NEGLECTED SERIAL CORRELATION TESTS IN UCARIMA MODELS

Gabriele Fiorentini and Enrique Sentana

CEMFI Working Paper No. 1406

October 2014

CEMFI
Casado del Alisal 5; 28014 Madrid
Tel. (34) 914 290 551 Fax (34) 914 291 056
Internet: www.cemfi.es

We are grateful to Andrew Harvey for helpful discussions. Financial support from MIUR through the project "Multivariate statistical models for risk assessment" (Fiorentini) and the Spanish Ministry of Science and Innovation through grant ECO 2011-26342 (Sentana) is gratefully acknowledged.

CEMFI Working Paper 1406
October 2014

NEGLECTED SERIAL CORRELATION TESTS IN UCARIMA MODELS

Abstract

We derive computationally simple and intuitive score tests of neglected serial correlation in unobserved component univariate models using frequency domain techniques. In some common situations in which the information matrix is singular under the null we derive extremum tests that are asymptotically equivalent to likelihood ratio tests, which become one-sided, and explain how to compute reliable Wald tests. We also explicitly relate the incidence of those problems to the model identification conditions and compare our tests with tests based on the reduced form prediction errors. Our Monte Carlo exercises assess the finite sample reliability and power of our proposed tests.

JEL Codes: C22, C52, C12.
Keywords: Extremum tests, Kalman filter, LM tests, singular information matrix, spectral maximum likelihood, Wiener-Kolmogorov filter.

Gabriele Fiorentini
Università di Firenze
fiorentini@disia.unifi.it

Enrique Sentana
CEMFI
sentana@cemfi.com

1 Introduction

The superposition of ARIMA time series models forms the basis of two dominant approaches to the classical decomposition of a univariate time series into trend, cyclical, seasonal and irregular components: the reduced form "model-based" decomposition analysed by Box, Hillmer and Tiao (1978) and Pierce (1978) and further extended by Agustín Maravall and his co-authors, and the so-called "structural time series" models studied by Nerlove (1967), Engle (1978) and Nerlove, Grether and Carvalho (1979) and subsequently developed by Andrew Harvey and his co-authors.

In both cases, the model parameters are estimated by maximising the Gaussian log-likelihood function of the observed data, which can be readily obtained either as a by-product of the Kalman filter prediction equations or from Whittle's (1962) frequency domain asymptotic approximation. Once the parameters have been estimated, filtered values of the unobserved components can be extracted by means of the Kalman smoother or its Wiener-Kolmogorov counterpart. These estimation and filtering issues are well understood (see Harvey (1989) and Durbin and Koopman (2012) for textbook treatments), and the same can be said of their efficient numerical implementation (see Commandeur, Koopman and Ooms (2011) and the references therein).

In contrast, specification tests for these models are far less known. While sophisticated users will often look at several diagnostics, formal tests are hardly ever reported in empirical work. One particularly relevant issue is the correct specification of the parametric ARIMA models for the unobserved components, as the various outputs of the model could be misleading under misspecified dynamics.

The objective of our paper is precisely to derive tests for neglected serial correlation in the unobserved components. For computational reasons, we focus most of our discussion on score tests, which only require estimation of the model under the null. As is well known, though, in standard situations likelihood ratio (LR), Wald and Lagrange multiplier (LM) tests are asymptotically equivalent under the null and sequences of local alternatives, and therefore they share their optimality properties. Another important advantage of score tests is that they often coincide with tests of easy to interpret moment conditions (see Newey (1985) and Tauchen (1985)), which will continue to have non-trivial power even in situations for which they are not optimal.

Earlier work on specification testing in unobserved component models include Engle and Watson (1980), who explained how to apply the LM testing principle in the time domain for dynamic factor models with static factor loadings, Harvey (1989), who provides a detailed discussion of time domain and frequency domain testing methods in the context of univariate "structural time series" models, and Fernández (1990), who applied the LM principle in the frequency domain to a multivariate structural time series model. More recently, in a companion paper (Fiorentini and Sentana (2013)) we have derived tests for neglected serial correlation in the latent variables of dynamic factor models using frequency domain techniques.

In the context of univariate unobserved components (UCARIMA) models, the contribution of this paper is threefold.

First, we propose dynamic specification test which are very simple to implement, and even simpler to interpret. Once an model has been specified and estimated, the tests that we propose can be routinely computed from simple statistics of the smoothed values of the different components. And even though our theoretical derivations make extensive use of spectral methods for time series, we provide both time domain and frequency domain interpretations of the relevant scores, so researchers who strongly prefer one method over the other could apply them without abandoning their favourite estimation techniques.

Second, we provide a thorough discussion of some common situations in which the standard form of LM tests cannot be computed because the information matrix is singular under the null. In those irregular cases, we derive versions of the score tests that remain asymptotically equivalent to the LR tests, which become one-sided, and explain how to compute asymptotically reliable Wald tests. We also explicitly relate the incidence of those problems to the identification conditions for UCARIMA models, and highlight that they contradict the widely held view that increases in the MA and AR polynomials of the same order provide locally equivalent alternatives in univariate tests for serial correlation (see e.g. Godfrey (1988)).

Third, we compare dynamic specification tests for the underlying components with tests based on the reduced form prediction errors. In this regard, we study their relative power and discuss some cases in which they are numerical equivalent.

The rest of the paper is organised as follows. In section 2, we review the properties of UCARIMA models, their estimators and filters. Then, we derive our tests in section 3, discuss potential pitfalls in section 4 and compare them to reduced form tests in section 5. This is followed by a Monte Carlo evaluation of their finite sample behaviour in section 6. Finally, our conclusions can be found in section 7. Auxiliary results are gathered in appendices.

2 Theoretical background

2.1 UCARIMA models

To keep the notation to a minimum throughout the paper we focus on models for a univariate observed series, y_t that can be defined in the time domain by the equations:

$$
\begin{aligned}
y_t &= \mu + x_t + u_t, && (1)\\
\alpha_x(L)x_t &= \beta_x(L)f_t, && (2)\\
\alpha_u(L)u_t &= \beta_u(L)v_t, && (3)\\
\begin{pmatrix} f_t \\ v_t \end{pmatrix} |I_{t-1};\mu,\boldsymbol{\theta} &\sim N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_f^2 & 0 \\ 0 & \sigma_v^2 \end{pmatrix}\right], && (4)
\end{aligned}
$$

where x_t is the "signal" component, u_t the orthogonal "non-signal" component, $\alpha_x(L)$ and $\alpha_u(L)$ are one-sided polynomials of orders p_x and p_u, respectively, while $\beta_x(L)$ and $\beta_u(L)$ are one-sided polynomials of orders q_x and q_u coprime with $\alpha_x(L)$ and $\alpha_u(L)$, respectively, I_{t-1} is an information set that contains the values of y_t and x_t up to, and including time $t-1$, μ is the unconditional mean and $\boldsymbol{\theta}$ refers to all the remaining model parameters.

Importantly, we maintain the assumption that the researcher makes sure that the parameters $\boldsymbol{\theta}$ are identified before estimating the model under the null.[1] Hotta (1989) provides a systematic way to check for identification (see Maravall (1979) for closely related results). Specifically, let c denote the degree of the polynomial greatest common divisor of $\alpha_x(L)$ and $\alpha_u(L)$, so that they share c common roots. Then, the UCARIMA model above will be identified (except at a set of parameter values of measure 0) when there are no restrictions on the AR and MA polynomials if and only if either $p_x \geq q_x+c+1$ or $p_u \geq q_u+c+1$, so that at least one of the components must be a "top-heavy" ARMA process in the terminology of Burman (1980).[2] Given the exchangeability of signal and non-signal components in the formulation above, in what follows we assume without loss of generality that this identification condition is satisfied by the signal component.

In this paper we are interested in hypothesis tests for $p_x = p_x^0$ vs $p_x = p_x^0 + k_x$ or $p_u = p_u^0$ vs $p_u = p_u^0 + k_u$, or the analogous hypotheses for q_x and q_{u_i}. For simplicity, we focus most of the discussion in those cases in which k_x and k_u are in fact 1, which leads to the following four hypothesis of interest:

1. SAR1: ARMA(p_x+1, q_x)+ARMA(p_u, q_u)

[1] But see section 7 for a brief discussion of models that are underidentified under the null but identified under the alternative.

[2] Although strictly speaking Proposition 2 in Hotta (1989) applies to stationary models, the emphasis on common roots is particularly important in the presence of integrated components, in which case p_x and p_u would represent the total number of AR roots, including those on the unit circle (see Harvey (1989) for further details).

2. Sma1: Arma$(p_x, q_x + 1)$+Arma(p_u, q_u)

3. Nar1: Arma(p_x, q_x)+Arma$(p_u + 1, q_u)$

4. Nma1: Arma(p_x, q_x)+Arma$(p_u, q_u + 1)$

Extensions to higher k_x and k_u are briefly discussed in section 3.1 below, as well as in our concluding remarks.

2.2 The reduced form model

Unobserved component models can readily handle integrated variables, but for simplicity of exposition in what follows we maintain the assumption that y_t is a covariance stationary process, possibly after suitable differencing, as in appendix A.

Under stationarity, the spectral density of the observed variable is proportional to

$$\begin{aligned} g_{yy}(\lambda) &= g_{xx}(\lambda) + g_{uu}(\lambda), \\ g_{xx}(\lambda) &= \sigma_f^2 \frac{\beta_x(e^{-i\lambda t})\beta_x(e^{i\lambda t})}{\alpha_x(e^{-i\lambda t})\alpha_x(e^{i\lambda t})}, \\ g_{uu}(\lambda) &= \sigma_v^2 \frac{\beta_u(e^{-i\lambda})\beta_u(e^{i\lambda t})}{\alpha_u(e^{-i\lambda t})\alpha_u(e^{i\lambda t})}. \end{aligned}$$

Given that

$$\begin{aligned} g_{yy}(\lambda) &= \frac{\sigma_f^2\beta_x(e^{-i\lambda t})\beta_x(e^{i\lambda t})\alpha_u(e^{-i\lambda t})\alpha_u(e^{i\lambda t}) + \sigma_v^2\beta_u(e^{-i\lambda})\beta_u(e^{i\lambda t})\alpha_x(e^{-i\lambda t})\alpha_x(e^{i\lambda t})}{\alpha_x(e^{-i\lambda t})\alpha_x(e^{i\lambda t})\alpha_u(e^{-i\lambda t})\alpha_u(e^{i\lambda t})} \\ &= \sigma_a^2 \frac{\beta_y(e^{-i\lambda t})\beta_y(e^{i\lambda t})}{\alpha_y(e^{-i\lambda t})\alpha_y(e^{i\lambda t})}, \end{aligned}$$

it follows that the reduced form model will be an Arma process with maximum orders $p_y = p_x + p_u$ for the Ar polynomial $\alpha_y(.) = \alpha_x(.)\alpha_u(.)$ and $q_y = \max(p_x + q_u, q_x + p_u)$ for the Ma polynomial $\beta_y(.)$. Cancellation will trivially occur when $\alpha_x(.)$ and $\alpha_u(.)$ share c common roots, but there could also be other cases (see Granger and Morris (1976) for further details). The coefficients of $\beta_y(L)$, as well as σ_a^2, which is the variance of the univariate Wold innovations, a_t, are obtained by matching autocovariances (see Fiorentini and Planas (1998) for a comparison of numerical methods). Assuming strict invertibility of the Ma part, we could then obtain the reduced form innovations a_t from the observed process by means of the one-sided filter

$$\alpha_y(e^{-i\lambda t})/\beta_y(e^{-i\lambda t}).$$

But as is well known, these reduced form residuals can also be obtained from the prediction equations of the Kalman filter without making use of the expressions for $\alpha_y(.)$ or $\beta_y(.)$.

2.3 Maximum likelihood estimation in the frequency domain

Let

$$I_{yy}(\lambda) = \frac{1}{2\pi T}\sum_{t=1}^{T}\sum_{s=1}^{T}(y_t - \mu)(y_s - \mu)' e^{-i(t-s)\lambda} \tag{5}$$

denote the periodogram of y_t and $\lambda_j = 2\pi j/T$ $(j = 0, \ldots T-1)$ the usual Fourier frequencies. If we assume that $g_{yy}(\lambda)$ is not zero at any frequency, the so-called Whittle (discrete) spectral approximation to the log-likelihood function is[3]

$$-\frac{NT}{2}\ln(2\pi) - \frac{1}{2}\sum_{j=0}^{T-1}\ln|g_{yy}(\lambda_j)| - \frac{1}{2}\sum_{j=0}^{T-1}\frac{2\pi I_{yy}(\lambda_j)}{g_{yy}(\lambda_j)}. \tag{6}$$

The MLE of μ, which only enters through $I_{yy}(\lambda)$, is the sample mean, so in what follows we focus on demeaned variables. In turn, the score with respect to all the remaining parameters is

$$\mathbf{s}_{\boldsymbol{\theta}}(\boldsymbol{\theta}) = \frac{1}{2}\sum_{j=0}^{T-1}\frac{\partial g_{yy}(\lambda_j)]}{\partial\boldsymbol{\theta}}M(\lambda_j)m(\lambda_j), \tag{7}$$

$$m(\lambda) = 2\pi I_{yy}(\lambda) - g_{yy}(\lambda),$$

$$M(\lambda) = g_{yy}^{-2}(\lambda).$$

The information matrix is block diagonal between μ and the elements of $\boldsymbol{\theta}$, with the (1,1)-element being $g_{yy}(0)$ and the (2,2)-block

$$\mathbf{Q} = \frac{1}{4\pi}\int_{-\pi}^{\pi}\frac{\partial g_{yy}(\lambda_j)]}{\partial\boldsymbol{\theta}}M(\lambda)\left\{\frac{\partial g_{yy}(\lambda_j)]}{\partial\boldsymbol{\theta}}\right\}^{*}d\lambda, \tag{8}$$

where * denotes the conjugate transpose of a matrix. A consistent estimator will be provided either by the outer product of the score or by

$$\boldsymbol{\Phi}(\boldsymbol{\theta}) = \frac{1}{2}\sum_{j=0}^{T-1}\frac{\partial g_{yy}(\lambda_j)]}{\partial\boldsymbol{\theta}}M(\lambda_j)\left\{\frac{\partial g_{yy}(\lambda_j)]}{\partial\boldsymbol{\theta}}\right\}^{*}.$$

Formal results showing the strong consistency and asymptotic normality of the resulting ML estimators under suitable regularity conditions have been provided by Dunsmuir and Hannan (1976) and Dunsmuir (1979), who also show their asymptotic equivalence to the time domain ML estimators.[4]

[3] There is also a continuous version which replaces sums by integrals (see Dusmuir and Hannan (1976)).

[4] This equivalence is not surprising in view of the contiguity of the Whittle measure in the Gaussian case (see Choudhuri, Ghosal and Roy (2004)).

2.4 The (Kalman-)Wiener-Kolmogorov filter

By working in the frequency domain we can easily obtain smoothed estimators of the latent variables too. Specifically, let

$$\begin{aligned} y_t - \mu &= \int_{-\pi}^{\pi} e^{i\lambda t} dZ_y(\lambda), \\ V[dZ_y(\lambda)] &= g_{yy}(\lambda) d\lambda \end{aligned}$$

denote the spectral decomposition of the observed process. The Wiener-Kolmogorov two-sided filter for the signal x_t at each frequency is given by

$$g_{xx}(\lambda) g_{yy}^{-1}(\lambda) dZ_y(\lambda).$$

Hence, the spectral density of the smoother $x_{t|T}^K$ as $T \to \infty$[5] will be

$$g_{x^K x^K}(\lambda) = \frac{g_{xx}^2(\lambda)}{g_{yy}(\lambda)} = \frac{g_{xx}(\lambda)}{g_{xx}(\lambda) + g_{uu}(\lambda)} g_{xx}(\lambda) = R_{xx}^2(\lambda) g_{xx}(\lambda), \tag{9}$$

while the spectral density of the final estimation error $x_t - x_{t|\infty}^K$ will be given by

$$g_{xx}(\lambda) - g_{x^K x^K}(\lambda) = [1 - R_{xx}^2(\lambda)] g_{xx}(\lambda) = \omega_{xx}(\lambda). \tag{10}$$

It is easily seen that $g_{x^K x^K}(\lambda)$ will approach $g_{xx}(\lambda)$ at those frequencies for which $g_{xx}(\lambda)$ is large relatively to $g_{uu}(\lambda)$, i.e. frequencies with a high signal to noise ratio. In this regard, we can view $R_{xx}^2(\lambda)$ as a frequency-by-frequency coefficient of determination.

Having smoothed y_t to estimate x_t, we can easily obtain the smoother for f_t, $f_{t|\infty}^K$, by applying to $x_{t|\infty}^K$ the one-sided filter

$$\alpha_x(e^{-i\lambda}) / \beta_x(e^{-i\lambda}). \tag{11}$$

Likewise, we can derive its spectral density, as well as the spectral density of its final estimation error $f_t - f_{t|\infty}^K$.

Entirely analogous derivations apply to the non-signal component u_t, with the peculiarity that

$$x_{t|\infty}^K + u_{t|\infty}^K = y_t$$

so that

$$R_{xx}^2(\lambda) + R_{uu}^2(\lambda) = 1 \quad \forall \lambda.$$

[5] The main difference between the Wiener-Kolmogorov filtered values, $x_{t|\infty}^K$, and the Kalman filter smoothed values, $x_{t|T}^K$, results from the dependence of the former on a double infinite sequence of observations (but see Levinson (1947)). As shown by Fiorentini (1995) and Gómez (1999), though, they can be made numerically identical by replacing both pre- and post- sample observations by their least squares projections onto the linear span of the sample observations.

Finally, we can obtain the autocovariances of $x^K_{t|\infty}$, $f^K_{t|\infty}$, $u^K_{t|\infty}$, $v^K_{t|\infty}$ and their final estimation errors by applying the usual inverse Fourier transformation

$$\gamma_{zz}(k) = cov(z_t, z_{t-k}) = \int_{-\pi}^{\pi} e^{i\lambda k} g_{zz}(\lambda) d\lambda.$$

2.5 Autocorrelation structure of the smoothed variables

As we have seen in the previous section, smoothed values of the latent variables are the result of optimal symmetric two-sided filters. As a consequence, their serial correlation structure is generally different from that of the unobserved state variables. To see the difference between the spectra of the signal and its estimators, recall that (9) implies that $g_{x^K x^K}(\lambda) < g_{xx}(\lambda)$ for any $\lambda \in (-\pi, \pi)$ for which $g_{uu}(\lambda) > 0$. Therefore, the variance of the optimal estimator will underestimate the variance of the unobserved signal, as expected.

As discussed in Maravall (1999), the serial dependence structure of the estimators of the unobserved components can be a useful tool for model diagnostic. Large discrepancies between theoretical and empirical autocovariance functions of those estimators can be interpreted as indication of model misspecification. As we shall in section 3.2, our LM tests carry out this comparison in a very precise statistical sense.

Given (9), we can write the spectral density of x^K as

$$g_{x^K x^K}(\lambda) = \frac{\sigma_f^4 \beta_x^2(e^{-i\lambda})\beta_x^2(e^{i\lambda})\alpha_u(e^{-i\lambda t})\alpha_u(e^{i\lambda t})}{\sigma_a^2 \alpha_x(e^{-i\lambda})\alpha_x(e^{i\lambda})\beta_y(e^{-i\lambda})\beta_y(e^{i\lambda})},$$

which corresponds to an $\textsc{Arma}(p_x + q_y, p_u + 2q_x)$ process in the absence of cancellation. Hence, the spectral density of the final estimation error $x_t - x^K_{t|\infty}$ in (10) will be

$$\omega_{xx}(\lambda) = \frac{\sigma_f^2 \sigma_v^2 \beta_x(e^{-i\lambda})\beta_x(e^{i\lambda})\beta_u(e^{-i\lambda t})\beta_u(e^{i\lambda t})}{\sigma_a^2 \beta_y(e^{-i\lambda})\beta_y(e^{i\lambda})},$$

which shares the structure of an $\textsc{Arma}(q_y, q_x + q_u)$ under the same circumstances.

In turn, the application of (11) to $x^K_{t|\infty}$ implies that the spectral density of $f^K_{t|\infty}$ will be

$$g_{f^K f^K}(\lambda) = \frac{\sigma_f^4 \beta_x(e^{-i\lambda})\beta_x(e^{i\lambda})\alpha_u(e^{-i\lambda t})\alpha_u(e^{i\lambda t})}{\sigma_a^2 \beta_y(e^{-i\lambda})\beta_y(e^{i\lambda})},$$

which suggests an $\textsc{Arma}(q_y, p_u + q_x)$ process, while

$$\omega_{ff}(\lambda) = \sigma_f^2 - g_{f^K f^K}(\lambda) = \frac{\sigma_f^2 \sigma_v^2 \beta_u(e^{-i\lambda})\beta_u(e^{i\lambda})\alpha_x(e^{-i\lambda t})\alpha_x(e^{i\lambda t})}{\sigma_a^2 \beta_y(e^{-i\lambda})\beta_y(e^{i\lambda})}$$

points out instead to an $\textsc{Arma}(q_y, p_x + q_u)$ for the final estimation error $f_t - f^K_{t|\infty}$.

There are special cases, however, in which the resulting models for the smoothed values of the unobserved variables and their innovations are much simpler. For example, if the signal follows a purely autoregressive process and the non-signal component is white noise, so that $\beta_x(L) = \alpha_u(L) = \beta_u(L) = 1$, then

$$
\begin{aligned}
g_{x^K x^K}(\lambda) &= \frac{\sigma_f^2 \sigma_v^2}{\sigma_a^2 \alpha_x(e^{-i\lambda})\alpha_x(e^{i\lambda})\beta_y(e^{-i\lambda})\beta_y(e^{i\lambda})}, \\
\omega_{xx}(\lambda) &= \frac{\sigma_f^2 \sigma_v^2}{\sigma_a^2 \beta_y(e^{-i\lambda})\beta_y(e^{i\lambda})}, \\
g_{f^K f^K}(\lambda) &= \frac{\sigma_f^4}{\sigma_a^2 \beta_y(e^{-i\lambda})\beta_y(e^{i\lambda})},
\end{aligned}
$$

and

$$
\omega_{ff}(\lambda) = \frac{\sigma_f^2 \sigma_v^2 \alpha_x(e^{-i\lambda t})\alpha_x(e^{i\lambda t})}{\sigma_a^2 \beta_y(e^{-i\lambda})\beta_y(e^{i\lambda})},
$$

with $p_y = q_y = p_x$.

Once again, entirely analogous derivations apply to the non-signal component $u^K_{t|\infty}$.

3 Neglected serial correlation tests

3.1 Testing for serial correlation in univariate observable processes

For pedagogical purposes, let us initially assume that x_t is an observable time series that has been modelled as an AR(2) process. A natural generalisation is

$$
(1 - \psi_{x1} L)(1 - \alpha_{x1} L - \alpha_{x2} L^2) x_t = f_t,
$$

so that the null becomes $H_0 : \psi_x = 0$.[6] Under the alternative, the spectral density of x_t is

$$
g_{xx}(\lambda|\sigma_f^2, \alpha_{x1}, \alpha_{x2}, \psi_x) = \frac{1}{1 + \psi_x^2 - 2\psi_x \cos\lambda} \cdot g_{xx}(\lambda|\sigma_f^2, \alpha_{x1}, \alpha_{x2}, 0),
$$

where

$$
g_{xx}(\lambda|\sigma_f^2, \alpha_{x1}, \alpha_{x2}, 0) = \frac{\sigma_f^2}{1 + \alpha_{x1}^2 + \alpha_{x2}^2 - 2\alpha_{x1}(1 - \alpha_{x2})\cos\lambda - 2\alpha_{x2}\cos 2\lambda}.
$$

Hence, the derivative of $g_{xx}(\lambda)$ with respect to ψ_x under the null is

$$
\frac{\partial g_{xx}(\lambda|\sigma_f^2, \alpha_{x1}, \alpha_{x2}, 0)}{\partial \psi_x} = 2\cos\lambda \cdot g_{xx}(\lambda|\sigma_f^2, \alpha_{x1}, \alpha_{x2}, 0). \tag{12}
$$

As a result, the spectral version of the score with respect to ψ_x under H_0 is

$$
\sum_{j=0}^{T-1} \cos\lambda_j g_{xx}^{-1}(\lambda_j)[2\pi I_{xx}(\lambda_j) - g_{xx}(\lambda_j)] = \sum_{j=0}^{T-1} \cos\lambda_j [2\pi I_{ff}(\lambda_j)],
$$

[6] This is a multiplicative alternative. Instead, we could test $H_0 : \alpha_{x3} = 0$ in the additive alternative

$$
(1 - \alpha_{x1} L - \alpha_{x2} L^2 - \alpha_{x3} L^3) x_t = f_t.
$$

In that case, it would be more convenient to reparametrise the model in terms of partial autocorrelations (see Barndorff-Nielsen and Schou (1973)).

where we have exploited the fact that

$$\sum_{j=0}^{T-1} \frac{\partial g_{xx}(\lambda_j)}{\partial \psi_x} g_{xx}^{-1}(\lambda_j) = \sum_{j=0}^{T-1} \cos \lambda_j = 0. \tag{13}$$

Given that

$$I_{ff}(\lambda_j) = \hat{\gamma}_{ff}(0) + 2\sum_{k=1}^{T-1} \hat{\gamma}_{ff}(k) \cos(k\lambda_j),$$

the spectral version of the score becomes

$$\sum_{j=0}^{T-1} \cos \lambda_j [2\pi I_{ff}(\lambda_j)] = T[\hat{\gamma}_{ff}(1) + \hat{\gamma}_{ff}(T-1)]. \tag{14}$$

In turn, the time domain version of the score will be

$$\sum_t (x_t - \alpha_{x1}x_{t-1} - \alpha_{x2}x_{t-2})(x_{t-1} - \alpha_{x1}x_{t-2} - \alpha_{x2}x_{t-3}) = \sum_t f_t f_{t-1},$$

which is essentially identical because $\hat{\gamma}_{ff}(T-1) = T^{-1}f_T f_1 = o_p(1)$. Therefore, the spectral LM test of AR(2) versus AR(3) is simply checking that the first sample (circulant) autocovariance of f_t, which are the innovations in the observed process, coincides with its theoretical value under H_0, exactly like the usual Breusch (1978) - Godfrey (1978a) serial correlation LM test in the time domain (see also Breusch and Pagan (1980) or Godfrey (1988)).

Let us now consider the following alternative generalisation of an AR(2)

$$(1 - \alpha_{x1}L - \alpha_{x2}L^2)x_t = (1 - \psi_f L)f_t.$$

In this case, the null is $H_0 : \psi_f = 0$. In turn, the spectral density of x_t under this alternative is

$$(1 + \psi_f^2 - 2\psi_f \cos \lambda) \cdot g_{xx}(\lambda|\sigma_f^2, \alpha_{x1}, \alpha_{x2}, 0),$$

whose derivative with respect to ψ_f under the null is

$$\frac{\partial g_{xx}(\lambda)}{\partial \psi_f} = -2\cos \lambda \cdot g_{xx}(\lambda). \tag{15}$$

Therefore, the spectral LM test of AR(2) versus ARMA(2,1) will be numerically identical to the corresponding test of AR(2) versus AR(3), which confirms that these two alternative hypotheses are locally equivalent for observable time series (see e.g. Godfrey (1988)).

Generalisations to test ARMA(p,q) vs ARMA(p+k,q) for k>1 are straightforward, since they only involve higher order (circulant) autocovariances of f_t, as in Godfrey (1978b). Similarly, it is easy to show that ARMA(p+k,q) and ARMA(p,q+k) multiplicative alternatives are also locally equivalent[7] Finally, we could also consider (multiplicative) seasonal alternatives.

[7]It would also be possible to develop tests of ARMA(p,q) against ARMA(p+k,q+k) along the lines of Andrews and Ploberger (1996). We leave those tests, which will also depend on the differences between sample and population autocovariances of f_t, for future research.

3.2 Testing for neglected serial correlation in the unobserved components

Let us now consider unobserved components models, which is the objective of our study. Initially, we assume that the model is identified under each of the four alternatives stated in section 2.1, and therefore overidentified under the null, and postpone the discussion of the more general case to section 4. In view of Hotta's (1989) results, this requires that either $p_x \geq q_x+c+2$ or $p_u \geq q_u + c + 2$ when there are c common AR roots.

Let us start by considering neglected serial correlation in the signal. Under alternative SAR1 the model will be

$$\left.\begin{array}{c} y_t = \mu + x_t + u_t, \\ (1-\psi_x L)\alpha_x(L)x_t = \beta_x(L)f_t, \\ \alpha_u(L)u_t = \beta_u(L)v_t, \end{array}\right\} \tag{16}$$

so that the null hypothesis is $H_0 : \psi_x = 0$, as in section 3.1. Given

$$\frac{\partial g_{yy}(\lambda)}{\partial \psi_x} = \frac{\partial g_{xx}(\lambda)}{\partial \psi_x} \tag{17}$$

and (12), after some straighforward manipulations we can prove that the score of the spectral log-likelihood for the observed series y_t under the null will be given by

$$\begin{aligned} & 2\sum_{j=0}^{T-1} \cos\lambda_j g_{xx}(\lambda_j) g_{yy}^{-2}(\lambda_j)[2\pi I_{yy}(\lambda_j) - g_{yy}(\lambda_j)]. \\ = \; & 2\sum_{j=0}^{T-1} \cos\lambda_j g_{xx}^{-1}(\lambda_j)[2\pi I_{x^K x^K}(\lambda_j) - g_{x^K x^K}(\lambda_j)] \\ = \; & 2\sum_{j=0}^{T-1} \cos\lambda_j [2\pi I_{f^K f^K}(\lambda_j) - g_{f^K f^K}(\lambda_j)]. \end{aligned}$$

Once more, the time domain counterpart to the spectral score with respect to ψ_x is (asymptotically) proportional to the difference between the first sample autocovariance of $f_{t|\infty}^K$ and its theoretical counterpart under H_0. Therefore, the only difference with the observable case is that the autocovariance of $f_{t|\infty}^K$, which is a forward filter of the Wold innovations of y_t, is no longer 0 when $\psi_x = 0$, although it approaches 0 as the signal to noise ratio increases. In that case, our proposed tests would converge to the usual Breusch-Godfrey LM tests for neglected serial correlation discussed in section 3.1.[8]

Let us illustrate our test by means of a simple example. Imagine that x_t follows an AR(2) process while u_t is white noise. The results in section 2.5 imply that when $\psi_x = 0$ $f_{t|\infty}^K$ will follow an AR(2) with an autoregressive polynomial $\beta_y(L)$ that satisfies the condition

$$\sigma_a^2 \beta_y(L)\beta_y(L^{-1}) = \sigma_f^2 + \alpha_x(L)\alpha_x(L^{-1})\sigma_v^2,$$

[8] Given that $\sigma_f^2 = g_{f^K f^K}(\lambda) + \omega_{ff}(\lambda)$ for all λ, we can also write the score as $2\sum_{j=0}^{T-1} \cos\lambda_j[2\pi I_{f^K f^K}(\lambda_j) + \omega_{ff}(\lambda_j)]$ in view of (13). Therefore, the score with respect to ψ_x also has the interpretation of the expected value of (14), which is score when x_t is observed, conditional on the past, present and future values of y_t (see Fiorentini, Galesi and Sentana (2014) for further details).

so that the smaller is σ_v^2, the closer $f_{t|\infty}^K$ will be to white noise. In any case, the LM test of $H_0 : \psi_x = 0$ will simply compare the first sample autocovariance of $f_{t|\infty}^K$ with its theoretical value. The advantage of our frequency domain approach is that we obtain those autocovariances without explicitly computing σ_a^2, $\beta_y(L)$ or indeed $f_{t|\infty}^K$.

In turn, under alternative Sma1 the equation for the signal in (16) is replaced by

$$\alpha_x(L)x_t = (1 - \psi_f L)\beta_x(L)f_t,$$

so that the null hypothesis becomes $H_0 : \psi_f = 0$. Then, it is straightforward to prove that this test will numerically coincide with the test of $H_0 : \psi_x = 0$ in view of (17), (12) and (15).

On the other hand, under alternative Nar1 the model will be

$$\left.\begin{aligned} y_t = \mu + x_t + u_t, \\ \alpha_x(L)x_t = \beta_x(L)f_t, \\ (1 - \psi_u L)\alpha_u(L)u_t = \beta_u(L)v_t, \end{aligned}\right\}, \tag{18}$$

while the equation for the non-signal component in (18) will be replaced by

$$\alpha_u(L)u_t = (1 - \psi_v L)\beta_u(L)v_t$$

under alternative Nma1. The exchangeability of signal and non-signal implies that *mutatis mutandis* exactly the same derivations apply to tests of neglected serial correlation in u_t. Finally, joint tests that simultaneously look for neglected serial correlation in the signal and non-signal components can be easily obtained by combining the two scores involved.

3.3 Parameter uncertainty

So far we have implicitly assumed known model parameters. In practice, some of them will have to be estimated under the null. Maximum likelihood estimation of the state space model parameters can be done either in the time domain using the Kalman filter or in the frequency domain.

As we mentioned before, the sampling uncertainty surrounding the sample mean μ is asymptotically inconsequential because the information matrix is block diagonal. The sampling uncertainty surrounding the other parameters, say $\boldsymbol{\vartheta}$, is not necessarily so.

The solution is the standard one: replace the inverse of $\mathcal{I}_{\boldsymbol{\psi\psi}}$, which is the $(\boldsymbol{\psi}, \boldsymbol{\psi})$ block of the information matrix by the $(\boldsymbol{\psi}, \boldsymbol{\psi})$ block of the inverse information matrix $\mathcal{I}^{\boldsymbol{\psi\psi}} = (\mathcal{I}_{\boldsymbol{\psi\psi}} - \mathcal{I}_{\boldsymbol{\psi\vartheta}}\mathcal{I}_{\boldsymbol{\vartheta\vartheta}}^{-1}\mathcal{I}_{\boldsymbol{\vartheta\psi}})^{-1}$ in the quadratic form that defines the LM test. As usual, this is equivalent to orthogonalising the spectral log-likelihood scores corresponding to the parameters in $\boldsymbol{\psi}$ with respect to the scores corresponding to the parameters $\boldsymbol{\vartheta}$ estimated under the null.

4 Potential pitfalls

As we mentioned in section 2.1, we maintain the innocuous assumption that $p_x > q_x$ (when there are no common roots) so that the signal component is a "top-heavy" model. However, by increasing the order of the MA polynomial of the signal, as the SMA1 alternative hypothesis does, the extended UCARIMA model may become underidentified despite the original null model being identified. Assuming no common roots, this will happen when $p_x = q_x + 1$ but $p_u < q_u + 1$ so that the null model is just identified. An important example would be:

$$\left.\begin{array}{r} y_t = x_t + u_t \\ (1-\alpha L)x_t = (1-\psi_f L)f_t \end{array}\right\} \tag{19}$$

with f_t and u_t bivariate white noise orthogonal at all leads and lags. The null hypothesis of interest is $H_0 : \psi_f = 0$, so that the model under the null is a univariate AR(1) + white noise process, while the signal under the alternative is an ARMA(1,1) instead with moving average coefficient ψ_f. In this context, it is possible to formally prove that

Proposition 1 *The score with respect to ψ_f of model (19) reparametrised in terms of $\gamma_{yy}(0)$, $\gamma_{yy}(1)$, α and ψ_f is 0 when $\alpha \neq 0$ regardless of the value of ψ_f.*

Intuitively, the problem is that ψ_f cannot be identified because the reduced form model for the observed series is an ARMA(1,1) fully characterised by its variance, its first autocovariance and α under both the null and the alternative. As a result, the original and extended log-likelihood functions would be identical at their respective optima, which in turn implies that the LR and LM tests will be trivially 0.[9]

A more difficult to detect problem arises when the original model is identified under the null hypothesis and the extended model is identified under the alternative but the information matrix of the extended model is singular under the null. Following Sargan (1983), we shall refer to this situation as a first-order underidentified case because in effect the additional parameter is locally identified but the usual rank condition for identification breaks down.

Although this may seem as a curiosity, it turns out that this problem necessarily occurs with the SAR1 alternative hypothesis whenever alternative SMA1 leads to an underidentified model.

[9] The only possible exception arises when the model is exactly on the boundary of the admissibility region under the null but not under the alternative. However, such anomalies tend to be associated to uninteresting cases. For example, in the AR(1) plus noise model the null parameter configuration will be at the boundary of the admissible parameter space if and only if the non-signal component is identically 0 (see Harvey (1989) and Fiorentini and Planas (2001) for other examples of admissibility restrictions on the model parameters).

Let us study in more detail the $\text{Ar}(1)$ plus white noise example discussed in the previous paragraphs, for which (16) reduces to

$$\left.\begin{array}{c} y_t = x_t + u_t \\ (1-\psi_x L)(1-\alpha L)x_t = f_t \end{array}\right\} \tag{20}$$

with f_t and u_t being bivariate white noise orthogonal at all leads and lags. The null hypothesis of interest is $H_0 : \psi_x = 0$, so that the model under the null is still an $\text{Ar}(1)$ plus white noise, while the signal under the alternative follows an $\text{Ar}(2)$ process. We can then show that[10]

Proposition 2 *The information matrix of model (20) is singular under the null hypothesis $H_0 : \psi_x = 0$.*

As we saw in section 3.1, the intuition is that under the null the score of an additional Ar root is the opposite of the score of an additional Ma root, but the latter is identically 0 at the parameter values estimated for the original $\text{Ar}(1)$ plus white noise model in view of Proposition 1. Therefore, a standard LM test is infeasible. In contrast, there is no linear combination of the first three scores that is equal to 0 under H_0 when $\alpha \neq 0$, so we can consistently estimate α, σ_f^2 and σ_u^2 if we impose the null hypothesis when it is indeed true. Likewise, there is no linear combination of the four scores that is equal to 0 when the true values of α and ψ_x are both different from 0, so again we can consistently estimate σ_f^2, σ_u^2, α and ψ_x in those circumstances, unlike what happened with model (19). For those reasons, it seems intuitive to report instead either a Wald test or a LR one. However, intuitions sometimes prove misleading.

It turns out that one has to be very careful in computing the significance level for the LR test and especially the Wald test because, as we will discuss below, the asymptotic distribution of the ML estimator of ψ_x will be highly unusual under the null. In contrast, there is a readily available LM-type test along the lines of Lee and Chesher (1986). Specifically, these authors propose to replace the usual score test by what they call an "extremum test". Given that the first-order conditions are identically 0, their suggestion is to study the restrictions that the null imposes on higher order conditions. An equivalent procedure to deal with the singularity of the information matrix is to choose a suitable reparametrisation. We follow this second route because it will allow us to obtain asymptotically valid LR and Wald tests too.

Our approach is as follows. First, we replace σ_f^2 and σ_u^2 by $\gamma_{yy}(0)$ and $\gamma_{yy}(1)$, as in Proposition 1. As the following result shows, this change confines the singularity to the last element of the score.

Proposition 3 *The $\psi_x\psi_x$ element of the information matrix of model (20) reparametrised in terms of $\gamma_{yy}(0), \gamma_{yy}(1), \alpha$ and ψ_x is zero under the null hypothesis $H_0 : \psi_x = 0$.*

[10] Harvey (1989) proved the same result in the special case of $\alpha = 1$, which we discuss in detail in appendix A.

Second, we replace ψ_x by either $\sqrt{\varphi}$ (positive root) or $-\sqrt{\varphi}$ (negative root) and retain the value of φ and the sign of the transformation which leads to the largest likelihood function under the alternative. Using the results of Rothitzky et al (2000), we can show that under the null the asymptotic distribution of the ML estimator of φ will be that of a normal variable censored from below at 0. In contrast, the asymptotic distribution of the corresponding estimator of ψ_x will be non-standard, with a faster rate of convergence, half of its density at 0 and the other half equally divided between the positive and negative sides. In this context, the LR test of the null hypothesis $H_0 : \varphi = 0$ will be a 50:50 mixture of a χ^2_0, which is 0 with probability 1, and χ^2_1. As for the Wald test, the square *t-ratio* associated to the ML estimator of φ will share the same asymptotic distribution. In contrast, Wald tests based on ψ_x will have a rather non-standard distribution which will render the *t-ratio* usually reported for this coefficient very misleading.

The following result explains how to conduct the score-type test

Proposition 4 *The extremum test of the null hypothesis $H_0 : \psi_x = 0$ is based on the influence function*

$$\frac{2[\cos(2\lambda) - \alpha\cos(\lambda)]\left(1-\alpha^2\right)\gamma_{yy}(1)}{\alpha\left(1+\alpha^2-2\alpha\cos\lambda\right)g^2_{yy}(\lambda|\gamma_{yy}(0),\gamma_{yy}(1),\alpha,0)}\left[I_{yy}(\lambda) - g_{yy}(\lambda|\gamma_{yy}(0),\gamma_{yy}(1),\alpha,0)\right], \qquad (21)$$

where

$$g_{yy}(\lambda|\gamma_{yy}(0),\gamma_{yy}(1),\alpha,0) = \gamma_{yy}(0) + \frac{2(\cos\lambda - \alpha)}{(1+\alpha^2-2\alpha\cos\lambda)}\gamma_{yy}(1).$$

Given the scores for $\gamma_{yy}(0)$, $\gamma_{yy}(1)$ and α under the null, this means that the extremum test is effectively comparing the *second* sample autocovariance of $f^K_{t|\infty}$ with its theoretical value after taking into account the estimated nature of those model parameters. Nevertheless, the test must be one-sided because (i) $\varphi \geq 0$ under the alternative regardless of whether we reparametrise ψ_x as $\pm\sqrt{\varphi}$ and (ii) the score under the null is the same in both cases, which implies that the Kuhn-Tucker multiplier will also coincide.[11]

Finally, it is worth noting that although ψ_x is not first-order identified because the derivative of the log-likelihood function with respect to this parameter is identically 0 and the expected value of the second derivative under the null is also 0 from Proposition 4, it is locally identified through higher order derivatives.[12]

A somewhat surprising implication of our previous results is that in this instance the usual local equivalence between AR(1) and MA(1) alternatives hypotheses for the signal breaks down. In contrast, there are other seemingly locally equivalent alternatives. Specifically, consider the

[11] When $\alpha = 0$ the test statistic in Proposition 4 breaks down. As Fiorentini and Paruolo (2009) show for the case of observable processes, the distribution of the residual serial correlation in an AR(1) model becomes highly non-standard when the first autocorrelation is in fact 0.

[12] See the proof of Proposition 4 for details.

following variation on model (20):

$$\left.\begin{array}{r} y_t = x_t + u_t \\ (1-\delta_x L^2)(1-\alpha L)x_t = f_t \end{array}\right\}. \tag{22}$$

In this case the null hypothesis of interest is $H_0 : \delta_x = 0$, so that the model under the null is still an AR(1) signal plus white noise, while the signal under the alternative is a "seasonal" AR(3) with restricted autoregressive polynomial $1-\alpha L-\delta_x L^2+\alpha\delta_x L^3$. The "top-heavy" nature of the signal together with the restrictions on the coefficients imply the model under the alternative should be identified. We can then show that

Proposition 5 *The LM test of the null hypothesis $H_0 : \delta_x = 0$ in model (22) will numerically coincide with a two-sided version of the test discussed in Proposition 4 once we correct for the sampling uncertainty in the estimation of the model parameters under the null.*

Nevertheless, such a test is suboptimal for testing the null hypothesis $H_0 : \psi_x = 0$ because it ignores the effective one-sided nature of its alternative.

For reasons analogous to the ones explained in section (3.1), the test in Proposition 5 will also coincide with the LM test of $H_0 : \delta_f = 0$ in the alternative "seasonal" model

$$\left.\begin{array}{r} y_t = x_t + u_t \\ (1-\alpha L)x_t = (1-\delta_f L^2)f_t \end{array}\right\}, \tag{23}$$

which will again be two sided. This equivalence is less obvious than it may seem because the signal follows a "bottom-heavy" process under the alternative. Nevertheless, the fact that the first MA coefficient is 0 is sufficient to guarantee identifiability in this case.

Another seemingly locally equivalent alternative to the neglected AR(1) component in the signal arises when we are interested in testing for first order serial correlation in the non-signal component u_t. In that case the model under the alternative becomes

$$\left.\begin{array}{r} y_t = x_t + u_t \\ (1-\alpha L)x_t = f_t \\ (1-\psi_u L)u_t = v_t \end{array}\right\} \tag{24}$$

with f_t and v_t orthogonal at all leads and lags. The null hypothesis of interest is $H_0 : \psi_u = 0$. In addition, we do not expect any singularity to be present under the alternative, on the grounds that the contemporaneous aggregation of AR(1)+AR(1) is an AR(2,1). We can then show that

Proposition 6 *The LM test of the null hypothesis $H_0 : \psi_u = 0$ in model (24) will numerically coincide with a two-sided version of the test discussed in Proposition 4 once we correct for the sampling uncertainty in the estimation of the model parameters under the null.*

As expected, the LM test of the null hypothesis $H_0 : \psi_v = 0$ in the model

$$\left.\begin{aligned} y_t &= x_t + u_t \\ (1-\alpha L)x_t &= f_t \\ u_t &= (1-\psi_v L)v_t \end{aligned}\right\} \tag{25}$$

will also coincide because the derivatives of $g_{yy}(\lambda)$ with respect to ψ_v in model (25) and with respect to ψ_u in model (24) only differ in their signs.

5 Comparison with tests based on the reduced form residuals

In the context of univariate time series models written in state space form, both Harvey (1989) and Durbin and Koopman (2012) suggest the calculation of neglected serial correlations tests for the reduced form residuals, a_t, which should be white noise under the null of correct dynamic specification. For that reason, it is of some interest to compare such tests to the tests that we have derived in the previous sections. To do so, let us introduce the following two alternative hypothesis of interest:

5. Rar1: Arma(p_x+1, q_x)+Arma(p_u+1, q_u) with a common Ar root.

6. Rma1: Arma(p_x, q_x+1)+Arma(p_u, q_u+1) with a common Ma root.

In this context, we can prove the following result:

Proposition 7 *Testing for* Rar1 *in the* Ucarima *model (1)-(4) is equivalent to testing for* Ar*(1)-type neglected serial correlation in the reduced form innovations, while testing for* Rma1 *in the structural form is the same as testing for* Ma*(1)-type neglected serial correlation in a_t.*

This means that when we test for first order neglected serial correlation in the reduced form residuals the model under the alternative hypothesis is in effect:

$$\left.\begin{aligned} y_t &= \mu + x_t + u_t, \\ \alpha_x(L)(1-\psi_a L)x_t &= \beta_x(L)f_t, \\ \alpha_u(L)(1-\psi_a L)u_t &= \beta_u(L)v_t, \end{aligned}\right\}. \tag{26}$$

In contrast, a test for neglected serial correlation in the signal makes use of the alternative model (16), while a test for neglected serial correlation in the non-signal component relies on (18).

Thus, the relative power of those three tests will depend on the nature of the true model under the alternative. Specifically, if we represent ψ_x on the horizontal axis and ψ_u on the vertical axis, the reduced form test of the null hypothesis $H_0 : \psi_a = 0$ will have maximum power for alternatives along the 45° degree line $\psi_u = \psi_x$ since it is locally the best test of neglected serial correlation in that direction in view of Proposition 7. In contrast, the structural form tests of the null hypotheses $H_0 : \psi_x = 0$ and $H_0 : \psi_u = 0$ will have maximum power along their respective axis (see Demos and Sentana (1998) for a related discussion in the context of ARCH tests). For the intermediate parameter combinations, we could use local power calculations along the lines of appendix B in Fiorentini and Sentana (2014) to compare our LM tests, which are based on the smoothed innovations of the state variables, to the LM tests based on the reduced form innovations.[13] Specifically, we could obtain two *isopower* lines, defined as the locus of points in ψ_x, ψ_u space for which for which the non-centrality parameter of the reduced form test is exactly the same as the non-centrality parameter of the structural tests for $H_0 : \psi_x = 0$ and $H_0 : \psi_u = 0$.

In principle, we could consider the joint test of the compound null hypothesis $H_0 : \psi_x = \psi_u = 0$ mentioned at the end of section 3.2, which will generally have two degrees of freedom instead. For comparing the joint test against the simple tests, though, we would have to equate their local power directly since the number of degrees of freedom would be different.

In view of the discussion in section 4, though, in those situations in which the UCARIMA model becomes underidentified under alternative SMA1, the reduced form test and the two sided versions of the structural tests will be identical.

6 Monte Carlo simulation

6.1 A regular case

We first report the results of some simulation experiments based on a special case of the example discussed at the end of section 3.2, in which the autoregressive polynomial of the signal contains a unit root. In this way, we can assess the finite sample reliability of the size of our proposed tests and their power relative to the reduced form test in a realistic situation in which the model remains identified under each of the four alternatives stated in section 2.1.

[13] Unlike in Fiorentini and Sentana (2014), though, the scores with respect to ψ_x and ψ_u will not be orthogonal to the scores with respect to the remaining structural parameters, $\boldsymbol{\vartheta}$. For that reason, we should conduct the local power calculations with the orthogonalised scores, which are the residuals in the regression of the scores for ψ_x and ψ_u on the scores that define the estimated parameters, with the covariance matrices computed under the null. This procedure would not only reflect the fact that the quadratic form that defines the non-centrality parameter requires the relevant block of the inverse, as opposed to the inverse of the relevant block, but it would also take into account that the expected Jacobian of the other scores with respect to ψ_x and ψ_u will not be 0. Exploiting the information matrix equality, this effectively implies that the non-centrality parameter will be a quadratic form in the direction of departure from the null with a weighting matrix equal to $\mathcal{I}_{\boldsymbol{\psi}\boldsymbol{\psi}} - \mathcal{I}_{\boldsymbol{\psi}\boldsymbol{\vartheta}}\mathcal{I}_{\boldsymbol{\vartheta}\boldsymbol{\vartheta}}^{-1}\mathcal{I}_{\boldsymbol{\vartheta}\boldsymbol{\psi}}$.

6.1.1 Size experiment

To evaluate possible finite sample size distorsions, we generate 10,000 samples of length 200 (equivalent to 50 years of quarterly data) of the following model

$$\left.\begin{aligned} y_t = \mu + x_t + u_t, \\ (1-L)(1-\alpha L)x_t = f_t, \end{aligned}\right\} \tag{27}$$

with f_t and u_t being contemporaneously uncorrelated bivariate Gaussian white noise. Thus, the signal component follows an Ari(1,1) under the null, while the non-signal component is white noise. Given that μ is inconsequential, we fix its true value to 0. We also fix the variance of u_t to 1 without loss of generality. As for the remaining parameters, we choose $\sigma_f^2 = 1$ and $\alpha = .7$ to clearly differentiate this design from the model in section 6.2.

For each simulated sample, we use the first differences of the data to compute the following LM tests:

1. first-order neglected serial correlation in the signal (χ_1^2)

2. first-order neglected serial correlation in the non-signal (χ_1^2)

3. first-order neglected serial correlation in the reduced form residuals (χ_1^2)

4. Joint test of 1. and 2. (χ_2^2)

The finite sample sizes for the four tests are displayed in the first panel of Table 1. As can be seen, the actual rejection rates at the 10, 5 and 1% of all four tests fall within the corresponding asymptotic confidence intervals of (9.41,10.59), (4.57,5.43) and (.80,1.20), so one can reliably use them.

6.1.2 Power experiments

Next, we simulate and estimate 5,000 samples of length 200 of DGPs in which either the signal or the noise may have an additional autoregressive root, with everything else being unchanged. In particular, we consider the following four alternatives:

a. neglected serial correlation in the signal ($\psi_x = .5; \psi_u = 0$), for which the LM test in 1. should be optimal

b. neglected serial correlation in the noise ($\psi_x = 0; \psi_u = .5$), for which the LM test in 2. should be optimal

c. symmetric neglected serial correlation in signal and noise ($\psi_x = .5; \psi_u = .5$), for which the residual LM test in 3. should be optimal

d. asymmetric neglected serial correlation in signal and noise ($\psi_x = .6; \psi_u = .3$) for which the joint LM test in 4. should be optimal.

The raw rejection rates are reported in the last four panels of Table 1. For alternative a., the ranking of the tests is as expected. However, for alternative b. the LM test for signal is able to match the power of the LM test for noise, closely followed by the residual and joint LM tests. Therefore, misspecification in the serial correlation of the non-signal component seems to substantially alter the serial correlation pattern of the filtered values of the correctly specified signal component because the parameter estimators at which the filter is evaluated are biased and the filter weights would be the wrong ones even if we knew the true values of the estimated parameters.

The most surprising result corresponds to alternative c., in that the joint LM test has more power than the asymptotically optimal reduced form test. In contrast, the rejection rates for alternative d. conform to the theoretical predictions.

In summary, our results show that the tests that look for neglected serial correlation in the signal and the noise, either separately or jointly, tend to dominate in terms of power the traditional tests based on the reduced form innovations.

6.2 Local level model

Next we analyze the local level model in appendix A, which is a rather important practical example of the situation discussed in section 4.

6.2.1 Size experiment

To evaluate possible finite sample size distorsions, we generate 10,000 samples of length 200 of the following model

$$\left.\begin{array}{r} y_t = \mu + x_t + u_t, \\ (1-L)x_t = f_t, \end{array}\right\}$$

with f_t and u_t being contemporaneously uncorrelated bivariate Gaussian white noise. As before, we fix the true value of μ to 0 and the variance of u_t to 1 without loss of generality. Therefore, the design depends on a single parameter: the noise to signal ratio σ_f^2, which we choose to be 1. This choice implies a Mean Square Error of the final estimation error of f_t relative to σ_f^2 of 55.28% according to expression (A4), which is neither too low nor too high.

For each simulated sample, we use the first differences of the data to compute the following statistics:

1. one-sided versions of the extremum test for first-order neglected serial correlation in the signal
2. two-sided version of the same test
3. likelihood ratio version
4. Wald test based on φ
5. Wald test based on ψ_x
6. second-order neglected serial correlation in the signal
7. first-order neglected serial correlation in the non-signal
8. first-order neglected serial correlation in the reduced form residuals

As expected from the theoretical results in section 4, the test statistics for 2., 6., 7. and 8. are numerically identical, so we only report one of them under the label LM2S.

It is also important to emphasise that the statistics 3., 4. and 5. require the estimation of model (27). For the reasons described in section 4, this is a non-trivial numerical task because when its true value is 0 (i) approximately half of the ML estimators of ψ_x are identically 0; (ii) the log-likelihood function is extremely flat in a neighbourhood of 0, especially if we parametrise it in terms of ψ_x; and (iii) when the maximum is not 0 it tends to have two commensurate maxima for positive and negative values of ψ_x. To make sure we have obtained the proper ML estimate, we maximise the spectral log-likelihood of model (27) four times: for positive and negative values of ψ_x and with this parameter replaced by $\pm\sqrt{\varphi}$, retaining the maximum maximorum. A kernel density estimate of the mixed-type discrete-continuous distribution of the ML estimators is displayed in Figure 1, with its continuous part scaled so that it integrates to .48, which is the fraction of non-zero estimates of ψ_x. In addition to bimodality, the sampling distribution shows positive skewness, which nevertheless tends to disappear in non-reported experiments with $T = 10,000$. The remaining 52% of the estimates of ψ_x are 0, in which case the test statistics 1., 3., 4. and 5. will all be 0 too.

The rejection rates under the null for the tests at the 10, 5 and 1% are displayed in Table 2. The only procedure which seems to have a reliable size is the two-sided LM test. In contrast, its one-sided version is somewhat conservative, while the LR and especially the two Wald tests are liberal. Reassuringly, though, the size distortions of the one-sided LM test disappear fairly quickly in non-reported experiments with larger sample sizes, while the distortions of the LR and Wald tests for φ go down more slowly and are still noticeable even in samples as big as $T = 50,000$ despite the fact that the fraction of 0 estimates converges very quickly to 1/2. As expected, though, the distortions of the Wald test based on ψ_x persist no matter how big the sample size is because the information matrix for this parametrisation is singular.

6.2.2 Power experiments

Next, we simulate and estimate 5,000 samples of length 200 of four alternative DGPs analogous to the ones described in a.-d. of the previous section. However, since our focus is on tests of the null hypothesis $H_0 : \psi_x = 0$, we only estimate the model under the null and under the a. alternative. In this regard, an additional issue that we encounter in some desgins is that from time to time the estimated value of σ_u^2 is 0. In those "pile-up" cases we compute the LM and Wald tests excluding the corresponding row and column of the information matrix.

In view of the substantial size distortions under the null, we report not only raw rejection rates based on asymptotic critical values in Table 3a but also size-adjusted ones in Table 3b, which exploit the Monte Carlo critical values obtained in the simuation described in the previous subsection. If we focus on this second table, we can conclude that the tests that explicitly acknowledge the implicit one-sided nature of the alternative to $H_0 : \psi_x = 0$ dominate the two-sided test, except when $\psi_x = 0$ but $\psi_u =.5$, when they tend to be equally powerful. In particular, the one-sided tests for $H_0 : \psi_x = 0$ dominate the residual correlation tests even when $\psi_x = \psi_u = .5$.

We can also conclude that the relative ranking of the extremum test, the likelihood ratio test and the Wald test for $H_0 : \varphi = 0$ depends on the DGP, although when it coincides with the alternative for which they are asymptotically optimal, the extremum test dominates the LR test, which in turn dominates the Wald test.

7 Conclusions and extensions

We have derived computationally simple and intuitive expressions for score tests of neglected serial correlation in unobserved component univariate models using frequency domain methods. Our tests can focus on the state variables individually or jointly. The implicit orthogonality conditions are analogous to the conditions obtained by treating the Wiener-Kolmogorov-Kalman smoothed estimators of the innovations in the latent variables as if they were observed, but they account for their final estimation errors.

In some common situations in which the information matrix is singular under the null we show that contrary to popular belief it is possible to derive extremum tests that are asymptotically equivalent to likelihood ratio tests, which become one-sided. We also explain how to compute asymptotically reliable Wald tests. As a result, from now on empirical researchers would be able to report test statistics in those irregular situations too. Further, we explicitly relate the incidence of those problems to the model identification conditions and compare our tests with tests based on the reduced form prediction errors.

We conduct Monte Carlo exercises to study the finite sample reliability and power of our proposed tests. In the regular case of a latent Ari(1,1) process cloaked in white noise, our results show that the finite sample size of the different tests is reliable. They also imply that the tests that look for neglected serial correlation in the signal and the noise, either separately or jointly, dominate in terms of power the traditional tests based on the reduced form innovations.

When we look at neglected serial correlation tests in the irregular local level model, our simulation results confirm that the finite sample distribution of the ML estimator of the additional autoregressive root in the signal is highly unusual under the null of correct specification, with almost half its mass at 0 and two modes, one positive and one negative. Not suprisingly, a Wald test based on this parameter is highly unreliable, even asymptotically. We also find some size distortions for the asymptotically valid one-sided tests of $H_0 : \psi_x = 0$ (but not for the two-sided LM test), which nevertheless progressively disappear as the sample size increases. After correcting for those distortions, though, we find that the one-sided tests dominate the residual correlation tests even when $\psi_x = \psi_u = .5$, but the relative ranking of the extremum test, the likelihood ratio test and the Wald test depends on the DGP under the alternative.

Although we have considered reasonable Monte Carlo designs, a more throrough analysis of the determinants of the size and power properties of the different tests would constitute a valuable addition.

The testing procedures we have developed can be extended in several interesting directions. First, it would be tedious but straightforward to consider models with more than two components. More interestingly, we could also consider models with purely seasonal components (see Harvey (1989) for some examples). Tests of higher order serial correlation also deserve further consideration since they might involve singularity problems too. For example, the Ari(1,1) plus white noise process discussed in section 6.1, which yields standard test statistics for neglected first order serial correlation, gives rise to a singular information matrix when we consider tests against first and second order serial correlation simultaneously because those tests are numerically equivalent to tests against the underidentified alternative of Arima(1,1,2) plus white noise.

Second, we have assumed throughtout the paper that the model estimated under the null is parametrically identified. Nevertheless, Harvey (1989) discusses some examples in which an UCARIMA model is underidentified under the null but identified under the alternative. He formally tackles the problem by using the procedure proposed by Aitchinson and Silvey (1960), which effectively adds a matrix to the information matrix to make sure that it has full rank (see also Breusch (1986)).

We have also maintained the assumption of normality. To understand its implications, let $\mu_{t|t-1}$ and $\sigma^2_{t|t-1}$ denote the conditional mean and variance of y_t given its past alone, which can be obtained from the prediction equations of the Kalman filter. Given that the serial correlation parameters effectively enter through $\mu_{t|t-1}$ only, the information matrix equality should continue to hold for their scores.

Although we have only considered unobserved components with rational spectral densities, in principle our methods could be applied to long memory processes too. In this regard, it would be worth exploring the fractionally integrated alternatives considered by Robinson (1994). More generally, it would also be interesting to consider non-parametric alternatives such as the ones studied by Hong (1996), in which the lag length is implicitly determined by the choice of bandwidth parameter in a kernel-based estimator of a spectral density matrix. Another potential extension would directly deal with non-stationary models without transforming the observed variables to achieve stationarity. All these topics constitute fruitful avenues for future research.

References

Aitchison, J. and Silvey, S. D. (1960): "Maximum-likelihood estimation procedures and associated tests of significance", *Journal of the Royal Statistical Society Series B* 22, 154–71.

Andrews, D.W.K. and Ploberger, W. (1996): "Testing for serial correlation against an ARMA(1,1) process", *Journal of the American Statistical Association* 91, 1331–1342.

Barndorff-Nielsen, O. and Schou, G. (1973): "On the parametrization of autoregressive models by partial autocorrelations", *Journal of Multivariate Analysis* 3, 408–419.

Box, G.E.P., Hillmer, S. and Tiao, G.C. (1978): "Analysis and modeling of seasonal time series", in A. Zellner, ed., *Seasonal analysis of economic time series*, U.S. Department of Commerce, Bureau of the Census, Washington D.C., 309-344.

Breusch, T. S. (1978): "Testing for autocorrelation in dynamic linear models", *Australian Economic Papers* 17, 334–355.

Breusch, T.S. (1986): "Hypothesis testing in unidentified models", *Review of Economic Studies* 53, 635-651.

Breusch, T. S. and Pagan, A.R. (1980): "The Lagrange Multiplier test and its applications to model specification in econometrics", *Review of Economic Studies* 47, 239-253.

Burman, J. P. (1980): "Seasonal adjustment by signal extraction", *Journal of the Royal Statistical Society*, Ser. A 143, 321–337.

Choudhuri , N., Ghosal S. and Roy, A. (2004): "Contiguity of the Whittle measure for a Gaussian time series", *Biometrika* 91, 211-218.

Commandeur, J.J.F., Koopman, S.J. and Ooms, M. (2011): "Statistical software for state space methods", *Journal of Statistical Software* 41, 1-18

Demos, A. and Sentana, E. (1998): "Testing for GARCH effects: a one-sided approach", *Journal of Econometrics* 86, 97-127.

Dunsmuir, W. (1979): "A central limit theorem for parameter estimation in stationary vector time series and its application to models for a signal oberved with noise", *Annals of Statistics* 7, 490-506.

Dunsmuir, W. and Hannan, E.J. (1976): "Vector linear time series models", *Advances in Applied Probability* 8, 339-364.

Durbin, J. and Koopman, S.J. (2012): *Time series analysis by state space methods*, 2nd ed., Oxford University Press, Oxford.

Engle, R. F. (1978): "Estimating structural models of seasonality", in A. Zellner, ed., *Seasonal analysis of economic time series*, U.S. Department of Commerce, Bureau of the Census, Washington D.C., 281–297.

Engle, R.F. and Watson, M.W. (1980): "Formulation générale et estimation de models multidimensionnels temporels a facteurs explicatifs non observables", *Cahiers du Séminaire d'Économétrie* 22, 109-125.

Fernández, F.J. (1990): "Estimation and testing of a multivariate exponential smoothing model", *Journal of Time Series Analysis* 11, 89–105.

Fiorentini, G. (1995): *Conditional heteroskedasticity: some results on estimation, inference and signal extraction, with an application to seasonal adjustment*, unpublished Doctoral Dissertation, European University Institute.

Fiorentini, G., Galesi, A. and Sentana, E. (2014): "A spectral EM algorithm for dynamic factor models", mimeo, CEMFI.

Fiorentini, G. and Planas, C. (1998): "From autocovariances to moving average: an algorithm comparison", *Computational Statistics* 13, 477-484.

Fiorentini, G. and Planas, C. (2001): "Overcoming nonadmissibility in ARIMA-model-based signal extraction", *Journal of Business and Economic Statistics*, 19, 455-64.

Fiorentini, G. and Sentana, E. (2013): "Dynamic specification tests for dynamic factor models", CEMFI Working Paper 1306.

Fiorentini, G. and Sentana, E. (2014): "Tests for serial dependence in static, non-Gaussian factor models", forthcoming in S.J. Koopman and N. Shephard (eds.) *Unobserved components and time series econometrics*, Oxford University Press.

Godfrey, L. G. (1978a): "Testing against general autoregressive and moving average error models when the regressors include lagged dependent variables", *Econometrica* 46, 1293–1301.

Godfrey, L. G. (1978b): "Testing for higher order serial correlation in regression equations when the regressors include lagged dependent variables", *Econometrica* 46, 1303–10.

Godfrey, L.G. (1988): *Misspecification tests in econometrics: the Lagrange multiplier principle and other approaches*, Econometric Society Monographs, Cambridge University Press, Cambridge.

Gómez, V. (1999): "Three equivalent methods for filtering finite nonstationary time series", *Journal of Business and Economic Statistics* 17, 109-116.

Granger C. W. J. and M.J. Morris (1976): "Time series modelling and interpretation", *Journal of the Royal Statistical Society* Series A 139, 246–257.

Hannan, E.J. (1973): "The asymptotic theory of linear time series models", *Journal of Applied Probability* 10, 130-145 (Corrigendum 913).

Harvey, A.C. (1989): *Forecasting, structural models and the Kalman filter*, Cambridge University Press, Cambridge.

Hong, Y. (1996): "Consistent testing for serial correlation of unknown form" *Econometrica* 64, 837-864.

Hotta, L. K. (1989): "Identification of unobserved components models", *Journal of Time Series Analysis* 10, 259-270.

Lee, L. F., and A. Chesher (1986): "Specification testing when score test statistics are identically zero", *Journal of Econometrics* 31, 121–149.

Levinson, N. (1947): "The Wiener RMS error criterion in filter design and prediction", *Journal of Mathematics and Physics* 25, 261-278.

Maravall, A. (1979): *Identification in dynamic shock-error models*, Springer-Verlag, Berlin.

Maravall, A. (1999): "Unobserved components in economic time series" in M.H. Pesaran and M.R. Wickens (eds.) *Handbook of Applied Econometrics. Volume I: Macroeconomics*, Blackwell.

Nerlove, M. (1967): "Distributed lags and unobserved components of economic time series", in W. E. A. Fellner, ed., *Ten economic essays in the tradition of Irving Fisher*, Wiley: New York, 126–169.

Nerlove, M., Grether, D. & Carvalho, J. (1979): *Analysis of economic time series, a synthesis*, Academic Press: New York.

Newey, W.K. (1985): "Maximum likelihood specification testing and conditional moment tests", *Econometrica* 53, 1047-70.

Pierce, D.A. (1978): "Seasonal adjustment when both deterministic and stochastic seasonality are present", in A. Zellner, ed., *Seasonal analysis of economic time series*, U.S. Department of Commerce, Bureau of the Census, Washington D.C., 242-269.

Priestley, M.B. (1981): *Spectral analysis and time series* vol. I and II, Academic Press.

Robinson, P.M. (1994): "Efficient tests of nonstationary hypothesis", *Journal of the American Statistical Association* 89, 1420-1437.

Rotnitzky, A., Cox, D. R., Bottai, M. and Robins, J. (2000): "Likelihood-based inference with singular information matrix", *Bernoulli* 6, 243-284.

Sargan, J.D. (1983): "Identification and lack of identification" *Econometrica* 51, 1605-1634.

Tauchen, G. (1985): "Diagnostic testing and evaluation of maximum likelihood models", *Journal of Econometrics* 30, 415-443.

Whittle, P. (1962): "Gaussian estimation in stationary time series", *Bulletin of the International Statistical Institute* 39, 105-129.

Appendices

A Local level model

A.1 Testing for neglected serial correlation in the trend

A.1.1 Against AR(1) alternatives

Consider the following modified version of model (20)

$$\left.\begin{array}{r} y_t = x_t + u_t \\ (1-\psi_x L)(1-L)x_t = f_t \end{array}\right\}, \tag{A1}$$

with f_t and u_t orthogonal at all leads and lags. The main difference is that we have replaced the covariance stationarity hypotheis for the signal x_t by a unit root one. As before, the null hypothesis of interest remains $H_0 : \psi_x = 0$, so that the model under the null is simply a random walk signal plus white noise, while the signal under the alternative is an Ari(1,1) with autoregressive coefficient ψ_x.

In order to use spectral methods we need to take first differences of the observed variable to make it stationary, which yields

$$\Delta y_t = \frac{1}{1-\psi_x L} f_t + (1-L)u_t.$$

Hence, it is easy to see that

$$V(\Delta y_t) = \gamma_{\Delta y \Delta y}(0) = \frac{\sigma_f^2}{1-\psi_x^2} + 2\sigma_u^2, \tag{A2}$$

$$cov(\Delta y_t, \Delta y_{t-1}) = \gamma_{\Delta y \Delta y}(1) = \psi_x \frac{\sigma_f^2}{1-\psi_x^2} - \sigma_u^2, \tag{A3}$$

$$cov(\Delta y_t, \Delta y_{t-j}) = \gamma_{\Delta y \Delta y}(j) = \psi_x^j \frac{\sigma_f^2}{1-\psi_x^2} \quad j \geq 2.$$

Similarly, the spectral density of Δy_t will be

$$\begin{aligned} g_{\Delta y \Delta y}(\lambda) &= \frac{\sigma_f^2}{(1-\psi_x e^{-i\lambda})(1-\psi_x e^{i\lambda})} + (1-e^{-i\lambda})(1-e^{i\lambda})\sigma_u^2 \\ &= \frac{\sigma_f^2}{1+\psi_x^2 - 2\psi_x \cos\lambda} + 2(1-\cos\lambda)\sigma_u^2. \end{aligned}$$

The reduced form of Δy_t is an Ima(1,1) process with Ma coefficient β_y given by

$$\beta_y = \frac{1}{2}\left(\sqrt{q^2+4q} - 2 - q\right),$$

where $q = \sigma_f^2/\sigma_u^2$ denotes the signal to noise ratio, and residual variance

$$\sigma_a^2 = -\sigma_u^2/\beta_y.$$

As is well known (see e.g. Priestley (1981, section 10.3), the variance of the final estimation error of f_t will be given by

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}\left(g_{ff}(\lambda)-\frac{\left|g_{f\Delta_y}(\lambda)\right|^2}{g_{\Delta_y\Delta_y}(\lambda)}\right)d\lambda=\frac{1}{2\pi}\sigma_f^2\int_{-\pi}^{\pi}\left(1-\frac{q}{q+2(1-\cos\lambda)}\right)d\lambda=\sigma_f^2\left(1-\frac{q}{\sqrt{q^2+4q}}\right) \tag{A4}$$

because

$$\int\frac{q}{q+2(1-\cos\lambda)}d\lambda=\frac{2\sqrt{q}}{\sqrt{q+4}}\arctan\left(\frac{\sqrt{q+4}}{\sqrt{q}}\tan\left(\frac{\lambda}{2}\right)\right)$$

and

$$\lim_{\lambda\to(\pi/2)^-}\arctan\left(\frac{\sqrt{q+4}}{\sqrt{q}}\tan\left(\frac{\lambda}{2}\right)\right)-\lim_{\lambda\to(-\pi/2)^+}\arctan\left(\frac{\sqrt{q+4}}{\sqrt{q}}\tan\left(\frac{\lambda}{2}\right)\right)=\pi$$

Interestingly, we would obtain exactly the same expression by working with pseudo-spectral densities in levels because

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}\left(g_{xx}(\lambda)-\frac{\left|g_{x\Delta_y}(\lambda)\right|^2}{g_{yy}(\lambda)}\right)d\lambda=\frac{1}{2\pi}\int_{-\pi}^{\pi}\left(\frac{\sigma_f^2}{2(1-\cos\lambda)}-\frac{\left(\frac{\sigma_f^2}{2(1-\cos\lambda)}\right)^2}{\frac{\sigma_f^2}{2(1-\cos\lambda)}+\sigma_u^2}\right)d\lambda$$

$$=\frac{1}{2\pi}\int_{-\pi}^{\pi}\left(\frac{\frac{\sigma_f^2\sigma_u^2}{2(1-\cos\lambda)}}{\frac{\sigma_f^2}{2(1-\cos\lambda)}+\sigma_u^2}\right)d\lambda=\frac{1}{2\pi}\sigma_f^2\int_{-\pi}^{\pi}\left(1-\frac{q}{q+2(1-\cos\lambda)}\right)d\lambda=\sigma_f^2\left(1-\frac{q}{\sqrt{q^2+4q}}\right).$$

The partial derivatives of this spectral density are

$$\begin{aligned}
\frac{\partial g_{\Delta y\Delta y}(\lambda)}{\partial\sigma_f^2} &= \frac{1}{1+\psi_x^2-2\psi_x\cos\lambda},\\
\frac{\partial g_{\Delta y\Delta y}(\lambda)}{\partial\sigma_u^2} &= 2(1-\cos\lambda),\\
\frac{\partial g_{\Delta y\Delta y}(\lambda)}{\partial\psi_x} &= \frac{2\sigma_f^2(\cos\lambda-\psi_x)}{(1+\psi_x^2-2\psi_x\cos\lambda)^2}.
\end{aligned}$$

Under the the null of $H_0:\psi_x=0$ those derivatives become

$$\begin{aligned}
\frac{\partial g_{\Delta y\Delta y}(\lambda)}{\partial\sigma_f^2} &= 1,\\
\frac{\partial g_{\Delta y\Delta y}(\lambda)}{\partial\sigma_u^2} &= 2(1-\cos\lambda),\\
\frac{\partial g_{\Delta y\Delta y}(\lambda)}{\partial\psi_x} &= 2\sigma_f^2\cos\lambda,
\end{aligned}$$

which implies that

$$\sigma_f^2\left[\frac{\partial g_{\Delta y\Delta y}(\lambda)}{\partial\sigma_u^2}-2\frac{\partial g_{\Delta y\Delta y}(\lambda)}{\partial\sigma_f^2}\right]+\frac{\partial g_{\Delta y\Delta y}(\lambda)}{\partial\psi_x}=0 \tag{A5}$$

for all λ. Obviously, exactly the same linear combination of the elements of $g_{yy}^{-1}(\lambda)\partial g_{yy}(\lambda)/\partial\boldsymbol{\theta}$ will be singular too. Therefore, the information matrix of the model, which is given by

$$\int_{-\pi}^{\pi}\frac{\partial g_{yy}(\lambda)}{\partial\boldsymbol{\theta}}\frac{1}{g_{yy}(\lambda)}\frac{1}{g_{yy}(\lambda)}\frac{\partial g_{yy}(\lambda)}{\partial\boldsymbol{\theta}'}d\lambda,$$

will only have rank 3 under the null. In view of this result, Harvey (1989) rightly concludes that a standard LM test is infeasible.

In contrast, there is no linear combination of the first two derivatives that is equal to 0 under H_0, so we can consistently estimate σ_f^2 and σ_u^2 if we impose the null hypothesis when it is indeed true. Likewise, there is no linear combination of the three derivatives that is equal to 0 under the alternative either, so again we can consistently estimate σ_f^2, σ_u^2 and ψ_x in those circumstances. For that reason, Harvey (1989) recommends reporting either a Wald test or a LR one, which for reasons explained in section 4 turns out not to be sound advice.

Nevertheless, an LM-type test is readily available once more along the same lines as in section 4. Specifically, we can tackle the problem created by (A5) by reparametrisation. First, we are going to replace σ_f^2 and σ_u^2 by $\gamma_{\Delta y\Delta y}(0)$ and $\gamma_{\Delta y\Delta y}(1)$. Thus, it is easy to see from (A2) and (A3) that

$$\begin{aligned}\sigma_u^2 &= \frac{1}{2\psi_x+1}\left[\psi_x\gamma_{\Delta y\Delta y}(0)-\gamma_{\Delta y\Delta y}(1)\right],\\ \sigma_f^2 &= \frac{1-\psi_x^2}{2\psi_x+1}\left[\gamma_{\Delta y\Delta y}(0)+2\gamma_{\Delta y\Delta y}(1)\right],\end{aligned}$$

which are well defined as long as $\psi_x \neq -\frac{1}{2}$ (or if $\gamma_{\Delta y\Delta y}(0)+2\gamma_{\Delta y\Delta y}(1)=0$ when $\psi_x \neq -\frac{1}{2}$).

With this notation, the spectral density becomes

$$\begin{aligned}g_{\Delta y\Delta y}(\lambda) \;=\;& \frac{1}{1+\psi_x^2-2\psi_x\cos\lambda}\frac{1-\psi_x^2}{2\psi_x+1}\left[\gamma_{\Delta y\Delta y}(0)+2\gamma_{\Delta y\Delta y}(1)\right]\\ &+2(1-\cos\lambda)\frac{1}{2\psi_x+1}\left[\psi_x\gamma_{\Delta y\Delta y}(0)-\gamma_{\Delta y\Delta y}(1)\right].\end{aligned}$$

The derivatives with respect to these new parameters are

$$\begin{aligned}\frac{\partial g_{\Delta y\Delta y}(\lambda)}{\partial\gamma_{\Delta y\Delta y}(0)} &= \frac{1}{1+\psi_x^2-2\psi_x\cos\lambda}\frac{1-\psi_x^2}{2\psi_x+1}+2(1-\cos\lambda)\frac{\psi_x}{2\psi_x+1}\\ \frac{\partial g_{\Delta y\Delta y}(\lambda)}{\partial\gamma_{\Delta y\Delta y}(1)} &= \frac{2}{1+\psi_x^2-2\psi_x\cos\lambda}\frac{1-\psi_x^2}{2\psi_x+1}-2(1-\cos\lambda)\frac{1}{2\psi_x+1}\\ \frac{\partial g_{\Delta y\Delta y}(\lambda)}{\partial\psi_x} &= -2\psi_x\frac{\left[\gamma_{\Delta y\Delta y}(0)+2\gamma_{\Delta y\Delta y}(1)\right]}{(2\psi_x+1)^2\left(1+\psi_x^2-2\psi_x\cos\lambda\right)^2}\\ &\times\left[\begin{array}{c}(\cos\lambda-2)\,\psi_x^3+\left(4\cos\lambda-4\cos^2\lambda\right)\psi_x^2\\ +\left(4\cos^3\lambda-4\cos^2\lambda+\cos\lambda+2\right)\psi_x+\left(2-4\cos^2\lambda\right)\end{array}\right].\end{aligned}$$

Under the null of $H_0 : \psi_x = 0$, these scores reduce to

$$\begin{aligned} \frac{\partial g_{\Delta y \Delta y}(\lambda)}{\partial \gamma_{\Delta y \Delta y}(0)} &= 1 \\ \frac{\partial g_{\Delta y \Delta y}(\lambda)}{\partial \gamma_{\Delta y \Delta y}(1)} &= 2\cos\lambda \\ \frac{\partial g_{\Delta y \Delta y}(\lambda)}{\partial \psi_x} &= 0. \end{aligned}$$

Although we have not yet eliminated the singularity, we have at least confined it to the last element of the score. If we further reparametrise ψ_x as $\pm\sqrt{\varphi}$, the spectral density becomes

$$\begin{aligned} g_{\Delta y \Delta y}(\lambda) &= \frac{1}{1+\varphi-2\sqrt{\varphi}\cos\lambda}\frac{1-\varphi}{2\sqrt{\varphi}+1}\left[\gamma_{\Delta y \Delta y}(0)+2\gamma_{\Delta y \Delta y}(1)\right] \\ &\quad +2(1-\cos\lambda)\frac{1}{2\sqrt{\varphi}+1}\left[\sqrt{\varphi}\gamma_{\Delta y \Delta y}(0)-\gamma_{\Delta y \Delta y}(1)\right]. \end{aligned}$$

Tedious algebra shows that the $\partial g_{\Delta y \Delta y}(\lambda)/\partial\varphi$ evaluated at $\varphi = 0$ will be equal to

$$2\sigma_f^2 \cos 2\lambda,$$

where we have used the fact that

$$\gamma_{\Delta y \Delta y}(0) + 2\gamma_{\Delta y \Delta y}(1) = \sigma_f^2$$

under the null. Hence, the extremum test for ψ_x, which coincides with the LM test for φ, is going to be based on the second autocovariance of the smoothed estimates of f_t. Importantly, Lee and Chesher (1986) show that the one-sided version of this extremum test continues to be asymptotically equivalent to both the LR and a one-sided version of the Wald test for φ.

A.1.2 Against MA(1) alternatives

Consider now the following variation on model (A1):

$$\left.\begin{aligned} y_t = x_t + u_t \\ (1-L)x_t = (1-\psi_f)f_t \end{aligned}\right\}, \tag{A6}$$

with f_t and u_t orthogonal at all leads and lags. The null hypothesis of interest is $H_0 : \psi_f = 0$, so that the model under the null is still a random walk signal plus white noise, while the signal under the alternative is an IMA(1,1) with moving average coefficient ψ_f.

In this case, the stationary model is

$$\Delta y_t = (1-\psi_f L)w_t + (1-L)u_t.$$

Hence, it is easy to see that

$$
\begin{aligned}
V(\Delta y_t) &= \gamma_{\Delta y \Delta y}(0) = (1+\psi_f^2)\sigma_f^2 + 2\sigma_u^2, && \text{(A7)}\\
cov(\Delta y_t, \Delta y_{t-1}) &= \gamma_{\Delta y \Delta y}(1) = -\psi_f \sigma_f^2 - \sigma_u^2, && \text{(A8)}\\
cov(\Delta y_t, \Delta y_{t-j}) &= \gamma_{\Delta y \Delta y}(j) = 0 \quad j \geq 2.
\end{aligned}
$$

Similarly, the spectral density of Δy_t will be

$$
\begin{aligned}
g_{\Delta y \Delta y}(\lambda) &= \left(1-\psi_f e^{-i\lambda}\right)\left(1-\psi_f e^{i\lambda}\right)\sigma_f^2 + (1-e^{-i\lambda})(1-e^{i\lambda})\sigma_u^2\\
&= (1+\psi_f^2 - 2\psi_f \cos\lambda)\sigma_f^2 + 2(1-\cos\lambda)\sigma_u^2.
\end{aligned}
$$

The partial derivatives of this spectral density are

$$
\begin{aligned}
\frac{\partial g_{\Delta y \Delta y}(\lambda)}{\partial \sigma_f^2} &= 1+\psi_f^2 - 2\psi_f\cos\lambda,\\
\frac{\partial g_{\Delta y \Delta y}(\lambda)}{\partial \sigma_u^2} &= 2(1-\cos\lambda),\\
\frac{\partial g_{\Delta y \Delta y}(\lambda)}{\partial \psi_f} &= 2\sigma_f^2(\cos\lambda - \psi_f).
\end{aligned}
$$

Under the the null of $H_0 : \psi_f = 0$ those derivatives become

$$
\begin{aligned}
\frac{\partial g_{\Delta y \Delta y}(\lambda)}{\partial \sigma_f^2} &= 1,\\
\frac{\partial g_{\Delta y \Delta y}(\lambda)}{\partial \sigma_u^2} &= 2(1-\cos\lambda),\\
\frac{\partial g_{\Delta y \Delta y}(\lambda)}{\partial \psi_f} &= 2\sigma_f^2\cos\lambda,
\end{aligned}
$$

which confirms that (A5) also holds for this model.

Let us now try and isolate the singularity in a single parameter by using the same procedure as in the previous section. First, we replace σ_f^2 and σ_u^2 by $\gamma_{\Delta y \Delta y}(0)$ and $\gamma_{\Delta y \Delta y}(1)$. Thus, it is easy to see from (A7) and (A8) that

$$
\begin{aligned}
\sigma_u^2 &= \frac{1}{(1-\psi_f)^2}\left[-\psi_f\gamma_{\Delta y \Delta y}(0) - (1+\psi_f^2)\gamma_{\Delta y \Delta y}(1)\right],\\
\sigma_f^2 &= \frac{1}{(1-\psi_f)^2}\left[\gamma_{\Delta y \Delta y}(0) + 2\gamma_{\Delta y \Delta y}(1)\right],
\end{aligned}
$$

which are well defined as long as $\psi_f \neq 1$.

With this notation, the spectral density becomes

$$
\begin{aligned}
g_{\Delta y \Delta y}(\lambda) &= \frac{(1+\psi_f^2 - 2\psi_f\cos\lambda)}{(1-\psi_f)^2}\left[\gamma_{\Delta y \Delta y}(0) + 2\gamma_{\Delta y \Delta y}(1)\right]\\
&\quad + 2(1-\cos\lambda)\frac{1}{(1-\psi_f)^2}\left[-\psi_f\gamma_{\Delta y \Delta y}(0) - (1+\psi_f^2)\gamma_{\Delta y \Delta y}(1)\right].
\end{aligned}
$$

The derivatives with respect to these new parameters are

$$\frac{\partial g_{\Delta y \Delta y}(\lambda)}{\partial \gamma_{\Delta y \Delta y}(0)} = \frac{(1+\psi_f^2 - 2\psi_f \cos\lambda)}{(1+\psi_f)^2} - 2(1-\cos\lambda)\frac{\psi_f}{(1-\psi_f)^2}$$
$$\frac{\partial g_{\Delta y \Delta y}(\lambda)}{\partial \gamma_{\Delta y \Delta y}(1)} = \frac{2(1+\psi_f^2 - 2\psi_f \cos\lambda)}{(1+\psi_f)^2} - 2(1-\cos\lambda)\frac{(1+\psi_f^2)}{(1-\psi_f)^2}$$
$$\frac{\partial g_{\Delta y \Delta y}(\lambda)}{\partial \psi_f} = 0.$$

Since this last derivative is 0 not only under the null but also under the alternative, ψ_f cannot be identified. Intuitively, the reason is that the process for Δy_t is an unrestricted Ma(1) under the alternative, which is fully characterised by $\gamma_{\Delta y \Delta y}(0)$ and $\gamma_{\Delta y \Delta y}(1)$.

Thus, the usual local equivalence between Ar(1) and Ma(1) alternatives hypothesis for the signal breaks down once again.

A.1.3 Against restricted MA(2) alternatives

Consider this alternative variation on model (A1):

$$\left.\begin{array}{r} y_t = x_t + u_t \\ (1-L)x_t = (1-\delta_f L^2) f_t \end{array}\right\}, \tag{A9}$$

with u_t and w_t orthogonal at all leads and lags. The null hypothesis of interest is $H_0 : \delta_f = 0$, so that the model under the null is still a random walk signal plus white noise, while the signal under the alternative is an Ima(1,2) with second moving average coefficient δ_f.

Therefore, the stationary model will be

$$\Delta y_t = (1-\delta_f L^2)w_t + (1-L)u_t,$$

whose spectral density is

$$\begin{aligned} g_{\Delta y \Delta y}(\lambda) &= \left(1-\delta_f e^{-i2\lambda}\right)\left(1-\delta_f e^{i2\lambda}\right)\sigma_f^2 + (1-e^{-i\lambda})(1-e^{i\lambda})\sigma_u^2 \\ &= (1+\delta_f^2 - 2\delta_f \cos 2\lambda)\sigma_f^2 + 2(1-\cos\lambda)\sigma_u^2. \end{aligned}$$

The partial derivatives of this spectral density are

$$\begin{aligned} \frac{\partial g_{\Delta y \Delta y}(\lambda)}{\partial \sigma_f^2} &= 1+\delta_f^2 - 2\delta_f \cos 2\lambda, \\ \frac{\partial g_{\Delta y \Delta y}(\lambda)}{\partial \sigma_u^2} &= 2(1-\cos\lambda), \\ \frac{\partial g_{\Delta y \Delta y}(\lambda)}{\partial \delta_f} &= 2\sigma_f^2(\cos 2\lambda - \delta_f). \end{aligned}$$

Under the the null of $H_0 : \delta_f = 0$ those derivatives become

$$\begin{aligned}
\frac{\partial g_{\Delta y \Delta y}(\lambda)}{\partial \sigma_f^2} &= 1, \\
\frac{\partial g_{\Delta y \Delta y}(\lambda)}{\partial \sigma_u^2} &= 2(1 - \cos\lambda), \\
\frac{\partial g_{\Delta y \Delta y}(\lambda)}{\partial \delta_f} &= 2\sigma_f^2 \cos 2\lambda.
\end{aligned}$$

Given that the linear span of $\partial g_{\Delta y \Delta y}(\lambda)/\partial \sigma_f^2$ and $\partial g_{\Delta y \Delta y}(\lambda)/\partial \sigma_u^2$ is the same as the linear span of $\partial g_{\Delta y \Delta y}(\lambda)/\partial \gamma_{\Delta y \Delta y}(0)$ and $\partial g_{\Delta y \Delta y}(\lambda)/\partial \gamma_{\Delta y \Delta y}(1)$, this test is going to coincide with the two-sided version of the extremum test against an AR(1) alternative.

A.1.4 Against restricted AR(2) alternatives

Consider yet another variation on model (A1):

$$\left.\begin{array}{c} y_t = x_t + u_t \\ (1 - \delta_x L^2)(1 - L)x_t = f_t \end{array}\right\}, \tag{A10}$$

with f_t and u_t orthogonal at all leads and lags. The null hypothesis of interest is $H_0 : \delta_x = 0$, so that the model under the null is still a random walk signal plus white noise, while the signal under the alternative is an ARI(2,1) with second autoregressive coefficient δ_x.

In this case, the spectral density of Δy_t will be

$$\begin{aligned}
g_{\Delta y \Delta y}(\lambda) &= \frac{\sigma_f^2}{(1 - \delta_x e^{-i2\lambda})(1 - \delta_x e^{i2\lambda})} + (1 - e^{-i\lambda})(1 - e^{i\lambda})\sigma_u^2 \\
&= \frac{\sigma_f^2}{1 + \delta_x^2 - 2\delta_x \cos 2\lambda} + 2(1 - \cos\lambda)\sigma_u^2.
\end{aligned}$$

The partial derivatives of this spectral density are

$$\begin{aligned}
\frac{\partial g_{\Delta y \Delta y}(\lambda)}{\partial \sigma_f^2} &= \frac{1}{1 + \delta_x^2 - 2\delta_x \cos 2\lambda}, \\
\frac{\partial g_{\Delta y \Delta y}(\lambda)}{\partial \sigma_u^2} &= 2(1 - \cos\lambda), \\
\frac{\partial g_{\Delta y \Delta y}(\lambda)}{\partial \delta_x} &= \frac{2\sigma_f^2(\cos 2\lambda - \delta_x)}{(1 + \delta_x^2 - 2\delta_x \cos 2\lambda)^2},
\end{aligned}$$

which under the the null of $H_0 : \delta_x = 0$ become

$$\begin{aligned}
\frac{\partial g_{\Delta y \Delta y}(\lambda)}{\partial \sigma_f^2} &= 1, \\
\frac{\partial g_{\Delta y \Delta y}(\lambda)}{\partial \sigma_u^2} &= 2(1 - \cos\lambda), \\
\frac{\partial g_{\Delta y \Delta y}(\lambda)}{\partial \delta_x} &= 2\sigma_f^2 \cos 2\lambda.
\end{aligned}$$

As expected, this test is locally equivalent to a test against a restricted MA(2), which is also locally equivalent to the two-sided version of the test against an unrestricted AR(1).

A.2 Testing for neglected serial correlation in the noise

Let us know see what happens if we are interested in testing for first order serial correlation in u_t. The model under the alternative becomes

$$\left.\begin{array}{c} y_t = x_t + u_t \\ (1-L)x_t = f_t \\ (1-\psi_u L)u_t = v_t \end{array}\right\}$$

with u_t and f_t orthogonal at all leads and lags. The null hypothesis of interest is $H_0 : \psi_u = 0$.

Taking first differences of the observed variables to make them stationary yields

$$\Delta y_t = f_t + \frac{1-L}{1-\psi_u L} v_t.$$

Using the expressions for the autocovariances of an ARMA(1,1) with a unit root in the MA part, it is easy to see that

$$\begin{aligned} V(\Delta u_t) &= \gamma_{\Delta u \Delta u}(0) = \frac{2}{1+\psi_u}\sigma_v^2 \\ cov(\Delta u_t, \Delta u_{t-1}) &= \gamma_{\Delta u \Delta u}(1) = -\frac{(1-\psi_u)}{1+\psi_u}\sigma_v^2 \\ cov(\Delta u_t, \Delta u_{t-j}) &= \gamma_{\Delta u \Delta u}(j) = \psi_u \gamma_{\Delta u \Delta u}(j-1) \quad j \geq 2, \end{aligned}$$

As a result,

$$\begin{aligned} V(\Delta y_t) &= \gamma_{\Delta y \Delta y}(0) = \sigma_f^2 + \frac{2}{1+\psi_u}\sigma_v^2, \\ cov(\Delta y_t, \Delta y_{t-1}) &= \gamma_{\Delta y \Delta y}(1) = -\frac{(1-\psi_u)}{1+\psi_u}\sigma_v^2, \\ cov(\Delta y_t, \Delta y_{t-j}) &= \gamma_{\Delta y \Delta y}(j) = \psi_u \gamma_{\Delta y \Delta y}(j-1) \quad j \geq 2. \end{aligned}$$

Similarly, the spectral density of Δy_t will be

$$\begin{aligned} g_{\Delta y \Delta y}(\lambda) &= \sigma_f^2 + \frac{(1-e^{-i\lambda})(1-e^{i\lambda})}{(1-\psi_u e^{-i\lambda})(1-\psi_u e^{i\lambda})}\sigma_v^2 \\ &= \sigma_f^2 + \frac{2(1-\cos\lambda)}{1+\psi_u^2 - 2\psi_u \cos\lambda}\sigma_v^2, \end{aligned}$$

and its partial derivatives

$$\begin{aligned} \frac{\partial g_{\Delta y \Delta y}(\lambda)}{\partial \sigma_f^2} &= 1 \\ \frac{\partial g_{\Delta y \Delta y}(\lambda)}{\partial \sigma_v^2} &= \frac{2(1-\cos\lambda)}{1+\psi_u^2 - 2\psi_u \cos\lambda} \\ \frac{\partial g_{\Delta y \Delta y}(\lambda)}{\partial \phi} &= \frac{4\sigma_f^2(\cos\lambda - \psi_u)(1-\cos\lambda)}{(1+\psi_u^2 - 2\psi_u \cos\lambda)^2} \end{aligned}$$

Under the the null of $H_0 : \psi_u = 0$ those derivatives become

$$\begin{aligned}
\frac{\partial g_{\Delta y \Delta y}(\lambda)}{\partial \sigma_f^2} &= 1 \\
\frac{\partial g_{\Delta y \Delta y}(\lambda)}{\partial \sigma_v^2} &= 2(1-\cos\lambda) \\
\frac{\partial g_{\Delta y \Delta y}(\lambda)}{\partial \phi} &= 4\sigma_f^2 \cos\lambda\,(1-\cos\lambda)
\end{aligned}$$

Given that the spectral density under the null is

$$\sigma_f^2 + 2(1-\cos\lambda)\sigma_v^2,$$

we can compute the information matrix by integrating the outerproduct of the following vector:

$$\begin{aligned}
\frac{\partial g_{\Delta y \Delta y}(\lambda)}{\partial \sigma_f^2}\frac{1}{g_{\Delta y \Delta y}(\lambda)} &= \frac{1}{\sigma_f^2 + 2(1-\cos\lambda)\sigma_v^2}, \\
\frac{\partial g_{\Delta y \Delta y}(\lambda)}{\partial \sigma_v^2}\frac{1}{g_{\Delta y \Delta y}(\lambda)} &= \frac{2(1-\cos\lambda)}{\sigma_f^2 + 2(1-\cos\lambda)\sigma_v^2}, \\
\frac{\partial g_{\Delta y \Delta y}(\lambda)}{\partial \psi_u}\frac{1}{g_{\Delta y \Delta y}(\lambda)} &= \frac{4\sigma_f^2 \cos\lambda\,(1-\cos\lambda)}{\sigma_f^2 + 2(1-\cos\lambda)\sigma_v^2}.
\end{aligned}$$

Unlike what happens in the test for $\psi_x = 0$, the information matrix will be regular when $\psi_u = 0$. Given that the score with respect to ψ_u involves a square cosine, which can always be expanded in terms of $\cos 2\lambda$ by using the trigronometric identity

$$\cos 2\lambda = 2\cos^2\lambda - 1, \tag{A11}$$

the test for neglected serial correlation in the noise will also coincide with the two-sided version of the extremum test.

Finally, it is easy to see that apart from a sign change, one would get the same derivative under the null if we were considering an MA(1) alternative for u_t.

B Proofs of propositions

Proposition 1

Given that $\psi_f \neq \alpha^{-1}$ if we choose an invertible MA polynomial, Lemma 1 allows us to replace σ_f^2 and σ_u^2 by the theoretical variance and first autocovariance of the observed series as follows:

$$\begin{aligned}
\sigma_f^2 &= \frac{(1-\alpha^2)}{(1-\alpha\psi_f)(\alpha-\psi_f)}\gamma_{yy}(1), \\
\sigma_u^2 &= \gamma_{yy}(0) - \frac{(1+\psi_f^2 - 2\alpha\psi_f)}{(1-\alpha\psi_f)(\alpha-\psi_f)}\gamma_{yy}(1)
\end{aligned}$$

under the assumption that $\psi_f \neq \alpha$, which is valid in a neighbourhood of $\psi_f = 0$ since we maintain the assumption that the true value of α is different from 0.

In this notation we can write the spectral density (C1) as

$$\begin{aligned} g_{yy}(\lambda) &= \frac{1+\psi_f^2-2\psi_f\cos\lambda}{1+\alpha^2-2\alpha\cos\lambda}\frac{\left(1-\alpha^2\right)}{(1-\alpha\psi_f)(\alpha-\psi_f)}\gamma_{yy}(1)+\gamma_{yy}(0)-\frac{(1+\psi_f^2-2\alpha\psi_f)}{(1-\alpha\psi_f)(\alpha-\psi_f)}\gamma_{yy}(1) \\ &= \gamma_{yy}(0)+\frac{2(\cos\lambda-\alpha)}{1+\alpha^2-2\alpha\cos\lambda}\gamma_{yy}(1), \end{aligned}$$

which does not depend on ψ_f. □

Proposition 2

The partial derivatives of the spectral density (C5) are:

$$\begin{aligned} \frac{\partial g_{yy}(\lambda|\sigma_f^2,\sigma_u^2,\alpha,\psi_x)}{\partial\sigma_f^2} &= \frac{1}{\left(1+\psi_x^2-2\psi_x\cos\lambda\right)\left(1+\alpha^2-2\alpha\cos\lambda\right)}, \\ \frac{\partial g_{yy}(\lambda|\sigma_f^2,\sigma_u^2,\alpha,\psi_x)}{\partial\sigma_u^2} &= 1, \\ \frac{\partial g_{yy}(\lambda|\sigma_f^2,\sigma_u^2,\alpha,\psi_x)}{\partial\alpha} &= \frac{2\sigma_f^2(\cos\lambda-\alpha)}{\left(1+\alpha^2-2\alpha\cos\lambda\right)^2\left(1+\psi_x^2-2\psi_x\cos\lambda\right)}, \\ \frac{\partial g_{yy}(\lambda|\sigma_f^2,\sigma_u^2,\alpha,\psi_x)}{\partial\psi_x} &= \frac{2\sigma_f^2(\cos\lambda-\psi_x)}{(1+\psi_x^2-2\psi_x\cos\lambda)^2\left(1+\alpha^2-2\alpha\cos\lambda\right)}. \end{aligned}$$

When $\psi_x = 0$, these derivatives reduce to

$$\begin{aligned} \frac{\partial g_{yy}(\lambda|\sigma_f^2,\sigma_u^2,\alpha,0)}{\partial\sigma_f^2} &= \frac{1}{(1+\alpha^2-2\alpha\cos\lambda)}, \\ \frac{\partial g_{yy}(\lambda|\sigma_f^2,\sigma_u^2,\alpha,0)}{\partial\sigma_u^2} &= 1, \\ \frac{\partial g_{yy}(\lambda|\sigma_f^2,\sigma_u^2,\alpha,0)}{\partial\alpha} &= \frac{2\sigma_f^2(\cos\lambda-\alpha)}{(1+\alpha^2-2\alpha\cos\lambda)^2}, \\ \frac{\partial g_{yy}(\lambda|\sigma_f^2,\sigma_u^2,\alpha,0)}{\partial\psi_x} &= \frac{2\sigma_f^2\cos\lambda}{(1+\alpha^2-2\alpha\cos\lambda)}. \end{aligned}$$

Given that the spectral density under the the null is

$$g_{yy}(\lambda|\sigma_f^2,\sigma_u^2,\alpha,0) = \frac{\sigma_f^2}{(1+\alpha^2-2\alpha\cos\lambda)}+\sigma_u^2,$$

and its reciprocal

$$g_{yy}^{-1}(\lambda|\sigma_f^2,\sigma_u^2,\alpha,0) = \frac{\left(1+\alpha^2-2\alpha\cos\lambda\right)}{\sigma_u^2\left(1+\alpha^2-2\alpha\cos\lambda\right)+\sigma_f^2},$$

we will have that for $\psi_x = 0$

$$\begin{aligned}
\frac{\partial g_{yy}(\lambda|\sigma_f^2, \sigma_u^2, \alpha, 0)}{\partial \sigma_f^2} \frac{1}{g_{yy}(\lambda|\sigma_f^2, \sigma_u^2, \alpha, 0)} &= \frac{1}{\sigma_u^2 \left(1 + \alpha^2 - 2\alpha \cos \lambda\right) + \sigma_f^2}, \\
\frac{\partial g_{yy}(\lambda|\sigma_f^2, \sigma_u^2, \alpha, 0)}{\partial \sigma_u^2} \frac{1}{g_{yy}(\lambda|\sigma_f^2, \sigma_u^2, \alpha, 0)} &= \frac{\left(1 + \alpha^2 - 2\alpha \cos \lambda\right)}{\sigma_u^2 \left(1 + \alpha^2 - 2\alpha \cos \lambda\right) + \sigma_f^2}, \\
\frac{\partial g_{yy}(\lambda|\sigma_f^2, \sigma_u^2, \alpha, 0)}{\partial \alpha} \frac{1}{g_{yy}(\lambda|\sigma_f^2, \sigma_u^2, \alpha, 0)} &= \frac{2\sigma_f^2(\cos \lambda - \alpha)}{(1 + \alpha^2 - 2\alpha \cos \lambda)} \frac{1}{\sigma_u^2 \left(1 + \alpha^2 - 2\alpha \cos \lambda\right) + \sigma_f^2}, \\
\frac{\partial g_{yy}(\lambda|\sigma_f^2, \sigma_u^2, \alpha, 0)}{\partial \psi_x} \frac{1}{g_{yy}(\lambda|\sigma_f^2, \sigma_u^2, \alpha, 0)} &= \frac{2\sigma_f^2 \cos \lambda}{\sigma_u^2 \left(1 + \alpha^2 - 2\alpha \cos \lambda\right) + \sigma_f^2}.
\end{aligned}$$

It is then easy to see that

$$\begin{aligned}
\sigma_f^2 &\left[\frac{\partial g_{yy}(\lambda|\sigma_f^2, \sigma_u^2, \alpha, 0)}{\partial \sigma_u^2} \frac{1}{g_{yy}(\lambda|\sigma_f^2, \sigma_u^2, \alpha, 0)} - (1 + \alpha^2) \frac{\partial g_{yy}(\lambda|\sigma_f^2, \sigma_u^2, \alpha, 0)}{\partial \sigma_f^2} \frac{1}{g_{yy}(\lambda|\sigma_f^2, \sigma_u^2, \alpha, 0)} \right] \\
&\quad + \alpha \frac{\partial g_{yy}(\lambda|\sigma_f^2, \sigma_u^2, \alpha, 0)}{\partial \psi_x} \frac{1}{g_{yy}(\lambda|\sigma_f^2, \sigma_u^2, \alpha, 0)} \\
&= \frac{\left(1 + \alpha^2 - 2\alpha \cos \lambda\right) \sigma_f^2}{\sigma_u^2 \left(1 + \alpha^2 - 2\alpha \cos \lambda\right) + \sigma_f^2} - \frac{(1 + \alpha^2)\sigma_f^2}{\sigma_u^2 \left(1 + \alpha^2 - 2\alpha \cos \lambda\right) + \sigma_f^2} \\
&\quad + \frac{2\sigma_f^2 \alpha \cos \lambda}{\sigma_u^2 \left(1 + \alpha^2 - 2\alpha \cos \lambda\right) + \sigma_f^2} = 0.
\end{aligned}$$

Given (8), this result implies that the information matrix of model (20) will only have rank 3 under the null when the true value of α is not zero. □

Proposition 3

Let us replace σ_f^2 and σ_u^2 by the variance and the first autocovariance of the observed series. Asuming that $\alpha + \psi_x \neq 0$, which is valid in a neighbourhood of $\psi_x = 0$ when the true value of α is different from 0, the solution will be

$$\begin{aligned}
\sigma_f^2 &= \frac{\left(1 - \alpha^2\right) \left(1 - \psi_x^2\right) (1 - \alpha\psi_x)}{\alpha + \psi_x} \gamma_{yy}(1), \\
\sigma_u^2 &= \gamma_{yy}(0) - \frac{(1 + \alpha\psi_x)}{\alpha + \psi_x} \gamma_{yy}(1).
\end{aligned}$$

so that

$$\begin{aligned}
\frac{\partial \sigma_f^2}{\partial \gamma_{yy}(0)} &= 0, \\
\frac{\partial \sigma_f^2}{\partial \gamma_{yy}(1)} &= \frac{\left(1 - \alpha^2\right) \left(1 - \psi_x^2\right) (1 - \alpha\psi_x)}{\alpha + \psi_x}, \\
\frac{\partial \sigma_f^2}{\partial \alpha} &= \frac{\left(1 - \psi_x^2\right)}{\left(\alpha + \psi_x\right)^2} \left[2\alpha^3\psi_x + \alpha^2(3\psi_x^2 - 1) - 2\alpha\psi_x - (1 + \psi_x^2)\right] \gamma_{yy}(1), \\
\frac{\partial \sigma_f^2}{\partial \psi_x} &= \frac{\left(1 - \alpha^2\right)}{\left(\alpha + \psi_x\right)^2} \left[2\psi_x^3\alpha + \psi_x^2(3\alpha^2 - 1) - 2\alpha\psi_x - (1 + \alpha^2)\right] \gamma_{yy}(1),
\end{aligned}$$

and

$$\begin{aligned}
\frac{\partial \sigma_u^2}{\partial \gamma_{yy}(0)} &= 1, \\
\frac{\partial \sigma_u^2}{\partial \gamma_{yy}(1)} &= -\frac{(1+\alpha\psi_x)}{\alpha+\psi_x}, \\
\frac{\partial \sigma_u^2}{\partial \alpha} &= \frac{1-\psi_x^2}{(\alpha+\psi_x)^2}\gamma_{yy}(1), \\
\frac{\partial \sigma_u^2}{\partial \psi_x} &= \frac{1-\alpha^2}{(\alpha+\psi_x)^2}\gamma_{yy}(1).
\end{aligned}$$

Under the null of $\psi_x = 0$ these derivatives simplify to

$$\begin{aligned}
\frac{\partial \sigma_f^2}{\partial \gamma_{yy}(0)} &= 0, \\
\frac{\partial \sigma_f^2}{\partial \gamma_{yy}(1)} &= \frac{1-\alpha^2}{\alpha}, \\
\frac{\partial \sigma_f^2}{\partial \alpha} &= -\frac{1+\alpha^2}{\alpha^2}\gamma_{yy}(1) = -\frac{1+\alpha^2}{\alpha}\frac{\sigma_f^2}{1-\alpha^2}, \\
\frac{\partial \sigma_f^2}{\partial \psi_x} &= -\frac{(1-\alpha^2)(1+\alpha^2)}{\alpha^2}\gamma_{yy}(1) = -\frac{(1+\alpha^2)}{\alpha}\sigma_f^2,
\end{aligned}$$

and

$$\begin{aligned}
\frac{\partial \sigma_u^2}{\partial \gamma_{yy}(0)} &= 1, \\
\frac{\partial \sigma_u^2}{\partial \gamma_{yy}(1)} &= -\frac{1}{\alpha}, \\
\frac{\partial \sigma_u^2}{\partial \alpha} &= \frac{1}{\alpha^2}\gamma_{yy}(1) = \frac{1}{\alpha}\frac{\sigma_f^2}{1-\alpha^2}, \\
\frac{\partial \sigma_u^2}{\partial \psi_x} &= \frac{1-\alpha^2}{\alpha^2}\gamma_{yy}(1) = \frac{1}{\alpha}\sigma_f^2,
\end{aligned}$$

where we have used the fact that when $\psi_x = 0$

$$\begin{aligned}
\gamma_{yy}(0) &= \sigma_u^2 + \frac{\sigma_f^2}{1-\alpha^2}, \\
\gamma_{yy}(1) &= \alpha\frac{\sigma_f^2}{1-\alpha^2}.
\end{aligned}$$

If we apply the chain rule to this reparametrisation, the new derivative wrt ψ_x evaluated at

$\psi_x = 0$ will be

$$\begin{aligned}
& \frac{\partial g_{yy}(\lambda|\sigma_f^2,\sigma_u^2,\alpha,0)}{\partial\sigma_f^2}\frac{1}{g_{yy}(\lambda|\sigma_f^2,\sigma_u^2,\alpha,0)}\frac{\partial\sigma_f^2}{\partial\psi_x}+\frac{\partial g_{yy}(\lambda|\sigma_f^2,\sigma_u^2,\alpha,0)}{\partial\sigma_u^2}\frac{1}{g_{yy}(\lambda|\sigma_f^2,\sigma_u^2,\alpha,0)}\frac{\partial\sigma_u^2}{\partial\psi_x} \\
& +\frac{\partial g_{yy}(\lambda|\sigma_f^2,\sigma_u^2,\alpha,0)}{\partial\psi_x}\frac{1}{g_{yy}(\lambda|\sigma_f^2,\sigma_u^2,\alpha,0)} \\
= & -\frac{(1+\alpha^2)}{\alpha}\sigma_f^2\frac{1}{\sigma_u^2\left(1+\alpha^2-2\alpha\cos\lambda\right)+\sigma_f^2} \\
& +\frac{1}{\alpha}\sigma_f^2\frac{\left(1+\alpha^2-2\alpha\cos\lambda\right)}{\sigma_u^2\left(1+\alpha^2-2\alpha\cos\lambda\right)+\sigma_f^2}+\frac{2\sigma_f^2\cos\lambda}{\sigma_u^2\left(1+\alpha^2-2\alpha\cos\lambda\right)+\sigma_f^2} \\
= & \frac{\sigma_f^2}{\sigma_u^2\left(1+\alpha^2-2\alpha\cos\lambda\right)+\sigma_f^2}\left(-\frac{(1+\alpha^2)}{\alpha}+\frac{1}{\alpha}\left(1+\alpha^2-2\alpha\cos\lambda\right)+2\cos\lambda\right) \\
= & \ 0,
\end{aligned}$$

as desired. Obviously, we would obtain exactly the same result had we expressed the spectral density of y_t in terms of $\gamma_{yy}(0)$, $\gamma_{yy}(1)$, α and ψ_x as

$$\begin{gathered}
g_{yy}(\lambda|\gamma_{yy}(0),\gamma_{yy}(1),\alpha,\psi_x)=\gamma_{yy}(0) \\
+\frac{\left(1-\alpha^2\right)\left(1-\psi_x^2\right)(1-\alpha\psi_x)-(1+\alpha\psi_x)\left(1+\alpha^2-2\alpha\cos\lambda\right)\left(1+\psi_x^2-2\psi_x\cos\lambda\right)}{(\alpha+\psi_x)\left(1+\alpha^2-2\alpha\cos\lambda\right)\left(1+\psi_x^2-2\psi_x\cos\lambda\right)}\gamma_{yy}(1),
\end{gathered} \tag{B1}$$

derived this expression with respect to ψ_x obtaining

$$\begin{aligned}
\frac{\partial g_{yy}(\lambda|\gamma_{yy}(0),\gamma_{yy}(1),\alpha,\psi_x)}{\partial\psi_x} &= -\frac{2\left(\alpha^2-1\right)\psi_x\gamma_{yy}(1)}{(\alpha+\psi_x)^2\left(1+\alpha^2-2\alpha\cos(\lambda)\right)\left(1+\psi_x^2-2\psi_x\cos\lambda\right)^2} \\
\times & \left(\begin{array}{c}\psi_x\left(\alpha^2\left(\psi_x^2+4\right)+4\alpha\psi_x+\psi_x^2\right)-(2\alpha+\psi_x)\left(2\alpha\psi_x^2+\alpha+2\psi_x\right)\cos(\lambda) \\ +(\alpha(\psi_x(\alpha+2\psi_x)+2)+\psi_x)\cos(2\lambda)-\alpha\psi_x\cos(3\lambda)\end{array}\right)
\end{aligned}$$

and evaluated this derivative at $\psi_x = 0$. □

Proposition 4

If we choose $\psi_x = +\sqrt{\varphi}$, the spectral density of y_t written in this form will be

$$\begin{gathered}
g_{yy}^{+}(\lambda|\gamma_{yy}(0),\gamma_{yy}(1),\alpha,\varphi)=\gamma_{yy}(0) \\
+\frac{\left(1-\alpha^2\right)(1-\varphi)\left(1-\alpha\sqrt{\varphi}\right)-\left(1+\alpha\sqrt{\varphi}\right)\left(1+\alpha^2-2\alpha\cos\lambda\right)\left(1+\varphi-2\sqrt{\varphi}\cos\lambda\right)}{\left(\alpha+\sqrt{\phi}\right)\left(1+\alpha^2-2\alpha\cos\lambda\right)\left(1+\varphi-2\sqrt{\varphi}\cos\lambda\right)}\gamma_{yy}(1)
\end{gathered}$$

while if we choose $\psi_x = -\sqrt{\varphi}$ it becomes

$$\begin{gathered}
g_{yy}^{-}(\lambda|\gamma_{yy}(0),\gamma_{yy}(1),\alpha,\varphi)=\gamma_{yy}(0) \\
+\frac{\left(1-\alpha^2\right)(1-\varphi)\left(1+\alpha\sqrt{\varphi}\right)-\left(1-\alpha\sqrt{\varphi}\right)\left(1+\alpha^2-2\alpha\cos\lambda\right)\left(1+\varphi+2\sqrt{\varphi}\cos\lambda\right)}{\left(\alpha-\sqrt{\phi}\right)\left(1+\alpha^2-2\alpha\cos\lambda\right)\left(1+\varphi-2\sqrt{\varphi}\cos\lambda\right).}\gamma_{yy}(1)
\end{gathered}$$

Next we must obtain the derivative under the alternative, and then evaluate it under the null. In this way we obtain

$$\frac{\partial g^{+}_{yy}(\lambda|\gamma_{yy}(0),\gamma_{yy}(1),\alpha,\varphi)}{\partial\varphi} = \frac{\left(\alpha^2-1\right)\gamma_{yy}(1)}{\left(\alpha+\sqrt{\varphi}\right)^2\left(1+\alpha^2-2\alpha\cos(\lambda)\right)\left(1+\varphi-2\sqrt{\varphi}\cos\lambda\right)^2}$$
$$\times\left(\begin{array}{c} -\left(\left(\alpha^2+1\right)\sqrt{\varphi}+2\alpha\varphi+2\alpha\right)\cos(2\lambda)-\sqrt{\varphi}\left(\alpha^2(\varphi+4)-\alpha\cos(3\lambda)+4\alpha\sqrt{\varphi}+\varphi\right)\\ +\left(2\alpha+\sqrt{\varphi}\right)\left(2\alpha\varphi+\alpha+2\sqrt{\varphi}\right)\cos\lambda \end{array}\right)$$

so that

$$\frac{\partial g^{+}_{yy}(\lambda|\gamma_{yy}(0),\gamma_{yy}(1),\alpha,0)}{\partial\varphi} = \frac{2\left(1-\alpha^2\right)\left(\cos(2\lambda)-\alpha\cos\lambda\right)}{\alpha\left(1+\alpha^2-2\alpha\cos(\lambda)\right)}\gamma_{yy}(1). \tag{B2}$$

Similarly,

$$\frac{\partial g^{-}_{yy}(\lambda|\gamma_{yy}(0),\gamma_{yy}(1),\alpha,\varphi)}{\partial\varphi} = \frac{\left(\alpha^2-1\right)\gamma_{yy}(1)}{\left(\alpha-\sqrt{\varphi}\right)^2\left(1+\alpha^2-2\alpha\cos(\lambda)\right)\left(1+\varphi+2\sqrt{\varphi}\cos\lambda\right)^2}$$
$$\times\left(\begin{array}{c} \left(\left(\alpha^2+1\right)\sqrt{\varphi}-2\alpha\varphi-2\alpha\right)\cos(2\lambda)+\sqrt{\varphi}\left(\alpha^2(\varphi+4)-\alpha\cos(3\lambda)-4\alpha\sqrt{\varphi}+\varphi\right)\\ +\left(2\alpha-\sqrt{\varphi}\right)\left(2\alpha\varphi+\alpha-2\sqrt{\varphi}\right)\cos(\lambda) \end{array}\right)$$

so that

$$\frac{\partial g^{-}_{yy}(\lambda|\gamma_{yy}(0),\gamma_{yy}(1),\alpha,0)}{\partial\varphi} = \frac{2\left(1-\alpha^2\right)\left(\cos(2\lambda)-\alpha\cos(\lambda)\right)}{\alpha\left(1+\alpha^2-2\alpha\cos(\lambda)\right)}\gamma_{yy}(1),$$

which coincides with (B2). Hence, the score test for the null hypothesis $H_0:\varphi:0$ will indeed be based on the "influence function" (21).

We can also try the alternative route proposed by Lee and Chesher (1986). Given that

$$\frac{\partial^2 g_{yy}(\lambda|\gamma_{yy}(0),\gamma_{yy}(1),\alpha,\psi_x)}{\partial\psi_x\partial\psi_x} = -\frac{4\left(\alpha^2-1\right)\gamma_{yy}(1)}{(\alpha+\psi_x)^3\left(1+\alpha^2-2\alpha\cos(\lambda)\right)\left(1+\psi_x^2-2\phi\cos\lambda\right)^3}$$
$$\times\left(\begin{array}{c} \left(-3\left(\alpha^2+1\right)\psi_x^4+\alpha\left(\alpha^2-10\right)\psi_x^3-3\alpha^2\psi_x^2+\alpha^2-3\alpha\phi^5\right)\cos(2\lambda)\\ +\left(-\alpha^3\left(3\phi^2+1\right)+\alpha^2\left(6\phi^4+8\phi^2-3\right)\psi_x+2\alpha\left(\psi_x^2+6\right)\psi_x^4+3\phi^5\right)\cos(\lambda)\\ -\psi_x\left(-3\alpha^3+\alpha^2\left(\psi_x^4+9\phi^2-3\right)\psi_x+6\alpha\phi^4+\psi_x^5\right)-\alpha\phi^3\cos(4\lambda)\\ +\psi_x^2(\alpha(\psi_x(\alpha+3\phi)+3)+\psi_x)\cos(3\lambda) \end{array}\right)$$

so that

$$\frac{\partial^2 g_{yy}(\lambda|\gamma_{yy}(0),\gamma_{yy}(1),\alpha,0)}{\partial\psi_x\partial\psi_x} = \frac{4\left(1-\alpha^2\right)\left(\cos(2\lambda)-\alpha\cos(\lambda)\right)}{\alpha\left(1+\alpha^2-2\alpha\cos(\lambda)\right)}\gamma_{yy}(1).$$

Having obtained the derivative of the original spectral density, we can obtain the second derivative of the spectral log-likelihood function with respect ψ_x by taking first derivatives of the score (7). But since we have seen that

$$\frac{\partial g_{yy}(\lambda|\gamma_{yy}(0),\gamma_{yy}(1),\alpha,0)}{\partial\psi_x}=0,$$

the second derivative of the log-likelihood function will be

$$\sum_{j=0}^{T-1}\gamma_{yy}(1)\frac{4\left(1-\alpha^2\right)\left(\cos(2\lambda_j)-\alpha\cos(\lambda_j)\right)}{\alpha\left(1+\alpha^2-2\alpha\cos(\lambda_j)\right)}\frac{\left[I_{yy}(\lambda_j)-g_{yy}(\lambda_j|\gamma_{yy}(0),\gamma_{yy}(1),\alpha,0)\right]}{g^2_{yy}(\lambda_j|\gamma_{yy}(0),\gamma_{yy}(1),\alpha,0)}$$

so that the tests will be indeed identical.

Finally, we can tediously show that

$$\frac{\partial^3 g_{yy}(\lambda|\gamma_{yy}(0),\gamma_{yy}(1),\alpha,\psi_x)}{(\partial\psi_x)^3} = \frac{12\left(\alpha^2-1\right)\gamma_{yy}(1)}{(\alpha+\psi_x)^4\left(\alpha^2-2\alpha\cos(\lambda)+1\right)\left(-2\psi_x\cos(\lambda)+\psi_x^2+1\right)^4}$$

$$\times\left(\begin{array}{c}\left(-\alpha^3\left(\psi_x^4+1\right)+4\alpha^2\left(\psi_x^4+\psi_x^2-1\right)\psi_x+\alpha\left(6\psi_x^2+17\right)\psi_x^4+4\psi_x^5\right)\cos(3\lambda)\\ +\left(\begin{array}{c}\alpha^4\left(\psi_x^4+1\right)+4\alpha^3\left(\psi_x^4+\psi_x^2+2\right)\psi_x+\alpha^2\left(-6\psi_x^6-15\psi_x^4+16\psi_x^2+1\right)\\ -4\alpha\left(\psi_x^2+7\right)\psi_x^5-6\psi_x^6\end{array}\right)\cos(2\lambda)\\ -\left(\begin{array}{c}4\alpha^4\left(\psi_x^3+\psi_x\right)+\alpha^3\left(17\psi_x^4+28\psi_x^2+1\right)+4\alpha^2\left(-2\psi_x^6-5\psi_x^4+7\psi_x^2+1\right)\psi_x\\ -2\alpha\left(\psi_x^2+11\right)\psi_x^6-4\psi_x^7\end{array}\right)\cos(\lambda)\\ +\psi_x\left(6\alpha^4\psi_x+4\alpha^3\left(7\psi_x^2+1\right)-\alpha^2\left(\psi_x^6+16\psi_x^4-16\psi_x^2-6\right)\psi_x-8\alpha\psi_x^6-\psi_x^7\right)+\alpha\psi_x^4\cos(5\lambda)\\ -\psi_x^3(\alpha(\psi_x(\alpha+4\psi_x)+4)+\psi_x)\cos(4\lambda)\end{array}\right),$$

so that

$$\frac{\partial^3 g_{yy}(\lambda|\gamma_{yy}(0),\gamma_{yy}(1),\alpha,0)}{(\partial\psi_x)^3} = \frac{12\left(\alpha^2-1\right)\cos(2\lambda)}{\alpha^2}\gamma_{yy}(1),$$

which in turn implies the local identifiability of ψ_x under the null. □

Proposition 5

The derivatives of the spectral density will be

$$\begin{aligned}
\frac{\partial g_{yy}(\lambda|\sigma_f^2,\sigma_u^2,\alpha,\delta_x)}{\partial\sigma_f^2} &= \frac{1}{(1+\alpha^2-2\alpha\cos\lambda)\left(1+\delta_x^2-2\delta_x\cos(2\lambda)\right)},\\
\frac{\partial g_{yy}(\lambda|\sigma_f^2,\sigma_u^2,\alpha,\delta_x)}{\partial\sigma_u^2} &= 1,\\
\frac{\partial g_{yy}(\lambda|\sigma_f^2,\sigma_u^2,\alpha,\delta_x)}{\partial\alpha} &= \frac{2(\cos\lambda-\alpha)}{(1+\alpha^2-2\alpha\cos\lambda)^2\left(1+\delta_x^2-2\delta_x\cos(2\lambda)\right)}\sigma_f^2,\\
\frac{\partial g_{yy}(\lambda|\sigma_f^2,\sigma_u^2,\alpha,\delta_x)}{\partial\delta_x} &= \frac{2(\cos(2\lambda)-\alpha)}{(1+\alpha^2-2\alpha\cos\lambda)\left(1+\delta_x^2-2\delta_x\cos(2\lambda)\right)^2}\sigma_f^2,
\end{aligned}$$

which under the null reduce to

$$\begin{aligned}
\frac{\partial g_{yy}(\lambda|\sigma_f^2,\sigma_u^2,\alpha,0)}{\partial\sigma_f^2} &= \frac{1}{(1+\alpha^2-2\alpha\cos\lambda)},\\
\frac{\partial g_{yy}(\lambda|\sigma_f^2,\sigma_u^2,\alpha,0)}{\partial\sigma_u^2} &= 1,\\
\frac{\partial g_{yy}(\lambda|\sigma_f^2,\sigma_u^2,\alpha,0)}{\partial\alpha} &= \frac{2(\cos\lambda-\alpha)}{(1+\alpha^2-2\alpha\cos\lambda)^2}\sigma_f^2,\\
\frac{\partial g_{yy}(\lambda|\sigma_f^2,\sigma_u^2,\alpha,0)}{\partial\delta_x} &= \frac{2\cos(2\lambda)}{(1+\alpha^2-2\alpha\cos\lambda)}\sigma_f^2.
\end{aligned}$$

Given that the spectral density under the the null is

$$g_{yy}(\lambda|\sigma_f^2,\sigma_u^2,\alpha,0) = \frac{\sigma_f^2}{(1+\alpha^2-2\alpha\cos\lambda)} + \sigma_u^2,$$

and its reciprocal

$$g_{yy}^{-1}(\lambda|\sigma_f^2,\sigma_u^2,\alpha,0)=\frac{\left(1+\alpha^2-2\alpha\cos\lambda\right)}{\sigma_u^2\left(1+\alpha^2-2\alpha\cos\lambda\right)+\sigma_f^2},$$

we will have that the contribution of frequency λ to the log-likelihood scores evaluated at $\delta_x=0$ will be

$$\begin{aligned}
s_{\sigma_f^2}(\lambda|\sigma_f^2,\sigma_u^2,\alpha,0) &= \frac{\left(1+\alpha^2-2\alpha\cos\lambda\right)}{\left(\sigma_u^2\left(1+\alpha^2-2\alpha\cos\lambda\right)+\sigma_f^2\right)^2}\left[2\pi I_{yy}(\lambda)-g_{yy}(\lambda|\sigma_f^2,\sigma_u^2,\alpha,0)\right], \\
s_{\sigma_u^2}(\lambda|\sigma_f^2,\sigma_u^2,\alpha,0) &= \frac{\left(1+\alpha^2-2\alpha\cos\lambda\right)^2}{\left(\sigma_u^2\left(1+\alpha^2-2\alpha\cos\lambda\right)+\sigma_f^2\right)^2}\left[2\pi I_{yy}(\lambda)-g_{yy}(\lambda|\sigma_f^2,\sigma_u^2,\alpha,0)\right], \\
s_{\alpha}(\lambda|\sigma_f^2,\sigma_u^2,\alpha,0) &= \frac{2(\cos\lambda-\alpha)\sigma_f^2}{\left(\sigma_u^2\left(1+\alpha^2-2\alpha\cos\lambda\right)+\sigma_f^2\right)^2}\left[2\pi I_{yy}(\lambda)-g_{yy}(\lambda|\sigma_f^2,\sigma_u^2,\alpha,0)\right], \\
s_{\delta_x}(\lambda|\sigma_f^2,\sigma_u^2,\alpha,0) &= \frac{2\cos(2\lambda)\left(1+\alpha^2-2\alpha\cos\lambda\right)\sigma_f^2}{\left(\sigma_u^2\left(1+\alpha^2-2\alpha\cos\lambda\right)+\sigma_f^2\right)^2}\left[2\pi I_{yy}(\lambda)-g_{yy}(\lambda|\sigma_f^2,\sigma_u^2,\alpha,0)\right].
\end{aligned}$$

Given that these scores are not orthogonal under the null, we will have to orthogonalise the last one with respect to the first three using the information matrix under the null, which will be given by (8), with the spectral derivatives obtained above. But given that the linear span of $\partial g_{yy}(\lambda|\sigma_f^2,\sigma_u^2,\alpha,0)/\partial\sigma_f^2$ and $\partial g_{yy}(\lambda|\sigma_f^2,\sigma_u^2,\alpha,0)/\partial\sigma_u^2$ is the same as the linear span of $\partial g_{yy}(\lambda|\gamma_{yy}(0),\gamma_{yy}(1),\alpha,0)/\partial\gamma_{yy}(0)$ and $\partial g_{yy}(\lambda|\gamma_{yy}(0),\gamma_{yy}(1),\alpha,0)/\partial\gamma_{yy}(1)$ when they are both evaluated under the null, the adjusted test is going to coincide with a two-sided version of the extremum test against an AR(1) alternative in Proposition 4. □

Proposition 6

As usual, it is convenient to reparametrise the model by replacing σ_f^2 and σ_u^2 by $\gamma_{yy}(0)$ and $\gamma_{yy}(1)$ from (C10) as follows

$$\sigma_f^2 = \frac{(1-\alpha^2)(\gamma_{yy}(1)-\psi_u\gamma_{yy}(0))}{\alpha-\psi_u} \tag{B3}$$

$$\sigma_u^2 = \frac{(1-\psi_u^2)(\gamma_{yy}(1)-\alpha\gamma_{yy}(0))}{\psi_u-\alpha} \tag{B4}$$

under the maintained assumption that $\alpha\neq\psi_u$. The spectral density then becomes

$$\begin{aligned}
g_{yy}(\lambda|\gamma_{yy}(0),\gamma_{yy}(1),\alpha,\psi_u) &= \frac{(1-\alpha^2)(\gamma_{yy}(1)-\psi_u\gamma_{yy}(0))}{(\alpha-\psi_u)\left(1+\alpha^2-2\alpha\cos\lambda\right)}+\frac{(1-\psi_u^2)(\gamma_{yy}(1)-\alpha\gamma_{yy}(0))}{(\psi_u-\alpha)\left(1+\psi_u^2-2\psi_u\cos\lambda\right)} \\
&= \left(\frac{(1-\psi_u^2)\alpha}{1+\psi_u^2-2\psi_u\cos\lambda}-\frac{(1-\alpha^2)\psi_u}{1+\alpha^2-2\alpha\cos\lambda}\right)\frac{\gamma_{yy}(0)}{\alpha-\psi_u} \\
&\quad+\left(\frac{(1-\psi_u^2)}{(1+\psi_u^2-2\psi_u\cos\lambda)}-\frac{(1-\alpha^2)}{(1+\alpha^2-2\alpha\cos\lambda)}\right)\frac{\gamma_{yy}(1)}{\psi_u-\alpha}.
\end{aligned}$$

Hence, the derivatives will be

$$
\begin{aligned}
\frac{\partial g_{yy}(\lambda|\gamma_{yy}(0),\gamma_{yy}(1),\alpha,\psi_u)}{\partial\gamma_{yy}(0)} &= \left(\frac{(1-\psi_u^2)\alpha}{1+\psi_u^2-2\psi_u\cos\lambda}-\frac{(1-\alpha^2)\psi_u}{1+\alpha^2-2\alpha\cos\lambda}\right)\frac{1}{\alpha-\psi_u},\\
\frac{\partial g_{yy}(\lambda|\gamma_{yy}(0),\gamma_{yy}(1),\alpha,\psi_u)}{\partial\gamma_{yy}(1)} &= \left(\frac{(1-\psi_u^2)}{(1+\psi_u^2-2\psi_u\cos\lambda)}-\frac{(1-\alpha^2)}{(1+\alpha^2-2\alpha\cos\lambda)}\right)\frac{1}{\psi_u-\alpha},\\
\frac{\partial g_{yy}(\lambda|\gamma_{yy}(0),\gamma_{yy}(1),\alpha,\psi_u)}{\partial\alpha} &= \left(\frac{(\alpha^4-1-4\alpha(\alpha-\psi_u))-2((1+\alpha^2)\psi_u-2\alpha)\cos\lambda}{(1+\alpha^2-2\alpha\cos\lambda)^2}\right.\\
&\quad\left.+\frac{1-\psi_u^2}{1+\psi_u^2-2\psi_u\cos\lambda}\right)\frac{(\gamma_{yy}(1)-\psi_u\gamma_{yy}(0))}{(\psi_u-\alpha)^2}\\
\frac{\partial g_{yy}(\lambda|\gamma_{yy}(0),\gamma_{yy}(1),\alpha,\psi_u)}{\partial\psi_u} &= \left(\frac{(\psi_u^4-1-4\psi_u(\psi_u-\alpha))-2((1+\psi_u^2)\alpha-2\psi_u)\cos\lambda}{\left(1+\psi_u^2-2\psi_u\cos\lambda\right)^2}\right.\\
&\quad\left.+\frac{1-\alpha^2}{1+\alpha^2-2\alpha\cos\lambda}\right)\frac{(\gamma_{yy}(1)-\alpha\gamma_{yy}(0))}{(\psi_u-\alpha)^2}
\end{aligned}
$$

Under the null hypothesis of $H_0:\psi_u=0$ the derivatives become

$$
\begin{aligned}
\frac{\partial g_{yy}(\lambda|\gamma_{yy}(0),\gamma_{yy}(1),\alpha,0)}{\partial\gamma_{yy}(0)} &= 1,\\
\frac{\partial g_{yy}(\lambda|\gamma_{yy}(0),\gamma_{yy}(1),\alpha,0)}{\partial\gamma_{yy}(1)} &= \frac{2(\cos\lambda-\alpha)}{1+\alpha^2-2\alpha\cos\lambda},\\
\frac{\partial g_{yy}(\lambda|\gamma_{yy}(0),\gamma_{yy}(1),\alpha,0)}{\partial\alpha} &= 2\frac{\left(2\cos^2\lambda-2\alpha\cos\lambda+\alpha^2-1\right)}{\left(1+\alpha^2-2\alpha\cos\lambda\right)^2}\gamma_{yy}(1),
\end{aligned}
$$

and

$$
\frac{\partial g_{yy}(\lambda|\gamma_{yy}(0),\gamma_{yy}(1),\alpha,0)}{\partial\psi_u}=2\frac{\left(2\cos^2\lambda-\alpha\cos\lambda-1\right)}{1+\alpha^2-2\alpha\cos\lambda}(\gamma_{yy}(1)-\alpha\gamma_{yy}(0))
$$

Let us double check these expressions using the chain rule. The partial derivatives of the spectral density (C9) with respect to the original parameters are:

$$
\begin{aligned}
\frac{\partial g_{yy}(\lambda|\sigma_f^2,\sigma_u^2,\alpha,\psi_u)}{\partial\sigma_f^2} &= \frac{1}{(1+\alpha^2-2\alpha\cos\lambda)},\\
\frac{\partial g_{yy}(\lambda|\sigma_f^2,\sigma_u^2,\alpha,\psi_u)}{\partial\sigma_u^2} &= \frac{1}{\left(1+\psi_u^2-2\psi_u\cos\lambda\right)},\\
\frac{\partial g_{yy}(\lambda|\sigma_f^2,\sigma_u^2,\alpha,\psi_u)}{\partial\alpha} &= \frac{2\sigma_f^2(\cos\lambda-\alpha)}{\left(1+\alpha^2-2\alpha\cos\lambda\right)^2},\\
\frac{\partial g_{yy}(\lambda|\sigma_f^2,\sigma_u^2,\alpha,\psi_u)}{\partial\psi_x} &= \frac{2\sigma_u^2(\cos\lambda-\psi_x)}{(1+\psi_x^2-2\psi_x\cos\lambda)^2}.
\end{aligned}
$$

When $\psi_u = 0$, these derivatives reduce to

$$\begin{aligned}
\frac{\partial g_{yy}(\lambda|\sigma_f^2, \sigma_u^2, \alpha, 0)}{\partial \sigma_f^2} &= \frac{1}{(1+\alpha^2 - 2\alpha\cos\lambda)}, \\
\frac{\partial g_{yy}(\lambda|\sigma_f^2, \sigma_u^2, \alpha, 0)}{\partial \sigma_u^2} &= 1, \\
\frac{\partial g_{yy}(\lambda|\sigma_f^2, \sigma_u^2, \alpha, 0)}{\partial \alpha} &= \frac{2\sigma_f^2(\cos\lambda - \alpha)}{(1+\alpha^2 - 2\alpha\cos\lambda)^2}, \\
\frac{\partial g_{yy}(\lambda|\sigma_f^2, \sigma_u^2, \alpha, 0)}{\partial \psi_u} &= 2\sigma_u^2 \cos\lambda.
\end{aligned}$$

In view of (B3) and (B4), the elements of the Jacobian matrix of the original parameters in terms of the new parameters will be

$$\begin{aligned}
\frac{\partial \sigma_f^2}{\partial \gamma_{yy}(0)} &= \frac{(1-\alpha^2)\psi_u}{\psi_u - \alpha} \\
\frac{\partial \sigma_f^2}{\partial \gamma_{yy}(1)} &= \frac{1-\alpha^2}{\alpha - \psi_u} \\
\frac{\partial \sigma_f^2}{\partial \alpha} &= \frac{2\alpha\psi_u - 1 - \alpha^2}{(\alpha-\psi)^2}(\gamma_{yy}(1) - \psi_u\gamma_{yy}(0)) \\
\frac{\partial \sigma_f^2}{\partial \psi_u} &= \frac{(1-\alpha^2)}{(\alpha-\psi_u)^2}(\gamma_{yy}(1) - \alpha\gamma_{yy}(0))
\end{aligned}$$

and

$$\begin{aligned}
\frac{\partial \sigma_u^2}{\partial \gamma_{yy}(0)} &= \frac{(1-\psi_u^2)\alpha}{\alpha - \psi_u} \\
\frac{\partial \sigma_u^2}{\partial \gamma_{yy}(1)} &= \frac{1-\psi_u^2}{\psi_u - \alpha} \\
\frac{\partial \sigma_u^2}{\partial \alpha} &= \frac{(1-\psi_u^2)}{(\alpha-\psi_u)^2}\left(\gamma_{yy}(1) - \psi_u\gamma_{yy}(0)\right) \\
\frac{\partial \sigma_u^2}{\partial \psi_u} &= \frac{2\alpha\psi_u - 1 - \psi_u^2}{(\alpha-\psi)^2}(\gamma_{yy}(1) - \alpha\gamma_{yy}(0))
\end{aligned}$$

which under the null become

$$\begin{aligned}
\frac{\partial \sigma_f^2}{\partial \gamma_{yy}(0)} &= 0 \\
\frac{\partial \sigma_f^2}{\partial \gamma_{yy}(1)} &= \frac{1-\alpha^2}{\alpha} \\
\frac{\partial \sigma_f^2}{\partial \alpha} &= -\frac{1+\alpha^2}{\alpha^2}\gamma_{yy}(1) \\
\frac{\partial \sigma_f^2}{\partial \psi_u} &= \frac{(1-\alpha^2)}{\alpha^2}(\gamma_{yy}(1) - \alpha\gamma_{yy}(0))
\end{aligned}$$

and

$$
\begin{aligned}
\frac{\partial \sigma_u^2}{\partial \gamma_{yy}(0)} &= 1 \\
\frac{\partial \sigma_u^2}{\partial \gamma_{yy}(1)} &= -\frac{1}{\alpha} \\
\frac{\partial \sigma_u^2}{\partial \alpha} &= \frac{1}{\alpha^2}\gamma_{yy}(1) \\
\frac{\partial \sigma_u^2}{\partial \psi_u} &= -\frac{1}{\alpha^2}(\gamma_{yy}(1) - \alpha\gamma_{yy}(0))
\end{aligned}
$$

The chain rule for derivatives then implies that

$$
\begin{aligned}
\frac{\partial g_{yy}(\lambda)}{\partial \sigma_f^2} &= \frac{1}{(1+\alpha^2 - 2\alpha\cos\lambda)}, \\
\frac{\partial g_{yy}(\lambda)}{\partial \sigma_u^2} &= 1, \\
\frac{\partial g_{yy}(\lambda)}{\partial \alpha} &= \frac{2\sigma_f^2(\cos\lambda - \alpha)}{(1+\alpha^2 - 2\alpha\cos\lambda)^2}, \\
\frac{\partial g_{yy}(\lambda)}{\partial \psi_u} &= 2\sigma_u^2 \cos\lambda.
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial g_{yy}(\lambda|\gamma_{yy}(0), \gamma_{yy}(1),\alpha, 0)}{\partial \gamma_{yy}(0)} &= \frac{\partial g_{yy}(\lambda|\sigma_f^2, \sigma_u^2, \alpha, 0)}{\partial \sigma_f^2}\frac{\partial \sigma_f^2}{\partial \gamma_{yy}(0)} + \frac{\partial g_{yy}(\lambda|\sigma_f^2, \sigma_u^2, \alpha, 0)}{\partial \sigma_u^2}\frac{\partial \sigma_u^2}{\partial \gamma_{yy}(0)} = 1, \\
\frac{\partial g_{yy}(\lambda|\gamma_{yy}(0), \gamma_{yy}(1),\alpha, 0)}{\partial \gamma_{yy}(1)} &= \frac{\partial g_{yy}(\lambda|\sigma_f^2, \sigma_u^2, \alpha, 0)}{\partial \sigma_f^2}\frac{\partial \sigma_f^2}{\partial \gamma_{yy}(1)} + \frac{\partial g_{yy}(\lambda|\sigma_f^2, \sigma_u^2, \alpha, 0)}{\partial \sigma_u^2}\frac{\partial \sigma_u^2}{\partial \gamma_{yy}(1)} \\
&= \frac{2(\cos\lambda - \alpha)}{1+\alpha^2 - 2\alpha\cos\lambda}, \\
\frac{\partial g_{yy}(\lambda|\gamma_{yy}(0), \gamma_{yy}(1),\alpha, 0)}{\partial \alpha} &= \frac{\partial g_{yy}(\lambda|\sigma_f^2, \sigma_u^2, \alpha, 0)}{\partial \sigma_f^2}\frac{\partial \sigma_f^2}{\partial \alpha} + \frac{\partial g_{yy}(\lambda|\sigma_f^2, \sigma_u^2, \alpha, 0)}{\partial \sigma_u^2}\frac{\partial \sigma_u^2}{\partial \alpha} \\
+\frac{\partial g_{yy}(\lambda|\sigma_f^2, \sigma_u^2, \alpha, 0)}{\partial \alpha} &= 2\frac{\left(2\cos^2\lambda - 2\alpha\cos\lambda + \alpha^2 - 1\right)}{(1+\alpha^2 - 2\alpha\cos\lambda)^2}\gamma_{yy}(1),
\end{aligned}
$$

and

$$
\begin{aligned}
\frac{\partial g_{yy}(\lambda|\gamma_{yy}(0), \gamma_{yy}(1),\alpha, 0)}{\partial \psi_u} &= \frac{\partial g_{yy}(\lambda|\sigma_f^2, \sigma_u^2, \alpha, 0)}{\partial \sigma_f^2}\frac{\partial \sigma_f^2}{\partial \psi_u} + \frac{\partial g_{yy}(\lambda|\sigma_f^2, \sigma_u^2, \alpha, 0)}{\partial \sigma_u^2}\frac{\partial \sigma_u^2}{\partial \psi_u} \\
+\frac{\partial g_{yy}(\lambda|\sigma_f^2, \sigma_u^2, \alpha, 0)}{\partial \psi_u} &= 2\frac{\left(2\cos^2\lambda - \alpha\cos\lambda - 1\right)}{1+\alpha^2 - 2\alpha\cos\lambda}(\gamma_{yy}(1) - \alpha\gamma_{yy}(0)),
\end{aligned}
$$

where we have used the fact that

$$
\begin{aligned}
\sigma_f^2 &= \frac{(1-\alpha^2)\gamma_{yy}(1)}{\alpha}, \\
\sigma_u^2 &= -\frac{(\gamma_{yy}(1) - \alpha\gamma_{yy}(0))}{\alpha}
\end{aligned}
$$

under the null. Obviously, the first three derivatives are the same for all the models which reduce to an AR(1) plus white noise under the corresponding null.

If we now scale them by the inverse spectral density under the null, we get

$$\begin{aligned}
\frac{\partial g_{yy}(\lambda)}{\partial \sigma_v^2} g_{yy}^{-1}(\lambda) &= \frac{1}{\sigma_u^2 (1+\alpha^2 - 2\alpha\cos\lambda) + \sigma_f^2}, \\
\frac{\partial g_{yy}(\lambda)}{\partial \sigma_\varepsilon^2} g_{yy}^{-1}(\lambda) &= \frac{\left(1+\alpha^2 - 2\alpha\cos\lambda\right)}{\sigma_u^2 (1+\alpha^2 - 2\alpha\cos\lambda) + \sigma_f^2}, \\
\frac{\partial g_{yy}(\lambda)}{\partial \alpha} g_{yy}^{-1}(\lambda) &= \frac{2\sigma_f^2(\cos\lambda - \alpha)}{(1+\alpha^2 - 2\alpha\cos\lambda)} \frac{1}{\sigma_u^2 (1+\alpha^2 - 2\alpha\cos\lambda) + \sigma_f^2}, \\
\frac{\partial g_{yy}(\lambda)}{\partial \psi_x} g_{yy}^{-1}(\lambda) &= 2\sigma_f^2 \cos\lambda \frac{\left(1+\alpha^2 - 2\alpha\cos\lambda\right)}{\sigma_u^2 (1+\alpha^2 - 2\alpha\cos\lambda) + \sigma_f^2}.
\end{aligned}$$

If we take the factor $[\sigma_u^2 \left(1+\alpha^2 - 2\alpha\cos\lambda\right) + \sigma_f^2]^{-1}$ out, we are left with

$$\begin{aligned}
[\sigma_u^2 \left(1+\alpha^2 - 2\alpha\cos\lambda\right) + \sigma_f^2] \frac{\partial g_{yy}(\lambda)}{\partial \sigma_v^2} g_{yy}^{-1}(\lambda) &= 1, \\
[\sigma_u^2 \left(1+\alpha^2 - 2\alpha\cos\lambda\right) + \sigma_f^2] \frac{\partial g_{yy}(\lambda)}{\partial \sigma_\varepsilon^2} g_{yy}^{-1}(\lambda) &= \left(1+\alpha^2 - 2\alpha\cos\lambda\right), \\
[\sigma_u^2 \left(1+\alpha^2 - 2\alpha\cos\lambda\right) + \sigma_f^2] \frac{\partial g_{yy}(\lambda)}{\partial \alpha} g_{yy}^{-1}(\lambda) &= \frac{2\sigma_f^2(\cos\lambda - \alpha)}{(1+\alpha^2 - 2\alpha\cos\lambda)}, \\
[\sigma_u^2 \left(1+\alpha^2 - 2\alpha\cos\lambda\right) + \sigma_f^2] \frac{\partial g_{yy}(\lambda)}{\partial \psi_x} g_{yy}^{-1}(\lambda) &= 2\sigma_f^2 \cos\lambda \left(1+\alpha^2 - 2\alpha\cos\lambda\right).
\end{aligned}$$

At first sight, it may seem that we no longer have an equivalent test. However, if we make use of the trigonometric identity (A11), we can write the last derivative as

$$2\sigma_f^2 \left(1+\alpha^2\right)\cos\lambda - \alpha(1+\cos 2\lambda)\, .$$

□

Proposition 7

Consider model (26). The spectral score with respect to ψ_a will be given by the sum of the spectral scores with respect to ψ_x and ψ_u evaluated at $\psi_x = \psi_u$. More specifically, given that

$$\frac{\partial g_{yy}(\lambda)}{\partial \psi_x} = \frac{\partial g_{xx}(\lambda)}{\partial \psi_x}, \quad \frac{\partial g_{yy}(\lambda)}{\partial \psi_u} = \frac{\partial g_{uu}(\lambda)}{\partial \psi_u}$$

and that

$$\frac{\partial g_{xx}(\lambda)}{\partial \psi_x} = 2\cos\lambda g_{xx}(\lambda), \quad \frac{\partial g_{uu}(\lambda)}{\partial \psi_u} = 2\cos\lambda g_{uu}(\lambda)$$

under the null of $H_0 : \psi_x = \psi_u = 0$, the score of the spectral log-likelihood for the observed series y_t will be given by

$$2\sum_{j=0}^{T-1} \cos\lambda_j [g_{xx}(\lambda_j) + g_{uu}(\lambda_j)] g_{yy}^{-2}(\lambda_j)[2\pi I_{yy}(\lambda_j) - g_{yy}(\lambda_j)] = 2\sum_{j=0}^{T-1} \cos\lambda_j 2\pi I_{aa}(\lambda_j),$$

which involves the first circulant autocorrelation of the reduced form residuals a_t. An analogous proof applies to the MA tests. □

C Auxiliary results

Lemma 1 *The spectral density of the observed process generated according to (19) will be*

$$g_{yy}(\lambda) = \frac{1+\psi_f^2 - 2\psi_f \cos\lambda}{1+\alpha^2 - 2\alpha\cos\lambda}\sigma_f^2 + \sigma_u^2 \tag{C1}$$

and its autocovariances

$$\begin{aligned}
\gamma_{yy}(0) &= \frac{(1+\psi_f^2 - 2\alpha\psi_f)}{(1-\alpha^2)}\sigma_f^2 + \sigma_u^2, && \text{(C2)}\\
\gamma_{yy}(1) &= \frac{(1-\alpha\psi_f)(\alpha-\psi_f)}{(1-\alpha^2)}\sigma_f^2, && \text{(C3)}\\
\gamma_{yy}(j) &= \alpha\gamma_{xx}(j-1), \quad j \geq 2. && \text{(C4)}
\end{aligned}$$

Proof. Since f_t and u_t are orthogonal at all leads and lags, the expression for the spectral density follows directly from the expressions for the spectral density of an ARMA(1,1) process. The same is true for the autocovariances, where we simply have to add σ_u^2 up to the zero order term. □

Lemma 2 *The spectral density of the observed process generated according to (20) will be*

$$g_{yy}(\lambda) = \frac{\sigma_f^2}{\left(1+\psi_x^2 - 2\psi_x\cos\lambda\right)(1+\alpha^2 - 2\alpha\cos\lambda)} + \sigma_u^2 \tag{C5}$$

and its autocovariances

$$\begin{aligned}
\gamma_{yy}(0) &= \frac{(\alpha\psi_x + 1)\sigma_f^2}{(1-\alpha^2)\left(1-\psi_x^2\right)(1-\alpha\psi_x)} + \sigma_u^2, && \text{(C6)}\\
\gamma_{yy}(1) &= \frac{(\alpha+\psi_x)\sigma_f^2}{(1-\alpha^2)\left(1-\psi_x^2\right)(1-\alpha\psi_x)}, && \text{(C7)}\\
\gamma_{yy}(2) &= (\alpha+\psi_x)\gamma_{xx}(j-1) - \alpha\psi_x\gamma_{xx}(j-2), \quad j \geq 2. && \text{(C8)}
\end{aligned}$$

Proof. Given that the autoregressive polynomial is $1-(\alpha+\psi_x)L + \alpha\psi_x L^2$, the first autocorrelation of the signal can be obtained from the Yule-Walker equation

$$\rho_{xx}(1) = (\alpha+\psi_x) - \alpha\psi_x\rho_{xx}(1),$$

which yields

$$\rho_{xx}(1) = \frac{\alpha+\psi_x}{\alpha\psi_x + 1},$$

while the remaining ones can be obtained from the recursion

$$\rho_{xx}(j) = (\alpha+\psi_x)\rho_{xx}(j-1) - \alpha\psi_x\rho_{xx}(j-2), \quad j \geq 1.$$

As for the unconditional variance, we can use the fact that

$$\gamma_{xx}(0)[1-(\alpha+\psi_x)\rho_{xx}(1) + \alpha\psi_x\rho_{xx}(2)] = \sigma_f^2,$$

with

$$[1-(\alpha+\psi_x)\rho_{xx}(1)+\alpha\psi_x\rho_{xx}(2)=\left(1-\alpha^2\right)\left(1-\psi_x^2\right)\frac{1-\alpha\psi_x}{\alpha\psi_x+1}$$

to obtain

$$\gamma_{xx}(0)=\frac{(\alpha\psi_x+1)\sigma_f^2}{(1-\alpha^2)\left(1-\psi_x^2\right)(1-\alpha\psi_x)}.$$

Similarly, the spectral density will be

$$g_{xx}(\lambda)=\frac{\sigma_f^2}{\left(1+\psi_x^2-2\psi_x\cos\lambda\right)(1+\alpha^2-2\alpha\cos\lambda)}.$$

Since x_t and u_t are orthogonal at all leads and lags, the result follows. □

Lemma 3 *The spectral density of the observed process generated according to (22) will be*

$$g_{yy}(\lambda)=\frac{\sigma_f^2}{(1+\alpha^2-2\alpha\cos\lambda)\left(1+\psi_x^2-2\psi_x\cos 2\lambda\right)}+\sigma_u^2.$$

Proof. The proof is entirely analogue to the proof of Lemma 1. □

Lemma 4 *The spectral density of the observed process generated according to (24) will be*

$$g_{yy}(\lambda)=\frac{\sigma_f^2}{(1+\alpha^2-2\alpha\cos\lambda)}+\frac{\sigma_u^2}{\left(1+\psi_u^2-2\psi_u\cos\lambda\right)},\tag{C9}$$

while the autocovariances become

$$\gamma_{yy}(j)=\frac{\alpha^j}{1-\alpha^2}\sigma_f^2+\frac{\psi_u^j}{1-\psi_u^2}\sigma_u^2,\quad j\geq 2.\tag{C10}$$

Proof. The autocovariances of the signal are

$$\gamma_{xx}(j)=\alpha^j\frac{\sigma_f^2}{1-\alpha^2},\quad j\geq 0$$

while its spectral density is

$$g_{xx}(\lambda)=\frac{\sigma_f^2}{(1+\alpha^2-2\alpha\cos\lambda)}.$$

Similarly, the autocovariances of the noise are

$$\gamma_{uu}(j)=\psi_u^j\frac{\sigma_v^2}{1-\psi_u^2},\quad j\geq 0$$

while its spectral density

$$g_{uu}(\lambda)=\frac{\sigma_v^2}{\left(1+\psi_u^2-2\psi_u\cos\lambda\right)}.$$

Since we are assuming that f_t and v_t are uncorrelated at all leads and lags, the autocovariances and the spectral density of y_t will be the sum of those of their underlying components. □

Lemma 5 *The spectral density of the observed process generated according to (25) will be*

$$g_{yy}(\lambda) = \frac{\sigma_f^2}{(1+\alpha^2 - 2\alpha\cos\lambda)} + \left(1+\psi_u^2 - 2\psi_u\cos\lambda\right)\sigma_v^2,$$

while the autocovariances become

$$\begin{aligned}
\gamma_{yy}(0) &= \frac{1}{1-\alpha^2}\sigma_f^2 + (1+\psi_u^2)\sigma_u^2 \\
\gamma_{yy}(1) &= \frac{\alpha}{1-\alpha^2}\sigma_f^2 - \psi_u\sigma_u^2 \\
\gamma_{yy}(j) &= \frac{\alpha^j}{1-\alpha^2}\sigma_f^2, \quad j \geq 2.
\end{aligned}$$

Proof. The proof is entirely analogue to the proof of Lemma 4. □

Table 1

Monte Carlo rejection rates (%) of LM tests at 10%, 5%, 1% significance levels

ψ_x	ψ_u	LM signal	LM noise	joint LM	LM resid
		10.30	9.99	10.26	9.78
0	**0**	5.04	4.94	5.07	4.87
		0.90	0.83	0.86	0.81
		32.54	24.22	28.28	17.90
.5	**0**	22.22	14.80	17.72	10.44
		7.56	4.46	5.38	2.88
		13.44	13.44	12.58	12.50
0	**.5**	7.24	7.10	6.58	7.16
		1.48	1.56	1.42	1.64
		11.42	9.50	13.14	12.18
.5	**.5**	6.08	4.70	7.22	6.86
		1.42	0.86	1.64	1.64
		19.62	12.38	22.98	15.92
.6	**.3**	12.06	6.62	14.48	9.42
		3.36	1.48	4.04	2.32

Table 2

Monte Carlo rejection rates (%) of tests at 10%, 5%, 1% significance levels

ψ_x	ψ_u	LM2S	LM1S	LR	W	Wnc	% zeros
		10.37	8.63	14.33	21.77	35.54	
0	**0**	4.93	4.28	7.17	14.81	31.37	52.00
		0.99	0.71	1.40	6.52	23.97	

Table 3a

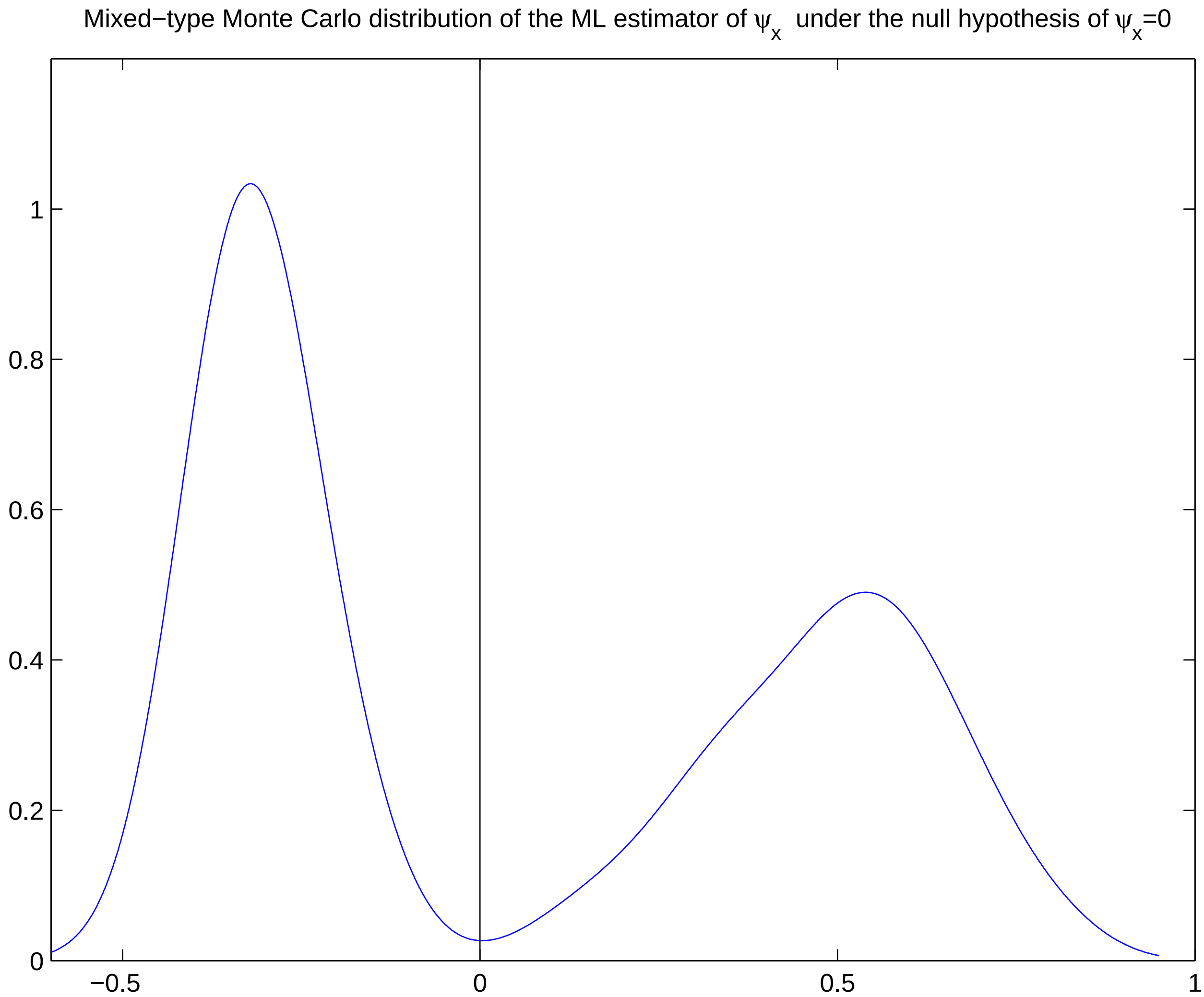
Monte Carlo rejection rates (%) of tests at 10%, 5%, 1% significance levels

ψ_x	ψ_u	LM2S	LM1S	LR	W	% zeros
		41.20	55.12	62.96	58.20	
.5	**0**	29.28	41.08	48.70	43.72	8.22
		12.34	17.78	24.08	22.52	
		9.42	8.38	14.52	24.52	
0	**.5**	4.80	4.00	7.00	19.16	53.60
		1.14	0.82	1.38	10.40	
		19.54	29.60	41.62	50.00	
.5	**.5**	11.94	18.72	28.14	39.66	22.16
		3.66	5.76	9.10	23.00	
		49.42	64.00	76.14	77.02	
.6	**.3**	37.14	49.32	63.68	64.78	5.54
		17.76	24.28	37.24	41.54	

Table 3b

Size-adjusted Monte Carlo rejection rates (%) of tests at 10%, 5%, 1%

ψ_x	ψ_u	LM2S	LM1S	LR	W
		40.38	58.32	54.82	31.32
.5	**0**	29.42	44.28	41.06	17.38
		12.66	20.72	20.16	4.56
		9.02	9.58	10.22	14.06
0	**.5**	4.90	4.80	4.72	8.36
		1.16	1.04	0.90	3.12
		18.92	32.52	34.08	30.64
.5	**.5**	12.00	20.84	21.70	18.02
		3.72	7.14	6.84	4.94
		48.56	66.52	69.74	52.16
.6	**.3**	37.40	52.92	55.40	34.32
		18.10	27.54	32.54	13.48


Mixed−type Monte Carlo distribution of the ML estimator of ψx under the null hypothesis of ψx=0
1
0.8
0.6
0.4
0.2
0
−0.5
0
0.5
1

CEMFI WORKING PAPERS

0801 *David Martinez-Miera and Rafael Repullo*: “Does competition reduce the risk of bank failure?”.

0802 *Joan Llull*: “The impact of immigration on productivity”.

0803 *Cristina López-Mayán:* “Microeconometric analysis of residential water demand”.

0804 *Javier Mencía and Enrique Sentana:* “Distributional tests in multivariate dynamic models with Normal and Student *t* innovations”.

0805 *Javier Mencía and Enrique Sentana:* “Multivariate location-scale mixtures of normals and mean-variance-skewness portfolio allocation”.

0806 *Dante Amengual and Enrique Sentana:* “A comparison of mean-variance efficiency tests”.

0807 *Enrique Sentana:* “The econometrics of mean-variance efficiency tests: A survey”.

0808 *Anne Layne-Farrar, Gerard Llobet and A. Jorge Padilla:* “Are joint negotiations in standard setting “reasonably necessary”?”.

0809 *Rafael Repullo and Javier Suarez:* “The procyclical effects of Basel II”.

0810 *Ildefonso Mendez:* “Promoting permanent employment: Lessons from Spain”.

0811 *Ildefonso Mendez:* “Intergenerational time transfers and internal migration: Accounting for low spatial mobility in Southern Europe”.

0812 *Francisco Maeso and Ildefonso Mendez:* “The role of partnership status and expectations on the emancipation behaviour of Spanish graduates”.

0813 *Rubén Hernández-Murillo, Gerard Llobet and Roberto Fuentes:* “Strategic online-banking adoption”.

0901 *Max Bruche and Javier Suarez:* “The macroeconomics of money market freezes”.

0902 *Max Bruche:* “Bankruptcy codes, liquidation timing, and debt valuation”.

0903 *Rafael Repullo, Jesús Saurina and Carlos Trucharte*: “Mitigating the procyclicality of Basel II”.

0904 *Manuel Arellano and Stéphane Bonhomme*: “Identifying distributional characteristics in random coefficients panel data models”.

0905 *Manuel Arellano, Lars Peter Hansen and Enrique Sentana*: “Underidentification?”.

0906 *Stéphane Bonhomme and Ulrich Sauder*: “Accounting for unobservables in comparing selective and comprehensive schooling”.

0907 *Roberto Serrano*: “On Watson’s non-forcing contracts and renegotiation”.

0908 *Roberto Serrano and Rajiv Vohra*: “Multiplicity of mixed equilibria in mechanisms: a unified approach to exact and approximate implementation”.

0909 *Roland Pongou and Roberto Serrano*: “A dynamic theory of fidelity networks with an application to the spread of HIV / AIDS”.

0910 *Josep Pijoan-Mas and Virginia Sánchez-Marcos:* “Spain is different: Falling trends of inequality”.

0911 *Yusuke Kamishiro and Roberto Serrano:* “Equilibrium blocking in large quasilinear economies”.

0912 *Gabriele Fiorentini and Enrique Sentana:* “Dynamic specification tests for static factor models”.

0913 *Javier Mencía and Enrique Sentana:* “Valuation of VIX derivatives”.

1001 *Gerard Llobet and Javier Suarez:* “Entrepreneurial innovation, patent protection and industry dynamics”.

1002 *Anne Layne-Farrar, Gerard Llobet and A. Jorge Padilla:* “An economic take on patent licensing: Understanding the implications of the “first sale patent exhaustion” doctrine.

1003 *Max Bruche and Gerard Llobet:* “Walking wounded or living dead? Making banks foreclose bad loans”.

1004 *Francisco Peñaranda and Enrique Sentana:* “A Unifying approach to the empirical evaluation of asset pricing models”.

1005 *Javier Suarez:* “The Spanish crisis: Background and policy challenges”.

1006 *Enrique Moral-Benito*: “Panel growth regressions with general predetermined variables: Likelihood-based estimation and Bayesian averaging”.

1007 *Laura Crespo and Pedro Mira:* “Caregiving to elderly parents and employment status of European mature women”.

1008 *Enrique Moral-Benito*: “Model averaging in economics”.

1009 *Samuel Bentolila, Pierre Cahuc, Juan J. Dolado and Thomas Le Barbanchon:* “Two-tier labor markets in the Great Recession: France vs. Spain”.

1010 *Manuel García-Santana and Josep Pijoan-Mas:* “Small Scale Reservation Laws and the misallocation of talent”.

1101 *Javier Díaz-Giménez and Josep Pijoan-Mas:* “Flat tax reforms: Investment expensing and progressivity”.

1102 *Rafael Repullo and Jesús Saurina:* “The countercyclical capital buffer of Basel III: A critical assessment”.

1103 *Luis García-Álvarez and Richard Luger:* “Dynamic correlations, estimation risk, and portfolio management during the financial crisis”.

1104 *Alicia Barroso and Gerard Llobet:* “Advertising and consumer awareness of new, differentiated products”.

1105 *Anatoli Segura and Javier Suarez:* “Dynamic maturity transformation”.

1106 *Samuel Bentolila, Juan J. Dolado and Juan F. Jimeno:* “Reforming an insider-outsider labor market: The Spanish experience”.

1201 *Dante Amengual, Gabriele Fiorentini and Enrique Sentana:* “Sequential estimation of shape parameters in multivariate dynamic models”.

1202 *Rafael Repullo and Javier Suarez:* “The procyclical effects of bank capital regulation”.

1203 *Anne Layne-Farrar, Gerard Llobet and Jorge Padilla:* “Payments and participation: The incentives to join cooperative standard setting efforts”.

1204 *Manuel Garcia-Santana and Roberto Ramos:* “Dissecting the size distribution of establishments across countries”.

1205 *Rafael Repullo:* “Cyclical adjustment of capital requirements: A simple framework”.

1206 *Enzo A. Cerletti and Josep Pijoan-Mas:* “Durable goods, borrowing constraints and consumption insurance”.

1207 *Juan José Ganuza and Fernando Gomez:* “Optional law for firms and consumers: An economic analysis of opting into the Common European Sales Law”.

1208 *Stéphane Bonhomme and Elena Manresa:* “Grouped patterns of heterogeneity in panel data”.

1209 *Stéphane Bonhomme and Laura Hospido:* “The cycle of earnings inequality: Evidence from Spanish Social Security data”.

1210 *Josep Pijoan-Mas and José-Víctor Ríos-Rull:* “Heterogeneity in expected longevities”.

1211 *Gabriele Fiorentini and Enrique Sentana:* “Tests for serial dependence in static, non-Gaussian factor models”.

1301 *Jorge De la Roca and Diego Puga:* “Learning by working in big cities”.

1302 *Monica Martinez-Bravo:* “The role of local officials in new democracies: Evidence from Indonesia”.

1303 *Max Bruche and Anatoli Segura:* “Debt maturity and the liquidity of secondary debt markets”.

1304 *Laura Crespo, Borja López-Noval and Pedro Mira:* “Compulsory schooling, education and mental health: New evidence from SHARELIFE”.

1305 *Lars Peter Hansen:* “Challenges in identifying and measuring systemic risk”.

1306 *Gabriele Fiorentini and Enrique Sentana:* “Dynamic specification tests for dynamic factor models”.

1307 *Diego Puga and Daniel Trefler:* “International trade and institutional change: Medieval Venice’s response to globalization”.

1308 *Gilles Duranton and Diego Puga:* “The growth of cities”.

1309 *Roberto Ramos:* “Banning US foreign bribery: Do US firms win?”.

1310 *Samuel Bentolila, Marcel Jansen, Gabriel Jiménez and Sonia Ruano:* “When credit dries up: Job losses in the Great Recession”.

1401 *Felipe Carozzi and Luca Repetto:* “Sending the pork home: Birth town bias in transfers to Italian municipalities”.

1402 *Anatoli Segura:* “Why did sponsor banks rescue their SIVs? A signaling model of rescues”.

1403 *Rosario Crinò and Laura Ogliari:* “Financial frictions, product quality, and international trade”.

1404 *Monica Martinez-Bravo:* “Educate to lead? The local political economy effects of school construction in Indonesia”.

1405 *Pablo Lavado:* “The effect of a child on female work when family planning may fail”.

1406 *Gabriele Fiorentini and Enrique Sentana:* “Neglected serial correlation tests in UCARIMA models”.