Highly irregular serial correlation tests*

Dante Amengual
CEMFI
<amengual@cemfi.es>

Xinyue Bei
Duke University
<xinyue.bei@duke.edu>

Enrique Sentana
CEMFI
<sentana@cemfi.es>

May 2023
Revised: January 2024

Abstract

Tests are developed for neglected serial correlation when the information matrix is repeatedly singular under the null hypothesis. Specifically, consideration is given to white noise against a multiplicative seasonal AR model, and a local-level model against a nesting UCARIMA one. The proposed tests, which involve higher-order derivatives, are asymptotically equivalent to the likelihood ratio test but only require estimation under the null. It is shown that the tests effectively check that certain autocorrelations of the observations are zero, so their asymptotic distribution is standard. Monte Carlo exercises examine finite sample size and power properties, with comparisons made to alternative approaches.

Keywords: Generalized extremum tests, Higher-order identifiability, Likelihood ratio test.

JEL: C22, C32, C52, C12

*Corresponding author: Enrique Sentana, CEMFI, Casado del Alisal 5, 28014 Madrid, Spain. <sentana@cemfi.es>

1 Introduction

The econometric literature on serial correlation tests, which can be traced back at least to Durbin and Watson (1950, 1951) and the references therein, is vast. Although in principle one could use any of the triad of classical hypothesis tests, given that Rao's (1948) score test and Silvey's (1959) numerically equivalent Lagrange multiplier (LM) statistic only require estimation of the model parameters under the null, they became the preferred choice for neglected serial correlation tests in econometric applications following Breusch (1978) and Godfrey (1978a,b), to the extent that they are nowadays routinely reported by all the popular regression packages. In addition to computational considerations, which continue to be very relevant for resampling procedures with relatively small sample sizes, two other important advantages of LM tests are that (i) rejections provide a clear indication of the specific directions along which modelling efforts should focus, and (ii) they are often easy to interpret as moment tests, so they remain informative for alternatives they are not designed for. Furthermore, under standard regularity conditions, they are asymptotically equivalent to the Likelihood ratio (LR) and Wald tests under the null and sequences of local alternatives, and thus they share their optimality properties.

One of those standard regularity conditions is a full rank information matrix of the unrestricted model parameters evaluated under the null. However, Fiorentini and Sentana (2016) highlighted some examples of neglected serial correlation tests in which this condition does not hold despite the fact that the model parameters are locally identified both under the null and the alternative hypotheses. To tackle this problem, they applied the "extremum" tests proposed by Lee and Chesher (1986). These authors studied situations in which one of the scores of the parameters of the model under the alternative is identically zero when evaluated under the null. Given that this renders standard LM tests infeasible, Lee and Chesher (1986) exploited the restrictions that the null hypothesis imposes on higher-order optimality conditions. Sometimes, the second derivative suffices, but it might be necessary to study the third or even higher-order ones. They proved the asymptotic equivalence between their extremum tests and the corresponding LR tests under the null and sequences of local alternatives in unrestricted contexts. Using earlier results by Cox and Hinkley (1974), this equivalence intuitively follows from the fact that the extremum tests can often be re-interpreted as standard LM tests of a suitable transformation of the parameter whose score is zero such that the new information matrix is no longer singular. For example, if the first two derivatives are identically zero when evaluated under the null but the third one is not, a cubic root provides an appropriate transformation which leads to a non-zero score after applying L'Hôpital's rule twice in succession. Naturally, the LR test is numerically invariant to this one-to-one transformation. In contrast, Wald tests are extremely sensitive to reparametrization under these circumstances.

Importantly, though, in all the examples Lee and Chesher (1986) and Fiorentini and Sentana (2016) discussed, the nullity of the information matrix of the alternative model under the null was one. The purpose of this paper is to develop tests for neglected serial correlation asymptotically equivalent to the LR test in some highly irregular situations in which the nullity of the information matrix is two or higher. To do so, we rely on the generalized extremum tests (GET) proposed in Amengual, Bei and Sentana (2023).

To understand our procedure, it is pedagogically convenient to consider the simplest possible situation of an information matrix with a zero 2×2 diagonal block because the scores of two of the parameters of the model under the alternative are identically zero when evaluated under the null. For simplicity, suppose the second-order derivatives are all different from zero. Unfortunately, a mere reparametrization will not solve the problem in this case because the number of distinct elements of the Hessian (three) exceeds the number of parameters affected by the singularities (two). For that reason, our solution involves two steps. First, we express the two parameters affected by the singularities in polar coordinates, which effectively correspond to the angle and length of their Cartesian representation on the real plane. For a fixed value of the angle, testing the null hypothesis is equivalent to testing that the Euclidean length of the parameter vector is zero, a unidimensional problem to which we could apply the Lee and Chesher (1986) solution. Unfortunately, the angle becomes underidentified under the null, so the second step of our solution relies on the supremum of their test statistic over all possible values of the angle as likelihood ratio analogue in the spirit of Davies' (1987).

More generally, GET is an LR-type test that compares the log-likelihood function under the null to the maximum of its lowest-order expansion under the alternative capable of identifying the restricted parameters.

For illustrative purposes, we use as examples two classes of univariate time series models very popular among practitioners:

1. the multiplicative seasonal Arima (Sarima) models put forward by Box and Jenkins (1970) to capture the autocorrelation of series with strong seasonal patterns, such as their famous airline passenger example, and

2. the unobserved components Arima (Ucarima) models, which constitute the basis of the "structural time series" models studied by Harvey (1989) as a way of performing the classical decomposition of a time series into trend, cyclical, seasonal and irregular components.

(see Lippi and Reichlin (1992) for an insightful comparison of some important characteristics of these two models).

We show that our proposed tests effectively check that certain autocorrelations of the observations are zero, which in turn implies that their asymptotic distribution is standard. This is somewhat remarkable because GET statistics typically have unusual asymptotic distributions (see e.g. Amengual, Bei and Sentana (2022)).

We conduct Monte Carlo exercises that study the finite sample size and power properties of our proposal and compare it to other tests for neglected serial correlation. We find that our suggested parametric bootstrap procedures yield very reliable test sizes for the small samples typically encountered in empirical applications to macroeconomic data. In addition, we confirm the power superiority of our tests over their competitors. Finally, we also confirm their substantial computational advantages over the corresponding LR tests, which require the maximization over the entire parameter space of an unrestricted log-likelihood function which is extremely flat around its maximum when the null hypothesis is true. These computational advantages are particularly pertinent for computing the bootstrap critical values mentioned above.

The rest of the paper is organized as follows. We derive our proposed tests for the two aforementioned examples in Sections 2 and 3, respectively, studying both their asymptotic properties and their finite sample ones. Next, we present our conclusions in Section 4, relegating proofs and some additional results to the appendices.

2 Multiplicative seasonal ARIMA models

The serial dependence structure of the popular multiplicative seasonal ARIMA models put forward by Box and Jenkins (1970) is perfectly understood, and the same is true of the properties of the maximum likelihood estimators (MLE) of their parameters in normal circumstances. Moreover, LM tests for neglected serial correlation in such models have been readily available for several decades. However, what it is far less known is that in some cases, the standard regularity conditions that guarantee the asymptotic validity of such tests do not hold. Next, we showcase the difficulties involved by means of a rather simple example.

2.1 The test statistic

Suppose that after taking regular and seasonal differences of an observed time series, a researcher would like to formally assess the need for a more complicated dependence structure. Specifically, assuming the data is observed at the quarterly frequency, and letting L denote the lag operator, one of the alternatives that a researcher might consider is the following AR(2)-SAR(2) process:

$$(1-\vartheta_1 L)(1-\vartheta_2 L)(1-\vartheta_3 L^4)(1-\vartheta_4 L^4)(y_t-\varphi_M)=\varepsilon_t, \tag{1}$$

with $E(\varepsilon_t) = 0$ and $V(\varepsilon_t) = \varphi_V$, where $y_t = \Delta\Delta_4 x_t$ with $\Delta = 1 - L$ and $\Delta_4 = 1 - L^4$ denoting the usual regular and seasonal difference operators, and x_t denoting the original data, so that $H_0 : \boldsymbol{\vartheta} = \mathbf{0}$, with $\boldsymbol{\vartheta} = (\vartheta_1, \vartheta_2, \vartheta_3, \vartheta_4)'$.

As usual, non-linear least squares estimation coincides with Gaussian ML, so that the criterion function will be

$$\sum_{t=1}^{T} l_t \quad \text{with} \quad l_t = -\frac{1}{2}\ln(2\pi) - \frac{1}{2}\ln\varphi_V - \frac{[y_t - \mu_t(\varphi_M, \boldsymbol{\vartheta})]^2}{2\varphi_V}, \tag{2}$$

where the conditional mean under the alternative is

$$\begin{aligned}\mu_t(\varphi_M, \boldsymbol{\vartheta}) =& \varphi_M + (\vartheta_1 + \vartheta_2)(y_{t-1} - \varphi_M) - \vartheta_1\vartheta_2(y_{t-2} - \varphi_M) + (\vartheta_3 + \vartheta_4)(y_{t-4} - \varphi_M) \\ & - (\vartheta_1 + \vartheta_2)(\vartheta_3 + \vartheta_4)(y_{t-5} - \varphi_M) + \vartheta_1\vartheta_2(\vartheta_3 + \vartheta_4)(y_{t-6} - \varphi_M) \\ & - \vartheta_3\vartheta_4(y_{t-8} - \varphi_M) + (\vartheta_1 + \vartheta_2)\vartheta_3\vartheta_4(y_{t-9} - \varphi_M) - \vartheta_1\vartheta_2\vartheta_3\vartheta_4(y_{t-10} - \varphi_M).\end{aligned}$$

The model parameters under the null are φ_M and φ_V, whose restricted MLEs coincide with the sample mean and variance (with denominator T) of y_t. Moreover, the MLEs of the parameters of the alternative model, which also include $\boldsymbol{\vartheta}$, usually converge to their true values at the standard $\sqrt{T}$ rate.

However, as we shall formally prove below, the information matrix of model (1) evaluated at $\boldsymbol{\vartheta} = \mathbf{0}$ has two zero eigenvalues because

$$\frac{\partial l_t}{\partial\vartheta_1} - \frac{\partial l_t}{\partial\vartheta_2} = 0 \text{ and } \frac{\partial l_t}{\partial\vartheta_3} - \frac{\partial l_t}{\partial\vartheta_4} = 0, \tag{3}$$

which makes this testing problem a highly irregular one. This is particularly relevant for Wald tests, which are extremely sensitive to reparametrizations in this context. For example, Fiorentini and Paruolo (2009) found that the rate of convergence of a sequential Cochrane-Orcutt-type estimator of what is effectively the product of the first two autocorrelations of y_t is T rather than $T^{\frac{1}{2}}$ or $T^{\frac{1}{4}}$ when $\vartheta_1 = \vartheta_2 = 0$ in a non-seasonal version of model (1) in which $\vartheta_3 = \vartheta_4 = 0$.

As we show in the proof of Proposition 1, we can find a suitable reparametrization relating $(\varphi_M, \varphi_V, \vartheta_1, \vartheta_2, \vartheta_3, \vartheta_4)$ and $(\phi_M, \phi_V, \theta_{i1}, \theta_{i2}, \theta_{u1}, \theta_{u2})$ that isolates the singularity in the last two parameters in such a way that the first derivatives of the log-likelihood function corresponding to θ_{u1} and θ_{u2} are both zero, where $\boldsymbol{\theta}_i' = (\theta_{i1}, \theta_{i2})$ contains the parameters of the alternative model that are first-order identified while $\boldsymbol{\theta}_u' = (\theta_{u1}, \theta_{u2})$ refers to those that are first-order underidentified but second-order identified in the terminology of Sargan (1983).

Fortunately, the assumptions of Theorem 1 in Amengual, Bei and Sentana (2023) apply to the second derivatives

$$\frac{\partial^2 l_t}{(\partial\theta_{u1})^2} = \frac{2(y_t - \phi_M)(y_{t-2} - \phi_M)}{\phi_V}, \quad \frac{\partial^2 l_t}{\partial\theta_{u1}\partial\theta_{u2}} = 0 \text{ and } \frac{\partial^2 l_t}{(\partial\theta_{u2})^2} = \frac{2(y_t - \phi_M)(y_{t-8} - \phi_M)}{\phi_V}, \tag{4}$$

because the asymptotic covariance matrix of

$$\left[\frac{\partial l_t}{\partial \phi_M}, \frac{\partial l_t}{\partial \phi_V}, \frac{\partial l_t}{\partial \theta_{i1}}, \frac{\partial l_t}{\partial \theta_{i2}}, \theta_{u1}^2 \frac{\partial^2 l_t}{(\partial \theta_{u1})^2} + \theta_{u2}^2 \frac{\partial^2 l_t}{(\partial \theta_{u2})^2} + 2\theta_{u1}\theta_{u2} \frac{\partial^2 l_t}{\partial \theta_{u1} \partial \theta_{u2}}\right]$$

scaled by $\sqrt{T}$ has full rank for any $(\theta_{u1}, \theta_{u2}) \neq (0, 0)$, which leads to the following result:

Proposition 1 *Under H_0,*

$$LR_T = GET_T + O_p(T^{-\frac{1}{4}}), \tag{5}$$

where LR_T is the likelihood ratio statistic based on (2), and

$$GET_T = T(\hat{r}_{1T}^2 + \hat{r}_{4T}^2 + \hat{r}_{2T}^2 \mathbf{1}[\hat{r}_{2T} \geq 0] + \hat{r}_{8T}^2 \mathbf{1}[\hat{r}_{8T} \geq 0]), \tag{6}$$

where $\mathbf{1}[.]$ is the usual indicator function and

$$\hat{r}_{jT} = \frac{1}{T} \sum\nolimits_t \frac{(y_t - \tilde{\phi}_M)(y_{t-j} - \tilde{\phi}_M)}{\tilde{\phi}_V},$$

with $\tilde{\phi}_M = T^{-1} \sum_t y_t$ and $\tilde{\phi}_V = T^{-1} \sum_t (y_t - \tilde{\phi}_M)^2$.

Therefore, the GET_T statistic is simply focusing on the first two regular sample autocorrelations and the first two seasonal ones, which is very intuitive in view of (1). Given that these estimated autocorrelations are asymptotically independent under the null of white noise, the asymptotic distribution of (6) will be a mixture of χ_2^2, χ_3^2 and χ_4^2 with weights $\frac{1}{4}$, $\frac{1}{2}$ and $\frac{1}{4}$, respectively. The partially one-sided nature of the test arises from the multiplicative nature of the alternative, which forces the roots to be always real. Additive alternatives, which allow for complex roots too, would give rise to two-sided tests.

Furthermore, we can show that the GET test statistic of white noise against the multiplicative AR(k)-SAR(k_s) model

$$\prod_{j=1}^{k} (1 - \vartheta_j L) \prod_{j=k+1}^{k+k_s} (1 - \vartheta_j L^4)(y_t - \varphi_M) = \varepsilon_t$$

for $k \geq 3$ or $k_s \geq 3$ will numerically coincide with the statistic in (6). The rationale is as follows. When the null is true, we can prove that the MLE of an additive AR(3) is such that all three roots of the lag polynomial are real with probability tending to zero, unless one of the roots is forced to be zero. Consequently, the LR for multiplicative AR(3) is asymptotically equivalent to the LR for AR(2), and the same applies to the corresponding GETs. Perhaps less surprisingly, we can also show that we would obtain exactly the same test statistic if we considered multiplicative MA alternatives instead.

To provide the intuition for the convergence rate in (5), it is convenient to look at expression (A3) in Appendix A. Specifically, given that we can write the LR test statistic in terms of a fourth-order Taylor expansion of the log-likelihood function whose leading terms coincide with the ones that appear in the expression for the GET test statistic, the $O_p(T^{-\frac{1}{4}})$ rate of the remainder immediate determines the rate of convergence of the difference between the two statistics under the null.

Finally, it is important to mention that our proposed test, which is based on sample autocorrelations, is numerically invariant to affine transformations of the observed series y_t. Effectively, this means that its finite sample distribution is pivotal with respect to $\boldsymbol{\varphi} = (\varphi_M, \varphi_V)'$. Therefore, one can estimate the sample mean and variance of y_t, and apply our test directly to the standardized series as if they were the observed variables.

2.2 Simulation evidence

Next, we study the finite sample size and power properties of the testing procedures we introduced in the previous subsection by means of several extensive Monte Carlo exercises. Given that no nuisance parameters are effectively involved under the null, we can set the unconditional mean and variance of the innovation ε_t to 0 and 1, respectively, both under the null and alternative hypotheses without any loss of generality.

Naturally, we estimate φ_M and φ_V under the null with the sample mean and variance, respectively. We recycle the sample mean as initial value for φ_M under the alternative. As for φ_V and $\boldsymbol{\vartheta}$, we use as starting values the ones we obtain by means of a minimum distance procedure that takes as objective function the Euclidean norm of the difference between the theoretical and sample values of the variance and the four autocovariances underlying our test statistic (6) in order to increase the chances that we obtain the right unrestricted ML estimates, and consequently, the correct LR test. Nevertheless, our results seem insensitive to this choice of initial values.

As alternative hypotheses we consider the covariance stationary models

$$\begin{aligned}(1-.1L-.1L^2-.1L^3-.1L^4)y_t &= \varepsilon_t \quad (H_{a_1}), \quad \text{and} \\ (1-.4L)(1+.4L)(1-.4L^4)(1+.4L^4)y_t &= \varepsilon_t \quad (H_{a_2}),\end{aligned}$$

for which the first, second, fourth and eighth autocorrelation coefficients in the population are (0.14,0.14,0.14,0.03) and (0,0.16,0.03,0.16). Note that two of the roots of the first process are complex conjugates, while our test is designed for the case of real roots.

We approximate the exact finite sample distribution using 10,000 simulated samples under the maintained hypothesis that the y_t's are *i.i.d.* as standard normals. In fact, we could thus obtain "exact" critical values for any sample size by increasing the number of simulations. Alternatively, one could consider a non-parametric bootstrap procedure that randomly draws with replacement from the observations, which would eliminate any time series dependence while allowing for any marginal distribution. Either way, we do not need to take into account the sensitivity of the critical values to $\tilde{\varphi}$ because the test statistics are numerically invariant to the values of these estimators.

In Table 1 we compare the results of our test with three alternative procedures: LM-AR(1) and LM-SAR(4), which denote standard LM tests based on the score of an AR(1) and a Wallis (1972)-style seasonal AR(4), respectively, and a moment test based on the first two regular sample autocorrelations and the first two seasonal ones (MT), which is effectively the two-sided version of (6), whose asymptotic distribution is χ^2_4 under the null. Specifically, Panel A of Table 1 contains the rejection rates based on asymptotic critical values for $T = 100$ (top) and $T = 400$ (bottom), while in Panel B we report the ones that rely on the parametric bootstrap.

The first three columns of each of those panels present the rejection rates at the 1%, 5% and 10% levels under the null. Given the number of replications, the 95% asymptotic confidence intervals for the Monte Carlo rejection probabilities under the null are (.80,1.20), (4.57,5.43) and (9.41,10.59) at the 1%, 5% and 10% levels. As can be seen in Panel A, all tests tend to be undersized at the usual nominal levels, with some significant size distortions across the board when $T = 100$. As expected, though, the rejection rates get much closer to the nominal sizes for $T = 400$. In contrast, the size of the tests becomes perfectly accurate by construction when we use the parametric bootstrap procedure described above.

In turn, the last six columns present the rejection rates at the 1%, 5% and 10% levels for the two alternatives we consider. The behavior of the different test statistics is in accordance with expectations. In particular, our proposal is the most powerful for H_{a_2}, which is not very surprising given that it is designed to direct power against such multiplicative alternatives with real roots. But it is also the top performer for H_{a_1} even though the process has two complex roots, which is perhaps not entirely surprising in view of the positivity of the relevant population autocorrelations. Predictably, the rejection rates in Panel B are slightly higher, which simply reflects the fact that all the tests tend to be conservative with the asymptotic critical values.

The scatterplot in Figure 1 visually illustrates the asymptotic equivalence under the null between LR_T and GET_T statistics stated in Proposition 1, with the Gaussian rank correlation coefficients between them being 0.932 and 0.986 across Monte Carlo samples of size $T = 100$ and 400, respectively. The Gaussian rank correlation coefficient between two variables is the usual Pearson correlation coefficient between the Gaussian scores of those variables, which are obtained by applying the inverse Gaussian cumulative distribution function transform to the ranks of the observations on each variable divided by $n + 1$ (see Amengual, Sentana and Tian

(2022)). Like the Spearman correlation coefficient, the Gaussian one is less sensitive to outliers than the Pearson one.

Finally, our results also indicate that the LR takes 755 (921) seconds of CPU time for 10,000 samples of length 100 (400), while computing GET only requires 0.20 (0.24) seconds, respectively, which makes a huge difference in the calculation of the bootstrap critical values.

3 UCARIMA models

These popular unobserved component models assume that the observed time series are the superposition of two or more latent ARIMA time series models, whose parameters can be estimated by maximizing 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 e.g. Harvey (1989) for a textbook treatment).

In contrast, tests that assess the correct specification of the parametric ARIMA models for the underlying components are far less well studied, even though the various outputs of an UCARIMA model could be misleading under misspecified dynamics. As mentioned in the introduction, Fiorentini and Sentana (2016) provided a thorough discussion of such tests, highlighting the popular local level model as an example in which the LM test cannot be computed in the usual way because the information matrix of the alternative model is sometimes singular under the null. Unfortunately, their solution based on Lee and Chesher (1986) cannot be applied when the nullity of the information matrix is two or more. Next, we study in detail a simple example of this situation.

3.1 The test statistic

The most popular UCARIMA model among practitioners is the local level process:

$$x_t = z_t + u_t, \tag{7}$$

$$\Delta z_t = f_t, \tag{8}$$

$$u_t = v_t \quad \text{and} \tag{9}$$

$$\begin{pmatrix} f_t \\ v_t \end{pmatrix} | I_{t-1} \sim N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_f^2 & 0 \\ 0 & \sigma_v^2 \end{pmatrix}\right], \tag{10}$$

where f_t and v_t follow two univariate white noise processes orthogonal at all leads and lags, I_{t-1} denotes the information set available at $t-1$ which contains past values of x_t, and σ_f^2 and σ_v^2 are both strictly positive to exclude degenerate cases. Thus, the observed series is simply a random walk plus noise, whose first differences $y_t = \Delta x_t$ follow an MA(1) process with coefficient

$$\beta_y = \frac{1}{2}(\sqrt{q^2+4q} - 2 - q) < 0, \tag{11}$$

where $q = \sigma_f^2/\sigma_v^2 > 0$ is the bounded signal to noise ratio, and residual variance $\sigma_a^2 = -\sigma_v^2/\beta_y$.

As is well known, this model justifies the popular Exponentially Weighted Moving Average (EWMA) prediction rule, which has proved remarkably successful in many applications ranging from macro time series to volatility forecasts. However, EWMA predictions become suboptimal if (8) or (9) are dynamically misspecified, so it makes sense to test them against some more general alternatives.

To illustrate the issues that may arise, we consider the following nesting model:

$$\left.\begin{array}{c}(1-\psi_1 L - \psi_2 L^2)\Delta z_t = f_t \\ (1-\alpha L)u_t = v_t\end{array}\right\} \tag{12}$$

in which the "signal" z_t follows an ARIMA(2,1,0) process while the "noise" u_t a stationary AR(1) process. As a result, the null hypothesis of interest is $H_0 : \alpha = \psi_1 = \psi_2 = 0$.

Once again, we can formally prove that the dimension of the nullspace of the information matrix of the parameters of model (12) evaluated under the null is two because the first-derivatives of the log-likelihood function corresponding to ψ_1 and ψ_2 are linear combinations of the ones corresponding to σ_f^2, σ_v^2 and α. In fact, we show in the proof of Proposition 2 that we can find a suitable reparametrization from $(\sigma_f^2, \sigma_v^2, \alpha, \psi_1, \psi_2)$ to $(\sigma_f^{2\dagger}, \sigma_v^{2\dagger}, \alpha^\dagger, \psi_1^\dagger, \psi_2^\dagger)$ that isolates the singularity in the last two parameters in such a way that the first-derivatives of the log-likelihood function corresponding to $\psi_1^\dagger$ and $\psi_2^\dagger$ are both zero.

Like Fiorentini and Sentana (2016), we can explicitly relate this singularity to the identification conditions for UCARIMA models in Hotta (1989). Specifically, although model (12) is generally identified, it is locally equivalent around the null to the following model:

$$\left.\begin{array}{c}\Delta z_t = (1-\psi_1 L - \psi_2 L^2) f_t \\ u_t = (1-\alpha L)v_t\end{array}\right\} \tag{13}$$

in the sense that the (absolute value of the) scores and information matrices are identical when H_0 holds. Unlike model (12), which generates the autocorrelation structure of a restricted ARMA(3,3) for y_t, model (13) generates the autocorrelation structure of an unrestricted MA(2), which depends on three parameters only, namely the two MA coefficients plus the variance of the reduced form innovations. In contrast, model (13) depends on five parameters, namely ψ_1, ψ_2 and α together with σ_f^2 and σ_v^2, which means that the MA(2) reduced form can only identify a manifold of dimension two of the structural parameters.

In addition, with the aforementioned reparametrization,

$$\frac{\partial^2 l_t}{(\partial\psi_1^\dagger)^2} = 0 \text{ and } \frac{\partial^3 l_t}{(\partial\psi_1^\dagger)^3} = 0, \text{ while } \frac{\partial^2 l_t}{(\partial\psi_2^\dagger)^2} \neq 0,$$

which means that these two parameters have different degrees of identification. Fortunately, the assumptions of the more general Theorem 2 in Amengual, Bei and Sentana (2023) apply, allowing us to obtain the following result:

Proposition 2 *Under H_0,*

$$LR_T = GET_T + O_p(T^{-\frac{1}{8}}), \tag{14}$$

where LR_T is the corresponding likelihood ratio statistic, and

$$\begin{aligned} GET_n &= \begin{pmatrix} \tilde{r}_{2T} & \tilde{r}_{3T} & \tilde{r}_{4T} \end{pmatrix} \mathcal{V}_{\boldsymbol{\rho}_a\boldsymbol{\rho}_a}^{-1} \begin{pmatrix} \tilde{r}_{2T} \\ \tilde{r}_{3T} \\ \tilde{r}_{4T} \end{pmatrix}, \quad \textit{with} \\ \tilde{r}_{jT} &= \frac{\sum_t y_t y_{t-j}}{\sum_t y_t^2} \textit{ and } \mathcal{V}_{\boldsymbol{\rho}_a\boldsymbol{\rho}_a} = \lim_{T\to\infty} V\left[\sqrt{T}\begin{pmatrix} \tilde{r}_{2T} \\ \tilde{r}_{3T} \\ \tilde{r}_{4T} \end{pmatrix}\right]. \end{aligned} \tag{15}$$

Therefore, both LR_T and GET_T are effectively testing that the second, third and fourth autocorrelations of y_t are zero. This result is not entirely surprising in view of the fact that y_t follows an MA(1) model under the null and an ARMA(3,3) under the alternative. Unlike what happened in the model discussed in Section 2, though, the sample autocorrelations are no longer asymptotically independent under the null, so we need their asymptotic covariance matrix to correct for sampling uncertainty, which is particularly simple to obtain in the frequency domain using the expressions in Appendix B.1, as we explain in the proof of Proposition 2.

Finally, expressions (A4)–(A8) in Appendix A provide intuition on the convergence rate in (14). Specifically, given that we can write the LR test statistic in terms of a eighth-order Taylor expansion of the log-likelihood function whose leading terms coincide with the terms that appear in the expression of the GET test, the $O_p(T^{-\frac{1}{8}})$ rate of the remainder immediate determines the rate of convergence of the two statistics under the null.

3.2 Simulation evidence

To assess the size properties of our proposed test, we generate 10,000 samples of lengths $T = 100$ and $T = 400$ of the local level model (7)-(10). Under the null, we simulate Gaussian shocks with $\sigma_f^2 = 1$ and $\sigma_v^2 = 0.5$, so that the signal to noise ratio is moderate.

We compute the spectral version of GET_T in (A10) using (B14) to estimate the information matrix (A9) and the fast Fourier transform to obtain the periodogram. It is important to emphasize that the LR_T statistic requires the estimation of model (12), which is a non-trivial numerical task for the reasons described in the introduction. To increase the chances that we obtain the correct unrestricted ML estimates, and consequently, the right LR test, we keep the maximum maximorum of the spectral log-likelihood of model (12) starting from two sets of initial values: the ones that maximize the log-likelihood function under the null, and another set that we obtain by means of a minimum distance procedure that takes as objective function the Euclidean norm of the difference between the theoretical and sample values of the variance and the first four autocovariances of the process, whose expressions we provide in Appendix B.2.

Although our main interest lies in the GET_T and LR_T statistics in Proposition 2, we also consider the following two moment tests for comparison purposes:

1. no second-order serial correlation in y_t,
2. no second- or third-order serial correlation in y_t.

Importantly, in computing these moment tests, we use the relevant elements of (A9) to obtain the adjusted asymptotic covariance matrix of the second and third sample autocovariances.

Unlike what happens in the multiplicative seasonal ARIMA model in Section 2 in which the autocorrelations did not depend on the mean and variance of y_t, the finite sample distribution of GET_T and LR_T is not pivotal with respect to the (unknown) value of the signal to noise ratio q, even though both statistics are numerically invariant to the scale of y_t. Intuitively, the value of q affects the autocorrelations on which our test is based, so we need to adjust for the sampling variability of its estimators under the null, both asymptotically and in the bootstrap. For that reason, we conduct a parametric bootstrap procedure whereby for each of those 10,000 simulated samples, we simulate another $NB-1$ samples in which we set σ_f^2 equal to one without loss of generality and $(1+q)^{-1}$ to its estimated value, so that we can automatically compute size-adjusted rejection rates, as forcefully argued by Horowitz and Savin (2000). In fact, the bounded support of $(1+q)^{-1}$ allows us to compute a table of "exact" critical values for a fine grid of values of this reduced-form MA coefficient before running the actual simulations (see Appendix D.1 in Amengual and Sentana (2015) for details). The same procedure works if we replace $(1+q)^{-1}$ by either β_y in (11) or the first-order autocorrelation of y_t, which are both between zero and minus one, but it is trickier to apply to q directly because this parameter can take any positive real value in the sample.

We present the rejection rates under the null for the tests at the 10%, 5% and 1% in the first three columns of Table 2. Once more, in Panel A we report the results based on asymptotic critical values for samples of length 100 (top) and 400 (bottom), and in Panel B those based on the bootstrap. Given that the number of Monte Carlo simulations is also 10,000, the 95% asymptotic confidence intervals for the rejection probabilities under the null are again (.80,1.20), (4.57,5.43) and (9.41,10.59) at the 1%, 5% and 10% levels.

In terms of size distortions, the same comments we made in the seasonal ARIMA example apply here too, with only two exceptions: (i) the LR test is now slightly oversized, and (ii) the bootstrap no longer provides exact critical values because of the need to estimate q. Nevertheless, all the testing procedures have reasonable sizes when we rely on bootstrap critical values regardless of the sample length. It is also worth mentioning that in a very small fraction of the samples of size $T = 100$ simulated under the null (0.62%), we encountered the "pile-up" problem associated to a positive first-order sample autocorrelation for y_t. In contrast, this never happened under either of the alternatives below, or indeed when $T = 400$.

In particular, we simulate and estimate 10,000 samples of the same length of the following two alternative data generation processes (DGPs):

$$\left.\begin{array}{r}(1+0.5L+0.4L^2)\Delta z_t = f_t \\ (1-0.5L)u_t = v_t\end{array}\right\} \quad (H_{a_1}) \quad \text{and}$$

$$\left.\begin{array}{r}(1-0.1L+0.5L^2)\Delta z_t = f_t \\ (1+0.5L)u_t = v_t\end{array}\right\} \quad (H_{a_2})$$

with the same σ_f^2 and σ_v^2 as in the null hypothesis. The first four autocorrelation coefficients of these processes in the population are (-0.32,-0.19,0.15,-0.04) and (-0.42,0.03,-0.15,0.15).

The corresponding rejection rates, which we report in the last six columns of Table 2, indicate that the behavior of the different test statistics is in accordance with expectations. For both alternatives, the GET and LR tests are more powerful than the competitors. Interestingly, LR is slightly more powerful than our proposal for both H_{a1} and H_{a_2}, which is in contrast with the ranking in the example in Section 2. Nevertheless, one should keep in mind that our equivalence result is an asymptotic one under the null and, presumably, suitable sequences of local alternatives, while the sample sizes we use in our simulations are moderately small and we are effectively considering fixed alternatives. In this respect, the scatterplot in Figure 2 visually illustrates the asymptotic equivalence under the null between LR_T and GET_T in Proposition 2, with the Gaussian rank correlation coefficients between the GET and LR test statistics across Monte Carlo samples of size $T = 100$ and 400 generated under the null being 0.743 and 0.807, respectively, reflecting the slower rate of convergence.

The simulation results also indicate that the LR takes 1,250 (1,763) seconds for 10,000 samples of length 100 (400), while computing GET only requires 4.5 (5.5) seconds, respectively, which once again makes a huge difference in the calculation of the bootstrap critical values.

4 Conclusions

We characterize the singularity of the information matrix of a multiplicative seasonal AR model à la Box and Jenkins under the null of white noise, as well as of a trend plus signal

Ucarima model that nests the popular local level process. Using the generalization in Amengual, Bei and Sentana (2023) of the extremum-type tests in Lee and Chesher (1986) to models in which the nullity of the information matrix under the null hypothesis is strictly larger than one, we explain how to obtain a score-type test based on higher-order derivatives which is asymptotically equivalent to the LR despite said singularity but only requires estimation under the null. This is particularly relevant for resampling-based inference because the fact that several log-likelihood derivatives are zero under the null implies that the LR requires the estimation of all the parameters that appear under the alternative in a model whose log-likelihood function is extremely flat.

Our proposed dynamic specification tests are simple to implement and even simpler to interpret. And although some of our theoretical derivations make extensive use of frequency domain methods for time series, we provide a simple time domain interpretation of the statistics, so that empirical researchers who are not familiar with spectral analysis can still apply them easily.

We conduct Monte Carlo exercises that study the finite sample size and power properties of our proposals and compare them to alternative approaches. We find that our suggested parametric bootstrap procedures work very well, and that our tests have more power than alternative procedures. We also find that the computational advantages of our GET procedures relative to the LR ones are very substantial.

In the two examples that we consider the model parameters are only identified up to higher-order when the null is true. As a result, a local power analysis of our proposed tests would necessarily involve sequences of those parameters converging to zero at unusually low rates. Nevertheless, given that in both cases our test statistics have χ^2-like asymptotic distributions under the null, they would approximately follow non-central χ^2 distributions in large samples if we ignore inequality constraints. Finding exact expressions for the non-centrality parameters constitutes an interesting avenue for further research.

5 Acknowledgements

We would like to thank the audience at the XIII Workshop in Time Series Econometrics (Zaragoza, March 2023) for useful comments and suggestions, as well as Gabriele Fiorentini, who also shared with us the code mapping Arma coefficients with autocorrelations in Fiorentini and Planas (1998). We are also grateful to the associate editor and two anonymous referees for useful feedback. Of course, the usual caveat applies. Amengual and Sentana gratefully acknowledge financial support from the Spanish Ministry of Science and Innovation through grant PID2021-128963NB-I00.

References

Amengual, D., Bei, X. and Sentana, E. (2022): "Normal but skewed?, *Journal of Applied Econometrics* 37, 1295–1313.

Amengual, D., Bei, X. and Sentana, E. (2023): "Hypothesis tests with a repeatedly singular information matrix", mimeo, CEMFI.

Amengual, D. and Sentana, E. (2015): "Is a normal copula the right copula?", CEMFI Working Paper 1504.

Amengual, D., Sentana, E. and Tian, Z. (2022): "Gaussian rank correlation and regression", chapter 11 in A. Chudik, C. Hsiao and A. Timmermann (eds.) *Essays in honor of M. Hashem Pesaran: Panel modeling, micro applications and econometric methodology*, *Advances in Econometrics* 43B, 269–306, Emerald.

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

Box, G. and Jenkins, G. (1970): *Time series analysis: forecasting and control.* Holden-Day.

Constantine, G.M. and Savits, T.H. (1996): "A multivariate Faa di Bruno formula with applications", *Transactions of the American Mathematical Society* 348, 503–520.

Cox, D. and Hinkley, D. (1974): *Theoretical statistics*, Chapman and Hall.

Davies, R.B. (1987): "Hypothesis testing when a nuisance parameter is present only under the alternatives", *Biometrika* 74, 33–43.

Dunsmuir, W. (1979): "A central limit theorem for parameter estimation in stationary vector time series and its application to models for a signal observed 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 Watson, G.S. (1950): "Testing for serial correlation in least squares regression, I", *Biometrika* 37, 409–428.

Durbin, J. and Watson, G.S. (1951): "Testing for serial correlation in least squares regression, II", *Biometrika* 38, 159–179.

Fiorentini, G. and Paruolo, P. (2009): "Testing residual autocorrelation in autoregressive processes with a zero root", mimeo, Università di Firenze.

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

Fiorentini, G. and Sentana, E. (2016): "Neglected serial correlation tests in Ucarima models", *SERIEs* 7, 121–178.

Fuller, W.A. (1995): *Introduction to time series analysis*, 2nd ed., Wiley.

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.

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

Horowitz, J. and Savin, N.E. (2000): "Empirically relevant critical values for hypothesis tests: a bootstrap approach", *Journal of Econometrics* 95, 375–389.

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

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

Lippi, M. and Reichlin, L. (1992): "On persistence of shocks to economic variables: a common misconception", *Journal of Monetary Economics* 29, 87–93.

Rao, C.R. (1948): "Large sample tests of statistical hypotheses concerning several parameters with applications to problems of estimation", *Mathematical Proceedings of the Cambridge Philosophical Society* 44, 50–57.

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

Silvey, S. D. (1959): "The Lagrangian multiplier test", *Annals of Mathematical Statistics* 30, 389–407.

Wallis, K.F. (1972): "Testing for fourth order autocorrelation in quarterly regression equations", *Econometrica* 40, 617–636.

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

Appendices

A Proofs

In this appendix, we thoroughly check that the multiplicative seasonal Arima and Ucarima models that we have considered satisfy the substantive assumptions required for the application of Theorems 1 and 2 in Amengual, Bei and Sentana (2023), respectively. For the sake of brevity, though, we do not include a detailed verification of the regularity conditions in their Assumptions 1 and 2.

Proof of Proposition 1

The scores evaluated under the null will be

$$\frac{\partial l_t}{\partial \varphi_M} = \frac{y_t - \varphi_M}{\varphi_V}, \; \frac{\partial l_t}{\partial \varphi_V} = \frac{(y_t - \varphi_M)^2 - \varphi_V}{2\varphi_V},$$

$$\frac{\partial l_t}{\partial \vartheta_1} = \frac{\partial l_t}{\partial \vartheta_2} = \frac{(y_t - \varphi_M)(y_{t-1} - \varphi_M)}{\varphi_V} \text{ and } \frac{\partial l_t}{\partial \vartheta_3} = \frac{\partial l_t}{\partial \vartheta_4} = \frac{(y_t - \varphi_M)(y_{t-4} - \varphi_M)}{\varphi_V}.$$

which immediately imply (3), thereby confirming that the nullity of the information matrix is two.

To isolate those singularities, consider the reparametrization from the original set of parameters $\boldsymbol{\varrho} = (\varphi_M, \varphi_V, \vartheta_1, \vartheta_2, \vartheta_3, \vartheta_4)'$ to a different set $\boldsymbol{\rho} = (\phi_M, \phi_V, \theta_{i1}, \theta_{i2}, \theta_{u1}, \theta_{u2})'$ defined by $\varphi_M = \phi_M$, $\varphi_V = \phi_V$, $\vartheta_1 = \theta_{i1} - \theta_{u1}$, $\vartheta_2 = \theta_{u1}$, $\vartheta_3 = \theta_{i2} - \theta_{u2}$ and $\vartheta_4 = \theta_{u2}$.

The corresponding first-order derivatives under the equivalent null hypothesis

$$H_0 : \theta_{i1} = \theta_{u1} = \theta_{i2} = \theta_{u2} = 0$$

are

$$\frac{\partial l_t}{\partial \theta_{i1}} = \frac{(y_t - \phi_M)(y_{t-1} - \phi_M)}{\phi_V}, \; \frac{\partial l_t}{\partial \theta_{i2}} = \frac{(y_t - \phi_M)(y_{t-4} - \phi_M)}{\phi_V}, \; \frac{\partial l_t}{\partial \theta_{u1}} = 0 \text{ and } \frac{\partial l_t}{\partial \theta_{u2}} = 0,$$

which verifies Assumption 3.1 in Amengual, Bei and Sentana (2023).

In turn, the second-order derivatives involving θ_{u1} and θ_{u2} are given in (4). Consequently,

$$\boldsymbol{\theta}_u^{\otimes 2\prime} \frac{\partial^2 l_t}{\partial \boldsymbol{\theta}_u^{\otimes 2}} = \theta_1^2 \frac{\partial^2 l_t}{\partial \theta_1^2} + \theta_2^2 \frac{\partial^2 l_t}{\partial \theta_2^2},$$

where $\boldsymbol{\theta}_u^{\otimes 2} = \left(\theta_1^2, \theta_1\theta_2, \theta_1\theta_2, \theta_2^2\right)'$ and

$$\frac{\partial^2 l_t}{\partial \boldsymbol{\theta}_u^{\otimes 2\prime}} = \left(\frac{\partial^2 l_t}{\partial \theta_1^2}, \frac{\partial^2 l_t}{\partial \theta_1 \partial \theta_2}, \frac{\partial^2 l_t}{\partial \theta_1 \partial \theta_2}, \frac{\partial^2 l_t}{\partial \theta_2^2}\right), \frac{\partial^2 l_t}{(\partial \theta_{u1})^2} = \frac{2(y_t - \phi_M)(y_{t-2} - \phi_M)}{\phi_V},$$

$$\frac{\partial^2 l_t}{\partial \theta_{u1} \partial \theta_{u2}} = 0 \text{ and } \frac{\partial^2 l_t}{(\partial \theta_{u2})^2} = \frac{2(y_t - \phi_M)(y_{t-8} - \phi_M)}{\phi_V}.$$

It is then easy to see that the asymptotic covariance matrix of

$$\left(\frac{\partial l_t}{\partial \phi_M}, \frac{\partial l_t}{\partial \phi_V}, \frac{\partial l_t}{\partial \theta_{i1}}, \frac{\partial l_t}{\partial \theta_{i2}}, \boldsymbol{\theta}_u^{\otimes 2\prime}\frac{\partial^2 l_t}{\partial \boldsymbol{\theta}_u^{\otimes 2}}\right)$$

under the null, namely $dg\{\varphi_V^{-1}, 1/2, 1, 1, 4\theta_1^2 + 4\theta_2^2\}$, has full rank, which verifies Assumption 3.2 in Amengual, Bei and Sentana (2023).

As a result, their Theorem 1 immediately implies that

$$LR_T = 2\left[L_T(\hat{\boldsymbol{\rho}}) - L_T(\tilde{\boldsymbol{\rho}})\right] = GET_T + O_p(T^{-\frac{1}{4}}),$$

where

$$\begin{aligned} GET_T &= \frac{1}{T}S'_{\boldsymbol{\theta}_1 T}(\tilde{\boldsymbol{\phi}}, \mathbf{0})V^{-1}_{\boldsymbol{\theta}_1\boldsymbol{\theta}_1}(\tilde{\boldsymbol{\phi}})S_{\boldsymbol{\theta}_1 T}(\tilde{\boldsymbol{\phi}}, \mathbf{0}) + \frac{1}{T}\sup_{\boldsymbol{\theta}_r \neq 0} Q_T(\boldsymbol{\theta}_r, \tilde{\boldsymbol{\phi}})\mathbf{1}\left[\boldsymbol{\theta}_r^{\otimes r\prime}D_{rT}(\tilde{\boldsymbol{\phi}}) \geq 0\right], \\ S'_{\boldsymbol{\theta}_1 t}(\tilde{\boldsymbol{\phi}}, \mathbf{0}) &= \left[\sum_{t=1}^{T}\frac{\partial l_t}{\partial \theta_{i1}}(\tilde{\phi}_M, 0), \sum_{t=1}^{T}\frac{\partial l_t}{\partial \theta_{i1}}(\tilde{\phi}_M, 0)\right] = T(\hat{r}_1, \hat{r}_2),\ V^{-1}_{\boldsymbol{\theta}_1\boldsymbol{\theta}_1} = \mathbf{I}_2 \end{aligned}$$

and

$$Q_T(\boldsymbol{\theta}_r, \tilde{\boldsymbol{\phi}}) = \frac{1}{T}\frac{1}{4\theta_{u1}^2 + 4\theta_{u2}^2}\left(\sum_{t=1}^{T}\theta_{u1}^2\frac{\partial^2 l_t}{\partial \theta_{u1}^2} + \theta_{u2}^2\frac{\partial^2 l_t}{\partial \theta_{u2}^2}\right)^2 = \frac{T}{\theta_{u1}^4 + \theta_{u2}^4}\left(\theta_{u1}^2\hat{r}_2 + \theta_{u2}^2\hat{r}_8\right)^2. \quad \text{(A1)}$$

Simple algebra then yields (6) because the value of $(\theta_{u1}, \theta_{u2})$ that maximizes (A1) is proportional to the vector $(\sqrt{\hat{r}_2\mathbf{1}\left[\hat{r}_2 \geq 0\right]}, \sqrt{\hat{r}_8\mathbf{1}\left[\hat{r}_8 \geq 0\right]})$ if $\hat{r}_2 \geq 0$ or $\hat{r}_8 \geq 0$, and to $(1,1)$ otherwise. □

To briefly illustrate the main idea behind Theorem 1 in Amengual, Bei and Sentana (2023) in this case, let us write θ_{u1} and θ_{u2} in polar coordinates so that $\theta_{u1} = \eta\upsilon_1$ and $\theta_{u2} = \eta\upsilon_2$ with $\upsilon_1^2 + \upsilon_2^2 = 1$, and consider the simplified null hypothesis $H_0 : \eta = 0$ for a fixed value of υ_1 and υ_2 determined by the relevant polar angle. In this context, the only relevant quantity associated to η is

$$\frac{\partial^2 l_t}{\partial \eta^2} = 2\upsilon_1^2\frac{(y_t - \phi_M)(y_{t-2} - \phi_M)}{\phi_V} + 2\upsilon_2^2\frac{(y_t - \phi_M)(y_{t-8} - \phi_M)}{\phi_V}.$$

Moreover, given that

$$E\left(\frac{\partial l_t}{\partial \boldsymbol{\phi}}\frac{\partial l_t}{\partial \boldsymbol{\theta}_i'}\right) = \mathbf{0} \text{ and } E\left[\frac{\partial l_t}{\partial \boldsymbol{\phi}}vech'\left(\frac{\partial^2 l_t}{\partial \boldsymbol{\theta}_u \partial \boldsymbol{\theta}_u'}\right)\right] = \mathbf{0}$$

under the null, we can ignore the parameter uncertainty in estimating ϕ_M and ϕ_V, at least asymptotically.

Next, letting

$$S_{\boldsymbol{\theta}_i}(\boldsymbol{\rho}) = \left[S_{\theta_{i1}}(\boldsymbol{\rho}), S_{\theta_{i2}}(\boldsymbol{\rho})\right]', \ \mathcal{H}_\eta(\boldsymbol{\phi}, \eta, \boldsymbol{\upsilon}) = \sum_{t=1}^{T}\frac{\partial^2 l_t}{\partial \eta^2}$$

and

$$\mathcal{V}(\boldsymbol{\phi}, \boldsymbol{\upsilon}) = Var\{T^{-\frac{1}{2}}[S'_{\boldsymbol{\theta}_i}(\boldsymbol{\phi}, \mathbf{0}), \mathcal{H}_\eta(\boldsymbol{\phi}, 0, \boldsymbol{\upsilon})]'|\boldsymbol{\phi}, \mathbf{0}\} = \left[\begin{array}{cc} \mathcal{V}_S(\boldsymbol{\phi}) & \mathbf{0} \\ \mathbf{0} & \mathcal{V}_H(\boldsymbol{\phi}, \boldsymbol{\upsilon}) \end{array}\right], \quad \text{(A2)}$$

a similar argument to the one used in the proof of Theorem 1 in Amengual, Bei, Sentana (2023) implies that

$$
\begin{aligned}
LR_T &= \sup_{(\boldsymbol{\theta}_i,\boldsymbol{\upsilon})} 2\begin{bmatrix} S_{\boldsymbol{\theta}_i}(\tilde{\boldsymbol{\phi}},0) \\ \mathcal{H}_\eta(\tilde{\boldsymbol{\phi}},0,\boldsymbol{\upsilon}) \end{bmatrix}\begin{pmatrix} \boldsymbol{\theta}_i \\ \eta^2 \end{pmatrix} - T\begin{pmatrix} \boldsymbol{\theta}_i \\ \eta^2 \end{pmatrix}'\mathcal{V}^{-1}(\tilde{\boldsymbol{\phi}},\boldsymbol{\upsilon})\begin{pmatrix} \boldsymbol{\theta}_i \\ \eta^2 \end{pmatrix} + O_p(T^{-\frac{1}{4}}) \\
&= \sup_{(\boldsymbol{\theta}_i,\boldsymbol{\upsilon})} 2S_{\boldsymbol{\theta}_i}(\tilde{\boldsymbol{\phi}},0)\boldsymbol{\theta}_i - T\theta_i'\mathcal{V}_S^{-1}(\tilde{\boldsymbol{\phi}})\theta_i + 2\mathcal{H}_\eta(\tilde{\boldsymbol{\phi}},0,\boldsymbol{\upsilon})\eta^2 - T\mathcal{V}_H^{-1}(\tilde{\boldsymbol{\phi}},\boldsymbol{\upsilon})\eta^4 + O_p(T^{-\frac{1}{4}}) \\
&= T^{-1}S_{\boldsymbol{\theta}_i}(\tilde{\boldsymbol{\phi}},0)\mathcal{V}_S^{-1}(\tilde{\boldsymbol{\phi}})S_{\boldsymbol{\theta}_i}(\tilde{\boldsymbol{\phi}},0) + \sup_{\boldsymbol{\upsilon}} T^{-1}\frac{\mathcal{H}_\eta(\tilde{\boldsymbol{\phi}},0,\boldsymbol{\upsilon})^2}{\mathcal{V}^{-1}(\tilde{\boldsymbol{\phi}},\boldsymbol{\upsilon})}\mathbf{1}\left[\mathcal{H}_\eta(\tilde{\boldsymbol{\phi}},0,\boldsymbol{\upsilon}) \geq 0\right] + O_p(T^{-\frac{1}{4}}),
\end{aligned} \tag{A3}
$$

where the first equality comes from a fourth-order Taylor expansion of the log-likelihood function whose first two terms are the leading ones and the rest is included in the $O_p(T^{-\frac{1}{4}})$ remainder, the second equality follows from (A2), and the last one is trivial.

Proof of Proposition 2

We can use expression (B11) in Appendix B.1 to compute the spectral approximation to the log-likelihood function of model (12) with $g_{yy}(\omega;\boldsymbol{\rho})$ given in (B15), $\boldsymbol{\rho} = (\boldsymbol{\phi}', \boldsymbol{\theta}')'$, $\boldsymbol{\phi} = (\sigma_f^2, \sigma_v^2)'$ and $\boldsymbol{\theta} = (\alpha, \psi_1, \psi_2)'$.

To simplify the notation, let us define the vector $\mathbf{C}(\omega) = [\mathbf{C}'_{\boldsymbol{\phi}}(\omega), \mathbf{C}'_{\boldsymbol{\theta}}(\omega)]'$ with

$$
\mathbf{C}_{\boldsymbol{\phi}}(\omega) = \frac{2\pi I_{yy}(\omega) - g_{yy}(\omega;\boldsymbol{\gamma})}{g^2_{yy}(\omega;\boldsymbol{\gamma})}\begin{bmatrix} 1 \\ \cos(\omega) \end{bmatrix} \text{ and } \mathbf{C}_{\boldsymbol{\theta}}(\omega) = \frac{2\pi I_{yy}(\omega) - g_{yy}(\omega;\boldsymbol{\gamma})}{g^2_{yy}(\omega;\boldsymbol{\gamma})}\begin{bmatrix} \cos(2\omega) \\ \cos(3\omega) \\ \cos(4\omega) \end{bmatrix},
$$

which correspond to the contribution of frequency ω to the spectral score of an MA(4) model parametrized in terms of its unconditional variance and first four autocovariances, say $\boldsymbol{\gamma} = (\gamma_0, \gamma_1, \gamma_2, \gamma_3, \gamma_4)$, evaluated at $\gamma_2 = \gamma_3 = \gamma_4 = 0$, as can be immediately seen from (B12). Importantly, $g_{yy}(\omega;\boldsymbol{\rho}) = g_{yy}(\omega;\boldsymbol{\gamma})$ for all ω under the locally equivalent null hypotheses

$$
H_0 : \alpha = \psi_1 = \psi_2 = 0 \text{ and } H_0 : \gamma_2 = \gamma_3 = \gamma_4 = 0
$$

when both σ_f^2 and σ_v^2 are strictly positive.

Therefore, we can write the contribution of frequency ω to the spectral score as

$$
\frac{\partial l}{\partial \sigma_f^2} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \end{pmatrix}\mathbf{C}(\omega), \quad \frac{\partial l}{\partial \sigma_v^2} = \begin{pmatrix} 2 & -2 & 0 & 0 & 0 \end{pmatrix}\mathbf{C}(\omega),
$$

$$
\frac{\partial l}{\partial \alpha} = \begin{pmatrix} -2\sigma_f^2 & 4\sigma_f^2 & -2\sigma_f^2 & 0 & 0 \end{pmatrix}\mathbf{C}(\omega), \quad \frac{\partial l}{\partial \psi_1} = \begin{pmatrix} 0 & 2\sigma_f^2 & 0 & 0 & 0 \end{pmatrix}\mathbf{C}(\omega)
$$

and

$$
\frac{\partial l}{\partial \psi_2} = \begin{pmatrix} 0 & 0 & 2\sigma_f^2 & 0 & 0 \end{pmatrix}\mathbf{C}(\omega).
$$

We can immediately notice that the last two elements of this score belong to the linear span of the first three, which confirms that the nullity of the information matrix is again two.

To isolate those singularities, we conduct a two-step reparametrization as follows. First, we consider

$$\sigma_f^2 = \sigma_f^{\diamond 2} - 2\sigma_f^{\diamond 2}\psi_1^{\diamond} + \sigma_f^{\diamond 2}\psi_1^{\diamond 2} - 2\sigma_f^{\diamond 2}\psi_2^{\diamond},\ \sigma_v^2 = \sigma_v^{\diamond 2} + \sigma_f^{\diamond 2}\psi_1^{\diamond} + 2\sigma_f^{\diamond 2}\psi_2^{\diamond},$$

$$\alpha = \alpha^{\diamond} + \frac{\sigma_f^{\diamond 2}}{\sigma_v^{\diamond 2}}\psi_1^{\diamond 2} + \frac{\sigma_f^{\diamond 2}}{\sigma_v^{\diamond 2}}\psi_2^{\diamond},\ \psi_1 = \psi_1^{\diamond} \text{ and } \psi_2 = \psi_2^{\diamond},$$

and then

$$\sigma_f^{\diamond 2} = \sigma_f^{\dagger 2} - \sigma_v^{\dagger 2}\psi_1^{\dagger 3},\ \sigma_v^{\diamond 2} = \sigma_f^{\dagger 2} + \frac{1}{2}\sigma_v^{\dagger 2}\psi_1^{\dagger 3},\ \alpha^{\diamond} = \alpha^{\dagger} - \frac{\sigma_f^{\dagger 4} + 2\sigma_v^{\dagger 2}\sigma_f^{\dagger 2}}{2\sigma_v^{\dagger 4}}\psi_1^{\dagger 3},\ \psi_1^{\diamond} = \psi_1^{\dagger} \text{ and } \psi_2^{\diamond} = \psi_2^{\dagger} - \frac{\psi_1^{\dagger 2}}{2}.$$

After this sequential reparametrization, the relevant derivatives evaluated under the null become

$$\begin{aligned}
\frac{\partial l}{\partial \sigma_f^{\dagger 2}} &= \frac{\partial l}{\partial \sigma_f^2} = \begin{pmatrix} 1 & 0 \end{pmatrix}\mathbf{C}_{\boldsymbol{\phi}}(\omega) \equiv a_1'\mathbf{C}_{\boldsymbol{\phi}}(\omega),\ \frac{\partial l}{\partial \sigma_v^{\dagger 2}} = \frac{\partial l}{\partial \sigma_v^2} = \begin{pmatrix} 2 & -2 \end{pmatrix}\mathbf{C}_{\boldsymbol{\phi}}(\omega) \equiv a_2'\mathbf{C}_{\boldsymbol{\phi}}(\omega),\\
\frac{\partial l}{\partial \alpha^{\dagger}} &= \frac{\partial l}{\partial \alpha} = \begin{pmatrix} -2\sigma_f^2 & 4\sigma_f^2 \end{pmatrix}\mathbf{C}_{\boldsymbol{\phi}}(\omega) + \begin{pmatrix} -2\sigma_f^2 & 0 & 0 \end{pmatrix}\mathbf{C}_{\boldsymbol{\theta}}(\omega) \equiv d_1'\mathbf{C}_{\boldsymbol{\phi}}(\omega) + b_1'\mathbf{C}_{\boldsymbol{\theta}}(\omega)
\end{aligned}$$

and

$$\frac{\partial l}{\partial \psi_1^{\dagger}} = \frac{\partial l}{\partial \psi_2'} = \frac{\partial^2 l}{(\partial \psi_1')^2} = \frac{\partial^3 l}{(\partial \psi_1')^3} = 0.$$

In addition, straightforward calculations deliver

$$\begin{aligned}
\frac{\partial^2 l}{(\partial \psi_2^{\dagger})^2} &= \begin{bmatrix} 2\sigma_f^2(\sigma_v^2 - 2\sigma_f^2)/\sigma_v^2 \\ 8\sigma_f^4/\sigma_v^2 \end{bmatrix}'\mathbf{C}_{\boldsymbol{\phi}}(\omega) + \begin{bmatrix} -8\sigma_f^2 \\ -4\sigma_f^4/\sigma_v^2 \\ 4\sigma_f^2 \end{bmatrix}'\mathbf{C}_{\boldsymbol{\theta}}(\omega)\\
&\equiv d_2'\mathbf{C}_{\boldsymbol{\phi}}(\omega) + b_2'\mathbf{C}_{\boldsymbol{\theta}}(\omega),
\end{aligned}$$

$$\begin{aligned}
\frac{\partial^2 l}{\partial \psi_1^{\dagger}\partial \psi_2^{\dagger}} &= \begin{bmatrix} -2\sigma_f^4/\sigma_v^2 \\ -2\sigma_f^2(\sigma_v^2 - 2\sigma_f^2)/\sigma_v^2 \end{bmatrix}'\mathbf{C}_{\boldsymbol{\phi}}(\omega) + \begin{bmatrix} -2\sigma_f^2(\sigma_f^2 + 2\sigma_v^2)/\sigma_v^2 \\ 4\sigma_f^2 \\ 0 \end{bmatrix}'\mathbf{C}_{\boldsymbol{\theta}}(\omega)\\
&\equiv d_3'\mathbf{C}_{\boldsymbol{\phi}}(\omega) + b_3'\mathbf{C}_{\boldsymbol{\theta}}(\omega)
\end{aligned}$$

and

$$\begin{aligned}
\frac{\partial^4 l}{(\partial \psi_1^{\dagger})^4} &= \begin{bmatrix} 6\sigma_f^2(4\sigma_f^4 + 14\sigma_f^2\sigma_v^2 + 9\sigma_v^4)/\sigma_v^4 \\ -24\sigma_f^2(2\sigma_f^4 + 7\sigma_f^2\sigma_v^2 + 2\sigma_v^4)/\sigma_v^4 \end{bmatrix}'\mathbf{C}_{\boldsymbol{\phi}}(\omega) + \begin{bmatrix} 24\sigma_f^2(\sigma_f^4 + 4\sigma_f^2\sigma_v^2 + 2\sigma_v^4)/\sigma_v^4 \\ -12\sigma_f^4/\sigma_v^4 \\ -12\sigma_f^2 \end{bmatrix}'\mathbf{C}_{\boldsymbol{\theta}}(\omega)\\
&\equiv d_4'\mathbf{C}_{\boldsymbol{\phi}}(\omega) + b_4'\mathbf{C}_{\boldsymbol{\theta}}(\omega).
\end{aligned}$$

Next, we carry out an eighth-order Taylor expansion of the reparametrized spectral log-likelihood function of the sample, $L_T(\boldsymbol{\rho}^{\dagger})$, around the true values of its parameters under the null, namely $\boldsymbol{\rho}_0^{\dagger\prime} = (\boldsymbol{\phi}_0^{\dagger\prime}, \mathbf{0}')$, which yields

$$L_T(\boldsymbol{\rho}^\dagger) - L_T(\boldsymbol{\rho}_0^\dagger) = \begin{bmatrix} \frac{\partial L}{\partial \boldsymbol{\phi}^\dagger} \\ \frac{\partial L}{\partial \alpha^\dagger} \\ \frac{1}{4!}\frac{\partial^4 L}{(\partial \psi_1^\dagger)^4} \\ \frac{\partial^2 L}{\partial \psi_1^\dagger \partial \psi_2^\dagger} \\ \frac{1}{2}\frac{\partial^2 L}{(\partial \psi_2^\dagger)^2} \end{bmatrix}' \begin{pmatrix} \boldsymbol{\phi}^\dagger - \boldsymbol{\phi}_0^\dagger \\ \alpha^\dagger \\ \psi_1^{\dagger 4} \\ \psi_1^\dagger \psi_2^\dagger \\ \psi_2^{\dagger 2} \end{pmatrix} - \frac{1}{2}T \begin{pmatrix} \boldsymbol{\phi}^\dagger - \boldsymbol{\phi}_0^\dagger \\ \alpha^\dagger \\ \psi_1^{\dagger 4} \\ \psi_1^\dagger \psi_2^\dagger \\ \psi_2^{\dagger 2} \end{pmatrix}' \mathcal{V} \begin{pmatrix} \boldsymbol{\phi}^\dagger - \boldsymbol{\phi}_0^\dagger \\ \alpha^\dagger \\ \psi_1^{\dagger 4} \\ \psi_1^\dagger \psi_2^\dagger \\ \psi_2^{\dagger 2} \end{pmatrix} \tag{A4}$$

$$+ \frac{1}{2}T \begin{pmatrix} \boldsymbol{\phi}^{\dagger\prime} - \boldsymbol{\phi}_0^{\dagger\prime} & \alpha^\dagger & \psi_1^{\dagger 4} & \psi_1^\dagger \psi_2^\dagger & \psi_2^{\dagger 2} \end{pmatrix}' (\mathcal{H} + \mathcal{V}) \begin{pmatrix} \boldsymbol{\phi}^\dagger - \boldsymbol{\phi}_0^\dagger \\ \alpha^\dagger \\ \psi_1^{\dagger 4} \\ \psi_1^\dagger \psi_2^\dagger \\ \psi_2^{\dagger 2} \end{pmatrix} \tag{A5}$$

$$+ \sum_{\boldsymbol{j} \in \mathcal{J}_1} \left(\frac{\prod j_i!}{\boldsymbol{j}!} \frac{1}{\sqrt{T}} \frac{\partial^{\iota' \boldsymbol{j}} L}{\partial \boldsymbol{\rho}^{\dagger \boldsymbol{j}}} \right) \left\{ \sqrt{T} (\boldsymbol{\rho}^\dagger - \boldsymbol{\rho}_0^\dagger)^{\boldsymbol{j}} \right\} \tag{A6}$$

$$+ \sum_{\boldsymbol{j} \in \mathcal{J}_2} \left(\frac{\prod j_i!}{\boldsymbol{j}!} \frac{1}{T} \frac{\partial^8 L}{\partial \boldsymbol{\rho}^{\dagger \boldsymbol{j}}} \right) \left\{ T (\boldsymbol{\rho}^\dagger - \boldsymbol{\rho}_0^\dagger)^{\boldsymbol{j}} \right\} \tag{A7}$$

$$+ \sum_{\iota' \boldsymbol{j} = 8} \left\{ \frac{\prod j_i!}{\boldsymbol{j}!} \frac{1}{T} \left[\frac{\partial^8 L}{\partial \boldsymbol{\rho}^{\dagger \boldsymbol{j}}} (\bar{\boldsymbol{\rho}}^\dagger) - \frac{\partial^8 L}{\partial \boldsymbol{\rho}^{\dagger \boldsymbol{j}}} (\boldsymbol{\rho}^\dagger) \right] \right\} \left\{ T (\boldsymbol{\rho}^\dagger - \boldsymbol{\rho}_0^\dagger)^{\boldsymbol{j}} \right\}, \tag{A8}$$

where

$$\mathcal{H} = \begin{bmatrix} \frac{\partial^2 L}{\partial \boldsymbol{\phi}^\dagger \partial \boldsymbol{\phi}^{\dagger\prime}} & \frac{\partial^2 L}{\partial \boldsymbol{\phi}^\dagger \partial \alpha^\dagger} & \frac{1}{4!}\frac{\partial^5 L}{\partial \boldsymbol{\phi}^\dagger (\partial \psi_1^\dagger)^4} & \frac{\partial^3 L}{\partial \boldsymbol{\phi}^\dagger \partial \psi_1^\dagger \partial \psi_2^\dagger} & \frac{1}{2}\frac{\partial^3 L}{\partial \boldsymbol{\phi}^\dagger (\partial \psi_2^\dagger)^2} \\ \frac{\partial^2 L}{\partial \alpha^\dagger \partial \boldsymbol{\phi}^{\dagger\prime}} & \frac{\partial^2 L}{\partial (\alpha^\dagger)^2} & \frac{1}{4!}\frac{\partial^5 L}{\partial \alpha^\dagger (\partial \psi_1^\dagger)^4} & \frac{\partial^3 L}{\partial \alpha^\dagger \partial \psi_1^\dagger \partial \psi_2^\dagger} & \frac{1}{2}\frac{\partial^3 L}{\partial \alpha^\dagger (\partial \psi_2^\dagger)^2} \\ \frac{1}{4!}\frac{\partial^5 L}{(\partial \psi_1^\dagger)^4 \partial \boldsymbol{\phi}^{\dagger\prime}} & \frac{1}{4!}\frac{\partial^5 L}{(\partial \psi_1^\dagger)^4 \partial \alpha^\dagger} & \frac{2}{8!}\frac{\partial^8 L}{(\partial \psi_1^\dagger)^8} & \frac{1}{4!}\frac{\partial^6 L}{(\partial \psi_1^\dagger)^4 \partial \psi_1^\dagger \partial \psi_2^\dagger} & \frac{1}{2}\frac{1}{4!}\frac{\partial^6 L}{(\partial \psi_1^\dagger)^4 (\partial \psi_2^\dagger)^2} \\ \frac{\partial^3 L}{\partial \psi_1^\dagger \partial \psi_2^\dagger \partial \boldsymbol{\phi}^{\dagger\prime}} & \frac{\partial^3 L}{\partial \psi_1^\dagger \partial \psi_2^\dagger \partial \alpha^\dagger} & \frac{1}{4!}\frac{\partial^6 L}{\partial \psi_1^\dagger \partial \psi_2^\dagger (\partial \psi_1^\dagger)^4} & \frac{1}{2}\frac{\partial^4 L}{\partial (\psi_1^\dagger)^2 \partial (\psi_2^\dagger)^2} & \frac{1}{3!}\frac{\partial^4 L}{\partial \psi_1^\dagger (\partial \psi_2^\dagger)^3} \\ \frac{1}{2}\frac{\partial^3 L}{(\partial \psi_2^\dagger)^2 \partial \boldsymbol{\phi}^{\dagger\prime}} & \frac{1}{2}\frac{\partial^3 L}{(\partial \psi_2^\dagger)^2 \partial \alpha^\dagger} & \frac{1}{2}\frac{1}{4!}\frac{\partial^6 L}{(\partial \psi_2^\dagger)^2 (\partial \psi_1^\dagger)^4} & \frac{1}{3!}\frac{\partial^4 L}{\partial \psi_1^\dagger (\partial \psi_2^\dagger)^3} & \frac{2}{4!}\frac{\partial^4 L}{(\partial \psi_2^\dagger)^4} \end{bmatrix},$$

$\mathcal{V} = E(\mathcal{H})$, $\boldsymbol{j}$ is a five-dimensional vector of indices, for example (1,0,2,0,3), in which case

$$\frac{\partial^{\iota' \boldsymbol{j}} L}{\partial \boldsymbol{\rho}^{\dagger \boldsymbol{j}}} = \frac{\partial^{\iota' \boldsymbol{j}} L}{\partial \sigma_f^{\dagger 4} \partial \alpha^\dagger \partial \psi_2^\dagger}, \; (\boldsymbol{\rho}^\dagger - \boldsymbol{\rho}_0^\dagger)^{\boldsymbol{j}} = (\sigma_f^{\dagger 2} - \sigma_{f0}^{\dagger 2}) \alpha^{\dagger 2} \psi_2^{\dagger 3},$$

$\mathcal{J}_0 = \{(2,2,2,8,0), (2,2,2,2,2), (2,2,2,0,4)\}$, $\mathcal{J}_1 = \left\{ \boldsymbol{j} : \boldsymbol{\iota}' \boldsymbol{j} \leq 8, \; \exists \boldsymbol{j}' \in \mathcal{J}_0 \text{ such that } \boldsymbol{j} < \boldsymbol{j}' \right\}$ and $\mathcal{J}_2 = \{\boldsymbol{j} : \boldsymbol{\iota}' \boldsymbol{j} \leq 8, \; \boldsymbol{j} \notin \mathcal{J}_1, \boldsymbol{j} \notin \mathcal{J}_0\}$, with $\boldsymbol{\iota}' = (1,1,1,1,1)$, are finite sets consisting of those indices.

On this basis, we will have that

$$\begin{pmatrix} \frac{\partial L}{\partial \boldsymbol{\phi}^{\dagger\prime}} & \frac{\partial L}{\partial \alpha^\dagger} & \frac{1}{4!}\frac{\partial^4 L}{(\partial \psi_1^\dagger)^4} & \frac{\partial^2 L}{\partial \psi_1^\dagger \partial \psi_2^\dagger} & \frac{1}{2}\frac{\partial^2 L}{(\partial \psi_2^\dagger)^2} \end{pmatrix}' \begin{pmatrix} \boldsymbol{\phi}^\dagger - \boldsymbol{\phi}_0^\dagger \\ \alpha^\dagger \\ \psi_1^{\dagger 4} \\ \psi_1^\dagger \psi_2^\dagger \\ \psi_2^{\dagger 2} \end{pmatrix} = \left\{ \begin{pmatrix} \mathbf{D} & \mathbf{0} \\ \mathbf{A} & \mathbf{B} \end{pmatrix} \begin{bmatrix} \mathbf{C}_{\boldsymbol{\phi}}(\omega) \\ \mathbf{C}_{\boldsymbol{\theta}}(\omega) \end{bmatrix} \right\}' \boldsymbol{\nu}$$

$$= \begin{bmatrix} \mathbf{D}\mathbf{C}_{\boldsymbol{\phi}}(\omega) \\ \mathbf{C}_{\boldsymbol{\theta}}(\omega) \end{bmatrix}' \begin{bmatrix} \mathbf{I}_2 & 0 \\ \mathbf{A}\mathbf{D}^{-1} & \mathbf{B} \end{bmatrix}' \boldsymbol{\nu} = \mathbf{S}(\boldsymbol{\phi}^*)' \boldsymbol{\lambda}(\boldsymbol{\phi}^\dagger, \boldsymbol{\theta}^\dagger),$$

where $\mathbf{I}_2$ is the identity matrix of order two,

$$\mathbf{D}=\begin{pmatrix} a_1' \\ a_2' \end{pmatrix},\ \mathbf{A}=\begin{pmatrix} a_3' \\ a_4' \\ a_5' \end{pmatrix},\ \mathbf{B}=\begin{pmatrix} b_1' \\ b_2' \\ b_3' \end{pmatrix},\ \boldsymbol{\nu}=\begin{pmatrix} \alpha^\dagger \\ \psi_1^{\dagger 4} \\ \psi_1^\dagger\psi_2^\dagger \\ \psi_2^{\dagger 2} \end{pmatrix},\ \mathbf{S}(\boldsymbol{\phi}^*)=\begin{bmatrix} \mathbf{DC}_{\boldsymbol{\phi}}(\omega) \\ \mathbf{C}_{\boldsymbol{\theta}}(\omega) \end{bmatrix}$$

$$\boldsymbol{\lambda}(\boldsymbol{\phi}^\dagger,\boldsymbol{\theta}^\dagger)=\begin{pmatrix} \mathbf{I}_2 & \mathbf{0} \\ \mathbf{AD}^{-1} & \mathbf{B} \end{pmatrix}'\boldsymbol{\nu}=\begin{bmatrix} (\boldsymbol{\phi}^\dagger-\boldsymbol{\phi}_0^\dagger)+\boldsymbol{\lambda}_{\boldsymbol{\phi}}(\boldsymbol{\theta}^\dagger) \\ \boldsymbol{\lambda}_{\boldsymbol{\theta}}(\boldsymbol{\theta}^\dagger) \end{bmatrix},\ \boldsymbol{\lambda}_{\boldsymbol{\phi}}(\boldsymbol{\theta}^\dagger)=\left(\mathbf{AD}^{-1}\right)'\begin{pmatrix} \alpha^\dagger \\ \psi_1^{\dagger 4} \\ \psi_1^\dagger\psi_2^\dagger \\ \psi_2^{\dagger 2} \end{pmatrix}$$

and

$$\boldsymbol{\lambda}_{\boldsymbol{\theta}}(\boldsymbol{\theta}^\dagger)=\mathbf{B}'\begin{pmatrix} \alpha^\dagger \\ \psi_1^{\dagger 4} \\ \psi_1^\dagger\psi_2^\dagger \\ \psi_2^{\dagger 2} \end{pmatrix}=\begin{pmatrix} -2\sigma_f^2\alpha^\dagger-4\sigma_f^2\psi_2^{\dagger 2}-\frac{2\sigma_f^2(\sigma_f^2+2\sigma_v^2)}{\sigma_v^2}\psi_1^\dagger\psi_2^\dagger+\frac{\sigma_f^2(\sigma_f^4+4\sigma_f^2\sigma_v^2+2\sigma_v^4)}{\sigma_v^4}\psi_1^{\dagger 4} \\ -\frac{2\sigma_f^4}{\sigma_v^2}\psi_2^{\dagger 2}+4\sigma_f^2\psi_1^\dagger\psi_2^\dagger-\frac{\sigma_f^4}{2\sigma_v^4}\psi_1^{\dagger 4} \\ 2\sigma_f^2\psi_2^{\dagger 2}-12\sigma_f^2\psi_1^{\dagger 4} \end{pmatrix}.$$

In addition, the interpretation of $\mathbf{C}(\omega)$ as a spectral log-likelihood score allows us to use expression (B13) to obtain the asymptotic variance of a suitably scaled version of

$$\mathcal{S}_T=\sum_{j=0}^{T-1}\mathbf{C}(\omega_j),$$

where $\omega_j=2\pi j/T$ (for $j=0,\ldots T-1$) are the usual Fourier frequencies. Specifically, if we partition the autocovariances into $\boldsymbol{\gamma}_n=(\gamma_0,\gamma_1)$ and $\boldsymbol{\gamma}_a=(\gamma_2,\gamma_3,\gamma_4)$, then we will have that

$$T^{-\frac{1}{2}}\mathcal{S}_T\xrightarrow{d}N\left[\begin{pmatrix} \mathbf{0} \\ \mathbf{0} \end{pmatrix},\begin{pmatrix} \mathcal{I}_{\gamma_n\gamma_n} & \mathcal{I}_{\gamma_n\gamma_a} \\ \mathcal{I}_{\gamma_n\gamma_a}' & \mathcal{I}_{\gamma_a\gamma_a} \end{pmatrix}\right]\equiv N(\mathbf{0},\mathcal{V}), \tag{A9}$$

with the different elements evaluated at $\boldsymbol{\gamma}_a=\mathbf{0}$. Consequently,

$$\begin{pmatrix} \boldsymbol{\phi}^{\dagger\prime}-\boldsymbol{\phi}_0^{\dagger\prime} & \alpha^\dagger & \psi_1^{\dagger 4} & \psi_1^\dagger\psi_2^\dagger & \psi_2^{\dagger 2} \end{pmatrix}\mathcal{V}\begin{pmatrix} \boldsymbol{\phi}^\dagger-\boldsymbol{\phi}_0^\dagger \\ \alpha^\dagger \\ \psi_1^{\dagger 4} \\ \psi_1^\dagger\psi_2^\dagger \\ \psi_2^{\dagger 2} \end{pmatrix}=\boldsymbol{\lambda}(\boldsymbol{\phi}^\dagger,\boldsymbol{\theta}^\dagger)'\mathcal{I}(\boldsymbol{\phi}^*)\boldsymbol{\lambda}(\boldsymbol{\phi}^\dagger,\boldsymbol{\theta}^\dagger).$$

On this basis, we can write $L_n(\boldsymbol{\rho}^\dagger)-L_n(\boldsymbol{\rho}_0^\dagger)$ as the local quadratic approximation in Theorem 2 of Amengual, Bei and Sentana (2023), with a remainder $R_n(\boldsymbol{\rho}^\dagger)$ equal to the sum of (A5)-(A8).

Next, we verify in detail Assumption 4 of that theorem.

Assumption 4.1: It is easy to see that $\boldsymbol{\lambda}(\boldsymbol{\phi}^\dagger,\boldsymbol{\theta}^\dagger)$ is continuous and $\boldsymbol{\lambda}(\boldsymbol{\phi}_0^\dagger,\mathbf{0})=\mathbf{0}$. In addition, $\boldsymbol{\lambda}(\boldsymbol{\phi}^\dagger,\boldsymbol{\theta}^\dagger)=\mathbf{0}$ if and only if

$$\begin{pmatrix} 2\sigma_f^2\psi_2^{\dagger 2}-12\sigma_f^2\psi_1^{\dagger 4} \\ -\frac{2\sigma_f^4}{\sigma_v^2}\psi_2^{\dagger 2}+4\sigma_f^2\psi_1^\dagger\psi_2^\dagger-\frac{\sigma_f^4}{2\sigma_v^4}\psi_1^{\dagger 4} \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

which gives the unique solution $\psi_1^\dagger = 0$, $\psi_2^\dagger = 0$ if we consider a local neighborhood around $(0,0)$ such that $|\psi_1^\dagger| \le \bar{\varepsilon}$ for $\bar{\varepsilon} > 0$ small enough.

Assumption 4.2: A central limit theorem for $T^{-1/2}\mathcal{S}_T$ holds because the special case of the MA(4) process to which this spectral score corresponds is covariance stationary and contains no discrete harmonic components. Given that the generalized spectral score is a full-rank linear transformation of this scaled average, a central limit theorem will apply to it too.

Assumption 4.3: The limiting variance of $T^{-1/2}\mathcal{S}_T$ will have full rank for analogous reasons, as long as we exclude processes whose MA polynomial contains unit roots. The full-rank mapping of the spectral log-likelihood score to this scaled average guarantees that the generalized spectral score will also have a full-rank limiting covariance matrix under the same circumstances.

Assumption 4.4: We verify this for each term in (A5)-(A8). In particular, (A5) follows from $\mathcal{H} + \mathcal{V} = O_p(T^{-\frac{1}{2}})$. As for (A6), we can easily check that: (i) the term in $\{.\}$ is $o_p(1 + \sqrt{T}||\boldsymbol{\lambda}(\boldsymbol{\phi}^\dagger, \boldsymbol{\theta}^\dagger)||)$, while (ii) the term in $[.]$ is $O_p(1)$ because the derivative corresponding to $j \in \mathcal{J}_1$ has zero expectation, which we can in turn verify by means of the multivariate Faa di Bruno's formula (see Constantine and Savits (1996) for details) or by calculating the required expectation directly. In turn, we can easily verify for (A7) that (i) the term in $\{.\}$ is $o_p(1 + T||\boldsymbol{\lambda}(\boldsymbol{\phi}^\dagger, \boldsymbol{\theta}^\dagger)||^2)$, while (ii) the term in $[.]$ is $O_p(1)$. Finally, the term in $\{.\}$ in (A8) is $O_p(1 + T||\boldsymbol{\lambda}(\boldsymbol{\phi}^\dagger, \boldsymbol{\theta}^\dagger)||^2)$, while the one in $[.]$ is $o_p(1)$.

Assumption 4.5: We can immediately see that

$$\frac{\partial \mathbf{C}(\omega)}{\partial \boldsymbol{\phi}'} = \frac{-4\pi I_{yy}(\omega) + g_{yy}(\omega, y)}{g_{yy}^3(\omega, y)} \begin{bmatrix} 1 \\ \cos(\omega) \\ \cos(2\omega) \\ \cos(3\omega) \\ \cos(4\omega) \end{bmatrix} \left\{ \begin{array}{cc} 1, & 2[1 - \cos(\omega)] \end{array} \right\}.$$

Hence, we will have that

$$\left| \frac{\partial \mathbf{C}(\omega)}{\partial \phi_k} \right| \le 2 \left| \frac{-4\pi I_{yy}(\omega)}{g_{yy}^3(\omega, y)} \right| + 2 \left| \frac{1}{g_{yy}^2(\omega, y)} \right| \le 2 \left| \frac{-4\pi I_{yy}(\omega)}{\left(\sigma_f^2 + 4\sigma_v^2\right)^3} \right| + 2 \left| \frac{1}{\left(\sigma_f^2 + 4\sigma_v^2\right)^2} \right| \equiv g(y),$$

where $g^2(y)$ is integrable. Consequently, Theorem 2 in Amengual, Bei and Sentana (2023) implies that

$$LR_T = GET_T + O_p(T^{-\frac{1}{8}}),$$

where

$$
\begin{aligned}
GET_T =& \sup_{\boldsymbol{\theta}} \left\{ 2\left[\mathcal{S}_{\boldsymbol{\theta},T}(\tilde{\boldsymbol{\phi}}) - \mathcal{I}_{\boldsymbol{\theta}\phi}(\tilde{\boldsymbol{\phi}})\mathcal{I}^{-1}_{\boldsymbol{\phi}\boldsymbol{\phi}}(\tilde{\boldsymbol{\phi}})\mathcal{S}_{\boldsymbol{\phi},T}(\tilde{\boldsymbol{\phi}}) \right]' \boldsymbol{\lambda}_{\boldsymbol{\theta}}(\boldsymbol{\theta}^{\dagger}) \right. \\
& \left. -n\boldsymbol{\lambda}'_{\boldsymbol{\theta}}(\boldsymbol{\theta}^{\dagger}) \left[\mathcal{I}_{\boldsymbol{\theta}\boldsymbol{\theta}}(\tilde{\boldsymbol{\phi}}) - \mathcal{I}_{\boldsymbol{\theta}\phi}(\tilde{\boldsymbol{\phi}})\mathcal{I}^{-1}_{\boldsymbol{\phi}\boldsymbol{\phi}}(\tilde{\boldsymbol{\phi}})\mathcal{I}_{\phi\boldsymbol{\theta}}(\tilde{\boldsymbol{\phi}}) \right] \boldsymbol{\lambda}_{\boldsymbol{\theta}}(\boldsymbol{\theta}^{\dagger}) \right\} \\
=& \left[T^{-\frac{1}{2}} \sum_{t=1}^{T} \mathbf{C}'_{\boldsymbol{\theta}}(\omega_t) \right] \left[\mathcal{I}_{\boldsymbol{\theta}\boldsymbol{\theta}}(\tilde{\boldsymbol{\phi}}) - \mathcal{I}_{\boldsymbol{\theta}\phi}(\tilde{\boldsymbol{\phi}})\mathcal{I}^{-1}_{\boldsymbol{\phi}\boldsymbol{\phi}}(\tilde{\boldsymbol{\phi}})\mathcal{I}_{\phi\boldsymbol{\theta}}(\tilde{\boldsymbol{\phi}}) \right]^{-1} \left[T^{-\frac{1}{2}} \sum_{t=1}^{T} \mathbf{C}'_{\boldsymbol{\theta}}(\omega_t) \right] \qquad \text{(A10)} \\
=& \begin{pmatrix} \tilde{r}_{2T} & \tilde{r}_{3T} & \tilde{r}_{4T} \end{pmatrix} \mathcal{V}^{-1}_{\boldsymbol{\rho}_a\boldsymbol{\rho}_a} \begin{pmatrix} \tilde{r}_{2T} \\ \tilde{r}_{3T} \\ \tilde{r}_{4T} \end{pmatrix},
\end{aligned}
$$

with the two equalities holding with probability approaching 1 and the second one following from $\Lambda_T = \{\sqrt{T}\lambda_{\boldsymbol{\theta}}\} \to \Lambda = \mathbb{R}^3$.

Assumption 4.6: Once again, we verify this for each term in (A5)-(A8). Specifically, (A5) follows from the fact that $\mathcal{H}+\mathcal{V} = O_p(T^{-\frac{1}{2}})$. In addition, note that $T^{\frac{1}{2}}\boldsymbol{\lambda}(\boldsymbol{\phi}_T, \boldsymbol{\theta}_T) = O(1)$ implies that

$$\phi_T = O(T^{-\frac{1}{2}}), \quad \alpha_T = O(T^{-\frac{1}{2}}), \quad \psi_{1T} = O(T^{-\frac{1}{8}}) \text{ and } \psi_{2T} = O(T^{-\frac{1}{4}}).$$

Thus, the slowest convergence rate is $T^{-\frac{1}{8}}$ because this is the rate of the $\{.\}$ terms in (A6) and (A7) and the [.] term in (A8).

The final step of the proof simply involves the application of the delta method to go from the autocovariances γ_j $(j = 0, ..., 4)$ to the autocorrelations $\rho_j = \gamma_j/\gamma_0$ $(j = 1, ..., 4)$, which delivers the expressions in the statement of the proposition. The intuition is that given that the restricted MLEs for σ_f^2 and σ_v^2 are such that in large samples the estimated model will perfectly match the sample variance and first autocovariance of y_t with probability approaching one, the first two components of $\mathcal{S}_{\boldsymbol{\theta},T}$ evaluated at $\tilde{\boldsymbol{\phi}}_T$ will be zero, which in turn implies that GET_T is effectively testing that the second, third and fourth autocovariances of y_t are simultaneously zero on the basis of their sample counterparts, but taking into account the sampling uncertainty in estimating those autocovariances when the true process is the local level model (7)-(10). □

B Additional results

B.1 Maximum likelihood estimation in the frequency domain

Henceforth, we assume that y_t is a covariance stationary series, which may require taking first or seasonal differences of the observations, as in the examples in Sections 2 and 3.

Let

$$I_{yy}(\omega) = \frac{1}{2\pi T} \sum_{t=1}^{T} \sum_{s=1}^{T} (y_t - \mu)(y_s - \mu)e^{-i(t-s)\omega}$$

denote the periodogram of y_t. If we assume that the spectral density $g_{yy}(\omega;\boldsymbol{\rho})$ is not zero at any of those frequencies, the so-called Whittle (1962)'s (discrete) spectral approximation to the log-likelihood function is

$$-\frac{T}{2}\ln(2\pi)-\frac{1}{2}\sum_{j=0}^{T-1}\ln|g_{yy}(\omega_j;\boldsymbol{\rho})|-\frac{1}{2}\sum_{j=0}^{T-1}\frac{2\pi I_{yy}(\omega_j)}{g_{yy}(\omega_j;\boldsymbol{\rho})}, \tag{B11}$$

where $\omega_j = 2\pi j/T$ (for $j = 0, \ldots T-1$) are the usual Fourier frequencies.

The MLE of μ, which only enters through $I_{yy}(\omega)$, 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

$$\frac{\partial l_t}{\partial \boldsymbol{\rho}}=\frac{1}{2}\sum_{j=0}^{T-1}\frac{\partial g_{yy}(\omega_j;\boldsymbol{\rho})}{\partial \boldsymbol{\rho}}M(\omega_j;\boldsymbol{\rho})m(\omega_j;\boldsymbol{\rho}), \tag{B12}$$

where $m(\omega;\boldsymbol{\rho}) = 2\pi I_{yy}(\omega) - g_{yy}(\omega;\boldsymbol{\rho})$ and $M(\omega;\boldsymbol{\rho}) = g_{yy}^{-2}(\omega;\boldsymbol{\rho})$.

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

$$\mathbf{Q}(\boldsymbol{\rho})=\frac{1}{4\pi}\int_{-\pi}^{\pi}\frac{\partial g_{yy}(\omega_j;\boldsymbol{\rho})}{\partial \boldsymbol{\rho}}M(\omega;\boldsymbol{\rho})\left\{\frac{\partial g_{yy}(\omega_j;\boldsymbol{\rho})}{\partial \boldsymbol{\rho}}\right\}^{*}d\omega, \tag{B13}$$

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

$$\boldsymbol{\Phi}(\boldsymbol{\rho})=\frac{1}{2}\sum_{j=0}^{T-1}\frac{\partial g_{yy}(\omega_j;\boldsymbol{\rho})}{\partial \boldsymbol{\rho}}M(\omega_j;\boldsymbol{\rho})\left\{\frac{\partial g_{yy}(\omega_j;\boldsymbol{\rho})}{\partial \boldsymbol{\rho}}\right\}^{*}. \tag{B14}$$

In fact, by selecting an artificially large value for T in (B14), one can approximate (B13) to any desired degree of accuracy. In addition, the univariate nature of y_t implies that both $g_{yy}(\omega_j;\boldsymbol{\rho})$ and its derivatives are real.

Formal results showing the strong consistency and asymptotic normality of the resulting ML estimators of dynamic latent variable models under suitable regularity conditions were provided by Dunsmuir (1979), who generalized earlier results for VARMA models by Dunsmuir and Hannan (1976). These authors also show the asymptotic equivalence between time and frequency domain MLEs.

B.2 The autocorrelation structure of the UCARIMA model

We can derive the autocovariance structure of $y_t = \Delta x_t$ by the usual inverse Fourier transformation $\gamma_{yy}(k) = cov(y_t, y_{t-k}) = \int_{-\pi}^{\pi} e^{i\omega k} g_{yy}(\omega)d\omega$ after exploiting that $g_{yy}(\omega)$ is the sum of the spectral densities of the signal and noise components, $s_t = \Delta z_t$ and $n_t = \Delta u_t$, respectively, which are cross-sectionally uncorrelated at all leads and lags. Specifically, we know that

$$y_t=\frac{1}{1-\psi_1 L-\psi_2 L^2}f_t+\frac{1-L}{1-\alpha L}u_t=s_t+n_t,$$

where the first component, s_t, is an AR(2) process while the second component, n_t, is an ARMA(1,1) with a unit root on the MA part. Thus,

$$\begin{aligned} g_{yy}(\omega;\boldsymbol{\rho}) &= g_{ss}(\omega;\boldsymbol{\rho}) + g_{nn}(\omega;\boldsymbol{\rho}) \\ &= \frac{\sigma_f^2}{(1-\psi_1 e^{-i\omega}-\psi_2 e^{-2i\omega})\,(1-\psi_1 e^{i\omega}-\psi_2 e^{2i\omega})} + \frac{(1-e^{-i\omega})(1-e^{i\omega})\sigma_v^2}{(1-\alpha e^{-i\omega})(1-\alpha e^{i\omega})} \\ &= \frac{\sigma_f^2}{(1+\psi_1^2+\psi_2^2)-2\psi_1(1-\psi_2)\cos\omega-2\psi_2\cos 2\omega} + \frac{2(1-\cos\omega)\sigma_v^2}{(1+\alpha^2)-2\alpha\cos\omega}. \quad \text{(B15)} \end{aligned}$$

However, the expressions for $\gamma_{yy}(k)$ are somewhat easier to obtain in the time domain as the sum of the autocovariances of the two underlying components.

The autocovariances of the AR(2) process for the signal are given by the usual Yule-Walker recursion

$$\gamma_{ss}(k) = \psi_1\gamma_s(k-1) + \psi_2\gamma_s(k-2), \tag{B16}$$

with initial conditions

$$\gamma_{ss}(0) = \left(\frac{1-\psi_2}{1+\psi_2}\right)\frac{\sigma_f^2}{(1-\psi_2)^2-\psi_1^2} \text{ and } \gamma_{ss}(1) = \left(\frac{\psi_1}{1+\psi_2}\right)\frac{\sigma_f^2}{(1-\psi_2)^2-\psi_1^2},$$

which yields

$$\gamma_{ss}(2) = \frac{\psi_1^2+\psi_2(1-\psi_2)}{1-\psi_2}\gamma_s(0),\ \gamma_{ss}(3) = \frac{\psi_1[\psi_1^2+\psi_2(2-\psi_2)]}{1-\psi_2}\gamma_s(0)$$

and

$$\gamma_{ss}(4) = \frac{\psi_1[\psi_1^3+\psi_1\psi_2(3-\psi_2)]+\psi_2^2(1-\psi_2)}{1-\psi_2}\gamma_s(0).$$

To find the solution for general k, it is convenient to find the roots of the characteristic equation (B16), which are given by $\delta_1 = \frac{1}{2}\psi_1 + \frac{1}{2}\sqrt{\psi_1^2+4\psi_2}$ and $\delta_2 = \frac{1}{2}\psi_1 - \frac{1}{2}\sqrt{\psi_1^2+4\psi_2}$.

When the roots are different (real or complex), the autocorrelation of order k will be given by

$$\gamma_{ss}(k) = \frac{\delta_1^{k+1}(1-\delta_2^2)-\delta_2^{k+1}(1-\delta_1^2)}{(\delta_1-\delta_2)(1+\delta_1\delta_2)}\gamma_s(0).$$

Applying L'Hôpital's rule, this simplifies to

$$\gamma_{ss}(k) = \left[1+k\frac{(1-\delta^2)}{(1+\delta^2)}\right]\delta^k\gamma_s(0)$$

when the two roots are equal, which happens for $\psi_2 = -\psi_1^2/4$ (see e.g. Fuller (1995)).

In turn, the autocovariances of the ARMA(1,1) process for the noise will be

$$\gamma_{nn}(0) = \sigma_v^2\left[1+\frac{(\alpha-1)^2}{1-\alpha^2}\right] = \frac{2\sigma_v^2}{\alpha+1},\ \gamma_{nn}(1) = \sigma_v^2\left[(\alpha-1)+\frac{(\alpha-1)^2\alpha}{1-\alpha^2}\right] = \frac{(\alpha-1)\sigma_v^2}{\alpha+1}$$

and $\gamma_{nn}(k) = [\alpha^{k-1}(\alpha-1)\sigma_v^2]/(\alpha+1)$. Finally, $\gamma_{yy}(k) = \gamma_{ss}(k) + \gamma_{nn}(k)$.

C Tables and figures

Table 1: Monte Carlo rejection rates (in %) under the null and alternative hypotheses for the white noise versus multiplicative seasonal AR test.

	Null hypothesis			Alternative hypotheses					
				H_{a_1}			H_{a_2}		
	1%	5%	10%	1%	5%	10%	1%	5%	10%
Panel A: Asymptotic critical values									
$T = 100$									
GET	1.1	4.1	8.2	29.8	45.9	55.2	27.3	48.7	60.9
LR	0.9	4.3	9.2	14.8	31.4	42.8	19.7	43.7	57.9
LM-AR(1)	0.7	4.1	9.0	14.9	29.9	39.6	2.6	9.3	16.0
LM-SAR(4)	0.5	3.7	8.0	11.6	27.7	38.4	2.5	9.5	16.1
MT	0.8	3.7	7.8	25.8	40.7	49.7	20.6	39.1	50.7
$T = 400$									
GET	1.0	4.9	9.8	87.9	94.7	96.8	92.5	97.8	99.1
LR	0.9	4.4	9.0	80.1	91.6	95.1	91.7	97.7	99.0
LM-AR(1)	0.9	4.7	9.5	58.1	76.1	83.5	3.1	10.4	17.2
LM-SAR(4)	1.1	5.0	9.7	58.4	78.2	85.5	5.1	13.7	21.7
MT	0.9	4.4	9.4	84.6	92.9	95.5	89.0	96.3	98.0
Panel B: Bootstrap critical values									
$T = 100$									
GET	*1.0*	*5.0*	*10.0*	29.3	48.2	57.7	26.5	51.9	64.2
LR	*1.0*	*5.0*	*10.0*	15.3	33.3	44.2	20.8	46.6	59.3
LM-AR(1)	*1.0*	*5.0*	*10.0*	16.8	32.3	40.9	3.1	10.8	17.0
LM-SAR(4)	*1.0*	*5.0*	*10.0*	15.4	31.7	41.6	3.8	11.6	18.4
MT	*1.0*	*5.0*	*10.0*	27.0	44.3	53.3	22.0	43.7	55.0
$T = 400$									
GET	*1.0*	*5.0*	*10.0*	87.4	94.7	96.8	92.1	97.9	99.1
LR	*1.0*	*5.0*	*10.0*	81.0	92.3	95.3	92.2	98.0	99.1
LM-AR(1)	*1.0*	*5.0*	*10.0*	60.1	76.9	84.1	3.7	10.8	17.7
LM-SAR(4)	*1.0*	*5.0*	*10.0*	57.3	78.2	86.0	4.9	13.7	22.1
MT	*1.0*	*5.0*	*10.0*	85.3	93.3	95.8	89.6	96.7	98.2

Notes: Results based on 10,000 samples. The mean and variance parameters φ_M and φ_V are estimated under the null using their sample analogs. GET is computed as defined in section 2.1. DGPs: the true unconditional mean and the variance of the innovations are set to 0 and 1, respectively, under both the null and alternative hypotheses. As for the alternative hypotheses,

$$(1 - 0.1L - 0.1L^2 - 0.1L^3 - 0.1L^4)y_t = \varepsilon_t \quad (H_{a_1})$$

and

$$(1 - 0.4L)(1 + 0.4L)(1 - 0.4L^4)(1 + 0.4L^4)y_t = \varepsilon_t \quad (H_{a_2}).$$

LM-AR(1) and LM-SAR(4) denote the Lagrange multiplier tests based on the score of an AR(1) and a seasonal SAR(4), respectively. MT refers to the two-sided version of GET. Finite sample critical values in Panel B are computed using a parametric bootstrap procedure.

Table 2: Monte Carlo rejection rates (in %) under the null and alternative hypotheses for the local level model versus the UCARIMA model (12) test.

	Null hypothesis			Alternative hypotheses H_{a_1}			Alternative hypotheses H_{a_2}		
	1%	5%	10%	1%	5%	10%	1%	5%	10%
Panel A: Asymptotic critical values									
$T = 100$									
GET	0.7	4.2	8.9	4.6	18.9	32.0	7.5	20.5	30.4
LR	1.0	5.2	10.4	11.3	30.3	43.7	10.8	27.4	39.9
2^{nd} autocorrelation	0.8	4.1	9.3	2.9	12.7	21.7	1.6	6.8	12.4
2^{nd} & 3^{rd} autocorrelation	0.8	4.1	8.8	4.1	17.2	29.0	1.5	6.6	11.9
$T = 400$									
GET	0.8	4.8	9.5	64.0	86.7	93.0	49.4	72.1	81.8
LR	0.8	5.4	10.4	75.2	90.7	95.3	62.9	82.8	89.8
2^{nd} autocorrelation	0.9	4.6	9.6	27.7	54.6	68.1	2.1	8.1	14.8
2^{nd} & 3^{rd} autocorrelation	0.9	4.5	9.6	48.4	75.4	85.6	2.0	7.8	14.0
Panel B: Bootstrap critical values									
$T = 100$									
GET	1.1	5.3	10.4	6.0	21.7	35.5	8.6	22.8	33.0
LR	1.9	5.3	10.4	11.0	27.0	38.9	10.3	24.1	35.1
2^{nd} autocorrelation	1.2	5.3	10.2	4.5	13.6	23.6	2.3	7.2	13.7
2^{nd} & 3^{rd} autocorrelation	1.3	5.5	10.0	5.5	19.8	31.4	2.2	7.9	13.2
$T = 400$									
GET	1.0	5.0	10.0	67,0	87,0	93,7	51,5	72,7	83,0
LR	1.2	5.2	10.2	71,4	88,1	92,9	59,0	78,8	86,4
2^{nd} autocorrelation	1.0	5.0	10.0	30,3	56,0	69,1	2,4	8,9	15,3
2^{nd} & 3^{rd} autocorrelation	1.0	5.0	10.1	52,8	76,8	85,8	2,5	8,4	14,2

Notes: Results based on 10,000 samples. The local level parameters σ_f^2 and σ_u^2 are estimated under the null. GET is computed as defined in section 3.1. DGPs: We simulate Gaussian shocks with $\sigma_f^2 = 1$ and $\sigma_v^2 = 0.5$ under both the null and the alternatives. Alternative hypotheses:

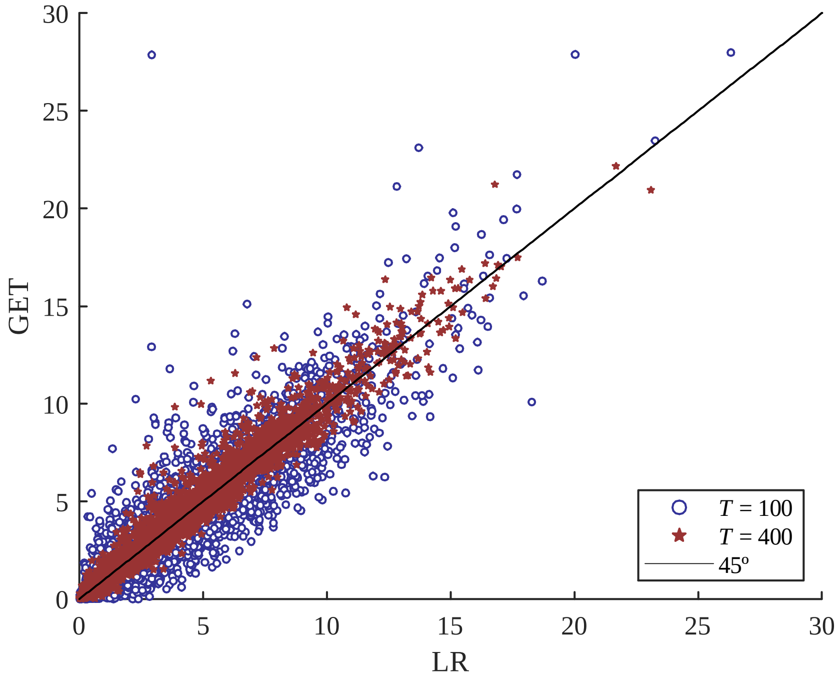
$$\left.\begin{array}{c}(1+0.5L+0.4L^2)\Delta z_t = f_t \\ (1-0.5L)u_t = v_t\end{array}\right\} \quad (H_{a_1})$$

and

$$\left.\begin{array}{c}(1-0.1L+0.5L^2)\Delta z_t = f_t \\ (1+0.5L)u_t = v_t\end{array}\right\} \quad (H_{a_2}).$$

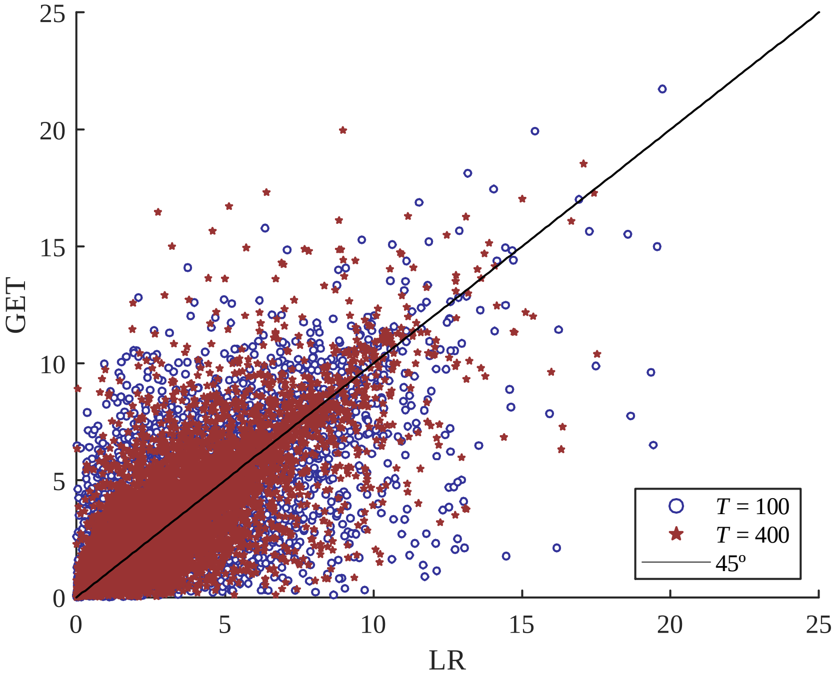
2^{nd} autocorrelation (2^{nd} & 3^{rd} autocorrelation) denote the moment test of no second-order (no second- or third-order) serial correlation in y_t. Finite sample critical values in Panel B are computed using a parametric bootstrap procedure.

Figure 1: Alignment of GET and LR under the null under null for the white noise versus multiplicative seasonal AR test.



Notes: Scatter plots of the GET_T and LR_T test statistics. Results based on 10,000 simulated samples of size T of $y \sim i.i.d.$ Gaussian. GET is computed as explain in section 2.1. The true mean and variance of the simulated data are set to 0 and 1, and the elements of φ are estimated using the sample mean and variance, respectively.

Figure 2: Alignment of GET and LR under the null for the local level model versus the UCARIMA model (12).



Notes: Scatter plots of the GET_T and LR_T test statistics. Results based on 10,000 simulated samples of size T of the model under the null with Gaussian shocks with $\sigma_f^2 = 1$ and $\sigma_v^2 = 0.5$. GET is computed as explained in section 3.1.