Original Article

Journal of Time Series Analysis

First version received September 2012

Published online in Wiley Online Library: 27 March 2013

(wileyonlinelibrary.com) DOI: 10.1111/jtsa.12026

Distributions for residual autocovariances in parsimonious periodic vector autoregressive models with applications

Pierre Duchesne[a,*,†] and Pierre Lafaye de Micheaux[a]

The asymptotic distribution of the residual autocovariance matrices in the class of periodic vector autoregressive time series models with structured parameterization is derived. Diagnostic checking with portmanteau test statistics represents a useful application of the result. Under the assumption that the periodic white noise process of the periodic vector autoregressive time series model is composed of independent random variables, we demonstrate that the finite sample distributions of the Hosking-Li-McLeod portmanteau test statistics can be approximated by those of weighted sums of independent chi-square random variables. The quantiles of the asymptotic distribution can be computed using the Imhof algorithm or other exact methods. Thus, using the (single) chi-square distribution for these test statistics appears inadequate in general, although it is often recommended in practice for diagnostic methods of that kind. A simulation study provides empirical evidence.

Keywords: Asymptotic distributions; Residual autocovariances; Parsimonious models; Periodic time series models; Portmanteau test statistics; Time series.

1. INTRODUCTION

The class of periodic time series models is useful to analyze time series data with periodic structure; see Holan *et al*. (2010) for a recent survey. However, a serious limitation of these models lies in the large number of parameters that need to be estimated. For example, a univariate model with daily data coming from climate applications may involve 365 free parameters. This is discussed in Lund *et al*. (2006). The problem is even more severe in the multivariate case, even for the very useful class of periodic vector autoregressive (PVAR) models. In a PVAR model, the maximal number of free parameters is given by $d^2\sum_{\nu=1}^{s} p(\nu)$, where d is the dimension of the process, s is the number of seasons and $p(\nu)$ are the autoregressive orders for each season. In the bivariate case, a first-order model (that is, $p(\nu) \equiv 1$) for quarterly, monthly and daily data would have 16, 48 and 1460 free parameters respectively. In fact, applications in the latter context have been limited to quarterly data, because for a larger number of seasons the number of parameters is too large. See the discussions in Franses and Paap (2004) and Lütkepohl (2005). To achieve parsimony, PVAR models with structured parameterization may be considered. By structured parameterization, we mean that the $d \times d$ autoregressive matrices are a function of a certain vector $\boldsymbol{\beta}$, whose dimension is often much smaller than $d^2\sum_{\nu=1}^{s} p(\nu)$. Thus, a structured parameterization can be interpreted as a transformation reducing the dimension of the parameter space. For example, by using Fourier representations for the coefficients, it is possible to model slow seasonal changes in the parameters. One argument in favour of such an approach is that in many circumstances, similar seasons should have relatively similar parameters. In modelling a univariate time series on monthly ozone data, Bloomfield *et al*. (1994) conclude that the twelve parameters in a periodic first-order autoregressive model can be reduced to three parameters, and the parameters between the seasons were allowed to change smoothly. See also Lund *et al*. (2006) for other examples of parameterization in univariate periodic models. In vector autoregressive models (VAR) with structured parameterization; see Ahn (1988) for general properties.

Diagnostic checking appears to be an important step in any statistical modelling application. In time series analysis, much effort has been devoted to obtain the asymptotic distribution of residual autocovariances under various conditions. See Li (2004) for a review, among others. In vector autoregressive-moving average models, the Hosking-Li-McLeod portmanteau test statistics are known to

[a]*Université de Montréal*

**Correspondence to: Pierre Duchesne, Université de Montréal; Département de mathématiques et de statistique; C.P. 6128 Succursale Centre-Ville; Montréal, Québec H3C 3J7; Canada.*

[†]*E-mail: duchesne@dms.umontreal.ca*

have approximate chi-square distributions. They have been introduced and studied in Hosking (1980) and Li and McLeod (1981). See also the monograph of Li (2004). When vector autoregressive-moving average time series models with structured parameterization are fitted, assuming that the white noise term is composed of independent random variables, the approximate chi-square distributions hold under the null hypothesis of correct specification, and the degrees of freedom simply need to be adjusted. See Ahn (1988) and Reinsel (1997). Under VAR and structured parameterized VAR models with uncorrelated white noise, the asymptotic distributions of the portmanteau test statistics are better approximated by those of weighted sums of chi-square variables. See Francq and Raïssi (2007) and Boubacar Mainassara (2011).

Ursu and Duchesne (2009) obtained the asymptotic distribution of the residual autocovariances in PVAR models with linear constraints on the model parameters. In their asymptotic result, the autoregressive parameters inside the seasons were allowed to satisfy constraints. For example, in a first-order PVAR model, for a known $d^2 \times K(v)$ matrix $\mathbf{R}(v)$ of rank $K(v)$, and a known $d^2 \times 1$ vector $\mathbf{b}(v)$, the autoregressive parameters $\Phi(v)$ were supposed to satisfy the relations $\phi(v) = \mathbf{R}(v)\beta(v) + \mathbf{b}(v)$, where $\phi(v) = \text{vec}\{\Phi(v)\}$, with vec$\{\cdot\}$ representing the operator that stacks the columns of a matrix into a vector. See Harville (1997) for the properties of the vec operator. Thus, the autoregressive parameters were not allowed to satisfy constraints between the seasons. In that framework, they justified the approximate chi-square distributions of the Hosking-Li-McLeod portmanteau test statistics when the periodic white noise is composed of independent random variables.

In this article, the asymptotic distributions of the residual autocovariances matrices in the class of PVAR models with structured parameterization are derived. Applications of the result include diagnostic checking with portmanteau test statistics. Quite surprisingly, the chi-square approximations appear inadequate in general, even under the assumption of an independent (periodic) white noise process. In fact, the exact distributions of the Hosking-Li-McLeod portmanteau test statistics are better approximated by those of weighted sums of chi-square random variables. In diagnostic checking VAR models, the asymptotic distribution of the (normalized) residual autocovariances under the null hypothesis is approximately idempotent, meaning that the weights in the asymptotic distributions of the classical portmanteau test statistics are approximately composed of ones and zeros. Our results show that in the general PVAR case, the weights in the asymptotic distributions of the Hosking-Li-McLeod procedures may be relatively far from zero, and thus, adjusting the degrees of freedom does not represent a solution in the present framework. In practical applications, it is thus preferable to compute the quantiles of the asymptotic distribution using Imhof's (1961) algorithm or other exact methods. See Duchesne and Lafaye de Micheaux (2010) for a recent account.

Our results are particularly useful in applications, given that the applicability of periodic models is limited due to the large number of parameters in these models. Here, the structured parameterization in PVAR models allows us to fit these models with a large number of seasons relatively easily, and diagnostic checking can be considered with the popular Hosking-Li-McLeod portmanteau test statistics under general assumptions on the model parameters.

The article is organized as follows. In Section 2, some notations and preliminaries are provided. Section 3 presents our main results. In Section 4, applications for diagnostic checking are given and in particular the asymptotic distributions of the portmanteau test statistics are derived. A simulation study is conducted in Section 5.

2. PRELIMINARIES

The PVAR model for a d-dimensional time series $\mathbf{Y}_t = (Y_t(1), \ldots, Y_t(d))^\top$ can be written as:

$$\mathbf{Y}_{ns+v} = \sum_{k=1}^{p(v)} \Phi_k(v;\beta_0)\mathbf{Y}_{ns+v-k} + \varepsilon_{ns+v}, \tag{1}$$

for season $v \in \{1, \ldots, s\}$, s being a predetermined value, at year n (say). The autoregressive model order at season v is noted $p(v)$, and $\Phi_k(v;\beta) = (\Phi_{k,ij}(v;\beta))_{i,j=1,\ldots,d}$, $k = 1, \ldots, p(v)$, are the matricial autoregressive model coefficients during season v. The components $\Phi_{k,ij}(v;\beta)$ are assumed to be twice continuously differentiable functions in an open neighbourhood of the $b \times 1$ vector of parameters β_0, where $b \le d^2 \sum_{v=1}^{s} p(v)$. The error process $\{\varepsilon_t, t \in \mathbb{Z}\}$ corresponds to a zero mean Gaussian periodic white noise, that is the $d \times 1$ random vectors $\varepsilon_t = (\varepsilon_t(1), \ldots, \varepsilon_t(d))^\top$ are independent and satisfy $E(\varepsilon_t) = 0$ and $E\left(\varepsilon_{ns+v}\varepsilon_{ns+v}^\top\right) = \Sigma_\varepsilon(v)$, where the error covariance matrix $\Sigma_\varepsilon(v)$ is defined periodically, that is $\Sigma_\varepsilon(v) = \Sigma_\varepsilon(ns+v)$; thus, $\Sigma_\varepsilon(v)$ depends on v but not on n and s. The error covariance matrices are assumed to be non-singular for all $v \in \{1, \ldots, s\}$. Model (1) is presumed to be stationary in the periodic sense. Periodic stationarity is discussed in Gladyšhev (1961). The seasonal autocovariance function of $\{\mathbf{Y}_t, t \in \mathbb{Z}\}$ is $\Gamma_Y(h;v) = \text{cov}(\mathbf{Y}_{ns+v}, \mathbf{Y}_{ns+v-h})$, which may depend on both lag h and season v but not on n. Periodic stationarity conditions are discussed in Lütkepohl (2005). For example, the condition for periodic stationarity in the special case $p(v) \equiv 1$ is that all eigenvalues of $\prod_{j=0}^{s-1} \Phi_1(s-j;\beta_0)$ are strictly smaller than one in modulus. Periodic stationarity implies that the process $\{\mathbf{Y}_t, t \in \mathbb{Z}\}$ has the following moving-average representation:

$$\mathbf{Y}_{ns+v} = \sum_{k=0}^{\infty} \Psi_k(v)\varepsilon_{ns+v-k}, \tag{2}$$

where $\Psi_0(v) = \mathbf{I}_d$. The seasonal matricial weights $\Psi_k(v)$ are supposed to satisfy the summability condition $\sum_{k=0}^{\infty} \|\Psi_k(v)\| < \infty$, where $\|\cdot\|$ denotes the Euclidian norm of a matrix. Using the expression (2), the autocovariance structure of the process satisfies:

$$\Gamma_Y(h;v) = \sum_{k=0}^{\infty} \Psi_{k+h}(v)\Sigma_\varepsilon(v-k-h)\Psi_k^\top(v-h), \tag{3}$$

where the covariance matrix $\Sigma_\varepsilon(v)$ is interpreted periodically in v with period s. See also Lütkepohl (2005) and Ursu and Duchesne (2009). The variance of the process is obtained by setting $h=0$ in expression (3). These expressions show that the autocovariances depend not only on h but also on v.

To fix ideas, consider the case where $p \equiv p(v)$. We denote this situation a PVAR (p) model. As for VAR stochastic processes, the autocovariance function of a PVAR (p) process can be calculated recursively:

$$\Gamma_Y(h;v) = \Phi_1(v)\Gamma_Y(h-1;v-1) + \Phi_2(v)\Gamma_Y(h-2;v-2) + \ldots + \Phi_p(v)\Gamma_Y(h-p;v-p). \tag{4}$$

Once the theoretical autocovariances $\Gamma_Y(h;v)$ are determined for $0 \le h \le p$ and seasons $v = 1, 2, \ldots, s$, the autocovariances for lags $h > p$ can be uniquely solved using the recursive relation (4). The expressions (4) correspond to periodic versions of the Yule–Walker equations.

For any $\boldsymbol{\beta}$, we introduce the quantities:

$$\mathbf{r}_{ns+v}(\beta) = \begin{cases} \mathbf{Y}_{ns+v} - \sum_{k=1}^{p(v)} \Phi_k(v;\beta)\mathbf{Y}_{ns+v-k}, & ns+v > p(v), \\ 0, & ns+v \le p(v), \end{cases} \tag{5}$$

which are well-defined for $n = 0, 1, \ldots, N-1$, N being the number of observed years (for simplicity, it is assumed that N complete years are observed; a similar assumption is made in Lund *et al.* (2006)). Applying the vec operator on both sides of (1), and using the rule $\mathrm{vec}(\mathbf{ABC}) = (\mathbf{C}^\top \otimes \mathbf{A})\mathrm{vec}(\mathbf{B})$, model (1) can be written alternatively as:

$$\varepsilon_{ns+v} = \mathbf{Y}_{ns+v} - \sum_{k=1}^{p(v)} \left(\mathbf{Y}_{ns+v-k}^\top \otimes \mathbf{I}_d\right)\phi_k(v;\beta),$$

where '$\otimes$' denotes the Kronecker product and $\phi_k(v;\beta) = \mathrm{vec}\{\Phi_k(v;\beta)\}$ corresponds to a $d^2 \times 1$ vector of parameters. See Harville (1997) for properties of the Kronecker product and of the vec operator. Let M be a maximal lag order, satisfying $1 \le M < N$, which is fixed with respect to the sample size Ns. We introduce the sample autocovariance matrices $\mathbf{C}_\mathbf{r}(h;v) = (C_{\mathbf{r},ij}(h;v))_{i,j=1,\ldots,d}$:

$$\mathbf{C}_\mathbf{r}(h;v) = \begin{cases} N^{-1}\sum_{n=h}^{N-1} \mathbf{r}_{ns+v}(\beta)\mathbf{r}_{ns+v-h}^\top(\beta), & h \ge 0, \\ \mathbf{C}_\mathbf{r}^\top(-h;v-h), & h < 0. \end{cases}$$

Similarly, the sample autocovariances are collected in a $(d^2M) \times 1$ vector:

$$\mathbf{c}_\mathbf{r}(v) = \left(\mathbf{c}_\mathbf{r}^\top(1;v), \ldots, \mathbf{c}_\mathbf{r}^\top(M;v)\right)^\top, \tag{6}$$

where $\mathbf{c}_\mathbf{r}(h;v) = \mathrm{vec}\{\mathbf{C}_\mathbf{r}(h;v)\}$. We also define similarly the sample autocovariances $\mathbf{C}_\varepsilon(h;v)$ and $\mathbf{c}_\varepsilon(v)$ on the basis of the unobservable error process $\{\varepsilon_t, t \in \mathbb{Z}\}$. Note that unless $\beta = \beta_0$, $\{\mathbf{r}_t(\beta), t \in \mathbb{Z}\}$ needs not be a white noise process.

We consider conditional maximum likelihood estimators for estimating β_0. That estimation method has a long history in time series analysis. See Tunnicliffe Wilson (1973), Anderson (1980), Hannan and Kavalieris (1984), Poskitt and Salau (1995), Reinsel (1997) and Lütkepohl (2005), among others. That approach allows us to derive explicit evaluation for the gradient and an asymptotic expression of the Hessian of the conditional log-likelihood function. Under certain general assumptions, conditional and exact maximum likelihood estimators are asymptotically equivalent. The conditional maximum likelihood criterion is often used because it is more easily tractable that the exact likelihood. When the results rely on an asymptotic theory, such as those presented in Sections 3 and 4, using conditional or exact likelihood estimators delivers the same conclusions under general assumptions, at least asymptotically. See Reinsel (1997, Chapter 5). Let $\hat{\beta}_N$ be the conditional Gaussian estimator of β_0 on the basis of the observed time series data $\mathbf{Y}_{ns+v}$, $n = 0, \ldots, N-1$ and $v = 1, \ldots, s$. Ignoring the normalizing factor involving $(2\pi)^{-(dNs)/2}$, $\hat{\beta}_N$ and $\hat{\Sigma}_\varepsilon(v)$, $v = 1, \ldots, s$ maximize the conditional Gaussian log-likelihood function:

$$\mathcal{L}\{\beta;\Sigma_\varepsilon(1),\ldots,\Sigma_\varepsilon(s)\} = -\frac{N}{2}\sum_{v=1}^{s} \log\det\{\Sigma_\varepsilon(v)\} - \frac{1}{2}\sum_{v=1}^{s}\sum_{n=0}^{N-1} \mathbf{r}_{ns+v}^\top(\beta)\Sigma_\varepsilon^{-1}(v)\mathbf{r}_{ns+v}(\beta).$$

When no confusion is possible, we let $\mathcal{L}\{\beta;\Sigma_\varepsilon(1),\ldots,\Sigma_\varepsilon(s)\} \equiv \mathcal{L}(\beta)$. To establish an asymptotic theory for $\hat{\beta}_N$, we adapt the arguments of Ahn and Reinsel (1988, p. 853). Initially, we study the consistency of the estimator in the framework of periodic models. In the univariate case, Lund *et al.*(2006, Theorem 1) state the asymptotic normal distribution of the maximum likelihood estimator and outline the proof of the result. In the multivariate case, first note that any periodic model can be expressed as a vector autoregressive model (see Franses and Paap (2004, Section 3.1)). A methodology using that companion representation is developed in Pagano (1978). It is often advisable to work directly with the periodic linear difference equations, see Basawa and Lund (2001, p. 654). Nevertheless, model (1) allows a multivariate autoregressive representation with autoregressive parameters twice continuously differentiable functions of $\boldsymbol{\beta}$. Because the resulting model is a multivariate autoregressive model with structured parameterization, the results of

Hannan and Deistler (1988) and Boubacar Mainassara and Francq (2011) can be used, and asymptotically efficient estimators of β_0 can thus be found. In particular $\hat{\beta}_N - \beta_0 = \mathbf{O}_P(N^{-1/2})$.

Using the differentiation rule $\partial\{\mathbf{a}^\top(\beta)\Omega\mathbf{a}(\beta)\}/\partial\beta^\top = 2\mathbf{a}^\top(\beta)\Omega\{\partial\mathbf{a}(\beta)/\partial\beta^\top\}$, where Ω is a symmetric matrix [see, e.g. Lutkepohl (2005, p. 667)], the derivative of the log-likelihood with respect to $\boldsymbol{\beta}$ is given by:

$$\frac{\partial\mathcal{L}(\beta)}{\partial\beta} = -\sum_{v=1}^{s}\sum_{n=0}^{N-1}\frac{\partial\mathbf{r}_{ns+v}^\top(\beta)}{\partial\beta}\boldsymbol{\Sigma}_\varepsilon^{-1}(v)\mathbf{r}_{ns+v}(\beta). \tag{7}$$

When $ns+v > p(v)$, it is not difficult to show the following formula:

$$\frac{\partial\mathbf{r}_{ns+v}(\beta_0)}{\partial\beta^\top} = -\sum_{k=1}^{p(v)}\left(\mathbf{Y}_{ns+v-k}^\top\otimes\mathbf{I}_d\right)\frac{\partial\phi_k(v;\beta_0)}{\partial\beta^\top}. \tag{8}$$

Note that we used the slight abuse of notation $\partial\mathbf{r}_{ns+v}(\beta_0)/\partial\beta^\top = \left\{\partial\mathbf{r}_{ns+v}(\beta)/\partial\beta^\top\right\}|_{\beta=\beta_0}$ and similarly for $\partial\phi_k(v;\beta_0)/\partial\beta^\top$. Using a martingale central limit theorem, it is possible to show that:

$$N^{-1/2}\frac{\partial\mathcal{L}(\beta_0)}{\partial\beta} = -N^{-1/2}\sum_{n=0}^{N-1}\sum_{v=1}^{s}\frac{\partial\mathbf{r}_{ns+v}^\top(\beta_0)}{\partial\beta}\boldsymbol{\Sigma}_\varepsilon^{-1}(v)\mathbf{r}_{ns+v}(\beta_0)\xrightarrow{\mathcal{L}}\mathcal{N}_b(0,\mathcal{I}(\beta_0)),$$

where $\partial\mathcal{L}(\beta_0)/\partial\beta = \{\partial\mathcal{L}(\beta)/\partial\beta\}|_{\beta=\beta_0}$ and $\mathcal{I}(\beta_0)$ denotes the information matrix of β_0. See Reinsel (1997, Section 5.1.4), Basawa and Lund (2001) and Lund *et al.* (2006) for similar arguments. Adopting arguments similar to Ahn (1988) and Ahn and Reinsel (1988), it is possible to show that the information matrix satisfies:

$$\begin{aligned}\mathcal{I}(\beta_0) &= \lim_{N\to\infty} -\frac{1}{N}E\left(\frac{\partial^2\mathcal{L}(\beta_0)}{\partial\beta\partial\beta^\top}\right),\\ &= \lim_{N\to\infty} N^{-1}\sum_{v=1}^{s}\sum_{n=0}^{N-1}E\left\{\frac{\partial\mathbf{r}_{ns+v}^\top(\beta_0)}{\partial\beta}\boldsymbol{\Sigma}_\varepsilon^{-1}(v)\frac{\partial\mathbf{r}_{ns+v}(\beta_0)}{\partial\beta^\top}\right\},\\ &= \sum_{v=1}^{s}\sum_{i=1}^{p(v)}\sum_{j=1}^{p(v)}\frac{\partial\phi_i^\top(v;\beta_0)}{\partial\beta}\left\{\Gamma_Y(j-i;v-i)\otimes\boldsymbol{\Sigma}_\varepsilon^{-1}(v)\right\}\frac{\partial\phi_j(v;\beta_0)}{\partial\beta^\top}.\end{aligned} \tag{9}$$

Note that the previous result generalizes the univariate result of Lund *et al.*(2006, p. 39) to the vector case. To show expression (9) we applied the rule stating that for any $m\times n$ matrix $\mathbf{A}$, $p\times q$ matrix $\mathbf{B}$, $n\times u$ matrix $\mathbf{C}$ and $q\times v$ matrix $\mathbf{D}$, $(\mathbf{A}\otimes\mathbf{B})(\mathbf{C}\otimes\mathbf{D}) = (\mathbf{AC})\otimes(\mathbf{BD})$ (see, e.g. Harville (1997, p. 337)). The second line of the previous calculation also used (7) and (8). Given that for any consistent sequence such that $\hat{\beta}_N^* \to^P \beta_0$,

$$-\left\{N^{-1}\frac{\partial^2\mathcal{L}(\beta)}{\partial\beta\partial\beta^\top}|_{\beta=\beta_N^*}\right\} = -N^{-1}\frac{\partial^2\mathcal{L}\left(\hat{\beta}_N^*\right)}{\partial\beta\partial\beta^\top} = \mathcal{I}(\beta_0) + \mathbf{o}_P(1),$$

a first-order Taylor expansion of the conditional maximum likelihood estimator is:

$$\begin{aligned}N^{1/2}\left(\hat{\beta}_N - \beta_0\right) &= -\left\{N^{-1}\frac{\partial\mathcal{L}\left(\hat{\beta}_N^*\right)}{\partial\beta\partial\beta^\top}\right\}^{-1}\left\{N^{-1/2}\frac{\partial\mathcal{L}(\beta_0)}{\partial\beta}\right\},\\ &= \mathcal{I}^{-1}(\beta_0)\left\{N^{-1/2}\frac{\partial\mathcal{L}(\beta_0)}{\partial\beta}\right\} + \mathbf{o}_P(1).\end{aligned}$$

Using Slutsky's Lemma, it follows that:

$$N^{1/2}\left(\hat{\beta}_N - \beta_0\right)\xrightarrow{\mathcal{L}}\mathcal{N}_b\left(0,\mathcal{I}^{-1}(\beta_0)\right).$$

See also Ahn and Reinsel (1988, p. 853).

From a practical point of view, it may be worth studying the asymptotic distribution of the coefficients $\hat{\Phi}_k(v)\equiv\Phi_k\left(v;\hat{\beta}_N\right)$ and $k=1,\ldots,p(v)$. It is more convenient to state the result in terms of the vectors $\phi_k\left(v;\hat{\beta}_N\right) = \mathrm{vec}\left\{\Phi_k\left(v;\hat{\beta}_N\right)\right\}$. The following results use the δ-method (see, e.g. Serfling (1980)). More precisely,

$$\phi_k\left(v;\hat{\beta}_N\right) - \phi_k(v;\beta_0) = \frac{\partial\phi_k(v;\beta_0)}{\partial\beta^\top}\left(\hat{\beta}_N - \beta_0\right) + \mathbf{O}_P(N^{-1}).$$

Hence:

$$N^{1/2}\left\{\phi_k\left(v;\hat{\beta}_N\right)-\phi_k(v;\beta_0)\right\}\xrightarrow{\mathcal{L}}\mathcal{N}_{d^2}\left(0,\frac{\partial\phi_k(v;\beta_0)}{\partial\beta^\top}\mathcal{I}^{-1}(\beta_0)\frac{\partial\phi_k^\top(v;\beta_0)}{\partial\beta}\right).$$

Having estimated β_0 by conditional maximum likelihood estimators, the model residuals $\hat{\varepsilon}_{ns+v}=\mathbf{r}_{ns+v}\left(\hat{\beta}_N\right)$, $n=0,\ldots,N-1$ and $v=1,\ldots,s$, can be calculated. Thus, we can also compute the residual autocovariances $\mathbf{C}_{\hat{\varepsilon}}(h;v)$, and we define naturally the vector of residual autocovariances $\mathbf{c}_{\hat{\varepsilon}}(v)$ for each season v.

The following example presents the key results of this section in the case of a PVAR (1) model.

Example 1 Consider the following model:

$$\mathbf{Y}_{ns+v}=\Phi(v;\beta_0)\mathbf{Y}_{ns+v-1}+\varepsilon_{ns+v}, \tag{10}$$

$v=1,\ldots,s$. Multiplying each side of (10) by $\mathbf{Y}_{ns+v-h}^\top$ and taking the mathematical expectation gives:

$$\Gamma_Y(h;v)=\Phi(v;\beta_0)\Gamma_Y(h-1;v-1)+E\left(\varepsilon_{ns+v}\mathbf{Y}_{ns+v-h}^\top\right). \tag{11}$$

For $h>0$, we obtain

$$\Gamma_Y(h;v)=\Phi(v;\beta_0)\Gamma_Y(h-1;v-1).$$

Setting $h=0$ in (11) gives

$$\begin{aligned}\Gamma_Y(0;v)&=\Phi(v;\beta_0)\Gamma_Y(-1;v-1)+\Sigma_\varepsilon(v),\\&=\Phi(v;\beta_0)\Gamma_Y^\top(1;v)+\Sigma_\varepsilon(v),\\&=\Phi(v;\beta_0)\Gamma_Y(0;v-1)\Phi^\top(v;\beta_0)+\Sigma_\varepsilon(v),\end{aligned}$$

which defines a linear system that can be solved for $\Gamma_Y(0;v)$ and $v\in\{1,\ldots,s\}$. To obtain the second equality in the last derivation, we used: $\Gamma_Y(-h;v)=\Gamma_Y^\top(h;v+h)$ and $h>0$. Alternatively, the moving average representation (2) and the expression (3) can be used to obtain explicit expressions. For example, the variance of the process at season v is given by:

$$\Gamma_Y(0;v)=\sum_{k=0}^{\infty}\Psi_k(v)\Sigma_\varepsilon(v-k)\Psi_k^\top(v).$$

The weights in (3) (defined in the periodic sense) need to be calculated. In the PVAR (1) model, these weights satisfy:

$$\Psi_k(v)=\begin{cases}\mathbf{I}_d, & \text{if } k=0,\\ \prod_{i=1}^{k}\Phi(v-i+1;\beta_0), & \text{if } k\geq1.\end{cases} \tag{12}$$

In our applications, see Section 5, the information matrix needs to be estimated. In the case of a PVAR (1) model, formula (9) simplifies to:

$$\mathcal{I}(\beta_0)=\sum_{v=1}^{s}\frac{\partial\phi^\top(v;\beta_0)}{\partial\beta}\left\{\Gamma_Y(0;v-1)\otimes\Sigma_\varepsilon^{-1}(v)\right\}\frac{\partial\phi(v;\beta_0)}{\partial\beta^\top}. \tag{13}$$

Note that the information matrix relies on $\Gamma_Y(0;v-1)$.

We study the asymptotic distributions of the residual autocovariances in Section 3.

3. ASYMPTOTIC DISTRIBUTION OF RESIDUAL AUTOCOVARIANCES

Let $\mathbf{c}_\varepsilon(v)=\left(\mathbf{c}_\varepsilon^\top(1;v),\ldots,\mathbf{c}_\varepsilon^\top(M;v)\right)^\top$, where $\mathbf{c}_\varepsilon(h;v)=\text{vec}\{\mathbf{C}_\varepsilon(h;v)\}$, be the vector of sample autocovariances of the white noise process $\{\varepsilon_t, t\in\mathbb{Z}\}$. In practical applications, our test procedures will rely on the residual autocovariances, but it is useful to present the asymptotic distribution of the random vector $\mathbf{c}_\varepsilon(v)$. Under the assumptions that the white noise is periodic and Gaussian, $N^{1/2}\mathbf{c}_\varepsilon(v)$ follows asymptotically a d^2M-variate normal distribution:

$$N^{1/2}\mathbf{c}_\varepsilon(v)\xrightarrow{\mathcal{L}}\mathcal{N}_{d^2M}(0,\mathbf{V}(v;M)\otimes\Sigma_\varepsilon(v)), \tag{14}$$

where $\mathbf{V}(v;M)$ corresponds to the $(dM)\times(dM)$ block diagonal matrix:

$$\mathbf{V}(v;M) = \begin{pmatrix} \Sigma_\varepsilon(v-1) & 0 & 0 & \dots & 0 \\ 0 & \Sigma_\varepsilon(v-2) & 0 & \dots & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \Sigma_\varepsilon(v-M) \end{pmatrix}. \tag{15}$$

This result can be obtained by considering that, using the property vec($\mathbf{ab}^\top$) = $\mathbf{b} \otimes \mathbf{a}$ for any vectors $\mathbf{a}$ and $\mathbf{b}$ (see Harville (1997, p. 340)), we have the useful relations:

$$\mathbf{c}_\varepsilon(j;v) = N^{-1}\sum_{n=j}^{N-1} \mathrm{vec}\left(\varepsilon_{ns+v}\varepsilon_{ns+v-j}^\top\right) = N^{-1}\sum_{n=j}^{N-1}\left(\varepsilon_{ns+v-j}\otimes\varepsilon_{ns+v}\right),$$

$j \in \{1,\ldots,M\}$. See also Ursu and Duchesne (2009) for additional details. By using (7) and (8), straightforward calculations give:

$$E\left\{\frac{\partial\mathcal{L}(\beta_0)}{\partial\beta}\mathbf{c}_\varepsilon^\top(j;v)\right\} = N^{-1}\sum_{n=j}^{N-1}\sum_{i=1}^{p(v)}\frac{\partial\phi_i^\top(v;\beta_0)}{\partial\beta}E\left\{\left(\mathbf{Y}_{ns+v-i}\varepsilon_{ns+v-j}^\top\right)\otimes\boldsymbol{\Sigma}_\varepsilon^{-1}(v)\mathbf{r}_{ns+v}(\beta_0)\varepsilon_{ns+v}^\top\right\}.$$

Noting that $E\left(\mathbf{Y}_{ns+v-i}\varepsilon_{ns+v-j}^\top\right) = \Psi_{j-i}(v-i;\beta_0)\Sigma_\varepsilon(v-j)$ and $\forall\, n$, it follows that, as $N\to\infty$,

$$E\left\{\frac{\partial\mathcal{L}(\beta_0)}{\partial\beta}\mathbf{c}_\varepsilon^\top(j;v)\right\} \to \sum_{i=1}^{p(v)}\frac{\partial\phi_i^\top(v;\beta_0)}{\partial\beta}\left\{\Psi_{j-i}(v-i;\beta_0)\Sigma_\varepsilon(v-j)\otimes\mathbf{I}_d\right\} \equiv \mathbf{A}(j;v),$$

where the dimensions of the matrices $\mathbf{A}(j;v)$ are $b\times d^2$. The seasonal weights are defined in the periodic sense, and they satisfy the relations $\Psi_0(v;\beta_0) = \mathbf{I}_d$ and $\Psi_k(v;\beta_0) = 0$, $k<0$, $v = 1,\ldots,s$. Using the differentiation rule

$$\frac{\partial\mathrm{vec}(\mathbf{AB})}{\partial\beta^\top} = \left(\mathbf{I}_q\otimes\mathbf{A}\right)\frac{\partial\mathrm{vec}(\mathbf{B})}{\partial\beta^\top} + \left(\mathbf{B}^\top\otimes\mathbf{I}_n\right)\frac{\partial\mathrm{vec}(\mathbf{A})}{\partial\beta^\top},$$

where $\mathbf{A}$ and $\mathbf{B}$ are $n\times p$ and $p\times q$ matrices respectively (see, e.g. Lütkepohl (2005, Appendix A.13)), it follows that:

$$\frac{\partial\mathbf{c}_\varepsilon(j;v)}{\partial\beta^\top} = N^{-1}\sum_{n=j}^{N-1}\left\{\left(\mathbf{I}_d\otimes\varepsilon_{ns+v}\right)\frac{\partial\varepsilon_{ns+v-j}}{\partial\beta^\top} + \left(\varepsilon_{ns+v-j}\otimes\mathbf{I}_d\right)\frac{\partial\varepsilon_{ns+v}}{\partial\beta^\top}\right\}.$$

Thus,

$$E\left\{\frac{\partial\mathbf{c}_\varepsilon(j;v)}{\partial\beta^\top}\right\} = N^{-1}\sum_{n=j}^{N-1}E\left\{\left(\varepsilon_{ns+v-j}\otimes\mathbf{I}_d\right)\frac{\partial\varepsilon_{ns+v}}{\partial\beta^\top}\right\}.$$

Using (8), it follows that, as $N\to\infty$,

$$E\left\{\frac{\partial\mathbf{c}_\varepsilon(j;v)}{\partial\beta^\top}\right\} \to -\sum_{i=1}^{p(v)}\left\{\Sigma_\varepsilon(v-j)\Psi_{j-i}^\top(v-i;\beta_0)\otimes\mathbf{I}_d\right\}\frac{\partial\phi_i(v;\beta_0)}{\partial\beta^\top} = -\mathbf{A}^\top(j;v).$$

A standard Taylor expansion gives:

$$\mathbf{c}_{\hat\varepsilon}(v) = \mathbf{c}_\varepsilon(v) + \frac{\partial\mathbf{c}_\varepsilon(v)}{\partial\beta^\top}\left(\hat\beta_N - \beta_0\right) + \mathbf{o}_p\left(N^{-1/2}\right).$$

For a similar development in VAR models with structured parameterization, see Ahn (1988, p. 591). Let

$$\mathbf{A}(v) = \left(\mathbf{A}(1;v),\cdots,\mathbf{A}(M;v)\right) \tag{16}$$

be the $b\times(d^2M)$ matrix corresponding to the asymptotic cross-covariance matrix between $N^{-1/2}\partial\mathcal{L}(\beta_0)/\partial\beta$ and $N^{1/2}\mathbf{c}_\varepsilon(v)$. It follows that:

$$E\left\{\frac{\partial\mathbf{c}_\varepsilon(v)}{\partial\beta^\top}\right\} \to -\mathbf{A}^\top(v), \text{as } N\to\infty. \tag{17}$$

Note that (17) is expressed in function of the derivatives of the autoregressive parameters associated with season v. Consequently, $N^{1/2}\mathbf{c}_{\hat\varepsilon}(v)$ and $N^{1/2}\left\{\mathbf{c}_\varepsilon(v) - \mathbf{A}^\top(v)\left(\hat\beta_N - \beta_0\right)\right\}$ share the same asymptotic distribution. A direct calculation gives:

$$N\text{var}\left\{\mathbf{c}_\varepsilon(v) - \mathbf{A}^\top(v)\left(\hat{\beta}_N - \beta_0\right)\right\} \to \mathbf{V}(v;M)\otimes\Sigma_\varepsilon(v) - \mathbf{A}^\top(v)\mathcal{I}^{-1}(\beta_0)\mathbf{A}(v), \tag{18}$$

as $N \to \infty$. Next, we state the joint asymptotic distribution of $N^{1/2}\left(\hat{\beta}_N - \beta_0\right)$ and $N^{1/2}\mathbf{c}_\varepsilon(v)$.

PROPOSITION 1 *Suppose that $\{\mathbf{Y}_t, t \in \mathbb{Z}\}$ denotes a PVAR process satisfying (1), with periodic Gaussian white noise $\{\varepsilon_t, t \in \mathbb{Z}\}$. Let $\hat{\beta}_N$ be the conditional maximum likelihood estimator of β_0. Consider a vector of M sample autocovariances collected in the vector $\mathbf{c}_\varepsilon(v)$. Then:*

$$\begin{pmatrix} N^{1/2}\left(\hat{\beta}_N - \beta_0\right) \\ N^{1/2}\mathbf{c}_\varepsilon(v) \end{pmatrix} \xrightarrow{\mathcal{L}} \mathcal{N}_{b+d^2M}\left(0, \begin{pmatrix} \mathcal{I}^{-1}(\beta_0) & \mathcal{I}^{-1}(\beta_0)\mathbf{A}(v) \\ \mathbf{A}^\top(v)\mathcal{I}^{-1}(\beta_0) & \mathbf{V}(v;M)\otimes\Sigma_\varepsilon(v) \end{pmatrix}\right),$$

where $\mathbf{V}(v;M)$ corresponds to the expression (15), $\mathcal{I}(\beta_0)$ is the information matrix (9) and $\mathbf{A}(v)$ is defined by (16).

The proof of Proposition 1 follows arguments similar to those in Ahn (1988), using periodic versions of limit theorems. See also Basawa and Lund (2001), Lund *et al.* (2006) and Ursu and Duchesne (2009). This leads us to the asymptotic distribution of the residual autocovariances.

PROPOSITION 2 *Suppose that $\{\mathbf{Y}_t, t \in \mathbb{Z}\}$ denotes a PVAR process satisfying (1), with a periodic Gaussian white noise $\{\varepsilon_t, t \in \mathbb{Z}\}$. Let $\hat{\beta}_N$ be the conditional Gaussian maximum likelihood estimator of β_0. Consider a vector of M residual autocovariances collected in the vector $\mathbf{c}_{\hat{\varepsilon}}(v)$ given by (6). Then:*

$$N^{1/2}\mathbf{c}_{\hat{\varepsilon}}(v) \xrightarrow{\mathcal{L}} \mathcal{N}_{d^2M}(0, \Omega_{vv}),$$

with $\Omega_{vv} = \mathbf{V}(v;M)\otimes\Sigma_\varepsilon(v) - \mathbf{A}^\top(v)\mathcal{I}^{-1}(\beta_0)\mathbf{A}(v)$, where $\mathbf{V}(v;M)$ corresponds to the expression (15), $\mathcal{I}(\beta_0)$ is the information matrix (9) and $\mathbf{A}(v)$ is defined by (16). Let $\mathbf{c}_{\hat{\varepsilon}} = \left(\mathbf{c}_{\hat{\varepsilon}}^\top(1), \ldots, \mathbf{c}_{\hat{\varepsilon}}^\top(s)\right)^\top$. Then

$$N^{1/2}\mathbf{c}_{\hat{\varepsilon}} \xrightarrow{\mathcal{L}} \mathcal{N}_{d^2Ms}(0, \Omega),$$

where the asymptotic covariance matrix Ω is a block matrix, with the asymptotic variances given by Ω_{vv}, $v = 1, \ldots, s$, and the asymptotic covariances given by:

$$\text{cov}\left(N^{1/2}\mathbf{c}_{\hat{\varepsilon}}(v), N^{1/2}\mathbf{c}_{\hat{\varepsilon}}\left(v'\right)\right) \to -\mathbf{A}^\top(v)\mathcal{I}^{-1}(\beta_0)\mathbf{A}\left(v'\right) = \Omega_{vv'},$$

as $N \to \infty$.

A particular case is when the parameters are indexed by the season v such that $\beta_0 = \left(\beta_0^\top(1), \ldots, \beta_0^\top(s)\right)^\top$. Suppose that the autoregressive parameters $\Phi_k(\beta_0;v)$, $k = 1, \ldots, p(v)$, are only functions of $\beta_0(v)$, and that the autoregressive parameters of season v are allowed to satisfy linear constraints inside that season. Under these conditions, it is possible to show that the conditional Gaussian maximum likelihood estimators of $\beta_0(v)$ and $\beta_0(v')$, $v \neq v'$ are asymptotically independent. See also Basawa and Lund (2001). Under these conditions, $N^{1/2}\mathbf{c}_{\hat{\varepsilon}}(v)$ and $N^{1/2}\mathbf{c}_{\hat{\varepsilon}}\left(v'\right)$ are thus asymptotically uncorrelated, $v \neq v'$, and we retrieve the results given in Ursu and Duchesne (2009) for PVAR models.

The results presented in Proposition 2 are valid in the more general case of parsimonious PVAR models. Examples include Fourier representations of the autoregressive coefficients:

$$\Phi_k(v;\beta_0) = \mathbf{A}_{0,k} + \sum_{l=1}^{r}\left\{\mathbf{B}_{l,k}sin(2\pi lv/s) + \mathbf{A}_{l,k}cos(2\pi lv/s)\right\}, \tag{19}$$

where $k = 1, \ldots, p(v)$, $v = 1, \ldots, s$. See also Lund *et al.* (2006) in the univariate case. Thus, in general, the parameters between the seasons may be functionally related. For parsimonious PVAR models, the autoregressive estimators $\Phi_k\left(\hat{\beta}_N; v\right)$ and $\Phi_l\left(\hat{\beta}_N; v'\right)$, $v \neq v'$, will not be asymptotically independent. From Proposition 2, the asymptotic covariances between $N^{1/2}\mathbf{c}_{\hat{\varepsilon}}(v)$ and $N^{1/2}\mathbf{c}_{\hat{\varepsilon}}\left(v'\right)$, $v \neq v'$ are given by $\Omega_{vv'}$, which do not vanish in general. Thus, the residual autocovariances are generally not independent. Compared with the results of McLeod (1994) and Ursu and Duchesne (2009), that result represents a substantial difference that complicates the asymptotic distributions of the Hosking-Li-McLeod test statistics, as discussed in Section 4.

4. APPLICATIONS TO DIAGNOSTIC CHECKING

Diagnostic checking of parsimonious PVAR models represents a useful application of Proposition 2. In the previous sections, it was assumed that $\{\varepsilon_t, t \in \mathbb{Z}\}$ was white noise. Here we want to test that hypothesis of model adequacy by looking at the remaining dependence in the residuals. Let $\Gamma_\varepsilon(h;v) = \mathrm{cov}(\varepsilon_{ns+v}\varepsilon_{ns+v-h})$ be the lag h theoretical autocovariance matrices at season v of $\{\varepsilon_t, t \in \mathbb{Z}\}$. These quantities are collected in the $(d^2M) \times 1$ vector:

$$\gamma_\varepsilon(v) = \left(\gamma_\varepsilon^\top(1;v), \gamma_\varepsilon^\top(2;v), \ldots, \gamma_\varepsilon^\top(M;v)\right)^\top, \tag{20}$$

where $\gamma_\varepsilon(h;v) = \mathrm{vec}\{\Gamma_\varepsilon(h;v)\}$. More formally, the null hypothesis of model adequacy is given by:

$$H_0 : \gamma_\varepsilon(v) = 0, v = 1, \ldots, s,$$

where $\gamma_\varepsilon(v)$ is defined by (20) and **0** corresponds to the $(d^2M) \times 1$ null vector. McLeod (1994) and Hipel and McLeod (1994) suggested that test statistics be used for each season v, $v = 1, \ldots, s$. They also considered global test statistics that take all the seasons into account. That strategy has been generalized to the PVAR models in Ursu and Duchesne (2009) and will be now developed for the PVAR models with structured parameterization.

Consider the matrices $\mathbf{P}_i(v)$, $i = 1, \ldots, M$, such that $\Sigma_\varepsilon^{-1}(v-i) \otimes \Sigma_\varepsilon^{-1}(v) = \mathbf{P}_i^\top(v)\mathbf{P}_i(v)$, $i = 1, \ldots, M$ and $\mathbf{P}_i(v)\{\Sigma_\varepsilon(v-i) \otimes \Sigma_\varepsilon(v)\}\mathbf{P}_i^\top(v) = \mathbf{I}_{d^2}$. Let the block-diagonal matrix $\mathbf{Q}_M(v)$ be defined by:

$$\mathbf{Q}_M(v) = \begin{pmatrix} \mathbf{P}_1(v) & 0 & 0 & \ldots & 0 \\ 0 & \mathbf{P}_2(v) & 0 & \ldots & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & \mathbf{P}_M(v) \end{pmatrix}. \tag{21}$$

The relation $\mathbf{Q}_M^\top(v)\mathbf{Q}_M(v) = \mathbf{V}^{-1}(v;M) \otimes \Sigma_\varepsilon^{-1}(v)$ follows. Let $\tilde{c}_{\hat\varepsilon}(v) = \mathbf{Q}_M(v)\mathbf{c}_{\hat\varepsilon}(v)$. Thus the asymptotic covariance matrix of $\tilde{\mathbf{c}}_{\hat\varepsilon}(v)$ is given by:

$$\lim_{N\to\infty} N\mathrm{var}\{\tilde{\mathbf{c}}_{\hat\varepsilon}(v)\} \equiv \tilde{\Omega}_{vv} = \mathbf{I}_{d^2M} - \mathbf{Q}_M(v)\left\{\mathbf{A}^\top(v)\mathcal{I}^{-1}(\beta_0)\mathbf{A}(v)\right\}\mathbf{Q}_M^\top(v), \tag{22}$$

and the asymptotic covariance between $\tilde{\mathbf{c}}_{\hat\varepsilon}(v)$ and $\tilde{\mathbf{c}}_{\hat\varepsilon}(v')$ is given by

$$\lim_{N\to\infty} N\mathrm{cov}\left\{\tilde{\mathbf{c}}_{\hat\varepsilon}(v), \tilde{\mathbf{c}}_{\hat\varepsilon}\left(v'\right)\right\} \equiv \tilde{\Omega}_{vv'} = -\mathbf{Q}_M(v)\left\{\mathbf{A}^\top(v)\mathcal{I}^{-1}(\beta_0)\mathbf{A}\left(v'\right)\right\}\mathbf{Q}_M^\top\left(v'\right). \tag{23}$$

Let $\hat{\mathbf{Q}}_M(v)$ be a consistent estimator of $\mathbf{Q}_M(v)$, obtained by estimating consistently $\mathbf{V}(v;M)$ and $\Sigma_\varepsilon(v-l)$, $l = 1, \ldots, M$ by $\hat{\mathbf{V}}(v;M)$ and $\hat{\Sigma}_\varepsilon(v-l)$, $l = 1, \ldots, M$.

The Hosking-Li-McLeod test statistic for a given season v is essentially based on the quadratic form $N\tilde{\mathbf{c}}_{\hat\varepsilon}^\top(v)\tilde{\mathbf{c}}_{\hat\varepsilon}(v)$. More precisely, it is defined by:

$$\mathcal{Q}_M(v) = N\sum_{l=1}^{M} \omega(l;v,N,s)\mathrm{tr}\left\{\mathbf{C}_{\hat\varepsilon}^\top(l;v)\hat{\Sigma}_\varepsilon^{-1}(v)\mathbf{C}_{\hat\varepsilon}(l;v)\hat{\Sigma}_\varepsilon^{-1}(v-l)\right\}. \tag{24}$$

The factor $\omega(l;v,N,s)$ needs to be specified. A possible natural choice is $\omega(l;v,N,s) \equiv 1$, but the Ljung–Box correction factor developed by McLeod (1994) defined by:

$$\omega(l;v,N,s) = \begin{cases} (N+2)/(N-l/s), & \text{if } l \equiv 0 \quad mod \quad s, \\ N/(N - \lfloor (l-v+s)/s \rfloor), & \text{otherwise}, \end{cases} \tag{25}$$

where $\lfloor a \rfloor$ denotes the integer part of the number a, is expected to improve the finite sample properties. See also Ursu and Duchesne (2009). We call (25) the Ljung-Box-McLeod correction factor. When the parameters are indexed by the season and estimated by conditional Gaussian maximum likelihood or least-squares estimators, the asymptotic covariance matrix (22) is approximatively idempotent. As a corollary, the asymptotic distribution of the test statistic $\mathcal{Q}_M(v)$ is well approximated under these conditions by a chi-square distribution with degrees of freedom depending on the rank of (22). The test statistic (24) is valid under more general assumptions. In the general case, classical results on quadratic forms enable one to demonstrate that the asymptotic distribution of $\mathcal{Q}_M(v)$ is given by:

$$\mathcal{Q}_M(v) \xrightarrow{\mathcal{L}} \sum_{i=1}^{d^2M} \lambda_i(v) Z_i^2, \tag{26}$$

where $Z_1, \ldots, Z_{d^2M}$ are independent $\mathcal{N}(0,1)$ random variables, and $\lambda_i(v)$, $i = 1, \ldots, d^2M$, correspond to the eigenvalues of the asymptotic covariance matrix (22).

A global portmanteau test statistic that takes into account all the seasons can be based on $N^{1/2}\tilde{\mathbf{c}}_{\hat\varepsilon} = N^{1/2}\left(\tilde{\mathbf{c}}_{\hat\varepsilon}^\top(1), \ldots, \tilde{\mathbf{c}}_{\hat\varepsilon}^\top(s)\right)^\top$, which has a block asymptotic covariance matrix $\tilde{\Omega}$, with diagonal blocks given by (22) and covariances satisfying (23). The global Hosking-Li-McLeod test statistic is:

$$\mathcal{Q}_M = \sum_{v=1}^{s} \mathcal{Q}_M(v). \tag{27}$$

The asymptotic distribution of $\mathcal{Q}_M$ is also a weighted sum of chi-square random variables:

$$\mathcal{Q}_M \xrightarrow{\mathcal{L}} \sum_{i=1}^{d^2Ms} \lambda_i Z_i^2, \tag{28}$$

where the λ_is denote the eigenvalues of the covariance matrix $\widetilde{\Omega}$.

The covariance matrix (22) may be arbitrarily far from an idempotent matrix, even when the periodic white noise is composed of independent random variables, and consequently the distribution of the portmanteau test statistic is not well approximated by a chi-square distribution. That result may be surprising, given the literature on portmanteau test statistics for diagnostic checking of time series models with independent white noise. Under certain conditions, it is well-known that the finite sample distribution of the Box-Pierce-Ljung test statistics is better approximated by the distribution of a weighted sum of chi-squares random variables. See Ljung (1986), among others. However, the chi-square approximation is generally satisfying, and it is of common use in practical applications. This is discussed in Li (2004). When the noise is composed of uncorrelated random variables, it is better to use the distribution of a weighted sum of chi-squares variables to approximate the finite sample distributions of Box-Pierce-Ljung test statistics (see Francq *et al.* (2005) and Francq and Raïssi (2007)). In the periodic case, the information matrix $\mathcal{I}(\beta)$ involves the derivatives of the autoregressive parameters across *all* seasons, whereas the matrices $\mathbf{A}(v)$ are always calculated for a given season (see Li and McLeod (1981, Theorem 3) and Ahn (1988, p. 592)).

It is possible to evaluate the distribution of the Gaussian quadratic form in (26) by means of Imhof's (1961) algorithm or other exact methods. More precisely, the Hosking-Li-McLeod test procedure $\mathcal{Q}_M(v)$ at season v relies on the following steps: (1) compute the eigenvalues $\hat{\lambda}_1(v), \ldots, \hat{\lambda}_{d^2M}(v)$ of $\widetilde{\Omega}_{N,vv}$, which provides a consistent estimator of $\widetilde{\Omega}_{vv}$. The asymptotic covariance matrix $\widetilde{\Omega}_{vv}$ is estimated by replacing unknown quantities in (22) by consistent estimators, following the traditional literature, (2) evaluate the $(1-\alpha)$-quantile $c_\alpha\left(\hat{\lambda}_1(v), \ldots, \hat{\lambda}_{d^2M}(v)\right)$ of $\sum_{i=1}^{d^2M} \hat{\lambda}_i N_i^2$ using Imhof's algorithm and (3) the null hypothesis is rejected when

$$\mathcal{Q}_M(v) \geq c_\alpha\left(\hat{\lambda}_1(v), \ldots, \hat{\lambda}_{d^2M}(v)\right).$$

The test procedure for $\mathcal{Q}_M$ is similar, but it is based on the d^2Ms empirical eigenvalues of a consistent estimator of the matrix $\widetilde{\Omega}$.

Interesting advantages of the proposed procedures are that the ranks of $\widetilde{\Omega}_{vv}$ or $\widetilde{\Omega}$ do not need to be known. Interestingly, the procedure is identical whether the asymptotic variance matrices are (approximately) singular (and idempotent) or non-singular. Thus, the test procedures are more general than the ones described in Ursu and Duchesne (2009): if the autoregressive parameters are indexed by the seasons and if they satisfy linear constraints inside the seasons, zero eigenvalues will be observed in the asymptotic limit. However, the general testing procedures remain unchanged.

5. SIMULATION EXPERIMENTS

In Section 4, we presented portmanteau test statistics. From a practical point of view, it seems natural to inquire about their finite sample properties. Here, we report the simulation results of a Monte Carlo experiment conducted to study the exact level of the test statistics calculated at each season $\mathcal{Q}_M(v)$, $v = 1, \ldots, s$, and that of the global test statistic $\mathcal{Q}_M$. To compare the exact distribution of the test statistics with their corresponding distribution, the following bivariate data-generating process was used:

$$\mathbf{Y}_{ns+v} = \Phi(v;\beta_0)\mathbf{Y}_{ns+v-1} + \varepsilon_{ns+v}. \tag{29}$$

We considered the case of twelve seasons, that is $s = 12$. Thus, we consider a bivariate PVAR (1). In general, that model relies on a maximum of $d^2s = 48$ free parameters. We used the simple Fourier representation of the PVAR parameters:

$$\Phi(v;\beta_0) = \mathbf{A}_0 + sin(2\pi v/12)\mathbf{B}_1 + cos(2\pi v/12)\mathbf{A}_1, v = 1, \ldots, 12,$$

with the model parameters $\beta_0 = (\text{vec}^\top(\mathbf{A}_0), \text{vec}^\top(\mathbf{B}_1), \text{vec}^\top(\mathbf{A}_1))^\top$, defined as:

$$\mathbf{A}_0 = \begin{pmatrix} 0.4 & 0.2 \\ 0.2 & 0.3 \end{pmatrix}, \mathbf{B}_1 = \begin{pmatrix} 0.1 & 0.4 \\ 0.4 & 0.2 \end{pmatrix}, \mathbf{A}_1 = \begin{pmatrix} 0.2 & 0.1 \\ 0.1 & 0.2 \end{pmatrix}.$$

The information matrix (9) reduces to the formula (13) with $s = 12$, where the 4×12 matrix of derivatives $\partial\phi(v;\beta_0)/\partial\beta^\top$ satisfies:

$$\frac{\partial\phi(v;\beta_0)}{\partial\beta^\top} = (\mathbf{I}_4 \vdots sin(2\pi v/12)\mathbf{I}_4 \vdots cos(2\pi v/12)\mathbf{I}_4),$$

Journal of
Time Series Analysis

Table 1. Empirical levels (number of rejection of the null hypothesis over 10,000 replications) at the 5% and 10% significance levels for the portmanteau test statistics $\mathcal{Q}_M(v)$ defined by (24) for $v = 1, \ldots, 12$ and $M = 1,2,3,6,8,10$. The model is given by (29). The number of years are equal to $N = 200$ and 400

	N = 200																							
$M \setminus v$	1		2		3		4		5		6		7		8		9		10		11		12	
	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%
1	457	967	497	1007	492	977	483	980	479	986	469	956	471	957	497	994	475	972	477	970	480	1015	502	984
2	459	938	456	1007	460	931	471	961	486	1006	473	899	448	966	455	962	463	960	534	1054	480	945	517	1021
3	444	910	459	942	490	976	468	962	44	900	485	951	468	946	460	937	463	962	469	995	466	933	466	945
6	441	889	440	919	473	944	428	871	413	859	417	846	421	860	446	906	435	855	445	959	420	889	437	901
8	404	881	406	866	396	871	384	834	384	836	409	858	389	818	418	852	404	804	403	856	391	797	361	820
10	382	830	362	843	368	809	367	806	379	802	358	737	347	759	367	783	391	765	396	797	345	771	350	757
$M \setminus v$	N = 400																							
	1		2		3		4		5		6		7		8		9		10		11		12	
	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%
1	526	1034	470	981	506	1012	465	964	479	1016	479	1003	450	984	461	976	502	999	465	949	472	934	543	1044
2	512	1000	503	1024	502	956	449	948	503	973	495	1041	497	982	486	984	514	1049	485	1008	483	979	532	1066
3	522	1050	517	1055	494	956	462	941	500	974	468	959	478	933	483	955	523	969	510	967	499	1021	454	964
6	483	997	449	942	477	928	428	907	450	937	484	963	439	894	470	922	459	940	489	1002	464	933	466	954
8	452	982	428	889	463	959	445	919	464	939	455	903	431	883	449	892	450	882	462	938	449	911	425	891
10	438	952	432	900	434	892	408	885	442	894	462	944	407	870	446	894	427	871	472	889	420	905	433	854

Table 2. Empirical levels (number of rejections of the null hypothesis over 10,000 replications) at the 5% and 10% significance levels for the portmanteau test statistics $\mathcal{Q}_M$ defined by (27) with $M=1,2,3,6,8,10$. The model is given by (29). The numbers of years are equal to $N=200$ and 400

M	$N=200$		$N=400$	
	5%	10%	5%	10%
1	487	969	477	980
2	440	948	488	1005
3	413	865	510	1001
6	357	726	410	867
8	275	623	370	748
10	205	498	320	691

The process $\varepsilon = \{\varepsilon_t, t \in \mathbb{Z}\}$ was assumed to be Gaussian white noise, composed of independent Gaussian random vectors with expectation **0** and common covariance matrix $\Sigma(v)=\mathbf{I}_2$, $s=1,\ldots,12$. To estimate $\mathbf{A}(v)$ defined by (16), explicit expressions of the seasonal weights $\Psi_k(v;\beta_0)$ are needed. These weights are given in Example 1.

We studied the empirical frequencies of rejection of the null hypothesis of adequacy at the 5% and 10% nominal levels, for each of two series length ($N=200$ and 400). For each experiment, 10,000 independent realizations were generated. Let $\hat{\beta}_N^{(0)} = \beta_0$ be a vector of initial values. For each realization of the PVAR (1) model defined by (29), the true model was estimated using the recursions:

$$\hat{\beta}_N^{(i+1)} = \hat{\beta}_N^{(i)} + \hat{\mathcal{I}}^{-1}\left(\hat{\beta}^{(i)}\right)\left\{\frac{N^{-1}\partial\hat{\mathcal{L}^{(i)}}}{\partial\beta^{\top}}\right\},$$

where the information matrix and the derivatives were naturally estimated. A tolerance of order 10^{-6} has been adopted on the sum of absolute differences $\sum_{j=1}^{12}\left|\hat{\beta}_N^{(i+1)}(j) - \hat{\beta}_N^{(i)}(j)\right|$, where $\hat{\beta}_N^{(i)} = \left(\hat{\beta}_N^{(i)}(1),\ldots,\hat{\beta}_N^{(i)}(12)\right)^{\top}$, to ensure the convergence of the algorithm. For each residual time series, the portmanteau test statistics $\mathcal{Q}_M(v)$, $v=1,\ldots,s$ and $\mathcal{Q}_M$ were calculated using the Ljung-Box-McLeod correction factor (25), for $M \in \{1,2,3,6,8,10\}$. Imhof's (1961) algorithm has been used to obtain the critical values; we used the R package CompQuadForm available from CRAN. All the simulation codes have been implemented in R. To attain computational efficiency, some parts have also been written using the C language and interfaced with the R software.

The number of rejections of the null hypothesis of adequacy are reported in Tables 1 and 2. The results presented in Table 1 indicate that some under-rejection has been observed for large values of M at both significance levels when $N=200$. Generally the results improved with N; when $N=400$, the empirical levels were close to the nominal levels at both significance levels. In unreported simulation experiments, we computed the P-values of the test statistics using the quantiles from the $\chi^2_{d^2M}$ and $\chi^2_{d^2(M-1)}$ distributions. The $\chi^2_{d^2M}$ distribution would be appropriate if the test statistics could be calculated using the (unknown) innovations, whereas Ursu and Duchesne (2009) argue that the $\chi^2_{d^2(M-1)}$ distribution is appropriate if an unrestricted PVAR (1) model is estimated. However, severe under-rejections occurred using the $\chi^2_{d^2M}$ distribution: at the 5% level, the empirical levels were systematically inferior to the results presented in Table 1. For example, when $M=6$ and $v=2$, we obtained 2.93% and 2.79% for $N=200$, 400 respectively, using the quantile of order 95th $\chi^2_{0.95,24} = 36.41$. However, we obtained over-rejection using the quantile $\chi^2_{0.95,20} = 31.41$ with $d^2(M-1)=20$ degrees of freedom. In that case, the empirical levels were equal to 9.66% and 10.12% under the sample sizes $N=200$, 400 respectively. In both situations, the results were significantly far from the 5% nominal level. Using Imhof's algorithm and the correct asymptotic distribution, we obtained from Table 1 empirical levels equal to 4.40% and 4.49%, which shows that the test statistic is slightly conservative.

The empirical levels of the global portmanteau test statistics are given in Table 2. Some under-rejections occurred for large values of M. When $N=400$, the results were generally reasonable at the 5% nominal level when $M \in \{1,2,3,6\}$. Considering the high dimension of the covariance estimator of the $(d^2sM) \times (d^2sM)$ matrix $\widetilde{\Omega}$, it is not surprising that large sample sizes are needed for the global test statistics. However, in practical applications, performing diagnostic checks season by season seems highly desirable, and the portmanteau test statistics $\mathcal{Q}_M(v)$, $v=1,\ldots,s$ are thus recommended for use.

Acknowledgements

The authors thank the editor, an associate editor and a referee for constructive suggestions and encouragements. The authors are most grateful to the referee, who went well beyond the call of duty in providing several constructive remarks and suggestions, which led to a significant improvement of the article. This work was supported by grants from the Natural Sciences and Engineering Research Council of Canada.

REFERENCES

Ahn, S. K. (1988) Distributions for residual autocovariance in multivariate autoregressive models with structured parameterization. *Biometrika* **75**, 590–593.

Ahn, S. K. and Reinsel, G. C. (1988) Nested reduced-rank autoregressive models for multiple time series. *Journal of the American Statistical Association* **83**, 849–856.

Anderson, T. W. (1980) Maximum likelihood estimation for vector autoregressive moving average models. In *Directions in Time Series*, (eds D. R. Brillinger and G. C. Tiao). Institute of Mathematical Statistics, pp. 49–59.

Basawa, I. V. and Lund, R. B. (2001) Large sample properties of parameter estimates for periodic ARMA models. *Journal of Time Series Analysis* **22**, 651–663.

Bloomfield, P., Hurd, H. L. and Lund, R. B. (1994) Periodic correlation in stratospheric ozone data. *Journal of Time Series Analysis* **15**, 127–150.

Boubacar Mainassara, Y. (2011) Multivariate portmanteau test for structural VARMA models with uncorrelated but non-independent error terms. *Journal of Statistical Planning and Inference* **141**, 2961–2975.

Boubacar Mainassara, Y. and Francq, C. (2011) Estimating structural VARMA models with uncorrelated but non-independent error terms. *Journal of Multivariate Analysis* **102**, 496–505.

Duchesne, P. and Lafaye de Micheaux, P. (2010) Computing the distribution of quadratic forms: Further comparisons between the Liu-Tang-Zhang approximation and exact methods. *Computational Statistics & Data Analysis* **54**, 858–862.

Francq, C. and Raïssi, H. (2007) Multivariate portmanteau test for autoregressive models with uncorrelated but non-independent errors. *Journal of Time Series Analysis* **28**, 454–470.

Francq, C., Roy, R. and Zakoïan, J.-M. (2005). Diagnostic checking in ARMA models with uncorrelated errors. *Journal of the American Statistical Association* **100**, 532–544.

Franses, P. H. and Paap, R. (2004) *Periodic Time Series Models*. New York: Oxford University Press.

Gladyšhev, E. G. (1961) Periodically correlated random sequences. *Soviet mathematics* **2**, 385–388.

Hannan, E. J. and Deistler, M. (1988) *The Statistical Analysis of Linear Systems*. New York: Wiley.

Hannan, E. J. and Kavalieris, L. (1984) Multivariate linear time series models. *Advances in Applied Probability* **16**, 492–561.

Harville, D. A. (1997) *Matrix Algebra from a Statistician's Perspective*. New York: Springer-Verlag.

Hipel, K. W. and McLeod, A. I. (1994) *Time Series Modelling of Water Resources and Environmental Systems*. Amsterdam: Elsevier.

Holan, S. H., Lund, R. and Davis, G. (2010) 'The ARMA alphabet soup: A tour of ARMA model variants. *Statistics Surveys* **4**, 232–274.

Hosking, J. (1980) The multivariate portmanteau statistic. *Journal of the American Statistical Association* **75**, 602–608.

Imhof, J. P. (1961) Computing the distribution of quadratic forms in normal variables. *Biometrika* **48**, 419–426.

Li, W. K. (2004) *Diagnostic Checks in Time Series*. New York: Chapman & Hall/CRC.

Li, W. K. and McLeod, A. I. (1981) Distribution of the residual autocorrelations in multivariate ARMA time series models. *Journal of the Royal Statistical Society, Series B* **43**, 231–239.

Ljung, G. M. (1986) Diagnostic testing of univariate time series models. *Biometrika* **73**, 725–730.

Lund, R. B., Shao, Q. and Basawa, I. V. (2006) Parsimonious periodic time series modeling. *Australian & New Zealand Journal of Statistics* **48**, 33–47.

Lütkepohl, H. (2005) *New Introduction to Multiple Time Series Analysis*. Berlin: Springer.

McLeod, A. I. (1994) Diagnostic checking periodic autoregression models with applications. *Journal of Time Series Analysis* **15**, 221–233.

Pagano, M. (1978) On periodic and multiple autoregressions. *The Annals of Statistics* **6**, 1310–1317.

Poskitt, D. S. and Salau, M. O. (1995) On the relationship between generalized least squares and Gaussian estimation of vector ARMA models. *Journal of Time Series Analysis* **16**, 617–645.

Reinsel, G. C. (1997) *Elements of Multivariate Time Series Analysis*, Second edition. New York: Springer-Verlag.

Serfling, R. J. (1980) *Approximation Theorems of Mathematical Statistics*. New York: Wiley.

Tunnicliffe Wilson, G. (1973) The estimation of parameters in multivariate time series models. *Journal of the Royal Statistical Society, Series B* **35**, 76–85.

Ursu, E. and Duchesne, P. (2009) On modeling and diagnostic checking of vector periodic autoregressive time series models. *Journal of Time Series Analysis* **30**, 70–96.