Revista Colombiana de Estadística - Theoretical Statistics
Julio 2020, volumen 43, no. 2, pp. 143 a 171
http://dx.doi.org/10.15446/rce.v43n2.83044

Optimal Detection of Bilinear Dependence in Short Panels of Regression Data

Detección óptima de dependencia bilineal en regresión con datos de panel cortos

Aziz Lmakri[1,a], Abdelhadi Akharif[1,b], Amal Mellouk[2,c]

[1]Department of Mathematics, Faculty of Sciences and Techniques, Abdelmalek Essaadi University, Tangier, Morocco

[2]Centre Régional des Métiers de l'Education et de la Formation, Tangier, Morocco

Abstract

In this paper, we propose parametric and nonparametric locally and asymptotically optimal tests for regression models with superdiagonal bilinear time series errors in short panel data (large n, small T). We establish a local asymptotic normality property– with respect to intercept μ, regression coefficient β, the scale parameter σ of the error, and the parameter b of panel superdiagonal bilinear model (which is the parameter of interest)– for a given density f_1 of the error terms. Rank-based versions of optimal parametric tests are provided. This result, which allows, by Hájek's representation theorem, the construction of locally asymptotically optimal rank-based tests for the null hypothesis $b = 0$ (absence of panel superdiagonal bilinear model). These tests –at specified innovation densities f_1– are optimal (most stringent), but remain valid under any actual underlying density. From contiguity, we obtain the limiting distribution of our test statistics under the null and local sequences of alternatives. The asymptotic relative efficiencies, with respect to the pseudo-Gaussian parametric tests, are derived. A Monte Carlo study confirms the good performance of the proposed tests.

***Key words*:** Bilinear process; local asymptotic normality; local asymptotic linearity; panel data; pseudo-Gaussian tests; rank tests.

Resumen

En este artículo, se proponen pruebas paramétricas y no paramétricas locales y asintóticamente óptimas para modelos de regresión con errores de series temporales bilineales superdiagonales en datos de panel cortos

[a]Ph.D. Student. E-mail: lmakriaziz@gmail.com
[b]Ph.D. E-mail: aakharif@gmail.com
[c]Ph.D. E-mail: prmelloukamal@gmail.com

(n grande, T pequeño). Se establece una propiedad de normalidad asintótica local con respecto a la intercepción μ, el coeficiente de regresión β, el parámetro de escala σ del error y el parámetro b del modelo bilineal superdiagonal con datos de panel (que es el parámetro de interés) para una densidad determinada f_1 de los términos de error. Se proporcionan versiones basadas en rangos de pruebas paramétricas óptimas. Este resultado permite, por el teorema de representación de Hájek, la construcción de pruebas locales basadas en rangos asintóticamente óptimas para la hipótesis nula $b = 0$ (ausencia del modelo bilineal superdiagonal con datos de panel). Estas pruebas, en densidades de innovación especicadas f_1, son óptimas (más estrictas), pero siguen siendo válidas en cualquier densidad subyacente. A partir de la contigüidad, se obtiene la distribución limitante de las estadísticas de prueba, bajo la hipótesis nula y una secuencia de alternativas locales. Se deriva eficiencia relativa asintótica de las pruebas, con respecto a las pruebas paramétricas pseudo-Gaussianas. Un análisis basado en simulaciones de Monte Carlo confirma el buen desempeño de las pruebas propuestas.

***Palabras clave*:** Datos de panel; Linealidad asintótica local; Normalidad asintótica local; Proceso bilineal; Prueba pseudo-gaussiana; Pruebas de rango.

1. Introduction

Recent evolution in theory and applications has provided very powerful convenient tools for the modelling of time series data, and in the last decades, we have seen a growing interest in nonlinear models. It has been shown that nonlinear time series models gives better approximations than higher-order linear ones simple in modelling nonlinear dynamic systems. One of the approaches to nonlinear time series modelling is the class of bilinear processes, introduced by Granger & Andersen (1978). Assuming (ε_t) is i.i.d. $(0, \sigma^2)$,

$$X_t = \sum_{j=1}^{p} a_j X_{t-j} + \sum_{j=1}^{q} c_j \varepsilon_{t-j} + \sum_{j=1}^{P} \sum_{k=1}^{Q} b_{jk} \varepsilon_{t-j} X_{t-k} + \varepsilon_t$$

defines the *bilinear process* (X_t) of order $(p, q; P, Q)$–shortly $BL(p, q; P, Q)$.

This interest is due to its widespread use in various fields, see for example, Maravall (1983), Rao & Gabr (1984), Weiss (1986). Regardless of theoretical difficulties, the fundamental probabilistic properties have been solved for several particular cases, for example, the stationarity and invertibility have been solved for first-order superdiagonal model by Guegan (1981). Testing problems and there power properties have been treated for null hypothesis of white noise against bilinear dependence in Hallin & Mélard (1988), Saikkonen & Luukkonen (1991), Benghabrit & Hallin (1992), Benghabrit & Hallin (1996) and Guegan & Pham (1992). The statistical problem of estimation of the parameters for some simple models have been considered in Pham & Tran (1981), Grahn (1995), Hristova (2005) and Tan & Wang (2015).

Regression models with correlated errors have been the focus of considerable attention in econometrics and statistics. Various manuscripts treat the problem of correlated errors in regression models in which the errors follow the linear models such as autoregressive (AR), moving average (MA) (e.g., Baltagi & Li, 1995), the mixed autoregressive and moving average (ARMA) models (e.g., Allal & El Melhaoui, 2006), or the nonlinear models such as RCAR, ARCH, fractional ARIMA and bilinear models (e.g., Hwang & Basawa, 1993, Dutta, 1999, Hallin, Taniguchi, Serroukh & Choy, 1999, and Elmezouar, Kadi & Gabr, 2012, etc.).

Consider the following panel data regression model in Pesaran (2015):

$$y_{i,t} = \mu + \beta' x_{i,t} + e_{i,t}, \quad i = 1, 2, \dots, n; \; t = 1, 2, \dots, T, \tag{1}$$

where $y_{i,t}$ is the observation on the i^{th} cross-sectional unit for the t^{th} time period, $x_{i,t}$ denotes the $K \times 1$ vector of observations on the non-stochastic regressors. $(\mu, \beta')' \in \mathbb{R}^{K+1}$ is the corresponding regression coefficients. Here, the error terms $e_{i,t}$ are assumed to follow a simple case of bilinear model with panel data, which takes the following form

$$e_{i,t} = b e_{i,t-l} \varepsilon_{i,t-k} + \varepsilon_{i,t} \qquad \text{with} \quad l > k \geqslant 1, \tag{2}$$

where $\varepsilon_{i,t} \sim i.i.d.(0, \sigma^2)$ for all i and t.

Test of homogeneity for panel bilinear time series model have been treated in Lee, Kim, & Lee (2013) and Kim (2014). Furthermore, probabilistic properties such as stationarity and invertibility have been studied in Quinn (1982) remains valid in panel bilinear model (2). Denote by $\mathcal{F}_{i,t}(\varepsilon)$ and $\mathcal{F}_{i,t}(e)$ the $\sigma-$algebras generated by $\{\varepsilon_{i,s} | s \leqslant t\}$ and $\{e_{i,s} | s \leqslant t\}$, respectively. Then,

1. Equation (2) admits a unique stationary solution $e_{i,t}$ if and only if $b^2\sigma^2 < 1$, and given by
$$e_{i,t} = \varepsilon_{i,t} + \sum_{j=1}^{\infty} b^j \varepsilon_{i,t-lj} \prod_{s=1}^{j} \varepsilon_{i,t-k-(s-1)l}.$$
2. Equation (2) is invertible if and only if $2b^2\sigma^2 < 1$, in this case, one can write
$$\varepsilon_{i,t} = e_{i,t} + \sum_{j=1}^{\infty} (-b)^j e_{i,t-kj} \prod_{s=1}^{j} e_{i,t-l-(s-1)k}.$$

Clearly, model (1) reduces to the classical multiple regression model

$$y_{i,t} = \mu + \beta' x_{i,t} + \varepsilon_{i,t},$$

with constant coefficients μ and β if and only if $b = 0$. The detection problem we are addressing consists of testing the null hypothesis $\mathcal{H}_0 : b = 0$ with unspecified μ, β, σ^2 and f_1 against the alternative $\mathcal{H}_1 : b \neq 0$. Clearly this testing problem corresponds to testing serial independence against bilinear serial dependence in model (1).

In this research, to derive optimal tests, the *uniform local asymptotic normality* (ULAN) property is established for a class of panel regression models with superdiagonal bilinear time series errors via the quadratic mean differentiability of $f^{1/2}$, where f is the density of $\varepsilon_{i,t}$. This last property (see, Le Cam & Yang, 2000), has recognized success in a variety of testing problems: see, Swensen (1985), Akharif & Hallin (2003), Cassart, Hallin & Paindaveine (2011), Bennala, Hallin & Paindaveine (2012) and Fihri, Akharif, Mellouk & Hallin (2020).

Our statistical tests are based on the ULAN property. These tests are shown to be asymptotically efficient and their asymptotic power is also derived.

ULAN plays a fundamental role in this treatment and leads us to construct locally and asymptotically optimal parametric tests. The special case of the pseudo-Gaussian tests (optimal under Gaussian densities but valid under finite-variance non-Gaussian ones) is derived, but unfortunately, their local asymptotic power under non-Gaussian g_1 (especially the skew and heavy-tailed ones), can be extremely poor, which leads us to construct a rank-based optimal tests (van der Waerden, Wilcoxon, Laplace, data-driven scores, etc.) based on the Hájek-Le Cam approach.

Asymptotic relative efficiencies with respect to the pseudo-Gaussian procedure show that the van der Waerden version of our rank-based tests uniformly dominates its pseudo-Gaussian counterpart.

The paper is organized as follows. In Section 2 we introduce notations, assumptions and state the ULAN property for model (1)-(2). Section 3 is devoted to prove a local asymptotic linearity property. These results are used in the derivation of locally asymptotically optimal (most stringent) tests, and in the computation of their asymptotic powers. The particular case of the pseudo-Gaussian tests is investigated in Section 3.2. Optimal rank tests are derived in Section 4 and some special cases (van der Waerden, Wilcoxon and Laplace scores) are considered in Section 4.3. Section 5.1 provides asymptotic relative efficiencies, and simulations are carried out in Section 5.2 to investigate the finite-sample performance of our tests. Finally, we provide some conclusions.

2. Uniform Local Asymptotic Normality

2.1. Notations and Main Assumptions

Denote by $\mathbf{P}^{(n)}_{\mu,\beta',\sigma^2,b;f_1}$ the probability distribution of the observations $\left[y_1^{(n)'}, y_2^{(n)'}, \ldots, y_n^{(n)'}\right]'$, where $y_i^{(n)} := (y_{i,1}, \ldots, y_{i,T})'$ is generated by (1) and (2), and by $\mathbf{P}^{(n)}_{\mu,\beta',\sigma^2,0;f_1}$ the probability distribution under the null hypothesis $e_{i,t} = \varepsilon_{i,t} \sim i.i.d.(0, \sigma^2)$. Assume that $\{\varepsilon_{i,t}\}_{(1\leq i\leq n, 1\leq t\leq T)}$ is unobservable sequence, with density $f : \varepsilon \mapsto f(\varepsilon) := (1/\sigma)f_1(\varepsilon/\sigma)$, where f_1 belongs to some adequate class of standardized densities (3). We suppose that the vector of starting values

$$e_0^{(n)} := \{(e_{i,1-l}^{(n)}\varepsilon_{i,1-k}, e_{i,2-l}^{(n)}\varepsilon_{i,2-k}, \ldots, e_{i,k-l}^{(n)}\varepsilon_{i,0}, e_{i,k+1-l}^{(n)}, \ldots, e_{i,-1}^{(n)}, e_{i,0}^{(n)}), i = 1, \ldots, n\}$$

is observable for each individual i.

Throughout this paper, we consider the class of standardized densities

$$\mathcal{F}_0 := \left\{ f_1 : f_1(u) > 0 \quad \forall u \in \mathbb{R}, \int_{-1}^{1} f_1(u)du = 0.5 = \int_{-\infty}^{0} f_1(u)du \right\}. \tag{3}$$

Note that, for f such that $f_1 \in \mathcal{F}_0$, the median and median absolute deviation are 0 and σ, respectively. This standardization, contrary to the usual one based on the mean and the standard deviation, avoids all moment assumptions; it plays the role of an identification constraint, and has no impact on subsequent results.

The main technical tool in our derivation of optimal tests is the *uniform local asymptotic normality*, with respect to $(\mu, \beta', \sigma^2, b)'$, at $(\mu, \beta', \sigma^2, 0)'$, of the families of distributions

$$\mathcal{P}_{f_1}^{(n)} := \left\{ \mathbf{P}_{\mu,\beta,\sigma^2,b;f_1}^{(n)} : (\mu, \beta) \in \mathbb{R}^{K+1}, \sigma^2 > 0, b \in \mathbb{R}^* \text{ and } 2b^2\sigma^2 < 1 \right\}. \tag{4}$$

Establishing ULAN property requires some technical assumptions about the innovation density f_1 (**Assumption (A)**) and the asymptotic behavior of the regressors (**Assumption (B)**).

Assumption (A)

(A.1) $f_1 \in \mathcal{F}_0$;

(A.2) f_1 is $\mathcal{C}^1$ on $\mathbb{R}$, with first derivative f_1' and letting $\Phi_{f_1} = -f_1'/f_1$, assume that $I(f_1) := \int_{\mathbb{R}} \Phi_{f_1}^2(u) f_1(u)du$, $J(f_1) := \int_{\mathbb{R}} u^2 \Phi_{f_1}^2(u) f_1(u)du$ and $K(f_1) := \int_{\mathbb{R}} u \Phi_{f_1}^2(u) f_1(u)du$ are finite.

Denote by $\mathcal{F}_A$ the set of all densities satisfying **Assumption (A)**.

Let $\mathbf{C}^{(n)} := \frac{1}{n} \sum_{i=1}^{n} \sum_{t=1}^{T} \mathbf{x}_{i,t}\mathbf{x}_{i,t}'$, the following assumption concern the asymptotic behavior of regression coefficients, it is standard in the context of rank-based inference.

Assumption (B)

(B.1) The limits $\lim_{n\to\infty} \mathbf{C}^{(n)} =: \mathbf{C}$, where $\mathbf{C}$ is positive definite. Letting $\mathbf{K}^{(n)} := (\mathbf{C}^{(n)})^{-1/2}$, note that $\lim_{n\to\infty} \mathbf{K}^{(n)} =: \mathbf{K} = \mathbf{C}^{-1/2}$;

(B.2) for all k and n, $\overline{x}_k^{(n)} := \frac{1}{nT} \sum_{i=1}^{n} \sum_{t=1}^{T} \mathbf{x}_{k;i,t} = 0$;

(B.3) the classical Noether (1949) conditions hold:

$$\lim_{n\to\infty} \frac{\max\limits_{1\leq i\leq n} x^2_{k;i,t}}{\sum\limits_{i=1}^{n}\sum\limits_{t=1}^{T} \mathbf{x}^2_{k;i,t}} = 0 \qquad k = 1, 2, \ldots, K, \quad t = 1, 2, \ldots, T.$$

Interesting special cases are

(i) The Student distributions (with $\nu > 2$ degrees of freedom), with standardized density

$$f_1(u) = f_{t_\nu}(u) := \frac{\Gamma((\nu+1)/2)}{\Gamma(\nu/2)} \sqrt{a_\nu/\pi\nu}(1 + a_\nu u^2/\nu)^{-(\nu+1)/2},$$

with $I(f_1) = a_\nu(\nu+1)/(\nu+3)$ and $J(f_1) = 3(\nu+1)/(\nu+3)$, the normalizing constant $a_\nu > 0$ is such that $f_{t_\nu} \in \mathcal{F}_A$.

(ii) The Gaussian distributions, with standardized density (with mean zero and variance $1/a$)

$$f_1(u) = f_{\mathcal{N}}(u) := \sqrt{a/2\pi}\exp(-au^2/2),$$

with $I(f_1) = a \simeq 0.4549$ and $J(f_1) = 3$; these values also can be obtained by taking limits, as $\nu \to \infty$, of the corresponding Student values since $a_\nu \to a$ as $\nu \to \infty$.

(iii) The double-exponential or Laplace distributions, with standardized density

$$f_1(u) = f_{\mathcal{L}}(u) := (1/2d)\exp(-|u|/d),$$

with $I(f_1) = 1/d^2$ and $J(f_1) = 2$, the normalizing constant $d = 1/\ln(2) \simeq 1.4426$ is such that $f_{\mathcal{L}} \in \mathcal{F}_A$.

(iv) The logistic distributions, with standardized density

$$f_1(u) = f_{Log}(u) := \sqrt{b}\exp(-\sqrt{b}u)/(1 + \exp(-\sqrt{b}u))^2,$$

with $I(f_1) = b/3$ and $J(f_1) = (12 + \pi^2)/9$, the normalizing constant $b = (\ln 3)^2 \simeq 1.2069$ is such that $f_{\mathcal{L}} \in \mathcal{F}_A$.

2.2. Uniform Local Asymptotic Normality

In this section, we shall state the uniform local asymptotic normality property for the model (1), with respect to intercept μ, regression coefficient β, scale parameter σ^2 and the parameter of interest b, for fixed density $f_1 \in \mathcal{F}_A$, the reader is referred to Le Cam & Yang (2000).

For this purpose, let $\tau^{(n)} := \left(\tau_1^{(n)}, \tau_2^{(n)'}, \tau_3^{(n)}, \tau_4^{(n)}\right)'$ be a sequence of real vectors in $\mathbb{R}^{K+3}$ such that $\tau^{(n)'}\tau^{(n)}$ is uniformly bounded as $n \to \infty$ and let $\theta :=$

$(\mu, \beta', \sigma^2, b = 0)'$. In addition, we consider sequences of local alternatives of the form $\theta + \nu^{(n)}\tau^{(n)}$ where

$$\nu^{(n)} := n^{-1/2} \begin{bmatrix} 1 & \mathbf{0} & 0 & 0 \\ \mathbf{0} & \mathbf{K}^{(n)} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

The test is equivalent to

$$\mathbf{P}^{(n)}_{\theta;f_1} : \tau_4^{(n)} = 0 \;\; \text{against} \;\; \mathbf{P}^{(n)}_{\theta+\nu^{(n)}\tau^{(n)};f_1} : \tau_4^{(n)} \neq 0.$$

Denote by $\Lambda^{(n)}_{\theta+\nu^{(n)}\tau^{(n)}/\theta;f}$ the logarithm of the likelihood ratio (conditional on $e_0^{(n)}$) for $\mathbf{P}^{(n)}_{\theta+\nu^{(n)}\tau^{(n)};f}$ against $\mathbf{P}^{(n)}_{\theta;f}$. Then,

$$\Lambda^{(n)}_{\theta+\nu^{(n)}\tau^{(n)}/\theta;f} = \sum_{i=1}^{n}\sum_{t=1}^{T}\left[\log f(e_{i,t} + \sum_{j=1}^{\infty}(-n^{-\frac{1}{2}}\tau_4^{(n)})^j e_{i,t-kj}\prod_{s=1}^{j} e_{i,t-l-(s-1)k}) - \log f(e_{i,t})\right]. \quad (5)$$

Define the standardized residuals as

$$Z_{i,t} = Z_{i,t}(\mu, \beta, \sigma^2) := \sigma^{-1}(y_{i,t} - \mu - \beta'\mathbf{x}_{i,t}),$$

for $i = 1, 2, \ldots, n; \quad t = 1, 2, \ldots, T$ and note that, under the null hypothesis, it coincides with $\varepsilon_{i,t}/\sigma$. We have then the following result.

Proposition 1. *Let Assumption (B) holds, fix $f_1 \in \mathcal{F}_A$. Then, the family $\mathcal{P}^{(n)}_{f_1}$ is ULAN (for $n \to \infty$ with T fixed) at any $\theta = (\mu, \beta', \sigma^2, 0)'$, with $(K+3)$-dimensional central sequence*

$$\Delta^{(n)}_{f_1}(\theta) := \begin{bmatrix} \Delta^{(n)}_{f_1;1}(\theta) \\ \Delta^{(n)}_{f_1;2}(\theta) \\ \Delta^{(n)}_{f_1;3}(\theta) \\ \Delta^{(n)}_{f_1;4}(\theta) \end{bmatrix} := \begin{bmatrix} \frac{n^{-1/2}}{\sigma}\sum_{i=1}^{n}\sum_{t=1}^{T}\Phi_{f_1}(Z_{i,t}) \\ \frac{n^{-1/2}}{\sigma}\sum_{i=1}^{n}\sum_{t=1}^{T}\Phi_{f_1}(Z_{i,t})(\boldsymbol{K}^{(n)})'x_{i,t} \\ \frac{n^{-1/2}}{2\sigma^2}\sum_{i=1}^{n}\sum_{t=1}^{T}\left[\Phi_{f_1}(Z_{i,t})Z_{i,t} - 1\right] \\ n^{-1/2}\sigma\sum_{i=1}^{n}\sum_{t=l+1}^{T}\Phi_{f_1}(Z_{i,t})Z_{i,t-k}Z_{i,t-l} \end{bmatrix}, \quad (6)$$

and $(K+3) \times (K+3)$-information matrix

$$\Gamma_{f_1}(\theta) := \big(\Gamma_{f_1;ij}\big)_{1\le i,j\le 4} = \begin{bmatrix} \frac{T}{\sigma^2}I(f_1) & 0 & \frac{T}{2\sigma^3}K(f_1) & 0 \\ \boldsymbol{0} & \frac{T}{\sigma^2}I(f_1)\boldsymbol{I}_K & \boldsymbol{0} & \boldsymbol{0} \\ \frac{T}{2\sigma^3}K(f_1) & 0 & \frac{T}{4\sigma^4}\big(J(f_1)-1\big) & 0 \\ 0 & 0 & 0 & (T-l)I(f_1)\sigma^2\sigma^4_{f_1} \end{bmatrix}. \tag{7}$$

More precisely, for any $\theta^{(n)} = (\mu^{(n)}, \beta^{(n)\prime}, \sigma^{2^{(n)}}, 0)'$ *such that* $\mu^{(n)} - \mu$, $(\boldsymbol{K}^{(n)})^{-1}(\beta^{(n)} - \beta)$ *and* $\sigma^{2^{(n)}} - \sigma^2$ *are* $O(n^{-1/2})$*, and for any bounded sequence* $\tau^{(n)} \in \mathbb{R}^{K+3}$*, we have under* $\boldsymbol{P}^{(n)}_{\theta^{(n)};f_1}$*, as* $n \to \infty$ *with* T *fixed ,*

$$\begin{aligned} \Lambda^{(n)}_{\theta^{(n)}+\nu^{(n)}\tau^{(n)}/\theta^{(n)};f_1} &:= \log\left(\frac{d\boldsymbol{P}^{(n)}_{\theta^{(n)}+\nu^{(n)}\tau^{(n)};f_1}}{d\boldsymbol{P}^{(n)}_{\theta^{(n)};f_1}}\right) \\ &= \tau^{(n)\prime}\Delta^{(n)}_{f_1}(\theta^{(n)}) - \tfrac{1}{2}\tau^{(n)\prime}\Gamma_{f_1}(\theta)\tau^{(n)} + o_p(1), \end{aligned} \tag{8}$$

and $\Delta^{(n)}_{f_1}(\theta^{(n)})$ *converges in distribution to a* $(K+3)^2$*-variate normal distribution with mean zero and covariance matrix* $\Gamma_{f_1}(\theta)$*.*

Proof. See appendix.

From this result, we have under $\mathbf{P}^{(n)}_{\theta;f_1}$,

$$\begin{bmatrix} \Delta^{(n)}_{f_1}(\theta) \\ \Lambda^{(n)}_{\theta^{(n)}+\nu^{(n)}\tau^{(n)}/\theta^{(n)};f_1} \end{bmatrix} \xrightarrow[n\to\infty]{\mathcal{D}} \mathcal{N}\left(\begin{bmatrix} 0 \\ -\frac{1}{2}\tau'\Gamma_{f_1}(\theta)\tau \end{bmatrix}, \begin{bmatrix} \Gamma_{f_1}(\theta) & \Gamma_{f_1}(\theta)\tau \\ \tau'\Gamma_{f_1}(\theta) & \tau'\Gamma_{f_1}(\theta)\tau \end{bmatrix}\right). \tag{9}$$

Consequently, since the hypotheses $\mathbf{P}^{(n)}_{\theta;f_1}$ and $\mathbf{P}^{(n)}_{\theta+\nu^{(n)}\tau^{(n)};f_1}$ are contiguous, Le Cam's third lemma leads to the convergence of $\Delta^{(n)}_{f_1}(\theta)$ under $\mathbf{P}^{(n)}_{\theta+\nu^{(n)}\tau^{(n)};f_1}$. In this case we have

$$\Delta^{(n)}_{f_1}(\theta) \xrightarrow[n\to\infty]{\mathcal{D}} \mathcal{N}(\Gamma_{f_1}(\theta)\tau, \Gamma_{f_1}(\theta)). \tag{10}$$

3. Locally Asymptotically Optimal Tests

In this section, we are interested in testing the null hypothesis $b = 0$ of randomness of the error regression model in (1), with unspecified error density $f_1 \in \mathcal{F}_0$, unspecified μ, β and unspecified error scale $\sigma-$ formally can be written as

$$\mathcal{H}_0^{(n)} := \bigcup_{g_1 \in \mathcal{F}_0} \mathcal{H}_0^{(n)}(g_1) := \bigcup_{g_1 \in \mathcal{F}_0} \bigcup_{\mu \in \mathbb{R}} \bigcup_{\beta' \in \mathbb{R}^K} \bigcup_{\sigma^2 > 0} \{\mathbf{P}^{(n)}_{\mu,\beta',\sigma^2,0;g_1}\}.$$

Parametric alternatives takes the form (for some fixed standardized $f_1 \in \mathcal{F}_A$)

$$\mathcal{H}_1^{(n)}(f_1) := \bigcup_{\mu \in \mathbb{R}} \bigcup_{\beta' \in \mathbb{R}^K} \bigcup_{\sigma^2 > 0} \bigcup_{b \in \mathbb{R}^*} \{\mathbf{P}^{(n)}_{\mu,\beta',\sigma^2,b;f_1}\}.$$

The parameters μ, β and σ^2 thus are nuisance parameters, while b is the parameter of interest. Before turning to this semiparametric hypothesis $\mathcal{H}_0^{(n)}$ (unspecified density), let us first investigate the parametric problem of testing $\mathcal{H}_0^{(n)}(f_1)$ (while f_1 remains specified) against $\mathcal{H}_1^{(n)}(f_1)$.

3.1. Optimal Parametric Tests

In this subsection, we construct a locally asymptotically optimal (namely, most stringent) tests in presence of nuisance parameters for testing serial independence in model (1). The notion of most stringency is a concept of optimality (see e.g., Wald (1943)). We suppose that the innovation density f_1 is specified, the main consequence of the ULAN results of Proposition 1 is that for each θ, and for given $f_1 \in \mathcal{F}_A$, the sequences of local experiments

$$\zeta_{f_1}^{(n)}(\theta) := \left(\mathbb{R}^n, B^n, \mathcal{P}^n = \left\{\mathbf{P}^{(n)}_{\theta+\nu^{(n)}\tau;f_1} | \tau \in \mathbb{R}^{K+3}\right\}\right)$$

converge weakly to the $(K+3)-$dimensional Gaussian shift experiments

$$\mathcal{G}_{f_1}(\theta) := \left(\mathbb{R}^{K+3}, B^{K+3}, \mathcal{P} = \left\{\mathcal{N}\big(\Gamma_{f_1}(\theta)\tau, \Gamma_{f_1}(\theta)\big) | \tau \in \mathbb{R}^{K+3}\right\}\right).$$

The classical theory of hypothesis testing in Gaussian shifts (see Section 11.9 of Le Cam, 1986) provides the general form for locally asymptotically most stringent tests of hypotheses in ULAN models. In this case, the null hypothesis $\mathcal{H}_0^{(n)}(f_1) =: \mathcal{H}_{f_1}^{(n)} = \bigcup_{\mu \in \mathbb{R}} \bigcup_{\beta' \in \mathbb{R}^K} \bigcup_{\sigma^2 > 0} \{\mathbf{P}^{(n)}_{\mu,\beta',\sigma^2,0;f_1}\}$ and the local alternative $\mathcal{H}_1^{(n)}(f_1)$ can be expressed as

$$\mathcal{H}_{f_1}^{(n)} : \tau \in \mathcal{M}(\Omega) \quad \text{against} \quad \mathcal{H}_1^{(n)}(f_1) : \tau \notin \mathcal{M}(\Omega),$$

where $M(\Omega)$ is the linear subspace of dimension $K+2$ of $\mathbb{R}^{K+3}$ generated by the matrix $\Omega' := (\mathbf{I}_{K+2}, \mathbf{0})$. Such tests, should be based on

$$\begin{aligned} Q_{f_1}^{(n)}(\theta) &:= \Delta_{f_1}^{(n)'}(\theta)\left[\Gamma_{f_1}^{-1}(\theta) - \Omega\big(\Omega'\Gamma_{f_1}(\theta)\Omega\big)^{-1}\Omega'\right]\Delta_{f_1}^{(n)}(\theta) \\ &= \Delta_{f_1;4}^{(n)^2}(\theta)/\Gamma_{f_1;44}(\theta). \end{aligned} \quad (11)$$

As θ remains unspecified under the null, we will need to replace it with some estimate. For this purpose, we assume the existence of $\widehat{\theta} := \widehat{\theta}_n$ satisfying the following assumption

Assumption (C). The estimate $\widehat{\theta}_n$ is such that

(i) $\widehat{\theta}_n$ is $\sqrt{n}$-consistent, i.e., for all $f_1 \in \mathcal{F}_A$ and all $\varepsilon > 0$, there exist $c := c(f_1, \theta, \varepsilon)$ and $N := N(f_1, \theta, \varepsilon)$ such that under $\mathbf{P}^{(n)}_{\mu,\beta',\sigma^2,0;f_1}$, we have

$$P(\sqrt{n} \parallel \widehat{\theta}_n - \theta \parallel > c) < \varepsilon \quad \forall n \geqslant N.$$

(ii) $\widehat{\theta}_n$ is locally asymptotically discrete, i.e., for all fixed value $s > 0$, the number of possible values of $\widehat{\theta}_n$ in

$$\mathcal{B} = \{u \in \mathbb{R}^{K+3}, \quad n^{-1/2} \parallel u - \theta \parallel \leqslant s\}$$

is bounded as $n \to \infty$.

Note that the condition (i) on the rate of convergence in probability of the estimates is satisfied by several estimates such as the maximum likelihood estimates, the Yule-Walker estimates, the M-estimates and the least square estimates; part (ii) has little practical implications.

The following proposition shows that substituting $\widehat{\theta}_n$ for θ does not influence the asymptotic behavior of the test statistic (11).

Proposition 2 (Asymptotic linearity)**.** *Suppose that Assumptions (A),(B) and (C) hold. Let $\widehat{\theta}_n$ be a deterministic sequence satisfying $n^{1/2}(\widehat{\theta}_n - \theta)$ is bounded by a constant $c > 0$. Then, under $\boldsymbol{P}^{(n)}_{\mu,\beta',\sigma^2,0;f_1}$, as $n \to \infty$, we have*

(i)

$$\Delta^{(n)}_{f_1}(\widehat{\theta}_n) - \Delta^{(n)}_{f_1}(\theta) + n^{1/2}\Gamma_{f_1}(\theta)(\widehat{\theta}_n - \theta) = o_P(1). \tag{12}$$

(ii)

$$\Delta^{(n)}_{f_1;4}(\widehat{\theta}_n) - \Delta^{(n)}_{f_1;4}(\theta) = o_P(1). \tag{13}$$

***Proof*.** See appendix. □

The following proposition then results from classical results on ULAN families (see, Le Cam, 1986, chapter 11).

Proposition 3. *Suppose that Assumptions (A), (B) and (C) hold. Then,*

(i) $Q^{(n)}_{f_1}(\widehat{\theta}_n) = Q^{(n)}_{f_1}(\theta) + o_P(1)$ is asymptotically chi-square, with 1 degrees of freedom under $\boldsymbol{P}^{(n)}_{\theta;f_1}$, and asymptotically noncentral chi-square, still with 1 degrees of freedom but with noncentrality parameter $\lambda_{f_1} := \tau_4^2\sigma^2(T - l)I(f_1)\sigma^4_{f_1}$ under $\boldsymbol{P}^{(n)}_{\theta+\nu^{(n)}\tau^{(n)};f_1}$;

(ii) *the sequence of tests rejecting the null hypothesis* $\mathcal{H}_{f_1}^{(n)}$ *(with standardized density* f_1*) whenever* $Q_{f_1}^{(n)}(\widehat{\theta}_n)$ *exceeds the* $(1-\alpha)$*-quantile of a chi-square distribution with one degree of freedom, is locally asymptotically optimal (most stringent), at asymptotic level* α*, for* $\mathcal{H}_{f_1}^{(n)}$ *against*

$$\bigcup_{\mu\in\mathbb{R}}\bigcup_{\beta'\in\mathbb{R}^K}\bigcup_{\sigma^2>0}\bigcup_{b\in\mathbb{R}^*}\{\boldsymbol{P}^{(n)}_{\mu,\beta',\sigma^2,b;f_1}\};$$

(iii) *the sequence of tests has asymptotic power* $1-F(\chi^2_{1,1-\alpha},\lambda_{f_1})$*, at* $\boldsymbol{P}^{(n)}_{\theta+\nu^{(n)}\tau^{(n)};f_1}$*, where* $F(.,\lambda_{f_1})$ *denotes the noncentral chi-square distribution function with one degree of freedom and non centrality parameter* λ_{f_1}*.*

Proof. See appendix. □

The Gaussian versions of $\Delta^{(n)}_{f_1;4}(\theta)$, $\Gamma_{f_1;44}(\theta)$ and $Q^{(n)}_{f_1}(\theta)$ are obtained by letting $f_1 = f_{\mathcal{N}}$ (standardized normal density $\mathcal{N}(0,1/a)$), this case is an exception, however, as $\Phi_{f_1}(u)$, $I(f_1)$ and $\sigma^4_{f_1}$ reduce to au, a and $1/a^2$, respectively, then one easily checks that

$$\Delta^{(n)}_{\mathcal{N};4}(\theta) = n^{-1/2}\sigma a\sum_{i=1}^{n}\sum_{t=l+1}^{T} Z_{i,t}Z_{i,t-k}Z_{i,t-l}, \qquad \Gamma_{\mathcal{N};44}(\theta) = \frac{\sigma^2}{a}(T-l),$$

and

$$Q^{(n)}_{\mathcal{N}}(\theta) = \frac{a^3}{T-l}\left[n^{\frac{-1}{2}}\sum_{i=1}^{n}\sum_{t=l+1}^{T} Z_{i,t}Z_{i,t-k}Z_{i,t-l}\right]^2, \tag{14}$$

respectively.

The Gaussian tests $Q^{(n)}_{\mathcal{N}}(\theta)$ unfortunately are valid under normal densities only, i.e., needs f_1 to be indicated as a standardized Gaussian one, then the parameter a also has to be fixed. In the following section, we demonstrate that a proper version–namely, pseudo-Gaussian test, that is, tests that are valid under a broad class of non-Gaussian densities with finite variance, while remaining optimal under Gaussian ones– in general, are preferable.

3.2. Pseudo-Gaussian Test

Herein, we construct a pseudo-Gaussian version of the Gaussian test $Q^{(n)}_{\mathcal{N}}(\theta)$. The Gaussian central sequence $\Delta^{(n)}_{\mathcal{N};4}(\theta)$ allows us to construct asymptotically optimal tests under $f_1 = f_{\mathcal{N}}$, hence for efficient detection of bilinear dependence in the parametric Gaussian model characterized by Gaussian disturbances. Extending the validity of the Gaussian optimal test to densities g_1 in a broad class of densities is of course highly desirable. Let us show that this is indeed possible and that a slight modification, $\Delta^{*(n)}_{\mathcal{N};4}$, say, of the efficient central sequence $\Delta^{(n)}_{\mathcal{N};4}$ leads to a *pseudo-Gaussian test* which remains valid when the actual density

$g_1 \in \mathcal{F}_A^{(2)}$ of all densities in $\mathcal{F}_A$ with finite variance. Note that $Z_{i,t}$ for $Z_{i,t}(\mu, \beta, \sigma^2)$ and let $m_1^{(n)} = m_1^{(n)}(\mu, \beta', \sigma^2) := \frac{1}{nT} \sum_{i=1}^{n} \sum_{t=1}^{T} Z_{i,t}(\mu, \beta', \sigma^2)$ is a $\sqrt{n}-$consistent estimator, under $\mathbf{P}_{\theta;g_1}^{(n)}$, of $\mu_1(g_1) := \int_{\mathbb{R}} u g_1(u) du$. Define

$$\Delta_{\mathcal{N};4}^{*(n)}(\theta) = n^{-1/2} \sigma a \sum_{i=1}^{n} \sum_{t=l+1}^{T} (Z_{i,t} - m_1^{(n)})(Z_{i,t-k} - m_1^{(n)})(Z_{i,t-l} - m_1^{(n)}). \quad (15)$$

Decomposing $Z_{i,t} - m_1^{(n)}$ into $(Z_{i,t} - \mu_1(g_1)) + (\mu_1(g_1) - m_1^{(n)})$, then, it easily follows from that, under $\mathbf{P}_{\theta;g_1}^{(n)}$, as $n \to \infty$ with T fixed

$$\Delta_{\mathcal{N};4}^{*(n)}(\theta) =$$
$$n^{-1/2} \sigma a \sum_{i=1}^{n} \sum_{t=l+1}^{T} (Z_{i,t} - \mu_1(g_1))(Z_{i,t-k} - \mu_1(g_1))(Z_{i,t-l} - \mu_1(g_1)) + o_p(1). \quad (16)$$

Then, under $\mathbf{P}_{\theta;g_1}^{(n)}$, $\Delta_{\mathcal{N};4}^{*(n)}(\theta)$ is asymptotically normal with mean 0 and variance $\Gamma_{\mathcal{N};g_1;44}^{*}(\theta) = a^2 \sigma^2 (T-l) \sigma_{g_1}^6$ where $\sigma_{g_1}^2 := \int_{\mathbb{R}} (z - \mu_1(g_1))^2 g_1(z) dz$.

On the other hand, it is easy to see that, still under $\mathbf{P}_{\theta+\nu^{(n)}\tau^{(n)};g_1}^{(n)}$, $\Delta_{\mathcal{N};4}^{*(n)}(\theta)$ and the log-likelihood $\Lambda_{\theta+\nu^{(n)}\tau^{(n)}/\theta;g_1}^{(n)}$ are jointly binormal; the desired result then follows from a routine application of Le Cam's third lemma. Since the intercept μ, the regression coefficients β, and the scale parameter σ^2 under the null hypothesis remain unspecified, some care has to be taken with the asymptotic impact of estimating μ, β, and σ^2 under unspecified density g_1.

Define the non-standardized centered residuals

$$V_{i,t}(\beta) := \sigma(Z_{i,t}(\mu, \beta, \sigma^2) - m_1^{(n)}) = y_{i,t} - \beta' \mathbf{x}_{i,t} - \frac{1}{nT} \sum_{i=1}^{n} \sum_{t=1}^{T} y_{i,t}.$$

A pseudo Gaussian test may then be based on a statistic of the form

$$\begin{aligned} Q_{\mathcal{N};g_1}^{*(n)}(\theta) &:= \Delta_{\mathcal{N};4}^{*(n)^2}(\theta) / \Gamma_{\mathcal{N};g_1;44}^{*}(\theta) \\ &:= \frac{1}{(T-l)\sigma_g^6} \left[n^{-1/2} \sum_{i=1}^{n} \sum_{t=l+1}^{T} V_{i,t}(\beta) V_{i,t-k}(\beta) V_{i,t-l}(\beta) \right]^2 \\ &=: Q_{\mathcal{N};g}^{*(n)}(\beta), \end{aligned} \quad (17)$$

with g is defined by $g(u) = (1/\sigma) g_1(u/\sigma)$. Clearly $Q_{\mathcal{N};g}^{*(n)}(\beta)$ depends only on β, which justifies the notation.

In practice, the pseudo-Gaussian tests will be based on the statistics

$$Q_{\mathcal{N}}^{\dagger(n)}(\widehat{\beta}) := \frac{1}{(T-l)s^6}\left[n^{-1/2}\sum_{i=1}^{n}\sum_{t=l+1}^{T} V_{i,t}(\widehat{\beta})V_{i,t-k}(\widehat{\beta})V_{i,t-l}(\widehat{\beta})\right]^2, \qquad (18)$$

where $\widehat{\beta}$ is an arbitrary $n^{1/2}(\mathbf{K}^{(n)})^{-1}$-consistent estimator of β and $s^2 := \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T} V_{i,t}^2(\widehat{\beta})$ is the empirical variance of the $V_{i,t}(\widehat{\beta})$'s. Consider the class $\mathcal{F}_A^{(2)}$ of all densities $g_1 \in \mathcal{F}_A$ such that $\sigma_{g_1}^2 < \infty$. Then under $\mathbf{P}_{\theta;g_1}^{(n)}$, and for any bounded sequence $\tau^{(n)} = \left(\tau_1^{(n)}, \tau_2^{(n)\prime}, \tau_3^{(n)}, 0\right)' \in \mathbb{R}^{K+3}$, as $n \to \infty$ with T fixed

$$\Delta_{\mathcal{N}}^{*(n)}(\theta + \nu^{(n)}\tau^{(n)}) - \Delta_{\mathcal{N}}^{*(n)}(\theta) = -\Gamma_{\mathcal{N};g_1}^{*}(\theta)\tau^{(n)} + o_P(1), \qquad (19)$$

with

$$\Gamma_{\mathcal{N};g_1}^{*}(\theta) := E\left[\Delta_{\mathcal{N}}^{*(n)}(\theta)\left(\Delta_{g_1}^{(n)}(\theta)\right)'\right]$$

$$= \begin{bmatrix} \frac{aT}{\sigma^2} & 0 & 0 & a(T-l)\mu_1^2(g_1) \\ \mathbf{0} & \frac{aT}{\sigma^2}\mathbf{I}_K & \mathbf{0} & \mathbf{0} \\ 0 & 0 & \frac{aT}{\sigma^4}\sigma_{g_1}^2 & 0 \\ 0 & 0 & 0 & a(T-l)\sigma^2\sigma_{g_1}^4 \end{bmatrix}. \qquad (20)$$

The following result is immediate from (19). Let Assumption (B) holds, assume that $\widehat{\theta}_n$ satisfies Assumptions (C) and fix $\theta \in \mathbb{R}^{K+3}$, we have

$$\Delta_{\mathcal{N};4}^{*(n)}(\widehat{\theta}_n) - \Delta_{\mathcal{N};4}^{*(n)}(\theta) = o_P(1). \qquad (21)$$

Showing that, under $\mathbf{P}_{\theta;g_1}^{(n)}$, as $n \to \infty$, with T fixed, $Q_{\mathcal{N}}^{\dagger(n)}(\widehat{\beta}) - Q_{\mathcal{N};g}^{*(n)}(\beta) = o_p(1)$.

The following result summarizes the asymptotic properties of the pseudo-Gaussian tests.

Proposition 4. *Let Assumptions (A), (B) and (C) hold, for any* $g_1 \in \mathcal{F}_A^{(2)}$. *Then,*

(i) $Q_{\mathcal{N}}^{\dagger(n)}(\widehat{\beta})$ *is asymptotically chi-square with* 1 *degrees of freedom under* $\boldsymbol{P}_{\mu,\beta,\sigma^2,0;g_1}^{(n)}$, *and asymptotically noncentral chi-square, still with* 1 *degrees of freedom but with noncentrality parameter* $\lambda_{\mathcal{N}} = (T-l)\sigma_g^2\tau_4^2$ *under*

$$\boldsymbol{P}_{\theta+n^{-1/2}\nu^{(n)}\tau;g_1}^{(n)};$$

(ii) the sequence of tests rejecting the null hypothesis

$$\mathcal{H}_A^{(n)^2} := \bigcup_{g_1 \in \mathcal{F}_A^{(2)}} \bigcup_{\mu} \bigcup_{\beta} \bigcup_{\sigma^2} \left\{ \boldsymbol{P}^{(n)}_{\mu,\beta,\sigma^2,0;g_1} \right\}$$

whenever $Q^{\dagger}_{\mathcal{N}} > \chi^2_{1,1-\alpha}$*, is locally asymptotically most stringent, at asymptotic probability level* α*, for* $\mathcal{H}_A^{(n)^2}$ *against alternatives of the form*

$$\bigcup_{\mu} \bigcup_{\beta} \bigcup_{\sigma} \bigcup_{b \in \mathbb{R}^*} \left\{ \boldsymbol{P}^{(n)}_{\mu,\beta,\sigma^2,b;f_{\mathcal{N}}} \right\};$$

(iii) the asymptotic power under $\boldsymbol{P}^{(n)}_{\theta+n^{-1/2}\nu^{(n)}\tau;g_1}$ *is* $1 - F(\chi^2_{1,1-\alpha}, \lambda_{\mathcal{N}})$.

The test statistic $Q^{\dagger(n)}_{\mathcal{N}}(\widehat{\beta})$ thus defines a pseudo-Gaussian test, that is, a test which is optimal under Gaussian assumptions but remains valid under a much broader class of densities.

4. Optimal Rank Tests

General results by Hallin & Werker (2003) indicate that semiparametrically efficient rank-based procedures can be obtained in relation to ranks being maximal invariants under model-generating groups of transformations $(\mathcal{G}^{(nT)}, \star)$. More precisely, note that the null hypothesis $\mathcal{H}_\beta^{(n)} := \bigcup_{g_1 \in \mathcal{F}_0} \bigcup_{\mu \in \mathbb{R}} \bigcup_{\sigma^2 > 0} \left\{ \mathbf{P}^{(n)}_{\mu,\beta,\sigma^2,0;g_1} \right\}$ is invariant under the action of the group $(\mathcal{G}^{(nT)}, \star)$ of all transformations $\mathcal{G}_h$ of $\mathbb{R}^{nT}$ such that

$$\mathcal{G}_h(y_{1,1}, \ldots, y_{n,T}) := (\beta' x_{1,1} + h(y_{1,1} - \beta' x_{1,1}), \ldots, \beta' x_{n,T} + h(y_{n,T} - \beta' x_{n,T})),$$

where $u \mapsto h(u)$ is continuous and increasing and $\lim_{u \to \pm\infty} h(u) = \pm\infty$. It is easy to check that $(\mathcal{G}^{(nT)}, \star)$ is actually a generating group for the null hypothesis $\mathcal{H}_\beta^{(n)}$.

4.1. Rank-Based Versions of Central Sequences

A maximal invariant for the group $(\mathcal{G}^{(nT)}, \star)$ is known to be the vector $R^{(n)} = R^{(n)}(\beta) := \left(R^{(n)}_{1,1}, \ldots, R^{(n)}_{n,T}\right)'$ where $R^{(n)}_{i,t} = R^{(n)}_{i,t}(\beta)$ denotes the rank of residual $Z^{(n)}_{i,t}(\beta)$ among $Z^{(n)}_{1,1}(\beta), \ldots, Z^{(n)}_{n,T}(\beta)$. Moreover, μ and σ^2 have no impact on residual ranks, hence we can assume that they are specified, which justifies the notation $Z^{(n)}_{i,t}(\beta)$ (instead of $Z^{(n)}_{i,t}(\mu, \beta, \sigma^2)$) and $R^{(n)}_{i,t}(\beta)$.

General results on semiparametric efficiency indicate that in such context, the expectation (under the null hypothesis) of the central sequence $\Delta^{(n)}_{f_1;4}(\theta)$ conditional on those ranks $R^{(n)}$ yields a version of the semiparametrically efficient central sequence (at f_1 and θ) given by:

$$\underset{\sim}{\Delta}^{(n)}_{f_1;4}(\theta) := E\big[\Delta^{(n)}_{f_1;4}(\theta) \mid R^{(n)}\big]. \tag{22}$$

In practice, the conditional expectation definition (22) of $\underset{\sim}{\Delta}^{(n)}_{f_1;4}(\theta)$ (an *exact-score* linear rank statistic) is not convenient, and the explicit *approximate-score* form (for simplicity, we are using the same notation as for the exact-score version) is preferable and given by (the notation $\underset{\sim}{\Delta}^{(n)}_{f_1;4}(\beta,\sigma)$ reflects the fact that it only depends on β and σ)

$$\underset{\sim}{\Delta}^{(n)}_{f_1;4}(\beta,\sigma) := \\ n^{-1/2}\sigma\sum_{i=1}^{n}\sum_{t=l+1}^{T}\left\{\varphi_{f_1}\left(\frac{R^{(n)}_{i,t}(\beta)}{N+1}\right)F_1^{-1}\left(\frac{R^{(n)}_{i,t-k}(\beta)}{N+1}\right)F_1^{-1}\left(\frac{R^{(n)}_{i,t-l}(\beta)}{N+1}\right)-\overline{m}_{f_1}\right\}, \tag{23}$$

with $N = n(T-l)$, $\varphi_{f_1} := \Phi_{f_1} \circ F_1^{-1}$ and

$$\overline{m}_{f_1} := \frac{1}{N(N-1)(N-2)}\sum\sum\sum_{1\leqslant t_1\neq t_2\neq t_3\leqslant N}\varphi_{f_1}\Big(\frac{t_1}{N+1}\Big)F_1^{-1}\Big(\frac{t_2}{N+1}\Big)F_1^{-1}\Big(\frac{t_3}{N+1}\Big).$$

Let

$$\begin{aligned}
s^{(n)^2}_{f_1} &:= \frac{1}{N(N-1)(N-2)}\sum\sum\sum_{1\leqslant t_1\neq t_2\neq t_3\leqslant N}\left[\varphi_{f_1}\Big(\frac{t_1}{N+1}\Big)F_1^{-1}\Big(\frac{t_2}{N+1}\Big)F_1^{-1}\Big(\frac{t_3}{N+1}\Big)\right]^2 \\
&+ \frac{2}{N(N-1)(N-2)(N-3)}\sum\sum\sum\sum_{1\leqslant t_1\neq t_2\neq t_3\neq t_4\leqslant N}\varphi_{f_1}\Big(\frac{t_1}{N+1}\Big)\varphi_{f_1}\Big(\frac{t_2}{N+1}\Big)F_1^{-1}\Big(\frac{t_2}{N+1}\Big) \\
&\qquad\times\left[F_1^{-1}\Big(\frac{t_3}{N+1}\Big)\right]^2F_1^{-1}\Big(\frac{t_4}{N+1}\Big) \\
&+ \frac{2}{N(N-1)(N-2)(N-3)(N-4)}\sum\sum\sum\sum\sum_{1\leqslant t_1\neq t_2\neq t_3\neq t_4\neq t_5\leqslant N}\varphi_{f_1}\Big(\frac{t_1}{N+1}\Big)F_1^{-1}\Big(\frac{t_2}{N+1}\Big)F_1^{-1}\Big(\frac{t_3}{N+1}\Big) \\
&\qquad\times\varphi_{f_1}\Big(\frac{t_3}{N+1}\Big)F_1^{-1}\Big(\frac{t_4}{N+1}\Big)F_1^{-1}\Big(\frac{t_5}{N+1}\Big) \\
&+ \frac{N-5}{N(N-1)(N-2)(N-3)(N-4)(N-5)}\sum\sum\sum\sum\sum\sum_{1\leqslant t_1\neq t_2\neq t_3\neq t_4\neq t_5\neq t_6\leqslant N}\varphi_{f_1}\Big(\frac{t_1}{N+1}\Big)F_1^{-1}\Big(\frac{t_2}{N+1}\Big)F_1^{-1}\Big(\frac{t_3}{N+1}\Big) \\
&\qquad\times\varphi_{f_1}\Big(\frac{t_4}{N+1}\Big)F_1^{-1}\Big(\frac{t_5}{N+1}\Big)F_1^{-1}\Big(\frac{t_6}{N+1}\Big) - N\overline{m}^2_{f_1}.
\end{aligned} \tag{24}$$

The following asymptotic representation result (25) shows that both (22) and (23) yield rank-based version of the central sequence $\Delta^{(n)}_{f_1;4}(\theta)$.

Proposition 5. *Fix $\theta = (\mu, \beta', \sigma^2, 0)'$, let f_1 and $g_1 \in \mathcal{F}_A$. Then,* *(i) under* $\boldsymbol{P}^{(n)}_{\theta;g_1}$*, as* $n \to \infty$ *with* T *fixed*

$$\begin{aligned}\underset{\sim}{\Delta}^{(n)}_{f_1;4}(\beta,\sigma) &:= E^{(n)}_{g_1}\big[\Delta^{(n)}_{f_1;4}(\theta) \mid R^{(n)}_{1,1}(\beta),\ldots,R^{(n)}_{n,T}(\beta)\big] + o_{L^2}(1) \\ &= \underline{\Delta}^{(n)}_{f_1,g_1;4}(\theta) + o_{L^2}(1),\end{aligned} \tag{25}$$

with

$$\underline{\Delta}^{(n)}_{f_1,g_1;4}(\theta) := n^{-1/2}\sigma\sum_{i=1}^{n}\sum_{t=l+1}^{T}\varphi_{f_1}\big(G_1(Z_{i,t})\big)F_1^{-1}\big(G_1(Z_{i,t-k})\big)F_1^{-1}\big(G_1(Z_{i,t-l})\big); \tag{26}$$

where G_1 *is the distribution function associated with* g_1*,*

(ii) under $\boldsymbol{P}^{(n)}_{\theta;g_1}$*,* $\underset{\sim}{\Delta}^{(n)}_{f_1;4}(\beta,\sigma)$ *has mean zero and variance*

$$\underset{\sim}{\Gamma}^{(n)}_{f_1;44}(\sigma^2) := (T-l)\sigma^2 s^{(n)^2}_{f_1} = \underline{\Gamma}_{f_1;44}(\sigma^2) + o(1)$$

as $n \to \infty$ *with* T *fixed, where* $\underline{\Gamma}_{f_1;44}(\sigma^2) := (T-l)I(f_1)\sigma^2\sigma^4_{f_1}$*;*

(iii) $\underline{\Delta}^{(n)}_{f_1,g_1;4}(\theta)$ *is asymptotically normal, with mean zero under* $\boldsymbol{P}^{(n)}_{\theta;g_1}$*, mean* $(T-l)\mathcal{I}(f_1,g_1)\sigma^2(f_1,g_1)\sigma^2\tau_4$ *under* $\boldsymbol{P}^{(n)}_{\theta+\nu^{(n)}\tau;g_1}$ *and variance* $\underline{\Gamma}_{f_1;44}(\sigma^2)$ *under both, with*

$$\mathcal{I}(f_1,g_1) = \int_0^1 \Phi_{f_1}(F_1^{-1}(u))\Phi_{g_1}(G_1^{-1}(u))du$$

and

$$\sigma(f_1,g_1) = \int_0^1 F_1^{-1}(v)G_1^{-1}(v)dv.$$

Proof. See appendix. □

4.2. Optimal Rank Tests

The parameters μ,β and σ^2 remain unspecified under the null, since β has only an influence on the ranks, a consistent estimator $\widehat{\beta} := \widehat{\beta}^{(n)}$ has to be substituted for the actual β value, yielding aligned ranks $R^{(n)}_{i,t}(\widehat{\beta}^{(n)})$. The effect of this alignment procedure is taken care of in a similar way as in Section 3, via the asymptotic linearity results of Propositions 6 and 7 below.

Proposition 6. *Let Assumption (B) holds and fix* $\mu \in \mathbb{R}, \beta \in \mathbb{R}^K, \sigma^2 > 0$*,* f_1 *and* $g_1 \in \mathcal{F}_A$*. Then, for any bounded sequence* $\tau_2^{(n)} \in \mathbb{R}^K$*, under* $\boldsymbol{P}^{(n)}_{\mu,\beta,\sigma^2,0;g_1}$*, as* $n \to \infty$ *with* T *fixed, we have*

$$\underset{\sim}{\Delta}^{(n)}_{f_1}(\mu,\beta+n^{-1/2}\boldsymbol{K}^{(n)}\tau_2^{(n)},\sigma^2,0) - \underset{\sim}{\Delta}^{(n)}_{f_1}(\mu,\beta,\sigma^2,0) = -\underset{\sim}{\Gamma}_{f_1;g_1}\tau^{(n)} + o_P(1), \tag{27}$$

with

$$\underset{\sim f_1;g_1}{\Gamma} := E\Big[\underline{\Delta}^{(n)}_{f_1,g_1}(\theta)\big(\Delta^{(n)}_{g_1}(\theta)\big)'\Big]$$

$$= \begin{bmatrix} \frac{T}{\sigma^2}\mathcal{I}(f_1,g_1)\boldsymbol{I}_K & \boldsymbol{0} \\ 0 & (T-l)\mathcal{I}(f_1,g_1)\sigma^2(f_1,g_1)\sigma^2 \end{bmatrix}. \tag{28}$$

The following proposition then is an immediate corollary of Proposition 6 and Lemma 4.4 in Kreiss (1987).

Proposition 7. *Let Assumption (B) holds, assume that $\widehat{\beta}$ satisfies Assumption (C) and fix $\mu \in \mathbb{R}, \beta \in \mathbb{R}^K, \sigma^2 > 0$, f_1 and $g_1 \in \mathcal{F}_A$. Then, under $\boldsymbol{P}^{(n)}_{\mu,\beta,\sigma^2,0;g_1}$, as $n \to \infty$ with T fixed*

$$\underset{\sim f_1;4}{\Delta^{(n)}}(\mu,\widehat{\beta}^{(n)},\sigma^2,0) - \underset{\sim f_1;4}{\Delta^{(n)}}(\mu,\beta,\sigma^2,0) = o_P(1). \tag{29}$$

Local asymptotic optimality with density f_1 is achieved by the test based on

$$\underset{\sim f_1}{Q^{(n)}}(\theta) := \underset{\sim f_1;4}{\Delta^{(n)}}^2(\theta)\underset{\sim f_1;44}{\Gamma}(\sigma^2) =$$

$$\frac{1}{(T-l)s^{(n)2}_{f_1}}\left[n^{-1/2}\sum_{i=1}^{n}\sum_{t=l+1}^{T}\left\{\varphi_{f_1}\Big(\frac{R^{(n)}_{i,t}(\beta)}{N+1}\Big)F_1^{-1}\Big(\frac{R^{(n)}_{i,t-k}(\beta)}{N+1}\Big)F_1^{-1}\Big(\frac{R^{(n)}_{i,t-l}(\beta)}{N+1}\Big)-\overline{m}_{f_1}\right\}\right]^2$$

$$=: \underset{\sim f_1}{Q^{(n)}}(\beta). \tag{30}$$

More precisely, we have the following result.

Proposition 8. *Let Assumptions (A), (B) and (C) hold. Then*

(i) $\underset{\sim f_1}{Q^{(n)}}(\widehat{\beta}) = \underset{\sim f_1}{Q^{(n)}}(\beta) + o_P(1)$ *is asymptotically chi-square, with* 1 *degrees of freedom under $\boldsymbol{P}^{(n)}_{\mu,\beta,\sigma^2,0;g_1}$, and asymptotically noncentral chi-square, still with* 1 *degrees of freedom but with noncentrality parameter*

$$\lambda_{f_1,g_1} := \frac{(T-l)\mathcal{I}^2(f_1,g_1)\sigma^4(f_1,g_1)\sigma^2}{I(f_1)\sigma^4_{f_1}}\tau_4^2 \quad \textit{under} \quad \boldsymbol{P}^{(n)}_{\theta+\nu^{(n)}\tau;g_1};$$

(ii) the sequence of tests rejecting the null hypothesis

$$\mathcal{H}^{(n)}_A := \bigcup_{g_1\in\mathcal{F}_A}\bigcup_{\mu\in\mathbb{R}}\bigcup_{\beta'\in\mathbb{R}^K}\bigcup_{\sigma^2>0}\{\boldsymbol{P}^{(n)}_{\mu,\beta',\sigma^2,0;g_1}\}$$

whenever $\underset{\sim f_1}{Q^{(n)}}(\widehat{\beta})$ exceeds the $(1-\alpha)$-quantile of a chi-square distribution with one degree of freedom, is locally asymptotically optimal (most stringent), at asymptotic probability level α, for $\mathcal{H}^{(n)}_A$ against

$$\bigcup_{\mu\in\mathbb{R}}\bigcup_{\beta'\in\mathbb{R}^K}\bigcup_{\sigma^2>0}\bigcup_{b\in\mathbb{R}^*}\{\boldsymbol{P}^{(n)}_{\mu,\beta',\sigma^2,b;f_1}\};$$

(iii) the sequence of tests has asymptotic power $1-F(\chi^2_{1,1-\alpha},\lambda_{f_1})$, *at* $\boldsymbol{P}^{(n)}_{\theta+\nu^{(n)}\tau;f_1}$, *where* $F(.,\lambda_{f_1,g_1})$ *denotes the noncentral chi-square distribution function with one degree of freedom and non centrality parameter* λ_{f_1,g_1}.

4.3. Important Particular Cases

The statistic $\underset{\sim f_1}{Q^{(n)}}(\widehat{\beta})$ is providing a general form for the optimal rank tests of the null hypothesis of serial independence of model (1). The three most important particular cases for the test statistic presented are the *van der Waerden* (normal scores), *Wilcoxon* (logistic scores) and *Laplace* (double exponential scores) test statistics, which are respectively optimal at normal, logistic and double exponential distributions.

(i) ***van der Waerden's*** test statistic is given by

$$\underset{\sim vdW}{Q^{(n)}}(\widehat{\beta}) := \frac{a^2}{(T-l)s^{(n)2}_{f_\mathcal{N}}}\left[n^{-1/2}\sum_{i=1}^{n}\sum_{t=l+1}^{T}\left\{\Psi^{-1}\left(\frac{R^{(n)}_{i,t}(\widehat{\beta})}{N+1}\right)\Psi^{-1}\left(\frac{R^{(n)}_{i,t-k}(\widehat{\beta})}{N+1}\right)\Psi^{-1}\left(\frac{R^{(n)}_{i,t-l}(\widehat{\beta})}{N+1}\right)-\overline{m}_{f_\mathcal{N}}\right\}\right]^2$$

with

$$\overline{m}_{f_\mathcal{N}} = \frac{1}{N(N-1)(N-2)}\sum\sum\sum_{1\leqslant t_1\neq t_2\neq t_3\leqslant N}\Psi^{-1}\left(\frac{t_1}{N+1}\right)\Psi^{-1}\left(\frac{t_2}{N+1}\right)\Psi^{-1}\left(\frac{t_3}{N+1}\right)$$

where Ψ is the standard normal distribution function.

(ii) ***Wilcoxon's*** test statistic is given by

$$\underset{\sim W}{Q^{(n)}}(\widehat{\beta}) := \frac{1}{(T-l)bs^{(n)2}_{l}}\left[n^{-1/2}\sum_{i=1}^{n}\sum_{t=l+1}^{T}\left\{\left(\frac{2R^{(n)}_{i,t}(\widehat{\beta})}{N+1}-1\right)\log\left(\frac{R^{(n)}_{i,t-k}(\widehat{\beta})}{N+1-R^{(n)}_{i,t-k}(\widehat{\beta})}\right)\times\log\left(\frac{R^{(n)}_{i,t-l}(\widehat{\beta})}{N+1-R^{(n)}_{i,t-l}(\widehat{\beta})}\right)-\overline{m}_l\right\}\right]^2$$

with

$$\overline{m}_l = \frac{1}{N(N-1)(N-2)}\sum\sum\sum_{1\leqslant t_1\neq t_2\neq t_3\leqslant N}\left(\frac{2t_1}{N+1}-1\right)\log\left(\frac{t_2}{N+1-t_2}\right)\log\left(\frac{t_3}{N+1-t_3}\right).$$

(*iii*) ***Laplace's*** test statistic is given by

$$\underset{\sim\mathcal{L}}{Q}^{(n)}(\widehat{\beta}) := \frac{d^2}{(T-l)s_{\mathcal{D}e}^{(n)^2}}\left[n^{-1/2}\sum_{i=1}^{n}\sum_{t=l+1}^{T}\left\{sign\left[F_1^{-1}\left(\frac{R_{i,t}^{(n)}(\widehat{\beta})}{N+1}\right)\right]F_1^{-1}\left(\frac{R_{i,t-k}^{(n)}(\widehat{\beta})}{N+1}\right)\right.\right.$$
$$\left.\left.\times F_1^{-1}\left(\frac{R_{i,t-l}^{(n)}(\widehat{\beta})}{N+1}\right) - \overline{m}_{\mathcal{D}e}\right\}\right]^2$$

with

$$\overline{m}_{\mathcal{D}e} = \frac{1}{N(N-1)(N-2)}\sum\sum\sum_{1\leqslant t_1\neq t_2\neq t_3\leqslant N} sign\left[F_1^{-1}\left(\frac{t_1}{N+1}\right)\right]F_1^{-1}\left(\frac{t_2}{N+1}\right)F_1^{-1}\left(\frac{t_3}{N+1}\right),$$

where F_1 is the distribution function of the double-exponential

$$F_1^{-1}(u) = \begin{cases} d\log(2u) & if \quad 0 < u \leqslant \frac{1}{2} \\ -d\log(2-2u) & if \quad \frac{1}{2} \leqslant u < 1. \end{cases}$$

It is worth noting that the scale factors a (for ***van der Waerden***), b (for ***Wilcoxon***) and d (for ***Laplace***) disappear in the final expression of the test statistics, due to the exact standardization by $s_{f_{\mathcal{N}}}^{(n)}$, $s_l^{(n)}$, and $s_{\mathcal{D}e}^{(n)}$ respectively. This confirms that the choice of the median of absolute deviations as a scale parameter in the definition of $\mathcal{F}_0$ has no impact on the results.

5. Power Comparison and Simulations

5.1. Asymptotic Relative Efficiencies

The Asymptotic Relative Efficiencies (AREs) of the rank-based tests $\underset{\sim f_1}{Q}^{(n)}(\widehat{\beta})$ with respect to the pseudo-Gaussian tests $Q_{\mathcal{N}}^{\dagger(n)}$ directly follow as ratios of noncentrality parameters under local alternatives. In order to compare the performance of the parametric and nonparametric tests presented, we calculate the AREs of nonparametric tests compared to the pseudo-Gaussian tests.

Proposition 9. *Let $f_1 \in \mathcal{F}_A$. Then, the AREs under $g_1 \in \mathcal{F}_A^{(2)}$, of the rank tests based on $\underset{\sim f_1}{Q}^{(n)}(\widehat{\beta})$ with respect to the pseudo-Gaussian tests based on $Q_{\mathcal{N}}^{\dagger(n)}$, when testing $\boldsymbol{P}_{\theta;g_1}^{(n)}$ against $\boldsymbol{P}_{\theta+\nu^{(n)}\tau;g_1}^{(n)}$, are*

$$ARE_{g_1}(\underset{\sim f_1}{Q}^{(n)}(\widehat{\beta})/Q_{\mathcal{N}}^{\dagger(n)}) = \left(\frac{\lambda_{f_1,g_1}}{\lambda_{\mathcal{N}}}\right)^2 = \left(\frac{\mathcal{I}^2(f_1,g_1)\sigma^4(f_1,g_1)}{I(f_1)\sigma_{g_1}^2\sigma_{f_1}^4}\right)^2. \quad (31)$$

The table (1) gives the numerical values of (31) for

$$\underset{\sim f_1}{Q^{(n)}}(\widehat{\beta}) = \underset{\sim vdW}{Q^{(n)}}(\widehat{\beta}),\quad \underset{\sim W}{Q^{(n)}}(\widehat{\beta}),\quad \underset{\sim La}{Q^{(n)}}(\widehat{\beta}),\quad \underset{\sim t_5}{Q^{(n)}}(\widehat{\beta}),\quad \underset{\sim s\mathcal{N}(10)}{Q^{(n)}}(\widehat{\beta}) \text{ and } \underset{\sim st_5(10)}{Q^{(n)}}(\widehat{\beta})$$

under the densities g : $\mathcal{N}$ormal, Logistic, $\mathcal{D}$ouble exponential, Student-t_5, skew normal $s\mathcal{N}(10)$ and skew Student $st_5(10)$ [1].

The results obtained are satisfactory and all are good, particularly so under heavy tails. Also, note that the AREs of the proposed van der Waerden tests with respect to the parametric Gaussian tests are uniformly larger than or equal to one for all distributions considered in Table 1, and are equal to one in the Gaussian case only (this result is proved in Chernoff & Savage (1958)), which means that rank-based tests are asymptotically more powerful than Gaussian tests. Note also that each value is maximum in its corresponding column. Thus, at each of the densities, nonparametric tests perform better, compared to pseudo-Gaussian tests, among the efficiencies achieved by the van der Waerden, Wilcoxon and Laplace tests.

Table 1: AREs, under normal, logistic, Double exponential, Student (5 degree of freedom), skew normal $s\mathcal{N}(\delta)$ and skew Student $st_5(\delta)$ (with $\delta = 10$) densities, of various rank tests based on van der Waerden, Wilcoxon, Laplace, Student, skew normal and skew Student scores, with respect to their pseudo-Gaussian counterpart.

Scores f_1 \ Actual density g_1	$\mathcal{N}$	l	$\mathcal{D}e$	t_5	$s\mathcal{N}(10)$	$st_5(10)$
Van der Waerden	1.0000	1.0613	1.3944	1.1435	2.0752	3.4375
Wilcoxon	0.8825	2.3115	2.1514	1.4804	4.1503	2.9597
Laplace	0.4486	1.6468	4.0000	1.8985	5.2279	4.0672
Student-t_5	0.7318	2.1393	2.3002	2.5625	3.7472	5.5718
skew normal $s\mathcal{N}(10)$	0.5117	1.8205	2.5651	1.9317	5.9721	4.1755
skew Student $st_5(10)$	0.7285	1.6414	1.8435	1.8540	3.2328	6.1059

5.2. Results of Monte Carlo Simulations

In this section, we conduct a Monte Carlo experiment to investigate the finite sample performance of the proposed procedures and behavior of our rank tests under a variety of error distributions. More precisely, we considered the model

$$\begin{cases} y_{i,t} &= \mu + \beta' x_{i,t} + e_{i,t} \\ e_{i,t} &= b e_{i,t-4}\varepsilon_{i,t-1} + \varepsilon_{i,t}, , \end{cases} \tag{32}$$

[1]See, for instance, Azzalini & Capitanio (2003) for a definition of skew normal and skew-t densities.

with

- $\star$ $i = 1, 2, \dots, 100$ and $t = 1, 2, \dots, 14$ [2],
- $\star$ $\mu = 1$, $\beta = (1, 1)'$, $x_{i,t} = \begin{bmatrix} x_{1;i,t} \\ x_{2;i,t} \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 10 & 1 \\ 1 & 20 \end{bmatrix} \right)$,
- $\star$ $b = 0$ for null hypothesis, and $b = 0.025, 0.05, 0.075, 0.1$ for increasingly several alternatives,
- $\star$ the $(\varepsilon_{i,t})$s are i.i.d. with a symmetric density - Gaussian ($\mathcal{N}$), logistic (l), double exponential ($\mathcal{D}e$), Student (t_5) - or with an asymmetric density - the skew normal $s\mathcal{N}(10)$ and skew Student $st_5(10)$ densities.

In order to examine the finite sample performances of the proposed procedures, we generated 2500 replications independent samples of size $N = n(T-4) = 1000$ from (32). For each replication, we performed the following tests at the asymptotic level $\alpha = 5\%$, the pseudo-Gaussian test based on $Q_{\mathcal{N}}^{\dagger(n)}$ in (18), the van der Waerden test, Wilcoxon test and Laplace test are based on, respectively, $\underset{\sim vdW}{Q}$, $\underset{\sim W}{Q}$ and $\underset{\sim L}{Q}$ in (4.3), as well the rank tests based on Student scores with 5 degrees of freedom and the rank tests with data-driven skew Student scores $st_{\hat{\nu}}(\hat{\delta})$.

A data-driven choice of the reference density adapting, for instance, to f's actual skewness and kurtosis. Hallin & Mehta (2015) propose selecting the reference density f by fitting a skew-t distribution (see Azzalini & Capitanio, 2003) with location zero, scale one, and density

$$f_{\delta,\nu}(z) = 2t_\nu(z)T_{\nu+1}\left(\delta z \left(\frac{\nu+1}{\nu+z^2}\right)^{1/2}\right),$$

where $\delta \in \mathbb{R}$ is a skewness parameter, $\nu > 0$ is the number of degrees of freedom governing the tails, t_ν and $T_{\nu+1}$ are the density distribution and cumulative distribution functions of the Student-t distributions with ν and $\nu + 1$ degrees of freedom, respectively. Estimators $\hat{\delta}$ and $\hat{\nu}$ are obtained from the residuals $Z_{i,t}^{(n)}$ using a maximum likelihood method (namely, maximizing a skew-t likelihood with respect to (δ, ν)). The f-score functions to be used in the testing procedure then are those associated with the skew-t density $f_{\hat{\delta},\hat{\nu}}$. This approach is also related to the theory of efficient (adaptive) estimation. Additionally, these data-driven scores tests as adaptive tests are valid and asymptotically optimal.

Rejection frequencies are reported in Table 2, they amply confirm the excellent overall performances of our rank-based procedure with data-driven scores. It also appears from the skew normal and skew Student simulations that asymmetry significantly improves the superiority of rank tests over the pseudo-Gaussian one.

---

[2] The use of large number of individuals and short period of time is the most common type of data in dynamic panel analysis. It is called *Short panels*, see Lillo & Torrecillas (2018)

TABLE 2: Rejection frequencies (out of 2500 replications), for $b = 0$ (null hypothesis) and various non-zero values of b (alternative hypotheses), with error density g_1 that is Gaussian ($\mathcal{N}$), logistic (l), double exponential ($\mathcal{D}e$), Student (t_5), the skew normal ($s\mathcal{N}(10)$) and skew Student t5 ($st_5(10)$) of the pseudo-Gaussian and rank tests based on van der Waerden, Wilcoxon, Laplace, Student-t_5 and data-driven scores.

Underlying densities g_1	Test	b				
		0	0.025	0.05	0.075	0.1
$\mathcal{N}$ormal	Pseudo Gaussien	0.0480	0.2040	0.6194	0.9522	0.9978
	Van der Waerden	0.0516	0.2142	0.6418	0.9718	0.9950
	Wilcoxon	0.0520	0.2208	0.6266	0.9320	0.9986
	Laplace	0.0506	0.2256	0.5536	0.8596	0.9880
	Student-t_5	0.0540	0.2120	0.6100	0.9236	0.9956
	Data-Driven	0.0428	0.2640	0.5400	0.7760	0.9200
logistique	Pseudo Gaussien	0.0480	0.3552	0.7420	0.9512	0.9908
	Van der Waerden	0.0500	0.2640	0.5884	0.9232	0.9960
	Wilcoxon	0.0516	0.2224	0.7272	0.9612	0.9984
	Laplace	0.0410	0.2628	0.6688	0.9324	0.9952
	Student-t_5	0.0456	0.2360	0.6096	0.9536	0.9972
	Data-Driven	0.0540	0.2400	0.5320	0.8560	0.9910
$\mathcal{D}$ouble exponentiel	Pseudo Gaussien	0.0536	0.3736	0.7316	0.8992	0.9964
	Van der Waerden	0.0492	0.3748	0.6376	0.9476	0.9888
	Wilcoxon	0.0480	0.2860	0.6156	0.9308	1.0000
	Laplace	0.0612	0.4604	0.7596	0.9952	1.0000
	Student-t_5	0.0518	0.3132	0.6248	0.9512	0.9996
	Data-Driven	0.0550	0.3800	0.6400	0.8040	0.9440
Student-t_5	Pseudo Gaussien	0.0440	0.2620	0.7456	0.9748	0.9988
	Van der Waerden	0.0502	0.2544	0.7368	0.9944	1.0000
	Wilcoxon	0.0486	0.3988	0.6996	0.9976	1.0000
	Laplace	0.0552	0.4204	0.7040	0.9964	1.0000
	Student-t_5	0.0528	0.3984	0.7192	0.9988	1.0000
	Data-Driven	0.0480	0.2040	0.5840	0.8280	0.9940
skew $\mathcal{N}$ormal-$s\mathcal{N}(10)$	Pseudo Gaussien	0.0510	0.4150	0.6994	0.9026	0.9915
	Van der Waerden	0.0532	0.3792	0.7092	0.8584	0.9906
	Wilcoxon	0.0518	0.2836	0.6160	0.8402	0.9902
	Laplace	0.0480	0.3228	0.6028	0.8006	0.9892
	Student-t_5	0.0506	0.3138	0.6180	0.8126	0.9880
	Data-Driven	0.0560	0.5200	0.7700	0.9820	0.9990
skew Student-$st_5(10)$	Pseudo Gaussien	0.0512	0.2276	0.6298	0.9064	0.9858
	Van der Waerden	0.0504	0.2096	0.6480	0.8984	1.0000
	Wilcoxon	0.0458	0.2136	0.6148	0.9092	0.9996
	Laplace	0.0468	0.2068	0.5504	0.8404	0.9932
	Student-t_5	0.0500	0.2544	0.6154	0.9128	0.9940
	Data-Driven	0.0476	0.4460	0.7968	0.9998	1.0000

6. Conclusions

In the present article, we derive a pseudo-Gaussian and rank-based tests for testing white noise against panel superdiagonal bilinear dependence in a multiple regression model for specified and unspecified error density. Moreover, the pseudo-Gaussian test appears to have quite poor performances under skewed and heavy-tailed distributions, which leads as to consider rank-based tests. These tests are nonparametric and they have better performance in terms of empirical power for van der Waerden, Wilcoxon, Laplace, Student t and data-driven scores.

Acknowledgements

The authors would like to express their gratitude to the two referees and the Editor for their constructive comments and suggestions which improved the presentation of the paper.

Appendix

Proof of Proposition 1. The proof of Proposition 1 relies on Swensen (1985) conditions 1.2 to 1.7 of Lemma 1, and the only delicate one actually is condition 1.2. The main point consists in showing that

$$(\mu, \beta, \sigma^2, b) \mapsto q^{\frac{1}{2}}_{\mu,\beta,\sigma^2,b;f_1}(y) = \left[\frac{1}{\sigma} f_1\left(\frac{1}{\sigma}\big(y-\mu-\beta' x+\sum_{j=1}^{\infty}(-b)^j u_{kj}\prod_{s=1}^{j} u_{l-(s-1)k}\big)\right)\right]^{\frac{1}{2}}$$

is quadratic mean differentiability, at any $(\mu, \beta, \sigma^2, 0)$, with $x := (x_1, x_2, \ldots, x_K)' \in \mathbb{R}^K$. In order to establish the quadratic mean differentiability of $q^{\frac{1}{2}}_{\mu,\beta,\sigma^2,b;f_1}$ it is sufficient to show that the four conditions (i)-(iv) of Lemma A.1 in Bennala et al. (2012) hold. This last is established in the following Lemma.

Lemma

Let Assumption (B) holds and fix $f_1 \in \mathcal{F}_1$. Define, for $y \in \mathbb{R}$,

$$\begin{aligned}
D_\mu q^{\frac{1}{2}}_{\mu,\beta,\sigma^2,0;f_1}(y) &= \frac{1}{2\sigma} q^{\frac{1}{2}}_{\mu,\beta,\sigma^2,0;f_1}(y)\Phi_{f_1}\big(\tfrac{y-\mu-\beta' x}{\sigma}\big),\\
D_\beta q^{\frac{1}{2}}_{\mu,\beta,\sigma^2,0;f_1}(y) &= \frac{1}{2\sigma} q^{\frac{1}{2}}_{\mu,\beta,\sigma^2,0;f_1}(y)\Phi_{f_1}\big(\tfrac{y-\mu-\beta' x}{\sigma}\big)(\mathbf{K}^{(n)})' x,\\
D_{\sigma^2} q^{\frac{1}{2}}_{\mu,\beta,\sigma^2,0;f_1}(y) &= \frac{1}{4\sigma^2} q^{\frac{1}{2}}_{\mu,\beta,\sigma^2,0;f_1}(y)\left(\tfrac{y-\mu-\beta' x}{\sigma}\Phi_{f_1}\big(\tfrac{y-\mu-\beta' x}{\sigma}\big) - 1\right),
\end{aligned}$$

and

$$D_b q^{\frac{1}{2}}_{\mu,\beta,\sigma^2,b;f_1}(y)_{|b=0} = \frac{1}{2\sigma} q^{\frac{1}{2}}_{\mu,\beta,\sigma^2,0;f_1}(y)\Phi_{f_1}\big(\tfrac{y-\mu-\beta' x}{\sigma}\big) u_k u_l.$$

Then, as τ, s, v and $r \to 0$,

1. $\int_{\mathbb{R}}\left[q^{\frac{1}{2}}_{\mu+\tau,\beta+\mathbf{s},\sigma^2+v,r;f_1}(y) - q^{\frac{1}{2}}_{\mu+\tau,\beta+\mathbf{s},\sigma^2+v,0;f_1}(y) - rD_b q^{\frac{1}{2}}_{\mu+\tau,\beta+\mathbf{s},\sigma^2+v,b;f_1}(y)_{|b=0}\right]^2 dy = o(r^2),$

2. $$\int_{\mathbb{R}}\left[q^{\frac{1}{2}}_{\mu+\tau,\beta+s,\sigma^2+v,0;f_1}(y) - q^{\frac{1}{2}}_{\mu,\beta,\sigma^2,0;f_1}(y) - \begin{pmatrix}\tau\\ s'\\ v\end{pmatrix}'\begin{pmatrix}D_\mu q^{\frac{1}{2}}_{\mu,\beta,\sigma^2,0;f_1}(y)\\ D_\beta q^{\frac{1}{2}}_{\mu,\beta,\sigma^2,0;f_1}(y)\\ D_{\sigma^2} q^{\frac{1}{2}}_{\mu,\beta,\sigma^2,0;f_1}(y)\end{pmatrix}\right]^2 dy = o\left(\left\|\begin{pmatrix}\tau\\ s'\\ v\end{pmatrix}\right\|^2\right),$$

and

3. $$\int_{\mathbb{R}}\left[q^{\frac{1}{2}}_{\mu+\tau,\beta+s,\sigma^2+v,r;f_1}(y) - q^{\frac{1}{2}}_{\mu,\beta,\sigma^2,0;f_1}(y) - \begin{pmatrix}\tau\\ s'\\ v\\ r\end{pmatrix}'\begin{pmatrix}D_\mu q^{\frac{1}{2}}_{\mu,\beta,\sigma^2,0;f_1}(y)\\ D_\beta q^{\frac{1}{2}}_{\mu,\beta,\sigma^2,0;f_1}(y)\\ D_{\sigma^2} q^{\frac{1}{2}}_{\mu,\beta,\sigma^2,0;f_1}(y)\\ D_b q^{\frac{1}{2}}_{\mu,\beta,\sigma^2,b;f_1}(y)|_{b=0}\end{pmatrix}\right]^2 dy = o\left(\left\|\begin{pmatrix}\tau\\ s'\\ v\\ r\end{pmatrix}\right\|^2\right).$$

Proof of Lemma

1. Let $\Upsilon(b) = \sum_{j=1}^{\infty}(-b)^j u_{kj}\prod_{s=1}^{j} u_{l-(s-1)k}$ and $y-(\mu+\tau)-(\beta+s)'x = z(\mu,\beta) := z$. Then (1) takes the form

$$\int_{\mathbb{R}}\left[f^{\frac{1}{2}}\big(z+\Upsilon(r)\big) - f^{\frac{1}{2}}(z) - \frac{r}{2}f^{\frac{1}{2}}(z)\Phi_f(z)u_k u_l\right]^2 dz = o(r^2),$$

is equivalent to

$$\int_{\mathbb{R}} r^2\left[\frac{1}{r}\left(f^{\frac{1}{2}}\big(z+\Upsilon(r)\big) - f^{\frac{1}{2}}(z)\right) - \frac{1}{2}f^{\frac{1}{2}}(z)\Phi_f(z)u_k u_l\right]^2 dz = o(r^2).$$

Then, it is sufficient to prove

$$\lim_{r\to 0}\int_{\mathbb{R}}\left[\frac{1}{r}\left(f^{\frac{1}{2}}\big(z+\Upsilon(r)\big) - f^{\frac{1}{2}}(z)\right) - \frac{1}{2}f^{\frac{1}{2}}(z)\Phi_f(z)u_k u_l\right]^2 dz = 0.$$

We have

$$\begin{aligned}\lim_{r\to 0}\frac{1}{r}\left(f^{\frac{1}{2}}\big(z+\Upsilon(r)\big) - f^{\frac{1}{2}}(z)\right) &= \lim_{r\to 0}\frac{1}{\Upsilon(r)}\left(f^{\frac{1}{2}}\big(z+\Upsilon(r)\big) - f^{\frac{1}{2}}(z)\right)\times\frac{\Upsilon(r)}{r}\\ &= \left(f^{\frac{1}{2}}(z)\right)'\times(-1)u_k u_l\\ &= \frac{1}{2}f^{\frac{1}{2}}(z)\Phi_f(z)u_k u_l,\end{aligned}$$

and just show

$$\int_{\mathbb{R}}\left[\frac{1}{r}\left(f^{\frac{1}{2}}\big(z+\Upsilon(r)\big)-f^{\frac{1}{2}}(z)\right)\right]^2 dz \leqslant \int_{\mathbb{R}}\left[\frac{1}{2}f^{\frac{1}{2}}(z)\Phi_f(z)u_k u_l\right]^2 dz<\infty.$$

We know that

$$f^{\frac{1}{2}}\big(z+\Upsilon(r)\big)-f^{\frac{1}{2}}(z)=\int_z^{z+\Upsilon(r)}\left(f^{\frac{1}{2}}(v)\right)' dv,$$

then

$$\begin{aligned}\int_{z\in\mathbb{R}}\left[\frac{1}{r}\left(f^{\frac{1}{2}}\big(z+\Upsilon(r)\big)-f^{\frac{1}{2}}(z)\right)\right]^2 dz &= \int_{z\in\mathbb{R}}\frac{1}{r^2}\left[\int_{v=z}^{z+\Upsilon(r)}\left(f^{\frac{1}{2}}(v)\right)' dv\right]^2 dz\\ &\leqslant \frac{\Upsilon(r)}{r^2}\int_{z\in\mathbb{R}}\int_{v=z}^{z+\Upsilon(r)}\left[\frac{1}{2}f^{\frac{1}{2}}(v)\Phi_f(v)\right]^2 dv\, dz\\ &\leqslant \frac{\Upsilon(r)}{r^2}\int_{v\in\mathbb{R}}\int_{z-\Upsilon(r)}^{z}\left[\frac{1}{2}f^{\frac{1}{2}}(v)\Phi_f(v)\right]^2 dz\, dv\\ &\leqslant \left[\frac{\Upsilon(r)}{r}\right]^2\int_{v\in\mathbb{R}}\left[\frac{1}{2}f^{\frac{1}{2}}(v)\Phi_f(v)\right]^2 dv\\ &\leqslant (-u_k u_l)^2\int_{v\in\mathbb{R}}\left[\frac{1}{2}f^{\frac{1}{2}}(v)\Phi_f(v)\right]^2 dv\\ &\leqslant \int_{\mathbb{R}}\left[\frac{1}{2}f^{\frac{1}{2}}(z))\Phi_f(z))u_k u_l\right]^2 dz.\end{aligned}$$

This completes the proof of part (1) of the Lemma.

2. The problem here reduces to the classical case of linear models considered by Swensen (1985).
3. The result here follows from (1) and (2) above. This completes the proof of Lemma. □

Lemma above and Lemma A.1 in Bennala et al. (2012) jointly imply the desired mean square differentiability property. The proof of Proposition 1 is thus complete.

Proof of Proposition 2 (*i*) Follows from Proposition 1 and the fact that

$$\Lambda^{(n)}_{\theta+\nu^{(n)}\tau^{(n)}/\theta;f_1}+\Lambda^{(n)}_{\theta/\theta+\nu^{(n)}\tau^{(n)};f_1}=o_P(1),$$

under $\mathbf{P}^{(n)}_{\mu,\beta',\sigma^2,0;f_1}$, as $n\to\infty$.

(*ii*) From (12) and by algebra calculations we obtain under $\mathbf{P}^{(n)}_{\mu,\beta',\sigma^2,0;f_1}$, as $n\to\infty$,

$$\Delta^{(n)}_{f_1;4}(\theta_n)-\Delta^{(n)}_{f_1;4}(\theta)=o_P(1).$$

Then, we can replace the deterministic sequence θ_n by the sequence of estimates $\widehat{\theta}_n$, so we have the result. $\square$

Proof of Proposition 3 (i) Letting $\Gamma_{f_1}(.)$ has been assumed continuous; $\Gamma_{f_1;44}(\widehat{\theta}_n) - \Gamma_{f_1;44}(\theta) = o_P(1)$. Then, Proposition 2 (ii) implies that, under $\mathbf{P}^{(n)}_{\theta;f_1}$, as $n \to \infty$,

$$\begin{aligned}
Q^{(n)}_{f_1}(\widehat{\theta}_n) - Q^{(n)}_{f_1}(\theta) &= \frac{\Delta^{(n)^2}_{f_1;4}(\widehat{\theta}_n)}{\Gamma_{f_1;44}(\widehat{\theta}_n)} - \frac{\Delta^{(n)^2}_{f_1;4}(\theta)}{\Gamma_{f_1;44}(\theta)} \\
&= \frac{\Delta^{(n)^2}_{f_1;4}(\widehat{\theta}_n)}{\Gamma_{f_1;44}(\widehat{\theta}_n)} - \frac{\Delta^{(n)^2}_{f_1;4}(\theta)}{\Gamma_{f_1;44}(\widehat{\theta}_n)} + \frac{\Delta^{(n)^2}_{f_1;4}(\theta)}{\Gamma_{f_1;44}(\widehat{\theta}_n)} - \frac{\Delta^{(n)^2}_{f_1;4}(\theta)}{\Gamma_{f_1;44}(\theta)} \\
&= \frac{1}{\Gamma_{f_1;44}(\widehat{\theta}_n)}\left(\Delta^{(n)^2}_{f_1;4}(\widehat{\theta}_n) - \Delta^{(n)^2}_{f_1;4}(\theta)\right) \\
&\qquad + \Delta^{(n)^2}_{f_1;4}(\theta)\left(\frac{1}{\Gamma_{f_1;44}(\widehat{\theta}_n)} - \frac{1}{\Gamma_{f_1;44}(\theta)}\right) \\
&= \frac{1}{\Gamma_{f_1;44}(\widehat{\theta}_n)}\left(\Delta^{(n)}_{f_1;4}(\widehat{\theta}_n) - \Delta^{(n)}_{f_1;4}(\theta)\right)\left(\Delta^{(n)}_{f_1;4}(\widehat{\theta}_n) + \Delta^{(n)}_{f_1;4}(\theta)\right) + o_P(1) \\
&= o_P(1).
\end{aligned}$$

From (10), we have $\Delta^{(n)}_{f_1;4}(\theta) \xrightarrow[n\to\infty]{\mathcal{D}} \mathcal{N}\big(\Gamma_{f_1;44}(\theta)\tau_4, \Gamma_{f_1;44}(\theta)\big)$, under $\mathbf{P}^{(n)}_{\theta+\nu^{(n)}\tau^{(n)};f_1}$. So that

$$\frac{\Delta^{(n)}_{f_1;4}(\theta)}{\sqrt{\Gamma_{f_1;44}(\theta)}} \xrightarrow[n\to\infty]{\mathcal{D}} \mathcal{N}(\Gamma^{1/2}_{f_1;44}(\theta)\tau_4, 1).$$

Cochran's Theorem leads to

$$\Delta^{(n)^2}_{f_1;4}(\theta)/\Gamma_{f_1;44}(\theta) = Q^{(n)}_{f_1}(\theta) \xrightarrow[n\to\infty]{\mathcal{D}} \chi^2_1(\lambda_{f_1});$$

with $\lambda_{f_1} = \left(\Gamma^{1/2}_{f_1;44}(\theta)\tau_4\right)^2 = \tau_4^2\sigma^2(T-l)I(f_1)\sigma^4_{f_1}$, which gives the desired result under $\mathbf{P}^{(n)}_{\theta+\nu^{(n)}\tau^{(n)};f_1}$.

(ii) Stringency is a consequence of the weak convergence of local experiments to Gaussian shifts (see Le Cam, 1986).

(iii) Follows from (i) and (ii). $\square$

Proof of Proposition 5. The proof of Part (i) of the proposition use the *Hájek's projection theorem* (see, Hájek & Šidák, 1967) and follows along the same lines as in Hallin & Mélard (1988), therefore it is omitted. Part (ii) is obtained by direct computation. As for Part (iii), under $\mathbf{P}^{(n)}_{\theta;g_1}$, the result straightforwardly follows from classical central limit theorems. On the other hand, it is easy to see that, still under $\mathbf{P}^{(n)}_{\theta+\nu^{(n)}\tau;g_1}$, $\underline{\Delta}^{(n)}_{f_1,g_1;4}(\theta)$ and the log-likelihood $\Lambda^{(n)}_{\theta+\nu^{(n)}\tau/\theta;g_1}$ are jointly multinormal. Then, the desired result follows from an application of Le Cam's third lemma. □

[Recibido: noviembre de 2019 — Aceptado: abril de 2020]

References

Akharif, A. & Hallin, M. (2003), 'Efficient detection of random coefficients in autoregressive models', *Annals of Statistics* **31**, 675–704.

Allal, J. & El Melhaoui, S. (2006), 'Tests de rangs adaptatifs pour les modèles de régression linéaire avec erreurs arma', *Annales des sciences mathématiques du Québec* **30**, 29–54.

Azzalini, A. & Capitanio, A. (2003), 'Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t-distribution', *Journal of the Royal Statistical Society, Series B* **65**, 367–389.

Baltagi, B. & Li, Q. (1995), 'Testing AR(1) against MA(1) disturbances in an error component model', *Journal of Econometrics* **68**, 133–151.

Benghabrit, Y. & Hallin, M. (1992), 'Optimal rank-based tests against first-order superdiagonal bilinear dependence', *Journal of Statistical Planning and Inference* **32**(1), 45–61.

Benghabrit, Y. & Hallin, M. (1996), 'Rank-based tests for autoregressive against bilinear serial dependence', *Journal of Nonparametric Statistics* **6**, 253–272.

Bennala, N., Hallin, M. & Paindaveine, D. (2012), 'Pseudo-gaussian and rank-based optimal tests for random individual effects in large n small t panels', *Journal of Econometrics* **170**(1), 50–67.

Cassart, D., Hallin, M. & Paindaveine, D. (2011), 'A class of optimal tests for symmetry based on local Edgeworth approximations', *Bernoulli* **17**, 1063–1094.

Chernoff, H. & Savage, I. R. (1958), 'Asymptotic normality and efficiency of certain nonparametric tests', *Annals of Mathematical Statistics* **29**, 972–994.

Dutta, H. (1999), 'Large sample tests for a regression model with autoregressive conditional heteroscedastic errors', *Communications in StatisticsTheory and Methods* **A28**, 105–117.

Elmezouar, Z. C., Kadi, A. M. & Gabr, M. M. (2012), 'Linear regression with bilinear time series errors', *Panamerican Mathematical Journal* **22**(1), 1–13.

Fihri, M., Akharif, A., Mellouk, A. & Hallin, M. (2020), 'Efficient pseudo-gaussian and rank-based detection of random regression coefficients', *Journal of Nonparametric Statistics* **32**(2), 367–402.

Grahn, T. (1995), 'A conditional least squares approach to bilinear time series', *Journal of Time Series Analysis* **16**, 509–529.

Granger, C. W. J. & Andersen, A. P. (1978), *An Introduction to Bilinear Time Series Models*, Vandenhoeck and Ruprecht, Göttingen.

Guegan, D. (1981), 'Etude d'un modèle non linéaire, le modèle superdiagonal d'ordre 1', *CRAS Série I,* **293**, 95–98.

Guegan, D. & Pham, D. T. (1992), 'Power of the score test against bilinear time series models', *Statistica Sinica* **2**, 157–170.

Hájek, J. & Šidák, Z. (1967), *Theory of Rank Tests*, Academic Press, New York.

Hallin, M. & Mehta, C. (2015), 'R-estimation for asymmetric independent component analysis', *Journal of the American Statistical Association* **110**(6), 218–232.

Hallin, M. & Mélard, G. (1988), 'Rank-based tests for randomness against first-order serial dependence', *Journal of the American Statistical Association* **83**(404), 1117–1128.

Hallin, M., Taniguchi, M., Serroukh, A. & Choy, K. (1999), 'Local asymptotic normality for regression models with long-memory disturbance', *The Annals of Statistics* **27**(6), 2054–2080.

Hallin, M. & Werker, B. J. M. (2003), 'Semi-parametric efficiency, distribution-freeness and invariance', *Bernoulli* **9**(1), 137–165.

Hristova, D. (2005), 'Maximum likelihood estimation of a unit root bilinear model with an application to prices', *Studies in Nonlinear Dynamics & Econometrics* **9**(1).

Hwang, S. Y. & Basawa, I. V. (1993), 'Parameter estimation in a regression model with random coefficient autoregressive errors', *Journal of Statistical Planning and Inference* **36**, 57–67.

Kim, I. (2014), 'A study on the test of homogeneity for nonlinear time series panel data using bilinear models', *Journal of Digital Convergence* **12**(7), 261–266.

Kreiss, J.-P. (1987), 'On adaptive estimation in stationary ARMA processes', *The Annals of Statistics* **15**(1), 112–133.

Le Cam, L. M. (1986), *Asymptotic Methods in Statistical Decision Theory*, Springer-Verlag, New York.

Le Cam, L. M. & Yang, G. L. (2000), *Asymptotics in Statistics: Some Basic Concepts*, 2 edn, Springer-Verlag, New York.

Lee, S. H., Kim, S. W., & Lee, S. D. (2013), 'Test of homogeneity for panel bilinear time series model', *The Korean Journal of Applied Statistics* **6**(3), 521–529.

Lillo, R. L. & Torrecillas, C. (2018), 'Estimating dynamic panel data. A practical approach to perform long panels', *Revista Colombiana de Estadística* **41**(1), 31–52.

Maravall, A. (1983), 'An application of nonlinear time series forecasting', *Journal of Business & Economic Statistics* **1**(1), 66–74.

Noether, G. E. (1949), 'On a theorem by wald and wolfowitz', *The Annals of Mathematical Statistics* **20**(3), 455–458.

Pesaran, H. (2015), *Time Series and Panel Data Econometrics*, Oxford University Press, Oxford,UK.

Pham, T. D. & Tran, L. T. (1981), 'On the first-order bilinear time series model', *Journal of Applied Probability* **18**, 617–627.

Quinn, B. G. (1982), 'Stationarity and invertibility of simple bilinear models', *Stochastic Processes and their Applications* **12**, 225–230.

Rao, T. S. & Gabr, M. (1984), *An introduction to bispectral analysis and bilinear time series models*, Springer Science & Business Media.

Saikkonen, P. & Luukkonen, R. (1991), 'Power properties of a time series linearity test against some simple bilinear alternatives', *Statistica Sinica* **1**, 453–464.

Swensen, A. R. (1985), 'The asymptotic distribution of the likelihood ratio for autoregressive time series with a regression trend', *Journal of Multivariate Analysis* **16**, 54–70.

Tan, L. & Wang, L. (2015), 'The lasso method for bilinear time series models', *Communications in Statistics - Simulation and Computation* **45**, 1–11.

Wald, A. (1943), 'Tests of statistical hypotheses concerning several parameters when the number of observations is large', *Transactions of the American Mathematical Society* **54**, 426–482.

Weiss, A. A. (1986), 'ARCH and bilinear time series models: Comparison and combination', *Journal of Business and Economic Statistics* **4**, 59–70.