AUTOCORRELATION BASED TESTS FOR VECTOR ERROR CORRECTION MODELS WITH UNCORRELATED BUT NONINDEPENDENT ERRORS

Hamdi Raïssi

EQUIPPE-GREMARS, Université Lille 3

Abstract

We consider in this paper the estimation and test-of-fit for vector error correction models with nonindependent innovations. The asymptotic properties of the residual sample autocorrelations are derived. It is shown that the asymptotic distribution can be quite different for models with iid innovations and models in which the innovations are nonindependent. Consequently, the usual chi-square distribution does not provide an adequate approximation of the distribution of the Box-Pierce goodness-of-fit portmanteau statistic in the presence of nonindependent innovations. We thus propose modified portmanteau and Lagrange Multiplier (LM) tests whose asymptotic distributions are a weighted sums of independent chi-squared random variables. Monte Carlo experiments illustrate the finite sample performance of the different tests.

Keywords: Cointegration; Weak error process; Portmanteau tests; Lagrange multiplier test; Vector Error Correction Model.

1. Introduction

The use of Vector Error Correction Models (VECM) in the study of multivariate non stationary time series has become of increased interest in recent years. The main reason is that VECM allow to describe the long run relationships and the short run relationships of non-stationary variables (see Johansen (1995) or Lütkepohl (2005) for the statistical analysis and illustrations of the use of the VECM). In VECM, the cointegrating rank gives the number of independent linear stationary combinations of a multivariate nonstationary process. The autoregressive order gives the number of short run relations. In the literature, useful tools have been developed to analyze the long run relationships, in particular the identification of the cointegrating rank. However less attention has been given to the short run relationships in the VECM framework, in particular the autoregressive order. Tests for the cointegrating rank depend on the autoregressive order p. Then it is important to choose the correct autoregressive order for the overall understanding of the VECM.

Duchesne (2005) studied tests for serial correlation in the case of vector autoregressive models with exogenous variables when some variables are cointegrated. Tests based on the residual autocorrelations in the framework of the VECM were considered by Brüggemann, Lütkepohl and Saikkonen (2006). An important output of their work is that the asymptotic distribution of the portmanteau tests statistics depends on

* Postal address: 20, avenue des buttes de Coësmes, CS 14315, F-35043 Rennes Cedex, France. mail: hamdi.raissi@insa-rennes.fr.

the cointegrating rank. They also proposed Lagrange multiplier tests for cointegrated series. A crucial assumption of their work is that they consider iid innovations.

However the assumption that the innovations are iid is often considered to be too restrictive. Indeed many economic data present dependent innovations. Some recent works pointed out that, in some cases, the assumption of iid errors is not valid (see e.g. Francq, Roy and Zakoïan (2005) or Romano and Thombs (1996)). There are models which are martingale differences as the class of multivariate GARCH models (see e.g. Bollerslev (1990), Engle and Kroner (1995) or Jeantheau (1998)). Some models produce processes which are uncorrelated but not a martingale difference in general as the class of the all-pass models (see e.g. Andrews, Davis and Breidt (2006)). In this paper we will study estimation of the adjustment and short run parameters. We will propose portmanteau tests and a Lagrange multiplier test which take into account nonindependent but uncorrelated innovations.

Francq and Raïssi (2007) considered portmanteau tests for stationary autoregressive multivariate processes. They relaxed the assumption of iid innovations and adjusted the critical values of the portmanteau tests in consequence. They found that the usual chi-square approximation can be misleading in presence of nonindependent errors or when the roots of the autoregressive polynomial are close to one. We will extend their methodology in the context of cointegrated variables.

The structure of the paper is as follows. In Section 2, the model we use in this paper is presented. In Section 3 we derive the estimators of the parameters of our model and we give the asymptotic behaviour of the residual autocorrelations. Section 4 studies the asymptotic behaviour of the residual autocorrelations. We obtain in Section 5 the asymptotic distribution of the portmanteau test for VECM with nonindependent innovations. In Section 6 we study the asymptotic distribution of the LM test statistic. In Section 7 Monte Carlo experiments are performed. The proofs are relegated to the appendix.

In the sequel the following notations are used. Weak convergence is denoted by $\Rightarrow$ and we denote by $\stackrel{P}{\to}$ the convergence in probability. Considering a $d \times r$-dimensional matrix A, we define the orthogonal complement $A_\perp$, which is a full column rank $d \times (d-r)$ matrix such that $A'A_\perp = 0$. The symbol $\otimes$ denotes the usual Kronecker product and $\mathrm{vec}(A)$ denotes the vector obtained by stacking the columns of the matrix A. The trace of the matrix A is denoted by Tr.

2. Characterization of the model

We consider the following vector error correction model

$$\Delta X_t = \alpha_0 \beta_0' X_{t-1} + \sum_{i=1}^{p-1} \Gamma_{0i} \Delta X_{t-i} + \epsilon_t \tag{2.1}$$

where α_0 and β_0 are full column rank matrices of dimension $d \times r$. The process (X_t) is of dimension d and $\Delta X_t := X_t - X_{t-1}$. The Γ_{0i}'s, $i \in \{1, ..., p-1\}$, are $d \times d$ short run parameter matrices. By convention the sum in (2.1) vanishes when $p = 1$. Note that if $r = 0$ the relation (2.1) is a vector autoregressive model for the process (ΔX_t). The error process (ϵ_t) is usually assumed iid Gaussian. We will consider in the sequel a weaker assumption for the error process. Let us denote by $\mid . \mid$ the determinant of a square matrix. In the rest of the paper we shall assume that $\mid B(z) \mid = 0$ implies that

$| z |> 1$ or $z = 1$ and that $| \alpha_{0\perp}' \Gamma_0 \beta_{0\perp} | \neq 0$, where

$$B(z) = (1-z)I_d - \alpha_0 \beta_0' z - \sum_{i=1}^{p-1} \Gamma_{0i}(1-z)z^i \quad \text{and} \quad \Gamma_0 = I_d - \sum_{i=1}^{p-1} \Gamma_{0i}.$$

Under these assumptions it follows from Granger's representation theorem given in Johansen (1995 p 49) that

$$X_t = C \sum_{i=1}^{t} \epsilon_i + Y_t + A, \tag{2.2}$$

where $C = \beta_{0\perp}(\alpha_{0\perp}' \Gamma_0 \beta_{0\perp})^{-1} \alpha_{0\perp}'$. The vector A depends on initial values and is such that $\beta_0' A = 0$. The stationary process (Y_t) is of the form

$$Y_t = \sum_{i=0}^{\infty} \varphi_{0i} \epsilon_{t-i}, \tag{2.3}$$

where the power series $C(z) = \sum_{i=0}^{\infty} \varphi_{0i} z^i$ is convergent for $| z | \leq 1 + \kappa$ for some $\kappa > 0$. From (2.2) we have

$$\beta_0' X_t = \beta_0' Y_{t-1}, \tag{2.4}$$

and

$$\Delta X_t = C \epsilon_t + \Delta Y_t. \tag{2.5}$$

In general the iid Gaussian assumption on (ϵ_t) may appear too strong for tests based on the autocorrelations of the process ϵ_t. Indeed there are numerous models which can be adjusted to the error process and do not satisfy the strong assumption of Gaussian iid innovations as for instance the multivariate GARCH models or all-pass models. The statistical analysis of these models have a recent important development (see e.g. Bauwens, Laurent and Rombouts (2006) for the GARCH models or Andrews, Davis and Breidt (2006) for the all-pass models). In addition it is well known that there are many situations where the standard assumption is not satisfied (see e.g. Francq and Zakoïan (1998) in the univariate case). Therefore we will consider the following weaker assumption which is more appropriate in many cases.

Assumption A1 The error process (ϵ_t) is strictly stationary ergodic with finite positive definite covariance matrix Σ_ϵ, such that $E(\epsilon_t) = 0$ and $Cov(\epsilon_t, \epsilon_{t-h}) = 0$, for all $h \neq 0$.

This assumption allows us to consider a wider class of error processes than the usual iid white noise. A sequence (ϵ_t) satisfying **A1** will be called weak white noise. The results of Granger's representation theorem still hold when the assumption of iid Gaussian innovations is replaced by **A1**. Then from (2.4) we see that the process $(\beta_0' X_t)$ is strictly stationary. The cointegrating rank r corresponds to the dimension of the process $(\beta_0' X_t)$, that is the number of independent linear combinations of the components of the process (X_t) which are strictly stationary. From (2.5) the process ΔX_t is strictly stationary and we say that (X_t) is $I(1)$. It is interesting to note that the interpretation of some events can change in our framework. Indeed consider the projection P_α on the space spanned by the columns α_0 and the projection $P_{\alpha_\perp}$ on the orthogonal space of the space spanned by the columns of α_0. Therefore the independence between the permanent shocks $P_{\alpha_\perp}\epsilon_t$ and the transitory shocks $P_\alpha\epsilon_t$ is no longer ensured if (ϵ_t) is not Gaussian.

Let us denote by S_ϵ the diagonal matrix with diagonal elements $\sqrt{E(\epsilon_{it}^2)}$, where ϵ_{it} is the i-th coordinate of ϵ_t. We also define the theoretical autocorrelations at lag h of the error process $R_\epsilon(h) = S_\epsilon^{-1}\Gamma_\epsilon(h)S_\epsilon^{-1}$ with $\Gamma_\epsilon(h) := E(\epsilon_t\epsilon'_{t-h})$ for all $h \neq 0$. To check the adequacy of the autoregressive order of the model (2.1), it is a common practice to test the following pair of hypotheses

$$H_0: \quad R_\epsilon(h) = 0 \quad \text{vs.} \quad H_1: \exists\ h \geq 1 \text{ such that } \quad R_\epsilon(h) \neq 0.$$

Note that in (2.1) the parameters α_0 and β_0 are not identified. Then to get rid of this problem one can consider the following normalization

$$\beta_{0c} = \beta_0(c'\beta_0)^{-1} \quad \text{and} \quad \alpha_{0c} = \alpha_0\beta_0'c, \tag{2.6}$$

where the matrix c is of dimension $d \times r$ and is such that $c'\beta_0$ has full rank. We have $\alpha_0\beta_0' = \alpha_{0c}\beta_{0c}'$. Once α_0 and β_0 have been normalized in (2.1), one can derive the estimators of the adjustment and cointegration spaces.

3. Asymptotic behaviour of the estimators

In this section we derive the estimators and state the results we need for our tests. Note that in our framework we use the quasi maximum likelihood method since we relaxed the iid Gaussian assumption. Following the estimation procedure described in Johansen (1995) the estimator of β_{0c} can be obtained by reduced rank regression. Consider the observations $X_1, \dots, X_T$. Let $Z_{0t} = \Delta X_t$, $Z_{1t} = X_{t-1}$ and $Z_{2t} = (\Delta X'_{t-1}, \dots, \Delta X'_{t-p+1})'$ where $X_t = 0$ for $t \leq 0$. We define the matrices $M_{ij} = T^{-1}\sum_{t=1}^{T} Z_{it}Z'_{jt}$ and $S_{ij} = M_{ij} - M_{i2}M_{22}^{-1}M_{2j}$ for $i, j \in \{0, 1\}$. Considering eigenvectors $\hat{v}_1, \dots, \hat{v}_d$ of the corresponding eigenvalues $\hat{\mu}_1 \geq \cdots \geq \hat{\mu}_d > 0$ of the matrix $S_{11}^{-\frac{1}{2}}S_{10}S_{00}^{-1}S_{01}S_{11}^{-\frac{1}{2}}$, we set

$$\hat{\beta} = S_{11}^{-\frac{1}{2}}(\hat{v}_1, \dots, \hat{v}_r).$$

For our asymptotic results we shall assume in the sequel that the cointegrating rank is well fitted. The normalized estimator of β_{0c} is given by $\hat{\beta}_c = \hat{\beta}(c'\hat{\beta})^{-1}$. Let us define the strong mixing coefficients $\alpha_a(h)$ for a given stationary process (a_t)

$$\alpha_a(h) = \sup_{A\in\sigma(a_u, u\leq t), B\in\sigma(a_u, u\geq t+h)} |P(A\cap B) - P(A)P(B)|,$$

which measures the temporal dependence of the process (a_t). We also define $\|a_t\|_q = (E\|a_t\|^q)^{1/q}$, where $\|.\|$ denotes the Euclidean norm with $E\|a_t\|^q < \infty$ and $q \geq 1$. Consider the following additional assumption.

Assumption A2 The process (ϵ_t) satisfies $\|\epsilon_t\|_{4+2\nu} < \infty$, and the mixing coefficients of the process (ϵ_t) are such that $\sum_{h=0}^{\infty}\{\alpha_\epsilon(h)\}^{\nu/(2+\nu)} < \infty$ for some $\nu > 0$.

The mixing assumption is not very restrictive in general. For instance note that **A2** is satisfied for exponential strongly mixing sequences. It is shown in Raïssi (2009) that under **A1** and **A2** we have

$$\hat{\beta}_c = \beta_{0c} + O_P(T^{-1}).$$

We will consider estimators of the parameters α_{0c} and the Γ_{0i}'s obtained as follows. Define $\theta_0 = (\alpha_{0c}, \Gamma_{01}, \dots, \Gamma_{0p-1})$ and $\xi_{t-1}(\beta) = ((\beta' Z_{1t})', Z_{2t}')'$, where β is a $d \times r$ dimensional matrix. The equation (2.1) becomes

$$Z_{0t} = \theta_0 \xi_{t-1}(\beta_{0c}) + \epsilon_t. \tag{3.1}$$

Then using (3.1), the least squares estimator of θ_0 is given by

$$\hat{\theta}(\beta_{0c}) = \hat{\Sigma}_z(\beta_{0c})\{\hat{\Sigma}_\xi(\beta_{0c})\}^{-1}, \tag{3.2}$$

where

$$\hat{\Sigma}_z(\beta) = T^{-1} \sum_{t=1}^{T} Z_{0t} \xi_{t-1}'(\beta) \quad \text{and} \quad \hat{\Sigma}_\xi(\beta) = T^{-1} \sum_{t=1}^{T} \xi_{t-1}(\beta) \xi_{t-1}'(\beta).$$

Note that the estimator $\hat{\theta}(\beta_{0c})$ is unfeasible since it depends on the unknown parameter β_{0c}. Since the processes $(\beta_{0c}' Z_{1t})$ and (Z_{0t}) are strictly stationary ergodic processes, $(\xi_t(\beta_{0c}))$ is also a strictly stationary ergodic process. Therefore the matrices $\hat{\Sigma}_z(\beta_{0c})$ and $\hat{\Sigma}_\xi(\beta_{0c})$ converge. We let

$$\hat{\theta} = \hat{\theta}(\hat{\beta}_c). \tag{3.3}$$

It is easy to see that under **A1** and **A2**, we have

$$\hat{\theta} = \hat{\theta}(\beta_{0c}) + o_P(T^{-\frac{1}{2}}). \tag{3.4}$$

The following proposition gives the asymptotic normality of the estimators of the parameter α_{0c} and the Γ_{0i}'s.

Proposition 3.1. *Under assumptions* **A1**-**A2**, *we have*

$$T^{\frac{1}{2}} \mathit{vec}\,(\hat{\theta} - \theta_0) \Rightarrow \mathcal{N}(0, \Sigma_\theta), \tag{3.5}$$

where

$$\Sigma_\theta = \sum_{h=-\infty}^{\infty} E\left\{ \Sigma_\xi(\beta_{0c})^{-1} \xi_{t-1}(\beta_{0c}) \xi_{t-h-1}'(\beta_{0c}) \Sigma_\xi(\beta_{0c})^{-1} \otimes \epsilon_t \epsilon_{t-h}' \right\}, \tag{3.6}$$

and $\Sigma_\xi(\beta_{0c}) = E(\xi_t(\beta_{0c}) \xi_t'(\beta_{0c}))$.

It is obvious that the asymptotic variance matrix Σ_θ depends on the choice of the normalization. When the process (ϵ_t) is iid we find $\Sigma_\theta = \Sigma_\xi(\beta_{0c})^{-1} \otimes \Sigma_\epsilon$. Note that the expression of the covariance matrix in the standard case can be very different from the expression (3.6).

4. Asymptotic distribution of the residual autocorrelations

Now we will establish the asymptotic normality of the residual autocorrelations. Define the residuals $\hat{\epsilon}_t = Z_{0t} - \hat{\theta}\xi_{t-1}(\hat{\beta}_c)$. Note that the residuals do not depend on the normalization. Let us consider the residual autocovariances

$$\hat{\Gamma}_\epsilon(h) = \frac{1}{T}\sum_{t=h+1}^{T} \hat{\epsilon}_t \hat{\epsilon}'_{t-h}, \quad \text{for} \quad 0 \leq h < T.$$

For a fixed integer m we define

$$\hat{\gamma}_m = \text{vec}\left\{\left(\hat{\Gamma}_\epsilon(1), \dots, \hat{\Gamma}_\epsilon(m)\right)\right\}.$$

Denote by $\hat{S}_\epsilon$ the diagonal matrix with, on the diagonal, the square roots of the diagonal elements of $\hat{\Gamma}_\epsilon(0)$. The residual autocorrelations at lag h are given by $\hat{R}_\epsilon(h) = \hat{S}_\epsilon^{-1}\hat{\Gamma}_\epsilon(h)\hat{S}_\epsilon^{-1}$. We consider the following vector of the m first residual autocorrelations

$$\hat{\rho}_m = \left\{I_m \otimes \left(\hat{S}_\epsilon \otimes \hat{S}_\epsilon\right)^{-1}\right\}\hat{\gamma}_m.$$

Define

$$\Phi_m = -E\{w_t \otimes \xi'_{t-1}(\beta_{0c}) \otimes I_d\},$$

where $w_t = (\epsilon'_{t-1}, \dots, \epsilon'_{t-m})'$. It can be shown that

$$\Phi_m = -\sum_{i=1}^{m}[\{e_m(i)e_p(1)' + e_m(i)e_p(2)' \times (1 - \mathbf{1}_{\{p=1\}})\} \otimes \Sigma_\epsilon](A^{i-1})'B' \otimes I_d, \tag{4.1}$$

where $e_m(i)$ denotes the m-dimensional vector with 1 as i-th element and 0 elsewhere, and $\mathbf{1}_{\{p=1\}} = 1$ for $p = 1$ and 0 for $p \neq 1$. The matrices B and A are given by

$$B = \begin{pmatrix} \beta'_{0c} & 0 & \dots & 0 \\ 0 & I_d & & \vdots \\ \vdots & & \ddots & \\ 0 & \dots & & I_d \end{pmatrix} \quad \text{and} \quad A = \begin{pmatrix} I + \alpha_{0c}\beta'_{0c} & \Gamma_{01} & \dots & & \Gamma_{p-1} \\ \alpha_{0c}\beta'_{0c} & \Gamma_{01} & \dots & & \Gamma_{p-1} \\ 0 & I_d & 0 & \dots & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & \dots & 0 & I_d & 0 \end{pmatrix}.$$

We also define

$$\Xi = \begin{pmatrix} \Sigma_{u_m} & \Sigma_{u_m,\theta} \\ \Sigma'_{u_m,\theta} & \Sigma_\theta \end{pmatrix} = \sum_{h=-\infty}^{\infty} E\Upsilon_t\Upsilon'_{t-h}, \quad \text{where} \quad \Upsilon_t = \begin{pmatrix} u_t \\ v_t \end{pmatrix}, \tag{4.2}$$

with $u_t = w_t \otimes \epsilon_t$ and $v_t = \Sigma_\xi(\beta_{0c})^{-1}\xi_{t-1}(\beta_{0c}) \otimes \epsilon_t$. The following proposition gives the limiting distribution of the residual autocovariances and autocorrelations.

Proposition 4.1. *Under assumptions* **A1**-**A2**, *we have*

$$\sqrt{T}\hat{\gamma}_m \Rightarrow \mathcal{N}(0, \Sigma_{\hat{\gamma}_m}), \tag{4.3}$$

where

$$\Sigma_{\hat{\gamma}_m} = \Sigma_{u_m} + \Phi_m \Sigma_\theta \Phi'_m + \Sigma_{u_m,\theta}\Phi'_m + \Phi_m \Sigma'_{u_m,\theta}. \tag{4.4}$$

In addition we have

$$\sqrt{T}\hat{\rho}_m \Rightarrow \mathcal{N}(0, \Sigma_{\hat{\rho}_m}), \tag{4.5}$$

where

$$\Sigma_{\hat{\rho}_m} = \left\{ I_m \otimes (S_\epsilon \otimes S_\epsilon)^{-1} \right\} \Sigma_{\hat{\gamma}_m} \left\{ I_m \otimes (S_\epsilon \otimes S_\epsilon)^{-1} \right\}.$$

Note that $\Sigma_{\hat{\gamma}_m}$ is invariant to the choice of the normalization matrix. In the strong case, we have $\Sigma_{u_m} = Eu_t u'_t = I_m \otimes \Sigma_\epsilon \otimes \Sigma_\epsilon$, $\Phi_m \Sigma_\theta \Phi'_m = \Phi_m \left\{ \Sigma_\xi(\beta_{0c})^{-1} \otimes \Sigma_\epsilon \right\} \Phi'_m$. Using the relation $(F \otimes G)(C \otimes D) = (FC) \otimes (GD)$ for matrices of appropriate dimension we also obtain

$$\Sigma_{u_m,\hat{\theta}_n} = -\Phi_m \Sigma_\theta.$$

Thus $\Sigma_{\hat{\gamma}_m} = \Sigma_{u_m} - \Phi_m \Sigma_{\hat{\theta}} \Phi'_m$. Straightforward computations show that

$$\Sigma_{\hat{\gamma}_m} = I_m \otimes \Sigma_\epsilon \otimes \Sigma_\epsilon - E\left[w_t \xi'_{t-1}(\beta_{0c}) \right] \Sigma_\xi(\beta_{0c})^{-1} E\left[w_t \xi'_{t-1}(\beta_{0c}) \right]' \otimes \Sigma_\epsilon, \tag{4.6}$$

so that we retrieve the result obtained by Brüggemann *et al* (2006) in the strong case. Note that the expression of the covariance matrix in (4.4) can be very different from the one of the standard case in (4.6).

5. Portmanteau test

In this section we consider the portmanteau tests to check the null hypothesis of uncorrelated errors. These tests were introduced in the univariate ARMA framework by Box and Pierce (1970) (BP hereafter) and Ljung and Box (1978) (LB hereafter). Basic forms of the multivariate version of the BP portmanteau test statistic are

$$\begin{aligned} Q_m &= T \sum_{h=1}^{m} \mathrm{Tr}\left(\hat{\Gamma}'_\epsilon(h) \hat{\Gamma}_\epsilon^{-1}(0) \hat{\Gamma}_\epsilon(h) \hat{\Gamma}_\epsilon^{-1}(0) \right) \\ &= T \sum_{h=1}^{m} \mathrm{Tr}\left(\mathrm{vec}\, \hat{\Gamma}_\epsilon(h) \right)' \left(\hat{\Gamma}_\epsilon^{-1}(0) \otimes \hat{\Gamma}_\epsilon^{-1}(0) \right) \left(\mathrm{vec}\, \hat{\Gamma}_\epsilon(h) \right) \\ &= T \hat{\gamma}'_m \left(I_m \otimes \hat{\Gamma}_\epsilon^{-1}(0) \otimes \hat{\Gamma}_\epsilon^{-1}(0) \right) \hat{\gamma}_m \\ &= T \hat{\rho}'_m \left(I_m \otimes \hat{R}_\epsilon^{-1}(0) \otimes \hat{R}_\epsilon^{-1}(0) \right) \hat{\rho}_m. \end{aligned}$$

This test statistic was introduced by Chitturi (1974) in the framework of vector autoregressive models. Similarly to the univariate LB portmanteau test statistic, Hosking (1980) defined the modified portmanteau statistic

$$\tilde{Q}_m = T^2 \sum_{h=1}^{m} (T-h)^{-1} \mathrm{Tr}\left(\hat{\Gamma}'_\epsilon(h) \hat{\Gamma}_\epsilon^{-1}(0) \hat{\Gamma}_\epsilon(h) \hat{\Gamma}_\epsilon^{-1}(0) \right),$$

which has better small sample properties when the error process is Gaussian. We derive the asymptotic distribution of the portmanteau tests statistics in our framework using the results of the previous section. Under the assumption that (ϵ_t) is an iid white noise Brüggemann *et al* (2006) found that the approximate distribution of the statistics Q_m and $\tilde{Q}_m$ is given by

$$Q_m \approx \chi^2(d^2(m-p+1)-dr). \tag{5.1}$$

Note that the distribution (5.1) depends on the cointegrating rank r. In the sequel the versions of the portmanteau tests proposed by Brüggemann *et al* (2006) will be denoted by BP_S and LB_S. The following proposition is a direct consequence of Proposition 4.1.

Proposition 5.1. *Under assumptions* **A1** *and* **A2**, *the statistics* Q_m *and* $\tilde{Q}_m$ *converge in distribution, as* $T \to \infty$, *to*

$$U_m(\delta_m) = \sum_{i=1}^{d^2 m} \delta_i U_i^2 \tag{5.2}$$

where $\delta_m = (\delta_1, \cdots, \delta_{d^2 m})'$ *is the vector of the eigenvalues of the matrix*

$$\Lambda_m = \left(I_m \otimes \Sigma_\epsilon^{-1/2} \otimes \Sigma_\epsilon^{-1/2}\right) \Sigma_{\hat{\gamma}_m} \left(I_m \otimes \Sigma_\epsilon^{-1/2} \otimes \Sigma_\epsilon^{-1/2}\right)$$

and the U_i*'s are independent* $\mathcal{N}(0,1)$ *variables.*

This Proposition shows that the asymptotic distribution of Q_m and $\tilde{Q}_m$ is a weighted sum of chi-squares. Contrary to the approximation in (5.1), the asymptotic distribution we obtain depends on the cointegrating relations since β_{0c} appears in the expression of the matrix $\Sigma_{\hat{\gamma}_m}$ in (4.4). However the distribution in (5.2) does not depend on the choice of the normalization matrix. Note also that the distribution (5.2) is given for fixed m, while the approximation (5.1) is obtained assuming that $m \to \infty$ as $T \to \infty$. In addition the tests we propose can be performed when $d^2(m-p+1)-dr \leq 0$. The following example shows that, for the asymptotic distribution of Q_m and $\tilde{Q}_m$, the $\chi^2(d^2(m-p+1)-dr)$ approximation is not valid in the framework of VECM with dependent but uncorrelated errors.

Example 5.1. Consider the following bivariate VECM with $p=2$ and $r=0$ with true parameter $\Gamma_{01}=0$

$$\Delta X_t = \Gamma_{01} \Delta X_{t-1} + \epsilon_t. \tag{5.3}$$

In this case there is no cointegration and the process (X_t) is a random walk. We can consider model (5.3) for instance to test linear Granger causality in mean relations between the components of (ΔX_t). In this case one has to check that the error process is a white noise in a first step.

Suppose that the innovation process (ϵ_t) is an ARCH(1) model with constant correlation proposed by Jeantheau (1998):

$$\begin{pmatrix} \epsilon_{1t} \\ \epsilon_{2t} \end{pmatrix} = \begin{pmatrix} \sigma_{11t} & 0 \\ 0 & \sigma_{22t} \end{pmatrix} \begin{pmatrix} \eta_{1t} \\ \eta_{2t} \end{pmatrix}$$

where

$$\begin{pmatrix} \sigma_{11t}^2 \\ \sigma_{22t}^2 \end{pmatrix} = \begin{pmatrix} 0.3 \\ 0.2 \end{pmatrix} + \begin{pmatrix} b_{11} & 0 \\ b_{21} & b_{22} \end{pmatrix} \begin{pmatrix} \epsilon_{1t-1}^2 \\ \epsilon_{2t-1}^2 \end{pmatrix}.$$

Assuming for simplicity that the variance of the Gaussian iid process $(\eta_{1t}, \eta_{2t})'$ is I_2, we obtain when $m=2$:

	Non zero eigenvalues of Λ_m	Distribution of $U_m\ (\delta_m)$
$b_{11} = b_{21} = b_{22} = 0$	(1,1,1,1)	χ^2_4
$b_{11} = 0.4$, $b_{21} = 0.15$, $b_{22} = 0.25$	(1.62, 1.52, 1.02, 1.51)	$1.62\chi^2_1 + 1.52\chi^2_1 + 1.02\chi^2_1 + 1.51\chi^2_1$

This table shows that the asymptotic distribution of the goodness-of-fit portmanteau tests may be quite different for VECM with ARCH innovations and VECM with iid the innovations.

The different matrices involved in the computation of $\hat{\delta}_m = (\hat{\delta}_1, \cdots, \hat{\delta}_{d^2m})'$ can be estimated as follows. A consistent estimator $\hat{\Phi}_m$ of the matrix Φ_m is obtained by replacing β_{0c}, α_{0c}, Σ_ϵ and the Γ_{0i}'s by their estimators in (4.1). We compute a consistent estimator of Ξ defined in (4.2) using the following method. Let us define $\hat{\Upsilon}_t = (\hat{u}'_t, \hat{v}'_t)'$ with $\hat{u}_t = \hat{w}_t \otimes \hat{\epsilon}_t$, $\hat{w}_t = (\hat{\epsilon}'_{t-1}, \dots, \hat{\epsilon}'_{t-m})'$ and $\hat{v}_t = \Sigma_\xi(\hat{\beta}_c)^{-1}\xi_{t-1}(\hat{\beta}_c) \otimes \hat{\epsilon}_t$. We also define $\hat{\mathcal{A}}_q(z) = I_{md^2+rd+d^2(p-1)} - \sum_{i=1}^q \hat{\mathcal{A}}_{q,i} z^i$, where $\hat{\mathcal{A}}_{q,1}, \cdots, \hat{\mathcal{A}}_{q,q}$ denote the coefficients of the LS regression of $\hat{\Upsilon}_t$ on $\hat{\Upsilon}_{t-1}, \cdots, \hat{\Upsilon}_{t-q}$. Let $\tilde{\epsilon}_{q,t}$ be the residuals of this regression and $\hat{\Sigma}_{\tilde{\epsilon}_q} = T^{-1}\sum_{t=1}^T \tilde{\epsilon}_{q,t}\tilde{\epsilon}'_{q,t}$. Under other additional assumptions it can be shown that

$$\hat{\Xi} := \hat{\mathcal{A}}_q^{-1}(1)\hat{\Sigma}_{\tilde{\epsilon}_q}\hat{\mathcal{A}}_q'^{-1}(1) \xrightarrow{P} \Xi$$

when $q = q(T) \to \infty$ and $q^3/T \to 0$ as $T \to \infty$. The order q can be chosen by considering an information criterion. Using $\hat{\Xi}$, $\hat{\Phi}_m$ and (4.4), one can compute a consistent estimator $\hat{\Sigma}_{\hat{\gamma}_m}$ of the matrix $\Sigma_{\hat{\gamma}_m}$. In addition from the consistency of $\hat{\theta}$ and the ergodic theorem, it is clear that we have $\hat{\Sigma}_\epsilon := \hat{\Gamma}_\epsilon(0) = \Sigma_\epsilon + o_p(1)$. Therefore we define

$$\hat{\Lambda}_m = \left(I_m \otimes \hat{\Sigma}_\epsilon^{-1/2} \otimes \hat{\Sigma}_\epsilon^{-1/2}\right)\hat{\Sigma}_{\hat{\gamma}_m}\left(I_m \otimes \hat{\Sigma}_\epsilon^{-1/2} \otimes \hat{\Sigma}_\epsilon^{-1/2}\right),$$

which is such that $\hat{\Lambda}_m = \Lambda_m + o_p(1)$. The components of $\hat{\delta}_m$ are given by the eigenvalues of $\hat{\Lambda}_m$.

At the asymptotic level υ, the $\mathrm{LB_W}$ test (resp. the $\mathrm{BP_W}$ test) consists in rejecting the null hypothesis of uncorrelated errors when

$$P\{U_m(\hat{\delta}_m) > Q_m\} < \upsilon \quad \text{(resp. } P\{U_m(\hat{\delta}_m) > \tilde{Q}_m\} < \upsilon), \tag{5.4}$$

The p-values in (5.4) can be evaluated using the Imhof algorithm (Imhof (1961)) or the saddlepoint approximation method (Kuonen (1999)). The reader is referred to Francq and Raïssi (2007, Section 4) for more details on the implementation of the BP_W and LB_W tests.

6. Lagrange multiplier test

Let us consider the following VAR model for the error process

$$\epsilon_t = \sum_{i=1}^m \psi_{0i}\epsilon_{t-i} + e_t,$$

where (e_t) is a white noise. Then the regression model (2.1) can be written as follows

$$\begin{aligned} Z_{0t} &= \theta_0\xi_{t-1}(\beta_{0c}) + \sum_{i=1}^{m}\psi_{0i}\epsilon_{t-i} + e_t \\ &= (\zeta'_{t-1}(\beta_{0c}) \otimes I_d)\psi_0 + e_t, \end{aligned} \tag{6.1}$$

with $\zeta_{t-1}(\beta) = (\xi'_{t-1}(\beta), \epsilon'_{t-1}, \dots, \epsilon'_{t-m})'$, $\psi_0 = (\text{vec}\,(\theta_0)', \psi_0'^{*})'$ and $\psi_0^* = \text{vec}\,(\psi_{01}, \dots, \psi_{0m})$. We define the matrix $R = (0_{d^2m\times(dr+d^2(p-1))}, I_{d^2m})$. Under the null hypothesis $R\psi_0 = 0$ and we have $\epsilon_t = e_t$.

Let us define the constrained estimator $\hat{\psi}^c = (\text{vec}\,(\hat{\theta})', 0, \dots, 0)'$. The score vector at the point $(\hat{\beta}_c, \hat{\psi}^c, \hat{\Sigma}_\epsilon)$ is such that

$$R\frac{\partial \ln(\mathcal{L}(\hat{\beta}_c, \hat{\psi}^c, \hat{\Sigma}_\epsilon))}{\partial \psi} = T(I_{dm} \otimes \hat{\Sigma}_\epsilon^{-1})\hat{\gamma}_m,$$

where $\ln(\mathcal{L}(\beta, \psi, \Sigma_\epsilon))$ is the log-likelihood function of the model (6.1). We introduce the modified LM test statistic

$$Q_{LM_w} = T\hat{\gamma}'_m(I_{dm} \otimes \hat{\Sigma}_\epsilon^{-2})\hat{\gamma}_m.$$

The following proposition is a consequence of Proposition 4.1.

Proposition 6.1. *Under assumptions* **A1** *and* **A2**, *the statistic Q_{LM_w} converges in distribution, as $T \to \infty$, to*

$$\check{U}_m\,(\iota_m) = \sum_{k=1}^{d^2m}\iota_k U_k^2 \tag{6.2}$$

where $\iota_m = (\iota_1, \cdots, \iota_{d^2m})$ is the vector of the eigenvalues of the matrix

$$\check{\Lambda}_m = \left(I_{dm} \otimes \Sigma_\epsilon^{-1}\right)\Sigma_{\hat{\gamma}_m}\left(I_{dm} \otimes \Sigma_\epsilon^{-1}\right)$$

and the U_k's are independent $\mathcal{N}(0,1)$ variables.

Similarly to the portmanteau tests, we find that the asymptotic distribution of the modified LM statistic is a weighted sum of chi-squares. However the weights in (6.2) are different from those in (5.2) in general. Note that the asymptotic distribution of the modified LM test depends on the cointegrating relations. At the asymptotic level υ, the LM_W test test rejects the null hypothesis when $P\{U_m(\hat{\iota}_m) > Q_{LM_w}\} < \upsilon$. The vector $\hat{\iota}_m$ is obtained by considering the matrix

$$\tilde{\Lambda}_m = \left(I_{dm} \otimes \hat{\Sigma}_\epsilon^{-1}\right)\hat{\Sigma}_{\hat{\gamma}_m}\left(I_{dm} \otimes \hat{\Sigma}_\epsilon^{-1}\right).$$

The modified LM test, denoted by LM_W, can be implemented in a similar way to the modified portmanteau tests.

Using the Breusch-Godfrey (Breusch (1978), Godfrey (1978)) approach we can consider the auxiliary model

$$\hat{\epsilon}_t = (\hat{\zeta}'_{t-1}(\hat{\beta}_c) \otimes I_d)\tilde{\psi}_0 + \tilde{e}_t,$$

with $\hat{\zeta}_{t-1}(\beta) = (\xi'_{t-1}(\beta), \hat{\epsilon}'_{t-1}, \dots, \hat{\epsilon}'_{t-m})'$, $\tilde{\psi}_0 = (\text{vec}\,(\theta_0), \tilde{\psi}_0^*)$ and $\tilde{\psi}_0^* = \text{vec}\,(\tilde{\psi}_{01}, \dots, \tilde{\psi}_{0m})$. Recall that the residuals are given by $\hat{\epsilon}_t = Z_{0t} - \hat{\theta}\xi_{t-1}(\hat{\beta}_c)$. The generalized least squares estimator of $\tilde{\psi}_0$ is

$$\begin{aligned}
\hat{\tilde{\psi}} &= \{T^{-1}\sum_{t=1}^{T}\hat{\zeta}_{t-1}(\hat{\beta}_c)\hat{\zeta}'_{t-1}(\hat{\beta}_c)\otimes I_d\}^{-1}\{T^{-1}\sum_{t=1}^{T}(\hat{\zeta}_{t-1}(\hat{\beta}_c)\otimes I_d)\hat{\epsilon}_t\} \\
&= \{T^{-1}\sum_{t=1}^{T}\hat{\zeta}_{t-1}(\hat{\beta}_c)\hat{\zeta}'_{t-1}(\hat{\beta}_c)\otimes I_d\}^{-1}\text{vec}\,(T^{-1}\sum_{t=1}^{T}\hat{\epsilon}_t\hat{\zeta}'_{t-1}(\hat{\beta}_c)).
\end{aligned}$$

Since $T^{-1}\sum_{t=1}^{T}\hat{\epsilon}_t\hat{\xi}'_{t-1}(\hat{\beta}_c) = 0$, we obtain

$$\hat{\tilde{\psi}}^* = (R\hat{J}^{-1}R')(I_{dm}\otimes\hat{\Sigma}_\epsilon^{-1})\hat{\gamma}_m.$$

Therefore in the standard case the following statistic is considered for the Breusch-Godfrey test

$$Q_{LM_s} = T\hat{\tilde{\psi}}'^*(R\hat{J}^{-1}R')^{-1}\hat{\tilde{\psi}}^*,$$

where

$$\hat{J} = \frac{\partial\ln(\mathcal{L}(\hat{\beta}_c, \hat{\psi}^c, \hat{\Sigma}_\epsilon))}{\partial\psi\partial\psi'} = T^{-1}\sum_{t=1}^{T}\zeta_{t-1}(\hat{\beta}_c)\zeta'_{t-1}(\hat{\beta}_c)\otimes\hat{\Sigma}_\epsilon^{-1}$$

is a consistent estimator of $J = E(\zeta_{t-1}(\hat{\beta}_c)\zeta'_{t-1}(\hat{\beta}_c))\otimes\Sigma_\epsilon^{-1}$. In the sequel when the Q_{LM_s} statistic is considered it is assumed that J is non-singular, so that $\hat{J}$ is invertible at least asymptotically. Under standard assumptions, Brüggemann *et al* (2006) showed that

$$Q_{LM_s} \Rightarrow \chi^2_{d^2m}. \tag{6.3}$$

The test proposed in Brüggemann *et al* (2006) will be denoted LM_S. The asymptotic distribution in (6.3) does not depend on the cointegrating relations while the statistic Q_{LM_s} depends on the cointegrating relations. Note that the statistics Q_{LM_w} and Q_{LM_s} depend on the number of residual autocorrelations m. Contrary to the modified and standard portmanteau tests, the statistics of the modified and standard LM tests are different. Note also that following the common approach we only use the first m residual autocorrelations to test H_0 in the different tests we consider in this paper.

7. Monte Carlo experiments

In this section we compare the small sample properties of the BP_S, LB_S, LM_S tests and the BP_W, LB_W, LM_W tests with bivariate and 3-dimensional simulated processes. We simulated $n = 1000$ independent trajectories of length $T = 100$, $T = 400$ and $T = 1000$ using parameters given in Table 1. Parameters (a) produce bivariate processes. We obtain 3-dimensional processes using parameters (b). Parameters (a) are chosen in the interior of the parameter space. Parameters (b) correspond to adjustment, long run and short run parameters considered in the simulation experiments in Brüggemann *et al* (2006). In practice the cointegrating rank is not known and can be tested using the likelihood ratio test given in Johansen (1995). However since we focus on the study of the short run relationships when the errors are dependent, our Monte Carlo experiments are carried out using the true cointegrating rank.

In this part we study the empirical size of the tests under comparison. We consider the case of iid innovations and the case of dependent but uncorrelated innovations. We first use for our study bivariate processes obtained using parameters (a) and test the null hypothesis at the asymptotic nominal level 5%. The error process is normally distributed with mean zero and variance matrix I_2 in the iid case. To assess the finite sample properties of the tests in presence of weak errors, we use the following ARCH innovations with constant correlation presented in Example 5.1,

$$\begin{pmatrix} \epsilon_{1t} \\ \epsilon_{2t} \end{pmatrix} = \begin{pmatrix} \sigma_{1t} & 0 \\ 0 & \sigma_{2t} \end{pmatrix} \begin{pmatrix} \eta_{1t} \\ \eta_{2t} \end{pmatrix} \tag{7.1}$$

where

$$\begin{pmatrix} \sigma_{1t}^2 \\ \sigma_{2t}^2 \end{pmatrix} = \begin{pmatrix} 0.1 \\ 0.1 \end{pmatrix} + \begin{pmatrix} 0.3 & 0.1 \\ 0.2 & 0.3 \end{pmatrix} \begin{pmatrix} \epsilon_{1t-1}^2 \\ \epsilon_{2t-1}^2 \end{pmatrix}$$

and the iid process $\eta_t = (\eta_{1t}, \eta_{2t})'$ is such that $\eta_t \sim \mathcal{N}(0, I_2)$. In this case the process (ϵ_t) is a martingale difference and presents conditional heteroscedasticity.

The results for processes generated using parameters (a) are presented in Table 2 when the innovations are iid, and in Table 3 when the error processes follow the weak white noise (7.1). For $n = 1000$ replications and under the hypothesis that the finite sample size of the tests is 5%, the relative rejection frequency should be between the significant limits 3.65% and 6.35% with probability 0.95. Therefore the relative rejection frequencies are displayed in bold type when they are outside the 5% significant limits 3.65% and 6.35%.

In Table 2 the results of the standard tests are better than those of the modified tests for samples of size $T = 100$. A possible explanation is that, on contrary to the standard tests, the computation of the critical values for the modified tests requires estimating parameters. Moreover a high dimension process is used for the estimation of Ξ. However for larger samples the results are similar for the tests under comparison. The empirical sizes decrease for large m.

For the weak white noise when $T = 100$, see Table 3, the same comments can be made as in the iid case. For large samples the standard tests are clearly oversized. From our simulation results the chi-square approximation of the standard portmanteau tests does not seem more satisfactory when the innovations are not independent. In general the results for the BP_S, LB_S and LM_S tests can be explained by the fact that dependent errors are not taken into account in the standard theoretical framework. When the sample is large, the modified tests perform well.

In order to study the behaviour of the different tests when the dimension increases, we consider parameters (b). The errors are iid normally distributed with variance matrix I_3. From Table 4 we can draw the conclusion that, in general, the modified tests perform much better than the standard tests for small m. For a large number of residual autocorrelations ($m = 20$), the modified tests seem to be more conservative than the standard tests.

In the empirical power part of this section, we study the ability of the different tests to detect the serial correlation of the error terms. We consider bivariate processes obtained using parameters (a) and test the null hypothesis at the asymptotic nominal level 5%. In a first experiment we use the following correlated error process

$$\epsilon_t = \cos(0.5 \arcsin(2\omega_1))\eta_t + \sin(0.5 \arcsin(2\omega_1))\eta_{t-1}.$$

It is easy to check that $Var(\epsilon_t) = I_2$ and $Corr(\epsilon_t, \epsilon_{t-1}) = \omega_1 I_2$. The different tests are performed for $m = 3$ and $m = 10$. We only simulated samples of length $T = 100$ and $T = 1000$. The results for the LB_W and LB_S tests are displayed in Figures 10.1 and 10.2. The results for the BP tests, which are similar, are not displayed. The results for the LM_W and LM_S tests are displayed in Figures 10.3 and 10.4. To study the case where the errors are correlated at higher lag, we use the following errors

$$\epsilon_{kt} = \eta_{kt} + \omega_2 \sum_{i=1}^{5} \eta_{kt-i}, \tag{7.2}$$

with $k = 1, 2$. In this case ϵ_t and ϵ_{t-i} are correlated for $i \leq 5$ and independent for $i > 5$ when $\omega_2 \neq 0$. We only displayed the results for $m = 10$ and $T = 1000$ in Table 5. Several values of ω_1 and ω_2 are considered in our experiments.

From Figure 10.1 we note that for small samples and large m the LB_S is more powerful than the LB_W. When m is small and the samples are small, the power of the LB_W and LB_S are similar. From Figure 10.3 the same remarks can be made for the LM_S and LM_W tests. This confirms that, in general, the standard tests have better finite sample properties than the modified tests when the samples are small. However we can remark from Figures 10.2 and 10.4 that the power of the different tests are similar for large samples. According to the results of Table 5, we can draw the same conclusion when the errors are correlated at higher lag. Therefore when the samples are large, it appears that the standard tests have no power advantage on the modified tests. As expected when we increase m, the power decreases noticeably for the different tests.

Some final remarks on the number of residual autocorrelations have to be made. When the samples are large, the rejection frequencies of the modified tests seem to be closer to the asymptotic nominal level for $m = 1, 5, 10$ than the case where we take $m = 20$ in Tables 2 and 3. Therefore we recommend to choose a small number of residual autocorrelations for the modified tests. Note also that the over rejections for small m of the standard portmanteau tests when the errors are dependent in Table 3 are not due to the choice of m. Indeed we can see from Table 2 that the standard portmanteau tests well control the error of first kind when the innovations are independent for $m = 5, 10$.

8. Conclusion

In this work we considered VECM with uncorrelated but nonindependent innovations. There exist numerous models in the literature which can be adjusted to the error process and do not satisfy the iid assumption. Thus it is important to consider innovations which are only uncorrelated. In this framework, we established the asymptotic normality of the estimators of the short run and adjustment parameters.

We also proposed a methodology to test the uncorrelatedness of dependent innovations when the variables are cointegrated. The tests introduced in this paper are a little more sophisticated than the standard tests. For this reason we do not recommend to use these tests when the sample is small. However we found that the usual chi-square approximation for the portmanteau tests is no more valid when the errors are not independent. Similarly the Lagrange multiplier test based on standard arguments seems not valid in presence of dependent errors for large samples. The tests we propose are built taking into account the possible dependence of the errors. In addition from the simulation results, the sample performances of the proposed tests are in general better or similar to those of the standard tests for large samples. Finally it should be noted that we often have no evidence on the independence of the innovations in practice. Thus we can draw the conclusion that it seems preferable to use the modified tests we propose when the samples are large.

9. Proofs

In order to prove Propositions 3.1 and 4.1, we have to state some intermediate asymptotic results. First we will state the following Lemma in which we use the mixing properties of the process (ϵ_t).

Lemma 1. *Under* **A1** *and* **A2** *we have*

$$\sup_{i,j} \sum_{h=-\infty}^{+\infty} \mid Cov(\epsilon_{m_1 t}\epsilon_{m_2 t-i}, \epsilon_{m_1' t-h}\epsilon_{m_2' t-j-h}) \mid < \infty,$$

where $m_1, m_2, m_1', m_2' \in \{1, \dots, d\}$.

Proof of Lemma 1. The result of Lemma 1 is a consequence of the Davydov inequality (Davydov (1968)) and is proved in Raïssi (2009).

Now define the linear process $V_t = \sum_{i=0}^{\infty} D_i \epsilon_{t-i}$, where $D(z) = \sum_{i=0}^{\infty} D_i z^i$ is convergent for $\mid z \mid \leq 1+\kappa$ for some $\kappa > 0$.

Lemma 2. *Under* **A1** *and* **A2** *we have*

$$T^{-\frac{1}{2}} \sum_{t=1}^{T} vec\,(\epsilon_t V_{t-1}') \Rightarrow \mathcal{N}(0, \Omega),$$

where Ω *is of the form*

$$\Omega = \sum_{h=-\infty}^{\infty} E\left\{ V_{t-1}V_{t-h-1}' \otimes \epsilon_t \epsilon_{t-h}' \right\}.$$

Proof of Lemma 2. Lemma 2 is a consequence of the CLT of Herrndorf (1984) and Lemma 1. This result is proved in Raïssi (2009).

Proof of Proposition 3.1. First note that from (3.4) we have $T^{\frac{1}{2}}\mathrm{vec}\,\{\hat{\theta} - \hat{\theta}(\beta_{0c})\} = o_P(1)$. Then it suffices to show the asymptotic normality of $T^{\frac{1}{2}}\mathrm{vec}\,\{\hat{\theta}(\beta_{0c}) - \theta_0\}$. From

(3.1) and (3.2) we have

$$T^{\frac{1}{2}}\mathrm{vec}\,\{(\hat{\theta}(\beta_{0c})-\theta_0)\} \quad = \quad T^{-\frac{1}{2}}\sum_{t=1}^{T}\mathrm{vec}\,\{\epsilon_t\xi_{t-1}'(\beta_{0c})\hat{\Sigma}_{\xi}(\beta_{0c})^{-1}\}.$$

Using the Slutsky lemma we obtain

$$T^{\frac{1}{2}}\mathrm{vec}\,\{(\hat{\theta}(\beta_{0c})-\theta_0)\} = T^{-\frac{1}{2}}\sum_{t=1}^{T}\mathrm{vec}\,\{\epsilon_t\xi_{t-1}'(\beta_{0c})\Sigma_{\xi}(\beta_{0c})^{-1}\} + o_P(1). \tag{9.1}$$

From (2.4) and (2.5) the vector $\Sigma_{\xi}(\beta_{0c})^{-1}\xi_t(\beta_{0c})$ can be written as

$$\Sigma_{\xi}(\beta_{0c})^{-1}\xi_t(\beta_{0c}) = \sum_{i=0}^{\infty}\vartheta_i\epsilon_{t-i}$$

where $\vartheta(z) = \sum_{i=0}^{\infty}\vartheta_i z^i$ is convergent for $\mid z \mid \leq 1+\delta$ for some $\delta > 0$. Then using Lemma 2 the result (3.5) follows. $\square$

Proof of Proposition 4.1 First let us define the white noise "empirical" autocovariances

$$C_h = \frac{1}{T}\sum_{t=h+1}^{T}\epsilon_t\epsilon_{t-h}' \quad \text{and} \quad c_m = \mathrm{vec}\,\{(C_1,\dots,C_m)\}\,.$$

We only show the asymptotic normality of $T^{\frac{1}{2}}(c_m', \mathrm{vec}\,(\hat{\theta}-\theta_0)')'$. The rest of the proof is similar to the proof of Theorem 1 of Francq and Raïssi (2007). Similarly to the proof of Proposition 3.1 this amounts to show the asymptotic normality of the joint distribution of c_m and $\mathrm{vec}\,(\hat{\theta}(\beta_{0c})-\theta_0)$.

Using the relation $\mathrm{vec}\,(gb') = b\otimes g$ where g and b are vectors, we write $c_m = T^{-1}\sum_{t=1}^{T}\{w_t\otimes\epsilon_t\}$ and

$$\begin{aligned}\mathrm{vec}\,\{(\hat{\theta}(\beta_{0c})-\theta_0)\} \quad &= T^{-1}\sum_{t=1}^{T}\mathrm{vec}\,\{\epsilon_t\xi_{t-1}'(\beta_{0c})\hat{\Sigma}_{\xi}(\beta_{0c})^{-1}\}\\ &= T^{-1}\sum_{t=1}^{T}\hat{\Sigma}_{\xi}(\beta_{0c})^{-1}\xi_{t-1}(\beta_{0c})\otimes\epsilon_t.\end{aligned}$$

Then from (9.1), we have

$$\sqrt{T}\begin{pmatrix} c_m \\ \mathrm{vec}\,\{\hat{\theta}(\beta_{0c})-\theta_0\}\end{pmatrix} = \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\begin{pmatrix} w_t\otimes\epsilon_t \\ \Sigma_{\xi}(\beta_{0c})^{-1}\xi_{t-1}(\beta_{0c})\otimes\epsilon_t\end{pmatrix} + o_P(1).$$

Using the relation $g\otimes b' = gb'$, it follows that

$$\begin{aligned}\sqrt{T}\begin{pmatrix} c_m \\ \mathrm{vec}\,\{\hat{\theta}(\beta_{0c})-\theta_0\}\end{pmatrix} \quad &= \quad \frac{1}{\sqrt{T}}\sum_{t=1}^{T}(y_{t-1}\otimes\epsilon_t) + o_P(1)\\ &= \quad \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\mathrm{vec}\,(\epsilon_t\otimes y_{t-1}') + o_P(1)\\ &= \quad \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\mathrm{vec}\,(\epsilon_t y_{t-1}') + o_P(1),\end{aligned}$$

where $y_{t-1}=\{\epsilon'_{t-1},\dots,\epsilon'_{t-m},\xi'_{t-1}(\beta_{0c})\Sigma_\xi(\beta_{0c})^{-1}\}'$. Using again Lemma 2 and noting that vec $(\epsilon_t y'_{t-1})=\Upsilon_t$, where (Υ_t) is defined in (4.2), we obtain

$$\sqrt{T}\left(\begin{array}{c} c_m \\ \mathrm{vec}\,\{\hat{\theta}(\beta_{0c})-\theta_0\} \end{array}\right)\Rightarrow\mathcal{N}(0,\Xi).$$

Considering C_h and $\hat{\Gamma}_\epsilon(h)$ as values of the same function at the points θ_0 and $\hat{\theta}(\beta_{0c})$ (see Francq and Raïssi (2007)), it can be shown that

$$\hat{\gamma}_m=c_m+\Phi_m[\mathrm{vec}\,\{\hat{\theta}(\beta_{0c})-\theta_0\}]+o_p(T^{-\frac{1}{2}}),$$

and then we obtain (4.3). The result (4.5) can be easily obtained from (4.3). □

Appendix A: Models with deterministic parameters

In this part we consider VECMs which allow for a nonzero mean or a linear trend for $\beta' X_t$. Let us consider the following VECM

$$\Delta x_t=\mu_{o0}+\mu_{o1}t+\alpha_0\beta_0' X_{t-1}+\sum_{i=1}^{p_0-1}\Gamma_{0i}\Delta X_{t-i}+\epsilon_t.$$

Using similar assumptions as in Section 2, it follows from Granger's representation theorem that

$$\begin{aligned} X_t &= C\sum_{i=1}^{t}(\epsilon_i+\mu_{o0}+\mu_{o1}t)+C(L)(\epsilon_t+\mu_{o0}+\mu_{o1}t)+A \\ &= C\sum_{i=1}^{t}\epsilon_i+\rho_{o1}t+\rho_{o0}+Y_t+A, \end{aligned} \tag{9.2}$$

The following restrictions are usually considered in the literature

$$\begin{array}{ll} R_l: & \mu_{o1}=\alpha\tau_{0l} \\ \tilde{R}_l: & \mu_{o1}=0 \\ \tilde{R}_s: & \mu_{o0}=\alpha\tau_{0s}\text{ and }\mu_{o1}=0 \\ R_s: & \mu_{o0}=\mu_{o1}=0. \end{array} \tag{9.3}$$

Note the relation $R_s\subset\tilde{R}_s\subset\tilde{R}_l\subset R_l$. The restriction R_s correspond to the case without deterministic terms studied in this paper. If we suppose that R_l hold, we may have $\rho_{o1}\neq 0$ and $\rho_{o0}\neq 0$ and in this case $(\beta_0' X_t)$ can be composed by a stationary process plus a linear trend. We say in this case that $(\beta_0' X_t)$ is trend stationary. If $\tilde{R}_l$ hold, again we may have $\rho_{o1}\neq 0$ and $\rho_{o0}\neq 0$, but $(\beta_0' X_t)$ is stationary. If $\tilde{R}_s$ hold we obtain $\rho_{o1}=0$ but we still may have $\rho_{o0}\neq 0$, and in this case it is allowed to have $E(\beta_0' X_t)\neq 0$. Finally the restriction R_s does not allow for any deterministic component for $(\beta_0' X_t)$ and (X_t). In these cases the number of independent linear combinations $\beta_0' X_t$ which are such that the random walk behaviour is vanished is the cointegrating rank. Note also that from (9.2) the process (ΔX_t) is stationary in all cases, so that (X_t) is $I(1)$. We do not discuss the unrestricted case which produce nonstationary processes with quadratic trend since it is rarely faced in practice.

We consider restrictions as in Section 2 to ensure identification. For restrictions R_l and R_s the parameters τ_{0cl} and τ_{0cs} are estimated by reduced rank regression defining $\beta_{0l} = (\beta_{0c}', \tau_{0cl})$ with $Z_{1t}^l = (X_{t-1}', t)'$ and $\beta_{0s} = (\beta_{0c}', \tau_{0cs})$ with $Z_{1t}^s = (X_{t-1}', 1)'$. For restriction R_l we define $Z_{2t}^\mu = (\Delta X_{t-1}', \dots, \Delta X_{t-p+1}', 1)'$ and $\xi_{t-1}^l = ((\beta_{0l}' Z_{1t}^l)', Z_{2t}'^\mu)'$. The LS estimator of $\theta_0^\mu = (\alpha_{0c}, \Gamma_{01} \dots \Gamma_{0p-1}, \mu_0)$ is obtained similarly as in (3.3). In the case of $\tilde{R}_l$ we consider $\tilde{\xi}_{t-1}^l = ((\beta_0' X_{t-1})', Z_{2t}'^\mu)'$. In the case of $\tilde{R}_s$ the estimator of $\theta_0 = (\alpha_{0c}, \Gamma_{01} \dots \Gamma_{0p-1})$ is obtained considering $\xi_{t-1}^s = ((\beta_{0s}' Z_{1t}^s)', Z_{2t}')'$ recalling that $Z_{2t} = (\Delta X_{t-1}', \dots, \Delta X_{t-p+1}')'$. In the sequel we denote these estimators by $\hat{\theta}^\mu$, $\tilde{\theta}^\mu$ and $\tilde{\theta}$ with obvious notations. Similarly to (3.2) we also define the unfeasible estimators $\tilde{\theta}^\mu(\beta_{0l})$, $\tilde{\theta}^\mu(\beta_{0c})$ and $\tilde{\theta}(\beta_{0s})$.

Considering similar arguments used in lemma 13.1 of Johansen (1995) for restriction $\tilde{R}_l$ (see also Raïssi (2007, p 110) for restriction R_l) it can be shown that

$$\hat{\theta}^\mu = \tilde{\theta}^\mu(\beta_{0l}) + o_p(T^{-\frac{1}{2}}), \quad \tilde{\theta}^\mu = \tilde{\theta}^\mu(\beta_{0c}) + o_p(T^{-\frac{1}{2}}) \text{ and } \tilde{\theta} = \tilde{\theta}(\beta_{0s}) + o_p(T^{-\frac{1}{2}}).$$

Then for the different restrictions considered in (9.3) the method for building tests based on the residual autocorrelations is completely similar to the case R_s. However note that for restrictions R_l, $\tilde{R}_l$ and $\tilde{R}_s$ we obtain

$$\Phi_m^l = -E\{w_t \otimes \xi_{t-1}^{l'} \otimes I_d\},$$

$$\tilde{\Phi}_m^l = -E\{w_t \otimes \tilde{\xi}_{t-1}^{l'} \otimes I_d\},$$

and

$$\tilde{\Phi}_m^s = -E\{w_t \otimes \tilde{\xi}_{t-1}^{s'} \otimes I_d\}.$$

Using (9.2) we write

$$\beta' X_t = \beta' \rho_{o1} t + \beta' \rho_{o0} + \beta' Y_t \quad \text{in the case of } R_l,$$

$$\beta' X_t = \beta' \rho_{o0} + \beta' Y_t$$

in the case of $\tilde{R}_l$ and

$$\beta' X_t = \beta' \rho_{o0} + \beta' Y_t \quad \text{in the case of } \tilde{R}_s.$$

Let us define $\check{\xi} = (X_{t-1}', \Delta X_{t-1}', \dots, \Delta X_{t-p+1}')'$. Since $E(\epsilon_t) = 0$, the deterministic terms vanishes in the expressions of Φ_m^l, $\tilde{\Phi}_m^l$ and $\tilde{\Phi}_m^s$ and we have

$$E\{w_t \otimes \xi_{t-1}^{l'}\} = E\{w_t \otimes \check{\xi}', 0_{md \times 1}\},$$

$$E\{w_t \otimes \tilde{\xi}^{l\prime}_{t-1}\} = E\{w_t \otimes \check{\xi}', 0_{md\times 1}\}$$

and

$$E\{w_t \otimes \tilde{\xi}^{s\prime}_{t-1}\} = E\{w_t \otimes \check{\xi}'\}.$$

Note also that we can write in all cases

$$\tilde{\xi}_t = A\tilde{\xi}_{t-1} + \tilde{\mu}_{o0} + \tilde{\mu}_{o1}t + \tilde{\epsilon}_t,$$

where $\tilde{\epsilon}_t = (\epsilon'_t, \epsilon'_t, 0, \dots, 0)'$, $\tilde{\mu}_{o0} = (\mu'_{o0}, \mu'_{o0}, 0, \dots, 0)'$ and $\tilde{\mu}_{o1} = (\mu'_{o1}, \mu'_{o1}, 0, \dots, 0)'$. Therefore the expression of $\tilde{\Phi}^s_m$ is the same to that of Φ_m. The expression of the dp first columns of Φ^l_m and $\tilde{\Phi}^l_m$ are the same to that of Φ_m. The last d columns of Φ^l_m and $\tilde{\Phi}^l_m$ are equal to zero.

Appendix B: Computation of the weights of the weighted sum of Chi squares

Given d-multivariate observations $X_1, \dots, X_T$, one can use the following steps to implement the tests introduced in this paper for testing the adequacy of the autoregressive order of a VECM.

1) Compute the estimates $\hat{\alpha}_c, \hat{\beta}_c, \hat{\Gamma}_1 \dots, \hat{\Gamma}_{p-1}$;
2) Compute the residuals $\hat{\epsilon}_t = \Delta X_t - \hat{\alpha}_c\hat{\beta}'_c X_{t-1} - \sum_{i=1}^{p-1}\hat{\Gamma}_i X_{t-i}$, for $t = p-1, \dots, T$, and the residual autocovariances $\hat{\Gamma}_\epsilon(0) = \hat{\Sigma}_\epsilon$ and $\hat{\Gamma}_\epsilon(h)$, for $h = 1, \dots, m$;
3) Compute
$$\widehat{\Upsilon}_t = \begin{pmatrix} \begin{pmatrix} \hat{\epsilon}_{t-1} \\ \vdots \\ \hat{\epsilon}_{t-m} \end{pmatrix} \otimes \hat{\epsilon}_t \\ \hat{\Sigma}_\xi(\hat{\beta}_c)^{-1}\xi'_{t-1}(\hat{\beta}_c) \otimes \hat{\epsilon}_t \end{pmatrix} \quad \text{for} \quad t = m+1, \dots, T;$$
4) Compute the first $p_0 + 1$ autocovariances of $\widehat{\Upsilon}_1, \dots, \widehat{\Upsilon}_T$;
5) Using the Durbin-Levinson algorithm, fit the AR(p_0) model
$$\left(I - \sum_{i=1}^{p_0}\hat{\mathcal{A}}_i B^i\right)\widehat{\Upsilon}_t = \hat{U}_t;$$
6) Define the estimator
$$\hat{\Xi} = \left(I - \sum_{i=1}^{p_0}\hat{\mathcal{A}}_i\right)^{-1}\hat{\Sigma}_U\left(I - \sum_{i=1}^{p_0}\hat{\mathcal{A}}_i\right)^{\prime -1}, \quad \hat{\Sigma}_U = \frac{1}{T}\sum_{t=1}^{T}\hat{U}_t\hat{U}'_t;$$
7) Define the estimator
$$\hat{\Sigma}_{\hat{\gamma}_m} = \hat{\Sigma}_{c_m} + \hat{\Phi}_m\hat{\Sigma}_{\hat{\theta}_n}\hat{\Phi}'_m + \hat{\Sigma}_{c_m,\hat{\theta}_n}\hat{\Phi}'_m + \hat{\Phi}_m\hat{\Sigma}'_{c_m,\hat{\theta}_n},$$
where
$$\hat{\Xi} = \begin{pmatrix} \hat{\Sigma}_{c_m} & \hat{\Sigma}_{c_m,\hat{\theta}_n} \\ \hat{\Sigma}'_{c_m,\hat{\theta}_n} & \hat{\Sigma}_{\hat{\theta}_n} \end{pmatrix}, \quad \hat{\Phi}_m = -\sum_{i=0}^{m-1}\left\{\mathbf{1}_{m\times p}(i+1,1)\otimes\hat{\Sigma}_\epsilon\right\}\left(\widehat{\tilde{A}}^i\right)'\otimes I_d;$$

8) Compute the eigenvalues $(\hat{\delta}_{1,d^2m}, \cdots, \hat{\delta}_{d^2m,d^2m})$ of

$$\hat{\Omega}_m = \left(I_m \otimes \hat{\Sigma}_\epsilon^{-1/2} \otimes \hat{\Sigma}_\epsilon^{-1/2}\right) \hat{\Sigma}_{\hat{\gamma}_m} \left(I_m \otimes \hat{\Sigma}_\epsilon^{-1/2} \otimes \hat{\Sigma}_\epsilon^{-1/2}\right);$$

References

Andrews, B., Davis, R.A., and Breidt, F.J. (2006) Maximum likelihood estimation for all-pass time series models. *Journal of Multivariate Analysis*, 97:1638-1659.

Bauwens, L., Laurent, S. and Rombouts, J.V.K. (2006) Multivariate GARCH models: A survey. *Journal of Applied Econometrics*, 21:79-109.

Box, G.E.P. and Pierce, D.A. (1970) Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. *Journal of American Statistical Association*, 65:1509-1526.

Bollerslev, T.P. (1990) Modeling the coherence in short-run nominal exchange rates: a multivariate generalized ARCH model. *Review of Economics and Statistics*, 72:498-505.

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

Brüggeman, R., Lütkepohl, H. and Saikkonen, P. (2006) Residual autocorrelation testing for error correction models. *Journal of Econometrics*, 134:579-604.

Chitturi, R. V. (1974). Distribution of residual autocorrelations in multiple autoregressive schemes. *Journal of American Statistical Association*, 69:928-934.

Davydov, Y. A. (1968) Convergence of distributions generated by stationary stochastic process. *Theory of Probability and Applications*, 13:691-96.

Duchesne, P. (2005) Testing for serial correlation of unknown form in cointegrated time series models. *Annals of the Institute of Statistical Mathematics*, 57:575-595.

Engle, R.F. and Kroner, K.F. (1995) Multivariate simultaneous generalized ARCH. *Econometric Theory*, 11:122-150.

Francq, C., and Raïssi, H. (2007) Multivariate portmanteau test for autoregressive models with uncorrelated but nonindependent 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 American Statistical Association*, 100:532-544.

Francq, C. and Zakoïan, J-M. (1998) Estimating linear representations of nonlinear processes. *Journal of Statistical Planning and Inference*, 68:145-165.

Godfrey, L.G. (1978) Testing for higher order serial correlation in regression equations when the regressors include lagged dependent variables. *Econometrica*, 46:1303-1313.

Herrndorf, N. (1984) A functional central limit theorem for weakly dependent sequences of random variables. *Annals of Probability*, 12:141-153.

Hosking, J. R. M. (1980) The multivariate portmanteau statistic. *Journal of American Statistical Association*, 75:343-386.

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

Jeantheau, T. (1998) Strong consistency of estimators for multivariate ARCH models. *Econometric Theory*, 14:70-86.

Johansen, S. (1995) *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models.* Oxford University Press, New York.

Kuonen, D. (1999) Saddlepoint approximations for distributions of quadratic forms in normal variables. *Biometrika*, 86:929-935.

Ljung, G.M. and Box, G.E.P. (1978) On measure of lack of fit in time series models. *Biometrika*, 65:297-303.

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

Raïssi, H. (2009) Testing the cointegrating rank with uncorrelated but dependent errors. *Stochastic Analysis and Applications*, 27:24-50.

Raïssi, H. (2008) Testing linear causality in mean in presence of other forms of causality. Working document Université Lille 3 and INSA Rennes. http://www.insa-rennes.fr/wpFichiers/25/82/ressources/File/causalgranger.pdf.

Romano, J. P. and Thombs, L. A. (1996) Inference for autocorrelations under weak assumptions. *Journal of American Statistical Association*, 91:590-600.

10. Tables and Figures

Table 1: Parameters used in the Monte Carlo experiments.

(a) $d = 2, \quad p = 2, \quad r = 1$

$$\alpha_{0c} = \begin{pmatrix} -0.3 \\ 0.4 \end{pmatrix} \qquad \beta_{0c} = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \qquad \Gamma = \begin{pmatrix} 0.3 & 0.2 \\ 0.1 & 0.4 \end{pmatrix}$$

(b) $d = 3, \quad p = 2, \quad r = 2$

$$\alpha_{0c} = \begin{pmatrix} 0.2 & 0 \\ -0.2 & 0.2 \\ 0 & 0 \end{pmatrix} \qquad \beta_{0c} = \begin{pmatrix} -1 & -1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad \Gamma = \begin{pmatrix} 0.5 & -0.2 & -0.2 \\ -0.2 & 0.5 & -0.2 \\ -0.2 & -0.2 & 0.5 \end{pmatrix}$$

Table 2: Empirical size (in %) of the portmanteau and LM tests for the DGP (a) in the strong case.

	$m = 1$			$m = 5$		
	$T = 100$	$T = 400$	$T = 1000$	$T = 100$	$T = 400$	$T = 1000$
BP_W	**6.8**	5.2	5.7	**13.0**	4.7	5.1
BP_S	n.a.	n.a.	n.a.	4.4	5.1	5.4
LB_W	**6.9**	5.2	5.7	**14.2**	5.0	5.2
LB_S	n.a.	n.a.	n.a.	5.4	5.3	5.5
LM_W	**7.3**	5.2	5.8	**13.4**	4.9	5.1
LM_S	5.9	5.0	5.0	5.5	5.3	**6.4**

	$m = 10$			$m = 20$		
	$T = 100$	$T = 400$	$T = 1000$	$T = 100$	$T = 400$	$T = 1000$
BP_W	**1.5**	**3.4**	4.7	**0.3**	**3.2**	3.9
BP_S	**2.9**	4.0	5.3	**1.3**	3.7	4.1
LB_W	**3.0**	4.3	5.2	**2.1**	4.3	4.1
LB_S	5.0	4.6	5.7	5.8	5.9	4.8
LM_W	**1.1**	**3.3**	4.7	**0.5**	**3.3**	3.7
LM_S	4.9	5.4	5.6	**2.1**	5.2	5.5

n.a.: not available.

Table 3: The same as for Table 2 but for the weak case (7.1).

	$m=1$			$m=5$		
	$T=100$	$T=400$	$T=1000$	$T=100$	$T=400$	$T=1000$
BP_W	6.1	5.1	4.8	**9.6**	4.4	5.3
BP_S	n.a.	n.a.	n.a.	5.7	**10.3**	**12.7**
LB_W	6.2	5.1	4.8	**10.7**	4.8	5.4
LB_S	n.a.	n.a.	n.a.	**7.0**	**10.4**	**13.1**
LM_W	**7.0**	5.2	5.1	**9.8**	5.5	5.9
LM_S	**10.1**	**11.1**	**12.2**	**8.9**	**10.6**	**12.4**
	$m=10$			$m=20$		
	$T=100$	$T=400$	$T=1000$	$T=100$	$T=400$	$T=1000$
BP_W	**1.1**	**2.4**	4.4	**0.3**	**2.6**	**3.5**
BP_S	**2.9**	6.3	**8.9**	**1.5**	4.4	**6.7**
LB_W	**2.6**	**3.6**	4.8	**1.7**	**3.4**	3.8
LB_S	5.3	**6.9**	**9.1**	6.2	**6.5**	**7.3**
LM_W	**1.2**	**2.8**	4.8	**0.4**	**2.7**	**3.5**
LM_S	**7.5**	**8.6**	**10.3**	**3.3**	6.2	**8.2**

n.a.: not available.

Table 4: Empirical size (in %) of the portmanteau and LM tests for the DGP (b) in the strong case.

	$m=1$			$m=5$		
	$T=100$	$T=400$	$T=1000$	$T=100$	$T=400$	$T=1000$
BP_W	**16.1**	4.8	6.2	5.1	4.1	4.5
BP_S	n.a.	n.a.	n.a.	**11.1**	**11.9**	**11.5**
LB_W	**16.8**	4.8	6.2	**6.4**	4.8	4.8
LB_S	n.a.	n.a.	n.a.	**14.7**	**12.8**	**11.7**
LM_W	**16.8**	5.3	**6.5**	6.0	4.6	4.7
LM_S	**14.7**	**6.8**	**6.4**	**13.7**	**12.4**	**10.9**
	$m=10$			$m=20$		
	$T=100$	$T=400$	$T=1000$	$T=100$	$T=400$	$T=1000$
BP_W	**2.1**	**2.3**	4.4	**0.1**	**0.9**	**1.6**
BP_S	3.7	5.0	5.7	**1.0**	**2.8**	**3.1**
LB_W	3.9	**3.4**	4.7	**3.1**	**1.7**	**2.2**
LB_S	**8.2**	5.9	6.2	**6.5**	4.7	4.7
LM_W	**2.4**	**2.4**	3.9	**0.1**	**0.7**	**1.9**
LM_S	**10.0**	**8.5**	**9.8**	**3.2**	6.1	6.0

n.a.: not available.

TABLE 5: Empirical power (in %) of the portmanteau and LM tests for the DGP (a) with errors given in (7.2), with $T = 1000$ and $m = 10$.

ω_2	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.9
BP_W	4.9	6.4	14.5	31.8	56.9	79.6	90.7	97.5	99.7
BP_S	5.3	6.8	15.6	33.9	58.0	80.5	91.1	97.6	99.7
LB_W	5.3	6.6	15.2	33.3	57.8	80.4	90.9	97.5	99.7
LB_S	5.7	7.1	16.1	35.0	58.9	81.0	91.5	97.8	99.7
LM_W	4.6	6.1	14.6	31.7	56.5	79.4	90.5	97.1	99.7
LM_S	5.7	8.1	16.6	35.7	60.1	81.2	92.5	97.8	99.8

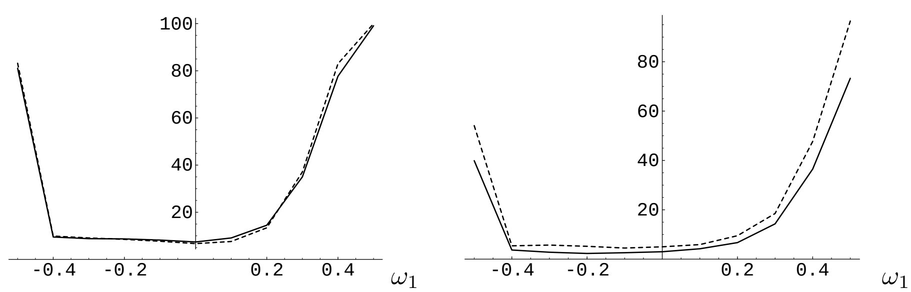

FIGURE 10.1: Empirical power (in %) of the LB_S test (dotted line) and LB_W test (full line) for the DGP (a) with $T = 100$. We take on the left $m = 3$ and on the right $m = 10$.

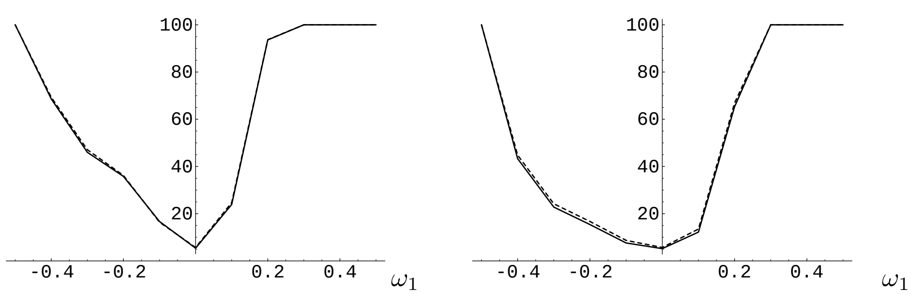

FIGURE 10.2: The same as for the Figure 10.1 but for $T = 1000$.

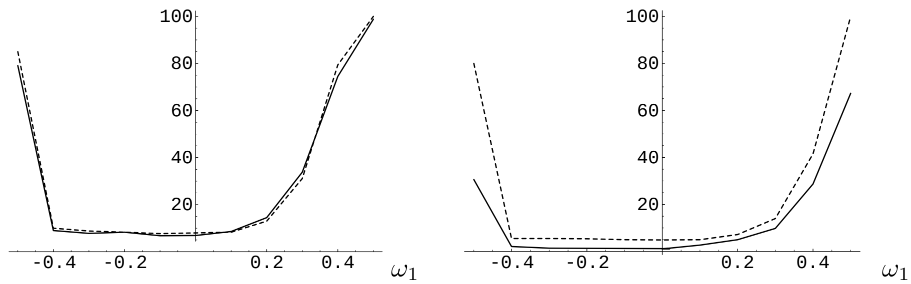

FIGURE 10.3: Empirical power (in %) of the LM_S test (dotted line) and LM_W test (full line) for the DGP (a) with $T = 100$. We take on the left $m = 3$ and on the right $m = 10$.

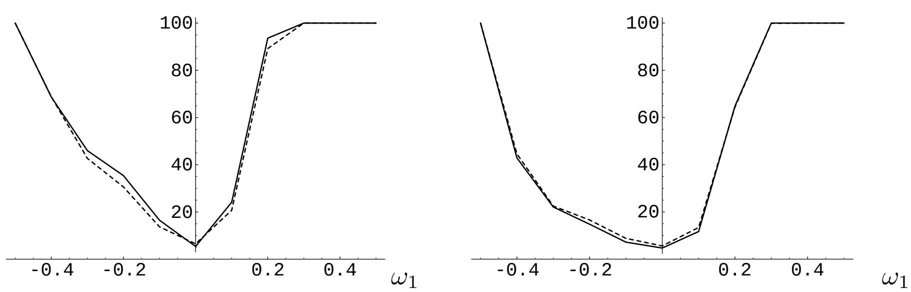

FIGURE 10.4: The same as for the Figure 10.3 but for $T = 1000$.