ARTICLE IN PRESS
JID: ECOSTA [m3Gsc;June 9, 2023;10:16]

Econometrics and Statistics xxx (xxxx) xxx

Contents lists available at ScienceDirect

Econometrics and Statistics

journal homepage: www.elsevier.com/locate/ecosta

Robust Fixed-*b* Inference in the Presence of Time-Varying Volatility

Matei Demetrescu[a], Christoph Hanck[b,*], Robinson Kruse-Becher[c]

[a] *Faculty of Statistics, TU Dortmund University*
[b] *Faculty of Economics and Business Administration, University of Duisburg-Essen*
[c] *University of Hagen, Faculty of Business Administration and Economics*

ARTICLE INFO

Article history:
Received 5 April 2022
Revised 25 May 2023
Accepted 25 May 2023
Available online xxx

JEL classification:
C12 (Hypothesis Testing)
C32 (Time-Series Models)

Keywords:
Hypothesis testing
HAC estimation
HAR testing
Volatility

ABSTRACT

Time-varying volatility arises in many macroeconomic and financial applications. While "fixed-*b*" arguments provide refinements in the use of estimators for the asymptotic variance of GMM estimators, the resulting fixed-*b* distributions of test statistics are not pivotal under time-varying volatility. Three approaches to robustify inference are investigated: (i) wild bootstrapping, (ii) time transformations and (iii) selection of test statistics and critical values according to the outcome of a pretest for heteroskedasticity. Simulations quantify the distortions from using the original fixed-*b* approach and compare the effectiveness of the proposed corrections. Overall, the wild bootstrap is to be recommended. An empirical application to the Fama & French five factor model illustrates the relevance of the procedures.

1. Motivation

The seminal contributions of Newey and West (1987) and Andrews (1991) provide Generalized Methods of Moments (GMM) (Hansen, 1982) tests robust to heteroskedasticity and serial dependence for stationary series. Relying on heteroskedasticity and autocorrelation consistent (HAC) estimators of long-run covariance matrices (LRCov) yields conventional asymptotic normal or χ^2 null distributions of test statistics. HAC estimators are typically nonparametric and involve a bandwidth B as a tuning parameter and a kernel k. Standard asymptotics turn out to deliver rather poor approximations to the actual finite-sample distributions (Kiefer and Vogelsang, 2005). Moreover, finite-sample distributions are sensitive to the choice of B and k. This poor finite-sample performance of HAC estimators can be explained by the "small-b" requirement that a vanishing fraction $b := B/T \to 0$ of autocovariance matrices, with T the sample size, be used for HAC estimation, while $b > 0$ in actual applications.

To tackle these issues with HAC estimation, Kiefer et al. (2000) and Kiefer and Vogelsang (2002a, 2002b, 2005) propose the so-called fixed-b asymptotic framework, which allows for $b \in (0, 1]$. This yields nonstandard heteroskedasticity- and autocorrelation *robust* (HAR) null distributions reflecting the choice of B and k even in the limit and leading to substantially better approximations to actual finite-sample distributions. The usefulness of such procedures has spawned a very active

* Corresponding author.
E-mail addresses: mdeme@statistik.tu-dortmund.de (M. Demetrescu), christoph.hanck@vwl.uni-due.de (C. Hanck), robinson.kruse-becher@fernuni-hagen.de (R. Kruse-Becher).

https://doi.org/10.1016/j.ecosta.2023.05.003
2452-3062/

Please cite this article as: M. Demetrescu, C. Hanck and R. Kruse-Becher, Robust Fixed-*b* Inference in the Presence of Time-Varying Volatility, Econometrics and Statistics, https://doi.org/10.1016/j.ecosta.2023.05.003

literature (e.g., Sun et al., 2008; Vogelsang and Wagner, 2013; 2014; Sun, 2014; Shao, 2015; Preinerstorfer and Pötscher, 2016; Pötscher and Preinerstorfer, 2018; Rho and Vogelsang, 2021; Demetrescu et al., 2022; Casini, 2022). Lazarus et al. (2018) provide recommendations for HAR inference in practice. Sun et al. (2008) and Lazarus et al. (2021) address more specifically the size-power tradeoff in selecting b in a HAR framework.

Crucially, under time-varying volatility (TVV), e.g., non-smooth permanent breaks or trends in the variances, the HAR limiting distribution additionally depends on the variance process and thus leads to nonpivotality, as posited by Ibragimov and Müller (2010) and demonstrated in Section 2 below; see also Xu (2012). Such TVV is present in many time series such as growth and inflation rates (the "Great Moderation"—e.g., Stock and Watson, 2002) and financial returns (e.g., Amado and Teräsvirta, 2013). For stationary and unit root autoregressions, see e.g. Phillips and Xu (2006) or Cavaliere and Taylor (2008b). Correspondingly, inference under TVV receives increased attention.

The paper discusses various corrections restoring the asymptotic pivotality of fixed-b limiting theory for hypothesis testing. Concretely, Section 2 characterizes the distortions of fixed-b test statistics due to TVV in a GMM setting. Inspired by Hansen (2000), Section 3 discusses a wild bootstrap for valid fixed-b inference under TVV and establishes its asymptotic validity as $T \to \infty$; see, e.g., Shao (2010) or Gonçalves and Vogelsang (2011) for relevant alternative bootstrap approaches in the present HAR setup. We furthermore consider a time-transformation approach (cf. Cavaliere and Taylor, 2008b), and a pretest for heteroskedasticity. In this sense, we complement recent work of Casini (2022, Secs. 4-5), who provides results for MSE-optimal bandwidth choices in a HAR framework under a broad class of "smooth" nonstationary processes. His results relate to least squares estimation and are mainly described in the spectral domain for a fairly general class of locally stationary processes using an in-fill asymptotic framework, and valid inference is achieved via a new "double" kernel. In contrast, our aim is to provide a quantification of the non-pivotality of fixed-b inference under deterministic patterns of TVV and an assessment of the effectiveness of the studied correction approaches.

Simulation results (Section 4) support the theory. HAC-based tests are size-distorted unless T is large; under homoskedasticity, standard fixed-b is a remedy. Fixed-b tests may however be heavily size distorted under TVV. The comparison shows that the corrections generally yield good finite-sample size. The time transformation is somewhat oversized for shorter series and strong autocorrelation, while the wild bootstrap is very accurate. The pretest has an intermediate position. Moreover, the wild bootstrap is the most powerful correction and is thus recommended. In an empirical application (Section 5), we use the wild bootstrap procedure for testing in the Fama-French five factor model. Proofs, additional simulation results and further remarks have been collected in the appendix.

2. The testing problem in a GMM framework

Consider the linear regression model

$$y_t = \boldsymbol{\beta}'\boldsymbol{x}_t + u_t, \quad t = 1, \dots, T, \tag{1}$$

where we are interested in inference on the vector $\boldsymbol{\beta} \in \mathbb{R}^K$ in a GMM framework using $L \geq K$ instruments $\boldsymbol{z}_t$ and a symmetric positive definite weighting matrix $\mathbf{W}_T$. The simplest case is OLS regression on an intercept, i.e. $x_t = z_t = 1\ \forall t$. The disturbances u_t exhibit TVV via a standard multiplicative component structure (following, e.g., Cavaliere, 2004) as specified by

Assumption 1. (i) $u_t = h_t\varepsilon_t$, $t = 1, \dots, T$, where $h_t = h(t/T) > 0$ is deterministic, piecewise Lipschitz with $\int_0^1 h^2(s)\mathrm{d}s = 1$. (ii) $\{\boldsymbol{x}_t, \boldsymbol{z}_t, \varepsilon_t\}$ is a strictly stationary sequence, $L_{4+2\delta}$-bounded for some $\delta > 0$, and strong mixing with coefficients α_j s.th. $\sum_{j\geq 0} \alpha_j^{1/p-1/(2+\delta)} < \infty$ for some $2 < p < 2+\delta$. (iii) $\boldsymbol{z}_t\varepsilon_t$ is zero-mean with positive definite LRCov $\boldsymbol{\Omega}$, and $\mathrm{E}\left(\boldsymbol{x}_t\boldsymbol{z}_t'\right) = \boldsymbol{\Upsilon}$, a full-rank $K \times L$ matrix. Finally, (iv) $\mathbf{W}_T \xrightarrow{p} \mathbf{W}$ with $\mathbf{W}$ a positive definite matrix.

Part (i) allows for general patterns of smoothly or abruptly changing unconditional variances, provided the latter are not too frequent (e.g., seasonally varying variances are excluded). The condition $\int_0^1 h^2(s)\mathrm{d}s = 1$ is a normalization; one may alternatively restrict $\boldsymbol{\Omega}$. Together with the Lipschitz continuity of h, parts (ii) and (iii) determine the behavior of partial sums specified below. In particular, short memory of the stochastic components is required for convergence to a limit process with independent increments, and imposing strong mixing conditions is a convenient way of imposing it here. For instance, mild serial correlation and also mild forms of conditional heteroskedasticity are allowed for under (ii) for the standardized disturbances ε_t. Parts (iii) and (iv) lay out standard regularity conditions for models like (1) in the linear GMM framework.

Letting $\mathbf{S}_{\boldsymbol{xz}} := \sum_{t=1}^T \boldsymbol{x}_t\boldsymbol{z}_t'$ and $\mathbf{M} = \mathbf{S}_{\boldsymbol{xz}}\mathbf{W}_T\mathbf{S}_{\boldsymbol{xz}}'$, the GMM estimator for $\boldsymbol{\beta}$ in (1) is

$$\hat{\boldsymbol{\beta}} = \mathbf{M}^{-1}\mathbf{S}_{\boldsymbol{xz}}\mathbf{W}_T\mathbf{S}_{y\boldsymbol{z}}', \tag{2}$$

with $\mathbf{S}_{y\boldsymbol{z}} = \sum_{t=1}^T y_t\boldsymbol{z}_t'$. The usual HAC covariance matrix estimator is given by

$$\widehat{\mathrm{Cov}}\left(\hat{\boldsymbol{\beta}}\right) = T\,\mathbf{M}^{-1}\mathbf{S}_{\boldsymbol{xz}}\mathbf{W}_T\hat{\boldsymbol{\Omega}}\mathbf{W}_T\mathbf{S}_{\boldsymbol{xz}}'\mathbf{M}'^{-1}, \tag{3}$$

where

$$\hat{\boldsymbol{\Omega}} = \sum_{j=-T+1}^{T-1} k(j/B)\hat{\boldsymbol{\Gamma}}_j \tag{4}$$

is an estimator of the LRCov of $\boldsymbol{z}_t u_t$: working with residuals $\hat{u}_t = y_t - \hat{\boldsymbol{\beta}}' \boldsymbol{x}_t$, we have $\hat{\boldsymbol{\Gamma}}_{|j|} = T^{-1} \sum_{t=|j|+1}^{T} \left(\boldsymbol{z}_t \hat{u}_t - \overline{\boldsymbol{z}\hat{u}} \right) \left(\boldsymbol{z}_{t-|j|} \hat{u}_{t-|j|} - \overline{\boldsymbol{z}\hat{u}} \right)'$, $\hat{\boldsymbol{\Gamma}}_{-|j|} = \hat{\boldsymbol{\Gamma}}'_{|j|}$.

Practitioners often compute the LRCov estimate with prewhitening (Andrews and Monahan, 1992). Fitting a VAR(p) process to $\boldsymbol{\psi}_t := \boldsymbol{z}_t \hat{u}_t$, $\boldsymbol{\psi}_t = \sum_{j=1}^{p} \tilde{\mathbf{A}}_j \boldsymbol{\psi}_{t-j} + \tilde{\boldsymbol{v}}_t$, one estimates the long-run covariance matrix $\hat{\boldsymbol{\Omega}}_{\boldsymbol{v}}$ of the VAR residuals along the same lines as in (4), then computes

$$\hat{\boldsymbol{\Omega}}_{pw} = \left(\mathbf{I} - \sum_{j=1}^{p} \hat{\mathbf{A}}_j \right)^{-1} \hat{\boldsymbol{\Omega}}_{\boldsymbol{v}} \left(\mathbf{I} - \sum_{j=1}^{p} \hat{\mathbf{A}}_j \right)^{\prime -1},$$

to be used alternatively to (4).

A Wald-type test for $H_0 : \boldsymbol{\beta} = \boldsymbol{\beta}_0$ is therefore given by

$$\mathcal{T}_K = \frac{1}{T} \left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0 \right)' \left(\mathbf{M}^{-1} \mathbf{S}_{\boldsymbol{xz}} \mathbf{W}_T \hat{\boldsymbol{\Omega}} \mathbf{W}_T \mathbf{S}'_{\boldsymbol{xz}} \mathbf{M}^{-1} \right)^{-1} \left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0 \right). \tag{5}$$

Crucially, fixed-b asymptotics rely on the limiting behavior of partial sums rather than on the small-b convergence in probability of $\hat{\boldsymbol{\Omega}}$ (or $\hat{\boldsymbol{\Omega}}_{pw}$). Assumption 1 implies $T^{-1/2} \sum_{t=1}^{[sT]} \boldsymbol{z}_t \varepsilon_t \Rightarrow \boldsymbol{\Omega}^{1/2} \boldsymbol{W}_L(s)$ with $\boldsymbol{W}_L(s)$ a vector of L independent standard Wiener processes (Davidson, 1994, Chapter 29). Thus

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{[sT]} \boldsymbol{z}_t u_t = \frac{1}{\sqrt{T}} \sum_{t=1}^{[sT]} h_t \boldsymbol{z}_t \varepsilon_t \Rightarrow \int_0^s h(r) \boldsymbol{\Omega}^{1/2} \mathrm{d}\boldsymbol{W}_L(r).$$

Notice that $\int_0^s h(r) \mathrm{d}\boldsymbol{W}_L(r)$ is a vector of L independent *time-transformed* Wiener processes. Denoting the variance profile $\eta(s) = \int_0^s h^2(r)\mathrm{d}r / \int_0^1 h^2(r)\mathrm{d}r$, it follows that $\int_0^s h(r)\mathrm{d}\boldsymbol{W}_L(r) \equiv \boldsymbol{W}_L(\eta(s))$, where $\equiv$ stands for equivalence in distribution; cf. Cavaliere (2004).

To characterize the limiting distributions, let, for any process $\boldsymbol{A}(s)$ with a.s. continuous paths,

$$\boldsymbol{\Theta}_{k,b}(\boldsymbol{A}(\cdot)) := \begin{cases} -\frac{1}{b^2} \int_0^1 \int_0^1 k'' \left(\frac{r-s}{b} \right) \bar{\boldsymbol{A}}(r) \bar{\boldsymbol{A}}(s)' \,\mathrm{d}r\mathrm{d}s & \text{smooth } k \\ \frac{1}{b} \left(2 \int_0^1 \bar{\boldsymbol{A}}(r) \bar{\boldsymbol{A}}(r)' \mathrm{d}r - \int_0^{1-b} \left[\bar{\boldsymbol{A}}(r+b) \bar{\boldsymbol{A}}(r)' + \bar{\boldsymbol{A}}(r) \bar{\boldsymbol{A}}(r+b)' \right] \mathrm{d}r \right) & \text{Bartlett,} \end{cases}$$

where $\bar{\boldsymbol{A}}(s) = \boldsymbol{A}(s) - s\boldsymbol{A}(1)$. We are now in the position to characterize the fixed-b limits of $\mathcal{T}_K$:

Proposition 1. *Under Assumption 1 and the sequence of local alternatives $\boldsymbol{\beta} = \boldsymbol{\beta}_0 + \boldsymbol{c}/\sqrt{T}$ with $\boldsymbol{c} \in \mathbb{R}^K$, it holds for $b \in (0, 1]$ as $T \to \infty$ that*

$$\mathcal{T}_K \xrightarrow{d} \mathcal{D}_{h,k,b,\boldsymbol{c}} := (\boldsymbol{W}_K(1) + \boldsymbol{\mu})' \boldsymbol{\Theta}_{k,b}^{-1} (\boldsymbol{W}_K(\eta(\cdot))) (\boldsymbol{W}_K(1) + \boldsymbol{\mu}),$$

with noncentrality parameter $\boldsymbol{\mu} = \left(\boldsymbol{\Upsilon} \mathbf{W} \boldsymbol{\Omega} \mathbf{W} \boldsymbol{\Upsilon}' \right)^{-1/2} \left(\boldsymbol{\Upsilon} \mathbf{W} \boldsymbol{\Upsilon}' \right) \boldsymbol{c}$.

Proof. See the Appendix. □

The result extends straightforwardly to tests of linear restrictions, in particular individual slope coefficient tests. The resulting distributions belong to the same family – the only difference being that K in $\boldsymbol{W}_K(\cdot)$ is replaced by the number of imposed restrictions.

Remark 1. The expressions may simplify in the case of optimally weighted GMM, where $\mathbf{W}_T = \hat{\boldsymbol{\Omega}}^{-1}$, although, as pointed out by a referee, care must be taken in such a two-step GMM procedure since, essentially, classical GMM optimality theory requires that $\mathbf{W}_T$ converges to a deterministic matrix – which of course need not be the case under fixed-b asymptotics when choosing $\hat{\boldsymbol{\Omega}}^{-1}$ as weighting matrix. Furthermore, Proposition 1 would require appropriate modifications. For instance, the noncentrality parameter $\boldsymbol{\mu}$ would depend on $\boldsymbol{\Theta}_{k,b}$ rather than $\boldsymbol{\Omega}$.

Under H_0 ($\boldsymbol{c} = \mathbf{0}$), we obtain the fixed-$b$ limiting distribution

$$\mathcal{T}_K \xrightarrow{d} \boldsymbol{W}'_K(1)\, \boldsymbol{\Theta}_{k,b}^{-1} (\boldsymbol{W}_K(\eta(\cdot)))\, \boldsymbol{W}_K(1). \tag{6}$$

This null distribution does not depend on $\boldsymbol{\Omega}$, but does depend on h – and only coincides with the corresponding fixed-b distribution of Kiefer and Vogelsang (2005) (say $\mathcal{D}_{k,b}$) if $h = \text{cst.}$ a.e. This result elaborates upon Ibragimov and Müller (2010, p. 545) that HAR testing is not robust to TVV by providing the fixed-b asymptotic distribution under TVV. Xu (2012) states lack of invariance of HAR in the special case $b = 1$ and the Bartlett kernel when testing for trend slopes.

Under additional conditions (e.g., Andrews, 1991) and in particular $b \to 0$ at suitable rates, $\hat{\boldsymbol{\Omega}} \xrightarrow{p} \boldsymbol{\Omega} \cdot \int_0^1 h^2(s)\mathrm{d}s = \boldsymbol{\Omega}$ even under TVV (Cavaliere, 2004, Lemma 4). Hence,

$$\mathcal{T}_K \xrightarrow{d} \chi^2_K \quad \text{under } H_0 \text{ and small } b, \tag{7}$$

so that HAC tests are robust to TVV. While the distribution of HAC-based $\mathcal{T}_K$ thus also does not depend on the choice of k and B, its finite-sample behavior hinges on both. In particular, the finite-sample quality of small-b asymptotics is poor (see also Section 4), so practitioners have to choose between two evils under uncertainty about h_t. Thus, the three novel corrections for fixed-b inference discussed in the following are relevant for econometric practice. (Ibragimov and Müller (2010) provide a partial solution to the present problem, constructing valid t-statistics when one can partition y_t into groups such that estimators for each are approximately independent, unbiased and Gaussian. The approach is a bit restrictive, however, relying on a result (Bakirov and Székely, 2005) guaranteeing size control only for $\alpha \leqslant 0.083$.)

Before proceeding, we note that Assumption 1 assumes the unconditional heteroskedasticity to be deterministic in nature. While deterministic volatility is a reasonable assumption for low-frequency data, this does rule out stochastic volatility. Results are available under which our results may be extended to cover stochastic volatility as well; see e.g. Cavaliere and Taylor (2008a) as well as more work (e.g., Boswijk et al., 2021). They however only apply for the bootstrap, and it is unclear how the time-transformation approach we also study would fit this extended framework. Therefore, a comparison of the corrections is not possible under stochastic volatility and we omit the topic.

Another possible extension concerns nonstationarity in the regressors or the instruments. Assumption 1 currently models unconditional heteroskedasticity as arising via the errors, with parts (ii) and (iii) not allowing for unconditional heteroskedasticity in the instruments and regressors. Clearly, this issue is moot in the above-mentioned leading case of a regression on a constant. In more general cases, we notice examining the proof of Proposition 1 that the result requires on the one hand the convergence $\frac{1}{T}\sum_{t=1}^{[sT]}\boldsymbol{z}_t\boldsymbol{x}_t' \xrightarrow{p} s\boldsymbol{\Upsilon}'$ (uniformly in s) as $T\to\infty$. So it appears that a constant cross-product moment matrix of $\boldsymbol{z}_t$ and $\boldsymbol{x}_t$ would not affect the result. A second key element of the proof is the weak convergence weak of the moment conditions to a time-transformed Brownian motion, $\frac{1}{\sqrt{T}}\sum_{t=1}^{[sT]}\boldsymbol{z}_t u_t \Rightarrow \boldsymbol{\Omega}^{1/2}\boldsymbol{W}_L(\eta(s))$. Unconditional heteroskedasticity in the instruments would lead to a more general Gaussian process as limit, which would again prevent us from comparing the three corrections in the following section. More specifically, while the wild bootstrap could be fixed by resorting to a more complex joint block wild bootstrap, it is not clear at all how to modify the time-transformation approach and we do not further pursue this topic here.

Also, Assumption 1 rules out strong persistence of the unit root-type, like most other work on HAR inference (see, e.g., Casini, 2022, and references cited therein). We however note that—although that is outside the scope of the present paper—the fixed-b framework is amenable to such persistence (e.g., Vogelsang and Wagner, 2013; 2014).

3. Approaches for robust inference under TVV

3.1. The wild bootstrap

We first exploit the wild bootstrap to estimate the actual null distribution of $\mathcal{T}_K$ under TVV. Its use to robustify against conditional or unconditional heteroskedasticity can be traced back to at least Hansen (2000, p. 106). See also Xu (2012) for the trend case and Cavaliere and Taylor (2008a) for the unit root context. The simplest, yet effective, fixed-regressor algorithm is as follows.

1. Estimate (1) and compute unrestricted residuals $\hat{u}_t = y_t - \hat{\boldsymbol{\beta}}'\boldsymbol{x}_t$.
2. Generate T iid standardized draws r_t^* (e.g., from $\mathcal{N}(0,1)$ or the Mammen distribution) and the bootstrap disturbances $u_t^* = r_t^*\hat{u}_t$. One may alternatively bootstrap the residuals from an AR(p) fit of $\hat{u}_t$, $\hat{u}_t = \sum_{j=1}^{p}\tilde{a}_j\hat{u}_{t-j} + \tilde{v}_t$ where $\tilde{}$ denotes OLS estimation (since the residuals—and hence also the bootstrap population of u_t^*—have zero mean whenever an intercept is included in the model, the autoregression does not require an intercept itself). Then, one may generate bootstrap errors $v_t^* = r_t^*\tilde{v}_t$, and either recolor them to obtain u_t^* or directly set $u_t^* = v_t^*$ without recoloring.
3. Generate the fixed-regressor wild bootstrap sample as $y_t^* = \boldsymbol{\beta}_0'\boldsymbol{x}_t + u_t^*$. For individual t-statistics, simply use the estimators of $\boldsymbol{\beta}$ imposing the individual null hypothesis.
4. Compute the bootstrap test statistic
$$\mathcal{T}_K^* = \frac{1}{T}\left(\hat{\boldsymbol{\beta}}^* - \boldsymbol{\beta}_0\right)'\left(\mathbf{M}^{-1}\mathbf{S}_{\boldsymbol{xz}}\mathbf{W}_T\hat{\boldsymbol{\Omega}}^*\mathbf{W}_T\mathbf{S}_{\boldsymbol{xz}}'\mathbf{M}^{-1}\right)^{-1}\left(\hat{\boldsymbol{\beta}}^* - \boldsymbol{\beta}_0\right),$$
with $\hat{\boldsymbol{\Omega}}^* = \sum_{j=-T+1}^{T-1}k(j/B)\hat{\boldsymbol{\Gamma}}_j^*$ and $\hat{\boldsymbol{\Gamma}}_{|j|}^* = T^{-1}\sum_{t=|j|+1}^{T}\left(\boldsymbol{z}_t\hat{u}_t^* - \overline{\boldsymbol{z}\hat{u}^*}\right)\left(\boldsymbol{z}_{t-|j|}\hat{u}_{t-|j|}^* - \overline{\boldsymbol{z}\hat{u}^*}\right)'$.
5. Use the quantiles of the resulting bootstrap distribution, $q_{1-\alpha}^*$, for inference.

Typically, one resorts to Monte Carlo simulation to approximate the desired bootstrap quantiles: repeat Steps 1-4 to obtain M resampled statistics $\left\{\mathcal{T}_{K,m}^*\right\}_{m=1,\dots,M}$ and use their sample $(1-\alpha)$-quantile, as critical value.

Proposition 2. *Under H_0, Assumption 1, $\mathrm{E}(|r_t^*|^k) < \infty\ \forall k\in\mathbb{N}$ and $T\to\infty$, $\Pr\left(\mathcal{T}_K > q_{1-\alpha}^*\right)\to\alpha$.*

Proof. See the Appendix. □

Proposition 2 shows that this wild bootstrap asymptotically controls size. Since the bootstrap generates critical values which are invariant to the true $\boldsymbol{\beta}$, the power under (local) alternatives of the test based on $\mathcal{T}_K$ remains as implied by Proposition 1.

Remark 2. As pointed out by a referee and discussed in, e.g., Davidson and MacKinnon (1999) and Djogbenou et al. (2019), the use of restricted residuals $u_t^0 = y_t - \boldsymbol{\beta}_0' \boldsymbol{x}_t$ in step 1 of the procedure may lead to better size control. We only resort to unrestricted residuals as the wild bootstrap generally already offers very good size control (see Section 4), with the main remaining distortions arising from serial correlation that such a restricted bootstrap variant would also not be likely to improve upon. We also refer to the discussion in Section 3.2 that the use of restricted residuals is, albeit in a slightly different setup, not without (power) issues of its own.

As, unlike u_t, the u_t^* are white noise if not recoloring v_t^*, it may be surprising at a first glance that the wild bootstrap is still valid. The key idea is that the wild bootstrap is intended to replicate $\mathcal{D}_{h,k,b,\mathbf{0}}$, which is invariant to serial correlation, but not to TVV. Thus, we need to replicate the variance profile $\eta(s)$ (which the bootstrap does achieve in all considered variations, see the proof), but not necessarily the serial correlation structure.

3.2. Time transformations

The second correction builds on Cavaliere and Taylor (2008b). It essentially transforms u_t back to homoskedasticity, making it valid to apply fixed-b methods to the transformed series. Their approach dealing with $I(1)$ needs to be adapted to our regression setup for $I(0)$ processes:

1. Compute regression errors under H_0, i.e. $u_t^0 = y_t - \boldsymbol{\beta}_0' \boldsymbol{x}_t$ and build $\boldsymbol{\zeta}_t = \sum_{j=1}^t \boldsymbol{z}_j u_j^0$. For t-statistics, use the estimator under the null, $\hat{\boldsymbol{\beta}}_{H_0}$ instead of $\boldsymbol{\beta}_0$.
2. Estimate the variance profile as $\hat{\eta}(s) = \sum_{t=1}^{[sT]} \hat{u}_t^2 / \sum_{t=1}^T \hat{u}_t^2$ with $\hat{u}_t = y_t - \hat{\boldsymbol{\beta}}' \boldsymbol{x}_t$ the unrestricted residuals.
3. Build the inverse $g(s)$ of $\hat{\eta}(s)$ and time transform $\boldsymbol{\zeta}$ via $\tilde{\boldsymbol{\zeta}}_t = \boldsymbol{\zeta}_{[Tg(t/T)]}$.
4. Compute $\Delta \tilde{\boldsymbol{\zeta}}_t$ and let $\tilde{\boldsymbol{\Omega}}$ estimate its LRCov with $B = [bT]$. The modified test statistic $\tilde{\mathcal{T}}_K$ is obtained by replacing $\hat{\boldsymbol{\Omega}}$ with $\tilde{\boldsymbol{\Omega}}$ in (3).

We now show that fixed-b asymptotics are recovered under H_0. Here, we make use of the result that $\hat{\eta}$ is uniformly consistent, such that $g(s)$ also converges uniformly to $\eta^{-1}(s)$. See Cavaliere and Taylor (2008b) for theoretical and computational details. Under local alternatives, however, the transformation has an asymptotic effect, unlike for the wild bootstrap:

Proposition 3. *Let $\mathcal{K}(s) := \boldsymbol{W}_K(s) + \eta^{-1}(s)\boldsymbol{\mu}$, with $\eta^{-1}(s)$ the inverse of $\eta(s)$. Under the assumptions of Proposition 1,*

$$\tilde{\mathcal{T}}_K \xrightarrow{d} \mathcal{K}'(1)\, \boldsymbol{\Theta}_{k,b}^{-1}(\mathcal{K}(\cdot))\, \mathcal{K}(1).$$

In particular, $\tilde{\mathcal{T}}_K \xrightarrow{d} \mathcal{D}_{k,b}$ for $\boldsymbol{c} = \mathbf{0}$.

Proof. See the Appendix. □

Proposition 3 leads to surprising behavior of $\tilde{\mathcal{T}}_K$ for large $||\boldsymbol{c}||$ unless $\eta^{-1}(s) = s$ (i.e., essentially for $h_t =$ cst.). Take e.g. the case of an intercept only, $z_t = x_t = 1$; then, $\tilde{\mathcal{T}}_1 \xrightarrow{p} 1/\Theta_{k,b}(\eta^{-1}(t)) =: G$ as $T \to \infty$ followed by $|c| \to \infty$, since c dominates in both numerator and in $\Theta_{k,b}(\eta^{-1}(t))$ whenever $\eta^{-1}(t)$ is nonlinear. Hence, $\tilde{\mathcal{T}}$ either always rejects (if G is in its critical region), or never (if G is not). Intuitively, the time transformation changes the shape of the local alternative, thus inducing a type of power breakdown previously observed for tests of structural change; cf. Vogelsang (1999).

One may be tempted to use the actual residuals $\hat{u}_t$ instead of the regression errors under the null u_t^0 in Step 1. This would eliminate the dependence of $\tilde{\boldsymbol{\Omega}}$ on $\boldsymbol{c}$ under the local alternative and avoid such power issues. At the same time, this would induce a change the null distribution, as is argued in Appendix B.

Possible WLS-type weighting of the observations would rely on standardizing y_t by a nonparametric estimate of h_t. Appendix B also argues that such variants may lack local power, too. We therefore do not further pursue weighting approaches here and advise against their use.

3.3. Pretesting

The third approach follows what applied researchers might do: only correct for a problem detected in the data. We first pretest for TVV and then work with either the fixed- or small-b approach, according to the outcome of the test. Concretely, we test the null of homoskedasticity in a first step. If the null is not rejected, we proceed with the fixed-b approach in the second step and compare the value of the test statistic $\mathcal{T}_K$ with fixed-b critical values derived from (6) assuming homoskedasticity, i.e., that $h =$ cst, such that $\eta(s) = s$ and $\boldsymbol{W}_K(\eta(\cdot)) = \boldsymbol{W}_K(\cdot)$ (see also Kiefer and Vogelsang, 2005, Thm. 3). If we do reject the null in the first step, we proceed with small-b inference in the second step and compare $\mathcal{T}_K$ with critical values from the χ_K^2 distribution. The intuition is that if the pretest does not reject, then under homoskedasticity the advantageous fixed-b asymptotics for $\mathcal{T}_K$ can be used. If the test rejects, small-b procedures may be preferable due to their robustness against TVV. As a pretest statistic we use

$$Q = \sup_{1 \le t \le T} \frac{1}{\sqrt{T}} \frac{\left| \sum_{j=1}^t \hat{u}_j^2 - \frac{t}{T} \sum_{j=1}^T \hat{u}_j^2 \right|}{\hat{\omega}_{\hat{u}^2}},$$

$\hat{\omega}^2_{\hat{u}^2}$ is a HAC estimator of the LRV of $\hat{u}^2_t$ (Deng and Perron, 2008). Under the null of homoskedasticity, $Q \Rightarrow \sup_{t\in[0,1]} |\bar{W}(t)|$. The test rejects for large values of Q and is consistent against a break in volatility. Clearly, other suitable tests for unconditional heteroskedasticity could also be used. We for example also experimented with the test suggested in Cavaliere and Taylor (2008b), with however slightly inferior small-sample properties.

Such pretesting—see above—aims to mimic and study a strategy that one regularly encounters in practice. Pretesting is often based on arguments along the following lines. Under unconditional homoskedasticity, both small-b and fixed-b procedures are valid and the pretest is not problematic in the limit. Similarly, under unconditional heteroskedasticity, one may build on consistency of the pretest and end up using the asymptotically valid small-b procedure with probability one in the limit. Of course, when Q has low power, it will relatively frequently, but wrongly (whenever the null of Q is false) suggest to use fixed-b critical values from (6) assuming $h = \text{cst}$ in the second step of the procedure. At the same time, Q is likely to have low power when heteroskedasticity is mild, such that the fixed-b critical values used in the second step may be "approximately" correct. Hence, this two-step strategy is intuitive to many practitioners and may provide relatively accurate inference. However, as is not uncommon, such pretesting strategies may induce distortions that are not shared by the alternatives whose (asymptotic) validity we establish in the previous two subsections. Concretely, the pretest procedure is prone to phenomena formalised in the literature on post-model selection inference (see e.g. Leeb and Pötscher, 2005, for a clarifying discussion of the impact of pretests on inference after variable selection in linear regression). Therefore, we further investigate the finite-sample properties of this approach in the following section.

4. Simulation evidence

4.1. Setup

In our simulations, we initally focus, similar to Kiefer and Vogelsang (2005), on the case with an intercept only. These simulations, for which results are reported in Section 4.2, will provide quantitative evidence illustrating the predictions of the above propositions on the (non-)robustness of the different approaches discussed here. Moreover, they compare the relative merits of the three corrections proposed in this paper. We study regression models in Section 4.3.

Thus, we test $H_0 : \beta = 0$ against $H_1 : \beta > 0$ in $y_t = \beta + u_t$ where $(1 - \phi L)u_t = h_t\varepsilon_t$, with $\varepsilon_t \sim iidN(0, 1)$ and $\phi = \{0.1, 0.85\}$. We also studied $\phi = 0.5$ and an ARMA(1,1), not reported for brevity. Obviously, the implied AR(∞) structure generally does not improve the performance of any test. However, it does not have an impact on the *relative* performance of the tests. As our focus is on TVV, we work with a relatively simple AR structure in what follows.

The DGPs for h_t follow Cavaliere and Taylor (2008b).

1. Downward break at $t = [0.2T]$ from $\sigma_0 = 5$ to $\sigma_1 = 1$.
2. Double break, upward (as in DGP3) at $[0.4T]$ and downward (back to initial level) at $[0.6T]$.
3. Upward break at $t = [0.8T]$ from $\sigma_0 = 1$ to $\sigma_1 = 5$.
4. Homoskedasticity ($h_t = 1$).

For power, we study $\beta_T = c\sqrt{\bar{\omega}^2/T}$. The average variance $\bar{\omega}^2$ depends on the particular DGP. Under homoskedasticity, $\bar{\omega}^2 = \sigma^2$, while e.g. $\bar{\omega}^2 = T^{-1}\sum_{t=1}^{T}(\sigma_0^2 + 1(t > [\tau T])\sigma_1^2) \rightarrow_{T\rightarrow\infty} \tau\sigma_0^2 + (1-\tau)\sigma_1^2$ under DGPs 1 and 3. We consider $c = \{2, 4, \dots, 16\}$, $\alpha = 0.1$, the quadratic spectral [QS] and Bartlett kernel and $b \in \{0, 0.1, \dots, 1\}$. We employ 5,000 Monte Carlo replications and set $M = 399$. The sample sizes are $T = \{100, 500\}$. To reduce the finite sample impact of short-run dynamics, one may also use prewhitened residuals in the bootstrap algorithm. In particular, we fit an AR(1) model such that $\hat{u}_t = y_t - \hat{a}_0 - \hat{a}_1 y_{t-1}$ are used instead of $y_t - \bar{y}$. It is not required to recolor the bootstrap shocks to obtain the correct limiting distribution, although it is possible. See the proof of Proposition 2 for details. In this sense, our wild bootstrap approach is related to Xu's (2012), who imposes a parametric (V)AR structure on the innovations and performs a sieve step. The prewhitening uses a data-driven choice for B (see Andrews and Monahan, 1992) for the Newey and West (1987, NW) approach.

4.2. Baseline location model

We focus on the Bartlett kernel for brevity; results for QS were similar. We waive to report qualitatively similar results for two additional DGPs with a down- and upward break and trending variances, as well as intermediate results for $T = 250$. Appendix D reports some additional results. The results for $b = 0$ in Figures 1 and 2 resemble that NW faces substantial size distortions for T as large as $T = 100$. These distortions are more pronounced the higher ϕ. Similar to Kiefer and Vogelsang (2005), the "homoskedasticity" panels (lower right) of Figures 1 and 2 further show that fixed-b asymptotics precisely approximate the finite-sample distribution of (5) for both T, all but eliminating the size distortions of NW. Similarly, the wild bootstrap and time transformation are accurate. 81011

DGPs 1 ("early down"), 2 ("up, then down") and 3 ("late up") illustrate (cf. Proposition 1) that fixed-b is not pivotal under TVV. In particular, the tests are conservative and increasingly so in b, except for DGP2 and large b. The finding that fixed-b works relatively well for small-b is expected, as it is then similar to the asymptotically valid NW approach (see above and Cavaliere, 2004). The wild bootstrap is again very accurate, with mild exceptions for large ϕ, small T and small b.

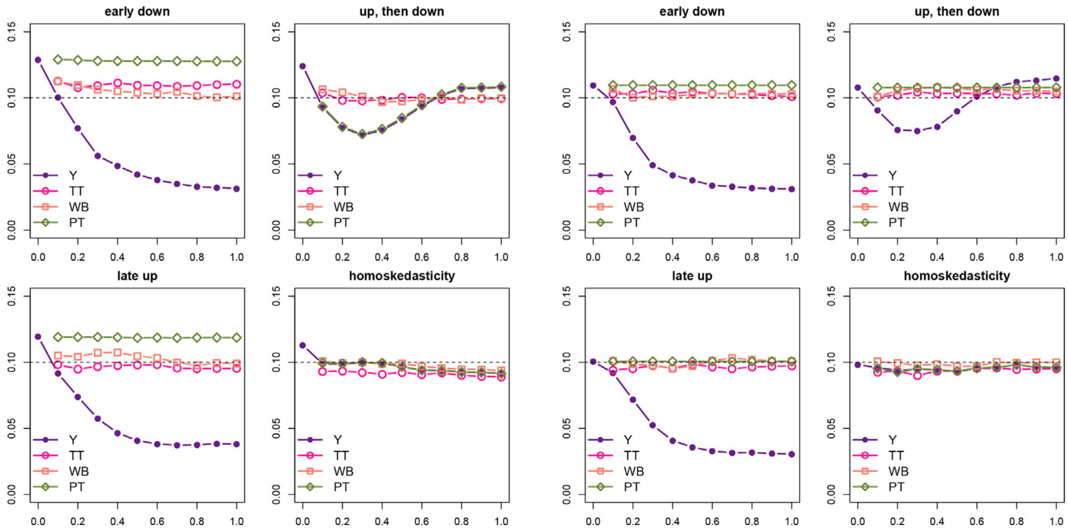

Fig. 1. Rejection rates plotted against b. Size, $\phi = 0.1$. Left block: $T = 100$, right block: $T = 500$. Y: standard fixed-b (NW for $b = 0$), TT: time transformation statistic (cf. Sec. 3.2), WB: wild bootstrap (cf. Prop. 2) and PT: pretest (cf. Section 3.3).

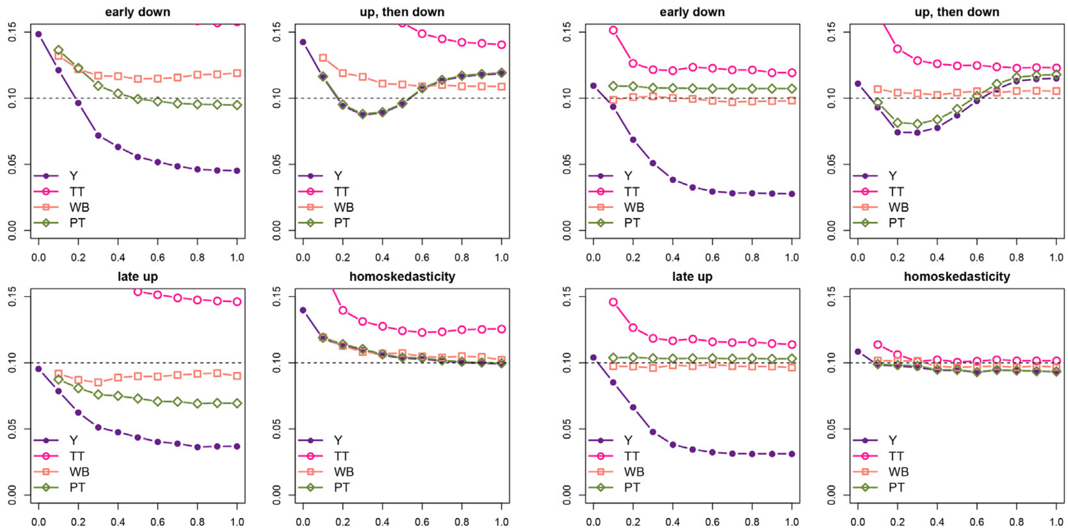

Fig. 2. Size, $\phi = 0.85$. Left: $T = 100$, right: $T = 500$. See notes to Figure 1.

The pretest's size ranges between the one of NW and the fixed-b approaches. When the pretest does not reject frequently, e.g., for small T and/or homoskedasticity and DGP2, its size tracks that of fixed-b for all b. The panels for different T for DGP2 confirm that higher power of the pretest results in a behavior more like that of NW. The time transformation performs well for large T (as predicted by Proposition 3) and small ϕ. Figure 2 reveals poor empirical size for $T = 100$. This is because the transformation may, for small T, produce stretches of identical observations that do not mimic the dynamics of the underlying series well.

The right panels of Figures 1 and 2 highlight the finite-sample character of the size distortions of NW and the wild bootstrap, which are removed for $T = 500$. These results further confirm that distortions for fixed-b are *not* of a finite-sample nature under TVV.

Figures 3 and 4 report power results for $T = 100$ (Appendix D reports additional qualitatively similar results with stronger serial correlation). As expected, power generally increases in c. The exception is the time transformation under TVV, where (see the discussion below Proposition 3) $\tilde{\mathcal{T}} \xrightarrow{p} G$ yields zero local power as $c \to \infty$ for the DGPs considered here. Overall, the wild bootstrap is most powerful. Essentially, it performs size adjustment for fixed-b, with a power curve starting near nominal size: unlike fixed-b it is not conservative (see $c = 0$) for "early down" and "late up." For $T = 100$, Q is not yet very powerful, such that the pretest sometimes sides with fixed-b, implying power ranging between the wild bootstrap and fixed-b.

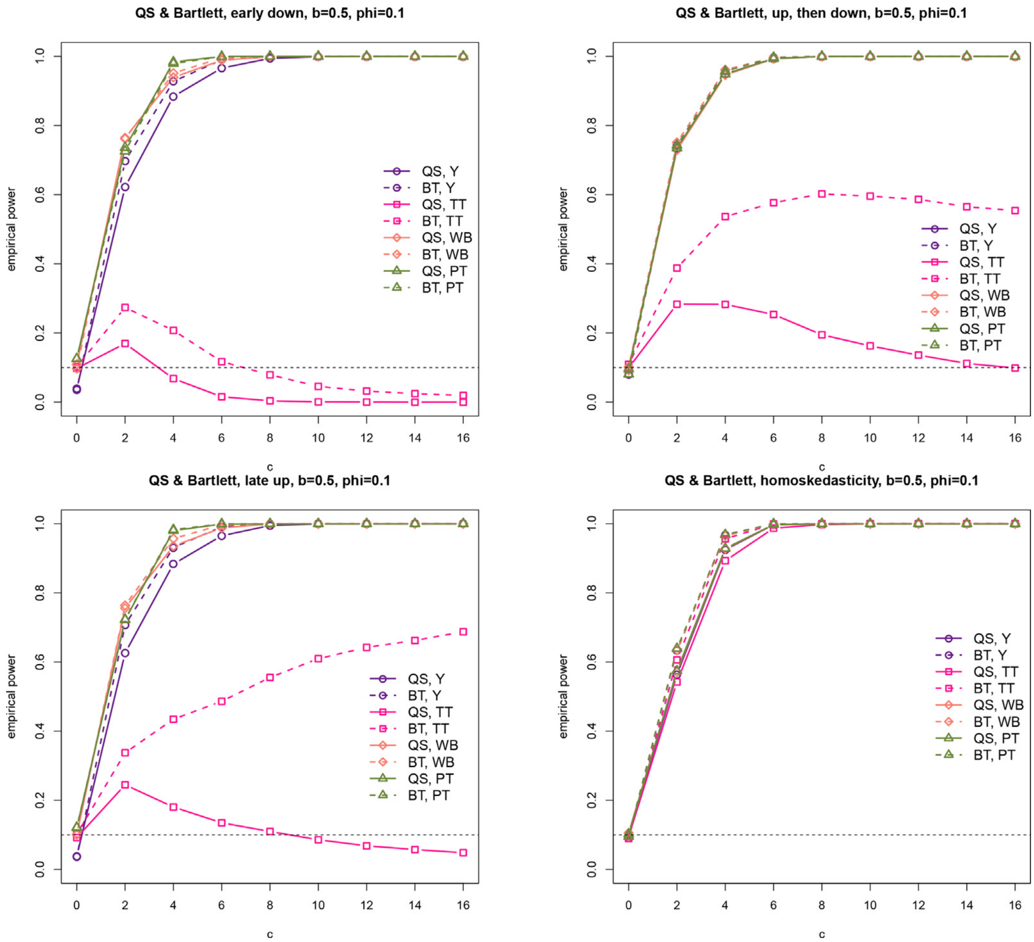

Fig. 3. Power vs. local coefficient c, $\phi = 0.1$, $b = 0.5$, $T = 100$. DGPs 1-4. See also notes to Figure 1.

The result that power varies slightly across types of TVV for fixed-b is due to size distortions. (The normalization employed here implies that the power of the robust tests is not affected by the particular type of TVV.) Figure 4 reveals that power does not vary markedly as a function of b for the different DGPs. Comparing dashed and solid lines suggests that the Bartlett kernel outperforms QS. The difference is only substantial for the time transformation, which is however dominated anyhow.

Our simulations therefore appear to suggest that the wild bootstrap with an intermediate value of b (say, $b = 0.4$) could be a useful choice in practice. (The pretest is a simple alternative for larger T.) Clearly and as usual, such recommendations emanating from simulation evidence—even if fairly robust here—cannot claim general validity and should hence be used with care. More specifically, it is well-known that bandwidth selection in HAR estimation (and elsewhere, see e.g. Andrews, 1991) may have an important effect on both size and power. What is more, an “optimal” bandwidth may be hard to find, given that it depends on, e.g., T, but also unknown population parameters like the spectral density (and hence, e.g., the strength of the dependence in the series) that need to be estimated themselves in turn. See Sun (2014, Sec. 6) for results on testing-optimal choice of b. Unlike our heuristic finding, his results do provide power-maximizing choices for b subject to a size constraint, albeit not in a setup allowing for unconditional heteroskedasticity. In the present setup, different variance profiles add another factor influencing the choice of an optimal bandwidth. Optimal choices under more general loss functions balancing type-I and -II errors is a challenging, but interesting avenue for further research (see also Lazarus et al., 2021).

In view of the simplicity and effectiveness of the wild bootstrap, we focus on this correction and its comparison to the standard Newey/West approach in the regression results to follow.

4.3. Regression results

This subsection presents simulation evidence for linear regression models estimated by OLS, as a leading special case of (1) and (2). The DGP follows that of the previous subsection as concerns the autoregressive error structure with pos-

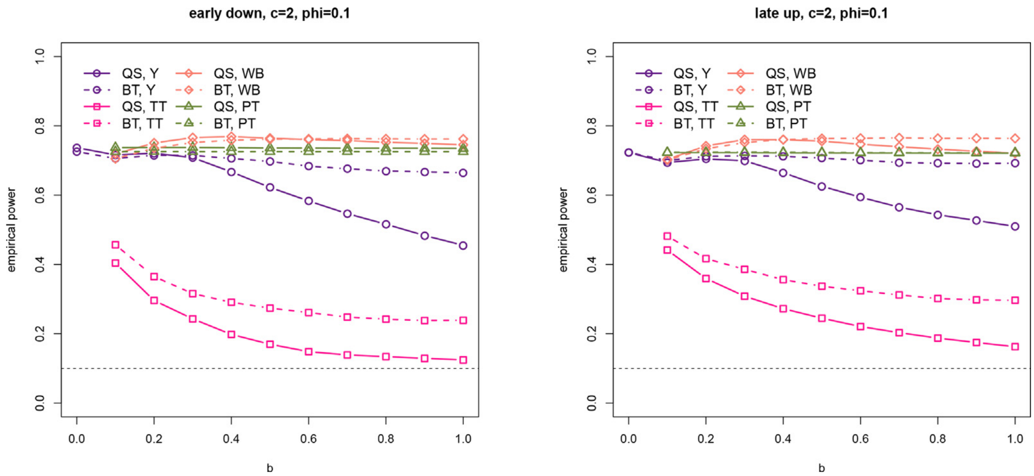

Fig. 4. Power vs. b, $\phi = 0.1$, $c = 2$, $T = 100$. DGPs 1 and 3. See also notes to Figure 3.

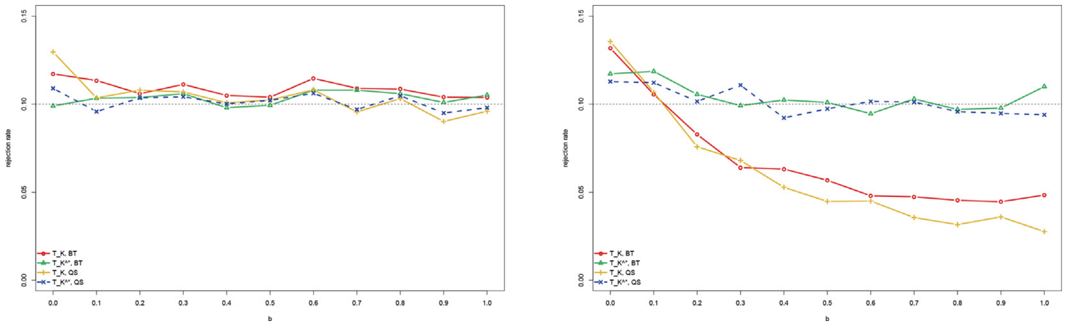

Fig. 5. Size vs. b, $\phi = 0.25$, $T = 100$. DGP 4 (left panel) and DGP 1 (right panel). See also notes to Figure 3.

sible variance breaks of different types. However, y_t is now, additionally to the constant, related to four standard normal regressors, viz.

$$y_t = \boldsymbol{\beta}'\boldsymbol{x}_t + u_t.$$

We test $H_0 : \beta_5 = 0$ against $H_1 : \beta_5 > 0$. We thus focus on testing a single restriction for brevity. Of course, more general linear restrictions may be considered in the usual fashion, cf. the hypothesis formulated above (5) as well the discussion below Proposition 1.

In the size experiments, $\beta_5 = 0$, while we take $\beta_5 = c/\sqrt{T}$, $c \in \{3, 6, \dots, 18\}$, for (local) power. We take $\beta_1 = \dots = \beta_4 = 0$ throughout.

Figure 5 reports a qualitatively similar behavior to that in the simpler location model. Again and as expected, the left panel confirms that both the wild bootstrap and standard fixed-b are effective under homoskedasticity. Small-b again overrejects even for $T = 100$ and a moderate degree of autocorrelation of $\phi = 0.25$ for u_t. The results in the right panel are for the heteroskedastic DGP1, on which we focus here for brevity given that Section 4.2 revealed that the qualitative findings are largely robust to the particular type of variance nonstationarity. Fixed-b loses its pivotalness, underrejecting with empirical sizes below 5% at a nominal level $\alpha = 0.1$ for b large. Overall, there are no systematic differences between the two kernels in terms of size.

Figure 6 provides corresponding power results. The fixed-b results are for an intermediate value of $b = 0.4$. This choice leads to moderate underrejections for standard fixed-b (cf. Figure 5). Unreported results demonstrate that the particular choice of b only has a small effect on power for the wild bootstrap, as in Section 4.2.

The Bartlett kernel, despite its comparable size, once more appears to outperform the QS kernel. The wild bootstrap outperforms the Newey/West and standard fixed-b approaches under heteroskedasticity (right panel), essentially because the bootstrap distribution adapts to the particular variance profile and hence (cf. Figure 5) effectively exhausts nominal size. As a consequence, it also produces fewer type-II errors. Accordingly, the small-b Newey/West approach has spuriously higher power due to being liberal under the null. This reasoning is supported by the left panel of Figure 6, where the homoskedastic DGP 4 leads to valid fixed-b and hence no underrejections under either the null or the alternative. Here, the local alternatives are not normalized by the average variance $\bar{\omega}^2$, to illustrate that such heteroskedasticity may, depending on

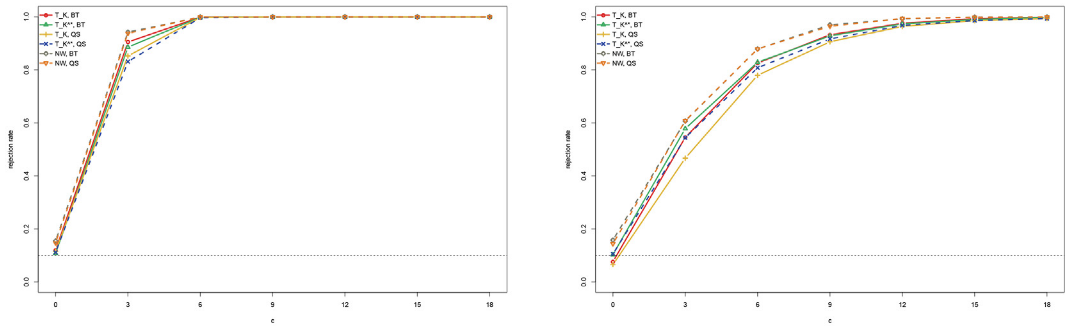

Fig. 6. Power vs. c, $\phi = 0.25$, $T = 50$. DGP 4 (left panel) and DGP 1 (right panel). See also notes to Figure 3.

the specific pattern, have implications for power relative to a benchmark like homoskedasticity. In either case, all approaches are powerful against alternatives sufficiently far away from the null (i.e., c large) already for $T = 50$.

5. Robust tests for the Fama-French five factor asset pricing model

We now apply the wild bootstrap procedure in the context of the Fama and French (2015) five factor asset pricing model. It can be expressed as a time series regression model for portfolio i

$$R_t - R_{Ft} = \beta_1 + \beta_2(R_{Mt} - R_{Ft}) + \beta_3 SMB_t + \beta_4 HML_t + \beta_5 RMW_t + \beta_6 CMA_t + u_t. \tag{8}$$

As R_{Ft} is the risk-free return, $R_t - R_{Ft}$ denotes the period t excess return on the portfolio. The value-weighted market portfolio excess return is labeled as $R_{Mt} - R_{Ft}$. The famous Captial Asset Pricing Model (CAPM) just includes this single factor. Following Fama and French (1993), two additional important factors shall be included: SMB_t measures the return on a diversified portfolio of small versus big stocks and HML_t is the difference between the returns on diversified portfolios of stocks with high and low book-to-market ratios. If the included factors captured all the variation in the expected portfolio returns, the intercept β_1 is zero for all portfolios i.

As a further extension, Fama and French (2015) advocate the inclusion of another two additional factors. While the extension of the CAPM by Fama and French (1993) is widely accepted, there is a controversy about the importance of the two extra factors in their five factor model. These two factors are RMW_t (the difference between returns of stocks with robust and weak profitability) and CMA_t (difference between returns of low and high investment companies). For exact factor definitions, see Fama and French (2015).

We consider tests on the individual significance of the intercept β_1 and the two parameters attached to the additional factors RMW_t and CMA_t (i.e. β_5 and β_6). This allows us to test interesting hypotheses regarding the validity of the five factor model and the importance of the profitability and the investment factors. In a related study, Ray and Savin (2008) consider robust fixed-b tests for the zero intercept null hypothesis in the Fama-French three factor model.

We use monthly data available on Kenneth French's website, available at http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/index.html. The dividend-adjusted value-weighted portfolio returns are sorted on the book-to-market ratio and we use the (i) first (1) and (ii) tenth (10) decile and (iii) also their spread (10-1). The series run from July 1963 to August 2018 with $T = 662$ observations in total. According to the simulation from Section 4, we set $b = 0.4$ and focus on the Bartlett kernel. The nominal significance level for the two-sided tests is set to one percent and the number of bootstrap replications equals 1,999.

Table 1 reports the main findings. The top left corner provides OLS estimates for the parameters β_1, β_5 and β_6 for the three different decile portfolios ((1), (10) and (10-1)). For completeness, we provide naive t-statistics (t_{β_1}, t_{β_5} and t_{β_6}) based on the white noise assumptions for the errors of the factor pricing regressions. In addition, Newey-West-based t-statistics with small-b are reported ($t_{\beta_1,NW}$, $t_{\beta_5,NW}$ and $t_{\beta_6,NW}$), next to the fixed-b statistics $t_{\beta_1,FB}$, $t_{\beta_5,FB}$ and $t_{\beta_6,FB}$. Rejections based

Table 1
Main estimation and testing results

OLS	β_1	β_5	β_6	t_{β_1}	t_{β_5}	t_{β_6}
(1)	-0.04	0.17	-0.19	-0.86	6.88	-5.49
(10)	-0.18	0.18	0.17	-3.43	7.09	4.72
(10-1)	-0.28	0.22	0.23	-5.30	8.57	6.20
Robust	$t_{\beta_1,NW}$	$t_{\beta_5,NW}$	$t_{\beta_6,NW}$	$t_{\beta_1,FB}$	$t_{\beta_5,FB}$	$t_{\beta_6,FB}$
(1)	-0.78	3.03	-3.53	-0.80	**3.10**	-3.62
(10)	-3.29	3.29	3.21	-3.94	**3.95**	3.85
(10-1)	-5.09	3.99	4.22	-3.25	**2.55**	2.69

on *asymptotic* critical values are indicated by a an underlined entry, while a rejection obtained through *wild bootstrap* critical values is signified by a boldfaced entry. Test results for the other three factors are omitted to save space. They are conclusive and suggest the relevance of the market factor, as well as SMB_t and HML_t. We test for unconditional heteroskedasticity in the regressors by applying the Q statistic of Deng and Perron (2008). The null of homoskedasticity is not rejected for any of the regressors.

The obtained results suggest the following: naive OLS and standard HAC inference (via Newey-West standard errors) lead to the same test decisions. These indicate the importance of both *RMW* and *CMA* for all portfolios. Moreover, in two cases even a significant intercept is found. But, when using asymptotic fixed-b inference, we do not find even a single rejection. This supports the three factor model and suggests the irrelevance of the two additional factors. However, when finally considering wild bootstrap inference, another clear result emerges: the *RMW* factor is found to be significant in all portfolios, while the opposite holds for the *CMA* factor. Given our theoretical results and insights from the simulation exercises, inference based on the wild bootstrap is most reliable. We, therefore, find evidence supporting the importance of the profitability factor, but not for the investment factor.

6. Concluding remarks

Fixed-b asymptotics are tremendously useful for more accurate inference for serially correlated data. Many important macroeconomic and financial time series are, however, also subject to time-varying volatility (TVV) such as variance breaks. We show that the standard fixed-b approach no longer yields pivotal tests under TVV and quantify the resulting distortions.

Based on wild bootstrap schemes (Cavaliere and Taylor, 2008a), on time transformations (Cavaliere and Taylor, 2008b) or on a pretest procedure, we provide corrections that restore size control of fixed-b methods even under TVV in a fairly general GMM setup. Simulations illustrate the useful size and power properties in particular of the wild bootstrap. The behavior of the pretest procedure evidently hinges on the power of the pretest to detect TVV. It is not reliable for small samples in which the pretest's power is too low. The time transformation may produce zero local power. The evidence provided here suggests quite plausibly that the multivariate wild bootstrap would provide a robust version of fixed-b tests in a composite-hypothesis situation as well. We expect the wild bootstrap to perform well for replicating distributions of statistics relying on self-normalization.

An empirical application to the Fama-French five factor asset pricing model reveals substantial differences in testing outcomes between asymptotic fixed-b and wild bootstrap inference. Our findings suggest that the additional factor on profitability included in the five factor model contributes significantly to the explanation of portfolio returns once accounting for heteroskedasticity via the wild bootstrap. However, the additional investment factor is found to be irrelevant.

Declaration of Competing Interest

We declare not to have any conflicts of interest. Funding information is provided and available from public organizations only.

Acknowledgments

The authors are grateful to two anonymous referees and an associate editor whose suggestions helped to substantially improve the paper. The first two authors gratefully acknowledge support of the German Research Foundation (DFG) through the projects DE 1617/4-2 and HA 6766/2-2. Hanck acknowledges support from MERCUR, project digital data in economics.

Appendix A. Proofs

Proof of Proposition 1

Let us first analyze the sampling properties of the GMM estimators. We have, with obvious notation $\mathbf{S}_{u\boldsymbol{z}} = \sum_{t=1}^{T} \boldsymbol{z}_t u_t$,

$$\hat{\boldsymbol{\beta}} - \boldsymbol{\beta} = \mathbf{M}^{-1}\mathbf{S}_{\boldsymbol{xz}}\mathbf{W}_T\mathbf{S}_{u\boldsymbol{z}}.$$

Now, $\boldsymbol{x}_t\boldsymbol{z}_t$ is itself strictly stationary and α-mixing, such that $T^{-1}\boldsymbol{S}_{\boldsymbol{xz}} \xrightarrow{p} \boldsymbol{\Upsilon}$ thanks to a LLN for α-mixing strictly stationary sequences (see Davidson, 1994, Section 20.6). The weak convergence $\frac{1}{\sqrt{T}}\sum_{t=1}^{[sT]} \boldsymbol{z}_t u_t \Rightarrow \boldsymbol{\Omega}^{1/2}\boldsymbol{W}_L(\eta(s))$ and the continuous mapping theorem [CMT] lead to

$$\sqrt{T}\left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}\right) \Rightarrow \left(\boldsymbol{\Upsilon}\mathbf{W}\boldsymbol{\Upsilon}'\right)^{-1}\boldsymbol{\Upsilon}\mathbf{W}\boldsymbol{\Omega}^{1/2}\boldsymbol{W}_L(\eta(1))$$

where $\eta(1) = 1$. We also have, along the same lines, $\sqrt{T}\left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0\right) \Rightarrow \left(\boldsymbol{\Upsilon}\mathbf{W}\boldsymbol{\Upsilon}'\right)^{-1}\boldsymbol{\Upsilon}\mathbf{W}\boldsymbol{\Omega}^{1/2}\boldsymbol{W}_L(1) + \boldsymbol{c}$.

Let $\mathcal{Z}_m := \frac{1}{\sqrt{T}} \sum_{t=1}^{m} \left(\boldsymbol{z}_t \hat{u}_t - \overline{\boldsymbol{z}\hat{u}}\right)$. Note that the arguments in the proof of Theorem 2 in Kiefer and Vogelsang (2005) can be used without further modification to conclude that

$$\hat{\boldsymbol{\Omega}} = -\frac{1}{T^2} \sum_{i=1}^{T-1} \sum_{j=1}^{T-1} \frac{T^2}{B^2} k''\left(\frac{i-j}{B}\right) \mathcal{Z}_i \mathcal{Z}_j' + o_p(1)$$

for kernels with smooth derivatives, or

$$\hat{\boldsymbol{\Omega}} = \frac{2}{bT} \sum_{i=1}^{T} \mathcal{Z}_i \mathcal{Z}_i' - \frac{1}{bT} \sum_{i=1}^{[(1-b)T]} \left(\mathcal{Z}_{i+[bT]} \mathcal{Z}_i' + \mathcal{Z}_i \mathcal{Z}_{i+[bT]}'\right) + o_p(1)$$

for the Bartlett kernel. These further simplify since $\overline{\boldsymbol{z}\hat{u}} = \mathbf{0}$ are the sample moment restrictions. Moreover,

$$\frac{1}{T}\mathbf{S}_{\boldsymbol{xz}}\mathbf{W}_T \frac{1}{\sqrt{T}} \sum_{t=1}^{[sT]} \boldsymbol{z}_t \hat{u}_t = \frac{1}{T}\mathbf{S}_{\boldsymbol{xz}}\mathbf{W}_T \frac{1}{\sqrt{T}} \sum_{t=1}^{[sT]} \boldsymbol{z}_t u_t - \frac{1}{T}\mathbf{S}_{\boldsymbol{xz}}\mathbf{W}_T \frac{1}{T} \sum_{t=1}^{[sT]} \boldsymbol{z}_t \boldsymbol{x}_t' \sqrt{T}\left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}\right).$$

Notice that $\frac{1}{T}\sum_{t=1}^{[sT]} \boldsymbol{z}_t \boldsymbol{x}_t' \xrightarrow{p} s\boldsymbol{\Upsilon}'$ uniformly (it is straightforward to establish stochastic equicontinuity of $\frac{1}{T}\sum_{t=1}^{[sT]} \boldsymbol{z}_t \boldsymbol{x}_t'$, so Theorem 21.9 in Davidson, 1994, applies), and we obtain

$$\frac{1}{T}\mathbf{S}_{\boldsymbol{xz}}\mathbf{W}_T \frac{1}{\sqrt{T}} \sum_{t=1}^{[sT]} \boldsymbol{z}_t \hat{u}_t \Rightarrow \boldsymbol{\Upsilon}\mathbf{W}\boldsymbol{\Omega}^{1/2}\boldsymbol{W}_L(\eta(s)) - s\boldsymbol{\Upsilon}\mathbf{W}\boldsymbol{\Omega}^{1/2}\boldsymbol{W}_L(1) \tag{9}$$

and therefore

$$\frac{1}{T^2}\mathbf{S}_{\boldsymbol{xz}}\mathbf{W}_T\hat{\boldsymbol{\Omega}}\mathbf{W}_T\mathbf{S}_{\boldsymbol{xz}}' \Rightarrow \boldsymbol{\Theta}_{k,b}(\boldsymbol{\Upsilon}\mathbf{W}\boldsymbol{\Omega}^{1/2}\boldsymbol{W}_L(\eta(\cdot))).$$

Summing up,

$$\mathcal{T}_K \Rightarrow \left(\boldsymbol{\Upsilon}\mathbf{W}\boldsymbol{\Omega}^{1/2}\boldsymbol{W}_L(1) + \left(\boldsymbol{\Upsilon}\mathbf{W}\boldsymbol{\Upsilon}'\right)\boldsymbol{c}\right)' \boldsymbol{\Theta}_{k,b}^{-1}\left(\boldsymbol{\Upsilon}\mathbf{W}\boldsymbol{\Omega}^{1/2}\boldsymbol{W}_L(\eta(\cdot))\right)\left(\boldsymbol{\Upsilon}\mathbf{W}\boldsymbol{\Omega}^{1/2}\boldsymbol{W}_L(1) + \left(\boldsymbol{\Upsilon}\mathbf{W}\boldsymbol{\Upsilon}'\right)\boldsymbol{c}\right).$$

Premultiply $\boldsymbol{\Theta}_{k,b}$ by $\mathbf{I} = \left(\boldsymbol{\Upsilon}\mathbf{W}\boldsymbol{\Omega}\mathbf{W}\boldsymbol{\Upsilon}'\right)^{1/2}\left(\boldsymbol{\Upsilon}\mathbf{W}\boldsymbol{\Omega}\mathbf{W}\boldsymbol{\Upsilon}'\right)^{-1/2}$ and postmultiply by the respective transpose. Then, since $\boldsymbol{\Upsilon}\mathbf{W}\boldsymbol{\Omega}^{1/2}\boldsymbol{W}_L(\eta(s))$ is itself a time-transformed Brownian motion with covariance matrix $\boldsymbol{\Upsilon}\mathbf{W}\boldsymbol{\Omega}\mathbf{W}\boldsymbol{\Upsilon}'$ (though of lower dimension K), it follows that $\left(\boldsymbol{\Upsilon}\mathbf{W}\boldsymbol{\Omega}\mathbf{W}\boldsymbol{\Upsilon}'\right)^{-1/2}\boldsymbol{\Upsilon}\mathbf{W}\boldsymbol{\Omega}^{1/2}\boldsymbol{W}_L(\eta(s))$ is the desired K-dimensional vector of independent time-transformed Wiener processes, $\boldsymbol{W}_K(\eta(s))$, and the result follows for the no prewhitening case.

Moving on to the case with prewhitening, use the Phillips-Solo decomposition and write

$$\tilde{\boldsymbol{v}}_t = \left(\mathbf{I} - \sum_{j=1}^{p}\tilde{\mathbf{A}}_j\right)\boldsymbol{\psi}_t + \sum_{j=1}^{p}\left(\sum_{k=j}^{p}\tilde{\mathbf{A}}_k\right)\Delta\boldsymbol{\psi}_{t-j+1},$$

such that

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{[sT]}\tilde{\boldsymbol{v}}_t = \left(\mathbf{I} - \sum_{j=1}^{p}\tilde{\mathbf{A}}_j\right)\frac{1}{\sqrt{T}}\sum_{t=1}^{[sT]}\boldsymbol{\psi}_t + \frac{1}{\sqrt{T}}\sum_{j=1}^{p}\left(\sum_{k=j}^{p}\tilde{\mathbf{A}}_k\right)\boldsymbol{\psi}_{[sT]-j+1} - \frac{1}{\sqrt{T}}\sum_{j=1}^{p}\left(\sum_{k=j}^{p}\tilde{\mathbf{A}}_k\right)\boldsymbol{\psi}_{-j+1}.$$

Now, $\sup_{s\in[0,1]}\left\|\boldsymbol{\psi}_{[sT]-j}\right\| \leq \sup_t \|\boldsymbol{z}_t u_t\| + \left\|\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}\right\| \sup_t \left\|\boldsymbol{z}_t \boldsymbol{x}_t'\right\| = o_p\left(T^{1/2}\right)$ such that

$$\hat{\boldsymbol{\Omega}}_v = \left(\mathbf{I} - \sum_{j=1}^{p}\tilde{\mathbf{A}}_j\right)\hat{\boldsymbol{\Omega}}\left(\mathbf{I} - \sum_{j=1}^{p}\tilde{\mathbf{A}}_j\right)' + o_p(1).$$

The behavior of the OLS estimators $\tilde{\mathbf{A}}_j$ in $\tilde{\boldsymbol{v}}_t = \boldsymbol{\psi}_t - \sum_{j=1}^{p}\tilde{\mathbf{A}}_j\boldsymbol{\psi}_{t-j}$ can be shown to be such that $\mathbf{I} - \sum_{j=1}^{p}\tilde{\mathbf{A}}_j$ is invertible in the limit, such that pre- and postmultiplying $\hat{\boldsymbol{\Omega}}_v$ with $\left(\mathbf{I} - \sum_{j=1}^{p}\tilde{\mathbf{A}}_j\right)^{-1}$ leads to the same asymptotic behavior of $\hat{\boldsymbol{\Omega}}_{pw}$ estimate as for $\hat{\boldsymbol{\Omega}}$ above and the result follows.

Proof of Proposition 2

Begin by noting that the bootstrapped $\boldsymbol{z}_t u_t^*$ satisfy the same moment and serial dependence restrictions as $\boldsymbol{z}_t u_t$ unconditionally (conditionally on the sample they are even independent). The arguments in the proof of Theorem 2 in Kiefer and Vogelsang (2005) lead for twice continuously differentiable kernels to

$$\hat{\boldsymbol{\Omega}}^* = -\frac{1}{T^2}\sum_{i=1}^{T-1}\sum_{j=1}^{T-1}\frac{T^2}{B^2}k''\left(\frac{i-j}{B}\right)\frac{1}{\sqrt{T}}\sum_{t=1}^{i}\left(\boldsymbol{z}_t u_t^* - \overline{\boldsymbol{z}u^*}\right)\frac{1}{\sqrt{P}}\sum_{t=1}^{j}\left(\boldsymbol{z}_t u_t^* - \overline{\boldsymbol{z}u^*}\right)' + o_p^*(1)$$

where $o_p^*(1)$ denotes asymptotic negligibility under the bootstrap measure, conditional on the data. The analogous result for the Bartlett kernel holds as well. Moreover, we have with $\boldsymbol{S}_{uz}^* = \sum_{t=1}^T \boldsymbol{z}_t u_t^*$ that

$$\hat{\boldsymbol{\beta}}^* - \boldsymbol{\beta}_0 = \left(\mathbf{S}_{xz}\mathbf{W}_T\mathbf{S}_{xz}'\right)^{-1}\mathbf{S}_{xz}\mathbf{W}_T\boldsymbol{S}_{uz}^*.$$

To guarantee size control in the limit, it therefore suffices to show that the bootstrap normalized partial sums $\boldsymbol{S}_T^*(s) := \frac{1}{\sqrt{T}}\sum_{t=1}^{[sT]}\boldsymbol{z}_t u_t^*$ converge weakly in probability to $\boldsymbol{\Sigma}^{1/2}\boldsymbol{W}_L(\eta(s))$ for some covariance matrix $\boldsymbol{\Sigma}$, since $\boldsymbol{\Sigma}$ would cancel out in the expression of $\mathcal{T}_K^*$ like in the proof of Proposition 1, and, therefore, the limiting null distribution $\mathcal{D}_{h,k,b,\mathbf{0}}$ would follow.

We examine first the case of Gaussian bootstrap variables r_t^* and bootstrapping $\hat{u}_t$ directly. Let $\boldsymbol{S}_T^*(s)$ denote the normalized partial sums of the cross-product of instrument and disturbance,

$$\boldsymbol{S}_T^*(s) = \frac{1}{\sqrt{T}}\sum_{t=1}^{[sT]}\boldsymbol{z}_t\left(u_t - \left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}\right)'\boldsymbol{x}_t\right)r_t^*.$$

Note that, conditional on the sample $\{\boldsymbol{x}_t, \boldsymbol{z}_t, y_t\}$, $t = 1, \ldots, T$, $\boldsymbol{S}_T^*(s)$ is a Gaussian process with independent increments. Its covariance kernel is given by

$$\begin{aligned}\mathrm{Cov}\left(\boldsymbol{S}_T^*(s), \boldsymbol{S}_T^*(r)\right) &= \frac{1}{T}\sum_{t=1}^{[\min\{s,r\}T]}\boldsymbol{z}_t\boldsymbol{z}_t'\left(u_t - \left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}\right)'\boldsymbol{x}_t\right)^2 \mathrm{E}\left((r_t^*)^2\right)\\ &= \frac{1}{T}\sum_{t=1}^{[\min\{s,r\}T]}\boldsymbol{z}_t\boldsymbol{z}_t'\left(u_t - \left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}\right)'\boldsymbol{x}_t\right)^2.\end{aligned}$$

Then, following the proof of Lemma A.5 in Cavaliere et al. (2010), it suffices to establish that

$$\frac{1}{T}\sum_{j=1}^{[sT]}\boldsymbol{z}_t\boldsymbol{z}_t'\left(u_t - \left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}\right)'\boldsymbol{x}_t\right)^2 \xrightarrow{p} \boldsymbol{\Sigma}\eta(s) \tag{10}$$

uniformly, i.e., that the wild bootstrap correctly replicates the variance profile $\eta(s)$ in the limit. Assumption 1(ii) implies the existence of the covariance matrix of $\boldsymbol{z}_t u_t$ (which we denote by $\boldsymbol{\Sigma}$), such that pointwise convergence of the l.h.s. follows via a Law of Large Numbers for strong mixing processes after exploiting that $\sup_{t\in\{1,\ldots,T\}}||\boldsymbol{z}_t\boldsymbol{z}_t'\boldsymbol{x}_t\boldsymbol{x}_t'|| = o_p(T)$ and $\hat{\boldsymbol{\beta}} - \boldsymbol{\beta} = O_p(T^{-1/2})$ under our assumptions.

To establish uniformity of (10), build quadratic forms to conclude that

$$\boldsymbol{a}'\left(\frac{1}{T}\sum_{j=1}^{[sT]}\boldsymbol{z}_t\boldsymbol{z}_t'\left(u_t - \left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}\right)'\boldsymbol{x}_t\right)^2\right)\boldsymbol{a} \xrightarrow{p} \eta(s)\boldsymbol{a}'\boldsymbol{\Sigma}\boldsymbol{a}.$$

However, given that the increments $\boldsymbol{a}'\left(\boldsymbol{z}_t\boldsymbol{z}_t'\left(u_t - \left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}\right)'\boldsymbol{x}_t\right)^2\right)\boldsymbol{a}$ are nonnegative $\forall t$ by construction, we may conclude that convergence of the quadratic forms is uniform for any choice of $\boldsymbol{a}$. Therefore, convergence in (10) must itself be uniform.

Finally, should r_t^* follow the Mammen distribution, say, $\boldsymbol{S}_T^*(s)$ is not Gaussian, but weak convergence to a Gaussian process (conditional on the sample) holds given that r_t^* are iid with finite moments of any order (see e.g. Davidson, 1994, Corollary 29.14), and the proof follows along the same lines.

We then examine the case of bootstrapping residuals from an autoregression of order p of $\hat{u}_t$, $\tilde{v}_t = \hat{u}_t - \sum_{j=1}^p \tilde{a}_j\hat{u}_{t-j}$. It is tedious, yet straightforward, to show that $\tilde{a}_j$ are $\sqrt{T}$ consistent for the coefficients of the best linear projection of ε_t on p lags, say a_j. See Phillips and Xu (2006) for details on autoregressions with unconditional heteroskedasticity. Let further $v_t := h_t\left(\varepsilon_t - \sum_{j=1}^p a_j\varepsilon_{t-j}\right)$ such that v_t satisfies indeed the same assumptions as u_t up to the (for this step irrelevant) different LRV. Then,

$$\boldsymbol{S}_T^*(s) = \frac{1}{\sqrt{T}}\sum_{t=1}^{[sT]}\boldsymbol{z}_t v_t r_t^* + \frac{1}{\sqrt{T}}\sum_{t=1}^{[sT]}\boldsymbol{z}_t\left(\hat{v}_t - v_t\right)r_t^*$$

and the result follows along the lines of the case without prewhitening if the second summand on the r.h.s. vanishes uniformly in $s \in [0, 1]$. Given that r_t^* are serially independent and independent of u_t and $\boldsymbol{x}_t$, this is implied by

$$\sup_{s\in[0,1]}\frac{1}{T}\sum_{t=1}^{[sT]}\left(\hat{v}_t - v_t\right)^2 \xrightarrow{p} 0$$

which is in turn implied by $\frac{1}{T}\sum_{t=1}^T\left|\hat{v}_t - v_t\right| \xrightarrow{p} 0$. Now, at all continuity points of h,

$$\tilde{v}_{[sT]} - v_{[sT]} = \hat{u}_{[sT]} - u_{[sT]} - \sum_{j=1}^p\left(\tilde{a}_j\hat{u}_{[sT]-j} - a_j\hat{u}_{[sT]-j}\right) - \sum_{j=1}^p\left(a_j\hat{u}_{[sT]-j} - a_j h_{[sT]}\varepsilon_{t-j}\right)$$

$$= -\left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}\right)' \boldsymbol{x}_{[sT]} - \sum_{j=1}^{p} \left(\tilde{a}_j - a_j\right)\left(u_{[sT]-j} - \left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}\right)' \boldsymbol{x}_{[sT]-j}\right)$$
$$- \sum_{j=1}^{p} a_j \left(-\left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}\right)' \boldsymbol{x}_{[sT]-j} - \left(h_{[sT]-j} - h_{[sT]}\right)\varepsilon_{t-j}\right)$$

where $\sup_{t\in\{1,\dots,T\}} |\varepsilon_t| = \sup_{t\in\{1,\dots,T\}} \|\boldsymbol{x}_t\| = o_p\left(T^{1/4}\right)$ thanks to the uniform $L_{4+2\delta}$ boundedness of ε_t and $\boldsymbol{x}_t$. Moreover, $h_{[sT]} - h_{[sT]-1} = O\left(T^{-1}\right)$ uniformly in s thanks to the piecewise Lipschitz condition on h. Hence $\sup_{s\in[0,1]\setminus D} \left|\hat{\nu}_{[sT]} - \nu_{[sT]}\right| \xrightarrow{p} 0$ with D the set of continuity points of h. At the discontinuities of $h(\cdot)$, we have that $h_{[sT]} - h_{[sT]-j} = O(1)$, but there are a finite number of discontinuities only, so their cumulated effect on $\frac{1}{T}\sum_{t=1}^{T} \left|\hat{\nu}_t - \nu_t\right|$ remains $O_p(T^{-1})$ and the result follows. Finally, notice that recoloring the bootstrapped u_t^* only changes the scale of the weak limit in probability of $\boldsymbol{S}_T^*(s)$, but not the variance profile.

Proof of Proposition 3

We start by deriving the result under the null. By applying the arguments of Cavaliere and Taylor (2008b, proof of Theorem 1) elementwise, it follows for the reverse time transformed $\tilde{\boldsymbol{\zeta}}_t$ that $\frac{1}{\sqrt{T}}\sum_{t=1}^{[sT]} \Delta\tilde{\boldsymbol{\zeta}}_t \Rightarrow \Omega^{1/2}\boldsymbol{W}_L(s)$ elementwise. Given independence of the elements of $\boldsymbol{W}_L(s)$, this may be extended to joint convergence. The result then follows along the lines of the proof of Proposition 1.

Next, consider local alternatives. Begin by writing

$$\boldsymbol{\zeta}_t = \sum_{j=1}^{t} \boldsymbol{z}_t\left(y_j - \boldsymbol{\beta}_0'\boldsymbol{x}_t\right) = \sum_{j=1}^{t} \boldsymbol{z}_t h_j \varepsilon_j + \frac{1}{\sqrt{T}}\sum_{j=1}^{t} \boldsymbol{z}_t \boldsymbol{x}_t' \boldsymbol{c}$$

and note that the variance profile estimate is invariant to β. Now, the time transformation leads to $\tilde{\boldsymbol{\zeta}}_t = \boldsymbol{\zeta}_{[Tg(t/T)]}$ (with g the inverse of $\hat{\eta}$), implying that

$$\sum_{j=1}^{t} \Delta\tilde{\boldsymbol{\zeta}}_j = \tilde{\boldsymbol{\zeta}}_t = \sum_{j=1}^{[Tg(t/T)]} \boldsymbol{z}_t h_j \varepsilon_j + \sqrt{T}\left(\frac{1}{T}\sum_{j=1}^{[Tg(t/T)]} \boldsymbol{x}_t \boldsymbol{z}_t'\right)' \boldsymbol{c}.$$

Since $\hat{\eta}$ converges uniformly to the variance profile η, its inverse converges weakly to the inverse η^{-1} of the variance profile, which is, like η, monotonic and continuous. Also, it is not difficult to show that $\frac{1}{T}\sum_{j=1}^{[Tg(t/T)]} \boldsymbol{x}_t \boldsymbol{z}_t' \xrightarrow{p} \boldsymbol{\Upsilon}\eta^{-1}(s)$ uniformly along the lines of the proof of Proposition 1. Hence, using the CMT, we obtain, as required for the result, that

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{[sT]} \Delta\tilde{\zeta}_j \Rightarrow \boldsymbol{\Omega}^{1/2}\boldsymbol{W}_L(s) + \eta^{-1}(s)\boldsymbol{\Upsilon}'\boldsymbol{c}.$$

Appendix B. Alternative time-transformation approaches

We give counter-examples for the case of an intercept only to save space. In this case, $\hat{\beta}$ is simply the sample average $\bar{y}$.

Standardizing the observations y_t by some (nonparametric) estimate of their standard deviation, $\hat{h}_t$, turns out to be of limited usefulness, given that in the WLS transformed model we would not have a regression with a constant anymore, but rather

$$\frac{y_t}{h_t} = \frac{\beta}{h_t} + \varepsilon_t.$$

While the error term is strictly stationary, time-varying volatility is re-introduced through the back door, given that the LS estimator for β in the transformed model is given as

$$\hat{\beta} = \beta + \frac{\sum_{t=1}^{T} \frac{\varepsilon_t}{h_t}}{\sum_{t=1}^{T} \frac{1}{h_t^2}}$$

and the term relevant for the limiting distribution of the test statistic is $T^{-1/2}\sum_{t=1}^{T} \frac{\varepsilon_t}{h_t}$. This also illustrates how time-varying volatility may be induced in regression models by nonstationary regressors or instruments. The fact that the true standard deviation was used for the WLS transformation does not affect the argument. One could alternatively standardize the series y_t under the null hypothesis, i.e. work with

$$\tilde{y}_t^w = \frac{y_t - \beta_0}{h_t}$$

and test the equivalent null hypothesis that $\mathrm{E}\left(\tilde{y}_t^w\right) = 0$. Denote the corresponding test statistic $\tilde{\mathcal{T}}^w$. The difference to testing the mean of y_t (either ignoring time-varying volatility or via WLS) is that $\tilde{y}_t^w$ is now strictly stationary under the null hypothesis and fixed-b inference would be applicable. The disadvantages of this approach appear under (local) alternatives, just like those of the time transformation, as illustrated in the following

Corollary 1. *Let $\bar{h}(s) = \int_0^s \frac{1}{h(r)}\mathrm{d}r$. Under the assumptions of Proposition 3 we have that*

$$\tilde{\mathcal{T}}^w \xrightarrow{d} \frac{W(1) + c\bar{h}(1)}{\sqrt{\Theta_{k,b}\left(W(s) + c\bar{h}(s)\right)}}.$$

Proof. Under $\beta = \beta_0 + c/\sqrt{T}$, we have that

$$\tilde{y}_t^w = \varepsilon_t + \frac{1}{h_t}\frac{c}{\sqrt{T}}.$$

This implies for the partial sums of $\tilde{y}_t^w$ that

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{[sT]}\tilde{y}_t^w = \frac{1}{\sqrt{T}}\sum_{t=1}^{[sT]}\varepsilon_t + \frac{c}{T}\sum_{t=1}^{[sT]}\frac{1}{h_t}.$$

With $h(\cdot)$ being piecewise Lipschitz, bounded and bounded away from zero, $1/h$ is itself piecewise Lipschitz, bounded and bounded away from zero, so we ultimately have

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{[sT]}\tilde{y}_t^w \Rightarrow W(s) + c\int_0^s \frac{1}{h(r)}\mathrm{d}r := W(s) + c\bar{h}(s)$$

as required for the result.

Note that $\bar{h}(s) = s$ only when h is constant, such that, in general, one would obtain a limiting distribution under local alternatives having the exact same properties as $c \to \pm\infty$ as the one in Proposition 3.

If using the estimates $\hat{\boldsymbol{\beta}}$ rather than the null values $\boldsymbol{\beta}_0$ in Step 1 of the time transformation approach, we obtain a different limiting distribution. Concretely, the behavior of $\hat{\boldsymbol{\Omega}}$ relies on the behavior of the time-transformed partial sums of $\boldsymbol{z}_t u_t^0$ (which, under the null, have a time-transformed Brownian motion as weak limit). These partial sums will be tied-down by using residuals; see (9). The reverse time transformation does not lead to a Brownian bridge as required, but rather to a functional involving $\eta^{-1}(s)$, so the resulting fixed-b distribution still depends on h under the null. □

Appendix C. Additional simulations on the asymptotic distributions

The QQ plots in Figure 7 demonstrate substantial differences for $b = \{0.1, 0.5, 0.9\}$, different patterns h_t and the Bartlett kernel (we simulate $\mathcal{D}_{h,k,b,c}$ for $c = 0$ with 50,000 replications and $T = 1,000$). DGP 1 has an early downward break in volatility, while DGP 2 exhibits a double break. (Section 4 gives details on the DGPs.) The larger b, the more pronounced are the discrepancies. They are most pronounced in the tails of the distributions, which matter for testing.

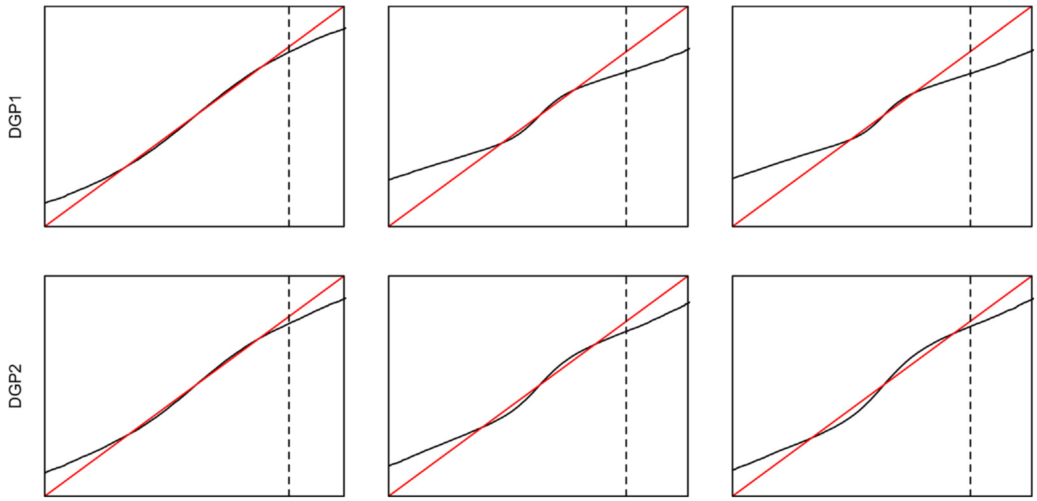

Fig. 7. QQ plots to compare $\mathcal{D}_{k,b}$ (x-axis) to the distributions $\mathcal{D}_{h,k,b}$ for $c = 0$ (y-axis) for different h_t and b (left: $b = 0.1$, center: $b = 0.5$, right: $b = 0.9$) and the Bartlett kernel. DGP1: early downward break in volatility; DGP2: double break in volatility. The dashed line is the 95% critical value from the $\mathcal{D}_{k,b}$ distribution.

Appendix D. Additional simulation results

Figures 8–11 provide additional size and power results for selected representative scenarios. Figure 9 omits results for the time transformation in view of the substantial upward size distortions.

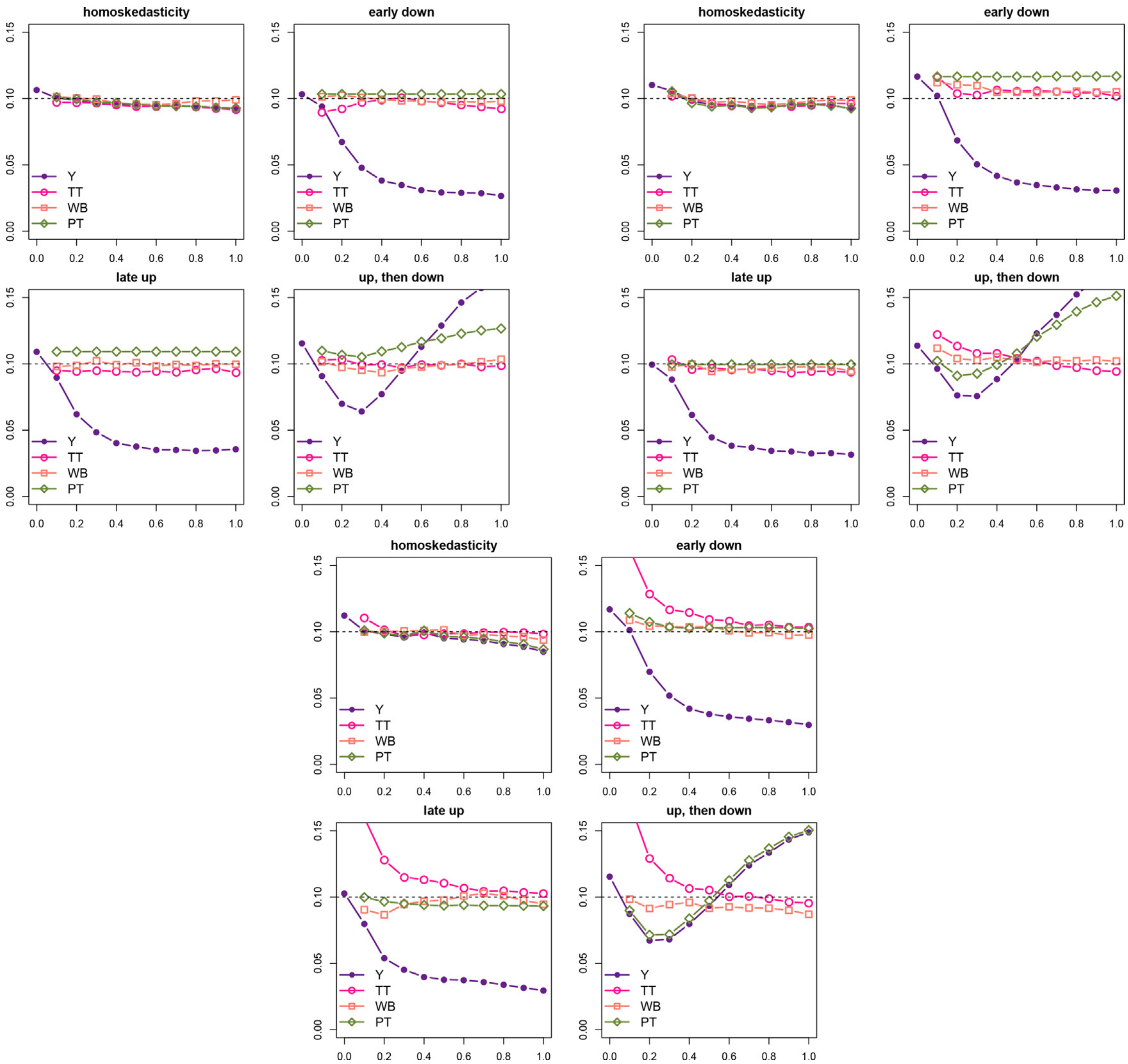

Fig. 8. Size, $T = 250$. Top left block: $\phi = 0.1$, top right: $\phi = 0.5$, bottom: $\phi = 0.85$. Y denotes the standard fixed-b approach, TT the time transformation statistic (cf. Sec. 3.2), WB the wild bootstrap approach (cf. Prop. 2) and PT the pretest (cf. Section 3.3). Rejection frequencies are given on the y-axis, while b-values are given on the x-axis.

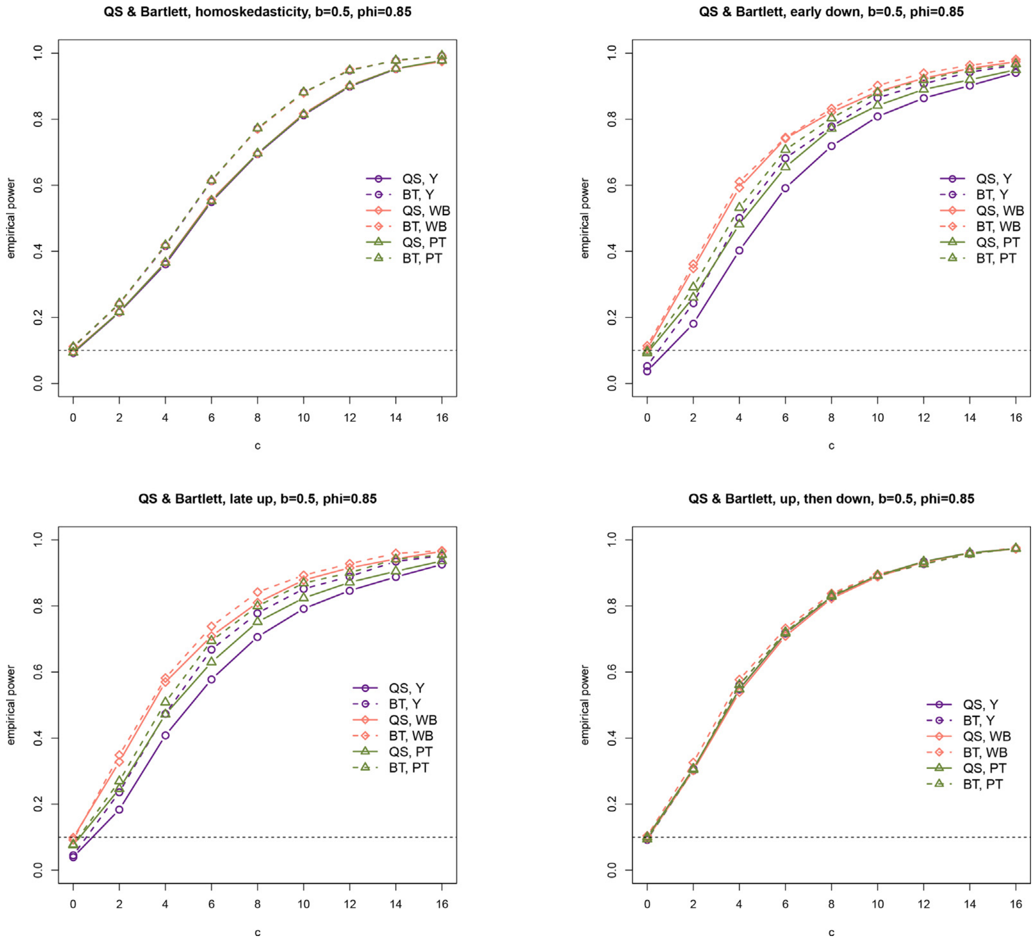

Fig. 9. Power vs. c, $\phi = 0.85$, $b = 0.5$, $T = 100$. See notes to Figure 3.

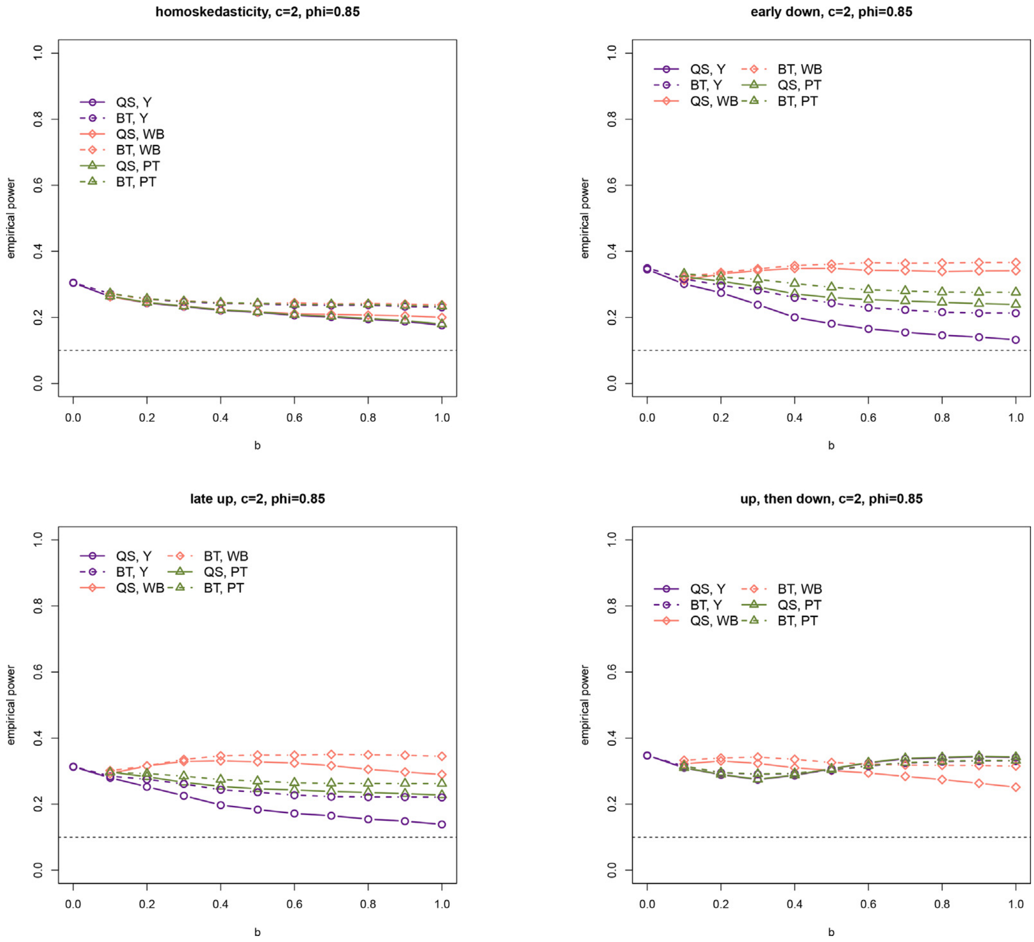

Fig. 10. Power vs. b, $\phi = 0.85$, $c = 2$, $T = 100$. See notes to Figure 4.

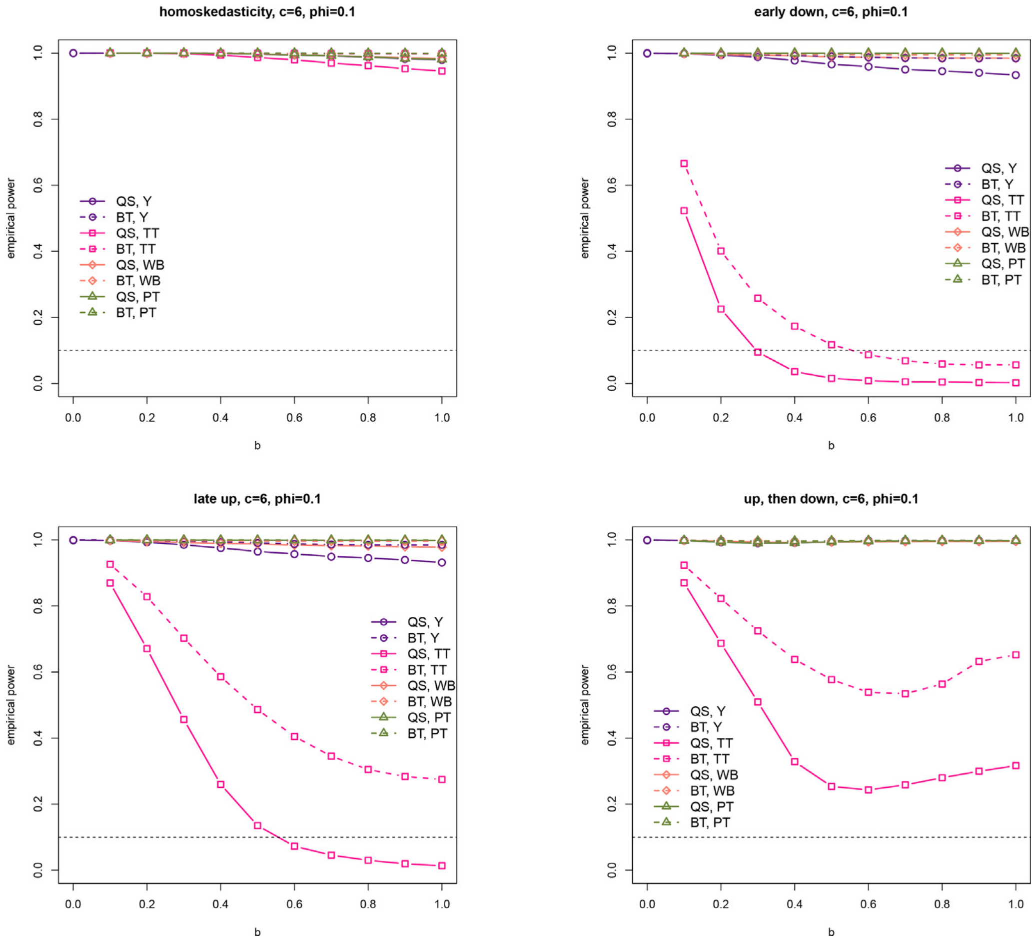

Fig. 11. Power vs. b, $\phi = 0.1$, $c = 6$, $T = 100$. See notes to Figure 4.

References

Amado, C., Teräsvirta, T., 2013. Modelling volatility by variance decomposition. Journal of Econometrics 175 (2), 142–153.
Andrews, D.W.K., 1991. Heteroskedasticity and autocorrelation consistent covariance matrix estimation. Econometrica 59 (3), 817–858.
Andrews, D.W.K., Monahan, J.C., 1992. An improved heteroskedasticity and autocorrelation consistent covariance matrix estimator. Econometrica 60 (4), 953–966.
Bakirov, N.K., Székely, G.J., 2005. Student's t-test for Gaussian scale mixtures. Zapiski Nauchnyh Seminarov POMI 328, 5–19.
Boswijk, H.P., Cavaliere, G., Georgiev, I., Rahbek, A., 2021. Bootstrapping non-stationary stochastic volatility. Journal of Econometrics 224 (1), 161–180.
Casini, A., 2022. Theory of evolutionary spectra for heteroskedasticity and autocorrelation robust inference in possibly misspecified and nonstationary models. Journal of Econometrics.
Cavaliere, G., 2004. Unit root tests under time-varying variances. Econometric Reviews 23 (3), 259–292.
Cavaliere, G., Rahbek, A., Taylor, A.M.R., 2010. Testing for co-integration in vector autoregressions with non-stationary volatility. Journal of Econometrics 158 (1), 7–24.
Cavaliere, G., Taylor, A.M.R., 2008a. Bootstrap unit root tests for time series with nonstationary volatility. Econometric Theory 24 (1), 43–71.
Cavaliere, G., Taylor, A.M.R., 2008b. Time-transformed unit root tests for models with non-stationary volatility. Journal of Time Series Analysis 29 (2), 300–330.
Davidson, J., 1994. Stochastic Limit Theory. Oxford University Press.
Davidson, R., MacKinnon, J.G., 1999. The size distortion of bootstrap tests. Econometric Theory 15, 361–376.
Demetrescu, M., Hanck, C., Kruse, R., 2022. Robust inference under time-varying volatility: A real-time evaluation of professional forecasters. Journal of Applied Econometrics 37 (5), 1010–1030.
Deng, A., Perron, P., 2008. The limit distribution of the CUSUM of squares test under general mixing conditions. Econometric Theory 24 (3), 809–822.
Djogbenou, A.A., MacKinnon, J.G., Nielsen, M.O., 2019. Asymptotic theory and wild bootstrap inference with clustered errors. Journal of Econometrics 212 (2), 393–412.
Fama, E., French, K., 1993. Common risk factors in the returns on stocks and bonds. Journal of Financial Economics 33 (1), 3–56.
Fama, E., French, K., 2015. A five-factor asset pricing model. Journal of Financial Economics 116 (1), 1–22.
Gonçalves, S., Vogelsang, T.J., 2011. Block bootstrap HAC robust tests: The sophistication of the naive bootstrap. Econometric Theory 27 (4), 745–791.
Hansen, B.E., 2000. Testing for structural change in conditional models. Journal of Econometrics 97 (1), 93–115.
Hansen, L.P., 1982. Large sample properties of generalized methods of moments estimators. Econometrica 50 (2), 1029–1054.
Ibragimov, R., Müller, U.K., 2010. t-statistic based correlation and heterogeneity robust inference. Journal of Business & Economic Statistics 28 (4), 453–468.
Kiefer, N.M., Vogelsang, T.J., 2002a. Heteroskedasticity-autocorrelation robust standard errors using the Bartlett kernel without truncation. Econometrica 70 (5), 2093–2095.

Kiefer, N.M., Vogelsang, T.J., 2002b. Heteroskedasticity-autocorrelation robust testing using bandwidth equal to sample size. Econometric Theory 18 (6), 1350–1366.
Kiefer, N.M., Vogelsang, T.J., 2005. A new asymptotic theory for heteroskedasticity-autocorrelation robust tests. Econometric Theory 21 (6), 1130–1164.
Kiefer, N.M., Vogelsang, T.J., Bunzel, H., 2000. Simple robust testing of regression hypotheses. Econometrica 68 (3), 695–714.
Lazarus, E., Lewis, D.J., Stock, J.H., 2021. The size-power tradeoff in HAR inference. Econometrica 89 (5), 2497–2516.
Lazarus, E., Lewis, D.J., Stock, J.H., Watson, M.W., 2018. HAR inference: Recommendations for practice. Journal of Business & Economic Statistics 36 (4), 541–559.
Leeb, H., Pötscher, B.M., 2005. Model selection and inference: Facts and fiction. Econometric Theory 21 (1), 21–59.
Newey, W.K., West, K.D., 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. Econometrica 55 (3), 703–708.
Phillips, P.C.B., Xu, K.L., 2006. Inference in autoregression under heteroskedasticity. Journal of Time Series Analysis 27 (2), 289–308.
Pötscher, B.M., Preinerstorfer, D., 2018. Controlling the size of autocorrelation robust tests. Journal of Econometrics 207 (2), 406–431.
Preinerstorfer, D., Pötscher, B.M., 2016. On size and power of heteroskedasticity and autocorrelation robust tests. Econometric Theory 32 (2), 261–358.
Ray, S., Savin, N., 2008. The performance of heteroskedasticity and autocorrelation robust tests: a Monte Carlo study with an application to the three-factor Fama–French asset-pricing model. Journal of Applied Econometrics 23 (1), 91–109.
Rho, S., Vogelsang, T.J., 2021. Inference in time series models using smoothed-clustered standard errors. Journal of Econometrics 224 (1), 113–133.
Shao, X., 2010. The dependent wild bootstrap. Journal of the American Statistical Association 105 (489), 218–235.
Shao, X., 2015. Self-normalization for time series: a review of recent developments. Journal of the American Statistical Association 110 (512), 1797–1817.
Stock, J.H., Watson, M.W., 2002. Has the business cycle changed and why? NBER Macroeconomics Annual 17 (1), 159–218.
Sun, Y., 2014. Fixed-smoothing asymptotics in a two-step generalized method of moments framework. Econometrica 82 (6), 2327–2370.
Sun, Y., Phillips, P.C.B., Jin, S., 2008. Optimal bandwidth selection in heteroskedasticity-autocorrelation robust testing. Econometrica 76 (1), 175–194.
Vogelsang, T.J., 1999. Sources of nonmonotonic power when testing for a shift in mean of a dynamic time series. Journal of Econometrics 88 (2), 283–299.
Vogelsang, T.J., Wagner, M., 2013. A fixed-b perspective on the Phillips-Perron unit root tests. Econometric Theory 29 (3), 609–628.
Vogelsang, T.J., Wagner, M., 2014. Integrated modified OLS estimation and fixed-b inference for cointegrating regressions. Journal of Econometrics 178 (2), 741–760.
Xu, K.L., 2012. Robustifying multivariate trend tests to nonstationary volatility. Journal of Econometrics 169, 147–154.