HAR TESTING FOR SPURIOUS REGRESSION IN TREND

By

Peter C. B. Phillips, Yonghui Zhang, and Xiaohu Wang

December 2018

COWLES FOUNDATION DISCUSSION PAPER NO. 2153

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
YALE UNIVERSITY
Box 208281
New Haven, Connecticut 06520-8281

http://cowles.yale.edu/

HAR Testing for Spurious Regression in Trend*

Peter C. B. Phillips[a], Yonghui Zhang[b], Xiaohu Wang[c]
[a]Yale University, University of Auckland,
University of Southampton & Singapore Management University
[b]Renmin University of China
[c]The Chinese University of Hong Kong

December 7, 2018

Abstract

The usual t test, the t test based on heteroskedasticity and autocorrelation consistent (HAC) covariance matrix estimators, and the heteroskedasticity and autocorrelation robust (HAR) test are three statistics that are widely used in applied econometric work. The use of these significance tests in trend regression is of particular interest given the potential for spurious relationships in trend formulations. Following a longstanding tradition in the spurious regression literature, this paper investigates the asymptotic and finite sample properties of these test statistics in several spurious regression contexts, including regression of stochastic trends on time polynomials and regressions among independent random walks. Concordant with existing theory (Phillips, 1986, 1998; Sun, 2004, 2014), the usual t test and HAC standardized test fail to control size as the sample size $n \to \infty$ in these spurious formulations, whereas HAR tests converge to well-defined limit distributions in each case and therefore have the capacity to be consistent and control size. However, it is shown that when the number of trend regressors $K \to \infty$, all three statistics, including the HAR test, diverge and fail to control size as $n \to \infty$. These findings are relevant to high dimensional nonstationary time series regressions.

JEL Classifications: C12, C14, C23

*This paper is written in tribute to Sir David Hendry and the vast contributions that he has made to econometrics over the past half century. The paper is based on a 2012 take-home examination set for graduate students in the School of Economics at Singapore Management University. An early draft of the main results was circulated in 2012. The paper was completed in December 2018. Phillips acknowledges support from the Kelly Foundation at the University of Auckland. Zhang acknowledges financial support from the National Natural Science Foundation of China (project No. 71401166). Corresponding author: Peter C. B. Phillips, Cowles Foundation for Research in Economics, Yale University, Box 208281, Yale Station, New Haven, Connecticut 06520-8281. Email: peter.phillips@yale.edu; Yonghui Zhang: yonghui.zhang@hotmail.com; Xiaohu Wang: xiaohu.wang@cuhk.edu.hk.

Keywords: HAR inference, Karhunen-Loève representation, Spurious regression, t-statistics.

It is meaningless to talk about 'confirming' theories when spurious results are so easily obtained. Hendry (1980)

1 Introduction

In a well-cited contribution that emphasized the importance of diagnostic testing in econometrics, David Hendry (1980) highlighted how easy it is to mistake spurious relationships as genuine when using trending data of the type that are so commonly encountered in econometric work, especially in macroeconomics. Spurious regressions occur when conventional significance tests are so seriously biased towards rejection of the null hypothesis of no relationship that the alternative of a genuine relationship is accepted when the variables have no meaningful relationship and may even be statistically independent. Hendry's article showcased the potential for nonsense regressions with the illustration of a regression between UK consumer prices and cumulative rainfall that displayed a high level of 'significance' and passed many - but not all - diagnostic tests.

Spurious regressions continue to attract considerable attention in econometric work, long after the original study by Yule (1926), the simulation experiments of Granger and Newbold (1974), and cautionary warnings made by David Hendry and many other writers since then. The limit theory of Phillips (1986) and Durlauf and Phillips (1988) provided the first analytic step forward by explaining the phenomena of persistent null hypothesis rejections in spurious regressions. These studies helped applied researchers understand the failure of conventional significance tests by showing that in regressions with independent or even correlated trending I(1) data the usual regression t- and F-ratio test statistics do not possess limiting distributions but actually diverge as the sample size $n \uparrow \infty$, leading inevitably to rejections of the null of no association. These studies formed the basis of a large subsequent literature that has analyzed spurious regressions among various classes of trend stationary, long memory, nonstationary, and near-nonstationary time series. A recent article by Ernst et al. (2017) provided further analysis by deriving an expression for the standard deviation of the sample correlation coefficient between two independent standard Brownian motions. While this expression does not explain the phenomenon of spurious regression betwen two independent random walks, it does reveal that the limiting correlation is not centred on the origin and is highly dispersed. This result complements the finding in Phillips (1986) and many subsequent papers that the coefficient of determination in a spurious regression has a well defined limit distribution and does not converge in probability to zero.

In later work, Phillips (1998) pointed out that spurious regressions typically reflect the fact that trending data may always be 'explained' by a coordinate system of other trending variables - which includes the example of UK price series being well-explained by cumulative rainfall that was used by David Hendry (1980). In this broad sense of interpretation, there are no spurious regressions for trending time series, just alternative 'valid' representations of the time series trajectories (and those of its limiting stochastic process, given a suitable normalization) in terms of other stochastic processes and deterministic functions of time.

The asymptotic theory in Phillips (1998) utilized the general representation of a stochastic process in terms of an orthonormal system and provided an extension of the Weierstrass theorem to include the approximation of continuous functions and stochastic processes by Wiener processes. That theory was applied to two classic examples of spurious regressions: regression of stochastic trends on time polynomials, and regressions among independent random walks. Such regressions were shown to reproduce asymptotically in part (and in whole as the regressor space expanded with sample size) the underlying valid representations of one trending process in terms of others, a coordinate system that is entirely analogous to orthonormal or Fourier series representations of a continuous function in terms of polynomials or other simple classes of functions over some interval. An important feature of these 'valid' trend relationships is that the coefficients in the representations, like those in the Karhunen-Loève representation of a general stochastic process, are themselves random variables. Randomness in the representation of time series trajectories is embodied in these coefficients. Much subsequent work has utilized these ideas and methods, either in justifying certain regression representations or in using partial versions of these regression representations to focus on certain features – such as long run features – of the data (Phillips, 2005, 2014; Müeller, 2007; Sun, 2004, 2014a,2014b, 2014c; Hwang and Sun, 2018; Müller and Watson, 2016, 2018).

An important element in Hendry's (1980) discussion of econometric practice was its emphasis on the value of diagnostic testing to ascertain limitations of regressions used in applications. In any empirical regression equation, the properties of the residuals depend inevitably on the properties of the data. To build upon a saying of the famous statistician John Tukey, in the regression equation $y = X\beta + u$ the empirical investigator chooses the variables y and X (possibly with the aid of an autometric regression or a machine learning algorithm) and god gives back u. Any misspecification in the relationship between y and X must therefore be manifest in the properties of u. This is precisely what occurs in a spurious regression – the residual embodies the consequences of a model's fundamental error of specification – as is revealed by the fact that tests for residual serial correlation such as the Durbin Watson statistic converge in probability to zero in such regressions (Phillips, 1986).

Accommodating departures in fitted relationships from conventional assumptions on the properties of regression errors and thereby some of the effects of misspecification has been a longstanding goal of econometrics. One of the great advances in econometric research over the last half century in response to this goal has been the development of methods of inference that are robust to some of the properties of the data and, particularly, those of the regression error. Such robustness can offer protection against specification error in validating inference. This research has led to the progressive development of heteroskedastic and autocorrelation consistent (HAC[1]) procedures and subsequently to heteroskedastic and autocorrelation robust (HAR[2]) methods. These methods control for the effects of serial dependence and heterogeneity in regression errors and they play a key role in achieving robustness in inference. One area where methods of achieving valid statistical inference via HAC procedures has proved especially important in practice are regressions that involve trending variables and cointegration. This goal motivated the early research on optimal semiparametric approaches to the estimation of cointegrating relationships (Phillips and Hansen, 1990) and continues to play a role in subsequent developments in this field (Phillips, 2014; Hwang and Sun, 2018).

HAC methods generally have good asymptotic properties but they are susceptible to large size distortions in practical work. Several alternative methods have been proposed in the recent literature to improve finite sample performance. Among these, the 'fixed-b' lag truncation rule (Kiefer and Vogelsang, 2002a, 2002b, 2005) has attracted considerable interest. The method uses a truncation lag M for including sample serial covariances that is proportional to the sample size n (i.e., $M \sim bn$ for some fixed $b \in (0,1)$) and sacrifices consistent variance matrix (and hence standard error) estimation in the interest of achieving improved performance in statistical testing by mirroring finite sample characteristics of test statistics in the new asymptotic theory of these tests. The formation of t ratio and Wald statistics based on HAC estimators without truncation belongs to the more general class of HAR test statistics. There are known analytic advantages to the fixed b approach, primarily related to controlling size distortion. In particular, research by Jansson (2004), Sun et al. (2009), and Sun (2014b) has shown evidence from Edgeworth expansions of enhanced higher order asymptotic size control in the use of these tests. Recently, Müller (2014), Lazarus, et al. (2018), and Sun (2018) have surveyed work in this literature and given recommendations for practical implementation.

[1] Heteroskedastic robust standard errors were introduced by Eicher (1967), Huber (1967) and White (1980). HAC estimators were introduced by White (1982) and have a long subsequent history of enhancement.

[2] Heteroskedastic and autocorrelation robust standard errors were introduced in Kiefer and Vogelsan (2002a, 2002b) and, following this lead, Phillips (2005a) used the HAR terminology to characterize a class of robust inferential procedures in an article concerned with the development of automated mechanisms of valid inference in econometrics. Other important early contributions concerning HAC covariance matrix estimators without truncation are given by Robinson (1998), Kiefer Vogelsang and Bunzel (2000), and Kiefer and Vogelsang (2005).

In studying spurious regression on trend phenomena, Phillips (1998) showed that the use of HAC methods attenuated the misleading divergence rate (under the null hypothesis of no association) by the extent to which the truncation lag $M \to \infty$. In particular, the divergence rate of the t statistic in a spurious regression involving independent $I(1)$ variables is $O_p\left(\sqrt{n/M}\right)$ rather than $O_p(\sqrt{n})$. Pursuing this philosophy further, Sun (2004) offered a new solution to deal with inference in spurious regressions. He argued that the divergence of the usual t-statistic arises from the use of a standard error estimator that underestimates the true variation of the ordinary least squares (OLS) estimator. He proposed use of a fixed-b HAR standard error estimator with a bandwidth proportional to the sample size (where $M \sim bn \to \infty$ at the same rate as n). The resulting t-statistic converges to a non-degenerate limiting distribution which depends on nuisance parameters. These discoveries revealed that prudent use of HAR techniques in regression testing might widen the range of inference to include spurious regression.

In the same spirit as Sun (2004, 2014), the present contribution analyzes possible advantages in using HAR test statistics in the context of simple trend regressions such as

$$x_t = at + u_t, \tag{1.1}$$

where u_t is $I(1)$. For trend assessment in models of this type it is of interest to test the null hypothesis $H_0 : a = 0$ of the absence of a deterministic trend in (1.1). This framework is a prototypical example of much more complex models where deterministic and stochastic trend components are present and valid testing is needed.

The paper considers three types of t test widely used in econometrics: the usual t test, the t test based on HAC covariance matrix estimators, and the fixed-b HAR test. We apply these t-statistics to three classic examples of spurious regressions: regression of stochastic trends on time polynomials, regression of stochastic trends on deterministic time trend and regression among independent random walks. The asymptotic behavior of these three different t-statistics are investigated. In the regression of stochastic trends on time polynomials and the regression among independent random walks, it is shown that the usual t test and HAC based t test are likely to indicate a significant relation with probability that goes to one as the sample size n goes to infinity. However, provided the number of regressors (K) is fixed, the HAR t-statistics converge to well-defined distributions free from nuisance parameters. As a result, when appropriate critical values are drawn from these limiting distributions, the HAR t-statistics would not diverge and valid inference on the regression coefficients would be possible, concordant with Sun (2004).

In contrast to these results and those of Sun (2004), we find that HAR t-statistics diverge at rate $\sqrt{K}$ as $K \to \infty$. Hence, the characteristics of spurious regression return even with the use of HAR test statistics in models with an increasing number of regressors. These findings seem relevant for machine learning and autometric model building methods which accommodate large numbers of regressors, including those of the $p > n$ variety where model searching often begins with more regressors than sample observations and penalized methods of estimation are needed to obtain even preliminary results.

Our results also reveal that the other two t-statistics (the usual t and HAC-based t) diverge at greater rates when $K \to \infty$ than when K is fixed. In the regression of stochastic trends on deterministic time trends, we derive the limiting distributions of the statistics under both the null and alternative hypotheses. The HAR test turns out to be the only test which is consistent and has controllable size. All the limit theory for these tests receives strong support in simulations. And, as will become evident, the appealing asymptotic properties of the HAR test in the fixed number of regressors case are manifest even in situations where some commonly-used regularity conditions in the construction of HAR tests are violated.

The rest of the paper is organized as follows. Section 2 examines regressions of stochastic trends on a complete orthonormal basis in $L_2[0,1]$ and establishes the limiting distributions of the three different t-statistics with explicit application to the prototypical case of a spurious linear trend regression. Section 3 examines the limit behavior of the t-statistics in regressions among independent random walks. Simulations are reported in Section 4. Section 5 concludes. All proofs are given in the Appendix.

2 Regression of Stochastic Trend on Time Polynomials

2.1 Model Details and Background

The development in this section concentrates on a simple unit root time series

$$X_t = \sum_{s=1}^{t} \mu_s, \tag{2.1}$$

whose increments μ_t form a stationary time series with zero mean, finite absolute moments to order $p > 2$, and continuous spectral density function $f_\mu(\lambda)$. We assume that X_t satisfies the functional central limit theorem (FCLT)

$$\frac{X_{\lfloor nr \rfloor}}{\sqrt{n}} \Rightarrow B(r) \equiv BM\left(\omega^2\right), \text{ with } \omega^2 = 2\pi f_\mu(0), \tag{2.2}$$

for which primitive conditions are well known (e.g., Phillips and Solo, 1992). The results that follow are illustrative and apply with suitable modification to more general nonstationary time series, such as near integrated or long memory series, which upon standardardization converge to limiting stochastic processes with sample paths that are continuous almost surely.

By the Karhunen- Loève (KL) expansion theorem (e.g., Loève, 1963, p.478), any function that is continuous in quadratic mean has a decomposition into a countable linear combination of orthogonal functions. The KL representation for the Brownian motion $B(r)$ is

$$B(r) = \sum_{k=1}^{\infty} \sqrt{\lambda_k} \varphi_k(r) \xi_k = \omega\sqrt{2} \sum_{k=1}^{\infty} \frac{\sin[(k-1/2)\pi r]}{(k-1/2)\pi} \xi_k, \tag{2.3}$$

where

$$\lambda_k = \frac{4\omega^2}{(2k-1)^2 \pi^2}, \quad \varphi_k(r) = \sqrt{2}\sin[(k-1/2)\pi r]$$

are eigenvalues and corresponding eigenfunctions of the Brownian motion's covariance kernel $\omega^2 (r \wedge s)$, and

$$\xi_k = \lambda_k^{-1/2} \int_0^1 B(s) \varphi_k(s) \, ds$$

are independently and identically distributed (iid) as $N(0,1)$. This series representation of $B(r)$ is convergent almost surely and uniformly in $r \in [0,1]$. Denoting $z_k = \sqrt{\lambda_k}\xi_k$ as the stochastic coefficients, the KL representation (2.3) could be rewritten as

$$B(r) = \sum_{k=1}^{\infty} z_k \varphi_k(r). \tag{2.4}$$

Starting from the KL representation of $B(r)$, Phillips (1998) studied the asymptotic properties of regressions of X_t on deterministic regressors of the type

$$X_t = \sum_{k=1}^{K} \hat{b}_k \varphi_k\left(\frac{t}{n}\right) + \hat{u}_t, \tag{2.5}$$

or, equivalently (with $\hat{a}_k = \hat{b}_k/\sqrt{n}$),

$$\frac{X_t}{\sqrt{n}} = \sum_{k=1}^{K} \hat{a}_k \varphi_k\left(\frac{t}{n}\right) + \frac{\hat{u}_t}{\sqrt{n}}. \tag{2.6}$$

Least squares estimation gives

$$\hat{\alpha}_K = (\hat{a}_1, ..., \hat{a}_K)' = \left(\Phi_K' \Phi_K\right)^{-1} \Phi_K' X/\sqrt{n},$$

where $\Phi_K = (\varphi_{K1}, ..., \varphi_{Kn})'$ with $\varphi_{Kt} = (\varphi_1(t/n), ..., \varphi_K(t/n))'$, and $X = (X_1, ..., X_n)'$. Let $C_K \in \mathbb{R}^K$ be any vector with $C_K' C_K = 1$. When K is fixed and $n \to \infty$, Phillips (1998) proved that

$$C_K' \hat{\alpha}_K \Rightarrow C_K' \int_0^1 \bar{\varphi}_K(r) B(r) dr \equiv N\left(0, C_K' \Lambda_K C_K\right),$$

where $\Lambda_K = \text{diag}(\lambda_1, ..., \lambda_K)$ and $\bar{\varphi}_K(r) = (\varphi_1(r), ..., \varphi_K(r))'$. In the expanding regressor case where $K = K(n) \to \infty$ and $K/n \to 0$, it was also shown in Phillips (1998) that

$$C_K' \hat{\alpha}_K \Rightarrow N\left(0, \sigma_c^2\right) \equiv c' Z = \sum_{k=1}^{\infty} c_k z_k,$$

where $c = (c_k) \in \mathbb{R}^\infty$ satisfies $c'c = 1$, $\Lambda = \text{diag}(\lambda_1, \lambda_2, \cdots)$, $\sigma_c^2 = c' \Lambda c$, and $Z = (z_k)_{k=1}^\infty$ are the random coefficients in the KL representation (2.4). Therefore, the fitted coefficients in regression (2.6) tend to random variables in the limit as $n \to \infty$ that match those in the KL representation of the limit process $B(\cdot)$. In other words, least squares regressions reproduce in part (when K is finite) and in whole (when $K \to \infty$) the underlying orthonormal representations.

2.2 Three t-statistics

Suppose interest centers on testing whether the regression coefficients are significant or more generally whether some linear combination $C_K' \beta_K$ of the underlying coefficients $\beta_K = (b_1, ..., b_K)'$ in the estimated regression (2.5) is equal to 0, that is

$$H_0 : C_K' \beta_K = 0 \text{ v.s. } H_1 : C_K' \beta_K \neq 0.$$

Three types of t-statistics are considered. The first is the usual t-ratio defined as

$$t_{C_K' \beta_K} = \frac{C_K' \hat{\beta}_K}{\left[s_b^2 C_K' \left(\Phi_K' \Phi_K\right)^{-1} C_K\right]^{1/2}} \tag{2.7}$$

with $s_b^2 = n^{-1} \sum_{t=1}^n \hat{u}_t^2 = n^{-1} \sum_{t=1}^n \left(X_t - \hat{\beta}_K' \varphi_{Kt}\right)^2$ the usual error variance estimate. The second t-statistic is constructed by using a HAC variance estimator and has the following representation

$$t^{HAC}_{C'_K\beta_K} = \frac{C'_K\hat{\beta}_K}{\hat{\omega}_{C'_K\beta_K}}, \tag{2.8}$$

where

$$\hat{\omega}^2_{C'_K\beta_K} = C'_K\left(\Phi'_K\Phi_K\right)^{-1}\left[n\widehat{\text{lrvar}}_{HAC}\left(\hat{u}_t\varphi_{Kt}\right)\right]\left(\Phi'_K\Phi_K\right)^{-1}C_K, \tag{2.9}$$

with

$$\widehat{\text{lrvar}}_{HAC}\left(\eta_t\right) = \sum_{j=-M}^{M} k\left(\frac{j}{M}\right)c\left(j,\eta\right),\ c\left(j,\eta\right) = \frac{1}{n}\sum_{1\leq t,t+j\leq n}\eta_t\eta'_{t+j}. \tag{2.10}$$

Here, $\widehat{\text{lrvar}}_{HAC}\left(\eta_t\right)$ is a kernel estimate of the long run variance of its argument, $k\left(\cdot\right)$ is a lag kernel, M is a bandwidth parameter satisfying $M/n+1/M\to 0$ as $n\to\infty$, and the argument $\eta_t = \hat{u}_t\varphi_{Kt}$ in (2.9).

If we choose a fixed $b\in(0,1]$ and set $M=\lfloor bn\rfloor$, the condition $M/n+1/M\to 0$ as $n\to\infty$ is violated. In that case, the long run variance estimate is a fixed-b estimate and leads to the HAR t-statistic

$$t^{HAR}_{C'_K\beta_K} = \frac{C'_K\hat{\beta}_K}{\check{\omega}_{C'_K\beta_K}}, \tag{2.11}$$

where

$$\check{\omega}^2_{C'_K\beta_K} = C'_K\left(\Phi'_K\Phi_K\right)^{-1}\left[n\widehat{\text{lrvar}}_{HAR}\left(\hat{u}_t\varphi_{Kt}\right)\right]\left(\Phi'_K\Phi_K\right)^{-1}C_K, \tag{2.12}$$

with

$$\widehat{\text{lrvar}}_{HAR}\left(\eta_t\right) = \sum_{j=-(n-1)}^{(n-1)} k_b\left(\frac{j}{n}\right)c\left(j,\eta\right),\ c\left(j,\eta\right) = \frac{1}{n}\sum_{1\leq t,t+j\leq n}\eta_t\eta'_{t+j}, \tag{2.13}$$

$k_b\left(\frac{j}{n}\right) = k\left(\frac{j}{nb}\right)$, and and $k\left(\cdot\right)$ is a lag kernel function as before.

With minor changes of the proof given in Phillips (1998), it is easy to deduce that for fixed K, $t_{C'_K\beta_K}\sim O_p\left(\sqrt{n}\right)$ and $t^{HAC}_{C'_K\beta_K}\sim O_p\left(\sqrt{n/M}\right)$ as $n\to\infty$, as discussed earlier. Therefore, such tests indicate statistically significant regression coefficients with probability that goes to one as $n\to\infty$. These results match what is now standard spurious regression limit theory for inference.

In addition, as we show in Theorem 2.3 below, the large regressor case where $K\to\infty$ leads to different results. In this case, both t-statistics $t_{C'_K\beta_K}$ and $t^{HAC}_{C'_K\beta_K}$ have greater rates of divergence that depend on the expansion rate of K, given by $t_{C'_K\beta_K} = O_p\left(\sqrt{nK}\right)$ and $t^{HAC}_{C'_K\beta_K} = O_p\left(\sqrt{nK/M}\right)$. Thus, with the addition of more regressors the combined effect of the regression coefficients – as well as that of the individual coefficents – appears more significant and diverges when $K\to\infty$ as $n\to\infty$. In consequence, large numbers of regressors effectively worsen the spurious regression problem.

Is there a test which does not always indicate that coefficients $\hat{\beta}_K$ are significant in the "spurious" regression (2.5)? As the results of Sun (2004) show, the answer is positive for the case where K is fixed. In this event, the HAR test is appealing in the sense that $t^{HAR}_{C'_K\beta_K} \sim O_p(1)$ when $n \to \infty$ and K is fixed, so that test size is controlled in the limit. Therefore, when appropriate critical values obtained from the limit distribution of $t^{HAR}_{C'_K\beta_K}$ are employed, the coefficients $\hat{\beta}_K$ do not inevitably signal significance as $n \to \infty$ and the usual misleading test implications of spurious regression do not manifest. However, in the important case where the regressor space expands and $K \to \infty$, the test statistic $t^{HAR}_{C'_K\beta_K}$ diverges to infinity at rate $O_p\left(\sqrt{K}\right)$ and the coefficients $\hat{\beta}_K$ become significant again even under HAR testing.

These results are collected in the following two theorems.

Theorem 2.1 *For fixed K, as $n \to \infty$ and $M/n + 1/M \to 0$, we have*

(i)

$$\frac{t_{C'_K\beta_K}}{\sqrt{n}} \Rightarrow \frac{C'_K Z_K}{\left[\int_0^1 \underline{B}^2_{\varphi_K}\right]^{1/2}};$$

(ii)

$$\sqrt{\frac{M}{n}}\, t^{HAC}_{C'_K\beta_K} \Rightarrow \frac{C'_K Z_K}{\left\{\int_{-1}^1 k(s)\, ds \int_0^1 \underline{B}^2_{\varphi_K}\left(C'_K\bar{\varphi}_K\right)^2\right\}^{1/2}};$$

(iii)

$$\begin{aligned} t^{HAR}_{C'_K\beta_K} &\Rightarrow \frac{C'_K Z_K}{\left\{\int_0^1\int_0^1 k_b(r-q)\,\underline{B}_{\varphi_K}(r)\,\underline{B}_{\varphi_K}(q)\left[C'_K\bar{\varphi}_K(r)\right]\left[C'_K\bar{\varphi}_K(q)\right] drdq\right\}^{1/2}} \\ &\equiv \frac{C'_K Z^W_K}{\left\{\int_0^1\int_0^1 k_b(r-q)\,\underline{W}_{\varphi_K}(r)\,\underline{W}_{\varphi_K}(q)\left[C'_K\bar{\varphi}_K(r)\right]\left[C'_K\bar{\varphi}_K(q)\right] drdp\right\}^{1/2}}, \end{aligned}$$

where $Z_K = (z_k)_{k=1}^K$ are the random coefficients in orthonormal representation (2.4), $\underline{B}_{\varphi_K}(r) = B(r) - Z'_K\bar{\varphi}_K(r)$, $Z^W_K = Z_K/\omega = \int_0^1 W\bar{\varphi}_K$, $W(\cdot) \equiv BM(1)$, $\omega^2 = 2\pi f_\mu(0)$, and $\underline{W}_{\varphi_K}(r) = \underline{B}_{\varphi_K}(r)/\omega = W(r) - \left(Z^W_K\right)'\bar{\varphi}_K(r)$.

Remark 2.2 *The fixed-b HAR based t-statistic $t^{HAR}_{C'_K\beta_K}$ asymptotically follows a well-defined limit distribution when the number of regressors K is fixed. The limit distribution is free from nuisance parameters and is easily computable but depends on the lag kernel as well as the form of the trend regressors, which influence the detrended standard Brownian motion process*

$\underline{W}_{\varphi_K}$. *The asymptotic critical values therefore differ from those of the usual standard normal limit distribution of a t-statistic. But the specific features of the limit distribution of* $t^{HAR}_{C_K'\beta_K}$, *which retain randomness in the denominator of the limiting statistic, help to control size in finite sample testing.*

Theorem 2.3 *As* $n, K \to \infty$, $M/n + 1/M \to 0$ *and* $K^{5/2}/n + K^{3/2}/n^{\frac{1}{2}-\frac{1}{p}} \to 0$, *the following results hold:*

(i)

$$\frac{t_{C_K'\beta_K}}{\sqrt{nK}} = \frac{C_K' Z_K}{\left[K \int_0^1 \underline{B}^2_{\varphi_K}\right]^{1/2}} + o_p(1) = O_p(1),$$

where $K \int_0^1 \underline{B}^2_{\varphi_K} = \omega^2/\pi^2 + o_p(1)$.

(ii)

$$\sqrt{\frac{M}{nK}} t^{HAC}_{C_K'\beta_K} = \frac{C_K' Z_K}{\left[K \int_{-1}^1 k(s)\, ds \int_0^1 B^2_{\varphi_K} \left[C_K' \bar{\varphi}_K\right]^2\right]^{1/2}} + o_p(1) = O_p(1),$$

(iii)

$$\frac{t^{HAR}_{C_K'\beta_K}}{\sqrt{K}} = \frac{C_K' Z_K}{\left\{K \int_0^1 \int_0^1 k_b(r-q)\, \underline{B}_{\varphi_K}(r)\, \underline{B}_{\varphi_K}(q)\, C_K' \bar{\varphi}_K(r)\, C_K' \bar{\varphi}_K(q)\, dr dq\right\}^{1/2}} + o_p(1) = O_p(1).$$

Remark 2.4 *Theorem 2.3 shows that all three t-statistics diverge as* $n \to \infty$ *but at different rates, each of which depends on* K. *The divergence rate of the fixed-b test statistic* $t^{HAR}_{C_K'\beta_K} = O_p\left(\sqrt{K}\right)$ *is the slowest and depends only on* K. *These results strengthen the finding in Phillips (1998) that attempts to deal with serial dependence in controlling size in significance testing generally fail when enough effort is put into the regression design to fit the trajectory. This failure now includes HAR testing when* $K \to \infty$. *All the tests are therefore ultimately confirmatory of the existence of a 'relationship' – in the present case a coordinate representation relationship among different types of trends, at least when a complete representation is attempted by allowing the number of regressors* K *to diverge with* n. *The results of the theorem may be interpreted to mean that when a serious attempt is made to model a stochastic trend using deterministic functions (either a large number of such regressors or regressors that are carefully chosen to provide a successful representation and trajectory fit) it will end up being successful even when a spurious regression robust method such as fixed-b HAR test is used.*

An additional matter concerning the form of these tests may usefully be highlighted. To construct the HAC and HAR t-statistics, the following condition

$$\mathrm{Var}\left(\frac{\Phi_K' X}{\sqrt{n}}\right) = \mathrm{Var}\left(\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\varphi_{Kt}X_t\right) = \Gamma_0 + \sum_{j=1}^{n-1}\left(1-\frac{j}{n}\right)\left(\Gamma_j+\Gamma_j'\right) \tag{2.14}$$

with $\Gamma_j = E\left(\varphi_{Kt}X_tX_{t-j}'\varphi_{K(t-j)}'\right)$ is usually imposed (e.g., Kiefer et al. 2000, Kiefer and Vogelsang 2002a, 2002b) as in standard approaches to robust covariance matrix estimation. In other words, the process $\{\varphi_{Kt}X_t\}$ is typically assumed to be unconditionally stationary or weakly dependent with uniformly bounded second moments so that series such as (2.14) converge. However, this condition is violated in both regressions (2.5) and (2.6) as

$$\mathbb{E}\left(\varphi_{Kt}X_tX_{t-j}'\varphi_{K(t-j)}'\right) = \varphi_{Kt}\mathbb{E}\left(\sum_{s=1}^{t}\mu_s\sum_{\tau=1}^{t-j}\mu_\tau\right)\varphi_{K(t-j)}'$$

depends on t. For example, when the components μ_s are $iid\left(0,\sigma^2\right)$ with partial sums satisfying (2.2) then

$$\mathbb{E}\left(\varphi_{Kt}X_tX_{t-j}'\varphi_{K(t-j)}'\right) = (t-j)\,\sigma^2\varphi_{Kt}\varphi_{K(t-j)}'$$

depends on t. Regardless of this violation, HAC and HAR t-statistics may still be constructed in the traditional way; and the HAR statistic, $t_{C_K'\beta_K}^{HAR}$ has nuisance parameter free asymptotic properties even though the above unconditional stationarity condition is not satisfied.

The above results apply straightforwardly to the simple case of a spurious linear regression on trend where the time series is a unit root process generated by

$$X_t = at + X_t^0,\ t = 1,...,n, \tag{2.15}$$

with $a=0$ and $X_t^0 = \sum_{s=1}^{t}\mu_s$ is the partial sum of a zero mean stationary process $\{\mu_s\}$ with continuous spectral density $f_\mu(\lambda)$. The standardized process $X_n(r) = n^{-1/2}X_{\lfloor nr\rfloor}^0$ satisfies the functional law

$$X_n(r) \Rightarrow B(r) \equiv BM\left(\omega^2\right),\ \omega^2 = 2\pi f_\mu(0) > 0.$$

The fitted regression model is

$$X_t = \hat{a}t + \hat{u}_t, \quad \text{or equivalently,} \quad \frac{X_t}{\sqrt{n}} = \left(\sqrt{n}\hat{a}\right)\frac{t}{n} + \frac{\hat{u}_t}{\sqrt{n}}, \tag{2.16}$$

where $\hat{a} = \sum_{t=1}^{n} tX_t / \sum_{t=1}^{n} t^2$ is the least squares (LS) estimate of a, which satisfies (Durlauf and Phillips, 1988)

$$\sqrt{n}\left(\hat{a} - a\right) = \frac{n^{-5/2}\sum_{t=1}^{n} tX_t^0}{n^{-3}\sum_{t=1}^{n} t^2} \Rightarrow 3\int_0^1 rB\left(r\right)dr \equiv N\left(0, \frac{6}{5}\omega^2\right), \tag{2.17}$$

so that $\hat{a}$ is consistent, including the case where $a = 0$. However, as is well known, the usual t-statistic has order $O_p\left(\sqrt{n}\right)$ and diverges as $n \to \infty$, indicating a significant relationship between $\{X_t\}$ and t in spite of the fact that $a = 0$. This outcome follows directly from Theorem 2.1 and the (alternate) representation for the standard Brownian motion $W\left(r\right)$ as

$$W\left(r\right) = r\xi_0 + \sqrt{2}\sum_{k=1}^{\infty}\frac{\sin\left[k\pi r\right]}{k\pi}\xi_k, \text{ with } \xi_k \equiv \text{iid } N\left(0,1\right), \tag{2.18}$$

which implies that

$$n^{-1/2}X_{\lfloor nr \rfloor}^0 \Rightarrow B\left(r\right) = \omega \cdot W\left(r\right) = \left(\omega\xi_0\right)r + \left(\omega\xi_k\right)\sqrt{2}\sum_{k=1}^{\infty}\frac{\sin\left[k\pi r\right]}{k\pi}.$$

Thus, when $a = 0$, the scaled LS estimator $\sqrt{n}\hat{a}$ has a random limit $\xi_a \equiv N\left(0, \frac{6}{5}\omega^2\right)$ from (2.17) that approximates but does not exactly reproduce the leading random coefficient term $\left(\omega\xi_0\right)$ in the representation (2.18). Importantly in this case, the deterministic functions in (2.18) are not orthonormal and there is dependence in $L_2\left[0,1\right]$ between the functions r and $\left\{\left(\sqrt{2}\sin\left[k\pi r\right]\right)/\left(k\pi\right)\right\}_{k=1}^{\infty}$. This dependence induces an asymptotic inefficiency in the trend coefficient estimate $\hat{a}$, since $\frac{6}{5}\omega^2 >\text{Var}(\omega\xi_0) = \omega^2$.

Next, in testing $H_0 : a = 0$ versus $H_1 : a \neq 0$, the following statistics are considered:

$$t_a = \frac{\hat{a}}{s_a} = \frac{\hat{a}}{\left\{\left[n^{-1}\sum_{t=1}^{n}\left(\hat{u}_t\right)^2\right]\left(\sum_{t=1}^{n} t^2\right)^{-1}\right\}^{1/2}}, \tag{2.19}$$

$$t_a^{HAC} = \frac{\hat{a}}{\hat{\omega}_a} = \frac{\hat{a}}{\left\{\left(\sum_{t=1}^{n} t^2\right)^{-1}\left[n\widehat{\text{lrvar}}_{HAC}\left(t\hat{u}_t\right)\right]\left(\sum_{t=1}^{n} t^2\right)^{-1}\right\}^{1/2}}, \tag{2.20}$$

$$t_a^{HAR} = \frac{\hat{a}}{\check{\omega}_a} = \frac{\hat{a}}{\left\{\left(\sum_{t=1}^{n} t^2\right)^{-1}\left[n\widehat{\text{lrvar}}_{HAR}\left(t\hat{u}_t\right)\right]\left(\sum_{t=1}^{n} t^2\right)^{-1}\right\}^{1/2}}, \tag{2.21}$$

where

$$\widehat{\mathrm{lrvar}}_{HAC}\left(t\hat{u}_t\right) = \sum_{j=-M}^{M} k\left(\frac{j}{M}\right)\left[\frac{1}{n}\sum_{1\le t,t+j\le M}\hat{u}_t\hat{u}_{t+j}t\left(t+j\right)\right], \text{ with } M/n+1/M \to 0 \text{ as } n\to\infty,$$

$$\widehat{\mathrm{lrvar}}_{HAR}\left(t\hat{u}_t\right) = \sum_{j=-(n-1)}^{(n-1)} k_b\left(\frac{j}{n}\right)\left[\frac{1}{n}\sum_{1\le t,t+j\le M}\hat{u}_t\hat{u}_{t+j}t\left(t+j\right)\right], \text{ for some fixed } b\in(0,1],$$

$k(\cdot)$ is a kernel function, $k_b(j/n) = k(j/(nb))$ and $\hat{u}_t = X_t - \hat{a}t$ for $t = 1,\cdots,n$. The asymptotic properties of these test statistics follow in the same way as before when $n \to \infty$ with $M/n + 1/M \to 0$, giving the following results.

(i) Under $H_0 : a = 0$,

$$\frac{t_a}{\sqrt{n}} \Rightarrow \frac{\sqrt{3}\int_0^1 rB}{\left\{\int_0^1 \underline{B}^2\right\}^{1/2}}, \tag{2.22}$$

$$\sqrt{\frac{M}{n}}t_a^{HAC} \Rightarrow \frac{\int_0^1 rB}{\left\{\int_{-1}^1 k(s)\,ds\int_0^1 r^2\underline{B}^2\right\}^{1/2}}, \tag{2.23}$$

$$t_a^{HAR} \Rightarrow \frac{\int_0^1 rB}{\left\{\int_0^1\int_0^1 k_b(r-q)\,\underline{B}(r)\,\underline{B}(q)\,rqdrdq\right\}^{1/2}} \equiv \frac{\int_0^1 rW}{\left\{\int_0^1\int_0^1 k_b(r-q)\,\underline{W}(r)\,\underline{W}(q)\,rqdrdq\right\}^{1/2}}; \tag{2.24}$$

(ii) Under $H_1 : a \neq 0$,

$$\frac{t_a}{n} \Rightarrow \frac{a}{\left(3\int_0^1 \underline{B}^2\right)^{1/2}}, \tag{2.25}$$

$$\frac{\sqrt{M}}{n}t_a^{HAC} \Rightarrow \frac{a}{\left[9\int_{-1}^1 k(s)\,ds\int_0^1 r^2\underline{B}^2\right]^{1/2}}, \tag{2.26}$$

$$\frac{t_a^{HAR}}{\sqrt{n}} \Rightarrow \frac{a}{\left[9\int_0^1\int_0^1 k_b(r-q)\,\underline{B}(r)\,\underline{B}(q)\,rqdrdq\right]^{1/2}}, \tag{2.27}$$

where $\underline{B}(r) := B(r) - 3\left(\int_0^1 sB(s)\,ds\right)r$ and $B(r) \equiv \omega W(r)$. Thus, under the null hypothesis both $t_a = O_p(\sqrt{n})$ and $t_a^{HAC} = O_p\left(\sqrt{n/M}\right)$ diverge but $t_a^{HAR} = O_p(1)$ and has a well defined nuisance parameter free limit distribution that may be used in statistical testing. Under the alternative hypothesis, all the listed statistics are divergent but at different rates. Only t_a^{HAR}

has effective discriminatory power, being consistent and having controllable size. These results match those in Sun (2004, 2014) showing that for simple trend misspecifications like that of a finite degree polynomial trend function in place of a stochastic trend, use of fixed-b HAR testing controls size and leads to a consistent test.

3 Regressions among independent random walks

This section extends these ideas to regressions among independent random walks. Let $B(\cdot)$ be a Brownian motion on the interval $[0,1]$. Phillips (1998) proved that there exist a sequence of independent standard Brownian motions $\{W_i\}_{i=1}^K$ that are independent of $B(\cdot)$, and a sequence of variables $\{d_i\}_{i=1}^K$ defined on an augmented probability space $(\Omega, \mathcal{F}, P)$ such that, as $K \to \infty$,

$$B(r) \sim \sum_{i=1}^{\infty} d_i W_i(r) \quad \text{in } L_2[0,1] \quad \text{a.s. } (P). \tag{3.1}$$

The random coefficients d_i are statistically dependent on $B(\cdot)$. Replacing the Wiener processes W_i by orthogonal functions $V_i(r)$ in $L_2[0,1]$ using the Gram-Schmidt process

$$\begin{aligned} V_1 &= W_1, \\ V_2 &= W_2 - \left(\int_0^1 W_2 V_1\right)\left(\int_0^1 V_1^2\right)^{-1} V_1, \\ V_3 &= W_3 - \left(\int_0^1 W_3 V_a'\right)\left(\int_0^1 V_a V_a'\right)^{-1} V_a, \ V_a' = [V_1, V_2], \text{ etc,} \end{aligned}$$

gives the representation

$$B(r) \sim \sum_{i=1}^{\infty} e_i V_i(r), \quad \text{with } \ e_i = \left(\int_0^1 BV_i\right)\left(\int_0^1 V_i^2\right)^{-1}. \tag{3.2}$$

In the following, we consider the unit root process $y_t = \sum_{s=1}^t \mu_s$ with mean zero stationary components $\{\mu_s\}$ with continuous spectral density $f_\mu(\lambda)$ and satisfying the functional law

$$n^{-1/2} y_{\lfloor nr \rfloor} \Rightarrow B(r) \equiv BM\left(\omega^2\right), \ \omega^2 = 2\pi f_\mu(0) > 0.$$

Let $x_t = (x_{kt}) = \left(\sum_{j=1}^t \mu_{kj}\right)_{k=1}^K$ be K independent standard Gaussian random walks, all of which are independent of y_t. Consider the linear regression $y_t = \hat{b}_x' x_t + \hat{u}_t$, based on $n > K$ observations of these series. The large n asymptotic behavior of $\hat{b}_x$ is (Phillips (1986))

$\hat{b}_x \Rightarrow \left(\int_0^1 W_x W_x'\right)^{-1} \left(\int_0^1 W_x B\right)$, where W_x is the vector standard Brownian motion weak limit of the standardized partial sum processes $n^{-1/2} x_{\lfloor n \cdot \rfloor}$.

Suppose we orthogonalize the regressors $\{x_{k\cdot} = (x_{kt})_{t=1}^n : k = 1, \cdots, K\}$ using the Gram-Schmidt process

$$\begin{aligned} z_{1t} &= x_{1t}, \\ z_{2t} &= x_{2t} - \left(x_{2\cdot}' x_{1\cdot}\right) \left(x_{1\cdot}' x_{1\cdot}\right)^{-1} x_{1t}, \\ z_{3t} &= x_{3t} - \left(x_{3\cdot}' X_a\right) \left(X_a' X_a\right)^{-1} x_{at}, \ X_a := [x_{1\cdot}, x_{2\cdot}] := \left[x_{a\cdot}'\right], \text{ etc.} \end{aligned}$$

By standard weak convergence arguments we have

$$n^{-1/2} z_{1\lfloor n \cdot \rfloor} \Rightarrow V_1(\cdot), \quad n^{-1/2} z_{2\lfloor n \cdot \rfloor} \Rightarrow V_2(\cdot), \quad n^{-1/2} z_{3\lfloor n \cdot \rfloor} \Rightarrow V_3(\cdot), \text{ etc.}$$

Now let $z_t = (z_{kt})_{k=1}^K$, and consider the regression

$$y_t = \hat{b}_{zK}' z_t + \hat{u}_t \tag{3.3}$$

The LS estimator $\hat{b}_{zK} = [\sum_{t=1}^n z_t z_t']^{-1} \sum_{t=1}^n z_t y_t$ has the limit

$$\hat{b}_{zK} \Rightarrow \left(\int_0^1 \bar{V}_K \bar{V}_K'\right)^{-1} \left(\int_0^1 \bar{V}_K B\right) \equiv E_K := (e_k)_{k=1}^K.$$

where $\bar{V}_K = (V_k)_{k=1}^K$ be a $K \times 1$ vector. Thus, the empirical regression of y_t on z_t reproduces the first K terms in the representation of the limit Brownian motion B in terms of an orthogonalized coordinate system formed from K independent standard Brownian motions.

Suppose now that we are interested in testing whether a linear combination of b_{zK} equals zero, viz.,

$$H_0 : C_K' b_{zK} = 0 \text{ v.s. } H_1 : C_K' b_{zK} \neq 0,$$

with $C_K \in \mathbb{R}^K$ satisfying $C_K' C_K = 1$. Again, three types of t-statistics are considered:

$$
\begin{aligned}
t_{b_{zK}} &= \frac{C_K' \hat{b}_{zK}}{s_{b_{zK}}} = \frac{C_K' \hat{b}_{zK}}{\left\{ C_K' \left[n^{-1} \sum_{t=1}^{n} (\hat{u}_t)^2 \right] \left(\sum_{t=1}^{n} z_t z_t' \right)^{-1} C_K \right\}^{1/2}}, && (3.4) \\
t_{b_{zK}}^{HAC} &= \frac{C_K' \hat{b}_{zK}}{\hat{\omega}_{b_{zK}}} = \frac{C_K' \hat{b}_{zK}}{\left\{ C_K' \left(\sum_{t=1}^{n} z_t z_t' \right)^{-1} \left[n \widehat{\text{lrvar}}_{HAC} (z_t \hat{u}_t) \right] \left(\sum_{t=1}^{n} z_t z_t' \right)^{-1} C_K \right\}^{1/2}}, && (3.5) \\
t_{b_{zK}}^{HAR} &= \frac{C_K' \hat{b}_{zK}}{\check{\omega}_{b_{zK}}} = \frac{C_K' \hat{b}_{zK}}{\left\{ C_K' \left(\sum_{t=1}^{n} z_t z_t' \right)^{-1} \left[n \widehat{\text{lrvar}}_{HAR} (z_t \hat{u}_t) \right] \left(\sum_{t=1}^{n} z_t z_t' \right)^{-1} C_K \right\}^{1/2}}, && (3.6)
\end{aligned}
$$

where

$$
\widehat{\text{lrvar}}_{HAC} (z_t \hat{u}_t) = \sum_{j=-M}^{M} k \left(\frac{j}{M} \right) \left[\frac{1}{n} \sum_{1 \le t, t+j \le M} z_t \hat{u}_t \hat{u}_{t+j} z_{t+j}' \right], \text{ with } M/n + 1/M \to 0, \text{ as } n \to \infty,
$$

$$
\widehat{\text{lrvar}}_{HAR} (z_t \hat{u}_t) = \sum_{j=-(n-1)}^{(n-1)} k_b \left(\frac{j}{n} \right) \left[\frac{1}{n} \sum_{1 \le t, t+j \le M} z_t \hat{u}_t \hat{u}_{t+j} z_{t+j}' \right], \text{ for some fixed } b \in (0, 1],
$$

$k(\cdot)$ is a kernel function, $k_b \left(\frac{j}{n} \right) = k \left(\frac{j}{nb} \right)$, and $\hat{u}_t = y_t - \hat{b}_{zK}' z_t$ for $t = 1, \cdots, n$.

The following theorem establishes the limiting distributions of these three t-statistics.

Theorem 3.1 *For fixed K, $n \to \infty$,*

(i)

$$
\frac{1}{\sqrt{n}} t_{b_{zK}} \Rightarrow \frac{C_K' E_K}{\left\{ C_K' \left(\int_0^1 \bar{V}_K \bar{V}_K' \right)^{-1} C_K \int_0^1 \underline{W}_{yK}^2 \right\}^{1/2}};
$$

(ii) When $1/M + M/n \to 0$,

$$
\sqrt{\frac{M}{n}} t_{b_{zK}}^{HAC} \Rightarrow \frac{C_K' E_K}{\left\{ C_K' \left(\int_0^1 \bar{V}_K \bar{V}_K' \right)^{-1} \left(\int_{-1}^1 k(s)\, ds \int_0^1 \underline{W}_{yK}^2 \bar{V}_K \bar{V}_K' \right) \left(\int_0^1 \bar{V}_K \bar{V}_K' \right)^{-1} C_K \right\}^{1/2}};
$$

(iii)

$$
t_{b_{zK}}^{HAR} \Rightarrow \frac{C_K' E_K}{\left\{ C_K' \left(\int_0^1 \bar{V}_K \bar{V}_K' \right)^{-1} H \left(\int_0^1 \bar{V}_K \bar{V}_K' \right)^{-1} C_K \right\}^{1/2}};
$$

where $\underline{W}_{yK}(r) = B(r) - E_K' \bar{V}_K(r)$, $H = \int_0^1 \int_0^1 k_b(r-q) \bar{V}_K(r) \underline{W}_{yK}(r) \underline{W}_{yK}(q) \bar{V}_K'(q)\, dr dq$.

Remark 3.2 *As it is shown in Theorem 3.1,* $t_{b_{zK}}$ *and* $t_{b_{zK}}^{HAC}$ *diverge at rate* $O_p\left(\sqrt{n}\right)$ *and* $O_p\left(\sqrt{n/M}\right)$*, respectively. Hence, such tests indicate inevitable significance of the regressors when* $n \to \infty$ *and* $1/M + M/n \to 0$*. However, the HAR based t-statistic* $t_{b_{zK}}^{HAR}$ *is convergent in distribution, which leads to valid statistical testing when appropriate critical values from the limit distribution of* $t_{b_{zK}}^{HAR}$ *are used. Note that* $B(r) \equiv BM\left(\omega^2\right) \equiv \omega W(r)$, $E_K = (e_k)_{k=1}^{K}$ *with*

$$e_k = \left(\int_0^1 BV_i\right)\left(\int_0^1 V_i^2\right)^{-1} = \omega\left(\int_0^1 WV_i\right)\left(\int_0^1 V_i^2\right)^{-1} \text{ for } k = 1, \cdots, \infty,$$

where $W(\cdot)$ *is a standard Brownian motion. Hence, the nuisance parameter* ω *appearing in the numerator and dominator of the limiting distribution of* $t_{b_{zK}}^{HAR}$ *cancels. The limit distribution of* $t_{b_{zK}}^{HAR}$ *is therefore free of nuisance parameter.*

Remark 3.3 *Even when* $\mu_s \sim_d iid(0,1)$*, we have*

$$E\left(y_t y_{t-j}\right) = E\left(\sum_{s=1}^{t}\mu_s \sum_{s=1}^{t-j}\mu_s\right) = t - j.$$

Thus $E\left(z_t y_t y_{t-j} z_{t-j}'\right) = E\left(y_t y_{t-j}\right) E\left(z_t z_{t-j}'\right)$ *depends on* t *in a similar way. Therefore, as we discussed earlier, the usual regularity conditions employed in constructing HAC and HAR t-statistics does not apply here.*

Remark 3.4 *In view of (3.2) and Theorem 4.3 in Phillips (1998),* $\underline{W}_{yK}(r) \to 0$ *almost surely and uniformly as* $K \to \infty$*. We can expect that the rates of divergence of* $t_{b_{zK}}$ *and* $t_{b_{zK}}^{HAC}$ *are greater in the case where* $K \to \infty$ *than they are when* K *is fixed. Moreover, similar to the earlier findingg in Theorem 2.3, the HAR statistic* $t_{b_{zK}}^{HAR}$ *will diverge at rate* $O_p\left(\sqrt{K}\right)$*. Details are omitted to save space. Hence, fitted coefficients of the spurious random walk regressors eventually be deemed significant when fixed critical values are employed in testing under all three t-statistics including* $t_{b_{zK}}^{HAR}$ *when both* K, $n \to \infty$*.*

4 Simulations

This section reports simulations to investigate the performance in finite samples of the different t-statistics in spurious trend regressions, simple time trend regression, and spurious regression among stochastic trends.

We first examine spurious regression of a stochastic trend on time polynomials. Consider the standard Gaussian random walk $X_t = \sum_{s=1}^{t} \mu_s$, where $u_s \sim_d iid\ N(0,1)^3$. Orthogonal basis functions $\{\varphi_k(\cdot)\}_{k=1}^{K}$, where $\varphi_k(r) = \sqrt{2}\sin[(k-0.5)\pi r]$, were used as regressors and fitted time trend regressions of the form $X_t = \varphi'_{Kt}\hat{\beta} + \hat{u}_t$ were run with $\varphi_{Kt} = [\varphi_1\left(\frac{t}{n}\right), ..., \varphi_K\left(\frac{t}{n}\right)]'$. We focus on the prototypical null hypothesis $H_0 : \beta_1 = 0$ in what follows. In the construction of the HAC and HAR t-statistics, a uniform kernel function was employed.

Figure 1 reports the kernel estimates of the probability densities for these t-statistics under different model scenarios based on 10,000 simulations. The first panel of the figure gives the results for the different t-statistics as the sample size n increases with fixed $K = 1$. It is evident that both the usual t-statistic and HAC t-statistic (with $M = \lfloor nb \rfloor = \lfloor n^{1/4} \rfloor$ and $b = n^{-3/4}$) diverge as n increases and the HAC statistic diverges at a slower rate. In contrast, the HAR t-statistic ($b = 0.2$) is evidently convergent to a well-defined probability distribution as the sample size expands. These results clearly corroborate Theorem 2.1.

The second panel presents the estimated densities of the three t-statistics as K increases for a fixed sample size $n = 200$. As K increases, all three t-statistics are clearly divergent but at different rates. For each statistic the increase in dispersion as K increases is evident. The last panel reports the results for the HAR t-statistic with $K = 1, 5, 20$ and bandwidth coefficient $b = 0, 0.1, 0.4, 0.6, 0.8, 1$. As K increases while maintaining the same bandwidth setting, the densities become more progressively dispersed. For fixed K, it is clear that the quantile is not a monotonic function of b. For $K = 1, 5$, when b is close to zero, the limiting distributions become more dispersed. When b is close to one, the limiting distributions also get dispersed for all three choices of K. As explained in Sun (2004), for small or moderate K, when b is close to zero, the behavior of the t-statistic may be better captured by conventional limit theory without taking into account the persistence of the regression residuals. But when b is close to unity, we can not expect the standard variance estimate to capture the strong autocorrelation. If we choose the kernel $k(x) = 1$ and use the full sample (i.e., setting $b = 1$), the long run variance estimate equals to zero by construction. We conjecture that for fixed K it may be possible to find an optimal bandwidth $b_{opt}(K)$ by following an approach similar to the method used in Sun, Phillips and Jin (2009) that controls for size and power. From the shape of the densities in the last panel of Figure 1, we would expect that any such optimal bandwidth $b_{opt}(K)$ will get closer to zero as K gets larger. Extension of robust testing techniques to machine learning regressions where K may be very large will likely require very careful bandwidth selection in significance testing that takes the magnitude of K into account.

Next, we consider a simple spurious linear trend regression of X_t on a time trend. Fig-

[3] Weakly dependent innovations in the form of an AR(1) error process, viz., $\mu_s = \rho\mu_{s-1} + \varepsilon_s$, with $\varepsilon_s \sim_d iid\ N(0,1)$, were also considered. The results were similar and so only the iid case is reported here.

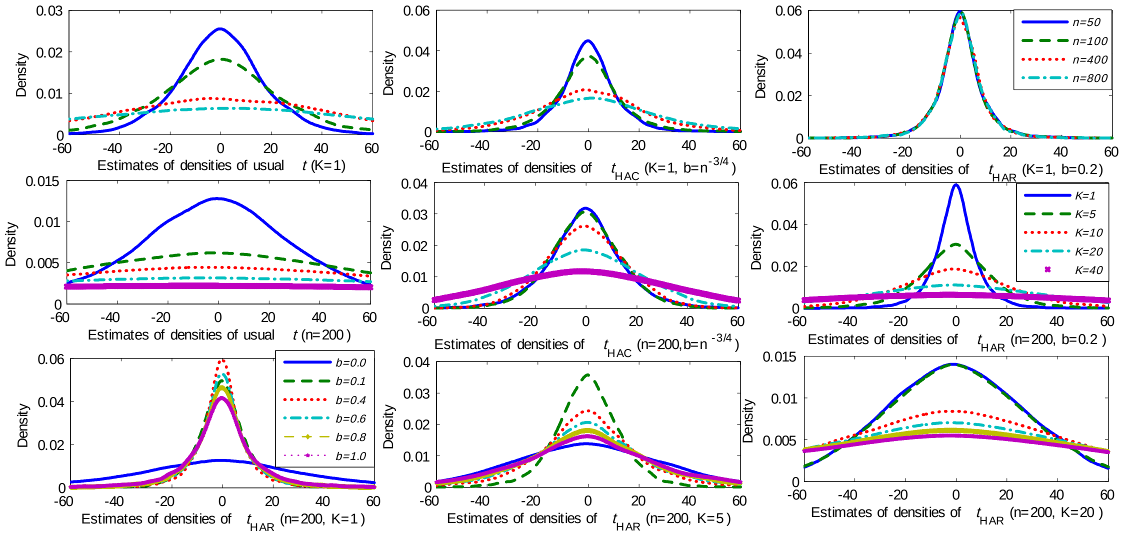

Figure 1: Densities of different t-statistics in spurious trending regression

ure 2 reports the sampling densities for different t-statistics based on 10,000 simulations. The first panel presents kernel estimates of densities of the t-statistics for sample sizes $n = 50, 100, 400, 800$. Again, the usual t-statistic and HAC statistic are divergent but at different rates. The HAR statistic is evidently convergent. The second panel in Figure 2 provides results for the HAR statistic with different bandwidth choices. It is clear that the distributions become more dispersed as b moves close to zero or close to one. In this respect the findings are similar to those of Figure 1 when $K = 1$.

Last, we consider spurious regressions of a standard Gaussian random walk process on independent Gaussian random walks. Figure 3 shows the kernel estimates of the probability densities for these t-statistics under different scenarios based on 10,000 simulations. The patterns exhibited are evidently similar to those in Figure 1. The same qualitative observations made for Figure 1 therefore apply to these regressions.

5 Conclusions

Robust inference in trend regression poses many challenges. Not least of these is the critical difficulty that a trending time series trajectory can be represented in a coordinate system by many different functions, be they relevant or irrelevant, stochastic or non-stochastic. Valid

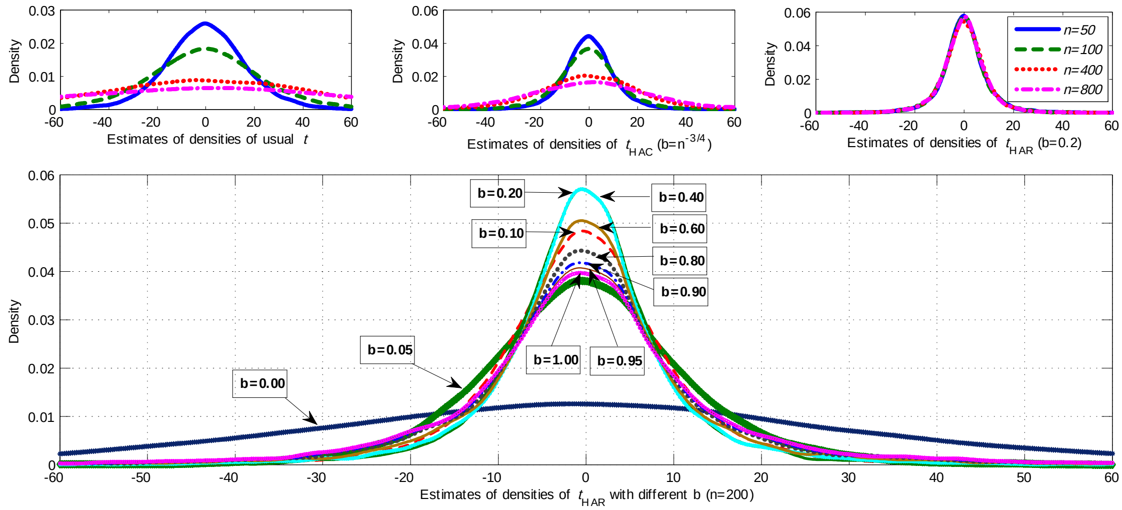

Figure 2: Densities of different t-statistics in simple spurious linear trend regression

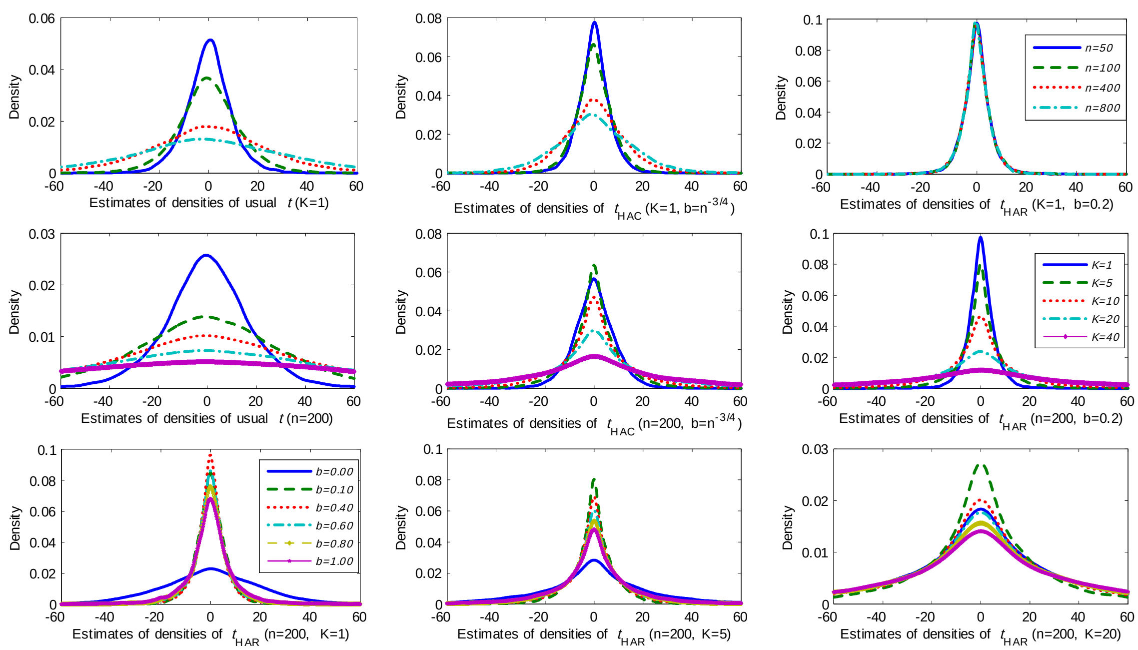

Figure 3: Densities of different t-statistics in spurious regression among random walks

significance testing in this context needs to allow for the fact that trend regression formulations inevitably fail to capture all the subtleties of reality and to a greater or lesser extent therefore involve some spurious components. The practical implications of this message is powerfully stated in the header by David Hendry that opens this article.

The present work has studied the asymptotic and finite sample performance of simple t statistics that seek to achieve some degree of robustness to misspecification in such settings. The analysis is based on three classic examples of spurious regressions, including regression of stochastic trends on time polynomials, regression of stochastic trends on a simple linear trend, and regression among independent random walks. Concordant with existing theory, the usual t-statistic and HAC standardized t-statistic both diverge and imply 'nonsense relationships' with probability going to one as the sample size tends to infinity. Also concordant with existing theory, when the number of regressors K is fixed, the HAR standardized t-statistics converge to non-degenerate distributions free from nuisance parameters, thereby controlling size and leading to valid significance tests in these spurious regressions. These findings reinforce the optimism expressed in earlier work that fixed-b methods of correction may fix inference problems in spurious regressions.

But when the number of trend regressors $K \to \infty$, the results are different. First, rates of divergence of the usual t-statistic and HAC t-test are greater by the factor $\sqrt{K}$ than when K is fixed. Second, the fixed-b HAR t-statistic is no longer convergent and instead diverges at the rate $\sqrt{K}$, leading to spurious inference of significance when $K \to \infty$. Thus, in the case of models with expanding regressor sets, none of these standard statistics produce valid consistent tests with controllable size. The failure of the HAR test in this setting is particularly important, given the growing use of machine learning algorithms in econometric work where large numbers of regressors are a normal feature in initial specifications. Future research might usefully focus on methods of controlling size and achieving consistent significant tests in such settings.

APPENDIX

A Proofs of Theorems in Section 2

Lemma A.1 *For any $r \in [0,1]$, let $\underline{B}_{\varphi_K}(r) = B(r) - Z_K \bar{\varphi}_K(r) = \sum_{k=K+1}^{\infty} \sqrt{\lambda_k}\varphi_k(r)\xi_k$ be the L_2-projection residual of B on $\varphi_K(r)$, with $\varphi_k(r) = \sqrt{2}\sin[(k-1/2)\pi r]$, $\lambda_k = \omega^2 / \left[(k-1/2)^2 \pi^2\right]$ and $\xi_k \equiv iid\ N(0,1)$. When $K \to \infty$,*

(i) $\underline{B}_{\varphi_K}(r) \sim O_p\left(1/\sqrt{K}\right)$ *uniformly in* $r \in [0,1]$,

(ii) $K\int_0^1 \underline{B}_{\varphi_K}^2 = \omega^2/\pi^2 + o_p(1)$ *with* $\omega^2 = 2\pi f_\mu(0)$,

(iii) $\int_0^1 \underline{B}_{\varphi_K}^2 \left[C_K'\bar{\varphi}_K\right]^2 \sim O_p(1/K)$,

(iv) $\int_0^1\int_0^1 k_b(r-q)\,\underline{B}_{\varphi_K}(r)\,\underline{B}_{\varphi_K}(q)\left[C_K'\bar{\varphi}_K(r)\right]\left[C_K'\bar{\varphi}_K(q)\right]drdq \sim O_p(1/K)$.

Lemma A.2 *When* $n\to\infty$, $K\to\infty$, $1/M + M/n \to 0$, *and* $K^{5/2}/n + K^{3/2}/n^{\frac{1}{2}-\frac{1}{p}} \to 0$,

(i) $C_K'\hat{\beta}_K/\sqrt{n} = C_K'\hat{\alpha}_K = C_K' Z_K + o_p(1)$,

(ii) $K\left(s_b^2 C_K'\left(\Phi_K'\Phi_K\right)^{-1} C_K\right) = K\int_0^1 \underline{B}_{\varphi_K}^2 + o_p(1)$,

(iii) $\frac{K}{M}\left(\hat{\omega}^2_{C_K'\beta_K}\right) = \left(\int_{-1}^1 k(s)\,ds\right) K\int_0^1 \underline{B}_{\varphi_K}^2\left[C_K'\bar{\varphi}_K\right]^2 + o_p(1)$,

(iv) $\frac{K}{n}\left(\check{\omega}^2_{C_K'\beta_K}\right) = K\int_0^1\int_0^1 k_b(r-q)\,\underline{B}_{\varphi_K}(r)\,\underline{B}_{\varphi_K}(q)\left[C_K'\bar{\varphi}_K(r)\right]\left[C_K'\bar{\varphi}_K(q)\right]drdq + o_p(1)$,

where $Z_K = (z_k)_{k=1}^K$ *are the random coefficients in the orthonormal representation (2.4),* s_b^2, $\hat{\omega}^2_{C_K'\beta_K}$ *and* $\check{\omega}^2_{C_K'\beta_K}$ *are defined as in formulae (2.7), (2.9) and (2.12), respectively.*

Proof of Lemma A.1. (i) It is easy to see that $E[\underline{B}_{\varphi_K}(r)] = 0$, and

$$\mathrm{Var}\left[\underline{B}_{\varphi_K}(r)\right] = \sum_{k=K+1}^{\infty}\lambda_k\varphi_k^2(r) = O\left(\sum_{k=K+1}^{\infty}\lambda_k\right) = O\left(\sum_{k=K+1}^{\infty}\frac{1}{k^2}\right) = O\left(\int_K^{\infty}\frac{1}{k^2}dk\right) = O\left(\frac{1}{K}\right)$$

uniformly in r. So by the Chebyshev's inequality, $\underline{B}_{\varphi_K}(r) = O_P\left(1/\sqrt{K}\right)$ uniformly in r.

(ii) See Phillips (2002), Lemma 3.1.

(iii)-(iv) The proofs of (iii) and (iv) are similar. Hence, only the proof of (iv) is given below. By noticing that $\xi_k \sim iid\ N(0,1)$ and for each $k = 1, ..., K$ the functions $\varphi_k(r) = \sqrt{2}\sin\left[(k-1/2)\pi r\right]$ are bounded uniformly in r, we have

$$\begin{aligned} E\left[\underline{B}_{\varphi_K}(r)\,\underline{B}_{\varphi_K}(q)\right] &= E\left[\sum_{k,l=K+1}^{\infty}\lambda_k^{1/2}\lambda_l^{1/2}\varphi_k(r)\,\varphi_l(q)\,\xi_k\xi_l\right] = \sum_{k=K+1}^{\infty}\lambda_k\varphi_k(r)\,\varphi_k(q) \\ &= O\left(\sum_{k=K+1}^{\infty}\lambda_k\right) = O\left(\frac{1}{K}\right) \quad \text{uniformly in } r\in[0,1] \text{ and } q\in[0,1]. \end{aligned}$$

Therefore,

$$
\begin{aligned}
& E\left\{\int_0^1 \int_0^1 k_b\,(r-q)\,\underline{B}_{\varphi_K}(r)\,\underline{B}_{\varphi_K}(q)\left[C_K'\bar{\varphi}_K(r)\right]\left[C_K'\bar{\varphi}_K(q)\right]drdq\right\} \\
= & \int_0^1 \int_0^1 k_b\,(r-q)\,E\left\{\underline{B}_{\varphi_K}(r)\,\underline{B}_{\varphi_K}(q)\right\}\left[C_K'\bar{\varphi}_K(r)\right]\left[C_K'\bar{\varphi}_K(q)\right]drdq \\
= & O\left(\frac{1}{K}\right)\int_0^1 \int_0^1 \left|C_K'\bar{\varphi}_K(r)\right|\left|C_K'\bar{\varphi}_K(q)\right|drdq \\
= & O\left(\frac{1}{K}\right)\left(\int_0^1 \left|C_K'\bar{\varphi}_K(r)\right|dr\right)^2 = O\left(\frac{1}{K}\right),
\end{aligned}
$$

since $k_b\,(r-q)$ is uniformly bounded and

$$
\left(\int_0^1 \left|C_K'\bar{\varphi}_K(r)\right|dr\right)^2 \le \int_0^1 \left[C_K'\bar{\varphi}_K(r)\right]^2 dr = C_K'\left(\int_0^1 \bar{\varphi}_K(r)\,\bar{\varphi}_K'(r)\,dr\right)C_K = C_K'C_K = 1,
$$

$$
\begin{aligned}
\left(\int_0^1 \left|C_K'\bar{\varphi}_K(r)\right|dr\right)^2 & \ge \left(\int_0^1 C_K'\bar{\varphi}_K(r)\,dr\right)^2 = \left(\int_0^1 \sum_{k=1}^K c_k\varphi_k(r)\,dr\right)^2 \\
& = \left(\sum_{k=1}^K c_k \left.\frac{-\sqrt{2}\cos\left[(k-1/2)\,\pi r\right]}{(k-1/2)\,\pi}\right|_0^1\right)^2 = \left(\sum_{k=1}^K \frac{\sqrt{2}c_k}{(k-1/2)\,\pi}\right)^2.
\end{aligned}
$$

Further

$$
\begin{aligned}
& E\left[\underline{B}_{\varphi_K}(r)\,\underline{B}_{\varphi_K}(q)\,\underline{B}_{\varphi_K}(s)\,\underline{B}_{\varphi_K}(\tau)\right] \\
= \;& E\left(\sum_{k=K+1}^{\infty}\lambda_k^2\varphi_k(r)\,\varphi_k(q)\,\varphi_k(s)\,\varphi_k(\tau)\,\xi_k^4+\sum_{\substack{h,k=K+1\\ h\neq k}}^{\infty}\lambda_k\lambda_h\varphi_k(r)\,\varphi_k(q)\,\varphi_h(s)\,\varphi_h(\tau)\,\xi_k^2\xi_h^2\right) \\
& +E\left(\sum_{\substack{l,k=K+1\\ l\neq k}}^{\infty}\lambda_k\lambda_l\varphi_k(r)\,\varphi_k(s)\,\varphi_l(q)\,\varphi_l(\tau)\,\xi_k^2\xi_l^2+\sum_{\substack{l,k=K+1\\ l\neq k}}^{\infty}\lambda_k\lambda_l\varphi_k(r)\,\varphi_k(\tau)\,\varphi_l(q)\,\varphi_l(s)\,\xi_k^2\xi_l^2\right) \\
= \;& 3\sum_{k=K+1}^{\infty}\lambda_k^2\varphi_k(r)\,\varphi_k(q)\,\varphi_k(s)\,\varphi_k(\tau)+\sum_{\substack{h,k=K+1\\ h\neq k}}^{\infty}\lambda_k\lambda_h\varphi_k(r)\,\varphi_k(q)\,\varphi_h(s)\,\varphi_h(\tau) \\
& +\sum_{\substack{l,k=K+1\\ l\neq k}}^{\infty}\lambda_k\lambda_l\varphi_k(r)\,\varphi_k(s)\,\varphi_l(q)\,\varphi_l(\tau)+\sum_{\substack{l,k=K+1\\ l\neq k}}^{\infty}\lambda_k\lambda_l\varphi_k(r)\,\varphi_k(\tau)\,\varphi_l(q)\,\varphi_l(s) \\
= \;& \sum_{k=K+1}^{\infty}\lambda_k\varphi_k(r)\,\varphi_k(q)\sum_{h=K+1}^{\infty}\lambda_h\varphi_h(s)\,\varphi_h(\tau)+\sum_{k=K+1}^{\infty}\lambda_k\varphi_k(r)\,\varphi_k(s)\sum_{l=K+1}^{\infty}\lambda_l\varphi_l(q)\,\varphi_l(\tau) \\
& +\sum_{k=K+1}^{\infty}\lambda_k\varphi_k(r)\,\varphi_k(\tau)\sum_{l=K+1}^{\infty}\lambda_l\varphi_l(q)\,\varphi_l(s) \\
= \;& 3\times O\left(\frac{1}{K}\right)\times O\left(\frac{1}{K}\right)=O\left(\frac{1}{K^2}\right)\quad\text{uniformly in }\;(r,q,s,\tau)\in[0,1]^4.
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
& E\left\{\left(\int_0^1\int_0^1 k_b(r-q)\,\underline{B}_{\varphi_K}(r)\,\underline{B}_{\varphi_K}(q)\left[C_K'\bar{\varphi}_K(r)\right]\left[C_K'\bar{\varphi}_K(q)\right]drdq\right)^2\right\} \\
= \;& O\left(\frac{1}{K^2}\right)\times\int_0^1\int_0^1\int_0^1\int_0^1\left|C_K'\bar{\varphi}_K(r)\right|\left|C_K'\bar{\varphi}_K(q)\right|\left|C_K'\bar{\varphi}_K(s)\right|\left|C_K'\bar{\varphi}_K(\tau)\right|drdqdsd\tau \\
= \;& O\left(\frac{1}{K^2}\right)\times\left(\int_0^1\left|C_K'\bar{\varphi}_K(r)\right|dr\right)^4=O\left(\frac{1}{K^2}\right).
\end{aligned}
$$

Finally, we get

$$
\text{Var}\left(\int_0^1\int_0^1 k_b(r-q)\,\underline{B}_{\varphi_K}(r)\,\underline{B}_{\varphi_K}(q)\left[C_K'\bar{\varphi}_K(r)\right]\left[C_K'\bar{\varphi}_K(q)\right]drdq\right)=O\left(\frac{1}{K^2}\right).
$$

■

Proof of Lemma A.2. (i) See Phillips (2002), Lemma 2.2.

(ii) Using the Hungarian strong approximation (e.g., Csörgõ and Horváth 1993), we can construct an expanded probability space with a Brownian motion $B(\cdot)$ for which

$$\sup_{1\leq t\leq n}|X_t - B(t)| = o_{a.s}\left(n^{1/p}\right),$$

or

$$\sup_{1\leq t\leq n}\left|\frac{X_t}{\sqrt{n}} - B\left(\frac{t}{n}\right)\right| = o_{a.s}\left(\frac{1}{n^{1/2-1/p}}\right).$$

Applying the matrix norm $\|A\| = \max_i \sum_{j=1}^{K}|a_{ij}|$, Phillips (2002) proved that

$$\frac{1}{n}\sum_{t=1}^{n}\varphi_{Kt}\varphi_{Kt}' = I_K + O\left(\frac{K}{n}\right),$$

and

$$\hat{\beta}_K/\sqrt{n} = \hat{\alpha}_K = \Lambda_K^{1/2}\tilde{\xi}_K + O_{a.s}\left(\frac{K}{n} + \frac{1}{n^{1/2-1/p}}\right) = Z_K + O_{a.s}\left(\frac{K}{n} + \frac{1}{n^{1/2-1/p}}\right),$$

where $\tilde{\xi}_K = (\xi_k)_{k=1}^{K}$, and $Z_K = (z_k)_{k=1}^{K}$ are the random coefficients in the orthonormal representation (2.4). Therefore, we have

$$\begin{aligned}
\sup_{1\leq t\leq n}\left|\frac{\hat{u}_t}{\sqrt{n}} - \underline{B}_{\varphi_K}\left(\frac{t}{n}\right)\right| &= \sup_{1\leq t\leq n}\left|\left(\frac{X_t}{\sqrt{n}} - \hat{\alpha}_K'\varphi_{Kt}\right) - \left(B\left(\frac{t}{n}\right) - Z_K'\varphi_{Kt}\right)\right| \\
&\leq \sup_{1\leq t\leq n}\left|\frac{X_t}{\sqrt{n}} - B\left(\frac{t}{n}\right)\right| + \sup_{1\leq t\leq n}\left|\varphi_{Kt}'(Z_K - \hat{\alpha}_K)\right| \\
&\leq o_{a.s}\left(\frac{1}{n^{1/2-1/p}}\right) + \sup_{1\leq t\leq n}\left\|\varphi_{Kt}'\right\|\left\|Z_K - \hat{\alpha}_K\right\| \\
&= o_{a.s}\left(\frac{1}{n^{1/2-1/p}}\right) + O_{a.s}\left(\frac{K}{n} + \frac{1}{n^{1/2-1/p}}\right)\sup_{1\leq t\leq n}\left\|\varphi_{Kt}'\right\| \\
&= O_{a.s}\left(\frac{K^2}{n} + \frac{K}{n^{1/2-1/p}}\right),
\end{aligned}$$

as $\sup_{1\leq t\leq n}\|\varphi_{Kt}'\| = \sup_{1\leq t\leq n}\sqrt{2}\sum_{k=1}^{K}|\sin[(k-1/2)\pi t/n]| = O(K)$. The second inequality comes from Hölder's inequality. Hence, when $K^{5/2}/n + K^{3/2}/n^{1/2-1/p} \to 0$, we have

$$
\begin{aligned}
\frac{K}{n} s_b^2 &= \frac{K}{n}\sum_{t=1}^{n}\left(\frac{\hat{u}_t}{\sqrt{n}}\right)^2 = \frac{1}{n}\sum_{t=1}^{n}\left(\sqrt{K}\underline{B}_{\varphi_K}\left(\frac{t}{n}\right) + O_{a.s}\left(\frac{K^{5/2}}{n} + \frac{K^{3/2}}{n^{1/2-1/p}}\right)\right)^2 \\
&= \frac{1}{n}\sum_{t=1}^{n}\left(\sqrt{K}\underline{B}_{\varphi_K}\left(\frac{t}{n}\right) + o_{a.s}(1)\right)^2 = \frac{K}{n}\sum_{t=1}^{n}\left(\underline{B}_{\varphi_K}\left(\frac{t}{n}\right)\right)^2 + o_{a.s}(1) \\
&= K\int_0^1 \underline{B}_{\varphi_K}^2(r)\,dr + o_p(1).
\end{aligned}
$$

It is straightforward to see that when $K/n \to 0$,

$$
C_K'\left(\frac{1}{n}\Phi_K'\Phi_K\right)^{-1} C_K = C_K'\left(I_K + O\left(\frac{K}{n}\right)\right) C_K = C_K' C_K + o(1) = 1 + o(1).
$$

Therefore, when $K^{5/2}/n + K^{3/2}/\left(n^{1/2-1/p}\right) \to 0$,

$$
K\left(s_b^2 C_K'\left(\Phi_K'\Phi_K\right)^{-1} C_K\right) = \frac{K}{n} s_b^2\left(C_K'\left(\frac{1}{n}\Phi_K'\Phi_K\right)^{-1} C_K\right) = K\int_0^1 \underline{B}_{\varphi_K}^2 + o_p(1).
$$

(iii), (iv) The proofs of (iii) and (iv) are similar, so only (iv) is proved here. When $K^{5/2}/n + K^{3/2}/n^{1/2-1/p} \to 0$,

$$
\begin{aligned}
&\frac{K}{n^2} C_K' \widehat{\text{lrvar}}_{\text{HAR}}\left(\hat{u}_t\varphi_{Kt}\right) C_K \\
=\ & \frac{1}{n^2}\sum_{j=-n+1}^{n-1} k_b\left(\frac{j}{n}\right)\frac{K}{n}\sum_{1\le t,t+j\le n}\hat{u}_t\hat{u}_{t+j} C_K'\varphi_{Kt}\varphi_{Kt+j}' C_K \\
=\ & \frac{1}{n^2}\sum_{s,t=1}^{n} k_b\left(\frac{t-s}{n}\right)\frac{\sqrt{K}\hat{u}_t}{\sqrt{n}}\frac{\sqrt{K}\hat{u}_s}{\sqrt{n}} C_K'\varphi_{Kt}\varphi_{Ks}' C_K \\
=\ & \frac{1}{n^2}\sum_{s,t=1}^{n} k_b\left(\frac{t-s}{n}\right)\left(\sqrt{K}\underline{B}_{\varphi_K}\left(\frac{t}{n}\right) + o_{a.s}(1)\right)\left(\sqrt{K}\underline{B}_{\varphi_K}\left(\frac{s}{n}\right) + o_{a.s}(1)\right) C_K'\varphi_{Kt}\varphi_{Ks}' C_K \\
=\ & \frac{1}{n^2}\sum_{s,t=1}^{n} k_b\left(\frac{t-s}{n}\right)\left(\sqrt{K}\underline{B}_{\varphi_K}\left(\frac{t}{n}\right)\right)\left(\sqrt{K}\underline{B}_{\varphi_K}\left(\frac{s}{n}\right)\right) C_K'\varphi_{Kt}\varphi_{Ks}' C_K + o_{a.s}(1) \\
=\ & K\int_0^1\int_0^1 k_b(r-q)\,\underline{B}_{\varphi_K}(r)\,\underline{B}_{\varphi_K}(q)\left[C_K'\bar{\varphi}_K(r)\right]\left[C_K'\bar{\varphi}_K(q)\right] dr\,dq + o_p(1).
\end{aligned}
$$

Therefore, when $K^{5/2}/n + K^{3/2}/\left(n^{1/2-1/p}\right) \to 0$,

$$\begin{aligned}
\frac{K}{n}\left(\breve{\omega}^2_{C'_K\beta_K}\right) &= C'_K\left(\frac{1}{n}\Phi'_K\Phi_K\right)^{-1}\left(\frac{K}{n^2}\widehat{\text{lrvar}}_{\text{HAR}}\left(\hat{u}_t\varphi_{Kt}\right)\right)\left(\frac{1}{n}\Phi'_K\Phi_K\right)^{-1}C_K \\
&= C'_K\left(\frac{K}{n^2}\widehat{\text{lrvar}}_{\text{HAR}}\left(\hat{u}_t\varphi_{Kt}\right)\right)C_K + o_{a.s.}(1) \\
&= K\int_0^1\int_0^1 k_b\left(r-q\right)\underline{B}_{\varphi_K}\left(r\right)\underline{B}_{\varphi_K}\left(q\right)\left[C'_K\bar{\varphi}_K\left(r\right)\right]\left[C'_K\bar{\varphi}_K\left(q\right)\right]drdq + o_p\left(1\right).
\end{aligned}$$

■

Proof of Theorem 2.1. (i)-(ii) The proofs are similar to those in Phillips (1998) and are omitted.

(iii) From Phillips (1998), when $n \to \infty$ and K is fixed, $n^{-1/2}\beta_K = \hat{\alpha}_K \Rightarrow Z_K$. Let $\varphi_{Kt} = \left(\varphi_1\left(t/n\right), ..., \varphi_K\left(t/n\right)\right)'$, we have

$$\frac{\hat{u}_{\lfloor n\cdot\rfloor}}{\sqrt{n}} = \frac{1}{\sqrt{n}}\left(X_{\lfloor n\cdot\rfloor} - \hat{\beta}'_K\varphi_{K\lfloor n\cdot\rfloor}\right)^2 \Rightarrow B\left(\cdot\right) - Z'_K\bar{\varphi}_K\left(\cdot\right) := \underline{B}_{\varphi_K}\left(\cdot\right).$$

The scaled long run variance estimator can be written as

$$\begin{aligned}
\frac{1}{n^2}\widehat{\text{lrvar}}_{\text{HAR}}\left(\hat{u}_t\varphi_{Kt}\right) &= \frac{1}{n^2}\sum_{j=-n+1}^{n-1}k_b\left(\frac{j}{n}\right)\frac{1}{n}\sum_{1\le t,t+j\le n}\hat{u}_t\hat{u}_{t+j}\varphi_{Kt}\varphi'_{Kt+j} \\
&= \frac{1}{n^2}\sum_{s,t=1}^{n}k_b\left(\frac{t-s}{n}\right)\frac{\hat{u}_t}{\sqrt{n}}\frac{\hat{u}_s}{\sqrt{n}}\varphi_{Kt}\varphi'_{Ks} \\
&\Rightarrow \int_0^1\int_0^1 k_b\left(r-q\right)\underline{B}_{\varphi_K}\left(r\right)\underline{B}_{\varphi_K}\left(q\right)\bar{\varphi}_K\left(r\right)\bar{\varphi}'_K\left(q\right)drdq.
\end{aligned}$$

Noticing that

$$\frac{1}{n}\Phi'_K\Phi_K = \frac{1}{n}\sum_{t=1}^{n}\varphi_{Kt}\varphi'_{Kt} \to \int_0^1\bar{\varphi}_K\left(r\right)\bar{\varphi}'_K\left(r\right)dr = I_K,$$

it follows that

$$\begin{aligned}
\frac{1}{n}\breve{\omega}^2_{C'_K\beta_K} &= C'_K\left\{\frac{1}{n}\Phi'_K\Phi_K\right\}^{-1}\left(\frac{1}{n^2}\widehat{\text{lrvar}}_{\text{HAR}}\left(\hat{u}_t\varphi_{Kt}\right)\right)\left\{\frac{1}{n}\Phi'_K\Phi_K\right\}^{-1}C_K \\
&\Rightarrow C'_K\int_0^1\int_0^1 k_b\left(r-q\right)\underline{B}_{\varphi_K}\left(r\right)\underline{B}_{\varphi_K}\left(q\right)\bar{\varphi}_K\left(r\right)\bar{\varphi}'_K\left(q\right)drdqC_K \\
&= \int_0^1\int_0^1 k_b\left(r-q\right)\underline{B}_{\varphi_K}\left(r\right)\underline{B}_{\varphi_K}\left(q\right)\left[C'_K\bar{\varphi}_K\left(r\right)\right]\left[C'_K\bar{\varphi}_K\left(q\right)\right]drdq.
\end{aligned}$$

Therefore,

$$
\begin{aligned}
t^{HAR}_{C_K'\beta_K} &= \frac{C_K'\hat{\beta}_K}{\check{\omega}_{C_K'\beta_K}} = \frac{n^{-1/2}C_K'\hat{\beta}_K}{\left\{n^{-1}\check{\omega}^2_{C_K'\beta_K}\right\}^{1/2}} \\
&\Rightarrow \frac{C_K' Z_K}{\left[\int_0^1\int_0^1 k_b(r-q)\,\underline{B}_{\varphi_K}(r)\,\underline{B}_{\varphi_K}(q)\left[C_K'\bar{\varphi}_K(r)\right]\left[C_K'\bar{\varphi}_K(q)\right]drdq\right]^{1/2}}.
\end{aligned}
$$

Let $Z_K^W = Z_K/\omega = \int_0^1 W\bar{\varphi}_K$, $W(r) = B(r)/\omega \equiv BM(1)$, $\omega^2 = 2\pi f_\mu(0)$, and $\underline{W}_{\varphi_K}(r) = \underline{B}_{\varphi_K}(r)/\omega = W(r) - \left(Z_K^W\right)'\bar{\varphi}_K(r)$. It then follows immediately that

$$
t^{HAR}_{C_K'\beta_K} \Rightarrow \frac{C_K' Z_K^W}{\left[\int_0^1\int_0^1 k_b(r-q)\,\underline{W}_{\varphi_K}(r)\,\underline{W}_{\varphi_K}(q)\left[C_K'\bar{\varphi}_K(r)\right]\left[C_K'\bar{\varphi}_K(q)\right]drdq\right]^{1/2}}.
$$

■

B Derivations leading to (2.22)-(2.27)

Lemma B.1 *For the regression model (2.16) let $B(\cdot) \equiv BM\left(\omega^2\right)$ with $\omega^2 = 2\pi f_\mu(0) > 0$. Irrespective of whether a is zero or not, when $n \to \infty$ and $1/M + M/n \to 0$, the following results hold:*

(i) for $r \in [0,1]$,

$$
\frac{\hat{u}_{\lfloor nr \rfloor}}{\sqrt{n}} \Rightarrow B(r) - 3\left(\int_0^1 sB(s)\,ds\right) r := \underline{B}(r) \ ;
$$

(ii)

$$
n^2 (s_a)^2 \Rightarrow 3\int_0^1 \underline{B}^2 \ ;
$$

(iii)

$$
\frac{n^2}{M}(\hat{\omega}_a)^2 \Rightarrow 9\int_{-1}^1 k(s)\,ds \int_0^1 r^2\underline{B}^2 \ ;
$$

(iv)

$$
n(\check{\omega}_a)^2 \Rightarrow 9\int_0^1\int_0^1 k_b(r-q)\,\underline{B}(r)\,\underline{B}(q)\,rqdrdq \ ;
$$

where s_a, $\hat{\omega}_a$, $\check{\omega}_a$ are defined as in (2.19), (2.20), (2.21), respectively, $\hat{u}_t = X_t - \hat{a}t$ for $t = 1, \cdots, n$, $k(\cdot)$ is a kernel function and $k_b\left(\frac{j}{n}\right) = k\left(\frac{j}{nb}\right)$.

Proof of Lemma B.1. (i) Using the functional law and continuous mapping it is straightforward to obtain

$$\sqrt{n}\left(\hat{a}-a\right)=\frac{n^{-5/2}\sum_{t=1}^{n}X_t^0}{n^{-3}\sum_{t=1}^{n}t^2}\Rightarrow\frac{\int_0^1 sB\left(s\right)ds}{\int_0^1 s^2ds}\equiv 3\int_0^1 sB\left(s\right)ds.$$

Therefore, for any $r\in[0,1]$,

$$\begin{aligned}\frac{\hat{u}_{\lfloor nr\rfloor}}{\sqrt{n}} &= \frac{X_{\lfloor nr\rfloor}-\hat{a}\cdot\lfloor nr\rfloor}{\sqrt{n}}=\frac{X_{\lfloor nr\rfloor}^0}{\sqrt{n}}-\sqrt{n}(\hat{a}-a)\frac{\lfloor nr\rfloor}{n}\\ &\Rightarrow B\left(r\right)-3\left(\int_0^1 sB\left(s\right)ds\right)r:=\underline{B}\left(r\right).\end{aligned}$$

(ii) From the expression of s_a given in (2.19), the following is immediate

$$\begin{aligned}n^2\left(s_a\right)^2 &= n^2\left[\frac{1}{n}\sum_{t=1}^{n}\left(X_t-\hat{a}t\right)^2\right]\left(\sum_{t=1}^{n}t^2\right)^{-1}=\left[\frac{1}{n}\sum_{t=1}^{n}\left(\frac{\hat{u}_t}{\sqrt{n}}\right)^2\right]\left(\frac{1}{n^3}\sum_{t=1}^{n}t^2\right)^{-1}\\ &\Rightarrow 3\int_0^1\underline{B}^2\left(s\right)ds.\end{aligned}$$

(iii) As $1/M+M/N\to 0$ when $n\to\infty$, we have for any $|j|\leq M$ and $r\in[0,1]$

$$\frac{\lfloor nr\rfloor+j}{n}=r+O\left(\frac{M}{n}\right)\to r\text{ as }n\to\infty.$$

Therefore, from the continuous mapping theorem, it follows that

$$\begin{aligned}\frac{1}{M}\frac{1}{n^3}\widehat{\text{lrvar}}_{HAC}\left(t\hat{u}_t\right) &= \frac{1}{M}\sum_{j=-M}^{M}k\left(\frac{j}{M}\right)\frac{1}{n}\sum_{1\leq t,t+j\leq n}\frac{\hat{u}_t}{\sqrt{n}}\frac{\hat{u}_{t+j}}{\sqrt{n}}\frac{t}{n}\frac{t+j}{n}\\ &\Rightarrow \int_{-1}^1 k\left(s\right)ds\int_0^1\underline{B}\left(r\right)^2r^2dr.\end{aligned}$$

Hence,

$$\begin{aligned}\frac{n^2}{M}\left(\hat{\omega}_a\right)^2 &= \left(\frac{1}{n^3}\sum_{t=1}^{n}t^2\right)^{-1}\left[\frac{1}{M}\frac{1}{n^3}\widehat{\text{lrvar}}_{HAC}\left(t\hat{u}_t\right)\right]\left(\frac{1}{n^3}\sum_{t=1}^{n}t^2\right)^{-1}\\ &\Rightarrow 9\int_{-1}^1 k\left(s\right)ds\int_0^1\underline{B}\left(r\right)^2r^2dr.\end{aligned}$$

(iv) For the HAR based test given in (2.21), we have

$$\begin{aligned}\widehat{\text{lrvar}}_{HAR}\left(t\hat{u}_t\right) &= \sum_{j=-n+1}^{n-1} k_b\left(\frac{j}{n}\right)\left(\frac{1}{n}\sum_{1\leq t,t+j\leq n} t\hat{u}_t\hat{u}_{t+j}\left(t+j\right)\right)\\ &= \frac{1}{n}\sum_{s,t=1}^{n} k_b\left(\frac{t-s}{n}\right)\hat{u}_t\hat{u}_s ts.\end{aligned}$$

By continuous mapping

$$\begin{aligned}\frac{1}{n^4}\widehat{\text{lrvar}}_{HAR}\left(t\hat{u}_t\right) &= \frac{1}{n^2}\sum_{s,t=1}^{n} k_b\left(\frac{t-s}{n}\right)\frac{\hat{u}_s}{\sqrt{n}}\frac{\hat{u}_t}{\sqrt{n}}\frac{s}{n}\frac{t}{n}\\ &\Rightarrow \int_0^1\int_0^1 k_b\left(r-q\right)\underline{B}\left(r\right)\underline{B}\left(q\right)rqdrdq.\end{aligned}$$

Then,

$$\begin{aligned}n\left(\breve{\omega}_a\right)^2 &= \left(\frac{1}{n^3}\sum_{t=1}^{n} t^2\right)^{-1}\left(\frac{1}{n^4}\widehat{\text{lrvar}}_{HAR}\left(t\hat{u}_t\right)\right)\left(\frac{1}{n^3}\sum_{t=1}^{n} t^2\right)^{-1}\\ &\Rightarrow 9\int_0^1\int_0^1 k_b\left(r-q\right)\underline{B}\left(r\right)\underline{B}\left(q\right)rqdrdq.\end{aligned}$$

■

Proof of (2.22)-(2.27). The stated results now follow directly from the above and the fact that $\sqrt{n}\hat{a} \Rightarrow 3\int_0^1 sB\left(s\right)ds$ under $H_0 : a = 0$, and $\hat{a} \xrightarrow{p} a$ under $H_1 : a \neq 0$. ■

C Proof of the Theorem in Section 3

Proof of Theorem 3.1. (i): In the regression (3.3), we already have that $n^{-1/2}y_{\lfloor nr\rfloor} \Rightarrow B\left(r\right)$, $n^{-1/2}z_{k\lfloor n\cdot\rfloor} \Rightarrow V_k\left(\cdot\right)$, for $k = 1, \cdots, K$, and $\hat{b}_{zK} \Rightarrow E_K := \left(e_k\right)_{k=1}^K$. Let $\bar{V}_K = \left(V_k\right)_{k=1}^K$. Based on continuous mapping, we have

$$n^{-2}\sum_{t=1}^{n} z_t z_t' = \frac{1}{n}\sum_{t=1}^{n}\frac{z_t}{\sqrt{n}}\frac{z_t'}{\sqrt{n}} \Rightarrow \int_0^1 \bar{V}_K\bar{V}_K'.$$

Noticing that

$$\frac{\hat{u}_{\lfloor n\cdot\rfloor}}{\sqrt{n}} = \frac{y_{\lfloor n\cdot\rfloor}}{\sqrt{n}} - \hat{b}_{zK}'\frac{z_{\lfloor n\cdot\rfloor}}{\sqrt{n}} \Rightarrow B\left(\cdot\right) - E_K'\bar{V}_K\left(\cdot\right) := \underline{W}_{yK}\left(\cdot\right),$$

we obtain

$$n^{-2}\sum_{t=1}^{n}(\hat{u}_t)^2=\frac{1}{n}\sum_{t=1}^{n}\left(\frac{\hat{u}_t}{\sqrt{n}}\right)^2\Rightarrow\int_0^1\underline{W}_{yK}^2.$$

Therefore

$$n\left(s_{b_{zK}}\right)^2=C_K'\left[n^{-2}\sum_{t=1}^{n}(\hat{u}_t)^2\right]\left(n^{-2}\sum_{t=1}^{n}z_tz_t'\right)^{-1}C_K\Rightarrow C_K'\left(\int_0^1\bar{V}_K\bar{V}_K'\right)^{-1}C_K\int_0^1\underline{W}_{yK}^2,$$

and

$$\frac{1}{\sqrt{n}}t_{b_{zK}}=\frac{C_K'\hat{b}_{zK}}{\left\{n\left(s_{b_{zK}}\right)^2\right\}^{1/2}}\Rightarrow\frac{C_K'E_K}{\left\{C_K'\left(\int_0^1\bar{V}_K\bar{V}_K'\right)^{-1}C_K\int_0^1\underline{W}_{yK}^2\right\}^{1/2}}.$$

(ii) As $1/M+M/n\to 0$ when $n\to\infty$, for any $|j|\leq M$ and $r\in[0,1]$, we have

$$\frac{\lfloor nr\rfloor+j}{n}=r+O\left(\frac{M}{n}\right)\to r\text{ as }n\to\infty.$$

Hence, for any $|j|\leq M$,

$$\frac{1}{n}\sum_{1\leq t,t+j\leq n}\frac{z_t}{\sqrt{n}}\frac{\hat{u}_t}{\sqrt{n}}\frac{\hat{u}_{t+j}}{\sqrt{n}}\frac{z_{t+j}'}{\sqrt{n}}\Rightarrow\int_0^1\underline{W}_{yK}^2\bar{V}_K\bar{V}_K',$$

and

$$\begin{aligned}\frac{1}{M}\frac{1}{n^2}\widehat{\text{lrvar}}_{HAC}\left(z_t\hat{u}_t\right) &= \frac{1}{M}\sum_{j=-M}^{M}k\left(\frac{j}{M}\right)\left[\frac{1}{n}\sum_{1\leq t,t+j\leq n}\frac{z_t}{\sqrt{n}}\frac{\hat{u}_t}{\sqrt{n}}\frac{\hat{u}_{t+j}}{\sqrt{n}}\frac{z_{t+j}'}{\sqrt{n}}\right]\\ &\Rightarrow \int_{-1}^{1}k\left(s\right)ds\int_0^1\underline{W}_{yK}^2\bar{V}_K\bar{V}_K'.\end{aligned}$$

Therefore,

$$\begin{aligned}\frac{n}{M}\left(\hat{\omega}_{b_{zK}}\right)^2 &= C_K'\left(n^{-2}\sum_{t=1}^{n}z_tz_t'\right)^{-1}\left[\frac{1}{M}\frac{1}{n^2}\widehat{\text{lrvar}}_{HAC}\left(z_t\hat{u}_t\right)\right]\left(n^{-2}\sum_{t=1}^{n}z_tz_t'\right)^{-1}C_K\\ &\Rightarrow C_K'\left(\int_0^1\bar{V}_K\bar{V}_K'\right)^{-1}\left[\int_{-1}^{1}k\left(s\right)ds\int_0^1\underline{W}_{yK}^2\bar{V}_K\bar{V}_K'\right]\left(\int_0^1\bar{V}_K\bar{V}_K'\right)^{-1}C_K,\end{aligned}$$

and

$$\begin{aligned}\sqrt{\frac{M}{n}} t_{b_{zK}}^{HAC} &= \frac{C_K' \hat{b}_{zK}}{\left\{\frac{n}{M}\left(\hat{\omega}_{b_{zK}}\right)^2\right\}^{1/2}} \\ &\Rightarrow \frac{C_K' E_K}{\left\{C_K'\left(\int_0^1 \bar{V}_K \bar{V}_K'\right)^{-1}\left[\int_{-1}^1 k(s)\, ds \int_0^1 \underline{W}_{yK}^2 \bar{V}_K \bar{V}_K'\right]\left(\int_0^1 \bar{V}_K \bar{V}_K'\right)^{-1} C_K\right\}^{1/2}}.\end{aligned}$$

(iii) Note that

$$\begin{aligned}\frac{1}{n^3}\widehat{\text{lrvar}}_{HAR}(z_t \hat{u}_t) &= \frac{1}{n^3}\sum_{j=-n+1}^{n-1} k_b\left(\frac{j}{n}\right)\left(\frac{1}{n}\sum_{1 \le t, t+j \le n} z_t \hat{u}_t \hat{u}_{t+j} z_{t+j}'\right) \\ &= \frac{1}{n^4}\sum_{s,t=1}^{n} k_b\left(\frac{t-s}{n}\right) z_t \hat{u}_t \hat{u}_s z_s' \\ &= \frac{1}{n^2}\sum_{s,t=1}^{n} k_b\left(\frac{t-s}{n}\right)\frac{z_t}{\sqrt{n}}\frac{\hat{u}_t}{\sqrt{n}}\frac{\hat{u}_s}{\sqrt{n}}\frac{z_s'}{\sqrt{n}} \\ &\Rightarrow \int_0^1\int_0^1 k_b(r-p)\bar{V}_K(r)\underline{W}_{yK}(r)\underline{W}_{yK}(p)\bar{V}_K'(p)\, dr dp := H.\end{aligned}$$

Therefore,

$$\begin{aligned}\left(\breve{\omega}_{b_{zK}}\right)^2 &= C_K'\left(n^{-2}\sum_{t=1}^{n} z_t z_t'\right)^{-1}\left(\frac{1}{n^3}\widehat{\text{lrvar}}_{HAR}(z_t \hat{u}_t)\right)\left(n^{-2}\sum_{t=1}^{n} z_t z_t'\right)^{-1} C_K \\ &\Rightarrow C_K'\left(\int_0^1 \bar{V}_K \bar{V}_K'\right)^{-1} H\left(\int_0^1 \bar{V}_K \bar{V}_K'\right)^{-1} C_K,\end{aligned}$$

and

$$t_{b_{zK}}^{HAR} = \frac{C_K' \hat{b}_{zK}}{\left\{\left(\hat{\omega}_{b_{zK}}\right)^2\right\}^{1/2}} \Rightarrow \frac{C_K' E_K}{\left\{C_K'\left(\int_0^1 \bar{V}_K \bar{V}_K'\right)^{-1} H\left(\int_0^1 \bar{V}_K \bar{V}_K'\right)^{-1} C_K\right\}^{1/2}}.$$

■

Notations

$L_2[0,1]$	space of square integrable functions on $[0,1]$.
$\Longrightarrow$	weak convergence.
$\lfloor \cdot \rfloor$	integer part of.
$:=$	definitional equality.
$o_p(1)$	tends to zero in probability.
$o_{a.s}(1)$	tends to zero almost surely.
$O_p(1)$	bounded in probability.
$\xrightarrow{p}$	converge in probability.
$r \wedge s$	$\min(r,s)$.
$\sim$	asymptotic equivalence.
$\equiv$	distributional equivalence.
$\sim_d$	distributed as

REFERENCES

Csörgõ, M., and Horváth, L. (1993). *Weighted Approximations in Probability and Statistics*. New York: Wiley.

Durlauf, S.N., and Phillips, P.C.B. (2018). Trends versus random walks in time series analysis. *Econometrica* 56: 1333-1354.

Eicker, F. (1967). Limit theorems for regression with unequal and dependent errors. *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*. pp. 59-82.

Fan, J., Shao, Q-M., and Zhou, W-Xin (2018). Are discoveries spurious? Distributions of maximum spurious correlations and their applications. *Annals of Statistics* 46: 989-1017.

Granger, C.W.J., and Newbold, P. (1974). Spurious regressions in econometrics. *Journal of Econometrics* 74: 111-120.

Hendry, D.F. (1980). Econometrics – alchemy or science. *Economica* 47: 387-406.

Hwang, J. and Sun, Y. (2018). Simple, robust, and accurate F and t tests in cointegrated systems, *Econometric Theory* 34: 949-984.

Jansson, M. (2004). The error in rejection probability of simple autocorrelation robust tests. *Econometrica* 72:937-946.

Kiefer, N.M., Vogelsang, T.J., and Bunzel, H. (2000). Simple robust testing of regression hypotheses. *Econometrica* 68:695-714.

Kiefer, N.M., and Vogelsang, T.J. (2002a). Heteroskedasticity-autocorrelation robust testing using bandwidth equal to sample size. *Econometric Theory* 18:1350-1366.

Kiefer, N.M., and Vogelsang, T.J. (2002b). Heteroskedasticity-autocorrelation robust standard errors using the bartlett kernel without truncation. *Econometrica* 70: 2093-2095.

Kiefer, N.M., Vogelsang, T.J. (2005). A new asymptotic theory for heteroskedasticity-autocorrelation robust tests. *Econometric Theory* 21: 1130-1164.

Loève, M. (1963). *Probability Theory.* 3rd Edition. New York: Van Nostrand.

Marmol, F. (1995) Spurious regressions between I(d) processes. *Journal of Time Series Analysis* 16: 313-321.

Marmol, F. and Velasco, C. (2002) Trend stationarity versus long-range dependence in time series analysis. *Journal of Econometrics* 108: 25-42.

Müller, U.K. (2007) A theory of robust long-run variance estimation. *Journal of Econometrics* 141: 1331-1352.

Müller, U.K. and Watson, M.W. (2018). Long-run covariability. *Econometrica* 86: 775-804.

Müller, U.K. and Watson, M.W. (2016). Low-frequency econometrics. In B. Honoré and L. Samuelson (Eds.) *Advances in Economics: Eleventh World Congress of the Econometric Society*, Vol. 2, pp. 53-94. Cambridge University Press.

Phillips, P.C.B. (1986). Understanding spurious regression in econometrics. *Journal of Econometrics* 33: 311-340.

Phillips, P.C.B. (1996). Spurious regression unmasked. *Cowles Foundation Discussion paper* 1135, Yale University.

Phillips, P.C.B. (1998). New tools for understanding spurious regressions. *Econometrica* 66:1299-1325.

Phillips, P.C.B. (2002). New unit root asymptotics in the presence of deterministic trends. *Journal of Econometrics* 111:323-353.

Phillips, P.C.B. (2014) Optimal estimation of cointegrated systems with irrelevant instruments. *Journal of Econometrics* 178: 210-224.

Phillips, P.C.B., and Solo, V. (1992). Asymptotics for linear processes. *The Annals of Statistics* 20: 971-1001.

Phillips, P.C.B., Sun, Y., and Jin, S. (2007). Consistent HAC estimation and robust regression testing using sharp origin kernels with no truncation. *Journal of Statistical Planning and Inference* 137: 985-1023.

Robinson, P.M. (1998). Inference-without-smoothing in the presence of nonparametric autocorrelation. *Econometrica* 66: 1163-1182.

Sun, Y. (2004). A convergent t-statistic in spurious regression. *Econometric Theory* 20: 943-962.

Sun, Y. (2014a). Let's fix it: Fixed-b asymptotics versus small-b asymptotics in heteroscedasticity and autocorrelation robust inference. *Journal of Econometrics* 178: 659-677.

Sun, Y. (2014b). Fixed-smoothing asymptotics in a two-step GMM framework. *Econometrica* 82: 2327-2370.

Sun, Y. (2014c). Fixed-smoothing asymptotics and asymptotic F and t tests in the presence of strong autocorrelation. *Advances in Econometrics* 33, Essays in Honor of Peter C.B. Phillips, 23-63.

Sun, Y., Phillips, P.C.B., and Jin, S. (2009). Power maximization and size control in heteroskedasticity and autocorrelation robust tests with exponentiated kernels. Forthcoming in *Econometric Theory*.

White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity, *Econometrica* 48: 817-838.

White, H. (1982). *Asymptotic Theory for Econometricians.* 1st Edition, Academic Press.

Yule, G.U. (1926). Why do we sometimes get nonsense correlations between time series? A study in sampling and the nature of time series. *Journal of the Royal Statistical Society* 89: 1-69.