

AN ALTERNATIVE BOOTSTRAP TO MOVING BLOCKS FOR TIME SERIES REGRESSION MODELS*

by

Javier Hidalgo

London School of Economics and Political Science

Contents:

Abstract

1. Introduction

2. Conditions and Main Results

3. Application of the Bootstrap

4. Monte Carlo Simulation

5. Proofs

6. A Technical Lemma

7. Conclusions

References

Tables 4.1 – 4.5

The Suntory Centre
Suntory and Toyota International Centres for
Economics and Related Disciplines
London School of Economics and Political Science
Houghton Street
London WC2A 2AE
Tel.: 020 7955 6698

Discussion Paper
No.EM/03/452
May 2003

* This article is based on research funded by the Economic and Social Research Council (ESRC) reference number: R000238212. I am grateful to an associate editor and two referees for their very helpful comments which have led to a substantially improved version. All remaining errors are, of course, my sole responsibility.

Abstract

The purpose of this paper is to introduce and examine two alternative, although similar, approaches to the Moving Blocks and subsampling Bootstraps to bootstrapping the estimator of the parameters for time series regression models. More specifically, the first bootstrap is based on resampling from the normalised *discrete Fourier transform* of the residuals of the model, whereas the second is from the residuals of the model itself. It is shown that the bootstraps are asymptotically valid under quite mild conditions. As a consequence of the result we are able to eliminate the apparent drawback of choosing the block length in empirical examples. A small Monte Carlo study of finite sample performance is included.

Keywords: Least squares estimation; long-range estimation; bootstrap methods.

JEL Nos.: C15, C22.

© by Javier Hidalgo. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without special permission provided that full credit, including © notice, is given to the source.

Contact address: Professor Javier Hidalgo, Department of Economics, London School of Economics and Political Science, Houghton Street, London WC2A 2AE.
Email: f.j.hidalgo@lse.ac.uk

1. INTRODUCTION

Since Efron's (1979) seminar paper, Bootstrap algorithms have attracted considerable effort to its development, being perhaps two the main motivations. First, bootstrap methods are capable of approximating the finite sample distribution of statistics more effectively than those based on their asymptotic counterparts. The second being that they allow computing valid asymptotic quantiles of the limiting distribution in situations where *a*) the limiting distribution is unknown or *b*) if known, the practitioner is unable to compute its quantiles. The basic idea of the bootstrap is, given a stretch of data $\mathcal{Z}_T = \{z_t, t = 1, \dots, T\}$ say, to treat the data as if it were the true population, and to carry out Monte-Carlo experiments in which pseudo-data is drawn from \mathcal{Z}_T . Based on the underlying distributional properties of $\{z_t, t = 1, \dots, T\}$, different schemes have been adopted and proposed.

To fix ideas, let us suppose that $z_t = (y_t, x_t)'$ follows the (linear) regression model

$$y_t = \mu_0 + \beta_0' x_t + u_t, \quad t = 1, \dots, T \quad (1.1)$$

where β_0 is a p -dimensional vector, and we are interested in making inferences on β_0 . When u_t is a sequence of independent identically distributed data, abbreviated henceforth as *iid*, the bootstrap entails to obtain a random sample $\{\hat{u}_t^*, t = 1, \dots, T\}$ from the empirical distribution of the (centered) *OLS* residuals $\hat{u}_t = \tilde{y}_t - \hat{\beta}' \tilde{x}_t$, $t = 1, \dots, T$, where

$$\hat{\beta} = \left(\sum_{t=1}^T \tilde{x}_t \tilde{x}_t' \right)^{-1} \sum_{t=1}^T \tilde{x}_t \tilde{y}_t, \quad (1.2)$$

and \tilde{w}_t stands for $w_t - \bar{w}$ with $\bar{w} = T^{-1} \sum_{t=1}^T w_t$. Next, generate the bootstrap version of (1.1) as

$$\tilde{y}_t^* = \hat{\beta}' \tilde{x}_t + \hat{u}_t^*, \quad t = 1, \dots, T, \quad (1.3)$$

and perform the *OLS* estimator in (1.3). This bootstrap is known as Residual Bootstrap, in contrast to the Paired Bootstrap, which entails to draw a random sample from the empirical distribution of $\tilde{z}_t = (\tilde{y}_t, \tilde{x}_t)'$, that is $\tilde{z}_t^* = (\tilde{y}_t^*, \tilde{x}_t^*)'$, and perform the *OLS* of \tilde{y}_t^* on \tilde{x}_t^* .

However, if the errors u_t were not *iid* the above scheme would not be valid, as it was first noted by Singh (1981) in the context of bootstrapping the sample mean of m -dependent data. To circumvent this problem, when the errors are heteroscedastic, Wu (1986) proposed the wild or external bootstrap. This bootstrap amounts to replace \hat{u}_t^* in (1.3) by $|\hat{u}_t| \xi_t$ where ξ_t is a sequence of *iid* $(0, 1)$ random variables and independent of z_t . In a time series setup, following ideas in Carlstein (1986), Künsch (1989) proposed to resample, not from z_t but from (overlapping) blocks of data, say $\mathcal{Z}_\ell = (z'_\ell, \dots, z'_{\ell+b-1})'$ where $\ell = 1, \dots, T - b + 1$. This bootstrap is known as the Moving Blocks Bootstrap (*MBB*). It is worth noting that Efron's (1979) bootstrap is identical to the *MBB* when $b = 1$.

Another approach is subsampling, see Politis and Romano (1994). The subsampling in the context of model (1.1) is as follows. Consider a subsample (block) of the data of size b , $Y_{b,\ell} = (\tilde{y}_\ell, \dots, \tilde{y}_{\ell+b-1})'$ and $X_{b,\ell} = (\tilde{x}_\ell, \dots, \tilde{x}_{\ell+b-1})'$, and compute the *OLS* estimator based on the subsample $Y_{b,\ell}$ and $X_{b,\ell}$, that is

$$\hat{\beta}_{b,\ell} = (X'_{b,\ell} X_{b,\ell})^{-1} X'_{b,\ell} Y_{b,\ell}.$$

Then, the distribution of $T^{1/2}(\hat{\beta} - \beta_0)$, say $\hat{F}_{\hat{\beta}}(z) = \Pr\{T^{1/2}(\hat{\beta} - \beta_0) \leq z\}$, is estimated by

$$\frac{1}{T-b+1} \sum_{\ell=1}^{T-b+1} \mathcal{I}\left(b^{1/2}(\hat{\beta}_{b,\ell} - \hat{\beta}) \leq z\right)$$

where $\mathcal{I}(A)$ denotes the indicator function of the set A . Both methods, subsampling and moving blocks, are similar in that they utilize blocks of data of size b . The important difference is that subsampling looks upon these blocks as "subseries" whereas moving blocks use the blocks as "building stones" to construct new pseudo-time series.

The last two methods have been advocated by Politis et al. (1997) or Fitzenberger (1998) in a model as (1.1) with time series data, motivated by the poor finite sample performance of inferences using the so-called *HAC* estimator. The latter estimator entails the estimation of the spectral matrix at frequency zero of $x_t u_t$, say Ω . In particular, following Parzen (1957), Ω is often estimated by

$$\check{\Omega} = \frac{1}{T} \sum_{r=1}^{T-1} w_{rT} \sum_{t(r)} \tilde{x}_t \hat{u}_t \tilde{x}'_{t+r} \hat{u}_{t+r} \quad (1.4)$$

where $\sum_{t(r)}$ denotes the sum over $1 \leq t, t+r \leq T$ and w_{rT} is a weight function which normally takes the form $w(r/m)$ and where m is a bandwidth parameter increasing slowly to infinity with T , that is $m^{-1} + mT^{-1} \rightarrow 0$. This approach, or versions thereof, has been extensively employed, see Andrews (1991) for a latter reference. All these methods date back to ideas in Jowett (1955) and Hannan (1957) for scalar series to "studentize" its sample mean and further developed by Brillinger (1979) in a multivariate setting. Recent surveys on the literature are den Haan and Levin (1997) and Robinson and Velasco (1997).

One potential drawback, however, of the *MBB* or the subsampling bootstrap is their implementation in empirical examples and in particular, the choice of the block-length b . This apparent drawback is motivated by the observation that, more than anything else, their performance depends rather critically on b , especially for moderate sample sizes. Although some automatic or semiautomatic procedures have recently appeared, see Hall et al. (1995) or Loh's (1987) calibration, the methods can be extremely expensive in computing time.

The purpose of this paper is, under some regularity conditions, to describe and analyse two different approaches which eliminate the problem of the choice of b and the choice of the bandwidth parameter m , as in Robinson (1998), but at the same time able to achieve better finite sample properties. The proposed procedures are easy to implement and computationally no more expensive than other bootstrap methods valid in the context of regression models where the errors are *iid* or heteroscedastic.

We now describe the main ideas of the bootstraps. In the frequency domain the *OLS* estimator given in (1.2) can be written as

$$\hat{\beta} = \left(\sum_{j=1}^{T-1} I_{xx}(\lambda_j) \right)^{-1} \sum_{j=1}^{T-1} I_{xy}(\lambda_j), \quad (1.5)$$

where

$$I_{xx}(\lambda) = w_x(\lambda) w'_x(-\lambda) \quad \text{and} \quad I_{xy}(\lambda) = w_x(\lambda) w'_y(-\lambda)$$

are the periodogram of x_t and cross-periodogram of x_t and y_t respectively, and

$$w_a(\lambda) = \frac{1}{(2\pi T)^{1/2}} \sum_{t=1}^T a_t e^{it\lambda} \quad (1.6)$$

denotes the discrete Fourier transform of a generic sequence of random variables a_t .

It should be noted that because $w_a(\lambda)$ is invariant to additive constants when evaluated at the Fourier frequencies $\lambda_j = (2\pi j)/T$ for integer j , we have that omission of the frequency $j = 0$ (and $j = T$) in (1.5) entails sample-mean correction.

A closer inspection of (1.5) suggests that $\widehat{\beta}$ can be regarded as the *OLS* estimator in the "regression" model

$$w_y(\lambda_j) = \beta'_0 w_x(\lambda_j) + w_u(\lambda_j) \quad j = 1, \dots, T-1, \quad (1.7)$$

where $w_x(\lambda_j)$ and $w_u(\lambda_j)$ play the role of being the regressors and error term respectively. One interesting property of $w_u(\lambda_j)$ is that they are asymptotically uncorrelated although, possibly, heteroscedastic, see Hannan (1970) or Brillinger (1981) among others. It is precisely this observation that motivated Hannan's (1963) (semi-parametric) generalized least squares estimator of β_0 for the model (1.1) and consequently extended to other useful models in econometrics, see Robinson (1991) and references therein. So, looking at (1.7), $v_u(\lambda_j) = w_u(\lambda_j)/|w_u(\lambda_j)|$ can be regarded as a sequence of zero mean and asymptotically independent homoscedastic random variables. This observation and writing (1.7) as

$$w_y(\lambda_j) = \beta'_0 w_x(\lambda_j) + |w_u(\lambda_j)| v_u(\lambda_j) \quad j = 1, \dots, T-1,$$

motivates the bootstrap schemes, differing in *STEP 2* below, which we now describe in the following four steps.

STEP 1 Obtain the *OLS* estimator via (1.2) or (1.5), and the ordinary least squares residuals

$$\widehat{u}_t = \widetilde{y}_t - \widehat{\beta}' \widetilde{x}_t, \quad t = 1, \dots, T.$$

STEP 2 (a) Compute the discrete Fourier transform of the residuals \widehat{u}_t , denoted $w_{\widehat{u}}(\lambda_j)$, and let $v_{\widehat{u}}(\lambda_j) = w_{\widehat{u}}(\lambda_j)/|w_{\widehat{u}}(\lambda_j)|$, $j = 1, \dots, [T/2]$. Draw independent bootstrap residuals $\eta_{j,1}^*$, $j = 1, \dots, [T/2]$, from the empirical distribution function of

$$\widetilde{v}_{\widehat{u}}(\lambda_j) = \widehat{\sigma}_v^{-1} (v_{\widehat{u}}(\lambda_j) - \bar{v}_{\widehat{u}}), \quad j = 1, \dots, [T/2]$$

where $\bar{v}_{\widehat{u}} = [T/2]^{-1} \sum_{j=1}^{[T/2]} v_{\widehat{u}}(\lambda_j)$ and $\widehat{\sigma}_v^2 = [T/2]^{-1} \sum_{j=1}^{[T/2]} |v_{\widehat{u}}(\lambda_j) - \bar{v}_{\widehat{u}}|^2$. That is, for all $j = 1, \dots, [T/2]$,

$$\Pr \{ \eta_{j,1}^* = \widetilde{v}_{\widehat{u}}(\lambda_k) \} = [T/2]^{-1}, \quad k = 1, \dots, [T/2].$$

(b) Let $\widetilde{u}^* = (\widetilde{u}_1^*, \widetilde{u}_2^*, \dots, \widetilde{u}_T^*)'$ be a random sample with replacement from the standardized residuals

$$\widetilde{u}_t = \widetilde{\sigma}_{\widehat{u}}^{-1} \widehat{u}_t; \quad \widetilde{\sigma}_{\widehat{u}}^2 = \frac{1}{T} \sum_{t=1}^T \widehat{u}_t^2,$$

and obtain the "discrete Fourier transform" of \widetilde{u}^* as

$$\eta_{j,2}^* = \frac{1}{T^{1/2}} \sum_{t=1}^T \widetilde{u}_t^* e^{-it\lambda_j}, \quad j = 1, \dots, [T/2].$$

Remark 1. Recall that because (1.1) contains an intercept, the sample mean of \widehat{u}_t is zero.

STEP 3 For $i = 1, 2$, obtain the bootstrap regression

$$w_{y^*,i}(\lambda_j) = \widehat{\beta}' w_x(\lambda_j) + |w_{\widehat{u}}(\lambda_j)| \eta_{j,i}^* \quad j = 1, \dots, [T/2]. \quad (1.8)$$

STEP 4 Compute the bootstrap estimator of $\widehat{\beta}$ as

$$\widehat{\beta}_i^* = \left(\sum_{j=1}^{[T/2]} I_{xx}(\lambda_j) \right)^{-1} \sum_{j=1}^{[T/2]} \text{Re}(I_{xy^*,i}(\lambda_j)), \quad i = 1, 2, \quad (1.9)$$

where $I_{xy^*,i}(\lambda_j) = w_x(\lambda_j) w_{y^*,i}'(-\lambda_j)$ and $\text{Re}(a)$ denotes the real part of the complex number a .

Remark 2. Observe that by symmetry of $I_{xx}(\lambda_j)$ and $I_{xy^*,1}(\lambda_j) = \overline{I_{xy^*,1}(-\lambda_j)}$, where \bar{a} indicates transposition combined with complex conjugation, the estimator $\widehat{\beta}_1^*$ in (1.9) becomes

$$\widehat{\beta}_1^* = \left(\sum_{j=1}^{T-1} I_{xx}(\lambda_j) \right)^{-1} \sum_{j=1}^{T-1} I_{xy^*,1}(\lambda_j)$$

where $\eta_{j,1}^* = \eta_{(T-j),1}^*$, for $j = 1, \dots, [T/2]$. In fact for the validity of the bootstrap, if the latter estimator $\widehat{\beta}_1^*$ were to be employed, it would be crucial to have $\eta_{j,1}^* = \eta_{(T-j),1}^*$. The reason is that if $\eta_{j,1}^*$ and $\eta_{(T-j),1}^*$ were generated as in STEP 2, e.g. independently, then by symmetry, the right side of the last displayed equation would be

$$\widehat{\beta}_1 + \left(2 \sum_{j=1}^{[T/2]-1} I_{xx}(\lambda_j) \right)^{-1} \sum_{j=1}^{[T/2]-1} \left(w_x(\lambda_j) |w_{\widehat{u}}(\lambda_j)| \left(\eta_{j,1}^* + \eta_{(T-j),1}^* \right) \right),$$

whose bootstrap variance would converge to $\Phi/2$ since $E \left| \eta_{j,1}^* + \eta_{(T-j),1}^* \right|^2 = 2$ instead of 4 as would be needed for the validity of the bootstrap.

Remark 3. It is worth mentioning that STEP 2 only requires $\eta_{j,i}^*$, conditional on the data, to be a sequence of iid $(0, 1)$ random variables in the unit sphere. So, Wu's (1986) wild/external bootstrap could be implemented. However, it appears that the proposed bootstrap performs better than the wild bootstrap. One possible reason is that $\eta_{j,i}^*$ is a random sample from the empirical distribution of $v_{\widehat{u}}(\lambda_j)$, so that it may capture better the finite sample distributional features of the model/data than if the sample were drawn from a population with the "same" distributional properties as those from the asymptotic distribution function of $v_{\widehat{u}}(\lambda_j)$ or a distribution as that used by Mammen (1993).

The proposed bootstraps described in STEPS 1-4 eliminate the need to choose the block-length b of Künsch's (1989) *MBB* or Politis and Romano's (1994) subsampling approach, see also Politis et al. (1997, 1999). It should be mentioned that resampling

methods in the frequency domain are not new, see for instance Franke and Härdle (1992), Politis and Romano (1992), Dahlhaus and Janas (1996) or Hidalgo and Kreiss (1999) among others. However, in all the latter aforementioned papers, the resampling techniques are based on the empirical distribution function of the periodogram ordinates and their motivation and emphasis are on the estimation of the spectral density function or functionals of the periodogram, such as the covariance.

The remainder of the paper is organized as follows. In the next section the conditions and main results of the paper are introduced. Section 3 illustrates how the bootstrap can be employed for statistical inferences on the parameters β_0 . Section 4 presents a small Monte-Carlo simulation to illustrate the small sample performance of the bootstraps. Section 5 gives the proofs of our results in Sections 2 and 3, which make use of a Lemma shown in Section 6. Finally Section 7 concludes and discusses extensions to more general models of interest in econometrics.

2. CONDITIONS AND MAIN RESULTS

Denote by $f_{xx}(\lambda)$ and $f_{uu}(\lambda)$ the spectral density matrix and function of x_t and u_t respectively, defined from the relationships

$$\begin{aligned}\gamma_{x,j} &= Cov(x_t, x_{t+j}) = \int_{-\pi}^{\pi} f_{xx}(\lambda) e^{-ij\lambda} d\lambda, \quad j = 0, \pm 1, \dots \\ \gamma_{u,j} &= Cov(u_t, u_{t+j}) = \int_{-\pi}^{\pi} f_{uu}(\lambda) e^{-ij\lambda} d\lambda, \quad j = 0, \pm 1, \dots\end{aligned}$$

Let us introduce the following regularity conditions:

C1 $\{x_t\}$ and $\{u_t\}$ are two covariance stationary linear processes defined as

$$x_t = \sum_{j=0}^{\infty} \zeta_j \xi_{t-j}, \quad \sum_{j=0}^{\infty} \|\zeta_j\|^2 < \infty \quad \text{and} \quad u_t = \sum_{j=0}^{\infty} \vartheta_j \varepsilon_{t-j}; \quad \sum_{j=0}^{\infty} \vartheta_j^2 < \infty,$$

respectively, where ζ_0 is the identity matrix, $\vartheta_0 = 1$, and $\|H\|$ denotes the norm of the matrix H . Moreover, the processes ε_t and ξ_s are uncorrelated for all $t, s = 0, \pm 1, \pm 2, \dots$

Let \mathcal{F}_t and \mathcal{G}_t be the σ -algebras of events generated by ε_s , $s \leq t$ and ξ_s , $s \leq t$, respectively.

C2 $\{\varepsilon_t\}$ is an ergodic process that satisfies (a) $E(\varepsilon_t | \mathcal{F}_{t-1} \cup \mathcal{G}_t) = 0$,

(b) $E(\varepsilon_t^2 | \mathcal{F}_{t-1} \cup \mathcal{G}_t) = E(\varepsilon_t^2) = \sigma_\varepsilon^2$ a.s. and (c) the joint fourth cumulant of ε_{t_i} , $i = 1, \dots, 4$ satisfies

$$cum(\varepsilon_{t_1}, \varepsilon_{t_2}, \varepsilon_{t_3}, \varepsilon_{t_4}) = \begin{cases} \kappa, & t_1 = t_2 = t_3 = t_4, \\ 0, & \text{otherwise,} \end{cases}$$

with $|\kappa| < \infty$.

C3 $\{\xi_t\}$ is a stochastic process that satisfies (a) $E(\xi_t | \mathcal{G}_{t-1} \cup \mathcal{F}_t) = 0$,

(b) $E(\xi_t \xi_t' | \mathcal{G}_{t-1} \cup \mathcal{F}_t) = E(\xi_t \xi_t') = \Xi$ a.s. and (c) the joint fourth cumulant of $\xi_{t_i j_i}$, $j_i = 1, \dots, p$ and $i = 1, \dots, 4$ satisfies

$$cum(\xi_{t_1 j_1}, \xi_{t_2 j_2}, \xi_{t_3 j_3}, \xi_{t_4 j_4}) = \begin{cases} \kappa_{\xi, j_1, j_2, j_3, j_4}, & t_1 = t_2 = t_3 = t_4, \\ 0, & \text{otherwise,} \end{cases}$$

with $\kappa_\xi = \max_{j_i=1,\dots,p,i=1,\dots,4} |\kappa_{\xi,j_1,j_2,j_3,j_4}| < \infty$.

C4 $|(\partial/\partial\lambda) A(\lambda)| = O(|A(\lambda)|/\lambda)$ and $\|(\partial/\partial\lambda) \Psi(\lambda)\| = O(\|\Psi(\lambda)\|/\lambda)$ as $\lambda \rightarrow 0+$, where

$$A(\lambda) = \sum_{j=0}^{\infty} \vartheta_j e^{ij\lambda} \quad \text{and} \quad \Psi(\lambda) = \sum_{j=0}^{\infty} \zeta_j e^{ij\lambda},$$

and such that $|A(\lambda)| > 0$ and $\|\Psi(\lambda)\| > 0$ for all $\lambda \in [0, \pi]$ and continuously differentiable in any open set outside the origin. In addition, for all $g = 1, \dots, p+1$, $f_{gg}^{-1/2}(\lambda) |\eta_g(\lambda)|$ is a non-zero finite vector, where $\eta_g(\lambda)$ denotes the g th row of $\text{diag}(\Psi(\lambda), A(\lambda))$ and $f_{gg}(\lambda)$ is the g th diagonal element of $f(\lambda) = \text{diag}(\text{diag}(f_{xx}(\lambda)), f_{uu}(\lambda))$.

C5 $\int_{-\pi}^{\pi} \|f_{xx}(\lambda) f_{uu}(\lambda)\| d\lambda < \infty$ and $E x_t x_t' = \Sigma > 0$.

We now comment on the conditions, which for the most part are similar to those used in Robinson (1995a,b,1998). The first three conditions are restricted in the linearity they impose but not otherwise. Part (a) of Conditions C2 and C3, together with the first part of C1, indicate that the best linear predictor is the best predictor, in the least squares sense. The last part of Condition C1 and part (b) of Condition C2 are presumably stronger than needed and some heterogeneity could be allowed, although it rules out heteroscedasticity. However, as the main motivation of the paper is to illustrate the possibility of avoiding the drawbacks of the *MBB* or subsampling bootstrap in a model like (1.1), we have preferred keeping them as they stand. Another motivation is due to the fact that these conditions allow us to consistently estimate the asymptotic covariance matrix of the *OLS* without resorting to any bandwidth or tuning parameter as Robinson (1988) proved. In particular that

$$\tilde{\Omega} = \frac{4\pi^2}{T} \sum_{j=1-[T/2]}^{[T/2]-1} I_{xx}(\lambda_j) I_{\hat{u}\hat{u}}(\lambda_j) \quad (2.1)$$

is a consistent estimator of $2\pi \int_{-\pi}^{\pi} f_{xx}(\lambda) f_{uu}(\lambda) d\lambda$.

Moreover, Conditions C1 – C3 also imply that $E(x_t u_t x_0 u_0) = E(x_t x_0) E(u_t u_0)$ for all t , so that the spectral density of $x_t u_t$ at frequency zero, $f_{xu}(0)$, is $2\pi \int_{-\pi}^{\pi} f_{xx}(\lambda) f_{uu}(\lambda) d\lambda$. It seems that for (2.1) to be a consistent estimator of $f_{xu}(0)$, the previous condition is required. Another set of conditions for which $E(x_t u_t x_0 u_0) = E(x_t x_0) E(u_t u_0)$ holds is if x_j and u_0 are independent for all negative and nonnegative values of j or that x_t and u_t are Gaussian and $\text{Cov}(x_j, u_0) = 0$ for all $j \leq 0$. In the terminology of Engle, Hendry and Richard (1983), they imply strong exogeneity of x_t or that x_t is predetermined respectively. So, we are implicitly assuming that the practitioner is willing to take seriously this assumption. However, as discussed by Robinson (1988), if this were not the case, then (2.1) would only converge to one of the components of the long run variance of $x_t u_t$, that is Ω , being the other component a function of the fourth cumulants. This component can be estimated using results in Taniguchi (1982) or Keenan (1987), although under long-memory no results are available. It should be mentioned that in this case, the proposed bootstrap is not valid. But, following ideas and arguments similar to those of Hall and Horowitz (1996), when analysing the t -statistic or the J -statistic for overidentifying restrictions in a *GMM* framework, similar corrections should apply in our context. However, we do not proceed as that will involve technicalities and it may obscure the main message of the paper. That

is, that even with time series data, there are setups encountered in applied work for which *MBB*/subsampling is not required. On the other hand, in some respects our conditions are quite general. For instance, it allows for long memory dependence for which the *MBB* has yet to be justified in a general setting as ours, although the latter algorithm is valid if x_t and u_t are strong mixing without assuming that x_t is strong exogenous or u_t homoscedastic.

Condition *C4* effectively allows for a possible singularity of the spectral density matrix and function of x_t and u_t respectively to be at frequency zero, but smooth elsewhere. This is done merely for notational convenience. The results do not depend on this assumption and they follow similarly if the singularity(ies) were located at some other(s) frequency(ies). So, we allow x_t and/or u_t to be, possibly, a long-memory process which has attracted immense attention in recent years in econometric literature. An example of a (scalar) process whose spectral density function satisfies *C4* is the Fractional Integrated Autoregressive Moving Average (*FARIMA*) model, see Granger and Joyeux (1980) or Hosking (1981). The *FARIMA* model has a spectral density function defined as

$$f(\lambda) = \frac{\sigma_\varepsilon^2}{2\pi} |1 - e^{i\lambda}|^{-d} \frac{|\Theta(e^{i\lambda}; \theta)|^2}{|\Delta(e^{i\lambda}; \theta)|^2}, \quad 0 \leq \lambda \leq \pi, \quad (2.2)$$

where $0 < C < \infty$ and $0 \leq d < 1/2$, and where Δ and Θ are the autoregressive and moving average polynomials respectively and such that they have no common roots and are outside the unit circle. An earlier example of a process exhibiting long-memory is the fractional Gaussian noise (*fgn*) process, introduced by Mandelbrot and Van Ness (1968), and whose autocorrelation structure is given by

$$\gamma_j = \frac{1}{2} \left(|j-1|^{1+2d} - 2|j|^{1+2d} + |j+1|^{1+2d} \right). \quad (2.3)$$

It is known that both models (2.2) and (2.3) satisfy

$$f(\lambda) \sim C\lambda^{-2d}, \text{ as } \lambda \rightarrow 0+.$$

The first part of Condition *C5*, that is $\int_{-\pi}^{\pi} \|f_{xx}(\lambda) f_{uu}(\lambda)\| d\lambda < \infty$, seems to be very mild and due to results in Robinson (1994) and Robinson and Hidalgo (1997), it also appears to be necessary and minimal for the central limit theorem of $\hat{\beta}$ to hold true. This condition is satisfied if, for instance, u_t follows a *FARIMA* or an *fgn* model with $d = d_u$ and x_t follows a model whose l th element is a *FARIMA* or an *fgn* model with $d = d_\ell$ satisfying the condition $2(d_\ell + d_u) < 1$ for all $\ell = 1, \dots, p$. Alternatively, *C5* could have been written in terms of the autocovariance function of x_t and u_t . *C5* could be replaced by assuming that $\gamma_{u,j} \gamma_{x,j}$ is summable under our framework. It entails that the spectral density function of $x_t u_t$ to be finite and positive at frequency zero. It is important to note that this is implied by summability of $|\gamma_{u,j} \gamma_{x,j}|$ but not the other way round. A standard example is when x_t exhibits strong "cyclical" behaviour as is the case with Gengebauer models. For the latter models, the autocovariance function behaves as $Cj^{2d-1} \cos(j\lambda)$ with $\lambda > 0$ and $0 \leq d < \frac{1}{2}$. It should be mentioned that if the possible singularities of f_x and f_u do not coincide, then d_ℓ and d_u can take any value smaller than $\frac{1}{2}$. Finally, the second part of Condition *C5* is a standard condition on the regressors x_t and needed for the identification of the parameters β_0 .

We then achieve the following:

Theorem 2.1. *Assuming C1-C5, as $T \rightarrow \infty$,*

$$T^{1/2} \left(\widehat{\beta} - \beta_0 \right) \xrightarrow{d} \mathcal{N} \left(0, \Sigma^{-1} \Omega \Sigma^{-1} \right).$$

We now focus on the properties of the bootstrap estimators $\widehat{\beta}_i^*$, $i = 1, 2$, given in (1.9). For the latter to be valid, the resampling scheme must be such that the conditional distribution of $T^{1/2} \left(\widehat{\beta}_i^* - \widehat{\beta} \right)$, for $i = 1, 2$, given $\left(\widetilde{x}', \widetilde{u}' \right)'$ where $\widetilde{x} = (x_1, \dots, x_T)'$ and $\widetilde{u} = (u_1, \dots, u_T)'$, consistently estimates the limiting distribution of $T^{1/2} \left(\widehat{\beta} - \beta_0 \right)$, that is

$$\Pr \left\{ T^{1/2} \left(\widehat{\beta}_i^* - \widehat{\beta} \right) \leq z \mid \left(\widetilde{x}', \widetilde{u}' \right)' \right\} \xrightarrow{P} \mathcal{G}(z), \quad \text{for } i = 1, 2$$

where $\mathcal{G}(z)$ is the probability distribution function of a $\mathcal{N} \left(0, \Sigma^{-1} \Omega \Sigma^{-1} \right)$ random variable, see Giné and Zinn (1989). Such a convergence will be denoted as “ $\xrightarrow{d^*}$ ”.

To prove the validity of the bootstraps described in *STEPS 1-4*, we need to strengthen slightly Condition C5.

C5' $\int_{-\pi}^{\pi} \|f_{xx}(\lambda) f_{uu}(\lambda)\| |\log(|\lambda|)| d\lambda < \infty$ and $E x_t x_t' = \Sigma > 0$.

Theorem 2.2. *Assuming C1-C4 and C5', for $i = 1, 2$ as $T \rightarrow \infty$*

$$T^{1/2} \left(\widehat{\beta}_i^* - \widehat{\beta} \right) \xrightarrow{d^*} \mathcal{N} \left(0, \Sigma^{-1} \Omega \Sigma^{-1} \right).$$

So Theorem 2 indicates that the bootstraps are consistent, and more importantly, that we can avoid the drawback of the *MBB* and the subsampling bootstrap, e.g. the choice of the block length b . Moreover, the results of Theorem 2 will allow us the implementation of tests for the parameters β_0 , say, such as that given in (3.1) below. This will be addressed in the next section.

3. APPLICATIONS OF THE BOOTSTRAP

Suppose that we are interested in the following hypothesis testing

$$H_0 : C\beta_0 = c \quad \text{against} \quad H_1 : C\beta_0 \neq c, \quad (3.1)$$

where C is a full rank $(r \times p)$ matrix and c is an $(r \times 1)$ vector. A common method to test H_0 is based on the *F - statistic*

$$F = T \left(C\widehat{\beta} - c \right)' \left(C\widehat{\Phi}C' \right)^{-1} \left(C\widehat{\beta} - c \right), \quad (3.2)$$

where $\widehat{\Phi} = \widehat{\Sigma}^{-1} \widetilde{\Omega} \widehat{\Sigma}^{-1}$ is a consistent estimator of $\Phi = \text{AsyVar} \left(T^{1/2} \left(\widehat{\beta} - \beta_0 \right) \right)$, with $\widetilde{\Omega}$ given in (2.1) and $\widehat{\Sigma} = (2\pi) T^{-1} \sum_{j=1}^{T-1} I_{xx}(\lambda_j)$.

Suppose that we want to bootstrap the test F in (3.2), being denoted by F^* . For F^* to be asymptotically valid, it should satisfy two basic requirements. First, under H_0 or local alternatives, e.g. $H_a : C\beta_0 = c + D/T^{1/2}$ for any vector D with bounded

norm, the conditional distribution of F^* , given the data $\left(\begin{smallmatrix} x' \\ \tilde{u} \end{smallmatrix}, u'\right)'$, should converge in probability to the limiting distribution of F given in (3.2). The second requirement on F^* is that, under H_1 , F^* should be bounded in probability. The latter is achieved once the bootstrap sample/model is obtained under H_0 .

More specifically, for $i = 1, 2$, let $\hat{\beta}_i^*$ denote the bootstrap estimator of β given in (3.6) below. Let us introduce $I_{\hat{u}^* \hat{u}^*, i}(\lambda_j) = |w_{\hat{u}^*, i}(\lambda_j)|^2$, where

$$w_{\hat{u}^*, i}(\lambda_j) = w_{y^*, i}(\lambda_j) - \hat{\beta}_i^{*'} w_x(\lambda_j), \quad j = 1, \dots, [T/2],$$

with $w_{y^*, i}(\lambda_j)$ as defined in (3.4) below. Then the bootstrap test F^* is defined as F in (3.2) but with $\hat{\beta}$ and $\hat{\Phi} = \hat{\Sigma}^{-1} \hat{\Omega} \hat{\Sigma}^{-1}$ being replaced by their bootstrap counterparts $\hat{\beta}_i^*$ and $\hat{\Phi}_i^* = \hat{\Sigma}^{-1} \hat{\Omega}_i^* \hat{\Sigma}^{-1}$ respectively, and where

$$\hat{\Omega}_i^* = \frac{4\pi^2}{T} \sum_{j=1}^{[T/2]-1} I_{xx}(\lambda_j) I_{\hat{u}^* \hat{u}^*, i}(\lambda_j), \quad i = 1, 2. \quad (3.3)$$

We need to show that

$$\Pr \left\{ T^{1/2} C \left(\hat{\beta}_i^* - \tilde{\beta} \right) \leq z \mid \left(\begin{smallmatrix} x \\ \tilde{u} \end{smallmatrix}, u \right) \right\} \xrightarrow{P} \lim_{T \rightarrow \infty} \Pr \left\{ T^{1/2} C \left(\hat{\beta} - \beta_0 \right) \leq z \right\}$$

when $C\beta_0 = c + D/T^{1/2}$ for any vector D with bounded norm, where $\tilde{\beta}$ is defined in (3.5) below, and

$$\Pr \left\{ T^{1/2} C \left(\hat{\beta}_i^* - \tilde{\beta} \right) \leq z \mid \left(\begin{smallmatrix} x \\ \tilde{u} \end{smallmatrix}, u \right) \right\} = O_{p^*}(1)$$

when $C\beta_0 \neq c + D/T^{1/2}$ for all $|D| \leq T^{1/2}$, that is under H_1 .

To achieve this goal, we proceed as in Mammen (1993) in the context of *iid* data. That is,

STEPS 1 AND 2 The same as those given in the introduction.

STEP 3 For $i = 1, 2$, obtain the bootstrap regression

$$w_{y^*, i}(\lambda_j) = \tilde{\beta}' w_x(\lambda_j) + |w_{\tilde{u}}(\lambda_j)| \eta_{j, i}^*, \quad j = 1, \dots, [T/2], \quad (3.4)$$

where $\tilde{\beta}$ is the restricted least squares, that is

$$\tilde{\beta} = \hat{\beta} + (X'X)^{-1} C' \left(C(X'X)^{-1} C' \right)^{-1} \left(c - C\hat{\beta} \right). \quad (3.5)$$

STEP 4 Compute the bootstrap counterpart of $\hat{\beta}$ as the *OLS* estimator from the regression model $w_{y^*, i}(\lambda_j)$ on $w_x(\lambda_j)$, that is

$$\hat{\beta}_i^* = \left(\sum_{j=1}^{[T/2]} I_{xx}(\lambda_j) \right)^{-1} \sum_{j=1}^{[T/2]} \text{Re}(I_{xy^*, i}(\lambda_j)), \quad (3.6)$$

where $I_{xy^*, i}(\lambda_j) = w_x(\lambda_j) w'_{y^*, i}(-\lambda_j)$ for $i = 1, 2$.

STEP 5 For $i = 1, 2$, the bootstrap test is given by

$$F_i^* = \varphi^* \left(T^{1/2} C \left(\widehat{\beta}_i^* - \widetilde{\beta} \right) \right),$$

where $\widehat{\Phi}_i^* = \widehat{\Sigma}^{-1} \widehat{\Omega}_i^* \widehat{\Sigma}^{-1}$ with $\widehat{\Omega}_i^*$ given in (3.3) and $\varphi^*(a) = a' \left(C \widehat{\Phi}_i^* C' \right)^{-1} a$.

STEP 6 For $i = 1, 2$, reject H_0 if $F \geq c_{(1-\alpha),i}^*$, where $c_{(1-\alpha),i}^*$ is the $(1-\alpha)$ th quantile of the bootstrap distribution of F_i^* , that is

$$\Pr \left\{ \varphi^* \left(T^{1/2} C \left(\widehat{\beta}_i^* - \widetilde{\beta} \right) \right) \leq z \mid \left(\underset{\sim}{x}, \underset{\sim}{u} \right) \right\}.$$

Theorem 3.1. *Assuming C1 – C4 and C5', under $H_0 \cup H_1$ in (3.1), for $i = 1, 2$ as $T \rightarrow \infty$*

$$F_i^* \xrightarrow{d^*} \chi_r^2,$$

where $r = \text{rank}(C)$.

The result of Theorem 3 allows us to implement the (bootstrap) test as follows. Denote $\varphi(a) = a' \left(C \widehat{\Phi} C' \right)^{-1} a$ and let $c_{T,(1-\alpha)}^f$ and $c_{(1-\alpha)}^a$ be such that

$$\Pr \left\{ \left| \varphi \left(T^{1/2} \left(C \widehat{\beta} - c \right) \right) \right| > c_{T,(1-\alpha)}^f \right\} = \alpha$$

and

$$\lim_{T \rightarrow \infty} \Pr \left\{ \left| \varphi \left(T^{1/2} \left(C \widehat{\beta} - c \right) \right) \right| > c_{(1-\alpha)}^a \right\} = \alpha,$$

respectively. Then, Theorem 1 indicates that $c_{T,(1-\alpha)}^f \rightarrow c_{(1-\alpha)}^a$, whereas Theorem 3 indicates that $c_{(1-\alpha),i}^*$ satisfies $c_{(1-\alpha),i}^* \xrightarrow{P} c_{(1-\alpha)}^a$ where

$$\Pr \left\{ \varphi \left(T^{1/2} C \left(\widehat{\beta}_i^* - \widetilde{\beta} \right) \right) > c_{(1-\alpha),i}^* \mid \left(\underset{\sim}{x}, \underset{\sim}{u} \right) \right\} = \alpha,$$

for $i = 1, 2$. (Observe that $C \widetilde{\beta} = c$.)

However, since the finite sample distribution of $T^{1/2} C \left(\widehat{\beta}_i^* - \widetilde{\beta} \right)$, and thus that of $\varphi^* \left(T^{1/2} C \left(\widehat{\beta}_i^* - \widetilde{\beta} \right) \right)$, is not available, $c_{(1-\alpha),i}^*$ is approximated as accurately as desired by a standard Monte-Carlo simulation algorithm. To that end, consider the bootstrap samples $\eta^{*(\ell)} = \left(\eta_1^{*(\ell)}, \dots, \eta_{\lfloor T/2 \rfloor}^{*(\ell)} \right)'$ for $\ell = 1, \dots, B$, and compute $\widehat{\beta}_1^{*(\ell)}$ as in (3.6) for each ℓ . Then, $c_{(1-\alpha),1}^*$ is approximated by the value $c_{(1-\alpha),1}^{*B}$ that satisfies

$$\frac{1}{B} \sum_{\ell=1}^B \mathcal{I} \left(\varphi \left(T^{1/2} C \left(\widehat{\beta}_1^{*(\ell)} - \widetilde{\beta} \right) \right) > c_{(1-\alpha),1}^{*B} \right) = \alpha,$$

whereas for the second bootstrap, consider the bootstrap samples $\widetilde{u}^{*(\ell)} = \left(\widetilde{u}_1^{*(\ell)}, \dots, \widetilde{u}_T^{*(\ell)} \right)'$ for $\ell = 1, \dots, B$, and compute $\widehat{\beta}_2^{*(\ell)}$ as in (3.6) for each ℓ . Then, $c_{(1-\alpha),2}^*$ is approximated by the value $c_{(1-\alpha),2}^{*B}$ that satisfies

$$\frac{1}{B} \sum_{\ell=1}^B \mathcal{I} \left(\varphi \left(T^{1/2} C \left(\widehat{\beta}_2^{*(\ell)} - \widetilde{\beta} \right) \right) > c_{(1-\alpha),2}^{*B} \right) = \alpha.$$

4. MONTE CARLO SIMULATION

In order to investigate how well the bootstrap estimators $\hat{\beta}_i^*$, for $i = 1, 2$, perform in finite samples, a small Monte Carlo experiment was conducted. In (1.1) we took $\mu_0 = \beta_0 = 1$. We have considered two sets of experiments. In the first set, x_t and/or u_t are long-memory whereas in the second set both x_t and u_t are strong-mixing sequences. The main motivation to consider the second set of experiments is to assess the relative performance of $\hat{\beta}_i^*$ to the *MBB* in a framework for which theoretical results are available for the latter bootstrap. So that a more "fair" comparison is given between the two methods. All throughout the experiments, we have calculated the size of the tests for the null hypothesis $H_0: \beta_0 = 1$. To that end, we have employed 5000 replications with samples sizes $T = 32, 64$ and 128 . In order to calculate the bootstrap for all the models and sample sizes considered, 2000 bootstrap samples were employed, that is we have chosen $B = 2000$.

In the first set of simulations, two groups of six models were generated. Specifically, for the first group, x_t was generated as a fractional Gaussian noise process (2.3) with parameter $d_x = 0.3$ and 0.45 respectively. On the other hand, u_t was generated as an *AR*(1) process

$$u_t = \rho u_{t-1} + \varepsilon_t \quad (4.1)$$

with $\rho = 0.0, 0.5$ and 0.9 respectively, with ε_t being a sequence of *iid* $\mathcal{N}(0, 1)$ random variables. For the second group, x_t was generated as a fractional Gaussian noise process (2.3) with parameter $d_x = 0.3$ and 0.4 respectively, whereas u_t was generated by (4.1) with ε_t as a fractional Gaussian noise process (2.3) with parameter $d_u = 0.15$ and 0.05 respectively, so that $d_u + d_x = 0.45$ in the latter case. Both x_t and ε_t were generated using an algorithm due to Davies and Harte (1987).

In addition, to assess the relative finite sample performance of the bootstrap test, we have compared the results with those obtained using (a) the χ_1^2 limiting distribution of

$$F = T^{1/2} \left(C\hat{\beta} - c \right) \left(C\hat{\Sigma}^{-1} \tilde{\Omega} \hat{\Sigma}^{-1} C' \right)^{-1} T^{1/2} \left(C\hat{\beta} - c \right) \quad (4.2)$$

where $\tilde{\Omega}$ was defined in (2.1), and (b) the *MBB* with four different block-lengths b , namely $b = 2, 4, 8$ and 16 . It is worth noting that $\tilde{\Omega}$ is the frequency domain counterpart of the more natural estimator of Ω , $\hat{\Omega} = \sum_{r=1}^{T-1} \hat{\gamma}_{x,r} \hat{\gamma}_{u,r}$, where $\hat{\gamma}_{x,r} = T^{-1} \sum_{t(r)} \tilde{x}_t \tilde{x}'_{t+r}$ and $\hat{\gamma}_{u,r} = T^{-1} \sum_{t(r)} \hat{u}_t \hat{u}'_{t+r}$. Following Robinson (1998), under $C1 - C5$, both $\hat{\Omega}$ and $\tilde{\Omega}$ are consistent estimators of Ω .

For the first set of experiments, the results for the first and second group of models are displayed in *TABLES 4.1 and 4.2* respectively, where *ASY* denotes the results of the test using (4.2), whereas *BOOT1* and *BOOT2* are those using the bootstrap approaches for $i = 1, 2$ respectively. Finally, the results for the *MBB* are only given for $T = 128$ and they are displayed in *TABLE 4.3*. The *MBB* algorithm was implemented as follows. Given the *OLS* residuals $\hat{u}_1, \dots, \hat{u}_T$, we considered blocks of size b , $\hat{U}_\ell = (\hat{u}_\ell, \dots, \hat{u}_{\ell+b-1})'$, $\ell = 1, \dots, [T/b]$. Then, a random sample from \hat{U}_ℓ , $\ell = 1, \dots, [T/b]$, was obtained and thus a bootstrap sample $\hat{U}^* = (\hat{u}_1^*, \dots, \hat{u}_T^*)'$. Given \hat{U}^* , we compute the *DFT*, denoted $w_{\hat{u}^*}(\lambda_j)$, and the bootstrap regression

$$w_{y^*}(\lambda_j) = \tilde{\beta}' w_x(\lambda_j) + w_{\hat{u}^*}(\lambda_j) \quad j = 1, \dots, [T/2]$$

and perform the *OLS*. Observe that we could have similarly computed the bootstrap

regression model as

$$y_t^* = \tilde{\beta}' \tilde{x}_t + \hat{u}_t^*, \quad t = 1, \dots, T,$$

and compute the *OLS* in the latter.

TABLES 4.1 TO 4.3 ABOUT HERE

The overall picture of *TABLES 4.1 to 4.3* indicates that inferences based on the proposed bootstraps work substantially better than inferences based on the asymptotic χ_1^2 distribution of the *OLS* estimator $\hat{\beta}$ and with Ω being estimated by (2.1) or those inferences based on the *MBB*. It seems that the performance of *BOOT1* tends to be better than *BOOT2* for larger sample sizes, whereas *BOOT2* tends to perform better than *BOOT1* for smaller sample sizes, although the difference is not very substantial. At the 1% level, the empirical sizes of the bootstrap tests are somewhat disappointing compared to that obtained using the limiting distribution of F . However, as the sample size increases, the performance of the *BOOT1* and *BOOT2* tend to be similar to that of the *ASY* at the 1% significance level. On the other hand, at the 5% and 10% levels, the *BOOT1* and *BOOT2* clearly dominate the performance of the *ASY* and even for sample sizes as small as 32, the performance of the *BOOT1* and *BOOT2* tests are outstanding. Finally, when comparing *BOOT1* and *BOOT2* with the *MBB*, the proposed bootstraps clearly outperform the *MBB*, indicating that the former ones possess better finite sample properties than the latter.

For our second set of experiments, we have considered six models. Specifically, x_t and u_t were generated as *AR*(1) with Gaussian innovations and parameters $\rho_x = 0.5, 0.9$ and $\rho_u = 0.0, 0.5, 0.9$ respectively. The results for *BOOT1*, *BOOT2* and *ASY* are displayed in *TABLE 4.4*, whereas those using the *MBB* are in *TABLE 4.5* below.

TABLES 4.4 AND 4.5 ABOUT HERE

The first conclusion we can draw from *TABLES 4.4 and 4.5* is that they are qualitatively similar to those obtained for the first set of experiments. When we compare the performance of the bootstraps *BOOT1* and *BOOT2* to *ASY* we observe that there is substantial gain by using the bootstrap approach, specially for very small sample sizes, being the gain more noticeable the higher the autocorrelation of x_t becomes. Moreover, it seems that the autocorrelation of the regressors have a more negative impact than that from the errors. In fact, it appears that the performance of the bootstrap is very uniform across the different correlation levels of u_t . Finally, as it happened with the first set of experiments, the results for the *MBB* are far worst than those obtained for *BOOT1* and *BOOT2*. Also, the results for the *MBB* appear to be in agreement with those obtained elsewhere, see for instance Fitzenberger (1988).

5. PROOFS

5.1. Proof of Theorem 1

The proof of the theorem will be split into two steps. Assuming that the integrable function $g(\lambda)$ satisfies the same conditions of $f_{xx}(\lambda)$ or $f_{uu}(\lambda)$ in C4, in Proposition 1 a Central Limit Theorem is achieved for the weighted cross-periodogram of ε_t and ξ_t with weights $g^{1/2}(\lambda_j)$. On the other hand, in Proposition 2 we show that once $g(\lambda)$ is identified as $f_{xx}(\lambda)f_{uu}(\lambda)$, the average cross-periodogram of x_t and u_t converges in distribution to a normal random variable. For notational convenience, in the proof

of the next two results we will assume, without loss of generality, that x_t is a scalar sequence of random variables.

Proposition 1. *Let $g(\lambda)$ be a symmetric integrable function in $[-\pi, \pi]$ which satisfies the same conditions of $f_{xx}(\lambda)$ and $f_{uu}(\lambda)$ in C_4 . Assuming $C1-C3$, as $T \rightarrow \infty$,*

$$\frac{2\pi}{T^{1/2}} \sum_{j=1}^{T-1} g^{1/2}(\lambda_j) w_\xi(-\lambda_j) w_\varepsilon(\lambda_j) \xrightarrow{d} \mathcal{N}\left(0, (2\pi)^{-1} \sigma_\xi^2 \sigma_\varepsilon^2 \int_{-\pi}^{\pi} g(\lambda) d\lambda\right), \quad (5.1)$$

where σ_ξ^2 and σ_ε^2 denote the variances of ξ_t and ε_t respectively.

Proof. The left side of (5.1) can be written as

$$\frac{1}{T^{1/2}} \sum_{t=1}^T \varepsilon_t \left(\sum_{s=1}^T \xi_s h_{t-s} \right)$$

where

$$h_t = \frac{1}{T} \sum_{j=1}^{T-1} g^{1/2}(\lambda_j) e^{it\lambda_j} \quad (5.2)$$

suppressing any reference to T in the definition of h_t . (5.1) is implied if the convergence holds true conditional on $\{\xi_t\}$ and we establish the latter. Denote $d_t = T^{-1/2} \sum_{s=1}^T \xi_s h_{t-s}$. Then following Scott (1973), see also Robinson and Hidalgo (1997), the proposition is proven if, as $T \rightarrow \infty$,

$$(a) \sum_{t=1}^T d_t^2 E(\varepsilon_t^2 | \mathcal{F}_{t-1} \cup \mathcal{G}_t) \xrightarrow{P} (2\pi)^{-1} \sigma_\xi^2 \sigma_\varepsilon^2 \int_{-\pi}^{\pi} g(\lambda) d\lambda \quad (5.3)$$

and for any $\psi > 0$,

$$(b) E\left(\sum_{t=1}^T d_t^2 E\left(\varepsilon_t^2 \mathcal{I}\left(|d_t \varepsilon_t|^2 > \psi\right) | \{\xi_t\}\right)\right) \rightarrow 0, \quad (5.4)$$

where $\mathcal{I}(A)$ denotes the indicator function of the set A . In fact, we can see that this is a *CLT* applied to $d_t \varepsilon_t$, a martingale difference sequence with respect to the σ -algebra generated by $\{\xi_i, i = 1, \dots, T; \varepsilon_j, j = 1, \dots, t-1\}$.

We begin with (a). By $C2$, the left side of (5.3) is

$$\sigma_\varepsilon^2 \sum_{t=1}^T d_t^2 = \frac{\sigma_\varepsilon^2}{T} \sum_{t=1}^T \left(\sum_{s=1}^T \xi_s h_{t-s} \right)^2, \quad (5.5)$$

whose first moment is, by $C3$,

$$\begin{aligned} \frac{\sigma_\varepsilon^2 \sigma_\xi^2}{T} \sum_{t=1}^T \sum_{s=1}^T h_{t-s}^2 &= \frac{\sigma_\varepsilon^2 \sigma_\xi^2}{T^3} \sum_{t,s=1}^T \sum_{j,k=1}^{T-1} g^{1/2}(\lambda_j) g^{1/2}(\lambda_k) e^{i(t-s)(\lambda_j - \lambda_k)} \\ &= \frac{\sigma_\varepsilon^2 \sigma_\xi^2}{T} \sum_{j=1}^{T-1} g(\lambda_j) \end{aligned}$$

since $\sum_{t=1}^T e^{it(\lambda_j - \lambda_k)} = 0$ unless $j = k$ in which case it is T . But, Lemma 6.1 implies that the right side of the last displayed equation converges to $(2\pi)^{-1} \sigma_\xi^2 \sigma_\varepsilon^2 \int_{-\pi}^{\pi} g(\lambda) d\lambda$.

Next, the second moment of the right side of (5.5) is

$$\begin{aligned} & \frac{\sigma_\varepsilon^4}{T^2} \sum_{t,r=1}^T E \left[\left(\sum_{s=1}^T \xi_s h_{t-s} \right)^2 \left(\sum_{s=1}^T \xi_s h_{r-s} \right)^2 \right] \\ &= \frac{\sigma_\varepsilon^4 \sigma_\xi^4}{T^2} \sum_{t,r=1}^T \left(\sum_{s=1}^T h_{t-s}^2 \right) \left(\sum_{s=1}^T h_{r-s}^2 \right) + \frac{2\sigma_\varepsilon^4 \sigma_\xi^4}{T^2} \sum_{t,r=1}^T \left(\sum_{s=1}^T h_{t-s} h_{r-s} \right)^2 \\ & \quad + \frac{\sigma_\varepsilon^4 \text{cum}(\xi_s, \xi_s, \xi_s, \xi_s)}{T^2} \sum_{t,r,s=1}^T h_{t-s}^2 h_{r-s}^2, \end{aligned} \quad (5.6)$$

by C3. The first term on the right of (5.6) is

$$\left(\frac{\sigma_\varepsilon^2 \sigma_\xi^2}{T} \sum_{t=1}^T \left(\sum_{s=1}^T h_{t-s}^2 \right) \right)^2$$

which converges to the square of $(2\pi)^{-1} \sigma_\xi^2 \sigma_\varepsilon^2 \int_{-\pi}^{\pi} g(\lambda) d\lambda$, proceeding as we did with the first term of (5.5). So, to complete the proof of part (a), it suffices to show that the last two terms on the right of (5.6) converge to 0, since the latter will imply that the second moment of (5.5) converges to the square of the first moment and so (5.5) converges in probability to the right of (5.3) by Markov inequality.

The third term on the right of (5.6) is, using the definition of h_t in (5.2), proportional to

$$\begin{aligned} & \frac{1}{T^6} \sum_{t,r,s=1}^T \sum_{j,k,\ell,q=1}^{T-1} g^{1/2}(\lambda_j) g^{1/2}(\lambda_k) g^{1/2}(\lambda_\ell) g^{1/2}(\lambda_q) e^{i((t-s)(\lambda_j - \lambda_k) - (r-s)(\lambda_\ell - \lambda_q))} \\ &= \frac{1}{T^2} \sum_{\substack{t,r,s=1 \\ t=r=s}}^T \left(\frac{1}{T} \sum_{j=1}^{T-1} g^{1/2}(\lambda_j) \right)^4 \\ & \quad + \frac{3}{T^3} \sum_{\substack{t,r,s=1 \\ t \neq r=s}}^T \left[\left(\frac{1}{T} \sum_{j=1}^{T-1} g(\lambda_j) \right) \left(\frac{1}{T^2} \sum_{\ell,q=1}^{T-1} g^{1/2}(\lambda_\ell) g^{1/2}(\lambda_q) \right) \right] \\ & \quad + \frac{1}{T^4} \sum_{\substack{t,r,s=1 \\ t \neq r \neq s}}^T \left(\frac{1}{T^2} \sum_{\ell,q=1}^{T-1} g(\lambda_\ell) g(\lambda_q) \right) \end{aligned}$$

which is bounded by

$$\frac{D}{T} \left[\left(\frac{1}{T} \sum_{j=1}^{T-1} g^{1/2}(\lambda_j) \right)^4 + \left(\frac{1}{T} \sum_{j=1}^{T-1} g^{1/2}(\lambda_j) \right)^2 \left(\frac{1}{T} \sum_{j=1}^{T-1} g(\lambda_j) \right) + \left(\frac{1}{T} \sum_{j=1}^{T-1} g(\lambda_j) \right)^2 \right].$$

But the last displayed expression is $O(T^{-1})$ since Lemma 6.1 implies that $T^{-1} \sum_{j=1}^{T-1} g(\lambda_j) = O(1)$ and Schwarz inequality that $\left| T^{-1} \sum_{j=1}^{T-1} g^{1/2}(\lambda_j) \right|^2 \leq T^{-1} \sum_{j=1}^{T-1} g(\lambda_j)$. The

second term on the right of (5.6), proceeding as with the third, is easily shown to be $O(T^{-1})$ also and so it is omitted. This concludes the proof of part (a).

Next, we show part (b). For any $\delta > 0$, the left side of (5.4) is bounded by

$$E \left[\sum_{t=1}^T d_t^2 E(\varepsilon_t^2 \mathcal{I}(\varepsilon_t^2 > \psi/\delta) | \{\xi_t\}) \right] + \Pr \left\{ \max_t d_t^2 > \delta \right\}. \quad (5.7)$$

The first term of (5.7) can be made arbitrarily small by choosing δ small enough, from the uniform integrability of ε_t^2 . (Recall that a sufficient condition for the latter property is that $E|\varepsilon_t|^{2+\kappa} < D$ for some $\kappa > 0$.) Next,

$$\max_t d_t^2 = \max_t T^{-1} \left| \sum_{s=1}^T \xi_s h_{t-s} \right|^2 \leq \left(\frac{D}{T^2} \sum_{t=1}^T \left| \sum_{s=1}^T \xi_s h_{t-s} \right|^4 \right)^{1/2}$$

because for any sequence ϕ_t , $(\max_t |\phi_t|)^2 = \max_t \phi_t^2 \leq \sum_{t=1}^T \phi_t^2$. But

$$\begin{aligned} \frac{1}{T^2} \sum_{t=1}^T E \left| \sum_{s=1}^T \xi_s h_{t-s} \right|^4 &= \frac{1}{T^2} \sum_{t=1}^T \sum_{s,r,q,v=1}^T E(\xi_s \xi_r \xi_q \xi_v h_{t-s} h_{t-r} h_{t-q} h_{t-v}) \\ &= \frac{3\sigma_\xi^4}{T^2} \sum_{t=1}^T \left(\sum_{s=1}^T h_{t-s}^2 \right)^2 + \frac{\text{cum}(\xi_s, \xi_s, \xi_s, \xi_s)}{T^2} \sum_{t=1}^T \left(\sum_{s=1}^T h_{t-s}^4 \right) \\ &= O(T^{-1}) \end{aligned}$$

proceeding as with the proof of part (a), (cf. (5.6)). So,

$$E \left(\max_t d_t^2 \right)^2 = O(T^{-1}).$$

Then by Markov inequality, the second term of (5.7) is $o(1)$, which completes the proof of part (b) and the proposition. \square

Proposition 2. *Assuming C1-C5, as $T \rightarrow \infty$,*

$$\frac{1}{T^{1/2}} \sum_{j=1}^{T-1} I_{xu}(\lambda_j) \xrightarrow{d} \mathcal{N}\left(0, (2\pi)^{-2} \Omega\right).$$

Proof. Since by C4, $f_{xx}(\lambda) f_{uu}(\lambda)$ has the same properties of $g(\lambda)$ in Proposition 1, then by symmetry it suffices to show that

$$\frac{1}{T^{1/2}} \sum_{j=1}^{[T/2]} f_{xx,j}^{1/2} f_{uu,j}^{1/2} \left(\frac{w_{x,j} w_{u,-j}}{f_{xx,j}^{1/2} f_{uu,j}^{1/2}} - (2\pi) \frac{w_{\xi,j}}{\sigma_\xi} \frac{w_{\varepsilon,-j}}{\sigma_\varepsilon} \right) \xrightarrow{P} 0,$$

which is the case, as we now show and where henceforth for a generic function $h(\lambda)$, h_j denotes $h(\lambda_j)$. The left side of the last displayed expression is, after standard calculations,

$$\frac{1}{T^{1/2}} \sum_{j=1}^{[T/2]} f_{xx,j}^{1/2} f_{uu,j}^{1/2} \left(\frac{w_{x,j}}{f_{xx,j}^{1/2}} - (2\pi)^{1/2} \frac{w_{\xi,j}}{\sigma_\xi} \right) \left(\frac{w_{u,-j}}{f_{uu,j}^{1/2}} - (2\pi)^{1/2} \frac{w_{\varepsilon,-j}}{\sigma_\varepsilon} \right)$$

$$\begin{aligned}
& + \frac{(2\pi)^{1/2}}{T^{1/2}} \sum_{j=1}^{\lfloor T/2 \rfloor} f_{xx,j}^{1/2} f_{uu,j}^{1/2} \left(\frac{w_{x,j}}{f_{xx,j}^{1/2}} - (2\pi)^{1/2} \frac{w_{\xi,j}}{\sigma_{\xi}} \right) \frac{w_{\varepsilon,-j}}{\sigma_{\varepsilon}} \\
& + \frac{(2\pi)^{1/2}}{T^{1/2}} \sum_{j=1}^{\lfloor T/2 \rfloor} f_{xx,j}^{1/2} f_{uu,j}^{1/2} \frac{w_{\xi,j}}{\sigma_{\xi}} \left(\frac{w_{u,-j}}{f_{uu,j}^{1/2}} - (2\pi)^{1/2} \frac{w_{\varepsilon,-j}}{\sigma_{\varepsilon}} \right).
\end{aligned} \tag{5.8}$$

First, observe that $C4$ and integrability of $f_{uu}(\lambda) f_{xx}(\lambda)$ imply that

$$f_{uu}^{1/2}(\lambda) f_{xx}^{1/2}(\lambda) \leq D \lambda^{-1/2} |\log |\lambda||^{-(1+\delta)/2} \tag{5.9}$$

for any arbitrarily small $\delta > 0$, and by Robinson's (1995a) Theorem 2 and its obvious extension to $[0, \pi]$, see Lemma 4.4 of Giraitis et al. (2001),

$$E \left| \frac{w_{x,j}}{f_{xx,j}^{1/2}} - (2\pi)^{1/2} \frac{w_{\xi,j}}{\sigma_{\xi}} \right|^2 \leq \frac{D \log j}{j}.$$

Then, by the Schwarz inequality, we obtain that the expectation of the first term of (5.8) is bounded by

$$D \sum_{j=1}^{\lfloor T/2 \rfloor} \frac{\log j}{j^{3/2} (\log T - \log j + \log 2\pi)^{(1+\delta)/2}} = O\left(\log^{-(1+\delta)/2} T\right).$$

Next we examine the second term of (5.8). Because by $C2$, $E(w_{\varepsilon,j} w_{\varepsilon,-k}) = \sigma_{\varepsilon}^2 \mathcal{I}(j = k)$, we have that the second moment of that term is bounded by

$$\begin{aligned}
& D \frac{1}{T} \sum_{j=1}^{\lfloor T/2 \rfloor} f_{xx,j} f_{uu,j} \frac{\log j}{j} \leq D \sum_{j=1}^{\lfloor T/2 \rfloor} \frac{\log j}{j^2 (\log T - \log j + \log (2\pi))^{1+\delta}} \\
& = O\left(\log^{-(1+\delta)} T\right),
\end{aligned}$$

by (5.9), so it is the third term of (5.8) proceeding as with the second term but using $C3$ instead of $C2$. This concludes the proof of the proposition. \square

Now we turn to the proof of Theorem 1, which follows by Proposition 2 and Cr amer's Theorem because

$$\frac{1}{T} \sum_{j=1}^{T-1} I_{xx,j} = \frac{1}{2\pi T} \sum_{t=1}^{T-1} (x_t - E x_t) (x_t - E x_t)' \xrightarrow{p} (2\pi)^{-1} \Sigma > 0 \tag{5.10}$$

by $C1$ and $C3$. \square

5.2. Proof of Theorem 2

The proof of this theorem is split into two propositions. Proposition 3 shows that the bootstrap second moment converges in probability to Φ , whereas Proposition 4 proves the Lindeberg's condition. These two propositions will imply the statement of the theorem. Henceforth, we shall denote $E^*(\cdot)$ as the bootstrap expectation, that is, for any random variable Y , $E^*(Y) = E\left(Y \left| \left\{ \tilde{x}, \tilde{u} \right\} \right.\right)$, where $\tilde{x} = (x_1, \dots, x_T)'$ and $\tilde{u} = (u_1, \dots, u_T)'$.

Proposition 3. Assuming $C1 - C4$ and $C5'$, as $T \rightarrow \infty$, for $i = 1, 2$

$$E^* \left(T^{1/2} \left(\widehat{\beta}_i^* - \widehat{\beta} \right) \right)^2 \xrightarrow{P} \Phi. \quad (5.11)$$

Proof. By the definition of $\widehat{\beta}_i^*$, we have that

$$T^{1/2} \left(\widehat{\beta}_i^* - \widehat{\beta} \right) = \left(\frac{1}{T} \sum_{j=1}^{T-1} I_{xx,j} \right)^{-1} \frac{1}{T^{1/2}} \sum_{j=1}^{T-1} w_{x,j} |w_{\widehat{u},j}| \eta_{j,i}^*.$$

Since for $i = 1$, $\eta_{j,1}^*$ is a sequence of uncorrelated distributed $(0, 1)$ random variables, it follows by standard manipulations that the left side of (5.11) is

$$\left(\frac{1}{T} \sum_{j=1}^{T-1} I_{xx,j} \right)^{-1} \left(\frac{1}{T} \sum_{j=1}^{T-1} I_{xx,j} I_{\widehat{u},j} \right) \left(\frac{1}{T} \sum_{j=1}^{T-1} I_{xx,j} \right)^{-1}. \quad (5.12)$$

On the other hand, for $i = 2$, because conditional on the sample, \widehat{u}_t^* is a sequence of independent identically distributed $(0, 1)$ random variables, it implies that

$$\begin{aligned} E^* \left(\eta_{j,2}^* \eta_{-k,2}^* \right) &= \frac{1}{T} E^* \left(\sum_{t=1}^T \widehat{u}_t^* e^{-it\lambda_j} \sum_{s=1}^T \widehat{u}_s^* e^{is\lambda_k} \right) \\ &= \frac{1}{T} \sum_{s=1}^T e^{it(\lambda_k - \lambda_j)} \\ &= \mathcal{I}(j = k). \end{aligned}$$

So for $i = 2$, the left side of (5.11) is also (5.12).

To conclude the proof we need to show that (5.12) converges in probability to Φ . First, by Slutsky's Theorem and (5.10), the first and third factors of (5.12) converge in probability to $(2\pi) \Sigma^{-1}$. So, the proof is completed if the second factor of (5.12) converges in probability to $(2\pi)^{-1} \int_{-\pi}^{\pi} f_{xx}(\lambda) f_{uu}(\lambda) d\lambda$. But that factor is

$$\frac{1}{T} \sum_{j=1}^{T-1} I_{xx,j} I_{uu,j} + \frac{1}{T} \sum_{j=1}^{T-1} I_{xx,j} (I_{\widehat{u},j} - I_{uu,j}), \quad (5.13)$$

whose first term, under $C1 - C4$ and $C5'$, converges in \mathcal{L}_1 -norm, and thus in probability, to $(2\pi)^{-1} \int_{-\pi}^{\pi} f_{xx}(\lambda) f_{uu}(\lambda) d\lambda$, see Robinson (1998). So, it remains to show that the second term of (5.13) converges to zero in probability. Using that

$$\widehat{u}_t = \widetilde{u}_t - \left(\widehat{\beta} - \beta_0 \right)' \widetilde{x}_t,$$

the second term of (5.13) is

$$\frac{1}{T} \sum_{j=1}^{T-1} I_{xx,j} \left(\left(\widehat{\beta} - \beta_0 \right)' I_{xx,j} \left(\widehat{\beta} - \beta_0 \right) - 2 \left(\widehat{\beta} - \beta_0 \right)' \operatorname{Re} \left(I_{xu,j} \right) \right) = o_p(1) \quad (5.14)$$

as we now show, where $\operatorname{Re}(z)$ denotes the real part of the complex number z .

Now,

$$\begin{aligned}
tr \left(\frac{1}{T} \sum_{j=1}^{T-1} I_{xx,j}^2 \right) &= \sum_{r=1}^p \left\{ \frac{1}{T} \sum_{j=1}^{T-1} f_{x_r, x_r, j}^2 \left(\frac{I_{x_r, x_r, j}}{f_{x_r, x_r, j}} - (2\pi) \frac{I_{\xi_r, \xi_r, j}}{\sigma_{\xi_r}^2} \right)^2 \right. \\
&\quad + \frac{4\pi}{\sigma_{\xi_r}^2 T} \sum_{j=1}^{T-1} f_{x_r, x_r, j}^2 \left(\frac{I_{x_r, x_r, j}}{f_{x_r, x_r, j}} - (2\pi) \frac{I_{\xi_r, \xi_r, j}}{\sigma_{\xi_r}^2} \right) I_{\xi_r, \xi_r, j} \\
&\quad \left. + \frac{4\pi^2}{\sigma_{\xi_r}^4 T} \sum_{j=1}^{T-1} f_{x_r, x_r, j}^2 I_{\xi_r, \xi_r, j}^2 \right\},
\end{aligned}$$

where $I_{\xi_r, \xi_r, j} = w_{\xi_r, j} w'_{\xi_r, -j}$, and z_r denotes the r th element of the vector z . But proceeding as in the proof of (4.8) in Robinson (1995b, pp. 1648-1651), the expectation of the right side of the last displayed equation is dominated by the expectation of the third term. But $E \left(I_{\xi_r, \xi_r, j}^2 \right) \leq D$ by Brockwell and Davis's (1991) Proposition 10.3.2. So, the expectation of the left side of the last displayed equation is bounded by

$$D \sum_{r=1}^p \frac{1}{T} \sum_{j=1}^{T-1} f_{x_r, x_r, j}^2 \leq D \frac{1}{T} \sum_{j=1}^{T-1} \lambda_j^{-2} |\log |\lambda_j||^{-2-2\delta} = o(T) \quad (5.15)$$

since by C4 and integrability of $f_{xx}(\lambda)$, for all $r = 1, \dots, p$, $f_{x_r, x_r, j} \leq D \lambda_j^{-1} |\log |\lambda||^{-1-\delta}$, for some $\delta > 0$. The right side of (5.15) together with Theorem 1, that is that $(\hat{\beta} - \beta_0) = O_p(T^{-1/2})$, implies that the first term of (5.14) is $o_p(1)$. Proceeding similarly with the second term of (5.14) and observing that $\text{Re}(I_{xu, j}) = 2^{-1}(I_{xu, j} + I_{xu, -j})$, we conclude that the second term of (5.14) is also $o_p(1)$, which completes the proof of the proposition. \square

Proposition 4. *Assuming C1-C4 and C5', for all $\psi > 0$ as $T \rightarrow \infty$, , for $i = 1, 2$*

$$E^* \left(\frac{1}{T} \sum_{j=1}^{T-1} \left\| (w_{x, j} | w_{\hat{u}, j} | \eta_j^*) \right\|^2 \mathcal{I} \left(T^{-1} \left\| (w_{x, j} | w_{\hat{u}, j} | \eta_{j, i}^*) \right\|^2 > \psi \right) \right) \xrightarrow{P} 0. \quad (5.16)$$

Proof. We first examine $\sup_j T^{-1} \|I_{xx, j} I_{\hat{u}\hat{u}, j}\|$ which, after standard inequalities, is bounded by

$$\begin{aligned}
&\sum_{r=1}^p \sup_{j=1, \dots, [T/2]} T^{-1} f_{x_r, x_r, j} f_{uu, j} \left| f_{x_r, x_r, j}^{-1/2} w_{x_r, j} f_{uu, j}^{-1/2} w_{u, -j} \right|^2 \\
&\quad + \sum_{r=1}^p \sup_{j=1, \dots, [T/2]} T^{-1} f_{x_r, x_r, j} \left| f_{x_r, x_r, j}^{-1/2} w_{x_r, j} \right|^2 |I_{\hat{u}\hat{u}, j} - I_{uu, j}|.
\end{aligned} \quad (5.17)$$

Next, by Robinson's (1995a) Theorem 2 and an obvious extension to all $1 \leq j \leq [T/2]$, see Lemma 4.4 of Giraitis et al. (2001),

$$\begin{aligned}
E \left| f_{uu, j}^{-1/2} w_{u, j} - (2\pi)^{1/2} \frac{w_{\varepsilon, j}}{\sigma_\varepsilon} \right|^2 &= O \left(\frac{\log j}{j} \right) \text{ and} \\
E \left| f_{x_r, x_r, j}^{-1/2} w_{x_r, j} - (2\pi)^{1/2} \frac{w_{\xi_r, j}}{\sigma_{\xi_r}} \right|^2 &= O \left(\frac{\log j}{j} \right) \text{ for } r = 1, \dots, p.
\end{aligned}$$

So, using the inequality $\sup_t |\phi_t|^2 \leq \sum_t |\phi_t|^2$, we have for instance that, for all $r = 1, \dots, p$,

$$\begin{aligned} & E \left(\sup_{j=1, \dots, [T^{1/2}]} \frac{1}{j} \left| f_{x_r, x_r, j}^{-1/2} w_{x_r, j} - (2\pi)^{1/2} \frac{w_{\xi_r, j}}{\sigma_{\xi_r}} \right|^2 \right) \\ & \leq \sum_{j=1}^{[T^{1/2}]} \frac{1}{j} E \left| f_{x_r, x_r, j}^{-1/2} w_{x_r, j} - (2\pi)^{1/2} \frac{w_{\xi_r, j}}{\sigma_{\xi_r}} \right|^2 = O(1), \end{aligned} \quad (5.18)$$

and

$$\begin{aligned} & E \left(\sup_{j=[T^{1/2}]+1, \dots, [T/2]} \frac{1}{j} \left| f_{x_r, x_r, j}^{-1/2} w_{x_r, j} - (2\pi)^{1/2} \frac{w_{\xi_r, j}}{\sigma_{\xi_r}} \right|^2 \right) \\ & \leq \sum_{j=[T^{1/2}]+1}^{[T/2]} \frac{1}{j} E \left| f_{x_r, x_r, j}^{-1/2} w_{x_r, j} - (2\pi)^{1/2} \frac{w_{\xi_r, j}}{\sigma_{\xi_r}} \right|^2 = O\left(T^{-1/2} \log T\right). \end{aligned} \quad (5.19)$$

On the other hand, by An et al. (1983),

$$\sup_{j=1, \dots, [T/2]} \left((2\pi) \sigma_\varepsilon^{-2} \log^{-1} T |w_{\varepsilon, j}|^2 \right) \leq 1 \quad \text{a.s.} \quad (5.20)$$

Then, combining (5.18), (5.19) and (5.20), the first term of (5.17) is bounded by

$$\begin{aligned} & D \sum_{r=1}^p \sup_{j=1, \dots, [T^{1/2}]} T^{-1} j f_{x_r, x_r, j} f_{uu, j} \sup_{j=1, \dots, [T^{1/2}]} \frac{1}{j} \left| f_{x_r, x_r, j}^{-1/2} w_{x_r, j} f_{uu, j}^{-1/2} w_{u, -j} \right|^2 \\ & + D \sum_{r=1}^p \sup_{j=1+[T^{1/2}], \dots, [T/2]} T^{-1} j f_{x_r, x_r, j} f_{uu, j} \sup_{j=1+[T^{1/2}], \dots, [T/2]} \frac{1}{j} \left| f_{x_r, x_r, j}^{-1/2} w_{x_r, j} f_{uu, j}^{-1/2} w_{u, -j} \right|^2 \\ & \leq D \sum_{r=1}^p \sup_{j=1, \dots, [T^{1/2}]} T^{-1} j f_{x_r, x_r, j} f_{uu, j} \log^2 T \\ & + D T^{-1/2} \log T \sum_{r=1}^p \sup_{j=1+[T^{1/2}], \dots, [T/2]} T^{-1} j f_{x_r, x_r, j} f_{uu, j} \log^2 T \\ & \leq D \log^{-\delta} T \end{aligned}$$

in probability since integrability of $\|f_{xx}(\lambda) f_{uu}(\lambda)\| |\log(|\lambda|)|$ implies that for some $\delta > 0$, $\|f_{xx, j} f_{uu, j}\| |\log |\lambda_j| \leq D \lambda_j^{-1} |\log |\lambda_j|^{-1-\delta}$, and hence $\|f_{xx, j} f_{uu, j}\| \leq D \lambda_j^{-1} |\log |\lambda_j|^{-2-\delta}$. On the other hand, because

$$I_{\widehat{u}, j} - I_{u, j} = (\beta_0 - \widehat{\beta})' I_{xx, j} (\beta_0 - \widehat{\beta}) + 2 (\beta_0 - \widehat{\beta})' \text{Re}(I_{xu, j}),$$

the second term of (5.17) is bounded by

$$\begin{aligned} & \sup_{j=1, \dots, [T/2]} T^{-1} f_{xx, j} \left| f_{xx, j}^{-1/2} w_{x, j} \right|^2 \left(\left\| \beta_0 - \widehat{\beta} \right\|^2 \left| f_{xx, j} \left| f_{xx, j}^{-1/2} w_{x, j} \right|^2 \right. \right. \\ & \quad \left. \left. + \left\| \beta_0 - \widehat{\beta} \right\| \left| f_{xx, j}^{1/2} f_{uu, j}^{1/2} \left| f_{xx, j}^{-1/2} w_{x, j} f_{uu, j}^{-1/2} w_{u, j} \right| \right) \right) \end{aligned}$$

which, by (5.18) and (5.20), is upper bounded by

$$\sup_{j=1, \dots, \lfloor T/2 \rfloor} T^{-1} f_{xx,j} \left(\left\| \beta_0 - \widehat{\beta} \right\|^2 f_{xx,j} + \left\| \beta_0 - \widehat{\beta} \right\| f_{xx,j}^{1/2} f_{uu,j}^{1/2} \right) \log^2 T \leq D \log^{-\delta} T$$

because $\left\| \beta_0 - \widehat{\beta} \right\| = O_p(T^{-1/2})$, $\|f_{xx,j}\| \leq D\lambda_j^{-1} |\log |\lambda_j||^{-1-\delta}$ and $\left\| f_{xx,j}^{1/2} f_{uu,j}^{1/2} \right\| \leq D\lambda_j^{-1/2} |\log |\lambda_j||^{-1-\delta/2}$ by integrability of $f_{xx}(\lambda)$ and $\|f_{xx}(\lambda) f_{uu}(\lambda)\| |\log(|\lambda|)|$ respectively. Thus, for $i = 1, 2$ we conclude that the left side of (5.16) is bounded by

$$\frac{D}{T} \sum_{j=1}^{T-1} I_{xx,j} I_{uu,j} E^* \left(|\eta_{j,i}^*|^2 \mathcal{I} \left(|\eta_{j,i}^*|^2 > \psi \log^\delta T \right) \right). \quad (5.21)$$

However, conditional on the data, $\eta_{j,1}^*$ is a sequence of *iid* random variables, so for $i = 1$ (5.21) is

$$E^* \left(|\eta_{j,1}^*|^2 \mathcal{I} \left(|\eta_{j,1}^*|^2 > \psi \log^\delta T \right) \right) \frac{D}{T} \sum_{j=1}^{T-1} I_{xx,j} I_{uu,j},$$

which converge to zero in probability since $\delta > 0$ and $\eta_{j,1}^*$ has finite second moments and by Proposition 3 the second factor of the last displayed expression converges in probability to $(2\pi)^{-1} \Omega < \infty$.

On the other hand, for $i = 2$, (5.21) is bounded by

$$\sup_j E^* \left(|\eta_{j,2}^*|^2 \mathcal{I} \left(|\eta_{j,2}^*|^2 > \psi \log^\delta T \right) \right) \left(\frac{D}{T} \sum_{j=1}^{T-1} I_{xx,j} I_{uu,j} \right),$$

which converges to zero in probability for all $\delta > 0$ as we now show. First, by Robinson (1988), the second factor converges in probability to $(2\pi)^{-1} \Omega < \infty$. On the other hand, standard algebra implies that

$$\begin{aligned} E^* \left(|\eta_{j,2}^*|^2 \mathcal{I} \left(|\eta_{j,2}^*|^2 > \psi \log^\delta T \right) \right) &\leq \frac{D}{\psi^2 \log^{2\delta} T} E^* |\eta_{j,2}^*|^4 \\ &\leq \frac{D}{\psi^2 \log^{2\delta} T} \left\{ \frac{1}{T} \sum_{t=1}^T u_t^4 + \left(\frac{1}{T} \sum_{t=1}^T u_t^2 \right)^2 \right\}. \end{aligned}$$

But, by a well-know argument, see Stout's (1974) Theorem 3.5.8, u_t is also ergodic which implies that the right side of the last displayed inequality converges to zero because $\delta > 0$. So, we conclude that (5.21) $\xrightarrow{P} 0$ and the proof of the proposition. \square

5.3. Proof of Theorem 3

We only sketch the proof since it follows immediately from Theorem 2 and the continuous mapping theorem. First, since $T^{1/2} (C\widehat{\beta}_i^* - c) = T^{1/2} C(\widehat{\beta}_i^* - \widetilde{\beta})$, Theorem 2 implies that

$$T^{1/2} C(\widehat{\beta}_i^* - \widetilde{\beta}) \xrightarrow{d^*} \mathcal{N}(0, C\Phi C')$$

for $i = 1, 2$. On the other hand, proceeding as in the proof of (5.10), $\widehat{\Sigma} \xrightarrow{P} (2\pi)^{-1} \Sigma > 0$ and if $\widetilde{\Omega}_i^* \xrightarrow{P^*} \Omega$ for $i = 1, 2$, we obtain that

$$C\widehat{\Sigma}^{-1}\widetilde{\Omega}_i^*\widehat{\Sigma}^{-1}C' \xrightarrow{P^*} C\Phi C'.$$

Then, the continuous mapping theorem implies that $F_i^* \xrightarrow{d^*} \chi_r^2$, for $i = 1, 2$. So, to conclude the proof of the theorem, we need to show that $\widetilde{\Omega}_i^* \xrightarrow{P^*} \Omega$. By definition, $\widetilde{\Omega}_i^* = (4\pi^2) T^{-1} \sum_{j=1}^{T-1} I_{xx,j} I_{\widehat{u}^* \widehat{u}^*,j}$. Next, because

$$\begin{aligned} w_{\widehat{u}^*}(\lambda_j) &= w_{y^*}(\lambda_j) - \widehat{\beta}_i^* w_x(\lambda_j) \\ &= (\widetilde{\beta} - \widehat{\beta}_i^*) w_x(\lambda_j) + |w_{\widetilde{u}}(\lambda_j)| \eta_{j,i}^*, \end{aligned}$$

we obtain that

$$\begin{aligned} \widetilde{\Omega}_i^* - \widetilde{\Omega} &= \frac{4\pi^2}{T} \sum_{j=1}^{T-1} I_{xx,j} \left((\widehat{\beta}_i^* - \widetilde{\beta})' I_{xx,j} (\widehat{\beta}_i^* - \widetilde{\beta}) - 2 (\widehat{\beta}_i^* - \widetilde{\beta})' \operatorname{Re} \left(|w_{\widetilde{u}}(\lambda_j)| \eta_{j,i}^* \overline{w_x(\lambda_j)} \right) \right) \\ &\quad + \frac{4\pi^2}{T} \sum_{j=1}^{T-1} I_{xx,j} I_{\widetilde{u}\widetilde{u},j} \left(|\eta_{j,i}^*|^2 - 1 \right). \end{aligned}$$

Now proceeding as in the proof of (5.14), after we observe that by Theorem 2, $(\widehat{\beta}_i^* - \widetilde{\beta}) = O_{p^*}(T^{-1/2})$, in particular that in Proposition 3 we have shown that $E^* \left(T^{1/2} (\widehat{\beta}_i^* - \widetilde{\beta}) \right)^2 = O_p(1)$, the first term on the right of the last displayed equation is $o_{p^*}(1)$. The second term on the right is also $o_{p^*}(1)$, by similar arguments and that $\eta_{j,i}^*$, conditional on the data, is a zero mean sequence of uncorrelated random variables. Finally, the third term on the right of the last displayed equation. Because $\widehat{u}_t = \widetilde{u}_t - (\widehat{\beta} - \beta_0) \widetilde{x}_t$, this term can be written as

$$\begin{aligned} &(\widehat{\beta} - \beta_0)' \frac{4\pi^2}{T} \sum_{j=1}^{T-1} I_{xx,j}^2 \left(|\eta_{j,i}^*|^2 - 1 \right) (\widehat{\beta} - \beta_0) \\ &- 2 (\widehat{\beta} - \beta_0)' \frac{4\pi^2}{T} \sum_{j=1}^{T-1} I_{xx,j} \operatorname{Re} (I_{xu,j}) \left(|\eta_{j,i}^*|^2 - 1 \right) + \frac{4\pi^2}{T} \sum_{j=1}^{T-1} I_{xx,j} I_{uu,j} \left(|\eta_{j,i}^*|^2 - 1 \right). \end{aligned}$$

Clearly the first two terms are $o_{p^*}(1)$, whereas the last term of the last displayed expression is also $o_{p^*}(1)$ using the same arguments as those employed to examine the behaviour of (5.21), since conditional on the data, $|\eta_{j,i}^*|^2 - 1$ is a sequence of zero mean uncorrelated random variables. Now, conclude the proof since we have already shown that $\widetilde{\Omega} \xrightarrow{P} \Omega$, e.g. (5.12). \square

6. A TECHNICAL LEMMA

Lemma 1. *Let $\varphi(\lambda)$ be an integrable function which satisfies (a) $\varphi(\lambda) > 0$ for all $\lambda \in [0, \pi]$, (b) $|(\partial/\partial\lambda)\varphi(\lambda)| = O(|\varphi(\lambda)|/\lambda)$ as $\lambda \rightarrow 0+$ and (c) it is continuously*

differentiable in any open set outside the origin. Then, as $T \rightarrow \infty$,

$$\frac{1}{T} \sum_{j=1}^{\lfloor T/2 \rfloor} \varphi(\lambda_j) - (2\pi)^{-1} \int_0^\pi \varphi(\lambda) d\lambda = o(1).$$

Proof. The left side of the last displayed expression is

$$\frac{1}{T} \varphi(\lambda_\pi) - \frac{1}{2\pi} \int_0^{\lambda_1} \varphi(\lambda) d\lambda + \frac{1}{2\pi} \sum_{j=1}^{\lfloor T/2 \rfloor - 1} \int_{\lambda_j < \lambda \leq \lambda_{j+1}} (\varphi(\lambda_j) - \varphi(\lambda)) d\lambda. \quad (6.1)$$

The first term of (6.1) is obviously $o(1)$, whereas the second term of (6.1) is also $o(1)$ because by integrability of $\varphi(\lambda)$, for any arbitrarily small $\delta > 0$,

$$\begin{aligned} \int_0^{\lambda_1} \varphi(\lambda) d\lambda &\leq D \int_0^{\lambda_1} \lambda^{-1} |\log \lambda|^{-1-\delta} d\lambda \\ &\leq D |\log \lambda_1|^{-\delta} = o(1). \end{aligned}$$

Next, the absolute value of the third term of (6.1) is bounded by

$$D \sum_{j=1}^{\lfloor T/2 \rfloor - 1} \int_{\lambda_j < \lambda \leq \lambda_{j+1}} \lambda_j^{-1} |\lambda_j - \lambda| \varphi(\bar{\lambda}) d\lambda$$

by the properties of $\varphi(\lambda)$ and the mean value theorem where $\bar{\lambda}$ is an intermediate point in the interval (λ_j, λ) . But the last displayed expression is bounded by

$$D \sum_{j=1}^{\lfloor T/2 \rfloor - 1} \frac{1}{j} \varphi(\lambda_j) \int_{\lambda_j < \lambda \leq \lambda_{j+1}} d\lambda \quad (6.2)$$

since $\sup_{\lambda_j < \lambda \leq \lambda_{j+1}} |\lambda_j - \lambda| \leq DT^{-1}$ and by (b) and (c), $\sup_{\lambda_j < \lambda \leq \lambda_{j+1}} \varphi(\bar{\lambda}) \leq D\varphi(\lambda_j)$. On the other hand, integrability of $\varphi(\lambda)$ implies that

$$\varphi(\lambda) \leq D\lambda^{-1} |\log(\lambda)|^{-1-\delta}$$

for any arbitrarily small $\delta > 0$. So, we conclude that (6.2) is bounded by

$$D \sum_{j=1}^{\lfloor T/2 \rfloor - 1} \frac{1}{j^2 (\log T - \log j + \log(2\pi))^{1+\delta}} = o(1)$$

by standard arguments, which completes the proof. \square

7. CONCLUSIONS

In a time series regression model, we have examined and discussed two similar bootstraps, which contrary to the moving blocks and subsampling bootstraps, do not require the choice of a tuning parameter, that is the block length b . The bootstraps are based on resampling from the discrete Fourier transform of the least squares residuals as in the original Efron's bootstrap or from the residuals themselves. So, we were

able to avoid the potential problem that the choice of the block length may have on finite samples. Although we have examined the algorithms in the context of linear regression models, the approach can be extended to more general models of interest in econometrics. More specifically in the context of instrumental variable/*GMM* estimators, which includes most of the available estimators, such as the *OLS*, *GLS* or *MLE*.

To that end, let $(y_t, x_t)'$ be a sequence of random variables observed at times $t = 1, \dots, T$, such that they follow the nonlinear model $u_t = u_t(y_t, x_t; \theta_0)$, where θ_0 is $p \times 1$ parameter vector. Suppose, for illustration purposes, that $u_t(\cdot)$ can be written as

$$h(y_t, \gamma) = f(x_t, \beta) + u_t. \quad (7.1)$$

A classical example of (7.1), and widely used, is the Box-Cox transformation, where

$$h(y_t, \gamma) = \frac{y_t^\gamma - 1}{\gamma},$$

and $\theta = (\gamma, \beta)'$

In this context, under the same conditions specified in Section 2, we propose the following bootstrap procedure.

STEP 1 Let $P(z_t)$ be a set of valid instruments, and obtain the *IVE* or *GMM* estimator of $\theta = (\gamma, \beta)'$ defined as

$$\begin{aligned} (\widehat{\gamma}, \widehat{\beta})' &= \widehat{\theta} = \arg \min_{\theta} \left(\sum_{t=1}^T u_t(\theta) P'(z_t) \right) \left(\sum_{t=1}^T P(z_t) P'(z_t) \right)^{-1} \left(\sum_{t=1}^T u_t(\theta) P(z_t) \right) \\ &= \arg \min_{\theta} \left(\sum_{j=1}^{T-1} w_{u(\theta)}(\lambda_j) \overline{w_{P(z)}(\lambda_j)} \right) \left(\sum_{j=1}^{T-1} w_{P(z)}(\lambda_j) \overline{w_{P(z)}(\lambda_j)} \right)^{-1} \\ &\quad \times \left(\sum_{j=1}^{T-1} w_{u(\theta)}(\lambda_j) w_{P(z)}(\lambda_j) \right) \end{aligned}$$

where $w_{u(\theta)}(\lambda_j)$ and $w_{P(z)}(\lambda_j)$ are the discrete Fourier transforms of $u_t(\theta) = h(y_t, \gamma) - f(x_t, \beta)$ and $P(z_t)$ respectively, and \bar{a} stands for transpose together with conjugation of a complex vector a . Next, obtain the residuals

$$\widehat{u}_t = h(y_t, \widehat{\gamma}) - f(x_t, \widehat{\beta}), \quad t = 1, \dots, T.$$

STEP 2 The same as *STEP 2* given in the introduction.

STEP 3 For $i = 1, 2$, using the discrete Riemann approximation of the identity

$$z_t = \left(\frac{T}{2\pi} \right)^{1/2} \int_{-\pi}^{\pi} w_z(\mu) e^{i\mu t} d\mu, \quad t = 1, \dots, T,$$

compute the bootstrap residuals

$$\widehat{u}_{t,i}^* = \left(\frac{2\pi}{T} \right)^{1/2} \sum_{j=1}^{T-1} e^{it\lambda_j} w_{\widehat{u}}(\lambda_j) \eta_{j,i}^*, \quad t = 1, \dots, T$$

and the bootstrap observations $y_{t,i}^*$ as

$$y_{t,i}^* = g_{\hat{\gamma}} \left(f \left(x_t, \hat{\beta} \right) + \hat{u}_{t,i}^* \right), \quad t = 1, \dots, T$$

where $g_a(z)$ satisfies that $h(g_a(z), a) = z$, e.g. $g_a(z)$ is the inverse function of $h(z, a)$.

STEP 4 For $i = 1, 2$, compute the bootstrap estimator as

$$\begin{aligned} \left(\hat{\gamma}_i^*, \hat{\beta}_i^{*'} \right)' &= \hat{\theta}^* = \arg \min_{\theta} \left(\sum_{j=1}^{T-1} w_{\hat{u}_i^*(\theta)}(\lambda_j) \overline{w_{P(z)}(\lambda_j)} \right) \left(\sum_{j=1}^{T-1} w_{P(z)}(\lambda_j) \overline{w_{P(z)}(\lambda_j)} \right)^{-1} \\ &\quad \times \left(\sum_{j=1}^{T-1} w_{\hat{u}_i^*(\theta)}(\lambda_j) w_{P(z)}(\lambda_j) \right), \end{aligned}$$

where $w_{\hat{u}_i^*(\theta)}(\lambda_j)$ denotes the discrete Fourier transform of $\hat{u}_{t,i}^* = h(y_{t,i}^*, \gamma) - f(x_t, \beta)$.

We conjecture that the preceding bootstrap should be valid. The intuition comes from the observation that because $w_{\hat{u}_i^*(\hat{\theta})} = w_{\hat{u}}(\lambda_j) \eta_{j,i}^*$, it implies that the covariance structure of the data/model is preserved, so the asymptotic behaviour of the bootstrap estimates should be that of $\hat{\theta}$. The technical details of the bootstrap will be more involved than those in the context examined in the paper. However, in view of the available results for extreme estimators, it is expected that the results given in Section 2 should hold true in the previous context/model. This topic and how we can implement the bootstrap in the situation of heteroscedastic models will be examined in a future paper.

References

- [1] An, H.-Z., Chen, Z.-G. and Hannan, E.J., 1983. The Maximum of the Periodogram, *Journal of Multivariate Analysis* 13, 383-400.
- [2] Andrews, D.W.K., 1991. Heteroscedasticity and Autocorrelation Consistent Covariance Matrix Estimation, *Econometrica* 59, 817-858.
- [3] Brillinger, D.R., 1979. Confidence Intervals for the Crosscovariance Function, *Selecta Statistica Canadiana* 5, 3-16.
- [4] Brillinger, D.R., 1981. Time Series, Data Analysis and Theory (*Holden-Day*, San Francisco).
- [5] Brockwell, P.J. and Davis, R.A., 1991. Time Series: Theory and Methods, (*Springer-Verlag*, New York).
- [6] Carlstein, E., 1986. The use of Subseries Values for Estimating the Variance of a General Statistic from a Stationary Sequence, *Annals of Statistics* 14, 1171-1179.
- [7] Dahlhaus, R. and Janas, D., 1996. A Frequency Domain Bootstrap for Ratio Statistics in Time Series Analysis, *Annals of Statistics* 24, 1934-1963.

- [8] Davies, R.A. and Harte, D.S., 1987. Tests for Hurst Effect, *Biometrika* 74, 95-102.
- [9] Den Haan, W. and Levin A., 1997. A Practitioners Guide to Robust Covariance Matrix Estimation, in: Maddala, G.S. and Rao, C.R. eds., *Handbook of Statistics, Vol. 15* (North-Holland, Amsterdam) 299-342.
- [10] Efron, B., 1979. Bootstrap Methods: Another Look at the Jackknife, *Annals of Statistics* 7, 1-26.
- [11] Engle, R.F., Hendry, D.F. and Richard, J-F., 1983. Exogeneity, *Econometrica* 51, 277-304.
- [12] Fitzenberger, B., 1998. The Moving Blocks Bootstrap and Robust Inference for Linear Least Squares and Quantiles Regressions, *Journal of Econometrics* 82, 235-287.
- [13] Franke, J. and Härdle, W., 1992. On Bootstrapping Kernel Spectral Estimates, *Annals of Statistics* 20, 121-145.
- [14] Giné, E. and Zinn, J., 1989. Necessary Conditions for the Bootstrap of the Mean, *Annals of Statistics* 17, 684-691.
- [15] Giraitis, L., Hidalgo, J. and Robinson, P.M., 2001. Gaussian Estimation of Parametric Spectral Density with Unknown Pole, *Annals of Statistics* 29, 987-1023.
- [16] Granger, C.W.J. and Joyeux, R., 1980. An Introduction to Long Memory Time Series and Fractional Differencing, *Journal of Time Series Analysis* 1, 424-438.
- [17] Hall, P. and Horowitz, J.L., 1996. Bootstrap Critical Values for Tests Based on Generalized-Method-of-Moments Estimators, *Econometrica* 64, 891-916.
- [18] Hall, P., Horowitz, J.L. and Jing, B.-Y., 1995. On Blocking Rules for the Bootstrap with Dependent Data, *Biometrika* 82, 561-574.
- [19] Hannan, E.J., 1957. The Variance of the Mean of a Stationary Process, *Journal of the Royal Statistical Society Series B* 19, 282-285.
- [20] Hannan, E.J., 1963. Regression for Time Series, in: M. Rosenblatt ed., *Time Series Analysis*, (Wiley, New York) 17-37..
- [21] Hannan, E.J., 1970. Multiple Time Series, (*Wiley and Sons*, New York).
- [22] Hidalgo, J. and Kreiss, J.-P., 1999. Bootstrap Specification Tests for Covariance Stationary Processes, Preprint.
- [23] Hosking, J., 1981. Fractional Differencing, *Biometrika* 68, 165-176.
- [24] Jowett, G.H., 1955. The Comparison of Means of Sets of Observations from Sections of Independent Stochastic Series, *Journal of the Royal Statistical Society Series B* 17, 208-227.
- [25] Keenan, D.M., 1987. Limiting Behavior of Functionals of Higher-order Sample Cumulant Spectra, *Annals of Statistics* 15, 134-151.
- [26] Künsch, H.-R., 1989. The Jackknife and the Bootstrap for General Stationary Observations, *Annals of Statistics* 17, 1217-1241.

- [27] Loh, W.Y., 1987. Calibrating Confidence Coefficients, *Journal of the American Statistical Association* 82, 155-162.
- [28] Mammen, E., 1993. Bootstrap and Wild Bootstrap for High Dimensional Linear Models, *Annals of Statistics* 21, 255-285.
- [29] Mandelbrot, B.B. and Van Ness, J.W., 1968. Fractional Brownian Motions, Fractional Noises and Applications, *SIAM Review* 10, 422-437.
- [30] Parzen, E., 1957. On Consistent Estimates of the Spectrum of a Stationary Time Series, *Annals of Mathematical Statistics* 28, 329-348.
- [31] Politis, D.N., Romano, J.P., 1992. A General Resampling Scheme for Triangular Arrays of α -mixing Random Variables with Application to the Problem of Spectral Density Estimation, *Annals of Statistics* 20, 1985-2007.
- [32] Politis, D.N., Romano, J.P., 1994. Large sample confidence regions based on subsamples under minimal assumptions, *Annals of Statistics* 22, 2031-2050.
- [33] Politis, D.N., Romano, J.P. and Wolf, W., 1997. Subsampling for Heteroskedastic Time Series, *Journal of Econometrics* 81, 281-317.
- [34] Politis, D.N., Romano, J.P. and Wolf, M., 1999. Subsampling, (*Springer-Verlag*, Berlin).
- [35] Robinson, P.M., 1991. Automatic Frequency Domain Inference on Semiparametric and Nonparametric Models, *Econometrica* 59, 1329-1363.
- [36] Robinson, P.M., 1994. Time Series with Strong Dependence. In: C.A. Sims, ed., *Advances in Econometrics: Sixth World Congress*, Vol. 1, (Cambridge University Press, Cambridge) 47-95.
- [37] Robinson, P.M., 1995a. Log-Periodogram Regression for Time Series with Long Range Dependence, *Annals of Statistics* 23, 1048-1072.
- [38] Robinson, P.M., 1995b. Gaussian Semiparametric Estimation of Long-Range Dependence, *Annals of Statistics* 23, 1630-1661.
- [39] Robinson, P.M., 1998. Inference-Without-Smoothing in the Presence of Nonparametric Autocorrelation, *Econometrica* 66, 1163-1182.
- [40] Robinson, P.M. and Hidalgo, J., 1997. Time Series Regression with Long Range Dependence, *Annals of Statistics* 25, 77-104.
- [41] Robinson, P.M. and Velasco, C., 1997. Autocorrelation-Robust Inference, in: Maddala, G.S. and Rao, C.R. eds., *Handbook of Statistics, Vol. 15* (North-Holland, Amsterdam) 267-298.
- [42] Scott, D.J., 1973. Central Limit Theorems for Martingales and for Processes with Stationary Increments using a Skorohod representation Approach, *Advances in Applied Probability* 5, 119-137.
- [43] Singh, K., 1981. On the Asymptotic Accuracy of Efron's Bootstrap, *Annals of Statistics* 9, 1187-1195.

- [44] Stout, W.F., 1974. Almost Sure Convergence, (*Academic Press*, New York).
- [45] Taniguchi, M., 1982. On Estimation of the Integrals of the Fourth Order Cumulant Spectral Density, *Biometrika* 69, 117-122.
- [46] Wu, C.F.J., 1986. Jackknife, Bootstrap and other Resampling Methods in Regression Analysis, *Annals of Statistics* 14, 1261-1295.

TABLE 4.1

Proportion of rejections in 5000 Monte Carlo experiments under $H_0 : \beta_0 = 0$ in model (1.1), when x_t is long memory with parameter d_x and u_t follows an $AR(1)$ model. Bootstrap critical values are computed using 2000 bootstrap samples.

d_x	Size	T	$\rho = .0$			$\rho = .5$			$\rho = .9$		
			32	64	128	32	64	128	32	64	128
0.30	10%	<i>ASY</i>	14.32	12.50	12.00	16.06	14.24	13.02	19.20	16.00	14.34
		<i>BOOT1</i>	12.04	10.34	10.36	12.56	10.72	10.70	13.56	12.16	10.14
		<i>BOOT2</i>	11.82	10.78	10.62	12.34	11.16	10.96	12.36	12.10	10.38
	5%	<i>ASY</i>	8.40	6.86	6.42	9.92	7.74	7.20	12.00	9.04	7.48
		<i>BOOT1</i>	7.04	5.54	5.72	6.98	5.96	5.58	7.36	6.78	5.02
		<i>BOOT2</i>	6.88	5.60	5.92	6.96	6.26	5.94	6.66	6.94	5.70
	1%	<i>ASY</i>	2.12	1.28	1.02	2.84	1.48	1.20	2.92	1.48	0.82
		<i>BOOT1</i>	1.96	1.42	1.34	2.06	1.64	1.40	2.12	1.58	1.32
		<i>BOOT2</i>	1.96	1.42	1.48	2.06	1.82	1.60	1.70	1.54	1.44
0.45	10%	<i>ASY</i>	15.22	12.96	13.16	18.06	15.44	14.60	23.34	20.00	17.08
		<i>BOOT1</i>	12.38	11.26	10.24	12.84	11.80	10.66	14.80	12.90	10.56
		<i>BOOT2</i>	12.36	11.84	10.84	13.32	12.42	11.34	14.66	13.76	11.60
	5%	<i>ASY</i>	9.56	7.56	7.44	11.60	9.40	8.54	15.82	12.86	10.32
		<i>BOOT1</i>	7.06	6.30	5.80	7.56	6.64	5.82	8.32	7.76	5.62
		<i>BOOT2</i>	7.08	6.52	6.38	7.74	7.34	6.44	8.04	8.30	6.52
	1%	<i>ASY</i>	2.74	1.60	1.28	3.84	2.08	2.14	5.16	3.22	2.30
		<i>BOOT1</i>	2.38	2.18	1.98	2.86	2.48	1.92	2.88	2.36	1.66
		<i>BOOT2</i>	2.56	2.18	1.92	2.54	2.46	2.08	2.48	2.34	1.92

TABLE 4.2

Proportion of rejections in 5000 Monte Carlo experiments under $H_0 : \beta_0 = 0$ in model (1.1), when x_t and u_t are long memory with parameters d_x and d_u respectively. Bootstrap critical values are computed using 2000 bootstrap samples.

d_x/d_u	Size	T	$\rho = .0$			$\rho = .5$			$\rho = .9$		
			32	64	128	32	64	128	32	64	128
.40/.05	10%	<i>ASY</i>	15.54	12.94	11.20	17.78	15.30	14.40	22.28	19.14	16.00
		<i>BOOT1</i>	12.30	11.20	10.38	12.62	11.78	10.56	14.66	12.62	10.40
		<i>BOOT2</i>	12.24	11.66	10.82	12.96	12.40	11.20	13.64	13.32	11.36
	5%	<i>ASY</i>	9.68	7.80	7.34	11.44	9.34	8.20	15.60	11.40	9.60
		<i>BOOT1</i>	7.16	5.92	5.88	7.20	6.60	5.84	8.08	7.54	5.50
		<i>BOOT2</i>	7.16	6.38	6.24	7.28	7.04	6.34	7.76	7.80	6.44
	1%	<i>ASY</i>	2.84	1.58	1.26	3.60	1.92	2.00	4.52	2.52	1.84
		<i>BOOT1</i>	2.10	1.84	1.64	2.56	2.04	1.70	2.52	2.00	1.34
		<i>BOOT2</i>	2.34	2.02	1.78	2.44	2.34	1.96	2.22	2.00	1.72
.30/.15	10%	<i>ASY</i>	15.36	13.76	13.00	16.72	14.84	13.62	19.92	16.70	14.20
		<i>BOOT1</i>	12.10	11.10	10.64	12.56	11.28	10.44	15.04	12.64	10.52
		<i>BOOT2</i>	12.00	11.28	10.82	12.62	12.08	11.14	12.70	12.42	10.82
	5%	<i>ASY</i>	9.50	7.86	6.90	10.44	8.36	7.64	11.82	9.20	7.60
		<i>BOOT1</i>	6.94	5.90	5.70	7.10	6.00	5.60	8.22	7.20	5.06
		<i>BOOT2</i>	6.66	6.04	6.02	6.90	6.24	6.08	6.82	7.10	5.76
	1%	<i>ASY</i>	2.52	1.42	1.34	2.92	1.52	1.40	2.80	1.44	0.72
		<i>BOOT1</i>	1.86	1.46	1.44	2.00	1.74	1.52	2.44	1.54	1.16
		<i>BOOT2</i>	1.98	1.44	1.64	1.98	1.76	1.50	1.64	1.56	1.18

TABLE 4.3

Proportion of rejections in 5000 Monte Carlo experiments under $H_0 : \beta_0 = 0$ in model (1.1), when x_t and u_t are long memory with parameters d_x and d_u respectively, using the moving block bootstrap algorithm. Bootstrap critical values are computed using 2000 bootstrap samples.

d_x/d	$b =$ $\rho =$	2			4			8			16		
		.0	.5	.9	.0	.5	.9	.0	.5	.9	.0	.5	.9
.30/.00	10%	14.54	15.76	17.54	14.52	15.86	18.04	14.52	15.98	18.72	14.36	16.08	18.78
	5%	10.04	11.38	11.88	10.08	11.58	12.58	9.94	11.56	13.16	9.94	11.80	13.50
	1%	4.88	5.84	6.02	4.78	6.06	6.78	4.74	6.06	7.58	4.64	6.22	7.86
.45/.00	10%	18.36	20.54	23.32	18.20	20.60	23.96	18.02	20.92	24.34	18.16	20.86	24.82
	5%	13.46	15.58	18.08	13.40	15.82	18.72	13.38	15.96	19.64	13.20	15.94	19.84
	1%	7.88	9.36	11.54	7.84	9.78	12.72	7.74	9.66	13.62	7.96	9.86	14.20
.40/.05	10%	17.14	19.22	21.74	17.18	19.34	22.30	17.08	19.80	22.88	17.22	19.72	22.98
	5%	12.54	14.10	16.22	12.62	14.42	16.88	12.68	14.68	17.64	12.48	14.82	18.00
	1%	7.16	8.26	9.66	7.18	8.52	10.52	7.26	8.82	11.58	7.12	8.92	12.22
.30/.15	10%	15.38	16.86	18.30	15.58	17.36	18.80	15.40	17.60	19.20	15.52	17.52	19.96
	5%	11.02	11.74	12.48	11.06	11.98	13.12	11.14	12.32	13.98	11.32	12.36	14.26
	1%	5.56	6.18	6.16	5.76	6.46	7.16	5.84	6.82	8.22	5.80	6.92	8.60

TABLE 4.4

Proportion of rejections in 5000 Monte Carlo experiments under $H_0 : \beta_0 = 0$ in model (1.1), when x_t and u_t were generated as $AR(1)$ models. Bootstrap critical values are computed using 2000 bootstrap samples.

Size	T		$\rho_u = .0$			$\rho_u = .5$			$\rho_u = .9$		
			32	64	128	32	64	128	32	64	128
$\rho_x = 0.50$	10%	<i>ASY</i>	14.88	12.42	11.26	15.80	13.62	11.32	17.12	14.00	11.50
		<i>BOOT1</i>	12.34	11.00	10.70	12.40	11.34	10.38	13.14	11.58	10.28
		<i>BOOT2</i>	11.96	10.86	10.76	12.64	11.48	10.48	12.12	11.68	10.24
	5%	<i>ASY</i>	8.50	7.24	5.86	9.80	7.68	6.52	10.06	7.28	6.46
		<i>BOOT1</i>	6.60	5.46	5.66	7.46	5.90	5.44	7.64	5.56	5.34
		<i>BOOT2</i>	6.94	5.58	5.78	7.90	6.14	5.74	6.98	5.70	5.64
	1%	<i>ASY</i>	1.96	1.42	0.96	2.64	1.44	0.98	2.36	1.26	0.86
		<i>BOOT1</i>	2.12	1.54	1.40	2.94	1.50	1.50	2.58	1.52	1.30
		<i>BOOT2</i>	2.28	1.72	1.48	3.04	1.54	1.74	2.04	1.44	1.48
$\rho_x = 0.90$	10%	<i>ASY</i>	18.62	16.00	15.00	23.22	18.58	16.04	29.60	24.04	19.38
		<i>BOOT1</i>	13.86	12.42	11.94	14.62	12.72	12.26	18.10	13.94	13.00
		<i>BOOT2</i>	14.42	13.48	12.90	15.88	13.84	13.00	18.58	14.92	13.36
	5%	<i>ASY</i>	12.76	10.04	8.70	16.48	12.68	10.00	22.70	17.42	12.54
		<i>BOOT1</i>	9.56	7.88	7.36	10.06	8.06	7.40	12.82	8.72	7.26
		<i>BOOT2</i>	9.56	8.50	8.04	10.28	8.76	8.06	12.72	9.36	8.20
	1%	<i>ASY</i>	4.46	3.20	2.08	6.64	4.60	2.92	10.92	7.36	4.38
		<i>BOOT1</i>	5.30	4.22	3.18	6.16	4.28	3.32	7.04	4.34	2.74
		<i>BOOT2</i>	4.50	4.02	3.38	5.30	3.90	3.52	5.62	4.02	2.88

TABLE 4.5

Proportion of rejections in 5000 Monte Carlo experiments under $H_0 : \beta_0 = 0$ in model (1.1), when x_t and u_t were generated as $AR(1)$ models, and using the moving block bootstrap algorithm. Bootstrap critical values are computed using 2000 bootstrap samples.

ρ_u/ρ_x	$b =$	2	4	8	16
.00/.50	10%	14.32	14.36	14.30	14.46
	5%	10.02	10.18	10.18	10.18
	1%	4.76	4.74	4.76	4.72
.00/.90	10%	26.16	26.18	26.20	25.92
	5%	21.88	21.66	21.74	21.40
	1%	15.74	15.98	15.66	15.48
.50/.50	10%	15.60	15.72	15.76	15.70
	5%	10.30	10.66	10.76	10.72
	1%	5.20	5.32	5.46	5.50
.50/.90	10%	28.08	28.20	28.22	28.18
	5%	23.72	23.84	23.92	23.98
	1%	17.36	17.62	17.74	17.70
.90/.50	10%	16.18	16.10	16.46	16.72
	5%	10.66	11.28	11.52	11.82
	1%	5.00	5.54	6.00	6.08
.90/.90	10%	31.30	31.84	32.42	32.84
	5%	27.16	27.50	28.38	28.72
	1%	20.88	21.52	22.18	23.08