

# Distribution-free specification tests for dynamic linear models

MIGUEL A. DELGADO<sup>†</sup>, JAVIER HIDALGO<sup>‡</sup> AND CARLOS VELASCO<sup>†</sup>

<sup>†</sup>*Universidad Carlos III, 28903 Madrid, Spain*

E-mail: delgado@est-econ.uc3m.es, carlos.velasco@uc3m.es

<sup>‡</sup>*London School of Economics, Houghton Street, London WC2A 2AE, UK*

E-mail: f.j.hidalgo@lse.ac.uk

**Summary** This article proposes goodness-of-fit tests for dynamic regression models, where regressors are allowed to be only weakly exogenous and arbitrarily correlated with past shocks. The null hypothesis is stated in terms of the lack of serial correlation of the errors of the model. The tests are based on a linear transformation of a Bartlett's  $T_p$ -process of the residuals. This transformation approximates the martingale component of the process so that it converges weakly to the standard Brownian motion under the null hypothesis. One feature of our setup is that we do not require to specify the dynamic structure of the regressors. Due to this, the transformation employs a semi-parametric correction that does not restrict the class of local alternatives that our tests can detect, in contrast with other works using smoothing techniques. A Monte Carlo study illustrates the finite sample performance of the tests.

**Keywords:** *Dynamic models, Empirical processes, Exogeneity, Goodness-of-fit, Local alternatives, Martingale decomposition.*

## 1. INTRODUCTION

Delgado et al. (2005) (DHV henceforth) proposed asymptotically distribution-free tests for the correct parametric specification of the autocorrelation structure of a time series process. The tests were based on a parametric transformation of Bartlett's (1954)  $T_p$ -process, which entails to consider its martingale component, so that asymptotically the transformed process converges to a standard Brownian motion. The tests were applied to observable data, so there was no need to compute the residuals of the model, and the martingale transformation only depended on a set of unknown parameters under the null hypothesis. The aim of this paper consists in extending the DHV procedure to test the specification of dynamic regression models. Here, we use the empirical spectral process of the residuals of the model, because, in the presence of general explanatory variables, regression models do not specify completely the dynamics of the dependent variable, unlike the linear models studied by DHV. The transformation of the corresponding  $T_p$ -process depends, despite the unknown parameters, on the non-parametric cross-spectrum between the regressors and the regression error term, which is non-constant and different from zero when regressors are only assumed to be weakly exogenous. A feasible transformation might be computed via a non-parametric smoothed estimator of this cross-spectrum. However, we show that we can avoid the smoothing in the feasible martingale

transformation by using directly the cross-periodogram, even though it is an inconsistent estimate of the cross-spectrum. In spite of this non-parametric aspect of our model, our tests have non-trivial power against local alternatives converging to the null at the parametric rate  $n^{1/2}$ .

The remainder of the paper is organized as follows. Section 2 introduces the model and describes the testing problem. Section 3 presents the transformation to obtain asymptotically distribution-free tests, whereas Section 4 discusses the power of our tests. Section 5 describes a Monte Carlo experiment to shed some light on the finite sample performance of our test and how it compares with Portmanteau tests based on non-parametric smoothing as well as directional and smooth tests. Finally, the proofs have been placed in the Appendix.

## 2. DYNAMIC MODELS

This section discusses methods for the correct specification of dynamic regression models

$$X_t = \mu_0 + \alpha_{01}X_{t-1} + \cdots + \alpha_{0p}X_{t-p} + \beta'_0 Z_t + \varepsilon_t, \quad (2.1)$$

where  $Z_t$  is a  $q$ -dimensional vector of deterministic and/or (weakly) exogenous variables and where the parameter vector  $\theta'_0 = (\mu_0, \alpha'_0, \beta'_0)$  is identified as the solution of the  $p + q + 1$  moment conditions

$$\mathbb{E} [W_t (X_t - \theta' W_t)] = 0, \quad (2.2)$$

where  $W'_t = (1, X_{t-1}, \dots, X_{t-p}, Z'_t)$  and  $\mathbb{E}(W_t W'_t)$  is a positive definite matrix.

The models considered in (2.1), also known as *ARX* models, are an important extension of those examined in DHV. Notice that in  $Z_t$  we can allow some of its components to be lagged values, for example  $Z_{kt} = Z_{jt-\ell}$  for some  $\ell \geq 1$ . In the context of model (2.1), a natural assumption is that

$$\mathbb{E} [\varepsilon_t | \mathcal{F} \{\varepsilon_s, Z_{s+1}, s < t\}] = 0, \quad (2.3)$$

where  $\mathcal{F} \{\varepsilon_s, Z_{s+1}, s < t\}$  is the  $\sigma$ -algebra generated by  $\{\varepsilon_s, Z_{s+1}, s < t\}$ . Equation (2.3) implies that  $\mathbb{E}[Z_t \varepsilon_s] = 0$  for all  $s \geq t$ , although it allows for feedback from  $\varepsilon_t$  to  $Z_{t+j}$ ,  $j > 0$ . The latter implies that it is possible that the cross-autocovariance of  $Z_t$  and  $\varepsilon_t$  satisfies that  $\gamma_{Z\varepsilon}(j) = \mathbb{E}[Z_{t+j} \varepsilon_t] \neq 0$  for some  $j > 0$ . Denoting herewith the cross-spectral density function between the sequences  $\{U_t\}_{t \in \mathbb{Z}}$  and  $\{V_t\}_{t \in \mathbb{Z}}$  by  $f_{UV}$ , we have that one consequence of the latter is that the cross-spectral density function between the sequences  $\{Z_t\}_{t \in \mathbb{Z}}$  and  $\{\varepsilon_t\}_{t \in \mathbb{Z}}$ ,  $f_{Z\varepsilon}$ , defined by

$$\gamma_{Z\varepsilon}(j) = \int_{-\pi}^{\pi} f_{Z\varepsilon}(\lambda) e^{ij\lambda} d\lambda, \quad j = 0, \pm 1, \pm 2, \dots,$$

is not a null function. That is, the sequence  $\{Z_t\}_{t \in \mathbb{Z}}$  is only predetermined in (2.1).

The null hypothesis of interest is that the errors  $\{\varepsilon_t\}_{t \in \mathbb{Z}}$  in (2.1) are not autocorrelated. In other words, that the regression model (2.1) captures the linear dynamic structure of  $\{X_t\}_{t \in \mathbb{Z}}$ . More specifically, for given  $\theta$ , define the residuals  $\{\varepsilon_t(\theta)\}_{t \in \mathbb{Z}}$  by

$$\varepsilon_t(\theta) := X_t - \theta' W_t, \quad (2.4)$$

and its autocovariance structure by  $\gamma_\varepsilon(j; \theta) := \mathbb{E}(\varepsilon_t(\theta) \varepsilon_{t+j}(\theta))$ . Then, our null hypothesis of interest is

$$H_0 : \gamma_\varepsilon(j; \theta_0) = 0, \quad \text{for all } |j| \geq 1 \quad \text{and some } \theta_0 \in \Theta \subset \mathbb{R}^{p+q+1}.$$

We are interested in omnibus tests, where the alternative hypothesis is the negation of the null. The compact set  $\Theta := A \times \mathbb{R}^{q+1}$  is chosen such that for all  $\alpha \in A$ , all the roots of the polynomial

$$\alpha(z) := 1 - \alpha_1 z - \dots - \alpha_p z^p \quad (2.5)$$

are outside the unit disk. Notice that the least-squares estimator of the parameters may be inconsistent if  $H_0$  does not hold even if the true value of  $\alpha$  is zero.

REMARK 2.1. It is worth mentioning that we could allow the so-called *ARMAX* models, i.e. (2.1), where

$$\varepsilon_t = \varrho_1 \varepsilon_{t-1} + \dots + \varrho_\ell \varepsilon_{t-\ell} + \eta_t.$$

In this latter scenario, our null hypothesis would be that  $\{\eta_t\}_{t \in \mathbb{Z}}$  follows a white-noise sequence. However, we shall consider (2.1) because of its generality and mathematical simplicity in terms of arguments and notation. Also, extensions of (2.1) to non-linear specifications is fairly straightforward and will not be pursued in this paper.

As in DHV, we can write the null hypothesis  $H_0$  in the frequency domain. Indeed, let  $f_\varepsilon(\lambda; \theta)$  denote the spectral density function of  $\{\varepsilon_t(\theta)\}_{t \in \mathbb{Z}}$  in (2.4), that is

$$\gamma_\varepsilon(j; \theta) = \int_{-\pi}^{\pi} f_\varepsilon(\lambda; \theta) \exp(ij\lambda) d\lambda, \quad j = 0, \pm 1, \dots,$$

and denote its spectral distribution function as  $F_\varepsilon(\lambda; \theta_0)$ , i.e.

$$F_\varepsilon(\lambda; \theta) := 2 \int_0^\lambda f_\varepsilon(\omega; \theta) d\omega.$$

Under  $H_0$  we have respectively the spectral density and distribution functions of  $\{\varepsilon_t(\theta_0)\}_{t \in \mathbb{Z}} = \{\varepsilon_t\}_{t \in \mathbb{Z}}$ . Then, we can equivalently write the null hypothesis  $H_0$  as

$$H_0 : \frac{F_\varepsilon(\lambda; \theta_0)}{F_\varepsilon(\pi; \theta_0)} = \frac{\lambda}{\pi} \quad \text{for all } \lambda \in [0, \pi] \quad \text{and some } \theta_0 \in \Theta, \quad (2.6)$$

being the alternative hypothesis  $H_1$  the negation of  $H_0$ . Thus, the null hypothesis  $H_0$  in (2.6) states that there exists a parameter value  $\theta_0 \in \Theta$  such that the sequence  $\{\varepsilon_t(\theta_0)\}_{t \in \mathbb{Z}}$  has a constant spectral density function, i.e. they are uncorrelated.

A natural estimator of  $F_\varepsilon(\lambda; \theta)$  is

$$\hat{F}_n(\lambda; \theta) := \frac{2\pi}{\tilde{n}} \sum_{j=1}^{[\tilde{n}\lambda/\pi]} I_{\varepsilon\varepsilon}(\lambda_j; \theta), \quad (2.7)$$

where  $\lambda_j := 2\pi j/n$ , for  $j = 1, \dots, \tilde{n}$ ,  $\tilde{n} := [n/2]$ ,  $[\cdot]$  denoting the integer part, and

$$I_{\varepsilon\varepsilon}(\lambda; \theta) := \frac{1}{2\pi n} \left| \sum_{t=1}^n \varepsilon_t(\theta) e^{it\lambda} \right|^2$$

is the periodogram of the sequence  $\{\varepsilon_t(\theta)\}_{t=1}^n$  defined in (2.4). In what follows, for a generic function  $g(\cdot; \theta)$ , we shall suppress any reference to  $\theta$  when the function is evaluated at the true value  $\theta_0$ . That is,  $g(\cdot; \theta_0) =: g(\cdot)$ . Observe that the estimator  $\hat{F}_n(\lambda; \theta)$  is location invariant, due to the omission of  $j = 0$  in (2.7). Thus, there is no need to centre the residuals or to estimate the mean  $\mu$  in (2.1). See Remark 2.2 below for a more explicit explanation and some implications.

If the true value of  $\theta, \theta_0$ , were known, or equivalently if we could observe the sequence  $\{\varepsilon_t\}_{t=1}^n$ , following Bartlett (1954), we might perform a goodness-of-fit test using the  $T_p$ -process

$$\hat{T}_n(\omega; \theta) := \tilde{n}^{1/2} \left( \frac{\hat{F}_n(\pi\omega; \theta)}{\hat{F}_n(\pi; \theta)} - \omega \right), \quad \omega \in [0, 1], \quad (2.8)$$

evaluated at  $\theta = \theta_0$ . Recall that in this case, we denote  $\hat{T}_n(\omega; \theta_0)$  by  $\hat{T}_n(\omega)$ . Before we present the properties of  $\hat{T}_n(\omega)$ , let us introduce the following regularity assumption.

**ASSUMPTION 2.1.**  $\{\varepsilon_t\}_{t \in \mathbb{Z}}$  is a zero mean sequence of random variables such that  $\mathbb{E}(\varepsilon_t \varepsilon_s) = \sigma_\varepsilon^2 \mathcal{I}(t = s)$  and that  $\mathbb{E}|\varepsilon(t)|^k | \mathcal{F}_{t-1} = \varkappa_k$ ,  $k = 1, \dots, 3$ , and  $\mathbb{E}|\varepsilon(t)|^k = \mu_k$ ,  $k = 3, \dots, 8$  with  $\mu_8 < \infty$ , where  $\mathcal{F}_{t-1}$  is the  $\sigma$ -algebra of events generated by  $\{\varepsilon_s, Z_{s+1}, s < t\}$ .

Herewith, we are denoting the indicator function by  $\mathcal{I}(\cdot)$ . Assumption 2.1 is similar to that given in Dahlhaus (1985) who only assumed constant conditional moments up to the third order. This implies that the fourth-order spectral density function of the process  $\{\varepsilon_t\}_{t \in \mathbb{Z}}$  is not necessarily constant (cf. lemma 2 in DHV).

Now, denoting by  $B(\omega)$  the standard Brownian bridge on  $[0, 1]$ , we have the following proposition.

**PROPOSITION 2.1.** *Under Assumption 2.1, we have that*

$$\hat{T}_n(\cdot) \xrightarrow{d} B(\cdot) \text{ in the Skorohod metric space } \mathbb{D}[0, 1].$$

The statistic given in (2.8) is not feasible as it depends on the unknown vector of parameters  $\theta_0$ . To be able to compute (2.8), and so the test, we shall replace  $\theta_0$  by, for example, the least-squares estimator, denoted  $\hat{\theta}_n$ .

**ASSUMPTION 2.2.** *Under  $H_0$ , it holds that  $\hat{\theta}_n - \theta_0 = O_p(n^{-1/2})$ .*

Sufficient conditions for Assumption 2.2 are the stationarity of  $\{Z_t\}_{t \in \mathbb{Z}}$ , (2.3) and that  $\gamma_{Z\varepsilon}(0) = 0$ . Notice that in contrast to DHV, Assumption 2.2 does not require a linear expansion of  $\hat{\theta}_n$ , only its rate of convergence. This is due to the explicit solution of the least-squares estimator. Also, we shall not give explicit conditions under which the sequence  $\{Z_t \varepsilon_t\}_{t \in \mathbb{Z}}$ , and so  $\hat{\theta}_n$ , satisfies the central limit theorem.

**REMARK 2.2.** It is worth noticing that the least-squares estimator of  $(\alpha', \beta)'$  is given by the minimization of  $\hat{F}_n(\pi; \theta)$ . That is,

$$\begin{aligned} (\hat{\alpha}'_n, \hat{\beta}'_n) &= \arg \min_{(\alpha', \beta')} \sum_{j=1}^{\tilde{n}} |w_X(\lambda_j) - \alpha' w_{X-}(\lambda_j) - \beta' w_Z(\lambda_j)|^2 \\ &= \arg \min_{(\alpha', \beta')} \hat{F}_n(\pi; \theta), \end{aligned} \quad (2.9)$$

where  $w_X(\lambda_j)$ ,  $w_{X-}(\lambda_j)$  and  $w_Z(\lambda_j)$  are respectively the discrete Fourier transform of  $\{X_t\}_{t=1}^n$ ,  $\{X_{t-1}, \dots, X_{t-p}\}_{t=1}^n$  and  $\{Z_t\}_{t=1}^n$ . So, observing that we do not employ the frequency  $\lambda_j = 0$  to compute  $\hat{F}_n(\pi; \hat{\theta}_n)$ ,  $\hat{F}_n(\pi; \hat{\theta}_n)$  is independent of the intercept estimator  $\hat{\mu}_n$ . The latter implies that the computation of  $\hat{T}_n(\omega; \hat{\theta}_n)$  is independent of the intercept  $\mu$ . For this reason and to simplify notation, in what follows, we shall assume that there is no intercept in (2.1) and accordingly

that  $W_t' = (X_{t-1}, \dots, X_{t-p}, Z_t')$  and  $\theta = (\alpha', \beta')$ . Moreover when we have trend regressors, such as polynomial trends, apart from a different rate of convergence of  $\hat{\theta}_n$ , we have that the distribution of  $\hat{T}_n(\omega; \hat{\theta}_n)$  is asymptotically independent of the estimation of the trend component of the regression model. Hence, in what follows we can consider the model

$$X_t = \alpha_{01}X_{t-1} + \dots + \alpha_{0p}X_{t-p} + \beta_0'Z_t + \varepsilon_t \quad (2.10)$$

without loss of generality. Also, notice that if we employed tapers,  $\hat{T}_n(\omega; \hat{\theta}_n)$  would be invariant to the trend as well as to the intercept.

Now, once we have an estimator of the unknown parameters  $\theta_0$ , we can obtain the residuals as  $\hat{\varepsilon}_t := \varepsilon_t(\hat{\theta}_n) = X_t - \hat{\theta}_n'W_t$ , and with  $I_{\hat{\varepsilon}\hat{\varepsilon}}(\lambda_j) := I_{\varepsilon\varepsilon}(\lambda_j; \hat{\theta}_n)$ , we set

$$\hat{F}_n(\omega; \hat{\theta}_n) := \frac{2\pi}{\tilde{n}} \sum_{j=1}^{[\tilde{n}\omega]} I_{\hat{\varepsilon}\hat{\varepsilon}}(\lambda_j).$$

So, the feasible  $T_p$ -process is defined as in (2.8) but with  $\hat{\theta}_n$  replacing  $\theta$ . That is,

$$\hat{T}_n(\omega; \hat{\theta}_n) = \tilde{n}^{1/2} \left( \frac{\hat{F}_n(\pi\omega; \hat{\theta}_n)}{\hat{F}_n(\pi; \hat{\theta}_n)} - \omega \right). \quad (2.11)$$

Before we describe the asymptotic properties of  $\hat{T}_n(\omega; \hat{\theta}_n)$ , we introduce the following regularity assumption.

**ASSUMPTION 2.3.** (i) *The cross-spectrum  $f_{Z\varepsilon}(\lambda)$  is differentiable at all  $\lambda \in [-\pi, \pi]$ .* (ii) *The spectral density matrix  $f_{ZZ}(\lambda)$  is continuous for all  $\lambda \in [-\pi, \pi]$ .* (iii) *The higher order (cross) spectral densities up to eighth order of  $\{Z_t\}_{t \in \mathbb{Z}}$  and  $\{\varepsilon_t\}_{t \in \mathbb{Z}}$  are bounded.*

Assumption 2.3(i) could be replaced by some Lipschitz condition, but that might complicate some of the technical arguments. Nevertheless, the assumption as it stands is very mild and it is satisfied for most models employed with real data. Next, because all the roots of the polynomial  $\alpha(z)$  in (2.5) are outside the unit disk, we obtain that the stationary solution of  $X_t$  is given by  $\alpha_0(L)^{-1}(\varepsilon_t - \beta_0'Z_t)$ , where  $\alpha_0(z)$  is defined in (2.5) with  $\alpha = \alpha_0$ . Thus, it follows that

$$f_{X-, \varepsilon}(\lambda) = \frac{\mathbf{L}_p(e^{i\lambda})}{\alpha_0(e^{i\lambda})} \left( \frac{\sigma_\varepsilon^2}{2\pi} - \beta_0' f_{Z\varepsilon}(\lambda) \right), \quad (2.12)$$

with  $\mathbf{L}_p(z) = (z, \dots, z^p)'$ , so that Assumption 2.3(i) implies that  $f_{W\varepsilon}(\lambda)$  is differentiable everywhere in  $\lambda \in [-\pi, \pi]$ .

One implication of (2.12) and Assumption 2.1 is that  $\Phi(1) = 0$ , where

$$\Phi(\omega) := \int_0^\omega \phi(v)dv, \quad \omega \in [0, 1]$$

and

$$\phi(\omega) = 4\pi \operatorname{Re} f_{W\varepsilon}(\pi\omega) = 4\pi \operatorname{Re}(f_{X-, \varepsilon}(\pi\omega)', f_{Z\varepsilon}(\pi\omega)'),$$

by orthogonality between  $\{W_t\}_{t \in \mathbb{Z}}$  and  $\{\varepsilon_t\}_{t \in \mathbb{Z}}$  and evenness (oddness) of the real (imaginary) part of  $f_{W\varepsilon}(\lambda)$ . However, it is important to emphasize that we are not assuming that  $f_{W\varepsilon}(\lambda) = 0$  for all  $\lambda$ . In fact, this is not the case because  $\mathbb{E}[Z_t \varepsilon_s]$  can be different than zero for some  $t > s$ . This is one of the main features of our specification in (2.1)/(2.10).

On the other hand, Assumptions 2.3(i)–(ii) imply that  $f_{WW}(\lambda)$  is bounded for all  $\lambda \in [-\pi, \pi]$  because

$$f_{WW}(\lambda) = \begin{pmatrix} f_{X-,X-}(\lambda) & f_{X-,Z}(\lambda) \\ f_{ZX-}(\lambda) & f_{ZZ}(\lambda) \end{pmatrix},$$

with

$$\begin{aligned} f_{X-,X-}(\lambda) &= \frac{\mathbf{L}_p(e^{i\lambda})}{\alpha_0(e^{i\lambda})} \left( \frac{\sigma_\varepsilon^2}{2\pi} + 2\beta'_0 \operatorname{Re} f_{Z\varepsilon}(\lambda) + \beta'_0 f_{ZZ}(\lambda) \beta_0 \right) \frac{\mathbf{L}_p(e^{-i\lambda})'}{\alpha_0(e^{-i\lambda})} \\ f_{X-,Z}(\lambda) &= \frac{\mathbf{L}_p(e^{i\lambda})}{\alpha_0(e^{i\lambda})} (f_{\varepsilon Z}(\lambda) + \beta'_0 f_{ZZ}(\lambda)). \end{aligned}$$

Finally, Assumption 2.3(iii) implies eight finite moments for  $Z_t$  and  $X_t$  as assumed for  $\varepsilon_t$  in Assumption 2.1. However, the requirement of higher order bounded spectra function of  $\{W_t\}_{t \in \mathbb{Z}}$  can be relaxed as in DHV at the expense of much lengthier arguments.

In what follows, for two vector sequences  $\{V_t\}_{t=1}^n$  and  $\{U_t\}_{t=1}^n$ , we denote its cross-periodogram by

$$I_{VU}(\lambda) := \frac{1}{2\pi n} \left( \sum_{t=1}^n V_t e^{it\lambda} \right) \left( \sum_{t=1}^n U_t e^{-it\lambda} \right)'.$$

PROPOSITION 2.2. *Assuming Assumptions 2.1–2.3, under  $H_0$  we have that*

$$\begin{aligned} \hat{T}_n(\omega; \hat{\theta}_n) &= \hat{T}_n(\omega) - \frac{4\pi}{\sigma_\varepsilon^2 \tilde{n}} \sum_{j=1}^{[\omega \tilde{n}]} \operatorname{Re} I_{\varepsilon W}(\lambda_j) \tilde{n}^{1/2} (\hat{\theta}_n - \theta_0) + o_p(1) \\ &= \hat{T}_n(\omega) - \Phi(\omega)' \tilde{n}^{1/2} (\hat{\theta}_n - \theta_0) / \sigma_\varepsilon^2 + o_p(1), \end{aligned} \quad (2.13)$$

where the  $o_p(1)$  is uniformly in  $\omega \in [0, 1]$ .

REMARK 2.3. The second equality in (2.13) follows because under weak regularity conditions, Brillinger (1981) implies that

$$\sup_{\omega} \left\| \frac{2\pi}{\tilde{n}} \sum_{j=1}^{[\omega \tilde{n}]} (I_{VU}(\lambda_j) - f_{VU}(\lambda_j)) \right\| = o_p(1).$$

REMARK 2.4. Proceeding as with the proof of Theorem 2 of DHV, Propositions 2.1 and 2.2 imply that the asymptotic distribution of  $\hat{T}_n(\omega; \hat{\theta}_n)$  depends, in general, on  $\hat{\theta}_n$  and so on the model as in other goodness-of-fit tests with estimated parameters. However, since the aim of the paper is to describe distribution-free (pivotal) tests, we will not explicitly examine the asymptotic distribution of  $\hat{T}_n(\omega; \hat{\theta}_n)$ .

REMARK 2.5. (*Strong*) *exogeneity and predetermined regressors.* When the regressors  $Z_t$  are (strong) exogenous, we have that  $f_{Z\varepsilon}(\lambda) = 0$  for all  $\lambda \in [0, \pi]$ , and hence  $\phi(\omega) = (\phi_1(\omega)', 0'_q)'$ , where  $\phi_1(\omega) := 4\pi \operatorname{Re} f_{X\varepsilon}(\pi\omega)$ . So, the latter together with (2.13) implies that

$$\hat{T}_n(\omega; \hat{\theta}_n) = \hat{T}_n(\omega) - \left( \int_0^\omega \phi_1'(v) dv \right) \tilde{n}^{1/2} (\hat{\alpha}_n - \alpha_0) / \sigma_\varepsilon^2 + o_p(1).$$

That is, similar to the case where regressors are deterministic, the estimation of  $\beta$  in (2.1) has no influence on the asymptotic distribution of  $\hat{T}_n(\omega; \hat{\theta}_n)$ , only the least-squares estimator of  $\alpha_0$ . Moreover, in this case the function  $\Phi(\omega)$  is known up to a set of parameters which can be consistently estimated by Assumption 2.2. But this case was already covered by DHV, and hence it is not of interest in this paper. On the other hand, it is worth mentioning that the null hypothesis that one particular component of  $Z_t$  is (strong) exogenous can be tested using the methods put forward in the paper.

From Proposition 2.2 and Remark 2.4, it is obvious that tests based on continuous functionals of  $\hat{T}_n(\omega; \hat{\theta}_n)$  are not pivotal, as their asymptotic distribution depends on the model specified under the null hypothesis  $H_0$  and on the unknown function  $\phi(\cdot)$ . The latter function not only depends on  $\theta_0$  but also on the joint dynamic properties of  $\{Z_t\}_{t \in \mathbb{Z}}$  and  $\{\varepsilon_t\}_{t \in \mathbb{Z}}$  described by  $f_{Z\varepsilon}$ , which is unknown to the practitioner. The next section introduces a linear transformation of  $\hat{T}_n(\omega; \hat{\theta}_n)$  which converges weakly, under  $H_0$ , to the standard Brownian motion, denoted  $B^0(\cdot)$ , whose critical values are readily available.

### 3. DISTRIBUTION-FREE TESTS

We are looking for a linear transformation, say  $\bar{\mathcal{L}}$ , such that  $\bar{\mathcal{L}}\hat{T}_n(\cdot; \hat{\theta}_n)$  converges weakly to the standard Brownian motion  $B^0$  under  $H_0$ . This transformation must remove the effect of  $\Phi(\omega)' \tilde{n}^{1/2}(\hat{\theta}_n - \theta_0)$  into the asymptotic linear expansion of  $\hat{T}_n(\omega; \hat{\theta}_n)$ ; see Proposition 2.2. As pointed out in Remarks 2.2 and 2.4, we shall only consider the interesting case where the regressors  $Z_t$  are only predetermined, but not strictly exogenous, so that the cross-spectral density  $f_{Z\varepsilon}(\lambda)$  is not constant.

Abbreviating for a generic function  $h(\cdot)$ ,  $h(\lambda_j)$  by  $h_j$ , and denoting  $m_j = 2\pi I_{\varepsilon\varepsilon, j} - \sigma_\varepsilon^2$ , we observe that, applying Proposition 2.2, we can write  $\hat{T}_n(\omega; \hat{\theta}_n)$ , up to terms of order  $o_p(1)$ , as

$$\frac{\tilde{n}^{-1/2}}{\hat{F}_n(\pi)} \sum_{j=1}^{[\omega\tilde{n}]} m_j - \frac{\Phi(\omega)' \tilde{n}^{1/2}}{\hat{F}_n(\pi)} \left( \sum_{j=1}^{\tilde{n}} I_{WW, j} \right)^{-1} \sum_{j=1}^{\tilde{n}} \text{Re} I_{W\varepsilon, j} - \frac{\omega \tilde{n}^{-1/2}}{\hat{F}_n(\pi)} \sum_{j=1}^{\tilde{n}} m_j, \quad (3.1)$$

which is similar to the corresponding expression given in DHV but with our generic definition of  $\phi(\omega)$ . However, unlike DHV, this expression (3.1) cannot be directly identified as a CUSUM of least-squares residuals. Nevertheless, a similar martingale transformation based on a forward projection on the function  $g(u) := (1, \phi(u)')'$  will remove the terms in (3.1) depending on  $\int_0^\omega g(u) du$ , i.e.  $\Phi(\omega)$  and  $\omega$ . The latter are the non-martingale components in the tied-down empirical process with estimated parameters  $\hat{T}_n(\omega; \hat{\theta}_n)$ .

So, following similar arguments to those in DHV, we propose as our transformation  $\bar{\mathcal{L}}$ ,

$$\bar{\mathcal{L}}\hat{T}_n(\omega; \hat{\theta}_n) := \hat{T}_n(\omega; \hat{\theta}_n) - \frac{\tilde{n}^{-1/2}}{\hat{F}_n(\pi; \hat{\theta}_n)} \sum_{j=1}^{[\omega\tilde{n}]} g'_j \left( \sum_{k=j+1}^{\tilde{n}} g_k g'_k \right)^{-1} \sum_{k=j+1}^{\tilde{n}} g_k \hat{m}_k, \quad (3.2)$$

where  $\hat{m}_k := 2\pi I_{\hat{\varepsilon}\hat{\varepsilon}, k} - \hat{F}_n(\pi; \hat{\theta}_n)$ ,  $\bar{n} = \tilde{n} - p - q - 1$ . The limiting continuous version of  $\bar{\mathcal{L}}$  is defined, for a generic function  $\xi : [0, 1] \mapsto \mathbb{R}$ , as

$$\mathcal{L}^0 \xi(\omega) := \xi(\omega) - \int_0^\omega g(v)' \Sigma^{-1}(v) \int_v^1 g(u) \xi(du) dv,$$

with  $\Sigma(v) := \int_v^1 g(u)g(u)'du$ .

Before we examine the properties of  $\bar{\mathcal{L}}\hat{T}_n(\omega; \hat{\theta}_n)$  in (3.2), we need to introduce the following assumption.

ASSUMPTION 3.1. *The matrix  $\tilde{n}^{-1} \sum_{k=\tilde{n}+1}^{\tilde{n}} g_k g_k'$  is non-singular.*

THEOREM 3.1. *Assume Assumptions 2.1–2.3 and 3.1. Then, under  $H_0$ ,*

$$\bar{\mathcal{L}}\hat{T}_n(\omega; \hat{\theta}_n) \xrightarrow{d} B^0 \text{ in the Skorohod metric space } \mathbb{D}[0, 1].$$

The transformation  $\bar{\mathcal{L}}$  is infeasible, as it depends on the unknown function  $g(u)$ . To construct a feasible version of  $\bar{\mathcal{L}}$ , we need to replace  $g(u)$  by some estimate. Recall that from (2.12),  $\phi(\omega) = 4\pi \text{Re}(f_{X-, \varepsilon}(\pi\omega)', f_{Z\varepsilon}(\pi\omega)')$  and because  $f_{Z\varepsilon}$  is an unknown function, we have that  $\phi$  is a non-parametric function. The latter is one of the main differences with DHV's paper. Because of that, we shall propose two feasible transformations. The first one employs the standard average periodogram estimator of the (scaled real part) of the cross-spectrum between  $\{W_t\}_{t \in \mathbb{Z}}$  and  $\{\varepsilon_t\}_{t \in \mathbb{Z}}$ , i.e.

$$\hat{\phi}_m(j/\tilde{n}) = \hat{\phi}_{m,j} := \frac{4\pi}{\bar{K}_m} \sum_{\ell=-m; \ell \neq 0}^m K_\ell \text{Re} I_{W\varepsilon, j+\ell}, \quad (3.3)$$

where  $K_\ell = K(\ell/m)$  and  $\bar{K}_m = \sum_{\ell=-m; \ell \neq 0}^m K_\ell$ . The second approach replaces  $f_{W\varepsilon}$  by the cross-periodogram. The latter is a much more delicate matter, as the periodogram is not a consistent estimator of  $f_{W\varepsilon}$ , only unbiased, unlike the former approach or that in DHV, where the function  $\phi(\cdot)$  was known up to a finite set of parameters.

ASSUMPTION 3.2. (i)  $K(x)$  is a non-negative continuous symmetric function in  $[-1, 1]$ . (ii)  $m^{-2} n^{1+\delta} + mn^{-1} \rightarrow 0$ , for some  $\delta > 0$ .

Non-parametric adjustment in related contexts has been also examined in Stute et al. (1998) and Stute and Zhu (2002). The estimator  $\hat{\phi}_{m,j}$  is of the leave-one-out type as it does not use the frequency  $\lambda_j$  in its computation. The latter is done to guarantee the orthogonality in finite samples of  $\hat{\phi}_{m,j}$  with respect to  $I_{\hat{\varepsilon}\hat{\varepsilon}, j}$  for all  $m$ , using the well-known result of the approximate orthogonality between the discrete Fourier transform of vector time series at different Fourier frequencies.

We need to strengthen Assumption 2.3.

ASSUMPTION 2.3'. *Assumption 2.3 holds and  $f_{W\varepsilon}(\lambda)$  has two bounded derivatives.*

Thus, in practice, we can take the discrete sample counterpart of  $\bar{\mathcal{L}}\hat{T}_n(\omega; \hat{\theta}_n)$ ,

$$\bar{\mathcal{L}}_n \hat{T}_n(\omega; \hat{\theta}_n) := \hat{T}_n(\omega; \hat{\theta}_n) - \frac{\tilde{n}^{-3/2}}{\hat{F}_n(\pi; \hat{\theta}_n)} \sum_{j=1}^{[\tilde{n}\omega]} \hat{g}_m\left(\frac{j}{\tilde{n}}\right)' \hat{\Sigma}_m\left(\frac{j}{\tilde{n}}\right)^{-1} \sum_{k=1+j}^{\tilde{n}} \hat{g}_m\left(\frac{k}{\tilde{n}}\right) \hat{m}_k, \quad (3.4)$$

where  $\hat{g}_m(\omega) := (1, \hat{\phi}_m(\omega)')$  and  $\hat{\Sigma}_m(\omega) := \tilde{n}^{-1} \sum_{j=1+[\tilde{n}\omega]}^{\tilde{n}} \hat{g}_{m,j} \hat{g}_{m,j}'$ .

THEOREM 3.2. *Assuming Assumptions 2.1–2.2, 2.3' and 3.1–3.2, under  $H_0$ ,*

$$\bar{\mathcal{L}}_n \hat{T}_n(\omega; \hat{\theta}_n) \xrightarrow{d} B^0 \text{ in the Skorohod metric space } \mathbb{D}[0, 1].$$



Note that the proof of this result does not show that  $\sup_{\omega \in [0,1]} |\bar{\mathcal{L}}_n \hat{T}_n(\omega; \hat{\theta}_n) - \bar{\mathcal{L}} \hat{T}_n(\omega)| = o_p(1)$  as it was necessary in DHV's proofs.

We now describe the unsmoothed version of the feasible transformation. Here the aim is to use the cross-periodogram instead of  $g_k$  or a consistent estimate of it. We propose to employ the transformation

$$\check{\mathcal{L}}_n \hat{T}_n(\omega; \hat{\theta}_n) = \hat{T}_n(\omega; \hat{\theta}_n) - \frac{\tilde{n}^{-1/2}}{\hat{F}_n(\pi; \hat{\theta}_n)} \sum_{j=1}^{[\tilde{n}\omega]} \hat{g}'_{j+1} \left( \sum_{k=j+1}^{\tilde{n}} \hat{g}_{k+2} \hat{g}'_{k+2} \right)^{-1} \sum_{k=j+1}^{\tilde{n}} \hat{g}_{k+2} \hat{m}_{k+1}, \quad (3.5)$$

where  $\hat{g}_j = I_{W\hat{\varepsilon}, j}$ ,  $j = 1, \dots, \tilde{n}$ .

The reason to employ, for example  $\sum_{k=j+1}^{\tilde{n}} g_{k+2} \hat{m}_{k+1}$  instead of  $\sum_{k=j+1}^{\tilde{n}} g_k \hat{m}_k$  as in (3.4) is because, contrary to the latter, there is leverage effect from  $g_{j+1}$  into  $\sum_{k=j+1}^{\tilde{n}} g_k \hat{m}_k$  which does not vanish sufficiently fast, as in the case with the smoothed version or in the case examined in DHV. At the same time, we guarantee that  $\hat{g}_{k+2} \hat{m}_{k+1}$  is approximately centred because  $\hat{g}$  and  $\hat{m}$  have different indices. Then, we have our next result.

**THEOREM 3.3.** *Assuming Assumptions 2.1–2.2, 2.3' and 3.1–3.2, under  $H_0$ , the unsmoothed transformation given in (3.5) satisfies that*

$$\check{\mathcal{L}}_n \hat{T}_n(\omega; \hat{\theta}_n) \xrightarrow{d} B^0 \text{ in the Skorohod metric space } \mathbb{D}[0, 1].$$

Theorems 3.2 and 3.3 justify asymptotic admissible tests based on continuous functionals of  $\check{\mathcal{L}}_n \hat{T}_n(\omega; \hat{\theta}_n)$ , as stated in the following corollary.

**COROLLARY 3.1.** *For any continuous functional  $\varphi : \mathbb{D}[0, 1] \mapsto \mathbb{R}^+$ , under  $H_0$  and assuming the same conditions of Theorem 3.3,*

$$\varphi(\check{\mathcal{L}}_n \hat{T}_n(\omega; \hat{\theta}_n)) \xrightarrow{d} \varphi(B^0).$$

Note that the non-parametric estimation does not affect first-order asymptotics of the tests, which have the same limiting behaviour as if  $g$  were known or parametrically modelled. However, the need to invert the  $(p + q + 1) \times (p + q + 1)$  matrix  $\hat{\Sigma}_m(\omega)$  in a discrete grid  $\omega = j/\tilde{n}$ , implies that this is only possible at  $j = 1, \dots, \tilde{n}$  due to the loss of degrees of freedom as we need to estimate the parameters in the regression model (2.1).

The distribution of  $\varphi(B^0)$  can be tabulated by Monte Carlo. For the main goodness-of-fit proposals, Kolmogorov–Smirnov and Cramér–von Mises,  $\varphi(B^0)$  is already tabulated, for instance in Shorack and Wellner (1986, pp. 34 and 748).

#### 4. LOCAL ALTERNATIVES AND CONSISTENCY

We consider two types of local alternatives, first a parametric one and secondly a more general non-parametric type of alternative which it may suggest or establish the origin of the possible misspecification of the model given in (2.1).

To study the power of our test let us consider local alternatives of the type

$$H_{an} : \alpha_{0,p+1} = \frac{c}{\tilde{n}^{1/2}} \quad \text{for some } c \neq 0. \quad (4.1)$$

Similar results are available for other forms of misspecification, including errors in the modelling of the relationship between the sequences  $\{Z_t\}_{t \in \mathbb{Z}}$  and  $\{X_t\}_{t \in \mathbb{Z}}$ .

THEOREM 4.1. *Assuming the same conditions as in Theorem 3.3, under  $H_{an}$ ,*

$$\bar{\mathcal{L}}_n \hat{\Gamma}_n \xrightarrow{d} B^0 + c\mathcal{L}^0\Psi \text{ in the Skorohod metric space } \mathbb{D}[0, 1], \quad (4.2)$$

where  $\Psi(\omega) := \sigma_\varepsilon^{-2} \int_0^\omega \phi_{p+1}(u) du$ , with

$$\phi_{p+1}(v) := 4\pi \operatorname{Re} f_{\varepsilon X_{-p-1}}(\pi v) = \operatorname{Re} \left\{ \frac{\exp(i(p+1)\pi v)}{\alpha_0(e^{i\pi v})} (2\sigma_\varepsilon^2 + 4\pi\beta'_0 \operatorname{Re} f_{Z_\varepsilon}(\pi v)) \right\}.$$

REMARK 4.1. Under the set of assumptions in the previous section, the proposed test does not have trivial power, as stated in the following theorem if  $Z_t$  cannot explain all the information contained in  $X_{t-p-1}$  at all frequencies. i.e. there is a set of positive Lebesgue measure where the spectral density matrix of  $(Z'_t, X_{t-p-1})'$  has full rank. This should imply that in a set of positive Lebesgue measure the cross-spectral density  $f_{X_{t-p-1}\varepsilon}(\lambda)$  is not a linear combination of the rows of  $f_{Z_\varepsilon}(\lambda)$ , which guarantees that  $\mathcal{L}^0\Psi$  is not zero for all  $\lambda$ .

Therefore, for a suitable continuous functional  $\varphi : \mathbb{D}[0, 1] \mapsto \mathbb{R}^+$ , such as the Cramér–von Mises or the Kolmogorov–Smirnov,  $\Pr[\varphi(B^0 + \mathcal{L}^0\Psi) > \varphi(B^0)] = 1$ , and the test will detect local departures from the null of the type  $H_{an}$  given in (4.1).

#### 4.2. Non-parametric alternatives

We now consider the case when  $\{\varepsilon_t\}_{t \in \mathbb{Z}}$  has not flat spectrum up to an  $n^{-1/2}$  factor. Notice that  $H_{an}$  implies that the spectral density function of  $\{\varepsilon_t\}_{t \in \mathbb{Z}}$ , where  $\theta$  does not include  $\alpha_{p+1}$ , is

$$\begin{aligned} f_\varepsilon(\lambda; \theta_0) &= \frac{\sigma_\varepsilon^2}{2\pi} + \frac{2c}{\tilde{n}^{1/2}} \operatorname{Re} f_{\varepsilon X_{-p-1}}(\lambda) + \frac{c^2}{\tilde{n}} f_{X_{-p-1}}(\lambda) \\ &= \frac{\sigma^2}{2\pi} \left\{ 1 + c \frac{\phi_{p+1}(\lambda/\pi)}{\sigma^2} \tilde{n}^{-1/2} + O(c^2 \tilde{n}^{-1}) \right\}. \end{aligned}$$

So, we could consider non-parametric alternatives of the type

$$H'_{an} : f_\varepsilon(\lambda; \theta_0) = \frac{\sigma^2}{2\pi} (1 + l(\lambda) \tilde{n}^{-1/2}) \quad \text{for some } \theta_0 \in \Theta,$$

where the function  $l(\cdot)$  is not in the space spanned by  $\phi(\cdot/\pi)$ . The latter implies that the correlation structure of  $\{\varepsilon_t\}_{t \in \mathbb{Z}}$  cannot be explained either by lag values of  $X_t$  or by any of the components of the variables  $Z_t$ . It is worth noticing that the test has maximum power against alternatives for which  $l(\cdot)$  belongs to the orthogonal space spanned by  $g$ . Then Theorem 4.1 holds for  $H'_{an}$  with  $\Psi(\omega) := \int_0^\omega l(\pi u) du$  and  $c = 1$  there.

The test is consistent in the direction of general fixed non-parametric or parametric alternatives in (4.1), such as  $\alpha_{p+1} = c$ ,  $c \neq 0$ . Though a precise justification under suitable regularity conditions is possible, this is beyond the scope of this paper and we will only provide a sketch of the main arguments. Assuming certain regularity conditions (such as that  $\alpha(L)$  has all roots outside the unit circle), Assumption 2.1 could be replaced by a linear process specification and Assumption 2.2 is satisfied under the alternative hypothesis  $H_1$ , where now  $\theta_0$  denotes the pseudo true value, defined by  $\theta_0 := \arg \min_{\theta \in \Theta} F(\pi; \theta)$ , which is such that the pseudo-innovations  $\{\varepsilon_t(\theta_0)\}$  are autocorrelated under  $H_1$ . Denote by  $f_\varepsilon(\lambda) := f_\varepsilon(\lambda; \theta_0)$  the (non-constant) spectral density of  $\{\varepsilon_t\}_{t \in \mathbb{Z}}$ . Indeed, proceeding as in DHV or Dahlhaus and Wefelmeyer (1996), we shall have that, for each  $\omega \in [0, 1]$ ,

$$\hat{T}_n(\omega; \hat{\theta}_n) = \hat{T}_n(\omega) + \Phi(\omega)' \frac{\tilde{n}^{1/2}(\hat{\theta}_n - \theta_0)}{\hat{F}_n(\pi)} + o_p(1).$$

Now,

$$\hat{T}_n(\omega) = \frac{1}{\tilde{n}^{1/2}} \frac{2\pi}{\hat{F}_n(\pi)} \sum_{j=1}^{[\tilde{n}\omega]} \left( \frac{I_{\varepsilon\varepsilon,j}}{f_{\varepsilon,j}} - 1 \right) f_{\varepsilon,j} + \tilde{n}^{1/2} \left( \frac{2\pi}{\hat{F}_n(\pi)} \frac{1}{\tilde{n}} \sum_{j=1}^{[\tilde{n}\omega]} f_{\varepsilon,j} - \omega \right),$$

where, under suitable regularity conditions, the first term on the right-hand side of the last display expression is  $O_p(1)$ , whereas the expression inside the parenthesis of the second term on the right-hand side converges to a constant for each  $\omega$ . Thus,  $|\hat{T}_n(\omega)|$  and  $|\tilde{\mathcal{L}}_n \hat{T}_n(\omega)|$  diverge to infinity at the rate  $n^{1/2}$ . From here, the consistency of the test follows by standard arguments.

Following the discussion in DHV, we can use Theorem 4.1 to derive optimal tests for  $H_0$  against the direction  $l$  given in  $H'_{an}$ . These test statistics are based on  $\tilde{\mathcal{L}}_n \hat{T}_n(\lambda)$  and thus they are also asymptotically distribution-free under  $H_0$ .

## 5. MONTE CARLO EXPERIMENT

This section presents a small simulation exercise to shed some light on the small sample behaviour of our tests. To that end, we have considered the ARX(1, 1) model

$$X_t = \alpha_1 X_{t-1} + \beta_1 Z_{1t} + \varepsilon_t, \quad t = 1, \dots, n, \quad (5.1)$$

where

$$\begin{aligned} Z_{1t} &= aZ_{1(t-1)} + u_t, \\ u_t &= (1 - b^2)^{1/2} v_t + b\varepsilon_{t-1}, \end{aligned}$$

and  $\{v_t\}_{t \in \mathbb{Z}}$  and  $\{\varepsilon_t\}_{t \in \mathbb{Z}}$  are mutually independent i.i.d.  $N(0, 1)$  variates. We have employed three sample sizes  $n = 100, 200, 400$ , and the following values of the parameters:

$$\beta_1 = \{0.2, 0.5, 1.0\}, \quad \alpha_1 \in \{0.2, 0.5, 0.8\}, \quad b \in \{0, 0.4, 0.8\},$$

whereas  $a = 0.5$  for all the combinations and sample sizes. The autoregressive parameters  $\alpha_1$  and  $a$  control partially the dependence structure of  $\{X_t\}_{t \in \mathbb{Z}}$  and  $\{Z_t\}_{t \in \mathbb{Z}}$ . On the other hand,  $b$  measures the ‘endogeneity’ of  $\{Z_t\}_{t \in \mathbb{Z}}$  in (5.1) (so that  $Z_t$  is strongly exogenous if  $b = 0$ ), together with the regression coefficient  $\beta_1$ .

We first estimate the parameters  $\alpha_1$ ,  $\beta_1$  and  $\sigma_\varepsilon^2$  in (5.1) by (2.9), and for a given feasible transformation  $\mathcal{L}_n^m$  of  $\hat{T}_n$  we compute the Cramér–von Mises statistic

$$C_n^m := \frac{1}{\tilde{n} - 3} \sum_{j=1}^{\tilde{n}-3} \left( \mathcal{L}_n^m \hat{T}_n \left( \frac{j}{\tilde{n}} \right) \right)^2,$$

where  $m$  indicates the type of approximation of  $\phi$  employed. We have considered three alternatives for the martingale transformation. The first one uses a non-consistent estimator of  $\phi$ , using the transformation  $\check{\mathcal{L}}_n$ , and it is denoted as  $C_n^0$  in Tables 1–5. For the cases where we estimate consistently  $\phi$ , we use the Tuckey–Hanning kernel in (3.3),

$$K_m(x) = \frac{1}{2} \left( 1 + \cos \left( \frac{\pi x}{m} \right) \right),$$

with bandwidth parameters  $m = [0.25n^{0.9}]$  and  $[0.30n^{0.9}]$ .

To be able to make comparisons we provide the results for the popular Ljung and Box's (1978) Portmanteau test

$$Q_p := n(n+2) \sum_{j=1}^p \frac{\hat{\rho}_\varepsilon(j)^2}{n-j},$$

where

$$\hat{\rho}_\varepsilon(j) := \left( \sum_{t=1}^n \hat{\varepsilon}_t^2 \right)^{-1} \sum_{t=j+1}^n \hat{\varepsilon}_t \hat{\varepsilon}_{t-j}, \quad j \geq 1,$$

are the sample autocorrelations of the residuals  $\{\hat{\varepsilon}_t\}_{t=1}^n$  for two choices for  $p$ . For  $n = 100, 200$ , we chose  $p = 10, 15$ , whereas for  $n = 400$ ,  $p = 15, 20$ . Those choices are close to  $n^{1/2}$ , which seems a reasonable compromise in terms of size and power. As in Hong (1996), we employ a standardized version of  $Q_p$  which we compare against the standard normal critical values.

For power comparisons we consider two local alternatives. The first one is based on the ARX(2, 1) model,

$$X_t = \alpha_1 X_{t-1} + \frac{5}{n^{0.5}} X_{t-1} + \beta_1 Z_{1t} + \varepsilon_t,$$

whereas the second local alternative is the ARMAX(1, 1, 1) model

$$X_t = \alpha_1 X_{t-1} + \beta_1 Z_{1t} + \frac{5}{n^{0.5}} \varepsilon_{t-1} + \varepsilon_t.$$

We report the percentage of rejections in 100,000 Monte Carlo replications.

The empirical size for tests based on  $C_n^0$  show an improvement with the sample size, but it also appears to depend on the model under consideration. More specifically, the percentage of rejections under  $H_0$  increases with  $\alpha_1$ ,  $b$  and  $\beta_1$  for all sample sizes. On the other hand, those for  $C_n^m$  are more stable, although there is some dependence on the value of  $b$ , perhaps due to some additional dependence on  $m$ .  $Q_p$  provides better sizes for the smaller values of  $n$  but similar for the larger ones. Here the choice of  $p$  seems to be quite important, with the number of rejections increasing with  $p$  and also with  $\alpha_1$ ,  $\beta_1$ , although it decreases with  $b$ .

For the power analysis we only report the simulations with  $n = 200$ , being the picture for other sample sizes similar, although perhaps for  $n = 100$ , the results show some instability due perhaps to the oversize of the tests for some parameter combinations. For AR(2) alternatives,

**Table 1.** Size of 5% tests,  $n = 100$ .

$\alpha_1$	$b = 0$						$b = 0.4$						$b = 0.8$					
	$C_n^0$	$C_n^{15}$	$C_n^{18}$	$Q_{10}$	$Q_{15}$		$C_n^0$	$C_n^{15}$	$C_n^{18}$	$Q_{10}$	$Q_{15}$		$C_n^0$	$C_n^{15}$	$C_n^{18}$	$Q_{10}$	$Q_{15}$	
			$\beta_1 = 0.2$			$\beta_1 = 0.2$			$\beta_1 = 0.2$			$\beta_1 = 0.2$			$\beta_1 = 0.5$			$\beta_1 = 0.5$
0.2	3.39	2.65	2.16	4.79	5.47		3.71	2.88	2.23	4.80	5.50		4.11	3.24	2.15	4.73	5.46	
0.5	5.65	3.50	2.23	5.04	5.77		6.41	3.96	2.34	5.13	5.82		7.22	4.05	2.21	4.95	5.66	
0.8	9.38	4.44	2.70	5.85	6.46		10.36	3.99	2.52	5.66	6.38		10.51	3.03	1.81	4.58	5.39	
0.2	4.40	2.68	2.22	5.20	5.87		5.23	3.22	2.25	5.25	5.99		6.34	4.02	2.35	5.07	5.79	
0.5	6.75	3.54	2.31	5.63	6.30		8.08	4.28	2.53	5.63	6.28		8.87	3.62	1.93	4.75	5.42	
0.8	10.06	4.27	2.70	6.47	7.10		11.02	3.40	2.45	6.11	6.83		11.10	2.72	1.69	4.54	5.30	
0.2	6.24	2.75	2.22	6.09	6.72		7.51	3.94	2.55	6.12	6.72		7.93	3.69	1.93	4.80	5.51	
0.5	8.28	3.52	2.39	6.63	7.22		9.46	3.99	2.65	6.30	6.93		9.55	3.23	1.76	4.52	5.23	
0.8	10.71	3.71	2.65	7.30	7.90		11.15	2.69	2.17	6.52	7.21		11.65	2.58	1.63	4.54	5.34	

**Table 2.** Size of 5% tests,  $n = 200$ .

$\alpha_1$	$b = 0$						$b = 0.4$						$b = 0.8$							
	$C_n^0$	$C_n^{29}$	$C_n^{35}$	$Q_{10}$	$Q_{15}$	$C_n^0$	$C_n^{29}$	$C_n^{35}$	$Q_{10}$	$Q_{15}$	$C_n^0$	$C_n^{29}$	$C_n^{35}$	$Q_{10}$	$Q_{15}$	$C_n^0$	$C_n^{29}$	$C_n^{35}$	$Q_{10}$	$Q_{15}$
	$\beta_1 = 0.2$						$\beta_1 = 0.2$						$\beta_1 = 0.2$							
0.2	2.20	3.86	3.46	4.70	5.20	2.41	4.12	3.44	4.68	5.20	2.71	4.59	3.37	4.57	5.10	2.71	4.59	3.37	4.57	5.10
0.5	3.71	4.83	3.47	4.86	5.37	4.20	5.36	3.53	4.88	5.35	4.65	6.01	3.53	4.68	5.17	4.65	6.01	3.53	4.68	5.17
0.8	6.40	6.59	4.13	5.37	5.77	7.07	6.10	4.00	5.17	5.60	6.67	5.02	3.13	4.19	4.69	6.67	5.02	3.13	4.19	4.69
	$\beta_1 = 0.5$						$\beta_1 = 0.5$						$\beta_1 = 0.5$							
0.2	2.79	3.89	3.48	5.06	5.57	3.38	4.51	3.53	5.10	5.57	4.19	5.54	3.59	4.86	5.32	4.19	5.54	3.59	4.86	5.32
0.5	4.29	4.87	3.48	5.33	5.83	5.29	6.19	3.84	5.30	5.70	5.54	5.67	3.13	4.43	4.93	5.54	5.67	3.13	4.43	4.93
0.8	6.71	6.57	4.23	5.95	6.33	7.41	5.10	3.84	5.47	5.95	6.75	4.64	3.05	4.13	4.70	6.75	4.64	3.05	4.13	4.70
	$\beta_1 = 1.0$						$\beta_1 = 1.0$						$\beta_1 = 1.0$							
0.2	4.03	3.97	3.48	5.83	6.29	4.99	5.49	3.83	5.75	6.17	4.91	5.40	3.00	4.52	5.00	4.91	5.40	3.00	4.52	5.00
0.5	5.46	4.99	3.59	6.22	6.60	6.29	5.91	4.08	5.85	6.20	5.50	5.27	2.92	6.89	7.76	5.50	5.27	2.92	6.89	7.76
0.8	7.22	6.01	4.20	6.79	6.96	7.41	3.99	3.39	5.88	6.27	6.88	4.48	2.98	4.15	4.73	6.88	4.48	2.98	4.15	4.73

**Table 3.** Size of 5% tests,  $n = 400$ .

$\alpha_1$	$b = 0$						$b = 0.4$						$b = 0.8$					
	$C_n^0$	$C_n^{54}$	$C_n^{65}$	$Q_{15}$	$Q_{20}$		$C_n^0$	$C_n^{54}$	$C_n^{65}$	$Q_{15}$	$Q_{20}$		$C_n^0$	$C_n^{54}$	$C_n^{65}$	$Q_{15}$	$Q_{20}$	
			$\beta_1 = 0.2$						$\beta_1 = 0.2$							$\beta_1 = 0.2$		
0.2	1.83	4.54	4.17	4.90	5.19		2.01	4.87	4.24	4.88	5.17		2.16	5.69	4.32	4.78	5.05	
0.5	2.91	5.54	4.23	5.00	5.25		3.34	6.23	4.39	5.00	5.23		3.74	7.68	4.81	4.83	5.11	
0.8	4.95	8.32	5.20	5.30	5.51		5.29	7.99	5.33	4.71	5.14		4.60	6.84	4.61	4.31	4.69	
			$\beta_1 = 0.5$						$\beta_1 = 0.5$						$\beta_1 = 0.5$			
0.2	2.21	4.56	4.17	5.21	5.48		2.69	3.99	4.16	5.19	5.46		3.33	6.78	4.44	5.02	5.22	
0.5	3.41	5.53	4.20	5.39	5.66		4.13	7.68	5.02	5.30	5.59		4.08	7.36	4.59	4.55	4.89	
0.8	5.08	8.25	5.31	5.83	6.02		5.53	6.45	4.94	5.46	6.58		4.56	6.34	4.51	4.28	4.66	
			$\beta_1 = 1.0$						$\beta_1 = 1.0$						$\beta_1 = 1.0$			
0.2	3.15	4.61	4.22	5.83	6.04		3.99	6.31	4.62	5.71	5.95		3.72	6.85	4.11	4.70	5.02	
0.5	4.27	5.61	4.30	6.09	6.25		4.86	7.38	5.17	5.71	5.96		3.86	6.91	4.41	4.40	4.75	
0.8	5.58	7.87	5.39	6.51	6.57		5.57	5.07	4.45	5.82	5.98		4.54	6.13	4.38	4.32	4.66	

**Table 4.** Power of 5% tests: AR(2) alternative,  $n = 200$ .

$\alpha_1$	$b = 0$						$b = 0.4$						$b = 0.8$					
	$C_n^0$	$C_n^{29}$	$C_n^{35}$	$Q_{10}$	$Q_{15}$		$C_n^0$	$C_n^{29}$	$C_n^{35}$	$Q_{10}$	$Q_{15}$		$C_n^0$	$C_n^{29}$	$C_n^{35}$	$Q_{10}$	$Q_{15}$	
			$\beta_1 = 0.2$						$\beta_1 = 0.2$							$\beta_1 = 0.2$		
0.2	57.38	64.79	55.01	94.48	91.03		58.34	62.55	51.11	94.65	91.25		60.73	90.54	80.66	94.57	91.34	
0.5	77.79	69.48	58.47	95.38	92.39		79.36	76.06	68.41	93.67	89.98		66.83	62.05	57.51	79.61	73.14	
0.8	99.30	36.56	31.19	93.53	89.47		95.28	25.01	22.99	85.32	79.40		51.79	28.27	21.45	51.25	45.22	
			$\beta_1 = 0.5$						$\beta_1 = 0.5$						$\beta_1 = 0.5$			
0.2	61.35	48.40	38.67	95.98	93.08		68.04	63.81	50.57	95.55	92.61		67.56	87.35	84.47	87.43	82.35	
0.5	84.89	57.04	46.93	96.61	94.06		86.90	69.50	61.77	91.88	87.52		58.12	22.78	18.31	53.39	47.29	
0.8	99.24	31.15	26.84	93.55	89.47		94.26	20.88	19.22	79.78	73.05		37.04	15.60	11.23	28.79	25.53	
			$\beta_1 = 1.0$						$\beta_1 = 1.0$						$\beta_1 = 1.0$			
0.2	73.05	28.29	21.39	97.90	96.11		81.58	58.94	45.97	94.93	91.53		59.38	28.50	25.46	52.09	45.70	
0.5	93.70	34.58	27.59	97.75	95.76		94.76	46.97	44.71	87.39	81.65		33.61	5.94	4.79	20.86	18.46	
0.8	98.98	22.10	20.25	93.75	89.69		91.80	19.62	17.19	70.89	63.35		17.45	5.15	3.89	12.70	12.19	



**Table 5.** Power of 5% tests: MA(1) alternative,  $n = 200$ .

$\alpha_1$	$b = 0$						$b = 0.4$						$b = 0.8$							
	$C_n^0$	$C_n^{29}$	$C_n^{35}$	$Q_{10}$	$Q_{15}$	$C_n^0$	$C_n^{29}$	$C_n^{35}$	$Q_{10}$	$Q_{15}$	$C_n^0$	$C_n^{29}$	$C_n^{35}$	$Q_{10}$	$Q_{15}$	$C_n^0$	$C_n^{29}$	$C_n^{35}$	$Q_{10}$	$Q_{15}$
	$\beta_1 = 0.2$						$\beta_1 = 0.2$						$\beta_1 = 0.2$							
0.2	21.82	12.59	12.90	30.99	27.21	31.08	13.89	14.92	38.29	33.28	34.95	10.58	13.33	43.22	37.55	34.95	10.58	13.33	43.22	37.55
0.5	66.50	19.65	17.12	60.72	53.10	74.56	14.97	11.21	64.91	56.96	55.70	5.23	2.35	47.95	41.27	55.70	5.23	2.35	47.95	41.27
0.8	95.45	23.68	17.12	83.19	76.34	93.93	13.21	7.52	79.37	71.86	53.61	11.34	5.20	37.45	32.27	53.61	11.34	5.20	37.45	32.27
	$\beta_1 = 0.5$						$\beta_1 = 0.5$						$\beta_1 = 0.5$							
0.2	32.96	12.01	12.05	32.41	28.24	58.36	14.59	14.28	49.60	42.88	53.72	3.96	3.63	47.68	41.08	53.72	3.96	3.63	47.68	41.08
0.5	76.50	19.84	16.93	62.30	54.40	88.15	15.55	10.09	71.44	63.02	56.49	6.69	2.86	38.68	34.01	56.49	6.69	2.86	38.68	34.01
0.8	97.04	25.12	18.36	84.33	77.23	95.57	14.04	8.30	80.96	73.20	47.55	12.85	6.67	33.53	28.92	47.55	12.85	6.67	33.53	28.92
	$\beta_1 = 1.0$						$\beta_1 = 1.0$						$\beta_1 = 1.0$							
0.2	68.60	11.07	10.50	49.52	42.18	89.21	15.78	12.40	68.86	60.06	54.80	3.70	2.01	38.14	32.83	54.80	3.70	2.01	38.14	32.83
0.5	91.95	19.85	16.42	73.64	64.86	95.49	16.08	9.81	79.12	70.74	48.98	8.79	4.19	29.98	26.11	48.98	8.79	4.19	29.98	26.11
0.8	98.65	27.35	20.72	87.69	80.81	95.91	14.77	9.41	82.38	74.55	40.27	13.85	8.07	30.86	26.66	40.27	13.85	8.07	30.86	26.66

$C_n^0$  shows highest power for models with high  $\alpha_1$  and  $b$ , otherwise  $C_n^m$  dominates, with power decreasing with  $m$ . Tests based on  $C_n^m$  are dominated in general by  $C_n^0$ , except for the least persistent models (with lowest  $\alpha_1$  and  $\beta_1$ ) for which  $\phi$  is rather flat and can be well estimated by kernel estimates with some oversmoothing as with the choices of  $m$  we employ. In general power increases with  $\alpha_1$  and  $\beta_1$  for small  $b$ , but the reverse situation arises for large value of  $b$ . For the MA(1) alternative,  $C_n^0$  dominates in almost every case, in some situations outperforming noticeably  $Q_p$ , while  $C_n^m$  displays much inferior results for all  $m$ .

## ACKNOWLEDGMENTS

The research of the first and third author is funded by the Spanish ‘Plan Nacional de I+D+I’, reference number SEJ2007-62908.

## REFERENCES

- Bartlett, M. S. (1954). Problems de l’analyse spectral des séries temporelles stationnaires. *Publications Institut de Statistique Université de Paris III*, 119–34.
- Billingsley, P. (1968). *Convergence of Probability Measures*. New York: John Wiley.
- Brillinger, D. R. (1981). *Time Series, Data Analysis and Theory*. San Francisco: Holden-Day.
- Dahlhaus, R. (1985). On the asymptotic distribution of Bartlett’s  $U_p$ -statistic. *Journal of Time Series Analysis* 6, 213–27.
- Dahlhaus, R. and W. Wefelmeyer (1996). Asymptotically optimal estimation in misspecified time series models. *Annals of Statistics* 24, 952–74.
- Delgado, M. A., J. Hidalgo and C. Velasco (2005). Distribution free goodness-of-fit tests for linear processes. *Annals of Statistics* 33, 2568–609.
- Hong, Y. (1996). Consistent testing for serial correlation of unknown form. *Econometrica* 64, 837–64.
- Ljung, G. M. and G. E. P. Box (1978). On a measure of lack of fit in time series models. *Biometrika* 65, 297–303.
- Shorack, G. R. and J. A. Wellner (1986). *Empirical Processes with Applications to Statistics*. New York: John Wiley.
- Stute, W., S. Thies and L.-X. Zhu (1998). Model checks for regression: an innovation process approach. *Annals of Statistics* 26, 1916–34.
- Stute, W. and L.-X. Zhu (2002). Model checks for generalized linear models. *Scandinavian Journal of Statistics* 29, 535–45.

## APPENDIX: PROOFS

We first state two general lemmas.

LEMMA A.1. *Let Assumptions 2.1–2.3 hold. Set  $\hat{g}_j = \hat{g}_m(j/\tilde{n})$ . Then under  $H_0$ , as  $n \rightarrow \infty$*

$$\sup_j \|\hat{g}_j - g_j\| = o_p(1), \quad \sup_j \|\hat{\Sigma}_j - \Sigma_j\| = o_p(1). \quad (\text{A.1})$$

**Proof:** We only prove the first part of (A.1) since the proof of  $\sup_j \|\hat{\Sigma}_j - \Sigma_j\| = o_p(1)$  follows by identical steps. Because  $g_j = (1, \phi_j)'$ , we ignore the first element. By the triangle inequality, we have that the left-hand side of (A.1) is bounded by

$$\sup_j \|\hat{\phi}_j - \tilde{\phi}_j\| + \sup_j \|\mathbb{E}\tilde{\phi}_j - \phi_j\| + \sup_j \|\tilde{\phi}_j - \mathbb{E}\tilde{\phi}_j\|, \quad (\text{A.2})$$

where, using the errors  $\varepsilon_t$ ,

$$\tilde{\phi}_j := \frac{4\pi}{\bar{K}^m} \sum_{\ell=-m; \ell \neq 0}^m K_\ell \operatorname{Re} I_{W\varepsilon, j+\ell}.$$

To simplify arguments, we shall take herewith  $K(u) = 1$ , so

$$\hat{\phi}_j = \frac{4\pi}{2m} \sum_{\ell=-m; \ell \neq 0}^m \operatorname{Re} I_{W\varepsilon, j+\ell}.$$

Now

$$\begin{aligned} \sup_j \|\hat{\phi}_j - \tilde{\phi}_j\| &\leq \|\hat{\theta}_n - \theta_0\| \sup_j \left\| \frac{1}{m} \sum_{\ell=-m}^m \operatorname{Re} I_{W\mathbf{w}, k+\ell} \right\| \\ &\leq n^{1/2} m^{-1} n^{1/2} \|\hat{\theta}_n - \theta_0\| \left\| n^{-1} \sum_{\ell=1}^{\tilde{n}} \operatorname{Re} I_{W\mathbf{w}, \ell} \right\| \\ &= O_p(m^{-1} n^{1/2}) = o_p(1) \end{aligned}$$

because  $n^{1/2} \|\hat{\theta}_n - \theta_0\| = O_p(1)$  by Assumption 2.2, and  $\|n^{-1} \sum_{\ell=1}^{\tilde{n}} \operatorname{Re} I_{W\mathbf{w}, \ell}\| = O_p(1)$  by Assumption 2.3(i). The second term in (A.2) is  $O(m^2 n^{-2} + n^{-1} \log n) = o(1)$  because of Assumption 2.3(i), whereas

$$\mathbb{E} \|\tilde{\phi}_j - \mathbb{E} \tilde{\phi}_j\|^4 = \frac{1}{16m^4} \sum_{a=-m}^m \sum_{b=-m}^m \sum_{c=-m}^m \sum_{d=-m}^m \mathbb{E}[h_{j+a} h_{j+b} h_{j+c} h_{j+d}],$$

where we have considered  $h_j := \operatorname{Re} I_{W\varepsilon, j} - \mathbb{E} \operatorname{Re} I_{W\varepsilon, j}$  as scalar to simplify notation. Now

$$\begin{aligned} \mathbb{E}[h_{j+a} h_{j+b} h_{j+c} h_{j+d}] &= \mathbb{E}[h_{j+a} h_{j+b}] \mathbb{E}[h_{j+c} h_{j+d}] + \mathbb{E}[h_{j+a} h_{j+c}] \mathbb{E}[h_{j+b} h_{j+d}] \\ &\quad + \mathbb{E}[h_{j+a} h_{j,d}] \mathbb{E}[h_{j+b} h_{j+c}] + \operatorname{cum}[h_{j+a}, h_{j+b}, h_{j+c}, h_{j+d}]. \end{aligned}$$

But, for all  $a, b$ ,  $\mathbb{E}[h_{j+a} h_{j+b}] = O(n^{-1} \log^3 n + \mathcal{I}(a = b))$ , whereas, distinguishing the contribution from higher-order cumulants and second-order cumulants (see Brillinger, 1981, p. 20 and Theorem 2.6.1),

$$\begin{aligned} \operatorname{cum}[h_{j+a}, h_{j+b}, h_{j+c}, h_{j+d}] &= O(n^{-2} \log^6 n + \delta_{a,b,c,d} n^{-1} \log^2 n + \delta_{a,b,c,d}^2 n^{-1} \log^3 n + \delta_{a,b,c,d}^3) \\ &= O(m^{-1} n^{-1} \log^2 n + m^{-3}), \end{aligned}$$

where  $\delta_{a,b,c,d}$  indicates a restriction among the indices  $a, b, c, d$ . Thus,

$$\mathbb{E} \|\tilde{\phi}_j - \mathbb{E} \tilde{\phi}_j\|^4 = O(n^{-2} \log^6 n + m^{-2}).$$

From here, we can conclude easily that  $\sup_j \|\tilde{\phi}_j - \mathbb{E} \tilde{\phi}_j\| = o_p(1)$  using that

$$\Pr \left( \sup_j \|\tilde{\phi}_j - \mathbb{E} \tilde{\phi}_j\| > c \right) \leq \sum_{j=1}^{\tilde{n}} \Pr(\|\tilde{\phi}_j - \mathbb{E} \tilde{\phi}_j\| > c) \leq c^{-4} \sum_{j=1}^{\tilde{n}} \mathbb{E} \|\tilde{\phi}_j - \mathbb{E} \tilde{\phi}_j\|^4$$

and that  $m^{-2} n = o(1)$ . □

LEMMA A.2. *Under the assumptions of Theorem 3.2,*

$$\sup_{\omega \in (0, \pi)} \left\| \frac{1}{\tilde{n}^{1/2}} \sum_{j=1}^{[\tilde{n}\omega]} (\hat{\phi}_j - \phi_j) m_j \right\| = o_p(1). \quad (\text{A.3})$$

**Proof:** To simplify arguments we will assume that  $K(u) = \mathcal{I}(|u| \leq 1)$ . Because  $\mathbb{E}\tilde{\phi}_j - 4\pi \operatorname{Re} f_{\mathbf{w}_{\varepsilon,j}}$  is  $O(n^{-2}m^2)$  uniformly in  $j$ , it is easy to show that

$$\sup_{\omega \in (0, \pi)} \left\| \frac{1}{\tilde{n}^{1/2}} \sum_{j=1}^{\lfloor \tilde{n}\omega \rfloor} (\mathbb{E}\tilde{\phi}_j - 4\pi \operatorname{Re} f_{\mathbf{w}_{\varepsilon,j}}) m_j \right\| = o_p(1),$$

assuming finite second derivatives of  $f_{\mathbf{w}_{\varepsilon,j}}$  in Assumption 2.3', and that

$$\sup_{\omega \in (0, \pi)} \left\| \frac{1}{\tilde{n}^{1/2}} \sum_{j=1}^{\lfloor \tilde{n}\omega \rfloor} (\hat{\phi}_j - \tilde{\phi}_j) m_j \right\| = o_p(1),$$

using Assumptions 2.2 and 2.3 as in Lemma A.1. The lemma now follows by Propositions A.1 and A.2.

**PROPOSITION A.1.** *Under the assumptions of Theorem 3.2, for all  $\omega \in [0, \pi]$ ,*

$$\frac{1}{\tilde{n}^{1/2}} \sum_{j=1}^{\lfloor \tilde{n}\omega \rfloor} (\tilde{\phi}_j - \mathbb{E}\tilde{\phi}_j) m_j = o_p(1). \quad (\text{A.4})$$

**Proof:** Writing  $\tilde{\phi}_j - \mathbb{E}\tilde{\phi}_j = \frac{1}{2m} \sum_{\ell=-m; \ell \neq 0}^m h_{j+\ell}$ , by Abel summation by parts, we obtain that the left-hand side of (A.4) is

$$\begin{aligned} & \frac{1}{2\tilde{n}^{1/2}m} \sum_{j=1}^{\lfloor \tilde{n}\omega \rfloor} (h_{j-m} - h_{j+1+m} - h_j + h_{j+1}) \sum_{\ell=1}^j m_\ell \\ &= \frac{1}{2\tilde{n}^{1/2}m} \sum_{j=1}^{\lfloor \tilde{n}\omega \rfloor} (h_{j-m} - h_{j+1+m} - h_j + h_{j+1})(m_{j-m} + m_j) \end{aligned} \quad (\text{A.5})$$

$$+ \frac{1}{2\tilde{n}^{1/2}m} \sum_{j=1}^{\lfloor \tilde{n}\omega \rfloor} (h_{j-m} - h_{j+1+m} - h_j + h_{j+1}) \sum_{\ell=1; \ell \neq j-m}^{j-1} m_\ell. \quad (\text{A.6})$$

Equation (A.5) is  $o_p(1)$  because the Cauchy–Schwarz inequality implies that

$$\begin{aligned} & (\mathbb{E}|h_{j-m} - h_{j+1+m} - h_j + h_{j+1}| |m_{j-m} + m_j|)^2 \\ & \leq \mathbb{E}|h_{j-m} - h_{j+1+m} - h_j + h_{j+1}|^2 \mathbb{E}|m_{j-m} + m_j|^2 < D, \end{aligned}$$

where, in what follows,  $D$  denotes a finite and positive constant.

It is worth mentioning that this is the best rate we can obtain under our general assumptions, because lack of (strong) exogeneity implies that  $\mathbb{E}(h_j m_j) \neq 0$ .

Next, we examine (A.6). We employ that  $h_\bullet$  and  $m_\bullet$  do not have subindices in common. So, although the expectation is not zero, unless the fourth cumulant is, this is  $O(n^{-1} \log^3 n)$  at most. The expectation of (A.6) is  $O(m^{-1} n^{-1/2} \log^3 n)$  because

$$\mathbb{E} \left( (h_{j-m} - h_{j+1+m} - h_j + h_{j+1}) \sum_{\ell=1; \ell \neq j-m}^{j-1} m_\ell \right) = O(n^{-1} \log^3 n).$$

Note that under Gaussianity the expectation would have been exactly zero. Next, we examine the second moment of (A.6). By the Cauchy–Schwarz inequality, it suffices to examine the second moment of each of

the following four terms:

$$\begin{aligned}
& \frac{1}{2n^{1/2}m} \sum_{j=1}^{[\bar{n}\omega]} h_{j-m} \sum_{\ell=1; \ell \neq j-m}^{j-1} H_{\ell} - \frac{1}{2n^{1/2}m} \sum_{j=1}^{[\bar{n}\omega]} h_{j+1+m} \sum_{\ell=1; \ell \neq j-m}^{j-1} m_{\ell} \\
& - \frac{1}{2n^{1/2}m} \sum_{j=1}^{[\bar{n}\omega]} h_j \sum_{\ell=1; \ell \neq j-m}^{j-1} H_{\ell} + \frac{1}{2n^{1/2}m} \sum_{j=1}^{[\bar{n}\omega]} h_{j+1} \sum_{\ell=1; \ell \neq j-m}^{j-1} m_{\ell}.
\end{aligned} \tag{A.7}$$

We will study the contribution due to the first term, the other three terms are similarly handled. The second moment of the first term of (A.7) is proportional to

$$\begin{aligned}
& \frac{1}{\bar{n}m^2} \sum_{j_1=1}^{[\bar{n}\omega]} \sum_{j_2=1}^{j_1} \sum_{\ell_1=1; \ell_1 \neq j_1-m}^{j_1-1} \sum_{\ell_2=1; \ell_2 \neq j_2-m}^{j_2-1} \mathbb{E}(h_{j_1-m} h_{j_2-m} m_{\ell_1} m_{\ell_2}) \\
& = \frac{1}{\bar{n}m^2} \sum_{j_1=1}^{[\bar{n}\omega]} \sum_{j_2=1}^{j_1} \sum_{\ell_1=1; \ell_1 \neq j_1-m}^{j_1-1} \sum_{\ell_2=1; \ell_2 \neq j_2-m}^{j_2-1} \mathbb{E}(h_{j_1-m} h_{j_2-m}) \mathbb{E}(m_{\ell_1} m_{\ell_2}) \\
& + \frac{1}{\bar{n}m^2} \sum_{j_1=1}^{[\bar{n}\omega]} \sum_{j_2=1}^{j_1} \sum_{\ell_1=1; \ell_1 \neq j_1-m}^{j_1-1} \sum_{\ell_2=1; \ell_2 \neq j_2-m}^{j_2-1} \mathbb{E}(h_{j_1-m} m_{\ell_1}) \mathbb{E}(h_{j_2-m} m_{\ell_2}) \\
& + \frac{1}{\bar{n}m^2} \sum_{j_1=1}^{[\bar{n}\omega]} \sum_{j_2=1}^{j_1} \sum_{\ell_1=1; \ell_1 \neq j_1-m}^{j_1-1} \sum_{\ell_2=1; \ell_2 \neq j_2-m}^{j_2-1} \mathbb{E}(h_{j_2-m} m_{\ell_1}) \mathbb{E}(h_{j_1-m} m_{\ell_2}) \\
& + \frac{1}{\bar{n}m^2} \sum_{j_1=1}^{[\bar{n}\omega]} \sum_{j_2=1}^{j_1} \sum_{\ell_1=1; \ell_1 \neq j_1-m}^{j_1-1} \sum_{\ell_2=1; \ell_2 \neq j_2-m}^{j_2-1} \text{cum}(h_{j_1-m}, h_{j_2-m}, m_{\ell_1}, m_{\ell_2}).
\end{aligned} \tag{A.8}$$

Because  $\mathbb{E}(m_{\ell_1} m_{\ell_2}) = O(n^{-1}) + \mathcal{I}(\ell_1 = \ell_2)$ ,  $\mathbb{E}(h_{j_1-m} h_{j_2-m}) = O(n^{-1}) + \mathcal{I}(j_1 = j_2)$  and  $\sum_{j=1}^{[\bar{n}\omega]} j = O([\bar{n}\omega]^2)$ , the first term on the right-hand side of (A.8) is

$$O\left(\frac{n}{m^2}\right) + \frac{1}{nm^2} \sum_{j_1=1}^{[\bar{n}\omega]} \sum_{\ell=1; \ell \neq j_1-m}^{j_1-1} D = O\left(\frac{n}{m^2}\right).$$

Similarly, the second and third terms on the right-hand side of (A.8) are  $O(m^{-2}n)$ .

Finally, the fourth term on the right-hand side of (A.8). First, observe that

$$\begin{aligned}
& \text{cum}(h_{j_1-m}, h_{j_2-m}, m_{\ell_1}, m_{\ell_2}) \\
& = \text{cum}(w_{z, j_1-m} w_{\varepsilon, j_1-m}^*, w_{z, j_2-m} w_{\varepsilon, j_2-m}^*, w_{\varepsilon, \ell_1} w_{\varepsilon, \ell_1}^*, w_{\varepsilon, \ell_2} w_{\varepsilon, \ell_2}^*) \\
& = \sum_{\nu} \prod_{r=1}^q \text{cum}(w_{a, s_1} w_{b, s_2}; (s_1, s_2) \in \nu_r)
\end{aligned}$$

with  $s_1, s_2 = j_1 - m, j_2 - m, \ell_1, \ell_2$ ,  $a$  and  $b$  are  $Z$  and  $\varepsilon$  and where the summation in  $\nu$  is over all indecomposable partitions  $\nu = \nu_1 \cup \dots \cup \nu_q$ ,  $q = 1, \dots, 4$ , of the table

$w_{z, j_1-m}$	$w_{\varepsilon, j_1-m}^*$
$w_{z, j_2-m}$	$w_{\varepsilon, j_2-m}^*$
$w_{\varepsilon, \ell_1}$	$w_{\varepsilon, \ell_1}^*$
$w_{\varepsilon, \ell_2}$	$w_{\varepsilon, \ell_2}^*$

see Brillinger (1981, p. 20 and Theorem 2.6.1). So, a typical component of the fourth term on the right-hand side of (A.8) is

$$\frac{1}{nm^2} \sum_{j_1=1}^{[\tilde{n}\omega]} \sum_{j_2=1}^{j_1-1} \sum_{\ell_1=1; \ell_1 \neq j_1-m}^{j_1-1} \sum_{\ell_2=1; \ell_2 \neq j_2-m}^{j_2-1} \text{cum}(w_{a,s_1} w_{b,s_2}; (s_1, s_2) \in \nu_r) = O(n^{-1} m^{-1} \log^3 n).$$

So, we conclude that the second moment of (A.6) converges to zero, and (A.4) holds true by Markov inequality.  $\square$

PROPOSITION A.2. *Under the assumptions of Theorem 3.2, the process*

$$\mathcal{X}_n(\omega) = \frac{1}{\tilde{n}^{1/2}} \sum_{j=1}^{[\tilde{n}\omega]} (\tilde{\phi}_j - \mathbb{E}\tilde{\phi}_j) m_j, \quad \omega \in [0, 1]$$

is tight.

**Proof:** Proceeding as with Proposition A.1,  $\mathcal{X}_n(\omega)$  can be written as

$$\begin{aligned} \mathcal{X}_{n1}(\omega) + \mathcal{X}_{n2}(\omega) &:= \frac{1}{2\tilde{n}^{1/2}m} \sum_{j=1}^{[\tilde{n}\omega]} (h_{j-m} - h_{j+1+m} - h_j + h_{j+1})(m_{j-m} + m_j) \\ &\quad + \frac{1}{2\tilde{n}^{1/2}m} \sum_{j=1}^{[\tilde{n}\omega]} (h_{j-m} - h_{j+1+m} - h_j + h_{j+1}) \sum_{\ell=1; \ell \neq j-m}^{j-1} m_\ell. \end{aligned}$$

Following Billingsley (1968, Theorem 15.6), a sufficient condition for the tightness of  $\mathcal{X}_n(\omega)$  is

$$\mathbb{E}|\mathcal{X}_{nj}(\omega_2) - \mathcal{X}_{nj}(\omega_1)|^\tau \leq D(\omega_2 - \omega_1)^{1+\delta}, \quad j = 1, 2, \quad (\text{A.9})$$

where  $\tau, \delta > 0$ ,  $\omega_2 > \omega_1$ , and, where without loss of generality, we can assume that  $\tilde{n}^{-1} \leq \omega_2 - \omega_1$ .

We begin with  $\mathcal{X}_{n1}(\omega)$ . By definition,

$$\mathcal{X}_{n1}(\omega_2) - \mathcal{X}_{n1}(\omega_1) = \frac{1}{2\tilde{n}^{1/2}m} \sum_{j=[\tilde{n}\omega_1]+1}^{[\tilde{n}\omega_2]} (h_{j-m} - h_{j+1+m} - h_j + h_{j+1})(m_{j-m} + m_j).$$

So, by the triangle inequality and proceeding as with the estimation of the second moment of (A.5), we have that  $\mathbb{E}|\mathcal{X}_{n1}(\omega_2) - \mathcal{X}_{n1}(\omega_1)|$  is bounded by

$$D \frac{[\tilde{n}\omega_2] - [\tilde{n}\omega_1]}{m\tilde{n}^{1/2}} \leq D(\omega_2 - \omega_1) \frac{n^{1/2}}{m} \leq D(\omega_2 - \omega_1)^{1+\delta}$$

because by Assumption 3.2(ii),  $m^{-1}n^{1/2} = o(n^{-2\delta})$  for some  $\delta > 0$ .

To complete the proof, we need to show (A.9) for  $\mathcal{X}_{n2}(\omega)$ . We will only examine the contribution due to the first term of (A.7) into the left-hand side of (A.9), that is

$$\mathbb{E} \left\| \frac{1}{2\tilde{n}^{1/2}m} \sum_{j=[\tilde{n}\omega_1]+1}^{[\tilde{n}\omega_2]} h_{j-m} \sum_{\ell=1; \ell \neq j-m}^{j-1} m_\ell \right\|^\tau.$$

Choosing  $\tau = 2$ , we have that the last displayed expression is bounded by

$$\begin{aligned}
& \frac{1}{\tilde{n}m^2} \sum_{j_1=\lceil \tilde{n}\omega_1 \rceil + 1}^{\lceil \tilde{n}\omega_2 \rceil} \sum_{j_2=\lceil \tilde{n}\omega_1 \rceil + 1}^{j_1} \sum_{\ell_1=1; \ell_1 \neq j_1 - m}^{j_1 - 1} \sum_{\ell_2=1; \ell_2 \neq j_2 - m}^{j_2 - 1} \left| \mathbb{E} (h_{j_1 - m} h_{j_2 - m}) \mathbb{E} (m_{\ell_1} m_{\ell_2}) \right| \\
& + \frac{1}{\tilde{n}m^2} \sum_{j_1=\lceil \tilde{n}\omega_1 \rceil + 1}^{\lceil \tilde{n}\omega_2 \rceil} \sum_{j_2=\lceil \tilde{n}\omega_1 \rceil + 1}^{j_1} \sum_{\ell_1=1; \ell_1 \neq j_1 - m}^{j_1 - 1} \sum_{\ell_2=1; \ell_2 \neq j_2 - m}^{j_2 - 1} \left| \mathbb{E} (h_{j_1 - m} m_{\ell_1}) \mathbb{E} (h_{j_2 - m} m_{\ell_2}) \right| \\
& + \frac{1}{\tilde{n}m^2} \sum_{j_1=\lceil \tilde{n}\omega_1 \rceil + 1}^{\lceil \tilde{n}\omega_2 \rceil} \sum_{j_2=\lceil \tilde{n}\omega_1 \rceil + 1}^{j_1} \sum_{\ell_1=1; \ell_1 \neq j_1 - m}^{j_1 - 1} \sum_{\ell_2=1; \ell_2 \neq j_2 - m}^{j_2 - 1} \left| \mathbb{E} (h_{j_2 - m} m_{\ell_1}) \mathbb{E} (h_{j_1 - m} m_{\ell_2}) \right| \\
& + \frac{1}{\tilde{n}m^2} \sum_{j_1=\lceil \tilde{n}\omega_1 \rceil + 1}^{\lceil \tilde{n}\omega_2 \rceil} \sum_{j_2=\lceil \tilde{n}\omega_1 \rceil + 1}^{j_1} \sum_{\ell_1=1; \ell_1 \neq j_1 - m}^{j_1 - 1} \sum_{\ell_2=1; \ell_2 \neq j_2 - m}^{j_2 - 1} \left| \text{cum} (h_{j_1 - m}, h_{j_2 - m}, m_{\ell_1}, m_{\ell_2}) \right|.
\end{aligned}$$

However, the last expression is bounded by  $D(\omega_2 - \omega_1)^{1+\delta}$  because proceeding as with the proof of (A.8), they are bounded by

$$D \frac{[\tilde{n}\omega_2]^2 - [\tilde{n}\omega_1]^2}{n^3 m^2} \log^3 n \leq D(\omega_2 - \omega_1) \frac{n}{m^2} \log^3 n \leq D(\omega_2 - \omega_1)^{1+2\delta},$$

as  $\tilde{n}^{-1} \leq \omega_2 - \omega_1$ . This completes the proof.  $\square$

LEMMA A.3. Let  $\xi(u) : [0, 1] \rightarrow \mathbb{R}^{p+q+1}$  be continuous. Assuming Assumption 2.1, we have that in  $\times_1^{p+q+1} \mathbb{D}[0, 1]$ ,

$$\left\{ \int_0^\omega \xi(u) \hat{T}_n(du) : \omega \in [0, 1] \right\} \text{ converges in distribution to } \left\{ \int_0^\omega \xi(u) dB(u) : \omega \in [0, 1] \right\}.$$

**Proof:** The proof is much simpler than that of Lemma 2 in DHV, so it is omitted.  $\square$

**Proof of Proposition 2.1:** The proof proceeds as that of Lemma 7 in DHV, and so it is omitted.  $\square$

**Proof of Proposition 2.2:** First it can be shown that  $\hat{F}_n(\pi) \rightarrow_p 1$  by Assumption 2.1 under  $H_0$ . So, we can write

$$\hat{F}_n(\omega\pi; \hat{\theta}_n) = \hat{F}_n(\omega\pi) - (\hat{\theta}_n - \theta_0)' \frac{4\pi}{\tilde{n}} \sum_{j=1}^{\lceil \omega\tilde{n} \rceil} \text{Re} I_{W\varepsilon, j} + (\hat{\theta}_n - \theta_0)' \frac{4\pi}{\tilde{n}} \sum_{j=1}^{\lceil \omega\tilde{n} \rceil} \text{Re} I_{Ww, j}(\hat{\theta}_n - \theta_0),$$

where

$$\sup_{\omega \in [0, 1]} \left\| \frac{4\pi}{\tilde{n}} \sum_{j=1}^{\lceil \omega\tilde{n} \rceil} \text{Re} I_{Ww, j} \right\| \leq \left\| \frac{4\pi}{\tilde{n}} \sum_{j=1}^{\tilde{n}} \text{Re} I_{Ww, j} \right\| = O_p(1) \quad (\text{A.10})$$

because of Assumption 2.3(ii). Then  $\hat{F}_n(\pi; \hat{\theta}_n) \rightarrow_p 1$  by Assumption 2.2 and because as we now show

$$A_n(\omega) := \frac{4\pi}{\tilde{n}} \sum_{j=1}^{\lceil \omega\tilde{n} \rceil} \text{Re} I_{W\varepsilon, j} = \Phi(\omega) + o_p(1) \quad (\text{A.11})$$

uniformly in  $\omega$ , and  $\Phi(1) = 0$ . By Assumption 2.3(i) we obtain that  $\mathbb{E}\hat{F}_n(\omega) = \Phi(\omega) + o(1)$  uniformly in  $\omega$  and by Assumption 2.3(iii),  $\hat{\Phi}_n(\omega) - \Phi(\omega) = o_p(1)$  for each  $\omega$ . Then we have to check the tightness of  $\bar{A}_n(\omega) := A_n(\omega) - \mathbb{E}[A_n(\omega)]$ .

Following Billingsley (1968, Theorem 15.6), a sufficient condition is that, for some  $\delta > 0$ ,  $0 \leq \omega_1 < \omega_2 \leq 1$ ,

$$\mathbb{E}|\bar{A}_n(\omega_2) - \bar{A}_n(\omega_1)|^2 = \mathbb{E} \left| \frac{2}{\bar{n}} \sum_{j=1+\lceil \omega_1 \bar{n} \rceil}^{\lceil \omega_2 \bar{n} \rceil} h_j \right|^2 \leq D(\omega_2 - \omega_1)^{1+\delta}. \quad (\text{A.12})$$

Without loss of generality, we consider only  $\bar{n}^{-1} \leq (\omega_2 - \omega_1)$ . Then using Assumption 2.1 and Assumption 2.3(i)–(iii), the left-hand side of (A.10) is bounded by

$$\mathbb{E}|\bar{A}_n(\omega_2) - \bar{A}_n(\omega_1)|^2 \leq D\bar{n}^{-1} \int_{\omega_1}^{\omega_2} d\lambda + D\bar{n}^{-1} \log^3 n \left( \int_{\omega_1}^{\omega_2} d\lambda \right)^2 \leq D(\omega_2 - \omega_1)^2.$$

Now (A.11) follows from (A.10) and Assumption 2.2, while (2.13) follows from (A.11) and Assumption 2.2.  $\square$

**Proof of Theorem 3.1:** Using the arguments in the proofs of Theorems 3 and 1 in DHV, we only need to consider convergence in intervals  $[0, \omega_0]$ , for any  $\omega_0 < 1$ . Since it is trivially satisfied that  $\sup_{\omega \in [0, \omega_0]} \|\tilde{\mathcal{L}}G(\omega)\| = 0$ , the theorem is a consequence of

$$\sup_{\omega \in [0, \omega_0]} |\tilde{\mathcal{L}}(\hat{T}_n(\omega; \hat{\theta}_n) - \hat{T}_n(\omega))| = o_p(1), \quad (\text{A.13})$$

$$\tilde{\mathcal{L}}\hat{T}_n(\omega) \xrightarrow{d} B^0 \text{ in the space } \mathbb{D}[0, \omega_0]. \quad (\text{A.14})$$

By definition,  $\tilde{\mathcal{L}}(\hat{T}_n(\omega; \hat{\theta}_n) - \hat{T}_n(\omega))$  is

$$\hat{T}_n(\omega; \hat{\theta}_n) - \hat{T}_n(\omega) - \int_0^\omega g(u)' \Sigma^{-1}(u) \int_u^1 g(v)(\hat{T}_n(dv; \hat{\theta}_n) - \hat{T}_n(dv)) du. \quad (\text{A.15})$$

By Proposition 2.2, the first two terms in (A.15) are equal to  $-\bar{n}^{1/2} \Phi(\omega)'(\hat{\theta}_n - \theta_0) + o_p(1)$  uniformly in  $\omega$ , whereas the third term is

$$\bar{n}^{1/2} \int_0^\omega g(u)' \Sigma^{-1}(u) \int_u^1 g(v)g(v)' dv du (\hat{\theta}_n - \theta_0) + o_p(1) = \bar{n}^{1/2} \Phi(\omega)'(\hat{\theta}_n - \theta_0) + o_p(1),$$

which shows (A.13).

To complete the proof we need to show (A.14). Fidi's convergence follows as in Proposition 2.1 or Lemma A.3. Then, it suffices to prove tightness. Since  $\hat{T}_n(\omega)$  is tight, we only need to show the tightness condition of

$$P_n(r) := \int_0^r H(u) \Psi_n(u) du,$$



where  $H(u) := g(u)' \Sigma(u)^{-1}$  and  $\Psi_n(u) := \tilde{n}^{-1/2} \sum_{j=1+\lceil \tilde{n}u \rceil}^{\tilde{n}} g_j m_j$ . Because by Lemma A.3,  $\sup_{u \in [0, \omega_0]} \|\Psi_n(u)\| = O_p(1)$  and  $\mathbb{E} \|\Psi_n(u)\|^2 < D$ ,

$$\begin{aligned} \mathbb{E} |P_n(r) - P_n(s)|^2 &= \int_s^r \int_s^r H(u_1) H(u_2) \frac{1}{\tilde{n}} \sum_{j=1+\lceil \tilde{n}u_1 \rceil}^{\tilde{n}} \sum_{k=1+\lceil \tilde{n}u_2 \rceil}^{\tilde{n}} g_j g_k \mathbb{E}(m_j m_k) du_1 du_2 \\ &\leq D \int_s^r \int_s^r \|H(u_1)\| \|H(u_2)\| du_1 du_2 \\ &= D |L(r) - L(s)|^2, \end{aligned}$$

where  $L(\cdot) = \int_0^\cdot \|H(u)\| du$  is a monotonic continuous and non-decreasing function.  $\square$

**Proof of Theorem 3.2:** Setting  $\hat{\phi}_j = \hat{\phi}_m(j/\tilde{n})$ ,  $\bar{\mathcal{L}}_n \hat{T}_n(\omega; \hat{\theta}_n)$  is, up to terms  $o_p(1)$  uniformly in  $\omega$ ,

$$\frac{1}{\tilde{n}^{1/2}} \sum_{j=1}^{\lceil \tilde{n}\omega \rceil} m_j - \frac{1}{\tilde{n}} \sum_{j=1}^{\lceil \tilde{n}\omega \rceil} \left( \frac{1}{\hat{\phi}_j} \right)' \hat{\Sigma}_{j+1}^{-1} \frac{1}{\tilde{n}^{1/2}} \sum_{\ell=j+1}^{\tilde{n}} \left( \frac{1}{\hat{\phi}_\ell} \right) m_\ell \quad (\text{A.16})$$

$$+ \left( \frac{1}{\tilde{n}} \sum_{j=1}^{\lceil \tilde{n}\omega \rceil} \phi'_j - \frac{1}{\tilde{n}} \sum_{j=1}^{\lceil \tilde{n}\omega \rceil} \left( \frac{1}{\hat{\phi}_j} \right)' \hat{\Sigma}_{j+1}^{-1} \frac{1}{\tilde{n}} \sum_{\ell=j+1}^{\tilde{n}} \left( \frac{1}{\hat{\phi}_\ell} \right) \phi'_\ell \right) \tilde{n}^{1/2} (\hat{\theta}_n - \theta_0), \quad (\text{A.17})$$

using Proposition 2.2.

Since  $\Sigma_j$  is assumed non-singular for all  $j = 1, \dots, \tilde{n}$ , using Lemma A.1 and Assumption 2.2 we obtain that (A.17) is  $o_p(1)$ , which is what it is required to conclude that the asymptotic behaviour of  $\bar{\mathcal{L}}_n \hat{T}_n(\omega)$  is given by that of (A.16).

We now show the weak convergence of (A.16) with  $\hat{\phi}_\ell$  replaced by  $\phi_\ell$ . In Lemma A.2 we show that the difference is negligible.

First, the expectation is clearly zero because  $\mathbb{E}(I_{\varepsilon\varepsilon, j} - 1) = 0$ ,  $j = 1, \dots, \tilde{n}$ . Next, we study the covariance structure. Let  $\omega_1 \leq \omega_2$ . Our aim is to show that

$$\mathbb{E}(a_n(\omega_1) a_n(\omega_2)) \xrightarrow{n \rightarrow \infty} \omega_1, \quad (\text{A.18})$$

where

$$\begin{aligned} a_n(\omega) &= \frac{1}{\tilde{n}} \sum_{j=1}^{\lceil \tilde{n}\omega \rceil} m_j - \frac{1}{\tilde{n}} \sum_{j=1}^{\lceil \tilde{n}\omega \rceil} \left( \frac{1}{\phi_j} \right)' \Sigma_{j+1}^{-1} \frac{1}{\tilde{n}} \sum_{\ell=j+1}^{\tilde{n}} \left( \frac{1}{\phi_\ell} \right) m_\ell \\ &:= a_{n1}(\omega) - a_{n2}(\omega), \end{aligned}$$

since (A.18) implies that if  $a_n(\omega)$  converges to a Gaussian process, this would be the standard Brownian motion.

Because  $\mathbb{E}(a_{n1}(\omega_1) a_{n1}(\omega_2)) = \omega_1$ , we have that (A.18) holds true if

$$\mathbb{E}(a_{n2}(\omega_1) a_{n2}(\omega_2)) = \mathbb{E}(a_{n1}(\omega_1) a_{n2}(\omega_2)) + \mathbb{E}(a_{n1}(\omega_2) a_{n2}(\omega_1)). \quad (\text{A.19})$$

First, it is easy to check that the right-hand side of (A.19) is

$$\frac{1}{\tilde{n}} \left\{ \sum_{j_1=1}^{\lceil \tilde{n}\omega_1 \rceil} \sum_{j_2=1}^{\lceil \tilde{n}\omega_2 \rceil} + \sum_{j_1=1}^{\lceil \tilde{n}\omega_2 \rceil} \sum_{j_2=1}^{\lceil \tilde{n}\omega_1 \rceil} \right\} \left( \frac{1}{\phi_{j_2}} \right)' \Sigma_{j_2}^{-1} \left( \frac{1}{\phi_{j_1}} \right).$$

Next, we examine the left-hand side of (A.19), which is

$$\begin{aligned} & \frac{1}{\tilde{n}} \sum_{j_1=1}^{[\tilde{n}\omega_1]} \sum_{j_2=1}^{[\tilde{n}\omega_2]} \left( \begin{array}{c} 1 \\ \phi_{j_1} \end{array} \right)' \Sigma_{j_1}^{-1} \frac{1}{\tilde{n}} \sum_{\ell_1=1+j_1} \sum_{\ell_2=1+j_2} \left\{ \left( \begin{array}{c} 1 \\ \phi_{\ell_1} \end{array} \right) \right. \\ & \left. \times \mathbb{E} \left( m_{\ell_1} m_{\ell_2} \right) \left( \begin{array}{c} 1 \\ \phi_{\ell_2} \end{array} \right)' \right\} \Sigma_{j_2}^{-1} \left( \begin{array}{c} 1 \\ \phi_{j_2} \end{array} \right), \end{aligned}$$

showing (A.19).

Since the fids of (A.16) converge to those of a Brownian motion, we only need to examine the tightness of  $a_{n_2}(\omega)$ , as it is already known that  $a_{n_1}(\omega)$  is tight. But we have that  $a_{n_2}(\omega_2) - a_{n_2}(\omega_1)$  is

$$\begin{aligned} & \left\{ \frac{1}{\tilde{n}} \sum_{j=[\tilde{n}\omega_1]+1}^{[\tilde{n}\omega_2]} \left( \begin{array}{c} 1 \\ \phi_j \end{array} \right)' \Sigma_{j+1}^{-1} \right\} \frac{1}{\tilde{n}^{1/2}} \sum_{\ell=[\tilde{n}\omega_2]+1}^{\tilde{n}} \left( \begin{array}{c} 1 \\ \phi_\ell \end{array} \right) m_\ell \\ & + \frac{1}{\tilde{n}^{1/2}} \sum_{\ell=[\tilde{n}\omega_1]+1}^{[\tilde{n}\omega_2]} \left( \begin{array}{c} 1 \\ \phi_\ell \end{array} \right)' \left( \frac{1}{\tilde{n}} \sum_{j=[\tilde{n}\omega_1]+1}^{\ell} \Sigma_{j+1}^{-1} \left( \begin{array}{c} 1 \\ \phi_j \end{array} \right) \right) m_\ell, \end{aligned}$$

from where it is easy to show that  $a_{n_2}(\omega)$  is tight. Observe that, for instance, the first term has again the structure  $(\zeta(\omega_2) - \zeta(\omega_1))\mathcal{Z}$ , where  $\mathcal{Z}$  is a random variable with at least finite second moments.  $\square$

**Proof of Theorem 3.3:** We first analyse an unfeasible version of the transformation  $\check{\mathcal{L}}_n, \check{\mathcal{L}}$ , assuming that we observe  $\theta_0$  and hence replacing  $\hat{g}_j$  by  $g_j = \text{Re } I_{W_{\varepsilon, j}}$  and  $\hat{m}_j$  by  $m_j, j = 1, \dots, \tilde{n}$ ,

$$\check{\mathcal{L}}\hat{T}_n(\omega; \hat{\theta}_n) = \frac{\tilde{n}^{-1/2}}{\hat{F}_n(\pi)} \sum_{j=1}^{[\omega\tilde{n}]} \left\{ m_j - \tilde{n}g'_{j+1} \left( \sum_{k=j+1}^{\tilde{n}} g_{k+2}g'_{k+2} \right)^{-1} \sum_{k=j+1}^{\tilde{n}} g_{k+2}m_{k+1} \right\}, \quad (\text{A.20})$$

and show that under the same conditions of the theorem,

$$\check{\mathcal{L}}\hat{T}_n(\omega; \hat{\theta}_n) \xrightarrow{d} B^0 \text{ in the Skorohod metric space } \mathbb{D}[0, \omega_0],$$

for any  $\omega_0 < 1$ . Then the proof of Theorem 3.3 is standard after we notice that  $\hat{\varepsilon}_t = \varepsilon_t + (\hat{\theta}_n - \theta_0)'W_t$ , Assumption 2.2 implies that  $\hat{\theta}_n - \theta_0 = O_p(n^{-1/2})$  and the arguments in the proofs of Theorems 3 and 4 in DHV.

We shall abbreviate  $g_{n,k}$  by  $g_k$  to simplify the notation. Now, because  $\hat{F}_n(\pi, \hat{\theta}_n) - \sigma_\varepsilon^2 = o_p(1)$ , recall that we can assume that  $\sigma_\varepsilon^2 = 1$  without loss of generality, we obtain that

$$\check{\mathcal{L}}\hat{T}_n(\omega) = \hat{T}_n(\omega) - \frac{1}{\tilde{n}^{1/2}} \sum_{j=1}^{[\tilde{n}\omega]} g'_{j+1} \left( \sum_{k=j+1}^{\tilde{n}} g_{k+2}g'_{k+2} \right)^{-1} \sum_{k=j+1}^{\tilde{n}} g_{k+2}m_{k+1} + o_p(1). \quad (\text{A.21})$$

So, except the  $o_p(1)$ , the right-hand side of  $\check{\mathcal{L}}\hat{T}_n(\omega)$  is

$$\frac{1}{\tilde{n}^{1/2}} \sum_{j=1}^{[\tilde{n}\omega]} \left\{ m_j - g'_{j+1} \left( \sum_{k=j+1}^{\tilde{n}} g_{k+2}g'_{k+2} \right)^{-1} \sum_{k=j+1}^{\tilde{n}} g_{k+2}m_{k+1} \right\}. \quad (\text{A.22})$$

Now, we could replace  $G_{j,n} = \frac{1}{\bar{n}} \sum_{k=j+1}^{\bar{n}} g_{k+2} g'_{k+2}$  by  $G_j = \frac{1}{\bar{n}} \sum_{k=j+1}^{\bar{n}} \mathbb{E}(g_{k+2} g'_{k+2})$ . Indeed,

$$\begin{aligned} & \frac{1}{\bar{n}^{1/2}} \sum_{j=1}^{[\bar{n}\omega]} \left\{ g'_{j+1} (G_j^{-1} - G_{j,n}^{-1}) \frac{1}{\bar{n}} \sum_{k=j+1}^{\bar{n}} g_{k+2} m_{k+1} \right\} \\ &= \frac{1}{\bar{n}^{3/2}} \sum_{j=1}^{[\bar{n}\omega]} \left\{ g'_{j+1} G_j^{-1} (G_{j,n} - G_j) G_{j,n}^{-1} \sum_{k=j+1}^{\bar{n}} g_{k+2} m_{k+1} \right\}. \end{aligned} \quad (\text{A.23})$$

However, Brillinger's (1981) Theorem 7.6.3, see also the proof of Lemma A.1, implies that uniformly in  $j$ ,

$$G_{j,n} - G_j = o_p(n^{-1/4}); \quad \sum_{k=j+1}^{\bar{n}} g_{k+2} m_{k+1} = O_p(n^{3/4}),$$

so that the right-hand side of (A.23) is  $o_p(1)$ , and hence the asymptotic distribution of (A.22) is given by that of

$$\begin{aligned} & \frac{1}{\bar{n}^{1/2}} \sum_{j=1}^{[\bar{n}\omega]} \left\{ m_j - g'_{j+1} G_j^{-1} \frac{1}{\bar{n}} \sum_{p=j+1}^{\bar{n}} g_{p+2} m_{p+1} \right\} \\ &= \frac{1}{\bar{n}^{1/2}} \sum_{j=1}^{[\bar{n}\omega]} \left\{ m_j - \mathbb{E}(g'_{j+1}) G_j^{-1} \frac{1}{\bar{n}} \sum_{k=j+1}^{\bar{n}} g_{k+2} m_{k+1} \right\} + o_p(1), \end{aligned} \quad (\text{A.24})$$

as we now show. Writing  $\tilde{g}_j = g_j - \mathbb{E}(g_j)$ , the difference between left-hand side and the first term on the right-hand side of (A.24) is

$$\frac{1}{\bar{n}^{3/2}} \sum_{j=1}^{[\bar{n}\omega]} \tilde{g}'_{j+1} G_j^{-1} \sum_{k=j+1}^{\bar{n}} g_{k+2} m_{k+1}.$$

Next, the second moment of the right-hand side of the last displayed equality is

$$\frac{1}{\bar{n}^3} \sum_{1=j \leq \ell}^{[\bar{n}\omega]} \mathbb{E} \left\{ \left( \tilde{g}'_{j+1} G_j^{-1} \sum_{k=j+1}^{\bar{n}} g_{k+2} m_{k+1} \right) \left( \tilde{g}'_{\ell+1} G_\ell^{-1} \sum_{q=\ell+1}^{\bar{n}} g_{q+2} m_{q+1} \right) \right\}. \quad (\text{A.25})$$

Now because  $\|G_j^{-1}\| < D$ , the expectation term in (A.25) is governed by

$$\begin{aligned} & \mathbb{E} \left( \sum_{q=\ell+1}^{\bar{n}} g_{q+2} m_{q+1} \sum_{k=j+1}^{\bar{n}} g_{k+2} m_{k+1} \right) \mathbb{E}(\tilde{g}'_{j+1} \tilde{g}_{\ell+1}) \\ &+ \mathbb{E} \left( \tilde{g}'_{j+1} \sum_{k=j+1}^{\bar{n}} g_{k+2} m_{k+1} \right) \mathbb{E} \left( \tilde{g}'_{\ell+1} \sum_{q=\ell+1}^{\bar{n}} g_{q+2} m_{q+1} \right) \\ &+ \mathbb{E} \left( \tilde{g}'_{\ell+1} \sum_{q=\ell+1}^{\bar{n}} g_{q+2} m_{q+1} \right) \mathbb{E} \left( \tilde{g}'_{j+1} \sum_{k=j+1}^{\bar{n}} g_{k+2} m_{k+1} \right) \\ &+ \sum_{k=j+1}^{\bar{n}} \sum_{q=\ell+1}^{\bar{n}} \text{cum}(\tilde{g}'_{j+1}, g_{k+2} m_{k+1}, \tilde{g}'_{\ell+1}, g_{q+2} m_{q+1}). \end{aligned}$$

Now, because for example  $\text{Cov}(\tilde{g}_{j+1}, \tilde{g}_{k+1}) = \mathcal{I}(j=k) + O(n^{-1})$ ,  $\text{Cov}(\tilde{g}_{j+1}, m_{k+1}) = \mathcal{I}(j=k) + O(n^{-1})$  and by Brillinger (1981, p. 20 and Theorem 4.3.2), the last displayed expression is  $O(1)$ , and

hence (A.25) is  $O(n^{-1})$ . So, we conclude that

$$\check{\mathcal{L}}\hat{T}_n(\omega) = \frac{1}{\tilde{n}^{1/2}} \sum_{j=1}^{[\tilde{n}\omega]} \left\{ m_j - \mathring{G}_j \frac{1}{\tilde{n}} \sum_{k=j+1}^{\tilde{n}} g_{k+2} m_{k+1} \right\} + o_p(1),$$

where

$$\mathring{G}_j = \mathbb{E}(g'_{j+1})G_j^{-1}. \quad (\text{A.26})$$

So, it suffices to examine the asymptotic behaviour of

$$\check{\mathcal{L}}\hat{T}_n(\omega) = \frac{1}{\tilde{n}^{1/2}} \sum_{j=1}^{[\tilde{n}\omega]} \left\{ m_j - \mathring{G}_j \frac{1}{\tilde{n}} \sum_{k=j+1}^{\tilde{n}} g_{k+2} m_{k+1} \right\}, \quad (\text{A.27})$$

and more specifically that (a)  $|\mathbb{E}\check{\mathcal{L}}\hat{T}_n(\omega)| = o(1)$ , (b)  $\text{Cov}(\check{\mathcal{L}}\hat{T}_n(\omega_1), \check{\mathcal{L}}\hat{T}_n(\omega_2)) = (\omega_1 \wedge \omega_2)\pi^{-1} + o(1)$  and (c) the tightness of the process  $\check{\mathcal{L}}\hat{T}_n(\omega)$ .

We begin with part (a). Now, because  $\mathbb{E}m_j = 0$  and  $\|\mathring{G}_j\| < D$ , we have that

$$|\mathbb{E}\check{\mathcal{L}}\hat{T}_n(\omega)| \leq \frac{D}{\tilde{n}^{1/2}} \sum_{j=1}^{[\tilde{n}\omega]} \left\| \frac{1}{\tilde{n}} \sum_{k=j+1}^{\tilde{n}} \mathbb{E}(g_{k+2} m_{k+1}) \right\| = O(n^{-1/2}),$$

because  $\mathbb{E}(g_{k+2} m_{k+1}) = \text{Cov}(I_{\varepsilon, k+1}, g_{k+2}) = O(n^{-1})$ .

Now, we examine part (b). To that end it suffices to show that

$$(i) \mathbb{E} \left( \frac{1}{\tilde{n}^{1/2}} \sum_{j=1}^{[\tilde{n}\omega_1]} m_j \frac{1}{\tilde{n}^{1/2}} \sum_{j=1}^{[\tilde{n}\omega_2]} m_j \right)^2 + o(1) = \frac{(\omega_1 \wedge \omega_2)}{\pi} + o(1); \quad (\text{A.28})$$

(ii) that the contribution of the other three terms in  $\text{Cov}(\check{\mathcal{L}}\hat{T}_n(\omega_1), \check{\mathcal{L}}\hat{T}_n(\omega_2)) = o(1)$ .

That (A.28) holds true is standard. See for instance DHV's Lemma 7. Now, regarding part (ii), it suffices to see that

$$\begin{aligned} & -\frac{1}{\tilde{n}^2} \sum_{j \leq \ell}^{[\tilde{n}\omega]} \mathbb{E} \left\{ m_j \mathring{G}_\ell \sum_{k=\ell+1}^{\tilde{n}} g_{k+2} m_{k+1} \right\} - \frac{1}{\tilde{n}^2} \sum_{j \leq \ell}^{[\tilde{n}\omega]} \mathbb{E} \left\{ m_\ell \mathring{G}_j \sum_{k=j+1}^{\tilde{n}} g_{k+2} m_{k+1} \right\} \\ & + \frac{1}{\tilde{n}^3} \sum_{j \leq \ell}^{[\tilde{n}\omega]} \mathbb{E} \left\{ \left( \mathring{G}_\ell \sum_{k=\ell+1}^{\tilde{n}} g_{k+2} m_{k+1} \right) \left( \mathring{G}_j \sum_{k=j+1}^{\tilde{n}} g_{k+2} m_{k+1} \right) \right\} \end{aligned} \quad (\text{A.29})$$

is  $o(1)$ . Observe that this is the term we obtain when  $\omega_1 = \omega_2 = \omega$ . That (A.29) is  $o(1)$  follows because the first term on (A.29) is proportional to

$$\begin{aligned} & \frac{1}{\tilde{n}^2} \sum_{j \leq \ell}^{[\tilde{n}\omega]} \sum_{k=\ell+1}^{\tilde{n}} \mathbb{E}\{m_j g_{k+2} m_{k+1}\} = \frac{1}{\tilde{n}^2} \sum_{j \leq \ell}^{[\tilde{n}\omega]} \sum_{k=\ell+1}^{\tilde{n}} \{\mathbb{E}\{m_j\}\mathbb{E}\{g_{k+2} m_{k+1}\} \\ & + \mathbb{E}\{m_j g_{k+2}\}\mathbb{E}\{m_{k+1}\} + \mathbb{E}\{m_j m_{k+1}\}\mathbb{E}\{g_{k+2}\}\}, \end{aligned}$$

which is zero because  $\mathbb{E}\{m_j m_{p+1}\} = \mathbb{E}\{m_j\} = 0$ . Next, the second term of (A.29) is

$$-\frac{1}{\tilde{n}^2} \sum_{j \leq \ell}^{[\tilde{n}\omega]} \mathring{G}_j \sum_{k=j+1}^{\tilde{n}} \mathbb{E}\{m_\ell g_{k+2} m_{k+1}\} = -\frac{1}{\tilde{n}^2} \sum_{j \leq \ell}^{[\tilde{n}\omega]} \mathring{G}_j \mathbb{E}\{g_{\ell+1}\}$$

because  $\mathbb{E}\{m_j\} = 0$  and  $\mathbb{E}\{m_\ell m_{k+1}\} = \mathcal{I}(\ell = k + 1)$ . And finally, the third term of (A.29) is  $\tilde{n}^{-2} \sum_{j \leq \ell}^{\lfloor \tilde{n}\omega \rfloor} \mathring{G}_j \mathbb{E}\{g_{\ell+1}\} + O(n^{-1})$ . Indeed, proceeding as before using Brillinger (1981, p. 20 and theorem 4.3.2) as in the proof of Proposition A.1, that for instance  $\mathbb{E}\{m_{k+1} m_{q+1}\} = \mathcal{I}(p = q)$  and then using the definition of  $\mathring{G}_j$  in (A.26), the third term of (A.29) is

$$\frac{1}{\tilde{n}^3} \sum_{j \leq \ell}^{\lfloor \tilde{n}\omega \rfloor} \mathring{G}_j \mathring{G}_\ell \sum_{k=\ell+1}^{\tilde{n}} \mathbb{E}(g_{k+2}^2) + O(n^{-1}) = \frac{1}{\tilde{n}^2} \sum_{j \leq \ell}^{\lfloor \tilde{n}\omega \rfloor} \mathring{G}_j \mathbb{E}\{g_{\ell+1}\} + o(1).$$

So, we conclude part (b) that  $\text{Cov}(\check{\mathcal{L}}\hat{T}_n(\omega_1), \check{\mathcal{L}}\hat{T}_n(\omega_2)) = (\omega_1 \wedge \omega_2)\pi^{-1} + o(1)$ .

To complete the proof we need to show part (c). From the definition of  $\check{\mathcal{L}}\hat{T}_n(\omega)$  in (A.27), it suffices to examine the tightness of

$$\frac{1}{\tilde{n}^{1/2}} \sum_{j=1}^{\lfloor \tilde{n}\omega \rfloor} \mathring{G}_j \frac{1}{\tilde{n}} \sum_{k=j+1}^{\tilde{n}} g_{k+2} m_{k+1}$$

as  $\tilde{n}^{-1/2} \sum_{j=1}^{\lfloor \tilde{n}\omega \rfloor} m_j$  is known to be tight. See, for instance, DHV. Now, because by Assumption 2.3,  $\|\mathring{G}_j - \mathring{G}_{j+1}\| = O(n^{-1})$ , it suffices to examine the tightness of

$$\begin{aligned} \frac{1}{\tilde{n}^{1/2}} \sum_{j=1}^{\lfloor \tilde{n}\omega \rfloor} \frac{1}{\tilde{n}} \sum_{k=j+1}^{\tilde{n}} g_{k+2} m_{k+1} &= (1 - \omega) \frac{1}{\tilde{n}^{1/2}} \sum_{\ell=1+\lfloor \tilde{n}\omega \rfloor}^{\tilde{n}} g_{\ell+2} m_{\ell+1} \\ &\quad + \frac{1}{\tilde{n}^{1/2}} \sum_{\ell=1}^{\lfloor \tilde{n}\omega \rfloor} \left(1 - \frac{\ell}{\tilde{n}}\right) g_{\ell+2} m_{\ell+1}. \end{aligned}$$

We shall examine the second term on the right-hand side being the first one similarly handled. Now by standard arguments, see Billingsley (1968), we only need to show that

$$\mathbb{E} \left\| \frac{1}{\tilde{n}^{1/2}} \sum_{\ell=1+\lfloor \tilde{n}\omega_1 \rfloor}^{\lfloor \tilde{n}\omega_2 \rfloor} g_{\ell+2} m_{\ell+1} \right\|^4 \leq D(\omega_2 - \omega_1)^{1+\delta}$$

for some  $\delta > 0$ . Now, the left-hand side of the last displayed expression is

$$\begin{aligned} &\frac{1}{\tilde{n}^2} \sum_{1+\lfloor \tilde{n}\omega_1 \rfloor = \ell_1, \ell_2, \ell_3, \ell_4}^{\lfloor \tilde{n}\omega_2 \rfloor} \mathbb{E}(g_{\ell_1+2} m_{\ell_1+1} g_{\ell_2+2} m_{\ell_2+1} g_{\ell_3+2} m_{\ell_3+1} g_{\ell_4+2} m_{\ell_4+1}) \\ &= 3 \frac{1}{\tilde{n}^2} \left( \sum_{1+\lfloor \tilde{n}\omega_1 \rfloor = \ell_1, \ell_2}^{\lfloor \tilde{n}\omega_2 \rfloor} \mathbb{E}(g_{\ell_1+2} m_{\ell_1+1} g_{\ell_2+2} m_{\ell_2+1}) \right)^2 \\ &\quad + \frac{1}{\tilde{n}^2} \sum_{1+\lfloor \tilde{n}\omega_1 \rfloor = \ell_1, \ell_2, \ell_3, \ell_4}^{\lfloor \tilde{n}\omega_2 \rfloor} \text{cum}(g_{\ell_1+2} m_{\ell_1+1}, g_{\ell_2+2} m_{\ell_2+1}, g_{\ell_3+2} m_{\ell_3+1}, g_{\ell_4+2} m_{\ell_4+1}). \end{aligned}$$

Now proceed as we did in part (b) to conclude that the right-hand side of the last displayed expression is bounded by  $D(\omega_2 - \omega_1)^2$  after we notice that we can always take  $\omega_1$  and  $\omega_2$  such that  $\tilde{n}^{-1} \leq (\omega_2 - \omega_1)$ . This completes the proof of the theorem.  $\square$

**Proof of Theorem 4.1:** From the definition of  $\hat{F}_n(\omega, \hat{\theta}_n)$  in (2.7), under  $H_{an}$ , and proceeding as in Proposition 2.2, we have that

$$\begin{aligned}\hat{F}_n(\omega, \hat{\theta}_n) &= \hat{F}_n(\omega) - \frac{4\pi}{\tilde{n}^{3/2}} \sum_{j=1}^{[\omega\tilde{n}]} \operatorname{Re} I_{\varepsilon \mathbf{w}, j} \tilde{n}^{1/2} (\hat{\theta}_n - \theta_0) \\ &\quad - c \frac{4\pi}{\tilde{n}^{3/2}} \sum_{j=1}^{[\omega\tilde{n}]} \operatorname{Re} I_{\varepsilon X(-p-1), j} + o_p(1) \\ &= \hat{F}_n(\omega) - \tilde{n}^{-1/2} \{ \Phi(\omega)' (\theta_n - \theta_0) + c\sigma^2 \Psi(\omega) \} + o_p(1),\end{aligned}$$

uniformly in  $\omega \in [0, 1]$ . From here, (4.2) follows repeating the same steps of Theorems 3.1 and 3.2, but noting the additional term given by  $\Psi(\omega) := \int_0^\omega l(\pi u) du$  in the general case. So, under  $H_{an}$ , the  $B^0 + \mathcal{L}^0\Psi$  is a non-centred Gaussian process, being the ‘non-centrality function’ given by  $\mathcal{L}^0\Psi$ . Now, the test will have non-trivial power under  $H_{an}$  if  $\mathcal{L}^0\Psi(\omega) \neq 0$  in a set, say  $\Pi(L)$ , with Lebesgue measure greater than zero. From the definitions of  $\mathcal{L}^0$  and  $\Psi$  and that  $\int_0^1 \phi(v) dv = 0$ , it is easily seen that

$$\mathcal{L}^0\Psi(\omega) = \int_0^\omega \left\{ l(\pi u) - g(u)' \Sigma(u)^{-1} \int_u^1 g(v) l(\pi v) dv \right\} du.$$

However, the expression in braces is just the residuals from the least-squares projection of  $l(\pi u)$  on  $g(u) = (1, \phi(u)')'$ , which obviously is different than zero unless  $l(\pi u)$  is in the space spanned by  $g(u)$ . But the latter is ruled out, which concludes the proof.  $\square$