

A Service of

ZBW

Leibniz-Informationszentrum Wirtschaft Leibniz Information Centre for Economics

Hanck, Christoph; Demetrescu, Matei; Kruse, Robinson

# Conference Paper Fixed-b Asymptotics for t-Statistics in the Presence of Time-Varying Volatility

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2015: Ökonomische Entwicklung -Theorie und Politik - Session: Time Series Econometrics, No. A23-V3

### **Provided in Cooperation with:**

Verein für Socialpolitik / German Economic Association

*Suggested Citation:* Hanck, Christoph; Demetrescu, Matei; Kruse, Robinson (2015) : Fixed-b Asymptotics for t-Statistics in the Presence of Time-Varying Volatility, Beiträge zur Jahrestagung des Vereins für Socialpolitik 2015: Ökonomische Entwicklung - Theorie und Politik - Session: Time Series Econometrics, No. A23-V3, ZBW - Deutsche Zentralbibliothek für Wirtschaftswissenschaften, Leibniz-Informationszentrum Wirtschaft

This Version is available at: https://hdl.handle.net/10419/112916

#### Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

#### Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



# WWW.ECONSTOR.EU

# Fixed-b Asymptotics for t-Statistics in the Presence of Time-Varying Volatility\*

Matei Demetrescu,<sup>a</sup> Christoph Hanck<sup>b</sup> and Robinson Kruse<sup>c</sup>

<sup>*a*</sup>Christian-Albrechts-University of Kiel<sup> $\dagger$ </sup> <sup>*b*</sup>University of Duisburg-Essen<sup> $\ddagger$ </sup> <sup>*c*</sup>CREATES, Aarhus University and Leibniz University Hannover<sup>§</sup>

Preliminary version: January 19, 2015

#### Abstract

The fixed-*b* asymptotic framework provides refinements in the use of heteroskedasticity and autocorrelation consistent variance estimators. We show however that the fixed-*b* limiting distributions of *t*-statistics are not pivotal when the variance of the underlying data generating process changes over time. To regain pivotal fixed-*b* inference under such time heteroskedasticity, we discuss three alternative approaches. We employ (1) the wild bootstrap (Cavaliere and Taylor, 2008, ET), (2) resort to time transformations (Cavaliere and Taylor, 2008, JTSA) and (3) suggest to pick suitable the asymptotics according to the outcome of a heteroskedasticity test, since small-*b* asymptotics deliver standard limiting distributions irrespective of the so-called variance profile of the series. We quantify the degree of size distortions from using the standard fixed-*b* approach and compare the effectiveness of the corrections via simulations. We also provide an empirical application to excess returns.

Keywords: Hypothesis tests, HAC estimation, HAR testing, Bandwidth, Robustness

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

<sup>\*</sup>The first two authors gratefully acknowledge the support of the German Research Foundation (DFG) through the projects DE 1617/4-1 and HA 6766/2-1. Robinson Kruse gratefully acknowledges financial support from CREATES funded by the Danish National Research Foundation.

<sup>&</sup>lt;sup>†</sup>Institute for Statistics and Econometrics, Christian-Albrechts-University of Kiel, Olshausenstr. 40-60, D-24118 Kiel, Germany, email: mdeme@stat-econ.uni-kiel.de.

<sup>&</sup>lt;sup>‡</sup>Corresponding author: Faculty of Economics and Business Administration, University of Duisburg-Essen, Universitätsstraße 12, D-45117 Essen, Germany, e-mail address: christoph.hanck@vwl.uni-due.de.

<sup>&</sup>lt;sup>8</sup>CREATES, Aarhus University, School of Economics and Management, Fuglesangs Allé 4, DK-8210 Aarhus V, Denmark, e-mail address: rkruse@creates.au.dk. Leibniz University Hannover, School of Economics and Management, Institute for Statistics, Königsworther Platz 1, D-30167 Hannover, Germany, e-mail address: kruse@statistik.uni-hannover.de.

## 1 Introduction

Most macroeconomic and financial variables are observed over time. When conducting statistical inference using such data, it is therefore crucial to allow for serial dependence in the series. For weakly stationary series, the seminal contribution of Newey and West (1987) (see also Andrews, 1991) allows to asymptotically robustify the class of GMM (Hansen, 1982) hypothesis tests to the potential presence of serial correlation. Relying on a heteroskedasticity- and autocorrelation consistent [HAC] estimator of the variance of the estimators, this framework allows for normal or  $\chi^2$  asymptotics. The asymptotic distributions turned out, however, to be fairly poor approximations to the actual finite-sample distributions, often leading to substantial size distortions in applied work. As a consequence, test results are often sensitive to the choice of bandwidth B and kernel k employed for estimating the variance. Also, the asymptotics require that a vanishing fraction  $b := B/T \to 0$  of the number of time series observations T is used. In actual applications, b must of course be positive.

To tackle these finite-sample issues with HAC variance estimation, a series of contributions, including Kiefer et al. (2000) and Kiefer and Vogelsang (2002a,b, 2005), proposes a new asymptotic framework, labelled fixed-*b* asymptotics, in which it is not required that  $b \rightarrow 0$ . This leads to new asymptotically heteroskedasticity- and autocorrelation robust [HAR] distributions (reviewed in more detail below) for the standard *t* and Wald-type test statistics in the GMM framework. Conveniently, the distributions reflect the choice of bandwidth and kernel even in the limit. The above-cited papers convincingly demonstrate that the asymptotic distributions may provide substantially better approximations to actual finite-sample distributions. The usefulness of such procedures has spawned a very active literature. An incomplete list of recent contributions includes Yang and Vogelsang (2011), Vogelsang and Wagner (2013) or Sun (2014a,b).

Our first contribution is to show in 3 that fixed-*b* asymptotics unfortunately do not automatically lead to finite-sample improvements in all empirically relevant settings and should hence only be used with additional care. In particular, variances—as a measure of volatility—varying in time (i.e. time heteroskedasticity) affect limiting distributions in the fixed-*b* framework and thus lead to a loss of asymptotic pivotality. Time-varying variances are present in many financial (see among others Guidolin and Timmermann, 2006; Amado and Teräsvirta, 2014; Teräsvirta and Zhao, 2011; Amado and Teräsvirta, 2013) and macroeconomic (see e.g. Stock and Watson, 2002; Sensier and van Dijk, 2004; Clark, 2009, 2011; Justiniano and Primiceri, 2008) time series such as excess returns, economic growth or inflation. Time-varying variances include, but are not limited to, permanent breaks or trends in the variance properties of (the innovations of) the series.<sup>1</sup> Correspondingly, consequences of and remedies for time heteroskedasticity for inference with dependent data have received substantial attention in recent years.<sup>2</sup> Yet, if one lets  $b \rightarrow 0$  as in

 $<sup>^{1}</sup>$ In the macroeconomic literature, a particular such phase of declining volatility at the end of the millennium is known as the "Great Moderation."

<sup>&</sup>lt;sup>2</sup>For stationary autoregressions see e.g. Phillips and Xu (2006) or Xu (2008); for unit root autoregressions, see Cavaliere and Taylor (2008b) or Cavaliere and Taylor (2009). The effects of time-varying volatility are amplified in panels of (nonstationary) series, making corrections all the more necessary; see e.g. Demetrescu and Hanck (2012) or Westerlund (2014).

Newey and West (1987) time-varying variance does not have an asymptotic effect (see Cavaliere, 2004). Practitioners thus face a trade-off in the precision of the critical values provided by the fixed-b approach, a trade-off which is determined by the strength of the time heteroskedasticity in the data generating process [DGP].

Therefore, the second contribution of the paper is to discuss methods for correctly sized fixed-b inference in the presence of time-varying variances, thus making the trade-off irrelevant. To achieve this, we make use of suitably modified techniques from the unit root testing literature (Section 4). More specifically, we build on the work of Cavaliere and Taylor (2008a) that suggests to employ a wild bootstrap scheme. Alternatively, we propose to time-transform heteroskedastic series as in Cavaliere and Taylor (2008b) so as to recover homoskedasticity prior to conducting the test. Since Cavaliere and Taylor work with integrated series directly, while our setup assumes integration of order zero, a modification of the transformation algorithm is needed. Third, we demonstrate that a pretest for time-varying variance can also be used for robustification: depending on the outcome of the test, one either uses small-b methods valid under heteroskedasticity or fixed-b methods requiring homoskedasticity.

Simulation results presented in Section 5 confirm our main analytical predictions: First and as is well-known, the standard HAC tests are size distorted in finite samples when there is serial correlation, a distortion which can be remedied using fixed-*b* methods under homoskedasticity. The latter are however size distorted under heteroskedasticity. Second, the corrections suggested here yield better finite-sample size under heteroskedasticity. They also show good performance under homoscedasticity. The time-transformation is somewhat conservative, while the wild bootstrap is sometimes slightly upward size distorted. Third and as one would expect, the pre-test has an intermediate position. Fourth, the wild bootstrap turns out to be more powerful than the time transformation.

Section 6 provides an empirical application to excess returns of US stocks and 10-year bonds over 30-day US treasury bills, illustrating the potential empirical effect of using testing procedures which are and are not robust to time-varying variances. Section 7 concludes. The appendices collect proofs and other derivations.

# 2 Fixed-b HAR testing

In this paper, we focus on the simple and prototypical case of tests for the mean of a series  $y_t$ ,  $E(y_t) = \mu$ . That is, we test  $H_0$ :  $\mu = \mu_0$ . The findings to be presented however generalize readily to other testing problems, e.g. to the case of testing moment restrictions. Our goal is to provide tests which are robust to the potential presence of both heteroskedasticity and autocorrelation.

The classical *t*-test for  $\mu$  relies on the normalized sample mean,

$$\sqrt{T}\left(\frac{\bar{y}-\mu_0}{\omega}\right)$$

with  $\bar{y} = \frac{1}{T} \sum_{t=1}^{T} y_t$  the sample average of  $y_t$  and  $\omega^2 = \lim_{T \to \infty} \operatorname{Var}(\sqrt{T}(\bar{y} - \mu))$ , the so-called long-run variance of  $y_t$ . We thus restrict ourselves to the case of  $\sqrt{T}$ -consistent sample averages, which are given for independent, both identically and heterogeneously distributed random variables, as well as serially correlated short memory series. The long-run variance—as opposed to the variance of  $y_t$ —captures the effect of possible serial correlation or heteroskedasticity on the sample average, hence the acronym HAC for its estimate.

Should  $y_t$  be weakly stationary with absolutely summable autocovariances  $\gamma_j = \text{Cov}(y_t, y_{t-j})$ , it holds that  $\omega^2 = \sum_{j=-\infty}^{\infty} \gamma_j$ . Regularity conditions assumed,<sup>3</sup> a central limit theorem applies for  $\bar{y}$  and

$$\sqrt{T}\left(\frac{\bar{y}-\mu_0}{\omega}\right) \stackrel{d}{\to} \mathcal{N}(0,1)$$

under the null.

In practice, the long-run variance  $\omega^2$  is unknown and has to be estimated, leading to the feasible *t*-ratio

$$\mathcal{T} = \sqrt{T} \left( \frac{\bar{y} - \mu_0}{\hat{\omega}} \right). \tag{1}$$

The most popular HAC estimators  $\hat{\omega}^2$  rely on suitably weighted sums of sample autocovariances; see Newey and West (1987) and Andrews (1991).<sup>4</sup> Thus,

$$\hat{\omega}^2 = \sum_{j=-T+1}^{T-1} k\left(\frac{j}{B}\right) \hat{\gamma}_j$$

where k is the employed kernel function, B denotes the so-called bandwidth and  $\hat{\gamma}_j$  is the *j*thorder sample autocovariance,  $\hat{\gamma}_j = \frac{1}{T} \sum_{t=j+1}^T (y_t - \bar{y}) (y_{t-j} - \bar{y})$ . Under additional regularity conditions (see e.g. Andrews, 1991), and in particular  $b = B/T \to 0$  at suitable rates, consistency follows,  $\hat{\omega} \xrightarrow{p} \omega$ , and

$$\mathcal{T} \stackrel{d}{\to} \mathcal{N}(0,1).$$

under  $H_0$ . Although the asymptotics are the same for any suitable kernel and bandwidth choice, the finite-sample behavior of  $\mathcal{T}$  does depend on the kernel and bandwidth chosen in the test situation at hand. The quality of the asymptotic approximation thus depends on user input. To make this dependence explicit, Kiefer and Vogelsang (2005) let  $B/T = b \in (0, 1]$  for the asymptotic analysis. While the resulting limiting distribution is free of nuisance parameters, it is nonstandard. But, more interestingly, it depends directly on the kernel k and indirectly (via b) on the bandwidth B, thus offering second-order refinements to the usual, small-b asymptotics where  $b \to 0$ ; see Sun (2014b). Concretely,

$$\mathcal{T} \stackrel{d}{\to} \mathcal{B}(k,b)$$

<sup>&</sup>lt;sup>3</sup>See e.g. Davidson (1994, Chapter 24) for sets of suitable assumptions.

<sup>&</sup>lt;sup>4</sup>Semiparametric estimates based on AR approximations (e.g. Berk, 1974) or on so-called steep origin kernels (Phillips et al., 2006) are also available in the literature.

where, assuming for simplicity that we work with kernels with smooth derivatives (of which Andrews' quadratic spectral (QS) kernel is an example),

$$\mathcal{B}(k,b) = \frac{W(1)}{\sqrt{-\int_0^1 \int_0^1 \frac{1}{b^2} k''\left(\frac{r-s}{b}\right) \left(W(r) - rW(1)\right) \left(W(s) - sW(1)\right) \mathrm{d}r\mathrm{d}s}}$$
(2)

with W(s) a standard Wiener process. The corresponding critical values for  $\mathcal{T}$  are tabulated as a function of k and b in Kiefer and Vogelsang (2005). For  $b \to 0$ ,  $\mathcal{B}(k, b) \xrightarrow{d} \mathcal{N}(0, 1)$  and small-b asymptotics are, in a sense, a particular case of the fixed-b approach.

Note that the functional  $\mathcal{B}(k, b)$  depends on the entire path of the Wiener process and not only on W(1), as is the case with the small-*b* approach. This has consequences when the volatility of  $y_t$  varies in time, as we shall see in the following section.

# 3 Failure of fixed-b HAR tests under time-varying variance

In order to analyze fixed-b asymptotics of  $\mathcal{T}$  under time-varying variance, we assume a multiplicative component structure of the series to be tested.

**Assumption 1** Let the observed series  $y_t$  be generated as

$$y_t = h_t v_t + \mu, \qquad t = 1, \dots, T,$$

where the stochastic component  $v_t$  is zero-mean stationary as specified by below, and the time heteroskedasticity is induced by the function h, also specified below.

This multiplicative structure is common in the literature; see e.g. Cavaliere (2004). This makes  $y_t$  a uniformly modulated process (Priestley, 1988, p. 165).<sup>5</sup> To conduct the asymptotic analysis, we assume the stochastic component to have short memory in the following sense.

**Assumption 2** Let  $v_t$  be a zero-mean strictly stationary series with unity long-run variance,  $L_{2+\delta}$ -bounded for some  $\delta > 0$ , and strong mixing with coefficients  $\alpha(j)$  for which

$$\sum_{j\geq 0} \alpha\left(j\right)^{1/p - 1/(2+\delta)} < \infty$$

for some 2 .

Strong mixing with coefficients  $\alpha$  satisfying some summability condition is a standard way of imposing short memory onto  $v_t$  (and thus  $y_t$ ); cf. e.g. Phillips and Durlauf (1986). Restricting the long-run variance to equal unity is not leading to any loss of generality, and allows one to

<sup>&</sup>lt;sup>5</sup>Other contributions model  $v_t$  explicitly as a linear process with modulated innovations; see e.g. Cavaliere and Taylor (2008a,b). Demetrescu and Sibbertsen (2014) argue that the two DGPs are essentially equivalent.

interpret  $h_t^2$  as the long-run variance of the series  $y_t$ . The assumption yields (see e.g. Davidson, 1994, Chapter 29) weak convergence of the partial sums of  $v_t$  to a standard Wiener process,

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{\left[sT\right]}v_{t} \Rightarrow W\left(s\right),$$

so  $v_t$  is integrated of order 0. While  $y_t$ , being a modulated version of  $v_t$ , is also strong mixing, its partial sums exhibit a limiting behavior depending on the modulating function  $h_t$ .

**Assumption 3** Let  $h_t = h(t/T)$  with  $h(\cdot)$  a deterministic, piecewise Lipschitz function, positive at all  $s \in [0, 1]$ .

This allows for general patterns of smoothly or abruptly changing variance, as long as the abrupt changes are not too frequent.<sup>6</sup>

Under the conditions spelled out by the above assumptions, we have (for details see Cavaliere, 2004, Lemma 3)

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{[sT]} (y_t - \mu) \Rightarrow \int_0^s h(v) \, \mathrm{d}W(v) \equiv B_h(s) \, .$$

The process  $B_h(s)$  is a Gaussian process, but not a Brownian motion: the covariance kernel of  $B_h$  is given by

$$\operatorname{Cov}\left(B_{h}\left(s\right),B_{h}\left(r\right)\right)=\int_{0}^{\min\{s,r\}}h^{2}\left(v\right)\mathrm{d}v$$

which is not proportional to min  $\{s, r\}$ , the covariance kernel of the standard Wiener process.

It is precisely the fact that the normalized partial sums of the centered  $y_t$  do not converge weakly to a Brownian motion that affects the fixed-*b* asymptotics of  $\mathcal{T}$ . Under a time-heteroskedastic DGP, the limiting distribution given by fixed-*b* asymptotics is stated in the following proposition.

**Proposition 1** Under  $H_0$  and Assumptions 1-3, it holds for kernels with smooth derivatives and  $b \in (0, 1]$  that

$$\mathcal{T} \stackrel{d}{\rightarrow} \mathcal{B}\left(h,k,b
ight) \equiv rac{B_{h}\left(1
ight)}{\sqrt{Q_{h,k,b}}}$$

as  $T \to \infty$ , where

$$Q_{h,k,b} = -\int_0^1 \int_0^1 \frac{1}{b^2} k'' \left(\frac{r-s}{b}\right) \left(B_h\left(r\right) - rB_h\left(1\right)\right) \left(B_h\left(s\right) - sB_h\left(1\right)\right) \, \mathrm{d}r \mathrm{d}s$$

**Proof:** See the Appendix.

<sup>&</sup>lt;sup>6</sup>Seasonally varying variances are excluded, for instance. This is not critical, however, since the work of Burridge and Taylor (2001) suggests that seasonally varying variances actually average out and do not affect convergence to Wiener process.

Proposition 1 demonstrates that HAR testing is not robust to time heteroskedasticity for fixed b. Although  $B_h(1)$  is normal with mean zero and variance  $\bar{\omega}^2 = \int_0^1 h^2(s) \, ds$ , the distribution of  $\mathcal{B}(h, k, b)$  is different from that of  $\mathcal{B}(k, b)$  whenever h is not constant almost everywhere. This is because  $Q_{h,k,b}$  is essentially different from the denominator of (2) under time-varying volatility.

Figure 1 illustrates this lack of pivotality, showing quantile-quantile (QQ) plots for the distributions  $\mathcal{B}(h, k, b)$  for  $b = \{0.1, 0.5, 0.9\}$  and four different h. We take T = 1000 and simulate the  $\mathcal{B}(h, k, b)$  with 50,000 replications. The kernel k is taken to be the Quadratic spectral kernel. Under DGP1, volatility is constant over time. This case is reported as a benchmark, where we compare the quantiles of  $\mathcal{B}(k, b)$  with themselves. The first row of graphs in Figure 1 show the results for a small, medium and large value of b. The negligible deviations are due to small Monte Carlo error. An early downward break in volatility is present in DGP2. Here, we compare the quantiles of  $\mathcal{B}(h, k, b)$  on the y-axis with the corresponding quantiles of  $\mathcal{B}(k, b)$  on the x-axis. The results shown in the second row of Figure 1 clearly demonstrate differences between the two distributions. The larger b, the more pronounced is the discrepancy. For the third DGP (exhibiting a late upward break), differences are present but less visible, except for extreme quantiles. As indicated by the vertical bars representing the 5% critical values, this non-pivotality might be overlooked by a test at that level. The results for the linear upward trend in volatility in DGP4 nicely illustrates the difference between the distributions as well as the role of b.

For b = 0, however, robustness to time-varying volatility is recovered. In a nutshell,  $B_h(1)$  follows a normal distribution with mean zero and variance  $\bar{\omega}^2$ , which can be interpreted as the average long-run variance of the series. Moreover, Cavaliere (2004) shows that, under mild conditions on the rate at which *b* vanishes, the HAC variance estimator is consistent precisely for this variance,  $plim \hat{\omega}^2 = \bar{\omega}^2$ . Hence,

$$\mathcal{T} \xrightarrow{d} \mathcal{N}(0,1) \quad \text{for } b \to 0.$$

In other words, small-*b* methods asymptotically lead to pivotality under time-varying volatility as it does under weak stationarity. Recall however that the finite-sample quality of the asymptotic approximation in the small-*b* framework is meager, so practitioners essentially have to choose between the devil and the deep blue sea when not knowing the properties of the DGP.

Section 5 quantifies the size distortions resulting from ignoring time-varying variance when using fixed-*b* asymptotic approximations. It will also recall that, despite asymptotic robustness, the small-*b* approach will often exhibit fairly strong size distortions in small samples under both homo- and heteroskedasticity. Hence, corrections are required.

#### 4 Robust inference under time-varying volatility

The critical issue about the failing asymptotics is that the partial sums of  $y_t$  do not converge to a Brownian motion anymore. In the following, three different corrections are discussed.

First, we build on Cavaliere and Taylor (2008a) who propose the wild bootstrap as a way to deal with the heteroskedasticity issue in a unit root testing context. Heuristically, we therefore



Figure 1: Quantile-quantile plots to compare  $\mathcal{B}(k, b)$  (x-axis) to the distributions  $\mathcal{B}(h, k, b)$  under various variance profiles h and for different b. DGP1: constant volatility; DGP2: early downward break in volatility; DGP3: late upward break in volatility; DGP4: linear upward trend in volatility. The quadratic spectral kernel is employed. The dashed vertical line is the 95% critical value from the  $\mathcal{B}(k, b)$  distribution.

exploit the bootstrap to mimic the actual null distribution of  $\mathcal{T}$  under heteroskedasticity.<sup>7</sup> While Cavaliere and Taylor (2008a) bootstrap residuals from a fitted autoregression, it suffices here to wild bootstrap  $y_t$  directly. The algorithm is as follows.

- 1. Generate T iid standard normal random variables  $r_t^*$ .
- 2. Generate the wild bootstrap sample as  $y_t^* = r_t^* (y_t \bar{y})$ .
- 3. Compute the bootstrap test statistic  $\mathcal{T}^*$  based on the resampled series  $y_t^*$ .
- 4. Repeat steps 1-3 to obtain a set of M resampled statistics  $\mathcal{T}_m^*$ ,  $m = 1, \ldots, M$ .
- 5. Use the  $1 \alpha$ -quantile of  $\{\mathcal{T}_m^*\}_{m=1,\dots,M}$ , say  $q_{1-\alpha}^*$ , as critical value for the test.

The following proposition shows that the wild bootstrap procedure gives size control in the limit.

**Proposition 2** Under  $H_0$  and Assumptions 1-3, it holds as  $T \to \infty$  and  $B/T \to b \in (0,1]$  that

$$\Pr\left(\mathcal{T} > q_{1-\alpha}^*\right) \to \alpha$$

#### **Proof:** See the Appendix.

Alternatively to step 5, one could of course use bootstrap *p*-values for a test decision; it can be seen from the proof of the proposition that the wild bootstrap *p*-values converge weakly to a uniform distribution U[0, 1].

But the wild bootstrap is computationally demanding, even in the above simplified version. Our second correction therefore elaborates on the approach provided by Cavaliere and Taylor (2008b) which modifies the data in such a way that the series are in a sense transformed back to homoskedasticity. Hence, it will be valid to apply fixed-b methods applied to the transformed series. Their time transformation approach needs to be adapted to our setup, though, since Cavaliere and Taylor (2008b) deal with I(1) processes under the null when testing for a unit root, whereas we deal with I(0) processes. The procedure is as follows.

1. Subtract the mean of  $y_t$  under the null and build the cumulated sums

$$x_t = \sum_{j=1}^t (y_j - \mu_0)$$

- 2. Estimate the variance profile of  $x_t$ ,  $\hat{\eta}(s) = \frac{\sum_{t=1}^{[sT]} (y_j \mu_0)^2}{\sum_{t=1}^T (y_j \mu_0)^2}$ , and build its inverse g(s).
- 3. Time-transform  $x_t$  via

$$\tilde{x}_t = x_{[Tg(t/T)]}$$

<sup>&</sup>lt;sup>7</sup>Cavaliere and Taylor (2009) show the wild bootstrap to cope with stochastic volatility as well. The multivariate case has been dealt with in a series of papers starting with Cavaliere et al. (2010).

4. Base the test on the differenced series,  $\tilde{y}_t = \Delta \tilde{x}_t$ , i.e. compute

$$\tilde{\mathcal{T}} = \sqrt{T} \frac{\tilde{y}}{\tilde{\omega}}$$

where  $\tilde{\omega}^2$  is an estimator of the long-run variance of  $\tilde{y}_t$  using a bandwidth B = bT.

The following proposition shows that fixed-b asymptotics are recovered.

**Proposition 3** Under  $H_0$  and Assumptions 1-3, it holds as  $T \to \infty$  and  $B/T \to b \in (0, 1]$  that

$$\tilde{\mathcal{T}} \stackrel{d}{\to} \mathcal{B}(k,b)$$
.

**Proof:** See the Appendix.

Although autocorrelated errors do not enter the asymptotic distribution, they may impact the empirical size in small samples. In the following, we follow Andrews and Monahan (1992) and use a simple procedure for reducing the effect of short-memory dynamics while maintaining the same limiting distribution for the test statistic. It involves estimating a low-order ARMA model and filtering the series using the estimated coefficients. We need not assume that the true DGP is in fact an ARMA process, but rather use it as a reasonable approximation. Lag orders are selected via an information criterion. The time-transformation is applied to the series  $\hat{u}_t - \mu_0$ , where  $\hat{u}_t$  denotes the residual from the regression  $\Theta(L)y_t = \Phi(L)u_t$ ;  $\Theta(L)$  and  $\Phi(L)$  are MAand AR-polynomials, respectively.

The third correction for time-varying volatility we propose is closely related to what practitioners often do: only correct for a problem you have detected in the data. Essentially, we propose to first test for heteroskedasticity, and then work with either fixed-b or small-b asymptotics according to the outcome of the test. The intuition is that if a test for time-varying variance does not reject, then the departures from constant variances cannot be strong enough to seriously distort the fixed-b asymptotics of  $\mathcal{T}$ . If on the other hand the test rejects, then the small bandwidth-choice procedures may be preferable, or alternatively either of the two robust versions discussed above. The success of such a testing strategy obviously depends on the properties of the pre-test.

To this end, we resort to the test proposed by Deng and Perron (2008). In a recent simulation study, Bertram and Grote (2014) demonstrate that the test has decent size and power properties, see also Xu (2013). It is based on the series  $z_t = y_t^2$ . The test statistic is given by

$$Q = \sup_{1 \le t \le T} \frac{1}{\sqrt{T}} \frac{|D_t|}{\hat{\omega}_z}$$

where  $D_t = \sum_{j=1}^t z_j - \frac{t}{T} \sum_{j=1}^T z_j$  and  $\hat{\omega}_z$  is a HAC estimator of the long-run variance of  $z_t$ . The test rejects for large values of Q.

## 5 Simulation evidence

This section studies the finite-sample behavior for the various statistics discussed above in different settings.

We consider one-sided tests of  $H_0: \mu = 0$  against  $H_1: \mu > 0$ . The DGP is given by

$$y_t = \mu + v_t \tag{3}$$

$$(1 - \phi L)v_t = (1 - \theta L)h_t\varepsilon_t \tag{4}$$

with  $\varepsilon_t \sim i.i.d.N(0,1)$ . We consider an AR(1)-process with  $(\phi,\theta) = (0.85,0)$  as well as an ARMA(1,1)-process with  $(\phi,\theta) = (0.5, -0.45)$  and  $h_t$  as before. The following deterministic volatility DGPs for  $h_t$  are studied

DGP1: Constant volatility  $(h_t = 1)$ ;

- DGP2: Downward break in volatility at t = [0.2T] from  $\sigma_0 = 5$  to  $\sigma_1 = 1$ ;
- DGP3: Upward break in volatility at t = [0.8T] from  $\sigma_0 = 1$  to  $\sigma_1 = 5$ ;

DGP4: Linear upward trend in volatility:  $h_t = \sigma_0 + (\sigma_1 - \sigma_0)(t/T)$  with  $\sigma_0 = 1$  and  $\sigma_1 = 5$ .

For power results, we take  $\mu_T = c(\bar{V}/T)^{1/2}$  with  $\bar{V}$  being the average variance depending on the particular DGP1-4. By doing so, we achieve comparable results among different DGPs. Under homoskedasticity (DGP1), we have  $\bar{V} = \sigma^2$ , while  $\bar{V} = T^{-1} \sum_{t=1}^T (\sigma_0^2 + 1(t > [\tau T])\sigma_1^2)$  under DGPs 2 and 3 for instance. The parameter c > 0 is a localizing constant and takes the values 5 and 10.

Critical values for the fixed-*b* approach are taken from Kiefer and Vogelsang (2005), Table 1. The nominal significance level is 5%. We use the QS kernel and values of *b* ranging from 0.1 to 1 in increments of 0.1. The number of wild bootstrap replications equals M = 399, while the number of Monte Carlo replications is 5,000. The sample sizes are T = 100 and T = 500. For the time-transformation, an AR(1) finite-sample correction is applied for all DGPs.

First, we present size results. Our leading case is the one with an AR(1) component; the ARMA(1,1) results are discussed afterwards. Under the AR(1) DGP, the procedure applies the correct parametric pre-whitening correction to the series.

Under homoskedasticity (DGP1), the top-left entry of Table 1 reveals that, as is well-known, NW based on a Bartlett kernel with automatic bandwidth selection (see Andrews, 1991) faces substantial size distortions for T as large as T = 100. Confirming the results of Kiefer and Vogelsang (2005), the remainder of the first row shows that fixed-b asymptotics provide a very good approximation to the finite-sample distribution of the t-ratio given in equation (1), all but eliminating the size distortions of NW. Unsurprisingly, the size of the pretest is intermediate between that of NW and the fixed-b approaches. Similarly, the corrections based on time transformations and the wild bootstrap provide accurate tests under homoskedasticity. DGPs 2 and 4 confirm our analytical prediction from Section 3 that, in general, fixed-b asymptotics are not pivotal under heteroskedasticity. In particular, the tests seem to be conservative and increasingly so in b. That fixed-b asymptotics work relatively well for small b is not unexpected, as they then operate very similarly to the standard NW approach which would be valid asymptotically (see above and Cavaliere, 2004). Our suggested corrections are generally effective in removing the size distortions, although the wild bootstrap has some upward size distortions for small b. Again, the pretest strikes a useful compromise between fixed-b and NW.

Table 2 highlights the finite-sample character of the size distortions of NW, which are largely removed for T = 500. Similarly, the upward size distortions of the wild bootstrap have all but disappeared for T = 500. Table 2 moreover confirms that the distortions for fixed-*b* are *not* of a finite-sample nature.

The results for ARMA errors (Table 3) are qualitatively similar, although, unsurprisingly, the empirical sizes generally become somewhat less accurate when the prewhitening procedure does not soak up errors that follow the same parametric model as that applied in the procedure.

Tables 4 and 5 report power results. As expected, power increases in c. The particular type of variance break yields some variation in power, which may however be partly be explained by size distortions. Typically, power increases with decreasing values of b. This is plausible, as critical values increase in b. Tests applied to the original series (i.e. Y) and the wild bootstrap analogues (YB) yield relatively high power. The pre-testing strategy performs similarly well. When the time-transformation is applied, power is generally lower than for competing statistics.

This finding may be motivated as follows: under the alternative, the mean of the time-transformed series  $\tilde{y}_t$  is non-zero but not constant in general. Hence usual demeaning of the series  $\tilde{y}_t$  to compute a long run variance estimate does not eliminate its non-zero mean implying an inflated variance estimator due to the neglected deterministic component; see Appendix B for details.

In view of the size distortions observed for the wild bootstrap for small values of b as well as decreasing power in b, the present simulations would therefore suggest to use the wild bootstrap with an intermediate value of b (say, b = 0.4) in practice. The pretest would be a useful alternative.

#### 6 Excess returns

We illustrate the robustified procedures by an empirical application to financial excess returns  $z_t$ . These are defined as the difference between speculative portfolio returns and risk-free returns. A primary research question in financial econometrics involves testing against positive excess returns (on average) for speculative investments. Typically, monthly excess returns show a mild degree of autocorrelation and strong heteroskedasticity. Thus, when testing the null that  $E(y_t) = 0$  one needs to account for both features, such that the procedures discussed in this paper may be valuable.

(1987) standard errors; fixed-bandwidth parameter	med series. X means that a time-transformation is	$(\alpha = 5\%)$ . DGP1: constant volatility; DGP2: early	l trend in volatility.
Table 1: Size results under an AR(1) component, $T = 100$ . NW stands for Newey and	b takes values from $b = 0.1$ to $b = 1$ . Y indicates that the test is carried for the un	applied. YB stands for wild bootstrap versions. PT is a practitioner's pre-testing st	downward break in volatility; DGP3: late upward break in volatility; DGP4: linear

0.9 1	0.035	0.048	0.046 0.046	38 0.037		0.011 0.011	0.037 0.037	0.063	)40 0.039		)46 0.045	)41 0.040	0.062	0.046 0.046		0.025 0.025	0.041 0.041	0.043 0.043	
0.8	0.039 0.0	0.052 0.0	0.048 0.0	0.041 0.0		0.013 0.0	0.038 0.0	0.059 0.0	0.040 0.0		0.047 $0.0$	0.043 $0.0$	0.0 0.0	0.049 0.0		0.027 0.0	0.046 0.0	0.046 0.0	0 0 0 0 0
0.7	0.042 (	0.052 (	0.048 (	0.043 (		0.014 (	0.039 (	0.056 (	0.041 (		0.048 (	0.041 (	0.059 (	0.050 (		0.029 (	0.047 (	0.045 (	0 000
$\frac{1}{0.6}$	0.045	0.052	0.050	0.046		0.015	0.038	0.055	0.041		0.049	0.042	0.058	0.051		0.030	0.047	0.045	0.094
$\frac{\text{fixed-}l}{0.5}$	0.047	0.052	0.050	0.048		0.018	0.038	0.054	0.043		0.048	0.043	0.055	0.050		0.032	0.051	0.046	0.096
0.4	0.047	0.052	0.052	0.049		0.017	0.037	0.057	0.042		0.049	0.044	0.057	0.051		0.033	0.052	0.050	0.097
0.3	0.049	0.052	0.057	0.051		0.020	0.034	0.057	0.043		0.049	0.043	0.061	0.051		0.036	0.049	0.051	0.000
0.2	0.053	0.049	0.059	0.053		0.028	0.031	0.064	0.049		0.053	0.044	0.063	0.054		0.037	0.045	0.051	0.041
0.1	0.058	0.047	0.065	0.058		0.050	0.021	0.076	0.061		0.062	0.041	0.073	0.063		0.049	0.040	0.061	0.051
MN	0.091	0.052	0.085			0.080	0.034	0.084			0.089	0.044	0.085			0.070	0.041	0.071	
	$\frac{DGPI}{Y}$	Х	YB	$\mathbf{PT}$	DGP2	Y	X	YB	$\mathbf{PT}$	DGP3	Y	Х	YB	$\mathbf{PT}$	DGPI	Y	Х	YB	ГQ

West (1987) standard errors; fixed-bandwidth parameter	nsformed series. X means that a time-transformation is	tegy ( $\alpha = 5\%$ ). DGP1: constant volatility; DGP2: early	pward trend in volatility.
Table 2: Size results under an AR(1) component, $T = 500$ . NW stands for Newey and West (1987) standard errc	b takes values from $b = 0.1$ to $b = 1$ . Y indicates that the test is carried for the untransformed series. X means	applied. YB stands for wild bootstrap versions. PT is a practitioner's pre-testing strategy ( $\alpha = 5\%$ ). DGP1: cc	downward break in volatility; DGP3: late upward break in volatility; DGP4: linear upward trend in volatility.

1	0.037	0.041 0.038	0.010	0.052	$0.051 \\ 0.056$		0.046	0.047	0.053	0.050		0.028	0.047	0.046	0.054
0.9	0.038 0.046	0.044 0.039	0.011	0.053	0.050 0.056		0.046	0.047	0.053	0.051		0.029	0.048	0.047	0.055
0.8	0.040	0.046 0.041	0.011	0.052	0.050 0.056		0.049	0.047	0.053	0.052		0.030	0.051	0.050	0.055
0.7	0.042 0.048	0.046 0.043	0.012	0.053	0.053 0.056		0.050	0.047	0.051	0.053		0.032	0.051	0.049	0.055
$\frac{\text{values}}{0.6}$	0.044	0.046 0.045 0.045	0.013	0.051	0.052 0.056		0.049	0.049	0.053	0.052		0.034	0.052	0.051	0.056
$\frac{\text{fixed-}b}{0.5}$	0.045	0.047 0.046 0.046	0.014	0.050	$0.054 \\ 0.056$		0.050	0.049	0.053	0.053		0.036	0.053	0.053	0.056
0.4	0.045	0.048 0.046 0.046	0.015	0.050	$0.054 \\ 0.056$		0.049	0.047	0.053	0.052		0.038	0.051	0.055	0.056
0.3	0.046	$0.051 \\ 0.047$	0.017	0.051	$0.054 \\ 0.057$		0.045	0.047	0.054	0.050		0.038	0.052	0.055	0.056
0.2	0.047	0.051 0.047	0.020	0.050	$0.054 \\ 0.057$		0.046	0.047	0.052	0.050		0.039	0.051	0.051	0.056
0.1	0.043	0.048 0.044	0.030	0.047	0.061 0.058		0.044	0.046	0.052	0.048		0.041	0.050	0.050	0.056
MM	0.057	0.054	0.061	0.045	0.061		0.059	0.049	0.055			0.058	0.053	0.059	
	DGP1 Y X	YB PT	DGP2 Y	X	YB PT	60.7U	Y	Х	YB	$\mathbf{PT}$	DGP4	Υ	Х	YB	ΡT

	NIM/	10	с U	6 U		$\frac{\text{fixed-}b}{0.5}$	values 0.6	2 0	2	0 0	<del>, -</del>
DGP1								5			1
Y	0.159	0.091	0.067	0.065	0.060	0.055	0.054	0.052	0.049	0.047	0.045
×	0.154	0.101	0.078	0.071	0.069	0.066	0.063	0.062	0.060	0.056	0.054
YB	0.154	0.101	0.076	0.073	0.068	0.063	0.063	0.063	0.062	0.061	0.062
ΓŢ		0.098	0.077	0.077	0.073	0.068	0.067	0.065	0.062	0.060	0.058
$DGP_{2}$											
к Х	0.167	0.077	0.034	0.028	0.023	0.020	0.020	0.018	0.016	0.014	0.013
×	0.131	0.104	0.094	0.089	0.084	0.079	0.077	0.075	0.071	0.068	0.067
YB	0.173	0.119	0.084	0.073	0.069	0.068	0.067	0.070	0.071	0.074	0.076
T		0.144	0.135	0.133	0.133	0.133	0.132	0.132	0.132	0.132	0.132
DGP3											
Y	0.169	0.098	0.067	0.064	0.060	0.059	0.057	0.056	0.055	0.053	0.051
×	0.158	0.115	0.085	0.075	0.069	0.065	0.062	0.060	0.058	0.056	0.054
YB	0.165	0.107	0.080	0.075	0.070	0.068	0.068	0.067	0.068	0.070	0.072
ΓŢ		0.115	0.092	0.090	0.085	0.087	0.084	0.085	0.084	0.083	0.082
ופטע											
4 100 V	0 144	0.078	0.050	0.042	0.039	0.039	0.037	0.037	0.035	0.035	0.034
	0.143	0.106	0.081	0.074	0.071	0.066	0.064	0.062	0.060	0.058	0.055
ΥB	0.145	0.095	0.073	0.065	0.060	0.057	0.055	0.055	0.056	0.056	0.055
Ĺ		0.118	0.112	0.110	0.107	0.108	0.107	0.107	0.105	0.105	0.105

Table 3: Size results under an ARMA(1,1) component, T = 100. NW stands for Newey and West (1987) standard errors; fixed-bandwidth parameter  $\overline{b}$  takes values from b = 0.1 to b = 1. Y indicates that the test is carried for the untransformed series. X means that a timetransformation is applied. YB stands for wild bootstrap versions. PT is a practitioner's pre-testing strategy ( $\alpha = 5\%$ ). DGP1: constant volatility; DGP2: early downward break in volatility; DGP3: late upward break in volatility; DGP4: linear upward trend in volatility.

1		0.229	0.210	.279	.237			.164	0.057	0.505	0.241		0.270	0.167	0.324	0.274		0.263	0.117	0.353	0.275
0.9		0.239 (	0.215 0	0.284 (	0.247 0			0.177 0	0.058 (	0.512 0	0.252 (		0.286 (	0.176 0	0.331 (	0.289 (		0.271 0	0.129 (	0.367 (	0.283 (
0.8		0.254	0.228	0.290	0.262			0.195	0.058	0.513	0.266		0.301	0.177	0.340	0.304		0.284	0.137	0.380	0.296
0.7		0.270	0.242	0.301	0.278			0.216	0.057	0.516	0.285		0.306	0.183	0.339	0.309		0.296	0.149	0.388	0.304
values 0.6		0.281	0.245	0.304	0.289			0.240	0.059	0.527	0.302		0.316	0.187	0.346	0.319		0.303	0.157	0.394	0.312
$\frac{\text{fixed-}b}{0.5}$		0.285	0.259	0.324	0.293			0.260	0.067	0.543	0.318		0.330	0.197	0.353	0.333		0.317	0.172	0.407	0.325
0.4		0.313	0.266	0.345	0.320			0.304	0.071	0.554	0.353		0.338	0.208	0.368	0.340		0.342	0.187	0.422	0.351
0.3		0.337	0.276	0.364	0.344			0.373	0.068	0.573	0.408		0.356	0.211	0.380	0.358		0.382	0.197	0.443	0.390
0.2		0.367	0.273	0.388	0.373			0.444	0.062	0.583	0.466		0.372	0.222	0.405	0.375		0.413	0.204	0.465	0.418
0.1		0.412	0.247	0.435	0.415			0.495	0.084	0.545	0.509		0.401	0.206	0.434	0.401		0.463	0.166	0.493	0.466
MN		0.508	0.243	0.490				0.562	0.264	0.552			0.499	0.212	0.479			0.545	0.189	0.532	
	DGP1	Y	Х	YB	ΡT	( ( (	$DGP_{2}$	Х	Х	YB	$\mathbf{PT}$	DGP3	Υ	X	YB	$\mathbf{PT}$	DGP4	Υ	Х	YB	$\mathbf{PT}$

bandwidth parameter b takes values from b = 0.1 to b = 1. Y indicates that the test is carried for the untransformed series. X means that a time-transformation is applied. YB stands for wild bootstrap versions. PT is a practitioner's pre-testing strategy ( $\alpha = 5\%$ ). DGP1: Table 4: Power results for c = 5 under an AR(1) component, T = 100. NW stands for Newey and West (1987) standard errors; fixedconstant volatility; DGP2: early downward break in volatility; DGP3: late upward break in volatility; DGP4: linear upward trend in volatility.

						$\overline{\text{fixed-}p}$	values				
	ΜN	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
1											
	0.888	0.827	0.772	0.721	0.658	0.615	0.582	0.561	0.529	0.498	0.467
	0.403	0.456	0.525	0.520	0.481	0.441	0.419	0.397	0.375	0.357	0.340
	0.874	0.843	0.797	0.746	0.684	0.648	0.619	0.602	0.583	0.570	0.554
		0.827	0.774	0.728	0.667	0.626	0.593	0.572	0.540	0.509	0.479
Ø											
	0.870	0.825	0.800	0.757	0.690	0.629	0.571	0.530	0.503	0.467	0.436
	0.299	0.040	0.039	0.040	0.044	0.040	0.035	0.034	0.034	0.033	0.033
	0.852	0.850	0.859	0.846	0.828	0.813	0.804	0.795	0.794	0.789	0.784
		0.828	0.804	0.764	0.697	0.640	0.588	0.550	0.526	0.495	0.466
رە											
	0.878	0.829	0.779	0.733	0.699	0.678	0.654	0.633	0.614	0.596	0.575
	0.293	0.326	0.391	0.393	0.374	0.339	0.305	0.288	0.278	0.261	0.248
	0.867	0.834	0.796	0.758	0.722	0.697	0.685	0.668	0.661	0.654	0.646
		0.829	0.780	0.736	0.706	0.685	0.661	0.640	0.622	0.604	0.583
4											
	0.924	0.882	0.826	0.788	0.741	0.702	0.680	0.657	0.627	0.599	0.567
	0.203	0.229	0.281	0.256	0.226	0.198	0.178	0.158	0.140	0.131	0.115
	0.916	0.892	0.868	0.832	0.800	0.776	0.755	0.733	0.711	0.694	0.681
		0.882	0.827	0.793	0.748	0.713	0.691	0.669	0.639	0.611	0.579

bandwidth parameter b takes values from b = 0.1 to b = 1. Y indicates that the test is carried for the untransformed series. X means that a time-transformation is applied. YB stands for wild bootstrap versions. PT is a practitioner's pre-testing strategy ( $\alpha = 5\%$ ). DGP1: Table 5: Power results for c = 10 under an AR(1) component, T = 100. NW stands for Newey and West (1987) standard errors; fixedconstant volatility; DGP2: early downward break in volatility; DGP3: late upward break in volatility; DGP4: linear upward trend in volatility.

We use data from the Center for Research in Security Prices (CRSP) which have been analyzed in Guidolin and Timmermann (2006). The authors use multiple regime autoregressive Markov switching models to study the joint dynamics of stock and bond returns. They find strong evidence for time-varying volatility during different regimes. See also Amado and Laakkonen (2014) for evidence on time-varying volatility in bond markets. Three time series of continuously compounded excess returns are analyzed here (cf. Figure 2): (1) returns for small caps (first and second size-sorted CRSP decile portfolios) (2) returns for large caps (deciles 9 and 10 size-sorted CRSP decile portfolios) and (3) returns of a portfolio consisting of 10-year bonds. The 30-day T-bill rate is taken to be the risk-free rate. The sample period ranges from January 1954 to December 1999, resulting in T = 552 observations.

The sample averages  $\bar{z}$  of the three series are (1) 0.822% for small caps, (2) 0.657% for large caps and (3) 0.081% for the bond portfolio. The Deng and Perron (2008) pre-test against a structural change in volatility provides evidence against homoskedasticity. The test statistic is significant at the nominal ten percent level for small and large caps, while its *p*-value is below one percent for the bond portfolio series. Small and large caps show some degree of autocorrelation. Estimated first-order autocorrelation coefficients equal 0.190 and 0.078, respectively. Our test results for the null hypothesis of zero excess returns against its positiveness, i.e.  $H_0: E(y_t) = 0$ against  $H_1: E(y_t) > 0$ , are reported in Table 6. The finite-sample correction for autocorrelation is applied. For all three series, an AR(1) model is fitted. We use M = 2000 wild bootstrap replications.

Our findings for the portfolio consisting of small firms suggest a clear and economically meaningful pattern: All statistics are significant at least at the nominal ten percent level, while most of them are significant at the five percent as well. When using small-*b* asymptotics, we find highly significant statistics pointing towards positive excess returns. For the fixed-*b* approach, while we also find evidence against the null, the results are in line with those of the simulations: for larger values of *b* (i.e. b > 0.4), we find weaker evidence against  $H_0$ . Another finding is that the time-transformation procedure leads to smaller statistics at least for some values of *b*, which may be related to its lower power. As the pre-test does not reject at the five percent level, a practitioner would follow the decisions of the fixed-*b* approach.

The results for large firms are quite similar, although the evidence is weaker for choices of b > 0.2. The test statistics are highly significant only for a relatively small value of b = 0.1 and with Newey and West (1987) standard errors. Test decisions are very similar across methods and differ with b. Finally, no single statistic is significant for the 10-years bond portfolio, yielding no evidence against the null hypothesis. This is not too surprising given the relatively small sample average of 0.081% resulting from the numerous instances of inverted yield curves.

# 7 Concluding remarks

Fixed-b asymptotics are a tremendously useful device to enable more accurate finite-sample inference when dealing with serially correlated data. Serial correlation is however not the only



Figure 2: US Monthly excess returns (Jan 1954–Dec 1999, CRSP)

(1) Sr	nall caps					fixed-	<u>b</u> values				
	NW	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
Y	$2.684^{a}$	$3.260^{a}$	$5.311^{a}$	$7.385^{a}$	$7.553^{a}$	$7.812^{b}$	$8.498^{b}$	$9.414^{b}$	$10.448^{b}$	$11.542^{b}$	$12.685^{b}$
Х	$2.349^{a}$	$2.432^{b}$	$2.239^{c}$	$2.868^{b}$	$3.460^{b}$	$4.120^{b}$	$4.948^{b}$	$5.930^{b}$	$7.027^{b}$	$8.208^{b}$	$9.467^{b}$
YB	$2.684^{a}$	$3.260^{a}$	$5.311^{a}$	$7.385^{a}$	$7.553^{a}$	$7.812^{a}$	$8.498^{b}$	$9.414^{b}$	$10.448^{b}$	$11.542^{b}$	$12.685^{b}$
$\mathbf{PT}$		$3.260^{a}$	$5.311^{a}$	$7.385^{a}$	$7.553^{a}$	$7.812^{b}$	$8.498^{b}$	$9.414^{b}$	$10.448^{b}$	$11.542^{b}$	$12.685^{b}$
(2) La	arge cans					fixed-	h values				
(2) 10	NW	0.1	0.2	0.3	0.4	$\frac{n \chi c u}{0.5}$	0.6	0.7	0.8	0.9	1
V	2 6054	2 E1 1ª	2 0226	2 700°	0.770°	2 170 <sup>c</sup>	9.751 <sup>C</sup>	1 161C	E 9696	6 199°	7 071 <sup>c</sup>
ı v	0.090 0.100a	3.014	2.933 2.953	2.700 2.920c	2.119 2.019 <sup>c</sup>	0.170 0.0110	3.731 2.020	4.401 4 EOEC	0.202 E 270 <sup>c</sup>	0.152 6 915°	7.071
A VD	3.132 2 GOE <sup>a</sup>	3.093 $2 = 1.4^{a}$	2.800	2.030 2.700°	2.910 $2.770^{\circ}$	3.311 $2.170^{\circ}$	0.090 9.7510	4.090 4 4610	0.010 E 0600	0.210 6 120 <sup>c</sup>	7.107 $7.071^{\circ}$
YВ	3.095	3.514	2.933	2.700	2.119	$3.170^{\circ}$	3.731	4.401	5.202	0.132	7.071
PT		3.514-	2.933	2.700*	2.779*	3.170°	3.751°	4.461	5.262*	6.132*	1.071
(3) 10	Y bonds					fixed-	b values				
( )	NW	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
Y	0.824	0.872	0.861	0.811	0.815	0.860	0.931	1.018	1.114	1.216	1.324
Х	0.736	0.811	0.741	0.663	0.634	0.644	0.680	0.731	0.791	0.857	0.927
YB	0.824	0.872	0.861	0.811	0.815	0.860	0.931	1.018	1.114	1.216	1.324
$\mathbf{PT}$		0.824	0.824	0.824	0.824	0.824	0.824	0.824	0.824	0.824	0.824

Table 6: Empirical test results for  $E(y_t) = 0$  against  $H_1: E(y_t) > 0$ .

Note: NW stands for Newey and West (1987) standard errors; values ranging from b = 0.1 to b = 1 are fixed-bandwidth parameters. Y indicates that the test is carried for the untransformed series. X means that a time-transformation is applied. YB stands for wild bootstrap versions. PT is a practitioner's pre-testing strategy ( $\alpha = 5\%$ ). Superscripts 'a', 'b' and 'c' refer to significance at the one-, five- and ten-percent level, respectively. For further details see the text.

nuisance practitioners need to pay attention to when aiming to conduct reliable hypothesis tests: many important macroeconomic and financial time series are subject to time-varying variances such as variance breaks. We show that the standard fixed-*b* approach no longer yields pivotal tests under such heteroskedasticity. Based on wild bootstrap schemes (Cavaliere and Taylor, 2008a), on time transformations (Cavaliere and Taylor, 2008b) or on a pre-test procedure, we provide corrections that restore size control of fixed-*b* methods even under heteroskedasticity. Simulations illustrate the useful size and power properties of the corrections, in particular of the wild bootstrap approach. An empirical application to excess returns reveals the importance of properly accounting for time-varying variances in practice.

# Appendix

#### A Proofs

#### **Proof of Proposition 1**

Note that the arguments in the proof of Theorem 2 in Kiefer and Vogelsang (2005) can be used without further modification to conclude that

$$\mathcal{T} = \frac{\frac{1}{\sqrt{T}} \sum_{t=1}^{T} (y_t - \mu_0)}{\sqrt{-\frac{1}{T^2} \sum_{i=1}^{T-1} \sum_{j=1}^{T-1} \frac{T^2}{B^2} k'' \left(\frac{i-j}{B}\right) \frac{1}{\sqrt{T}} \sum_{t=1}^{i} (y_t - \bar{y}) \frac{1}{\sqrt{T}} \sum_{t=1}^{j} (y_t - \bar{y})} + o_p(1).$$

The weak convergence

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{[sT]} (y_t - \mu_0) \Rightarrow B_h(s)$$

then suffices for establishing the desired null distribution.

#### **Proof of Proposition 2**

Let  $S_T^*(s)$  denote the normalized partial sums of the bootstrapped centered sample,

$$S_T^*(s) = \frac{1}{\sqrt{T}} \sum_{t=1}^{[sT]} (y_t - \bar{y}) r_t^*$$

To guarantee size control in the limit, it suffices to show that the bootstrap partial sums converge weakly in probability to  $\operatorname{Var}(y_t) B_h(s)$ , since  $\operatorname{Var}(y_t)$  would cancel out in the bootstrapped *t*ratio. Note that, conditional on the sample  $y_t$ ,  $t = 1, \ldots, T$ ,  $S_T^*(s)$  is a Gaussian process with independent increments. Its covariance kernel is given by

$$\operatorname{Cov}\left(S_{T}^{*}\left(s\right), S_{T}^{*}\left(r\right)\right) = \frac{1}{T} \sum_{t=1}^{\left[\min\{s,r\}T\right]} \left(y_{t} - \bar{y}\right)^{2} E\left(\left(r_{t}^{*}\right)^{2}\right) = \frac{1}{T} \sum_{t=1}^{\left[\min\{s,r\}T\right]} \left(y_{t} - \bar{y}\right)^{2}.$$

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

$$\frac{1}{T}\sum_{j=1}^{[sT]} (y_j - \bar{y})^2 \Rightarrow \operatorname{Var}(y_t) \int_0^s h^2(r) \,\mathrm{d}r$$

i.e. that the wild bootstrap correctly replicates the variance profile of the sample  $y_t$  in the limit. Assumption 2 guarantees pointwise convergence of  $\frac{1}{T} \sum_{j=1}^{[sT]} (y_j - \bar{y})^2$  via a Law of Large Numbers for strong mixing processes (see Davidson, 1994, Section 20.6), and the monotonicity of the quadratic variation function leads to uniformity of the convergence, as required for the result.

#### **Proof of Proposition 3**

Cavaliere and Taylor (2008b, proof of Theorem 1) show that

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{[sT]}\tilde{y}_{t} \Rightarrow \tilde{\omega}W\left(s\right).$$

The result then follows like in the proof of Proposition 1.

## **B** Effect of time transformation under the alternative

To understand the mechanism, write

$$x_t = \sum_{j=1}^t (y_j - \mu_0) = \sum_{j=1}^t h_j v_j + t (\mu_1 - \mu_0).$$

Let  $\mu_1 - \mu_0 = \delta$ ; estimating the variance profile under the null implies that

$$\hat{\eta}(s) = \frac{\sum_{t=1}^{[sT]} (y_j - \mu_0)^2}{\sum_{t=1}^T (y_j - \mu_0)^2} \Rightarrow \frac{\operatorname{Var}(y_t) \int_0^s h^2(s) \, \mathrm{d}s + s\delta^2}{\operatorname{Var}(y_t) \int_0^1 h^2(s) \, \mathrm{d}s + \delta^2},$$

hence the estimate is asymptotically biased. This bias could be eliminated by demeaning  $y_t$  when estimating the variance profile, but this is, strictly speaking, not necessary from a robustness point of view since the estimator is not asymptotically biased under the null and size properties are not affected. Now, the transformation is

$$\tilde{x}_t = x_{[Tg(t/T)]}$$

with g the inverse of  $\hat{\eta}$ , which means that

$$\tilde{x}_t = \sum_{j=1}^{\left[Tg(t/T)\right]} h_j v_j + \left[Tg\left(\frac{t}{T}\right)\right]\delta.$$

The numerator of the test statistic is based on the time-transformed  $\tilde{x}_T = x_T$  having expectation  $T\delta$  (even with the biased variance profile estimate  $\hat{\eta}$ , given that  $\hat{\eta}(1) = 1$ ), so the power problem rests with the denominator of the test statistic. The HAC variance estimator is based on usual demeaning, while the mean of  $\Delta \tilde{x}_t$  is not constant in general,

$$E\left(\Delta \tilde{x}_t\right) = \delta\left(\left[T g\left(\frac{t}{T}\right)\right] - \left[T g\left(\frac{t-1}{T}\right)\right]\right) \approx \delta \tilde{g}'\left(\frac{t}{T}\right)$$

with  $\tilde{g}$  the weak limit of g.

#### References

- Amado, C. and H. Laakkonen (2014). Modeling time-varying volatility in financial returns: Evidence from the bond markets. In N. Haldrup, M. Meitz, and P. Saikkonen (Eds.), *Essays* in Nonlinear Time Series Econometrics, pp. 139–160. Oxford University Press.
- Amado, C. and T. Teräsvirta (2013). Modelling volatility by variance decomposition. Journal of Econometrics 175(2), 142–153.
- Amado, C. and T. Teräsvirta (2014). Modelling changes in the unconditional variance of long stock return series. Journal of Empirical Finance 25(1), 15–35.
- Andrews, D. W. (1991). Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica: Journal of the Econometric Society* 59(3), 817–858.
- Andrews, D. W. and J. C. Monahan (1992). An improved heteroskedasticity and autocorrelation consistent covariance matrix estimator. *Econometrica* 60(4), 953–966.
- Berk, K. N. (1974). Consistent autoregressive spectral estimates. *The Annals of Statistics* 2(3), 489–502.
- Bertram, P. and C. Grote (2014). A comparative study of volatility break tests. mimeo.
- Burridge, P. and A. M. R. Taylor (2001). On regression-based tests for seasonal unit roots in the presence of periodic heteroscedasticity. *Journal of Econometrics* 104(1), 91–117.
- Cavaliere, G. (2004). Unit root tests under time-varying variances. *Econometric Reviews* 23(3), 259–292.
- Cavaliere, G., A. Rahbek, and A. M. R. Taylor (2010). Testing for co-integration in vector autoregressions with non-stationary volatility. *Journal of Econometrics* 158(1), 7–24.
- Cavaliere, G. and A. M. R. Taylor (2008a). Bootstrap unit root tests for time series with nonstationary volatility. *Econometric Theory* 24(1), 43–71.
- Cavaliere, G. and A. M. R. Taylor (2008b). Time-transformed unit root tests for models with non-stationary volatility. *Journal of Time Series Analysis* 29(2), 300–330.
- Cavaliere, G. and A. M. R. Taylor (2009). Heteroskedastic time series with a unit root. *Econo*metric Theory 25(5), 1228–1276.
- Clark, T. E. (2009). Is the Great Moderation over? An empirical analysis. *Economic Review* 4, 5–42.
- Clark, T. E. (2011). Real-time density forecasts from bvars with stochastic volatility. Journal of

Business and Economic Statistics 29, 327–341.

- Davidson, J. (1994). Stochastic Limit Theory. Oxford university press.
- Demetrescu, M. and C. Hanck (2012). Unit root testing in heteroskedastic panels using the Cauchy estimator. Journal of Business & Economic Statistics 30, 256–264.
- Demetrescu, M. and P. Sibbertsen (2014). Inference on the long-memory properties of time series with non-stationary volatility. Technical Report 531, Leibniz Universität Hannover, Wirtschaftswissenschaftliche Fakultät.
- Deng, A. and P. Perron (2008). The limit distribution of the CUSUM of squares test under general mixing conditions. *Econometric Theory* 24(3), 809–822.
- Guidolin, M. and A. Timmermann (2006). An econometric model of nonlinear dynamics in the joint distribution of stock and bond returns. *Journal of Applied Econometrics* 21, 1–22.
- Hansen, L. P. (1982). Large sample properties of generalized methods of moments estimators. *Econometrica* 50(2), 1029–1054.
- Justiniano, A. and G. Primiceri (2008). The time-varying volatility of macroeconomic fluctuations. *American Economic Review* 98(3), 604–641.
- Kiefer, N. M. and T. J. Vogelsang (2002a). Heteroskedasticity-autocorrelation robust standard errors using the Bartlett kernel without truncation. *Econometrica* 70(5), 2093–2095.
- Kiefer, N. M. and T. J. Vogelsang (2002b). Heteroskedasticity-autocorrelation robust testing using bandwidth equal to sample size. *Econometric Theory* 18, 1350–1366.
- Kiefer, N. M. and T. J. Vogelsang (2005). A new asymptotic theory for heteroskedasticityautocorrelation robust tests. *Econometric Theory* 21(6), 1130–1164.
- Kiefer, N. M., T. J. Vogelsang, and H. Bunzel (2000). Simple robust testing of regression hypotheses. *Econometrica* 68(3), 695–714.
- Newey, W. K. and K. D. West (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica: Journal of the Econometric Society* 55(3), 703–708.
- Phillips, P. C. and S. N. Durlauf (1986). Multiple time series regression with integrated processes. The Review of Economic Studies 53(4), 473–495.
- Phillips, P. C. B., Y. Sun, and S. Jin (2006). Spectral density estimation and robust hypothesis testing using steep origin kernels without truncation. *International Economic Review* 47(3), 837–894.
- Phillips, P. C. B. and K. L. Xu (2006). Inference in autoregression under heteroskedasticity. Journal of Time Series Analysis 27(2), 289–308.
- Priestley, M. B. (1988). Non-linear and non-stationary time series analysis. Academic Press London.
- Sensier, M. and D. van Dijk (2004). Testing for volatility changes in U.S. macroeconomic time series. The Review of Economics and Statistics 86(3), 833–839.
- Stock, J. H. and M. W. Watson (2002). Has the business cycle changed and why? NBER Macroeconomics Annual 17(1), 159–218.
- Sun, Y. (2014a). Fixed-smoothing asymptotics in a two-step generalized method of moments

framework. *Econometrica* 82(6), 2327–2370.

- Sun, Y. (2014b). Let's fix it: Fixed-b asymptotics versus small-b asymptotics in heteroskedasticity and autocorrelation robust inference. *Journal of Econometrics* 178(3), 659–677.
- Teräsvirta, T. and Z. Zhao (2011). Stylized facts of return series, robust estimates and three popular models of volatility. *Applied Financial Economics* 21(1-2), 67–94.
- Vogelsang, T. J. and M. Wagner (2013). A fixed-b perspective on the Phillips-Perron unit root tests. *Econometric Theory 29*, 609–628.
- Westerlund, J. (2014). Heteroscedasticity robust panel unit root tests. Journal of Business & Economic Statistics 32(1), 112–135.
- Xu, K.-L. (2008). Bootstrapping autoregression under non-stationary volatility. *Econometrics Journal* 11, 1–26.
- Xu, K.-L. (2013). Powerful tests for structural changes in volatility. *Journal of Economet*rics 173(1), 126–142.
- Yang, J. and T. J. Vogelsang (2011). Fixed-b analysis of LM-type tests for a shift in mean. The Econometrics Journal 14, 438–456.