdf is in fact *random*. That is, not only the bootstrap cdf F_n^* converges weakly rather than in probability, but also the limiting cdf is random, as it depends on the random variable $V(1)$. Therefore, there is no reason to expect that the difference between the random function F_n^* and the non-random function F converges in probability to 0, as required for standard bootstrap consistency to apply.

The fact that the conditional cdf F_n^* converges weakly (in $\mathscr{D}(\mathbb{R})$), rather than in probability, to a random cdf, we label 'weak convergence in distribution', and denote as '$\xrightarrow{w}_w$'. More specifically, for sequences of random variables (Z_n, Y_n) and (Z, Y) (possibly defined on different probability spaces), the notation $Z_n|Y_n \xrightarrow{w}_w Z|Y$, when the conditional distribution of $Z|Y$ is diffuse (non-atomic), means that

$$F_n(\cdot|Y_n) := P(Z_n \le \cdot|Y_n) \xrightarrow{w} P(Z \le \cdot|Y) =: F(\cdot|Y).$$

A more general definition of $Z_n|Y_n \xrightarrow{w}_w Z|Y$, applicable to the case where (Z_n, Y_n) and (Z, Y) are random elements of a metric space $\mathcal{S}_Z \times \mathcal{S}_Y$ (and hence to stochastic processes), is that $E(g(Z_n)|Y_n) \xrightarrow{w} E(g(Z)|Y)$ for all bounded continuous functions $g : \mathcal{S}_Z \to \mathbb{R}$, see Cavaliere and Georgiev (2019) and the references therein.

When Z_n represents a bootstrap statistic and the conditioning set Y_n is the original data D_n, we use the notation '$\xrightarrow{w^*}_w$'. Hence, eq. (12) corresponds to the weak convergence in distribution

$$S_n^* \xrightarrow{w^*}_w N(0, V(1))|V(1).$$

Unless $V(1)$ is non-random (which is not the case under stochastic volatility as considered here), this convergence shows that the limit bootstrap measure is indeed a random measure. Hence, the bootstrap cannot be valid in the usual sense of weak convergence in probability of F_n^* to F.

3.2 Examples (continued)

The result in Section 3.1 applies to the other examples considered, except for the asymptotically pivotal statistic T_n.

Example 1 (cont'd). Consider the location model example and assume that the bootstrap data are generated as $y_t^* = \varepsilon_t^*$, with ε_t^* as defined above. The bootstrap test statistics are $S_n^* := \sqrt{n}\bar{\varepsilon}_n^*$ and $T_n^* := \sqrt{n}\bar{\varepsilon}_n^*/s_n^*$, $s_n^* = (n^{-1}\sum_{t=1}^n(\varepsilon_t^* - \bar{\varepsilon}_n^*)^2)^{1/2}$. Using the argument discussed above, we have that $S_n^* \xrightarrow{w^*}_w M^*(1)|V(1)$, where $M^*(u) := \int_0^u \sigma(s)\mathsf{d}B_z^*(s)$ with B_z^* a standard Brownian motion, stochastically independent of σ and B_z, and $V(1) := \int_0^1 \sigma(u)^2\mathsf{d}u$. The distribution of $M^*(1)|V(1)$ corresponds to the normal distribution $N(0, V(1))|V(1)$. In contrast, for T_n^* it holds that

$$T_n^* \xrightarrow{w^*}_w \left.\frac{M^*(1)}{\sqrt{V(1)}}\right| V(1),$$

which corresponds to a $N(0,1)$ random variable, as $M^*(1)|\,V(1) \overset{d}{=} N(0, V(1))|V(1)$. Since weak convergence of a conditional distribution to a non-random cdf corresponds to weak convergence in probability, in this special case $T_n^* \xrightarrow{w^*}_p N(0,1)$.

Two facts are worth stressing. First, in the above representations of the limit conditional distribution of the bootstrap statistic, σ and B_z^* are independent, even if the original processes σ and B_z are not. This result stems from the assumption that the wild bootstrap shocks w_t^* are independent of the original data. Second, if σ and B_z are stochastically independent, then $M^*(1)|V(1) \overset{d}{=} M(1)|V(1)$. This distributional equality is crucial to determine validity of the bootstrap.

Example 2 (cont'd). For the bootstrap CUSUM statistics, suppose that the bootstrap data are generated as $\varepsilon_t^* = \hat{\varepsilon}_t w_t^*$ where $\hat{\varepsilon}_t := y_t - \bar{y}_n$, such that when the null hypothesis is true $\hat{\varepsilon}_t = \varepsilon_t - \bar{\varepsilon}_n$. The bootstrap statistics are defined as

$$CS_n^* = \sup_{u\in[0,1]} |M_n^*(u) - uM_n^*(1)|, \quad CT_n^* = \frac{\sup_{u\in[0,1]} |M_n^*(u) - uM_n^*(1)|}{\sqrt{U_n^*(1) - n^{-1}M_n^*(1)^2}},$$

with $M_n^*(u) := n^{-1/2}\sum_{t=1}^{\lfloor n\cdot\rfloor}\varepsilon_t^*$ as above and $U_n^*(1) := n^{-1}\sum_{t=1}^{n}(\varepsilon_t^*)^2$. Under the null hypothesis,

$$n^{-1/2}\sum_{t=1}^{\lfloor n\cdot\rfloor}\varepsilon_t^* = n^{-1/2}\sum_{t=1}^{\lfloor n\cdot\rfloor}\hat{\varepsilon}_t w_t^* = n^{-1/2}\sum_{t=1}^{\lfloor n\cdot\rfloor}\varepsilon_t w_t^* + O_p(n^{-1/2}),$$

and it holds that

$$\begin{aligned} CS_n^* &\overset{w^*}{\to}_w \sup_{u\in[0,1]} |M^*(u) - uM^*(1)| \Big| \sigma, \\ CT_n &\overset{w^*}{\to}_w \frac{\sup_{u\in[0,1]}|M^*(u) - uM^*(1)|}{\sqrt{V(1)}} \Bigg| \sigma, \end{aligned}$$

where again $M^*(u) := \int_0^u \sigma(s)\mathsf{d}B_z^*(s)$ with B_z^* a standard Brownian motion, stochastically independent of σ and B_z, and $V(1) := \int_0^1 \sigma(u)^2\mathsf{d}u$. Both bootstrap statistics have a *random* non-pivotal asymptotic null distribution.

EXAMPLE 3 (CONT'D). Finally, consider the unit root example. To avoid the problems described in Basawa et al. (1991), the bootstrap data are generated with the unit root imposed, i.e. $y_t^* = y_{t-1}^* + \varepsilon_t^*$, with $y_0^* = 0$ and $\varepsilon_t^* := (\Delta y_t)w_t^*$; see e.g. Cavaliere and Taylor (2008). Under the null, clearly $\varepsilon_t^* := \varepsilon_t w_t^*$. As discussed earlier for the non-bootstrap case, we have that

$$R_n^* = \frac{\frac{1}{2}\left(M_n^{*2}(1) - U_n^*(1)\right)}{\int_0^1 M_n^{*2}(u)\mathsf{d}u}, \tag{13}$$

and, up to a negligible term,

$$W_n^* = \frac{\frac{1}{2}\left(M_n^{*2}(1) - U_n^*(1)\right)}{\sqrt{U_n^*(1)\int_0^1 M_n^{*2}(u)\mathsf{d}u}}.$$

In this case it holds that

$$R_n^* \overset{w^*}{\to}_w \frac{\int_0^1 M^*(u)\mathsf{d}M^*(u)}{\int_0^1 M^{*2}(u)\mathsf{d}u}\Bigg|\sigma, \quad W_n^* \overset{w^*}{\to}_w \frac{\int_0^1 M^*(u)\mathsf{d}M^*(u)}{\sqrt{V(1)\int_0^1 M^{*2}(u)\mathsf{d}u}}\Bigg|\sigma, \tag{14}$$

where $M^*(u) := \int_0^u \sigma(s)\mathsf{d}B_z^*(s)$ with B_z^* a standard Brownian motion, stochastically independent of σ and B_z, and $V(u) := \int_0^u \sigma(s)^2\mathsf{d}s$. The asymptotic distributions in (14) are random, except in the special case where σ is non-stochastic. □

Cavaliere and Georgiev (2019) provide a number of other examples where the bootstrap validity condition (10) fails for any non-random cdf F, and develop an alternative criterion for *conditional* bootstrap validity. We will apply this concept to the present situation and extend it to the analysis of consistency of the bootstrap tests in the next section.

4 Validity of the bootstrap

Despite the fact that under non-stationary stochastic volatility the bootstrap is unable to consistently estimate the limiting distribution of the original statistic, it can still be valid, in the sense that it delivers control over type one error probabilities as n diverges. This can be seen by focusing on the bootstrap p-value. Taking the statistic S_n and associated bootstrap analog S_n^* to illustrate, the bootstrap p-value is defined as

$$p_n^* := P(S_n^* \leq S_n | D_n) = F_n^*(x)|_{x=S_n},$$

where $F_n^*(\cdot)$ is the cdf of S_n^*, conditional on the data (this definition of the p-value assumes a left-tailed test, which will be assumed below unless indicated otherwise). As in Cavaliere and Georgiev (2019), we say that the bootstrap based on S_n, S_n^* is valid (conditionally on the volatility process $\sigma_n := \{\sigma_t\}_{t=1}^n$) if p_n^* is asymptotically $U(0,1)$ distributed conditionally on σ_n, i.e.

$$P(p_n^* \leq q|\sigma_n) \xrightarrow{p} q, \quad q \in (0,1). \tag{15}$$

If this is the case, even if (as shown in the previous section) the limiting conditional distribution of S_n^* is random, the bootstrap test can still be correctly sized in large samples. Moreover, proofs of validity in the form of (15) also imply that, unconditionally,

$$P(p_n^* \leq q) \to q, \quad q \in (0,1).$$

In the next subsections we discuss a set of sufficient conditions for (15) to hold. These are new in the literature on bootstrapping conditional mean models when the volatility can be stochastic. First, in Section 4.1 we provide our strategy to assess bootstrap validity. Our main results are given in Section 4.2. Application to our examples are provided in Section 4.3. Finally, the behavior under the alternative is analyzed and applied to our examples in Section 4.4.

4.1 Weak convergence in distribution and bootstrap validity

In this section, we summarize the approach developed by Cavaliere and Georgiev (2019), applied here to establish conditional bootstrap validity in the presence of non-stationary stochastic volatility.

Specifically, we consider a statistic τ_n which is a function of the data D_n, which in our general set-up may be represented by (M_n, U_n); that is, $\tau_n = \tau(M_n, U_n)$. Its bootstrap equivalent is $\tau_n^* = \tau(M_n^*, U_n^*)$, and we let $\tau = \tau(M, V)$, with (M, V) denoting the weak limit of (M_n, U_n), see Section 2.3. With $\sigma_n \in \mathscr{D}[0,1]$ the volatility process defined in (6) and $\sigma \in \mathscr{D}[0,1]$ its weak limit in Assumption 2, it follows by Cavaliere and Georgiev (2019, Corollary 3.2) that if the condition

$$(\tau_n|\sigma_n, \tau_n^*|D_n) \xrightarrow{w}_w (\tau|\sigma, \tau|\sigma) \tag{16}$$

is satisfied[3], with the random cdf of $\tau|\sigma$ being sample-path continuous, then

$$\sup_{x\in\mathbb{R}} |P(\tau_n \leq x|\sigma_n) - P(\tau_n^* \leq x|D_n)| \xrightarrow{p} 0.$$

[3]By (16), we mean that $(E(g(\tau_n)|\sigma_n), E(h(\tau_n)|D_n)) \xrightarrow{w} (E(g(\tau)|\sigma), E(h(\tau)|\sigma))$ jointly for all bounded continuous $g, h : \mathbb{R} \to \mathbb{R}$.

This means that the bootstrap consistently estimates the distribution of the original statistic conditional on the volatility process, which in turn implies that the bootstrap is conditionally valid, i.e.,

$$P(p_n^* \leq q|\sigma_n) \xrightarrow{p} q$$

for all $q \in (0, 1)$, where $p_n^* = P(\tau_n^* \leq \tau_n|D_n)$ is the bootstrap p-value. The key condition to verify is therefore the one given in (16), along with continuity of the limiting (random) cdf.

Because $\tau_n = \tau(M_n, U_n)$ with $(M_n, U_n) \in \mathscr{D}_2[0, 1]$, proving (16) involves proving conditional functional limit theorems. It is known, see Goggin (1994) and Crimaldi and Pratelli (2005), that joint weak convergence of e.g. $((M_n, U_n), \sigma_n)$ is not sufficient for conditional weak convergence. For example, Goggin (1994) shows that a sufficient condition is that (M_n, U_n) is independent of σ_n, or that a change of measure can be found (with weakly convergent Radon-Nikodym derivative) under which this independence holds. These conditions do not seem to be directly applicable to the present case. In contrast, our approach to proving (16) involves Skorokhod's representation theorem, and in particular the version of Kallenberg (1997), see Corollary A.1 in Appendix A. This allows us to obtain limit results "as if" the conditioning element $\sigma_n(\cdot)$ converges almost surely to $\sigma(\cdot)$. By restricting the dependence between (M_n, U_n) and σ_n, we may then fix a realization of the volatility process and prove an unconditional functional limit theorem for each realization (except on a set with measure zero).

4.2 Main results

Recall that the main assumption used to derive the limiting distribution of the original statistic and the limiting (conditional) distribution of the bootstrap statistic is that the errors form a mds with respect to the past information set. This condition, however, is not sufficient for conditional bootstrap validity, unless the volatility is deterministic or stationary. In the presence of non-stationary stochastic volatility, further conditions are required. A sufficient set of conditions is provided in the following assumption.

Assumption 3 *Define* $\mathcal{G}_{nt} := \sigma\left(\{z_i\}_{i=1}^t, \{\sigma_i\}_{i=1}^n\right)$ *and hence* $\mathcal{G}_{n0} := \sigma\left(\{\sigma_i\}_{i=1}^n\right)$, *and define* $\psi_{nt}^2 := E(z_t^2|\mathcal{G}_{n,t-1})$ *and* $v_{nt} := z_t/\psi_{nt}$. *Then:*

(a) for all n, $\{v_{nt}\}_{t=1}^n$ *is independent of* $\mathcal{G}_{n0}$, *and* $\{\psi_{nt}\}_{t=1}^n$ *is* $\mathcal{G}_{n0}$*-measurable.*

(b) $\{z_t, \mathcal{G}_{nt}\}_{1\leq t\leq n, n\geq 1}$ *is a martingale difference array (mda), satisfying for all* $\epsilon > 0$*:*

$$n^{-1}\sum_{t=1}^{n} E\left(z_t^2 \mathbb{I}_{\{|z_t|>\sqrt{n}\epsilon\}}|\mathcal{G}_{n0}\right) \xrightarrow{p} 0. \tag{17}$$

A few remarks are in order.

Remark 4.1 If $\{z_t\}_{t\geq 1}$ is independent of $\{\sigma_t\}_{t\geq 1}$, as in Example V.1, then Assumption 3 is trivially satisfied with $\psi_{nt} = 1$ and $v_{nt} = z_t$ (the Lindeberg condition (17) is implicitly assumed in Assumption 2, to guarantee $B_{z,n} \xrightarrow{w} B_z$). The dependence allowed by the assumption is needed to cover situations such as the GARCH process in Example V.2, where for all $t < n$,

$$z_t^2 = \frac{\sigma_{t+1}^2 - \omega_n - \beta_n\sigma_t^2}{\alpha_n\sigma_t^2},$$

which is known given $\{\sigma_t\}_{t=1}^n$, such that $\psi_{nt} = |z_t|$ and hence $v_{nt} = z_t/|z_t| = \text{sgn}(z_t)$, which because of symmetry of the Gaussian distribution will indeed be independent of $|z_t|$ and hence $\{\sigma_t\}_{t=1}^n$.

Remark 4.2 The mda assumption $E(z_t|\mathcal{G}_{n,t-1}) = 0$ rules out leverage effects, such as implied by non-zero correlation between z_t and the volatility shocks η_t in the stochastic volatility model of Example V.1. It may be possible to weaken this assumption for the results to follow, and allow for dependencies for finite n, as long as they vanish asymptotically (such that σ and B_z are independent). In the latter case it would be guaranteed that at least the unconditional validity property $P(p_n^* \leq q) \to q$ for $q \in (0,1)$ holds for the bootstrap, by Theorem 3.1 of Cavaliere and Georgiev (2019).

Remark 4.3 Part (a) of Assumption 3 implies that we may recover

$$M_n(\cdot) = n^{-1/2} \sum_{t=1}^{\lfloor n\cdot \rfloor} \sigma_t \psi_{nt} v_{nt}, \qquad U_n(\cdot) = n^{-1} \sum_{t=1}^{\lfloor n\cdot \rfloor} \sigma_t^2 \psi_{nt}^2 v_{nt}^2,$$

and $V_n(\cdot)$ from the two independent sequences $\{\sigma_t\}_{t=1}^n$ and $\{v_{nt}\}_{t=1}^n$. This independence facilitates the analysis of conditional distributions, as will be evident from the proof of Theorem 1. We conjecture that the asymptotic results which follow also hold when part (a) is replaced by (a') for all n, $\{\psi_{nt}\}_{t=1}^n$ is $\mathcal{G}_{n0}$-measurable. □

The main result in this section is given in Theorem 1 and Corollary 1 below.

Theorem 1 *Under Assumptions 1–3, we have as $n \to \infty$,*

$$((M_n, U_n)(\cdot)|\sigma_n, (M_n^*, U_n^*)(\cdot)|D_n) \xrightarrow{w}_w ((M, V)(\cdot)|\sigma, (M, V)(\cdot)|\sigma).$$

The key result of Theorem 1 is that the bootstrap processes M_n^* and U_n^*, conditionally on the data, replicate in the limit the distribution of the original processes M_n and U_n, conditionally on the volatility process σ_n. The implication of Theorem 1 on the behavior of the bootstrap p-values is provided in the following corollary, which applies to a statistic $\tau_n = \tau(M_n, U_n)$ (which, under the null, converges weakly to $\tau = \tau(M, V)$) and its bootstrap equivalent $\tau_n^* = \tau(M_n^*, U_n^*)$.

Corollary 1 *Under the conditions of Theorem 1, the bootstrap is valid conditionally on $\sigma_n(\cdot)$, i.e. with $p_n^* := P^*(\tau_n^* \leq \tau_n)$,*

$$p_n^*|\sigma_n \xrightarrow{w}_w U(0,1),$$

provided that the conditional distribution of $\tau = \tau(M, V)$ given σ is sample-path continuous and the function τ is itself continuous.

4.3 Examples revisited

In this section we check whether the conditions for bootstrap validity hold for the examples. We assume throughout that Assumptions 1 and 2, strengthened by 3, hold.

Example 1 (cont'd). As earlier, the bootstrap statistics are given by $S_n^* := \sqrt{n}\bar{\varepsilon}_n^*$ and $T_n^* := \sqrt{n}\bar{\varepsilon}_n^*/s_n^*$. Under the null hypothesis, the original statistics are given by $S_n^* =$

$\sqrt{n}\bar{\varepsilon}_n$ and $T_n^* = \sqrt{n}\bar{\varepsilon}_n/\hat{s}_n$. The original (bootstrap) statistics obtain as a continuous transformation of (M_n, U_n) (of (M_n^*, U_n^*)). Hence by Theorem 1 and a version of the CMT (see Cavaliere and Georgiev, 2019, Appendix A) it holds that

$$(S_n|\sigma_n, S_n^*|D_n) \xrightarrow{w}_w (M(1)|\sigma, M(1)|\sigma)\,,$$

with $M(1) = \int_0^1 \sigma(u)\mathsf{d}B_z(u)$. With $V(1) = \int_0^1 \sigma(s)^2\mathsf{d}s$, the cdf of $M(1)|\sigma$ is given by $\Phi(uV(1)^{-1/2})$, which is sample-path continuous with probability 1. Hence, by Corollary 1, the bootstrap is valid conditionally on the volatility path σ. For the studentized statistic it holds that

$$(T_n|\sigma_n, T_n^*|D_n) \xrightarrow{w}_w (Z|\sigma, Z|\sigma)\,, \tag{18}$$

where $Z \sim N(0,1)$, independent of σ; this implies that the bootstrap is conditionally valid. Notice that the convergence in (18) implies that

$$\sup_{x\in\mathbb{R}} |P(T_n \leq x|\sigma_n) - P^*(T_n^* \leq x)| \xrightarrow{p} 0,$$

which is the classic consistency result for the bootstrap. The same type of result does not hold for the bootstrap based on S_n, S_n^*; however, bootstrap conditional validity is guaranteed by Corollary 1.

Example 2 (cont'd). As for the previous example, since the CUSUM (bootstrap) statistics are continuous transformations of (M_n, U_n) (of (M_n^*, U_n^*)), from Theorem 1 and the CMT in Cavaliere and Georgiev (2019) we have that, for $\tau_S := \sup_{u\in[0,1]} |M(u) - uM(1)|$

$$(CS_n|\sigma_n, CS_n^*|D_n) \xrightarrow{w}_w (\tau_S|\sigma, \tau_S|\sigma)\,;$$

similarly, for $\tau_T := V(1)^{-1/2} \sup_{u\in[0,1]} |M(u) - uM(1)|$,

$$(CT_n|\sigma_n, CT_n^*|D_n) \xrightarrow{w}_w (\tau_T|\sigma, \tau_T|\sigma)\,.$$

Both conditional asymptotic distributions are continuous with probability 1. As discussed in Andrews (1997), this holds using the results in Lifshits (1982) because the limiting random distributions corresponds (up to an almost surely strictly positive term) to the supremum of a Gaussian process with covariance function which is nonsingular almost surely. Hence, by Corollary 1 the bootstrap is valid conditionally on the volatility path σ.

Example 3 (cont'd). Finally, in the unit root example we have that, under the stated assumption and if the null hypothesis holds, with $\tau_R := (\int_0^1 M^2(u)\mathsf{d}u)^{-1} \int_0^1 M(u)\mathsf{d}M(u)$,

$$(R_n|\sigma_n, R_n^*|D_n) \xrightarrow{w}_w (\tau_R|\sigma, \tau_R|\sigma)\,.$$

Similarly, for the t ratio test, with $\tau_W := (\int_0^1 M^2(u)\mathsf{d}u)^{-1/2} \int_0^1 M(u)\mathsf{d}M(u)$,

$$(W_n|\sigma_n, W_n^*|D_n) \xrightarrow{w}_w (\tau_W|\sigma, \tau_W|\sigma)\,.$$

As proved in Lemma A.1 in Appendix A, the limiting conditional cdfs have almost surely continuous sample paths, and hence by Corollary 1, the bootstrap is valid conditionally on σ. □

4.4 Power considerations

We now briefly discuss the behavior of the bootstrap tests under the alternative hypothesis when the stochastic volatility process induces randomness of the limiting distribution of the bootstrap statistic. As before, consider a left-sided test based on the statistic $\tau_n = \tau(M_n, U_n)$ and its bootstrap equivalent $\tau_n^* = \tau(M_n^*, U_n^*)$. Suppose that under the alternative the original statistic diverges, say to $-\infty$, while the bootstrap statistic satisfies

$$\tau_n^* \overset{w^*}{\to}_w \tau|\sigma \tag{19}$$

for some random element τ. Then, the following lemma holds.

Lemma 2 *Suppose that (19) holds and that $\tau_n \overset{p}{\to} -\infty$ as $n \to \infty$. Then, with $p_n^* := P^*(\tau_n^* \leq \tau_n)$, it holds that $p_n^* \overset{p}{\to} 0$.*

Lemma 2 shows that the fact that the limit distribution of the bootstrap statistic is random and depends on the volatility path does not affect the consistency of the test. Essentially, weak convergence in distribution of τ_n^* given the data implies that the bootstrap statistic is $O_p^*(1)$, in probability. If the original statistic diverges to $-\infty$, it then holds that the bootstrap test rejects with probability converging to 1.[4] We now apply this result to the three leading examples.

Example 1 (cont'd). Consider the location model example, where the econometrician is interested in testing the simple null hypothesis $\theta = \bar{\theta}$ when $\bar{\theta} > \theta_0$, θ_0 being the true parameter value. A wild bootstrap with the null imposed generates bootstrap data as $\varepsilon_t^* := \hat{\varepsilon}_t w_t^*$ with w_t^* i.i.d. $N(0,1)$ and $\hat{\varepsilon}_t := y_t - \bar{\theta} = \varepsilon_t + \delta$, $\delta := \theta_0 - \bar{\theta} < 0$. It follows that, conditionally on the data, $M_n^*(\cdot) := n^{-1/2} \sum_{t=1}^{\lfloor n\cdot \rfloor} \varepsilon_t^* \sim N(0, \hat{U}_n(\cdot))$, where

$$\hat{U}_n(\cdot) := n^{-1} \sum_{t=1}^{\lfloor n\cdot \rfloor} \hat{\varepsilon}_t^2 = U_n(\cdot) + \delta^2 + O_p(n^{-1/2}) \overset{w}{\to} V(\cdot) + \delta.$$

Hence, under Assumptions 1 and 2, we have, as $n \to \infty$, that $M_n^*(\cdot) \overset{w^*}{\to}_w \tilde{M}^*(\cdot)|\sigma$, where $\tilde{M}^*(\cdot) := \int_0^\cdot \tilde{\sigma}(u) \mathsf{d}B_z^*(u)$, with $\tilde{\sigma}(u) := (\sigma(u)^2 + \delta^2)^{1/2}$. This implies that $S_n^* \overset{w^*}{\to}_w \tilde{M}^*(1)|\sigma$. As $S_n \to -\infty$ as $n \to \infty$, the conditions of Lemma 2 are satisfied and $p_n^* \overset{p}{\to} 0$. Consistency of the test based on T_n follows by standard arguments as $T_n^* \overset{w^*}{\to}_p N(0,1)$.

Example 2 (cont'd). For the bootstrap CUSUM statistics, consider the alternative $\theta_t = \theta_1 + g(t/n)$, where $g : [0,1] \to \mathbb{R}$ is an arbitrary function satisfying $0 < \int_0^1 g^2(u) \mathsf{d}u < \infty$, see Ploberger and Krämer (1992). The wild bootstrap partial sum is $M_n^*(\cdot) := n^{-1/2} \sum_{t=1}^{\lfloor n\cdot \rfloor} \varepsilon_t^* \sim N(0, \hat{U}_n(\cdot))$, where in this case $\hat{U}_n(\cdot) := n^{-1} \sum_{t=1}^{\lfloor n\cdot \rfloor} \hat{\varepsilon}_t^2 = U_n(\cdot) + G_n(\cdot) + O_p(n^{-1/2})$, with

$$G_n(\cdot) := n^{-1} \sum_{t=1}^{\lfloor n\cdot \rfloor} \left(g(t/n) - n^{-1} \sum_{t=1}^{n} g(t/n) \right)^2 \to \int_0^\cdot \left(g(s) - \int_0^1 g(r) \mathsf{d}r \right)^2 \mathsf{d}s =: G(\cdot).$$

[4] Notice that a consistent right-sided test can be obtained by focusing on the bootstrap p-value $\tilde{p}_n^* := 1 - p_n^*$.

This implies that

$$M_n^*(\cdot) \xrightarrow{w^*}_w \tilde{M}^*(\cdot) := \int_0^{\cdot} \tilde{\sigma}(u)\mathsf{d}B_z^*(u)|\sigma,$$

with $\tilde{\sigma}(u) := (\sigma(u)^2 + g(u) - \int_0^1 g(u)\mathsf{d}u)^{1/2}$ and B_z^* a standard Brownian motion, stochastically independent of σ. The limiting distribution of the bootstrap statistic CS_n^* is then given by

$$CS_n^* \xrightarrow{w^*}_w \sup_{u\in[0,1]} |\tilde{M}^*(u) - u\tilde{M}^*(1)| \Big| \sigma.$$

In order to analyze the CT_n^* statistic, notice that its denominator satisfies, in probability,

$$U_n^*(1) - n^{-1}M_n^*(1)^2 = U_n^*(1) + O_p^*(n^{-1}) = \hat{U}_n(1) + O_p^*(n^{-1/2}) \xrightarrow{w} V(1) + G(1),$$

which implies that

$$CT_n^* \xrightarrow{w^*}_w \frac{\sup_{u\in[0,1]} |\tilde{M}^*(u) - u\tilde{M}^*(1)|}{\sqrt{V(1) + G(1)}} \Bigg| \sigma.$$

As both S_n and T_n diverge under the alternative considered, Lemma 2 applies and for both tests $p_n^* \xrightarrow{p} 0$.

Example 3 (cont'd). Consider the unit root example with wild bootstrap shocks generated with the null hypothesis, i.e. $\varepsilon_t^* := (\Delta y_t)w_t^*$. The bootstrap R_n^* statistic is given as in (13) with $M_n^*(\cdot) := n^{-1/2}\sum_{t=1}^{\lfloor n\cdot\rfloor}(\Delta y_t)w_t^*$ and $U_n^*(\cdot) := n^{-1/2}\sum_{t=1}^{\lfloor n\cdot\rfloor}(\Delta y_t)^2 w_t^{*2}$. Conditionally on the data M_n^* is a Gaussian process with independent increments and variance $\hat{U}_n(\cdot) := n^{-1}\sum_{t=1}^{\lfloor n\cdot\rfloor}(\Delta y_t)^2$. Under the alternative that $y_t = (1+\theta)y_{t-1} + \varepsilon_t$ with $\theta \in (-2, 0)$, Δy_t can be written as the linear process with exponentially decaying coefficients $\Delta y_t = \sum_{i=0}^{t-1}\psi_i\varepsilon_{t-i}$ with $\psi_0 = 0$ and $\psi_i = \theta(1+\theta)^{i-1}$, $i = 1, 2, \ldots$. Hence, by standard decompositions for squared stationary autoregressions it holds that (the proof is reported in the Appendix)

$$\hat{U}_n(\cdot) = \overline{\psi}^2 U_n(\cdot) + o_p(1), \qquad \overline{\psi}^2 := \sum_{i=0}^{\infty}\psi_i^2, \tag{20}$$

where the $o_p(1)$ term is uniform in $\cdot \in [0,1]$, which implies that $\hat{U}_n(\cdot) \xrightarrow{w} \overline{\psi}^2 V(\cdot)$. Hence, $M_n^*(\cdot) \xrightarrow{w^*}_w \overline{\psi}^2 M^*(\cdot)|\sigma$. Finally, using the fact that $U_n^*(\cdot) = \overline{\psi}^2\hat{U}_n(\cdot) + o_p^*(1)$, in probability, where $\hat{U}_n(\cdot) \xrightarrow{w} V(\cdot)$, it holds that

$$R_n^* \xrightarrow{w^*}_w \frac{\frac{1}{2}(\overline{\psi}^2 M^{*2}(1) - \overline{\psi}^2 V(1))}{\overline{\psi}^2 \int_0^1 M^{*2}(u)\mathsf{d}u} \Bigg| \sigma = \frac{\frac{1}{2}(M^{*2}(1) - V(1))}{\overline{\psi}^2 \int_0^1 M^{*2}(u)\mathsf{d}u} \Bigg| \sigma.$$

Hence, under the alternative the bootstrap replicates the null distribution of the original statistic conditional on the volatility process and consistency of the bootstrap test follows from Lemma 2. An identical result holds for the t-ratio test based on W_n. □

REMARK 4.4 While in this section we focused on asymptotic power against fixed alternatives, it is possible to extend our analysis to cover power against local alternatives. To illustrate, consider the test based on S_n for the hypothesis $\mathsf{H}_0 : \theta = \bar{\theta}$ in the location model $y_t = \theta + \varepsilon_t$ (Example 1). Under a sequence of local alternatives of the form $\mathsf{H}_n : \theta_n = \bar{\theta} + \delta_n$ with $\delta_n = n^{-1/2}c$, it is straightforward to show that $S_n = c + M_n(1)$, which converges weakly to $c+M(1)$ under Assumptions 1–2. For the bootstrap statistic, the results obtained above with $\delta_n \to 0$ imply $S_n^* \xrightarrow{w^*}_w M(1)^*|V(1) \stackrel{d}{=} V(1)^{1/2}Z^*|V(1)$, with Z^* being $N(0,1)$ (independent of $V(1)$); hence, the bootstrap distribution under the local alternative is the same as under the null. Suppose now that Assumption 3 holds; then, by Theorem 1,

$$(S_n|\sigma_n, S_n^*|D_n) \xrightarrow{w}_w (c + M(1)|\sigma, M(1)|\sigma) \, .$$

Hence the limiting cdf of $S_n|\sigma_n$ is given by $F_c(x) = \Phi((x-c)V(1)^{-1/2})$, which is continuous with probability 1, while the bootstrap cdf $F_n^*(x)$ converges weakly to $F^*(x) = \Phi(xV(1)^{-1/2})$. Then, by application of Theorem 3.3 in Cavaliere and Georgiev (2019) it holds that the power of the bootstrap test at the $100\alpha\%$ nominal level, conditionally on the volatility process, is given by

$$\begin{aligned} P(p_n^* \leq \alpha|\sigma_n) = P(F_n^*(S_n) \leq \alpha|\sigma_n) &\xrightarrow{w} F_c(F^{*-1}(\alpha)) \\ &= \Phi((V(1)^{1/2}\Phi^{-1}(\alpha) - c)V(1)^{-1/2}) \\ &= \Phi(\Phi^{-1}(\alpha) - cV(1)^{-1/2}). \end{aligned}$$

By construction, the local power function depends on c; we observe that it also depends on $V(1)$, and hence is random in the limit. In more general testing problems, the conditional local power function will depend on the entire volatility process. In the next section, we illustrate, by Monte Carlo simulations, the dependence on c as well as on (the limit of) $\sigma_n(\cdot)$. □

5 Numerical results

In this section we analyze finite sample size and power properties of bootstrap tests under non-stationary stochastic volatility using Monte Carlo simulations. To study the behavior of the tests from Examples 1–3 under the null hypothesis, we report the Monte Carlo (empirical) cdfs of bootstrap p-values, both unconditionally over all Monte Carlo replications and conditionally on specific simulated volatility paths. Following Cavaliere and Georgiev (2019), we report the results in the form of fan charts of the conditional cdfs, displayed together with the unconditional cdf and the theoretical cdf of the $U(0,1)$ distribution. Similarly, we display conditional power curves of the tests under local alternatives in fan charts.

In all experiments, we draw observations $\{\varepsilon_t\}_{t=1}^n$ from the GARCH(1,1) process from Example V.2, with

$$\omega_n = 1 - \alpha_n - \beta_n = n^{-1}\kappa, \qquad \alpha_n = (2n)^{-1/2}\sigma_\eta,$$

corresponding to a limit process $\sigma^2(u)$ with unit unconditional variance $\bar{\sigma}^2$, mean-reversion parameter κ, and volatility-of-volatility parameter σ_η. We set $\kappa = 5$ and

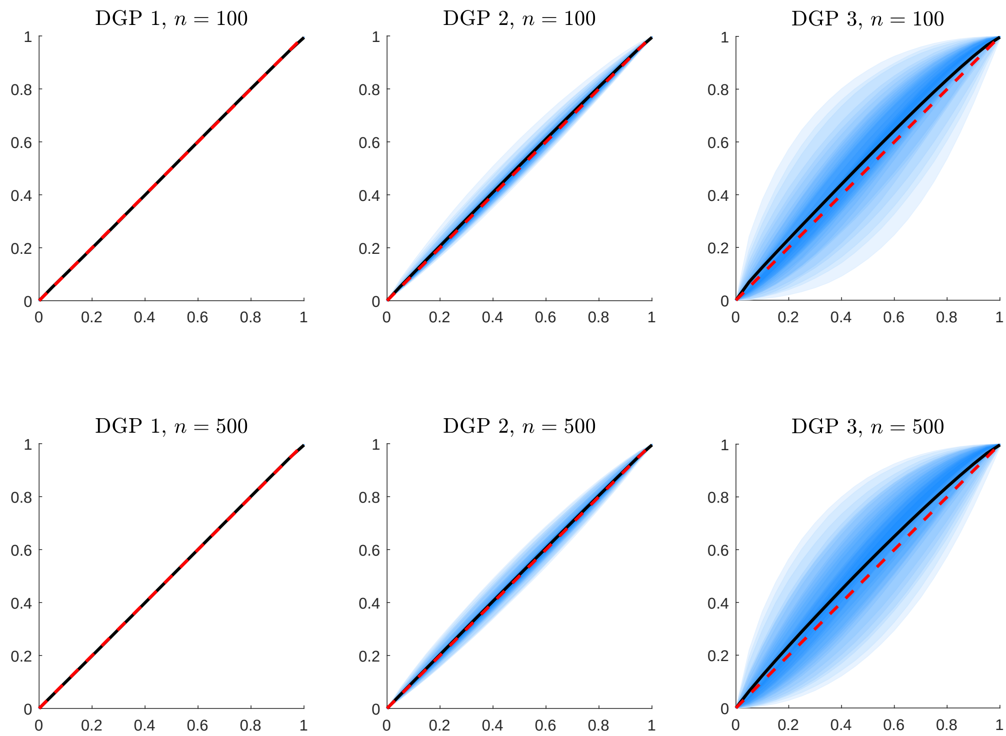

Figure 1: Monte Carlo conditional cdfs of bootstrap p-values of T_n

$\sigma_\eta = \sqrt{10}$, corresponding to a rather persistent volatility process with a reasonable amount of short-run variability of the volatility, which we know from earlier simulation studies to lead to substantial size distortions in tests using standard (constant-volatility) asymptotic critical values. We expect similar results from stochastic volatility processes (Example V.1) with the same type of persistence and volatility-of-volatility properties. We report results for two sample sizes, $n \in \{100, 500\}$. The standardized errors z_t are drawn from three different distributions, discussed below.

For each distribution and sample size, we first simulate 100 different realizations of the volatility path $\{\sigma_t\}_{t=1}^n$. For each of these paths, we draw $50{,}000$ replications from the conditional distribution of $\{\varepsilon_t = \sigma_t z_t\}_{t=1}^n$ given $\{\sigma_t\}_{t=1}^n$. As discussed in Remark 4.1, this is equivalent to drawing $v_{nt} = \text{sgn}(z_t)$ conditional on $\psi_{nt} = |z_t|$ for $t = 1, \ldots, n-1$, and drawing z_n from its unconditional distribution (independent of $\{\sigma_t\}_{t=1}^n$). For each choice of the distribution of $\{z_t\}_{t=1}^n$, we can check the conditions of Assumption 3 for conditional validity of the bootstrap.

The first data-generating process, labelled DGP 1, is defined by $z_t \sim N(0,1)$. In that case the conditional distribution of the signs v_{nt} is discrete uniform over $\{-1, 1\}$, independent of $|z_t|$. This in turn implies that the mda condition of Assumption 3 is satisfied (as well as the independence, measurability and Lindeberg conditions), such that the bootstrap is conditionally valid.

In DGP 2, z_t is drawn from the following mixed normal density

$$f(z) = \tfrac{1}{3}\phi(z; \mu_1, \sigma_1) + \tfrac{2}{3}\phi(z; \mu_2, \sigma_2), \tag{21}$$

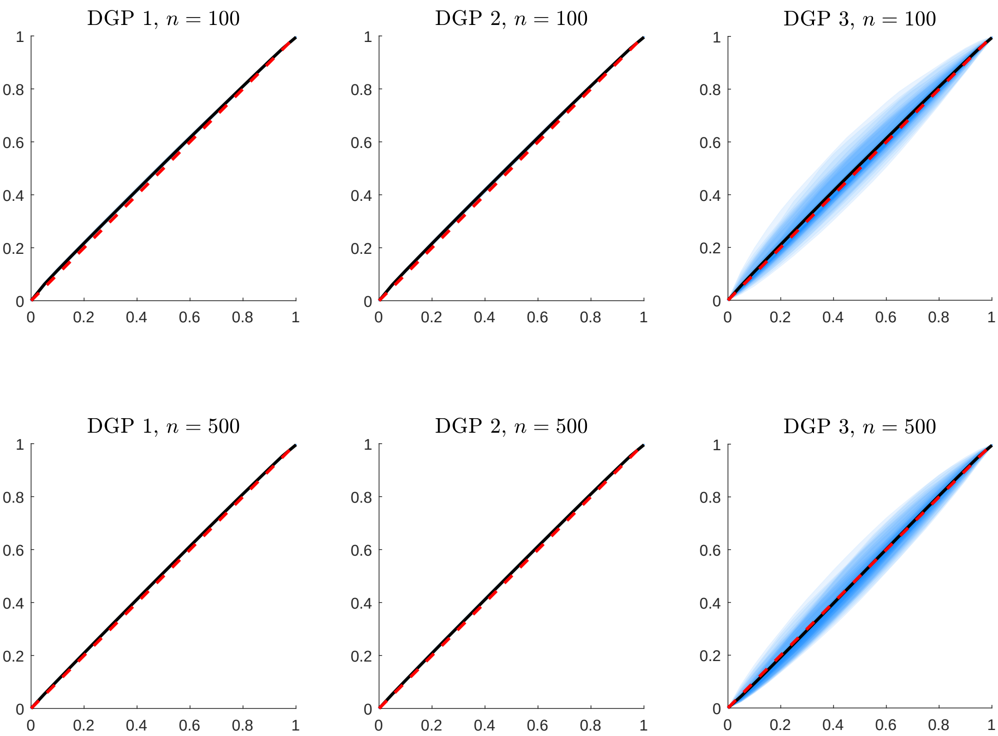

Figure 2: Monte Carlo conditional cdfs of bootstrap p-values of CT_n

where $\phi(z; \mu, \sigma)$ is the pdf of the $N(\mu, \sigma^2)$ distribution, and where $\mu_1 = -2a, \sigma_1 = a, \mu_2 = a, \sigma_2 = a\sqrt{2}$, with $a = \sqrt{3/11}$. This distribution was constructed by Meijer (2000) to be asymmetric but with skewness 0 (and with mean zero and unit variance). Because $v_{nt} = \text{sgn}(z_t)$ in this case has a conditional distribution depending on $\psi_{nt} = |z_t|$, with $P(v_{nt} = 1|\psi_{nt}) = f(\psi_{nt})/(f(\psi_{nt}) + f(-\psi_{nt})) \neq \frac{1}{2}$, it follows that Assumption 3 is violated, and conditional validity of the bootstrap is not guaranteed. On the other hand, the zero skewness implies that the limit result of Example V.2 still applies, with B_z independent of B_η and hence σ. As conjectured in Remark 4.2, we may expect unconditional bootstrap validity in this case.

In DGP 3, z_t is drawn from another version of (21), but now with $\mu_1 = -2b, \sigma_1 = b\sqrt{2}, \mu_2 = b, \sigma_2 = b$, with $b = \sqrt{3/10}$. This is a distribution with mean zero, unit variance and negative skewness, so that B_z and B_η in Example V.2 have a negative correlation, corresponding to leverage effects. This implies that the wild bootstrap is invalid in this case, both conditionally and unconditionally.

Figures 1–3 display the results for the behavior of bootstrap p-values (based on 199 bootstrap replications) under the null hypothesis, for the studentized tests based on T_n, CT_n and W_n, respectively. Unreported results for the other three test statistics S_n, CS_n and R_n are very similar to the results for the corresponding studentized tests.

From Figure 1, we observe that when the standardized errors z_t are standard normal (DGP 1, left panels), then the conditional distribution of the bootstrap p-values is very close to uniform, and appears to be independent of σ for both sample sizes considered. Thus the theoretical conditional validity of the bootstrap in this case is clearly reflected

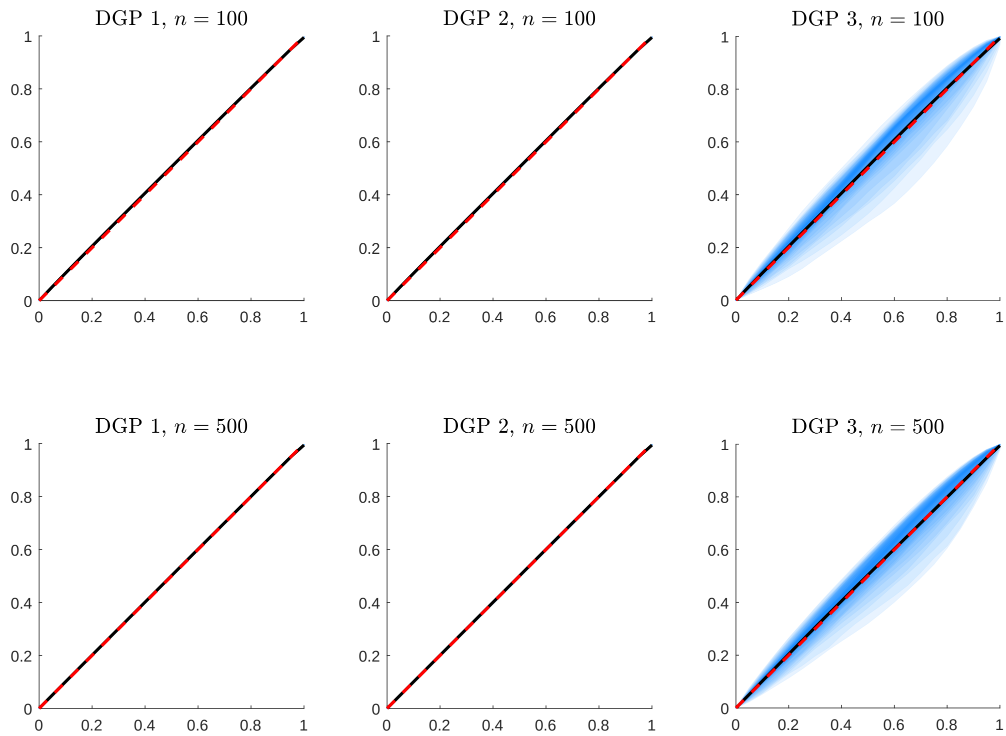

Figure 3: Monte Carlo conditional cdfs of bootstrap p-values of W_n

in finite-sample behavior. When the distribution of z_t is asymmetric with zero skewness (DGP 2, centre panels), then the bootstrap appears to be valid on average (indicated by the solid line almost coinciding with the $U(0,1)$ cdf, especially for $n = 500$), but the conditional cdfs of bootstrap p-values do depend on the volatility path and deviate from the uniform cdf, illustrating the conjectured violation of conditional bootstrap validity. Finally, for DGP 3 (right panels, skewed z_t), we observe more extreme dependence of bootstrap p-values on the volatility path. In this case the bootstrap does not appear to be valid on average either, as predicted by the dependence between B_z and σ implied by this DGP, which is not replicated by the wild bootstrap.

Figure 2 displays the results for the studentized CUSUM test based on CT_n. For this test, the finite-sample size distortion (indicated by the difference between the solid and dashed line) is more pronounced than for the location test, in particular for the smaller sample size ($n = 100$). Unreported additional simulations show that these size distortions are even stronger for the test based on CS_n. The results improve when the sample size increases, and it should be noted that the rejection frequencies at the 5% significance level are still fairly close to 0.05; the deviations are larger at the centre of the distribution. For this test, the dependence of the conditional cdf of p-values on σ is much weaker than for the location test. For DGP 2, we do not observe any deviation of conditional cdfs from their average; in case of DGP 3, there is a clear violation of conditional bootstrap validity, but the deviations are less pronounced than for T_n.

The results for the unit root test based on W_n are given in Figure 3. For DGP 1 and 2, the size distortions appear to be negligible for both sample sizes. Similarly to the

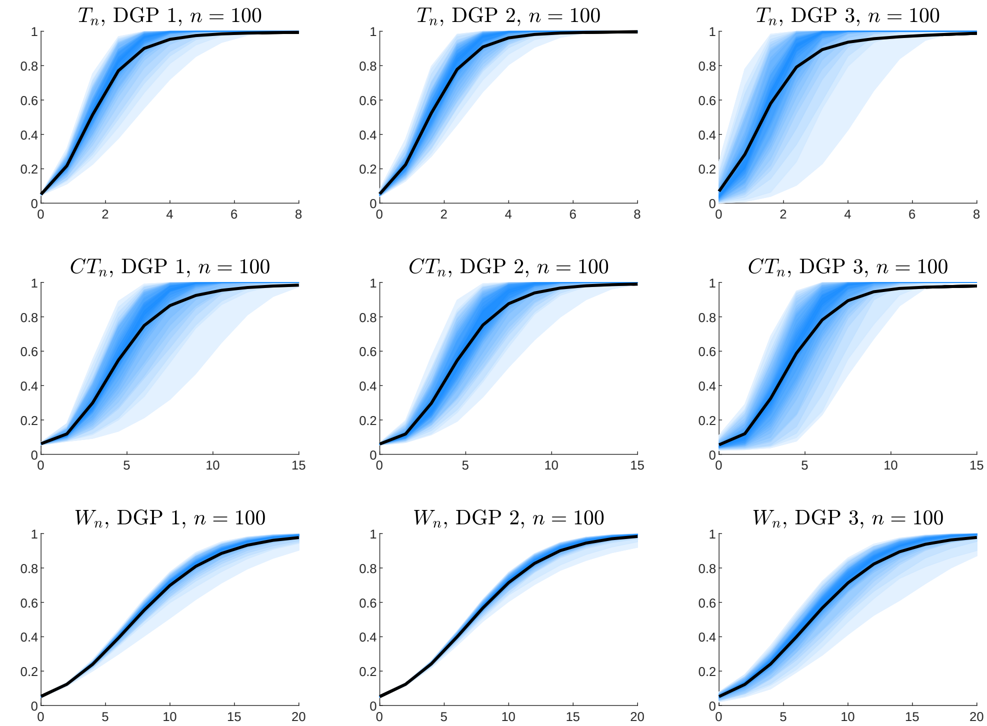

Figure 4: Monte Carlo conditional rejection frequencies of bootstrap tests

CUSUM test, the dependence of bootstrap p-values on the volatility path for DGP 1 and 2 appears to be very weak. On the other hand, for DGP 3 we find this dependence to be clearly present, illustrating again a violation of conditional bootstrap validity. On average, the bootstrap appears to be valid even for DGP 3, although theoretically we would not expect this to be the case because of the dependence between B_z and σ. Unreported simulations have shown that in case of stronger leverage effects (i.e., a stronger correlation between B_z and B_η), the unconditional cdf of bootstrap p-values does differ from the $U(0,1)$ cdf, as predicted by the theoretical results.

We conclude this section with some local power results of the bootstrap tests, again conditional on the same realizations of the volatility path as considered for the size of the tests. As in Remark 4.4, for the location tests we evaluate the rejection frequency of the test for $\mathsf{H}_0 : \theta = 0$ against local alternatives $\theta_n = -n^{-1/2}c$, with $c \in [0, 8]$. For the CUSUM tests, the local alternative is a break in the mean of the series, at $t = n/2$, from $\theta_{n,t} = 0$ to $\theta_{n,t} = n^{-1/2}c$, with $c \in [0, 15]$. For the unit root tests, we consider local alternatives $\theta_n = -n^{-1}c$, with $c \in [0, 20]$. We provide results for the tests based on the studentized statistics T_n, CT_n and W_n, and for the sample size $n = 100$.

Figure 4 displays the rejection frequencies of the bootstrap tests, based on 10,000 replications of the test for each volatility path, plotted against c. We observe that the conditional rejection probabilities under the alternative hypothesis depend on both the non-centrality parameter c and the volatility process σ, for each test and DGP 1–3. While the dependence on the volatility is as expected for DGP 2 and 3, we note that for DGP 1, where the rejections probabilities under the null hypothesis are conditionally independent of the volatility process, the power of the tests clearly depends on the volatility (as discussed in Remark 4.4).

6 Conclusions

In this paper we have analyzed the properties of wild bootstrap inference in time series models for the conditional mean under non-stationary stochastic volatility. Our results can be generalized in several directions. First, our applications deal with univariate time series models and it is naturally of interest to apply our results to multivariate (time series) models, where volatilities and correlations are time-varying, stochastic and non-stationary. In particular, in Boswijk *et al.* (2016) the bootstrap was considered for multivariate cointegrated vector autoregressions in the presence of stationary volatility, in combination with possible deterministic changes in the volatility; we conjecture that our results obtained here also apply to the case of non-stationary multivariate stochastic volatility. Second, it would be important to understand how to bootstrap conditional mean time series models in the presence of leverage. Although, as we have shown, the wild bootstrap is not valid in this context, our theory may be useful for assessing validity of other bootstrap methods when the volatility displays leverage effect.

A Mathematical Appendix

A.1 Auxiliary results

Throughout, we make use of the following version of Skorokhod's representation theorem, see A.1.

Theorem A.1 *[Kallenberg, 1997, Corollary 5.12] Let f and $\{f_n\}_{n\geq 1}$ be measurable functions from a Borel space $\mathcal{S}$ to a Polish space $\mathcal{T}$, and let ξ and $\{\xi_n\}_{n\geq 1}$ be random elements in $\mathcal{S}$ with $f_n(\xi_n) \overset{w}{\to} f(\xi)$. Then there exist some random elements $\tilde{\xi} \overset{d}{=} \xi$ and $\tilde{\xi}_n \overset{d}{=} \xi_n$ defined on a common probability space with $f_n(\tilde{\xi}_n) \overset{a.s.}{\to} f(\xi)$.*

The next lemma contains a result about the asymptotic continuity of the distribution function of Dickey-Fuller type-statistics under non-stationary stochastic volatility.

Lemma A.1 *Under Assumptions 1 and 2, let*

$$\tau_1 := \frac{\int_0^1 M(u)\mathsf{d}M(u)}{\int_0^1 M^2(u)\mathsf{d}u}, \qquad \tau_2 := \frac{\int_0^1 M(u)\mathsf{d}M(u)}{\sqrt{V(1)\int_0^1 M^2(u)\mathsf{d}u}}.$$

Then the random cdfs $F_1(\cdot) := P(\tau_1 \leq \cdot \,|\, \sigma)$ and $F_2(\cdot) := P(\tau_2 \leq \cdot \,|\, \sigma)$ are sample-path continuous a.s.

Proof of Lemma A.1. We reduce the proof to the following well-known result (Rao and Swift, 2006, pp. 472–473). Let $\{X(u)\}_{u\in[0,1]}$ be a Gaussian process with mean zero and a continuous covariance kernel, let $q : [0,1] \to \mathbb{R}$ be a square-integrable function and let $\alpha \in \mathbb{R}$ be arbitrary. Then the distribution of $\int_0^1 (X(u) + \alpha q(u))^2 \mathsf{d}u$ is that of an infinite series of independent non-central χ^2 random variables and, as a result, it has a continuous cdf.

The random cdfs F_1 and F_2 are determined, up to a modification, by the *distribution* of (B_z, σ), such that the structure of the probability space on which (B_z, σ) is defined is irrelevant for the claim of interest. We therefore assume, without loss of generality, that the independent processes B_z and σ are defined on a product probability space. Let $(\Omega_\sigma, \mathcal{F}_\sigma, P_\sigma)$ be the factor-space on which σ is defined. Fix $A \in \mathcal{F}_\sigma$ with $P_\sigma(A) = 1$ such that $V(\omega, \cdot) := \int_0^\cdot \sigma^2(\omega, u)\mathsf{d}u$ is well-defined, continuous and $0 < V(\omega, 1) < \infty$. Let $\Gamma := \{\sigma(\omega) : \omega \in A\}$ be the set of trajectories for σ when $\omega \in A$. For every $\gamma \in \Gamma$, the process $M_\gamma(\cdot) := \int_0^\cdot \gamma(u)\mathsf{d}B_z(u)$ is a.s. well-defined and $\int_0^1 M_\gamma^2(u)\mathsf{d}u > 0$ a.s. The result in the lemma will follow if the deterministic cdfs $P(\tau_{\gamma 1} \leq \cdot)$ and $P(\tau_{\gamma 2} \leq \cdot)$ are continuous for every $\gamma \in \Gamma$:

$$P(\tau_{\gamma 1} = x) = 0, \quad P(\tau_{\gamma 2} = x) = 0, \quad \forall x \in \mathbb{R}, \tag{A.1}$$

where

$$\tau_{\gamma 1} := \frac{\int_0^1 M_\gamma(u)\mathsf{d}M_\gamma(u)}{\int_0^1 M_\gamma^2(u)\mathsf{d}u}, \qquad \tau_{\gamma 2} := \frac{\int_0^1 M_\gamma(u)\mathsf{d}M_\gamma(u)}{\sqrt{V(1)\int_0^1 M_\gamma^2(u)\mathsf{d}u}}.$$

In fact, (A.1) implies that F_1 and F_2 have sample-path continuous modifications, and moreover, by continuity, F_1 and F_2 are indistinguishable from these modifications.

We turn to the proof of (A.1). For an arbitrary fixed $\gamma \in \Gamma$, define the time-changed 'bridge' process X_γ by

$$X_\gamma(u) := M_\gamma(u) - \frac{V_\gamma(u)}{V_\gamma(1)} M_\gamma(1), \quad u \in [0, 1].$$

Then X_γ and $M_\gamma(1)$ are independent, for they are jointly Gaussian with covariance function

$$\mathrm{Cov}(X_\gamma(u), M_\gamma(1)) = V_\gamma(u) - \frac{V_\gamma(u)}{V_\gamma(1)} V_\gamma(1) = 0, \quad u \in [0, 1].$$

In terms of X_γ and $M_\gamma(1)$, we find

$$\tau_{\gamma 1} = \frac{1}{2}\frac{M_\gamma(1)^2 - V_\gamma(1)}{\int_0^1 M_\gamma^2(u)\mathsf{d}u} = \frac{1}{2}\frac{M_\gamma(1)^2 - V_\gamma(1)}{\int_0^1 (X_\gamma(u) + M_\gamma(1)q_\gamma(u))^2\mathsf{d}u}$$

and

$$\tau_{\gamma 2} = \frac{1}{2}\frac{M_\gamma(1)^2 - V_\gamma(1)}{\sqrt{V_\gamma(1)\int_0^1 (X_\gamma(u) + M_\gamma(1)q_\gamma(u))^2\mathsf{d}u}},$$

for $q_\gamma(u) := V_\gamma(u)/V_\gamma(1)$. The equality

$$P(\tau_{\gamma i} = x) = E\left[P(\tau_{\gamma i} = x | M_\gamma(1))\right] = 0$$

will hold for $i = 1, 2$ and any $x \in \mathbb{R}$ iff

$$P(\tau_{\gamma i} = x | M_\gamma(1)) = 0 \text{ a.s.}$$

for $i = 1, 2$ and any $x \in \mathbb{R}$. In its turn, using the independence of $X_\gamma(u)$ and $M_\gamma(1)$, the latter will hold if

$$P\left(\frac{1}{2}\frac{\alpha^2 - V_\gamma(1)}{\int_0^1 (X_\gamma(1) + \alpha q_\gamma(u))^2 \mathsf{d}u} = x\right) = 0,$$

$$P\left(\frac{1}{2}\frac{\alpha^2 - V_\gamma(1)}{\sqrt{V_\gamma(1)\int_0^1 (X_\gamma(u) + \alpha q_\gamma(u))^2 \mathsf{d}u}} = x\right) = 0$$

holds for all $x \in \mathbb{R}$ and $\alpha \neq \pm\sqrt{V_\gamma(1)}$ (because $P(M_\gamma^2(1) = V_\gamma(1)) = 0$), which in its turn will hold if

$$P\left(\int_0^1 (X_\gamma(u) + \alpha q_\gamma(u))^2 \mathsf{d}u = x\right) = 0$$

for any $\alpha, x \in \mathbb{R}$. Since X_γ is a zero-mean Gaussian process with a continuous covariance and q_γ is square integrable, the equality in the previous display indeed holds, by Rao and Swift (2006, pp. 472–473). □

A.2 Proofs

Proof of Lemma 1. We follow the approach of the proof of Lemma 1 and other intermediate results in Cavaliere and Taylor (2009). First, defining $e_t = z_t^2 - 1$,

$$\sup_{u\in[0,1]} |U_n(u) - V_n(u)| = \sup_{u\in[0,1]} \left| n^{-1} \sum_{t=1}^{\lfloor nu \rfloor} \sigma_t^2 e_t \right| \xrightarrow{p} 0$$

by Theorem A.1 of Cavaliere and Taylor (2009), since $\{e_t, \mathcal{F}_t\}_{t\geq 1}$ is an mds by Assumption 1 and $\sigma^2_{\lfloor n\cdot\rfloor+1} = \sigma_n^2(\cdot) \xrightarrow{w} \sigma^2(\cdot)$ by Assumption 2 and the CMT; this proves (8), because convergence in the sup norm implies convergence in the Skorokhod metric, i.e., in $\mathscr{D}[0,1]$. Next, we apply Theorem 2.1 of Hansen (1992) to

$$M_n(\cdot) = \int_0^{\cdot} \sigma_n(u) \mathsf{d}B_{z,n}(u),$$

noting that Assumption 1 implies $\sup_{n\geq 1} n^{-1}\sum_{t=1}^n E(z_t^2) = 1$, so that using Assumption 2, we have

$$(\sigma_n(\cdot), B_{z,n}(\cdot), M_n(\cdot)) \xrightarrow{w} (\sigma(\cdot), B_z(\cdot), M(\cdot)).$$

The CMT together with (8) then implies (7), because

$$\int_0^u \sigma_n^2(s)\mathsf{d}s = \frac{1}{n}\sum_{t=1}^{\lfloor nu \rfloor} \sigma_t^2 + \sigma^2_{\lfloor nu\rfloor+1}(u - \lfloor nu \rfloor n^{-1}), \qquad u \in [0,1],$$

so that $U_n(\cdot) = V_n(\cdot) + o_p(1) = \int_0^{\cdot} \sigma_n^2(s)\mathsf{d}s + o_p(1)$, i.e., $U_n(\cdot)$ is a continuous functional of $\sigma_n(\cdot)$ plus an asymptotically negligible term. □

Proof of Theorem 1. The idea of the proof is to construct on a special probability space random elements distributed like $(\sigma_n, M_n, U_n, M_n^*, U_n^*)$ and such that on this probability space the convergence asserted in Theorem 1 holds weakly a.s.; on a general probability space it will then hold $\xrightarrow{w}_w$. Throughout, we use repeatedly the fact that for independent random elements ξ and η and for a measurable real ϕ such that $E(|\phi(\xi,\eta)|) < \infty$, it holds that $E(\phi(\xi,\eta)|\eta) = E(\phi(\xi,s))|_{s=\eta}$, with $E(\phi(\xi,s))$ defining a function of a non-random s; see Dudley (2004, p. 341).

By Assumption 3, ψ_{nt} are $\mathcal{G}_{n0}$-measurable and hence are measurable functions of $\sigma_n(\cdot)$ that we denote, with a slight abuse of notation, by $\psi_{nt}(\sigma_n(\cdot))$. Let

$$e_{nm}(\gamma) := E\left(v_{nt}^2\psi_{nt}^2(\gamma)\mathbb{I}_{\{|v_{nt}\psi_{nt}(\gamma)|>\sqrt{n}/m\}}\right),$$

for $m \in \mathbb{N}$ and a generic non-random γ; then $e_{nm}(\sigma_n(\cdot))$ is a version of the conditional expectation $E\left(z_t^2\mathbb{I}_{\{|z_t|>\sqrt{n}/m\}}|\sigma_n(\cdot)\right)$ because $\{v_{nt}\}_{t=1}^n$ and $\sigma_n(\cdot)$ are independent. Define $B_{v,n}(\cdot) := n^{-1/2}\sum_{t=1}^{\lfloor n\cdot\rfloor} v_{nt}$. We apply Theorem A.1 with $\xi_n = (\sigma_n, B_{v,n})$, $\xi = (\sigma, B_z)$,

$$f_n(\xi_n) = (\sigma_n, Q_{\psi,n}, Q_{z,n}, \mathcal{L}_n, L_n) \text{ and } f(\xi) = (\sigma, Q, Q, 0^\infty, 0^\infty),$$

where $Q_{\psi,n}(\cdot) = n^{-1}\sum_{t=1}^{\lfloor n\cdot\rfloor}\psi_{nt}^2$, $Q_{z,n}(\cdot) = n^{-1}\sum_{t=1}^{\lfloor n\cdot\rfloor} z_t^2$, $\mathcal{L}_n = \left\{n^{-1}\sum_{t=1}^n e_{nm}(\sigma_n)\right\}_{m\in\mathbb{N}} \in \mathbb{R}^\infty$, $L_n = \left\{n^{-1}\sum_{t=1}^n z_t^2\mathbb{I}_{\{|z_t|>\sqrt{n}/m\}}\right\}_{m\in\mathbb{N}} \in \mathbb{R}^\infty$, $Q(u) = u, u \in [0,1]$, and 0^∞ is the zero sequence in $\mathbb{R}^\infty$, the Frechet space. The domain of f_n and f is the Borel space $\mathscr{D}_2[0,1]$ with the Skorokhod metric and the induced Borel σ-algebra, and the codomain is the Polish space $\mathscr{D}_3[0,1]\times\mathbb{R}^\infty\times\mathbb{R}^\infty$ with the product of the Skorokhod and the Frechet metric. The assumptions imply $(Q_{\psi,n}, Q_{z,n}) \xrightarrow{p} (Q,Q)$, because $(Q_{\psi,n}-Q, Q_{z,n}-Q)$ is the partial sum process of $n^{-1}(\psi_{nt}^2-1, z_t^2-1)$, which is an mda with respect to $\mathcal{F}_t$ since $E(\psi_{nt}^2|\mathcal{F}_{t-1}) = E(z_t^2|\mathcal{F}_{t-1}) = 1$ by the tower property; this partial sum converges to the zero function in probability by the corollary to Theorem 3.3 of Hansen (1992). Noting that $L_n \xrightarrow{w} 0^\infty$ follows from the corresponding result for $\mathcal{L}_n = E(L_n|\mathcal{G}_{n0})$, applying Markov's inequality, the assumptions therefore imply $f_n(\xi_n) \xrightarrow{w} f(\xi)$.

Theorem A.1 then implies the existence of $\tilde{\xi}_n = (\tilde{\sigma}_n, \tilde{B}_{v,n}) \stackrel{d}{=} (\sigma_n, B_{v,n})$ and $\tilde{\xi} = (\tilde{\sigma}, \tilde{B}_z) \stackrel{d}{=} (\sigma, B_z)$, defined on a single probability space and such that

$$\left(\tilde{\sigma}_n, \tilde{Q}_{\psi,n}, \tilde{Q}_{z,n}, \tilde{\mathcal{L}}_n, \tilde{L}_n\right) := f_n(\tilde{\xi_n}) \overset{a.s.}{\to} f(\tilde{\xi}) = (\tilde{\sigma}, Q, Q, 0^\infty, 0^\infty). \tag{A.2}$$

Finally, we complete the set up by introducing a product extension of the previous probability space where a sequence $\{\tilde{w}_t^*\} \stackrel{d}{=} \{w_t^*\}$ and a standard Brownian motion $\tilde{B}_z^*$ are defined and are independent of $\{(\tilde{\sigma}_n, \tilde{B}_{v,n})\}_{n\geq1}$ and $(\tilde{\sigma}, \tilde{B}_z)$.

As $\tilde{B}_{v,n}$ and $\tilde{\sigma}_n$ are independent (because $B_{v,n}$ and σ_n are), it holds for any integrable random variable $h(\tilde{\sigma}_n, \tilde{B}_{v,n})$ that $E(h(\tilde{\sigma}_n, \tilde{B}_{v,n})|\tilde{\sigma}_n) = E(h(\gamma, \tilde{B}_{v,n}))|_{\gamma=\tilde{\sigma}_n}$. A similar equality holds for the independent $\tilde{B}_z$ and $\tilde{\sigma}$. Therefore, to prove any convergence of the form

$$E\left(h_n(\tilde{\sigma}_n, \tilde{B}_{v,n})|\tilde{\sigma}_n\right) \overset{a.s.}{\to} E\left(h(\tilde{\sigma}, \tilde{B}_z)|\sigma\right), \tag{A.3}$$

it is sufficient to prove that $E(h_n(\gamma_n, \tilde{B}_{v,n})) \to E(h(\gamma, \tilde{B}_z))$ for all deterministic sequences $\{\gamma_n\}_{n\geq1}$ in some set $\Gamma \subset \mathscr{D}_\infty[0,1]$ such that $P(\{\tilde{\sigma}_n\}_{n\geq1} \in \Gamma) = 1$. We now choose and fix Γ. Consider all the outcomes $\tilde{\omega}$ such that convergence (A.2) holds; the set of such outcomes $\tilde{\omega}$ has probability 1. Define $\Gamma \subset \mathscr{D}_\infty[0,1]$ as the set of sequences $\{\tilde{\sigma}_n(\cdot,\tilde{\omega})\}_{n\geq1}$ corresponding to such $\tilde{\omega}$, then $P(\{\tilde{\sigma}_n\}_{n\geq1} \in \Gamma) = 1$ as required.

As noted in Remark 4.3, we may recover (M_n, U_n) (and hence the original data D_n) from $(\sigma_n, B_{v,n})$ as some measurable transformation, say $m_n(\sigma_n, B_{v,n})$. Define accordingly $(\tilde{M}_n, \tilde{U}_n) := m_n(\tilde{\sigma}_n, \tilde{B}_{v,n})$ (and analogously $\tilde{D}_n$). With $\tilde{z}_{nt} := \tilde{\psi}_{nt}\tilde{v}_{nt}$, where $\tilde{\psi}_{nt} = \psi_{nt}(\tilde{\sigma}_n)$ and

$$\tilde{v}_{nt} := n^{1/2}\left(\tilde{B}_{v,n}(t/n) - \tilde{B}_{v,n}((t-1)/n)\right),$$

define also the process $\tilde{B}_{z,n}(\cdot) := n^{-1/2}\sum_{t=1}^{\lfloor n\cdot\rfloor}\tilde{z}_{nt} =: m_{z,n}(\tilde{\sigma}_n, \tilde{B}_{v,n})$, such that

$$(\tilde{\sigma}_n, \tilde{B}_{z,n}, \tilde{M}_n, \tilde{U}_n) \overset{d}{=} (\sigma_n, B_{z,n}, M_n, U_n).$$

We proceed to the convergence of $(\tilde{M}_n, \tilde{U}_n)$ conditional on $\tilde{\sigma}_n$ and prove that

$$E\left(g(\tilde{B}_{z,n}, \tilde{M}_n, \tilde{U}_n)\middle|\,\tilde{\sigma}_n\right) \overset{a.s.}{\to} E\left(g(\tilde{B}_z, \tilde{M}, \tilde{V})\middle|\,\tilde{\sigma}\right) \tag{A.4}$$

for continuous bounded real g of matching domain. This convergence is of the form (A.3) with $h_n = g \circ (m_{z,n}, m_n)$. By the discussion in the previous paragraph, (A.4) will follow from the standard weak convergence of $h_n(\gamma_n, \tilde{B}_{v,n})$, for all $\{\gamma_n\}_{n\geq 1} \in \Gamma$, that we establish next. In so doing, for any random element $Z = \phi(\tilde{\sigma}_n, \tilde{B}_{v,n})$ we write $Z(\gamma_n)$ for $\phi(\gamma_n, \tilde{B}_{v,n})$.

For $\{\tilde{\sigma}_n\}_{n\in\mathbb{N}}$ replaced by a fixed $\{\gamma_n\}_{n\geq 1} \in \Gamma$, $\tilde{z}_{nt}(\gamma_n) = \psi_{nt}(\gamma_n)\tilde{v}_{nt}$ is an mda satisfying the conditions of Brown (1971)'s functional central limit theorem. First, $E(\psi_{nt}(\gamma_n)\tilde{v}_{nt}|\{\tilde{v}_{ni}\}_{i=1}^{t-1}) = \psi_{nt}(\gamma_n)E\left(\tilde{v}_{nt}|\{\tilde{v}_{ni}\}_{i=1}^{t-1}\right) = 0$ because the mda property of $\tilde{v}_{nt}$ is inherited from the original probability space as $\{\tilde{v}_{ni}\}_{i=1}^{n} \overset{d}{=} \{v_{ni}\}_{i=1}^{n}$. Second, $n^{-1}\sum_{t=1}^{\lfloor n\cdot\rfloor} E(\psi_{nt}^2(\gamma_n)\tilde{v}_{nt}^2|\,\{\tilde{v}_{ni}\}_{i=1}^{t-1}) = n^{-1}\sum_{t=1}^{\lfloor n\cdot\rfloor}\psi_{nt}^2(\gamma_n) = \tilde{Q}_{\psi,n}(\cdot,\gamma_n) \to Q(\cdot)$, where the first equality is again inherited from the original probability space, and the convergence by the definition of Γ. Third, as $\tilde{\mathcal{L}}_n(\gamma_n) \to 0^\infty$ again by the choice of Γ, it holds that $n^{-1}\sum_{t=1}^{n} e_{nm}(\gamma_n) \to 0$ for all $m \in \mathbb{N}$, which is equivalent to

$$n^{-1}\sum_{t=1}^{n} E\left(\tilde{z}_{nt}^2(\gamma_n)\mathbb{I}_{\{|\tilde{z}_{nt}(\gamma_n)|>\sqrt{n}/m\}}\right) \to 0, \quad m \in \mathbb{N},$$

by the definition of e_{nm} and implies the Lindeberg condition in its usual form

$$n^{-1}\sum_{t=1}^{n} E\left(\tilde{z}_{nt}^2(\gamma_n)\mathbb{I}_{\{|\tilde{z}_{nt}(\gamma_n)|>\sqrt{n}\epsilon\}}\right) \to 0$$

for all $\epsilon > 0$. Therefore,

$$\tilde{B}_{z,n}(\cdot,\gamma_n) \overset{w}{\to} \tilde{B}_z(\cdot),$$

in the sense that $E(g(\tilde{B}_{z,n}(\cdot,\gamma_n))) \to E(g(\tilde{B}_z))$ for continuous bounded real g with matching domain. For the same fixed γ_n, this in turn implies that

$$\tilde{M}_n(\cdot,\gamma_n) = \int_0^{\cdot}\gamma_n(u)\mathsf{d}\tilde{B}_{z,n}(u,\gamma_n) \overset{w}{\to} \int_0^{\cdot}\gamma(u)\mathsf{d}\tilde{B}_z(u),$$

where $\gamma = \lim\gamma_n$ exists by the choice of γ_n. More precisely, by Theorem 2.1 of Hansen (1995), as $\sup_{n\geq 1}\sum_{t=1}^{n} E(\tilde{z}_{nt}^2(\gamma_n)) = \sup_{n\geq 1}\tilde{Q}_{\psi,n}(u,\gamma_n) < \infty$, the previous convergence holds jointly with that of $\tilde{B}_{z,n}$, such that $E(g(\tilde{B}_{z,n}(\cdot,\gamma_n),\tilde{M}_n(\cdot,\gamma_n)))\to E(g(\tilde{B}_z,\int_0^\cdot \gamma \mathsf{d}\tilde{B}_z))$ for continuous bounded real g. Furthermore, using

$$\tilde{U}_n(\cdot) = n^{-1}\sum_{t=1}^{\lfloor n\cdot\rfloor}\tilde{\sigma}_t^2\tilde{\psi}_{nt}^2 + n^{-1}\sum_{t=1}^{\lfloor n\cdot\rfloor}\tilde{\sigma}_t^2\left(\tilde{z}_{nt}^2-\tilde{\psi}_{nt}^2\right)$$
$$= \int_0^\cdot \tilde{\sigma}_n^2(u)\mathsf{d}\tilde{Q}_{\psi,n}(u) + n^{-1}\sum_{t=1}^{\lfloor n\cdot\rfloor}\tilde{\sigma}_t^2\tilde{e}_{nt},$$

it follows that $\tilde{U}_n(\cdot,\gamma_n)\xrightarrow{p}\int_0^\cdot \gamma^2(u)\mathsf{d}u$ by Theorem A.1 of Cavaliere and Taylor (2009), since $\tilde{z}_{nt}^2(\gamma_n)-\tilde{\psi}_{nt}^2(\gamma_n)$ is an mda. As convergence in probability to a constant is joint with any weak convergence of random elements defined on the same probability space, it follows that

$$E\left[g(\tilde{B}_{z,n}(\cdot,\gamma_n),\tilde{M}_n(\cdot,\gamma_n),\tilde{U}_n(\cdot,\gamma_n))\right]\to E\left[g\left(\tilde{B}_z,\int_0^\cdot \gamma\mathsf{d}\tilde{B}_z,\int_0^\cdot \gamma^2\right)\right]$$

for continuous bounded real g. Since $P(\Gamma)=1$, $(\tilde{B}_z,\int_0^\cdot\gamma\mathsf{d}\tilde{B}_z,\int_0^\cdot\gamma^2)|_{\gamma=\tilde{\sigma}}=(\tilde{B}_z,\tilde{M},\tilde{V})$ and $\tilde{B}_z$ is independent of $\tilde{\sigma}$, we can conclude that (A.4) holds.

We turn to the bootstrap processes. Define

$$\tilde{B}_{z,n}^*(\cdot):=n^{-1/2}\sum_{t=1}^{\lfloor n\cdot\rfloor}\tilde{z}_{nt}\tilde{w}_t^*,\quad \tilde{M}_n^*(\cdot):=n^{-1/2}\sum_{t=1}^{\lfloor n\cdot\rfloor}\tilde{\sigma}_t\tilde{z}_{nt}\tilde{w}_t^*,\quad \tilde{U}_n^*(\cdot):=n^{-1}\sum_{t=1}^{\lfloor n\cdot\rfloor}\tilde{\sigma}_t^2\tilde{z}_{nt}^2\tilde{w}_t^{*2}.$$

Here we show that

$$E\left(g(\tilde{B}_{z,n}^*,\tilde{M}_n^*,\tilde{U}_n^*)\middle|\,\tilde{\sigma}_n,\tilde{B}_{v,n}\right)\xrightarrow{a.s.}E\left(g(\tilde{B}_z^*,\tilde{M}^*,\tilde{V})\middle|\,\tilde{\sigma}\right)$$

for continuous bounded real g, where $\tilde{B}_z^*$ is a standard Brownian motion independent of $(\tilde{\sigma},\tilde{B}_z)$, and $\tilde{M}^*(\cdot):=\int_0^\cdot\tilde{\sigma}\mathsf{d}\tilde{B}_z^*$. Given that $\{\tilde{w}_t^*\}$ and $(\tilde{\sigma},\tilde{B}_z)$ are independent, as in the proof of (A.4), we could proceed by fixing $\{(\gamma_n,b_n)\}_{n\ge1}\in\Gamma\mathrm{B}$, where $\Gamma\mathrm{B}$ is an appropriate set with $P((\tilde{\sigma}_n,\tilde{B}_{v,n})_{n\ge1}\in\Gamma\mathrm{B})=1$, and then discuss the standard weak convergence of $(\tilde{B}_{z,n}^*,\tilde{M}_n^*,\tilde{U}_n^*)$ as a transformation of $(\gamma_n,b_n,\{\tilde{w}_t^*\})$ instead of $(\tilde{\sigma},\tilde{B}_z,\{\tilde{w}_t^*\})$. Since now $(\tilde{\sigma}_n,\tilde{B}_{v,n})$ and $\{\tilde{w}_t^*\}$ are defined on a product space, we implement this equivalently by fixing outcomes $\tilde{\omega}$ in the component space of $(\tilde{\sigma}_n,\tilde{B}_{v,n})$ and letting the outcome in the component space of $\{\tilde{w}_t^*\}$ be the only source of randomness. In what follows, fix an $\tilde{\omega}$ in a probability-one set where convergence (A.2) holds. Then

$$n^{-1/2}\sum_{t=1}^{\lfloor n\cdot\rfloor}\tilde{z}_{nt}(\tilde{\omega})\tilde{w}_t^*\xrightarrow{w}B_z^*(\cdot),$$

because $n^{-1}\sum_{t=1}^{\lfloor n\cdot\rfloor}\tilde{z}_{nt}^2(\tilde{\omega})=Q_{z,n}(\cdot,\tilde{\omega})\to Q(\cdot)$ and $L_n(\tilde{\omega})\to 0^\infty$ by the choice of $\tilde{\omega}$, where

$$L_n(\tilde{\omega})=\left\{n^{-1}\sum\nolimits_{t=1}^{\lfloor n\cdot\rfloor}\tilde{z}_{nt}^2(\tilde{\omega})\mathbb{I}(|\tilde{z}_{nt}(\tilde{\omega})|>\sqrt{n}/m)\right\}_{m\in\mathbb{N}}.$$

It follows that $\tilde{M}_n^*(\cdot,\tilde{\omega}) = n^{-1/2}\sum_{t=1}^{\lfloor n\cdot\rfloor}\tilde{\sigma}_t(\tilde{\omega})\tilde{z}_{nt}(\tilde{\omega})\tilde{w}_t^* \xrightarrow{w} \int_0^{\cdot}\tilde{\sigma}(\tilde{\omega})\mathsf{d}\tilde{B}_z^*$. Further,

$$\begin{aligned}\tilde{U}_n^*(\cdot,\tilde{\omega}) &= n^{-1}\sum_{t=1}^{\lfloor n\cdot\rfloor}\tilde{\sigma}_t^2(\tilde{\omega})\tilde{z}_{nt}^2(\tilde{\omega})\tilde{w}_t^{*2} \\ &= \tilde{U}_n(\cdot) + n^{-1}\sum_{t=1}^{\lfloor n\cdot\rfloor}\tilde{\sigma}_t^2(\tilde{\omega})\tilde{z}_{nt}^2(\tilde{\omega})(\tilde{w}_t^{*2}-1) \xrightarrow{p} \tilde{V}(\cdot,\tilde{\omega}),\end{aligned}$$

using Theorem A.1 of Cavaliere and Taylor (2009). Since $\tilde{V}(\cdot,\tilde{\omega})$ is non-random, the last two convergences are joint:

$$E\left[g\left(\tilde{M}_n^*(\cdot,\tilde{\omega}),\tilde{U}_n^*(\cdot,\tilde{\omega})\right)\right] \to E\left[g\left(\tilde{M}^*(\cdot,\tilde{\omega}),\tilde{V}(\cdot,\tilde{\omega})\right)\Big|\,\tilde{\sigma}\right]$$

for continuous and bounded real g. This implies, by the product structure of the probability space and the probability-one set of eligible outcomes $\tilde{\omega}$, that

$$E\left(g(\tilde{M}_n^*,\tilde{U}_n^*)|\tilde{\sigma}_n,\tilde{B}_{v,n}\right) \xrightarrow{a.s.} E\left(g(\tilde{M}^*,\tilde{V})|\tilde{\sigma}\right),$$

and eventually, as $(\tilde{M}^*,\tilde{V},\tilde{\sigma}) \overset{d}{=} (\tilde{M},\tilde{V},\tilde{\sigma})$, that

$$E\left(g(\tilde{M}_n^*,\tilde{U}_n^*)|\tilde{\sigma}_n,\tilde{B}_{v,n}\right) \xrightarrow{a.s.} E\left(g(\tilde{M},\tilde{V})|\tilde{\sigma}\right).$$

Notice that conditioning on $\tilde{\sigma}_n,\tilde{B}_{v,n}$ can be replaced by conditioning on $\tilde{D}_n$ because $(\tilde{M}_n^*,\tilde{U}_n^*)$ is a measurable function of $\tilde{D}_n$ and $\{\tilde{w}_t^*\}$.

We can conclude from (A.4) and this result that

$$\left(E\left[h(\tilde{M}_n,\tilde{U}_n)\Big|\,\tilde{\sigma}_n\right],E\left[g(\tilde{M}_n^*,\tilde{U}_n^*)\Big|\,\tilde{D}_n\right]\right) \xrightarrow{a.s.} \left(E\left[h(\tilde{M},\tilde{V})\Big|\,\tilde{\sigma}\right],E\left[g(\tilde{M},\tilde{V})\Big|\,\tilde{\sigma}\right]\right)$$

for all continuous and bounded real h,g, whereas on a general probability space

$$\left(E\left[h(M_n,U_n)|\,\sigma_n\right],E\left[g(M_n^*,U_n^*)|\,D_n\right]\right) \xrightarrow{w} \left(E\left[h(M,V)|\,\sigma\right],E\left[g(M,V)|\,\sigma\right]\right), \quad \text{(A.5)}$$

because $(\tilde{\sigma}_n,\tilde{M}_n,\tilde{U}_n,\tilde{D}_n,\tilde{M}_n^*,\tilde{U}_n^*) \overset{d}{=} (\sigma_n,M_n,U_n,D_n,M_n^*,U_n^*)$. This is precisely the definition of the joint $\xrightarrow{w}_w$ convergence in the theorem. □

Proof of Corollary 1. From (A.5) with $h=g$, if the random cdf $P(g(M,V)\le\cdot|\sigma)$ a.s. has continuous sample paths, conditional validity of the bootstrap as in Corollary 1 follows from Corollary 3.2 of Cavaliere and Georgiev (2019). □

Proof of Lemma 2. For any $K\in\mathbb{R}$, consider the continuous function $g_K:\mathbb{R}\to[0,1]$ defined by $g_K(x)=\mathbb{I}_{(-\infty,K]}(x)+(K+1-x)\mathbb{I}_{(K,K+1]}$. Then $\mathbb{I}_{(-\infty,K]}\le g_K\le\mathbb{I}_{(-\infty,K+1]}$ and the convergence $\tau_n^* \xrightarrow{w^*}_w \tau|\sigma$ implies that

$$F_n^*(K)\le E^*(g_K(\tau_n^*)) \xrightarrow{w} E(g_K(\tau)|\sigma)\le F_{\tau|\sigma}(K+1),$$

where $F_{\tau|\sigma}(K+1)=P(\tau\le K+1|\sigma)$. Therefore, for all $q\in(0,1)$,

$$\liminf_{n\to\infty}P(F_n^*(K)\le q)\ge P(F_{\tau|\sigma}(K+1)\le q).$$

As a result,

$$\begin{aligned}\liminf_{n\to\infty} P(F_n^*(\tau_n) \le q) &\ge \liminf_{n\to\infty} P(F_n^*(\tau_n) \le q, \tau_n \le K)\\ &\ge \liminf_{n\to\infty} P(F_n^*(K) \le q, \tau_n \le K)\\ &\ge \liminf_{n\to\infty} P(F_n^*(K) \le q) - \lim_{n\to\infty} P(\tau_n > K)\\ &\ge P(F_{\tau|\sigma}(K+1) \le q),\end{aligned}$$

since $\tau_n \xrightarrow{p} -\infty$ means that $\lim_{n\to\infty} P(\tau_n > K) = 0$ for all $K \in \mathbb{R}$. By Markov's inequality,

$$P(F_{\tau|\sigma}(K+1) \le q) \ge 1 - q^{-1}E(F_{\tau|\sigma}(K+1)),$$

and the proof is completed by letting $K \to -\infty$. □

Proof of eq. (20). Notice that

$$\begin{aligned}\hat{U}_n(\cdot) &= n^{-1}\sum_{t=1}^{\lfloor n\cdot\rfloor}\left(\sum_{i=0}^{t-1}\psi_i\varepsilon_{t-i}\right)^2\\ &= n^{-1}\sum_{t=1}^{\lfloor n\cdot\rfloor}\sum_{i=0}^{t-1}\psi_i^2\varepsilon_{t-i}^2 + n^{-1}\sum_{t=1}^{\lfloor n\cdot\rfloor}\sum_{i=0}^{t-1}\sum_{j=0}^{t-1}\psi_i\psi_j\varepsilon_{t-i}\varepsilon_{t-j}\\ &=: a_{1n}(\cdot) + a_{2n}(\cdot),\end{aligned}$$

with $a_{1n}(\cdot)$ and $a_{2n}(\cdot)$ implicitly defined. First, $a_{2n}(\cdot) = o_p(1)$ uniformly in $\cdot \in [0,1]$, similarly to Lemma A.7 in Cavaliere *et al.* (2010a). Second,

$$a_{1n}(\cdot) = n^{-1}\sum_{t=1}^{\lfloor n\cdot\rfloor}\varepsilon_t^2\left(\sum_{i=0}^{\lfloor n\cdot\rfloor - t}\psi_i^2\right) = \left(\sum_{i=0}^{\infty}\psi_i^2\right)U_n(\cdot) + b_n(\cdot),$$

with

$$b_n(\cdot) := n^{-1}\sum_{t=1}^{\lfloor n\cdot\rfloor}\varepsilon_t^2\left(\sum_{i=\lfloor n\cdot\rfloor - t+1}^{\infty}\psi_i^2\right).$$

Since the ψ_i's are exponentially decaying, there exist constants C and $\rho \in (0,1)$ such that $\sum_{i=\lfloor n\cdot\rfloor-t+1}^{\infty}\psi_i^2 \le C\rho^{\lfloor n\cdot\rfloor - t+1}$. Using the facts that $\max_{t=1,\dots,n}\sigma_t^2 = O_p(1)$ by Assumption 2 and $E(z_t^2) = 1$ by Assumption 1, it holds that

$$\begin{aligned}\sup_{u\in[0,1]} b_n(u) &\le Cn^{-1}\sup_{u\in[0,1]}\sum_{t=1}^{\lfloor nu\rfloor}\sigma_t^2 z_t^2\rho^{\lfloor n\cdot\rfloor - t+1}\\ &\le C\left(\max_{t=1,\dots,n}\sigma_t^2\right)\left(n^{-1}\max_{t=1,\dots,n} z_t^2\right)\sup_{u\in[0,1]}\left(\sum_{t=1}^{\lfloor n\cdot\rfloor}\rho^{\lfloor n\cdot\rfloor - t+1}\right)\\ &= O_p(1)o_p(1)\sum_{t=1}^{n}\rho^t = o_p(1).\end{aligned}$$

Hence, $\hat{U}_n(\cdot) = (\sum_{i=0}^{\infty}\psi_i^2)U_n(\cdot) + o_p(1)$. □

References

Andrews, D. W. K. (1997), "A Conditional Kolmogorov Test", *Econometrica* 65, 1097–1128.

Andrews, D. W. K. (2000), "Inconsistency of the Bootstrap When a Parameter Is on the Boundary of the Parameter Space", *Econometrica* 68, 399–405.

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

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.

Brown, B. M. (1971), "Martingale Central Limit Theorems", *Annals of Mathematical Statistics* 42, 59–66.

Casini, A. and P. Perron (2019), "Structural Breaks in Time Series", forthcoming in the *Oxford Research Encyclopedia of Economics and Finance.*

Cavaliere, G. and I. Georgiev (2019), "Inference Under Random Limit Bootstrap Measures", `arXiv:2930679`. Revision requested by *Econometrica.*

Cavaliere, G., H. B. Nielsen and A. Rahbek (2015), "Bootstrap Testing of Hypotheses on Cointegration Relations in VAR Models", *Econometrica* 83, 813–831.

Cavaliere, G., H. B. Nielsen and A. Rahbek (2017), "On the Consistency of Bootstrap Testing for a Parameter on the Boundary of the Parameter Space", *Journal of Time Series Analysis* 38, 513–534.

Cavaliere, G., A. Rahbek and A. M. R. Taylor (2010a), "Testing for Co-Integration in Vector Autoregressions with Non-Stationary Volatility", *Journal of Econometrics* 158, 7–24.

Cavaliere, G., A. Rahbek and A. M. R. Taylor (2010b), "Cointegration Rank Testing under Conditional Heteroskedasticity", *Econometric Theory* 26, 1719–1760.

Cavaliere, G. and A. M. R. Taylor (2007), "Testing for Unit Roots in Time Series Models with Non-Stationary Volatility", *Journal of Econometrics* 140, 919–947.

Cavaliere, G. and A. M. R. Taylor (2008), "Bootstrap Unit Root Tests for Time Series with Nonstationary Volatility", *Econometric Theory* 24, 43–71.

Cavaliere, G. and A. M. R. Taylor (2009), "Heteroskedastic Time Series with a Unit Root", *Econometric Theory* 25, 1228–1276.

Chan, N. H. and C. Z. Wei (1988), "Limiting Distributions of Least Squares Estimates of Unstable Autoregressive Processes," *Annals of Statistics* 16, 367–401.

Crimaldi, I. and L. Pratelli (2005), "Convergence Results for Conditional Expectations", *Bernoulli* 11, 737–745.

Deng, A. and P. Perron (2008), "A Non-Local Perspective on the Power Properties of the CUSUM and CUSUM of Squares Tests for Structural Change", *Journal of Econometrics* 142, 212–240.

Dudley, R. M. (2004), *Real Analysis and Probability.* Cambridge: Cambridge University Press.

Engle, R. F. and J. G. Rangel (2008), "The Spline-GARCH Model for Low-Frequency Volatility and Its Global Macroeconomic Causes", *The Review of Financial Studies* 21, 1187–1222.

Georgiev, I. (2018), "Asymptotics for Cointegrated Processes with Infrequent Stochastic Level Shifts and Outliers", *Econometric Theory* 24, 587–615.

Georgiev, I., D. I. Harvey, S. J. Leybourne and A. M. R. Taylor (2018), "Testing for Parameter Instability in Predictive Regression Models", *Journal of Econometrics* 204, 101–118.

Goggin, E. M. (1994), "Convergence in Distribution of Conditional Expectations", *Annals of Probability* 22, 1079–1114.

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

Gonçalves, S. and L. Kilian (2007), "Asymptotic and Bootstrap Inference for AR(∞) Processes with Conditional Heteroskedasticity", *Econometric Reviews* 26, 609–641.

Hansen, B. E. (1992), "Convergence to Stochastic Integrals for Dependent Heterogeneous Processes", *Econometric Theory* 8, 489–500.

Hansen, B. E. (1995), "Regression with Nonstationary Volatility", *Econometrica* 63, 1113–1132.

Hansen B. E. (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, 548–574.

Jacod, J. and A. N. Shiryaev (2003), *Limit Theorems for Stochastic Processes* (2nd ed.). Berlin: Springer.

Kallenberg, O. (1997), *Foundations of Modern Probability.* New York: Springer.

Kim, C. and C. Nelson (1999), "Has The U.S. Economy Become More Stable? A Bayesian Approach Based On A Markov-Switching Model Of The Business Cycle", *The Review of Economics and Statistics* 81, 608–616.

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

Lifshits, M. A. (1982) “On the Absolute Continuity of Distributions of Functionals of Random Processes”, *Theory of Probability and Its Applications* 27, 600–607.

Loretan, M. and P. C. B. Phillips (1994),“Testing the Covariance Stationarity of Heavy-Tailed Time Series: An Overview of the Theory with Applications to Several Financial Datasets”, *Journal of Empirical Finance* 1, 211–248.

McConnell, M. M. , and G. Perez-Quiros (2000), “Output Fluctuations in the United States: What Has Changed since the Early 1980’s?”, *American Economic Review* 90, 1464–1476.

Meijer, E. (2000), “An Asymmetric Distribution with Zero Skewness,” Working paper, University of Groningen, `http://dx.doi.org/10.2139/ssrn.2531847`.

Nelson, D. B. (1990), “ARCH Models as Diffusion Approximations”, *Journal of Econometrics* 45, 7–38.

Perron, P. (2006) “Dealing with Structural Breaks”, in Palgrave Handbook of Econometrics, Vol. 1: Econometric Theory, K. Patterson and T. C. Mills (eds.), Palgrave Macmillan, 2006, 278–352.

Ploberger, W. and W. Krämer (1992), “The Cusum Test with OLS Residuals”, *Econometrica* 60, 271–285.

Rao, M. M. and R. J. Swift (2006). *Probability Theory with Applications.* New York: Springer.

Sensier, M. and D. van Dijk (2004), “Testing for Volatility Changes in U.S. Macroeconomic Time Series”, *The Review of Economics and Statistics* 86, 833–839.

Xu, K.-L. and P. C. B. Phillips (2008), “Adaptive Estimation of Autoregressive Models with Time-Varying Variances”, *Journal of Econometrics* 142, 265–280.