ESCUELA DE NEGOCIOS Y ECONOMÍA

PONTIFICIA UNIVERSIDAD CATÓLICA DE VALPARAÍSO

Working Papers

2017-03

Testing the lag length of vector autoregressive models: A power comparison between portmanteau and Lagrange multiplier tests

Raja Ben Hajria

National Engineering School, University of Monastir

Salah Khardani

National Engineering School, University of Monastir

Hamdi Raïssi

Instituto de Estadística, Pontificia Universidad Católica de Valparaíso

ene.pucv.cl/wp

Testing the lag length of vector autoregressive models: A power comparison between portmanteau and Lagrange multiplier tests

Raja Ben Hajria*, Salah Khardani,† Hamdi Raïssi‡

November 17, 2017

Abstract: In this paper we provide an asymptotic theoretical power comparison in the Bahadur sense, between the portmanteau and Breusch-Godfrey Lagrange Multiplier (LM) tests for the goodness-of-fit checking of vector autoregressive (VAR) models. The merits and the drawbacks of the studied tests are illustrated using Monte Carlo experiments.

Keywords: VAR model; VECM model; Cointegration; Residual autocorrelations; Portmanteau tests, Lagrange Multiplier tests.

*National Engineering School of Monastir,University of Monastir, Tunisia.

†National Engineering School of Monastir,University of Monastir, Tunisia.

‡Corresponding author: Instituto de Estadística, PUCV, Av. Errazuriz 2734, Valparaíso, V región, Chile. E-mail: hamdi.raissi@pucv.cl. This paper was supported by CONICYT-FONDECYT under grant $N°1160527$.

1 Introduction

In empirical economic studies VAR models are often used to investigate relationships between variables. The reader is referred to Lütkepohl (2005) for the various tools that are commonly used in the context of VAR models. However it is widely known that a careful analysis relies on a well specified model. For instance Stock and Watson (1989) or Gonzalo and Pitarakis (1998), among many others, documented the importance of an adequate choice of the lag length for Granger causality or cointegration analyses. The dominant tests in the literature are the portmanteau tests, introduced first in Box and Pierce (1970) (BP hereafter) and Ljung and Box (1978) (LB hereafter), and the Breusch-Godfrey LM test proposed in Breusch (1978) and Godfrey (1978). These tests for univariate models were extended in the VAR framework (see Chitturi (1974) for testing the adequacy of stationary VAR models using the portmanteau test, and Brüggeman, Lütkepohl and Saikkonen (2006) in the case of cointegrated variables). Such kind of tests are routinely used in softwares as R, SAS or JMulTi. (see Pfaff (2008) for the implementation in R, or Lütkepohl and Krätzig (2004) for the software JMulTi).

Let denote by m the number of autocorrelations used in the test statistics (see equations (2.6), (2.7) and (2.9) below). Hatemi (2004) and Brüggeman *et al* (2006) carried out simulation studies to compare the finite sample properties of the LM and LB tests. They found that the LM test has a better control of the type I errors than the portmanteau tests when m is small. On the other hand Brüggeman *et al* (2006) underlined that the size properties of the portmanteau tests are better than those of the LM test when a large m is used. It also emerges from the above studies that the portmanteau tests are slightly more powerful than the LM test. Bearing in mind that the alternatives were linear in both papers, the latter result may seem surprising at first sight as the LM test is intended to detect such kinds of alternatives by nature (see equation (2.8) below). In this paper we give some asymptotic evidence for the power comparison between the LM and portmanteau tests. More precisely it is shown that the LM test is more powerful than the portmanteau tests in the Bahadur sense. However it turns out that this asymptotic power advantage is at the cost of the control of the type I errors.

This paper is structured as follows. In the next section, we introduce the portmanteau and the LM tests. The powers of the studied tests are compared in the Bahadur sense. In Section 3 Monte Carlo experiments are conducted to illustrate the theoretical findings.

The following notations will be used throughout the paper. For a multivariate random variable v, let $\|v\|_q = (\mathrm{E}\|v\|^q)^{1/q}$, where $\|.\|$ denotes the Euclidean norm with $\mathrm{E}\|v\|^q < \infty$ and $q \geq 1$. We denote by $A \otimes B$ the Kronecker product of two matrices A and B. The determinant of a square matrix A is denoted by $\det(A)$. The vector obtained by stacking the columns of A is denoted $\mathrm{vec}(A)$. The symbol $\Rightarrow$ denotes the convergence in distribution, and we denote by $\stackrel{P}{\to}$ the convergence in probability.

2 Testing the lag length of VAR processes

Let us consider the VAR model in its error correction form (VECM):

$$\Delta X_t = \Pi_0 X_{t-1} + \sum_{i=1}^{p_0-1} \Gamma_{0i} \Delta X_{t-i} + \epsilon_t \qquad (2.1)$$

where $\Delta X_t := X_t - X_{t-1}$. The Γ_{0i}, $i \in \{1, ..., p_0 - 1\}$, are $d \times d$ short run parameters matrices. By convention the sum vanishes in (2.1) when $p_0 = 1$. Let us denote by p the adjusted lag length. In the sequel we assume that $X_{-p}, \dots, X_0, X_1, \dots, X_T$ are observed.

Assumption A1

(a) The process (ϵ_t) is iid with mean zero and positive definite covariance matrix Σ_ϵ, and such that $\mathrm{E}(\|\epsilon_t\|^4) < \infty$.

(b) The matrix Π_0 is of rank $0 < r_0 < d$, so that Π_0 can be written as $\Pi_0 = \alpha_0\beta_0'$ where α_0 and β_0 are full column rank matrices of dimension $d \times r_0$.

(c) The autoregressive polynomial $A(z) = (1-z)I_d - \Pi_0 z - \sum_{i=1}^{p_0-1} \Gamma_{0i}(1-z)z^i$, is such that $\mid A(z) \mid = 0$ implies that $\mid z \mid > 1$ or $z = 1$.

(d) The matrix $\alpha_{0\perp}'\Gamma_0\beta_{0\perp}$ is of full rank $d - r_0$, where $\Gamma_0 = I_d - \sum_{i=1}^{p_0-1} \Gamma_{0i}$, and $\alpha_{0\perp}'\alpha_0 = 0$, $\beta_{0\perp}'\beta_0 = 0$.

In this paper the cointegrating rank is assumed to be known. Nevertheless this assumption is not realistic in practice since the cointegrating rank has to be estimated in general. Noting that the effects of estimated cointegrating rank on the lag length selection does not constitute the main scope of the paper, and we simply refer to Brüggeman *et al* (2006), Tables 4 and 5, and Figure 2, who studied this issue. If $r_0 = 0$ the model consists in a VAR for differentiated data. If $r_0 = d$ the process (X_t) follows a stationary VAR model. For ease of exposition we will not consider the $r_0 = 0$ and $r_0 = d$ cases. Also we do not consider deterministic terms in (2.1). However it is easy to see that all the results of the paper can be obtained in a similar way if $r_0 = 0$ or $r_0 = d$, or when deterministic terms are present. From **A1(b)** the system is cointegrated. The α_0 and β_0 correspond to the adjustment and long run parameters. Note that α_0 and β_0 should be identified in some appropriate way (see Johansen (1995p72). From Granger's representation theorem, the solution of (2.1) has the following form under **A1**

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

where $C = \beta_{0\perp}(\alpha'_{0\perp}\Gamma_0\beta_{0\perp})^{-1}\alpha'_{0\perp}$. The term A depends on initial values and is such that $\beta'_0 A = 0$. The stationary process (Y_t) is of the form

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

where $C(z) = \sum_{i=0}^{\infty} \varphi_{0i} z^i$ is convergent for $\mid z \mid \leq 1 + \delta$, for some $\delta > 0$. The stationary process (ΔX_t) can be written as

$$\Delta X_t = \sum_{i=0}^{\infty} \psi_i \epsilon_{t-i}, \tag{2.3}$$

where the ψ_i's can be obtained from (2.2).

For the adjusted lag length p that could be $p \neq p_0$, the model (2.1) is estimated by maximum likelihood (see Johansen (1995)). If we suppose that the lag length is well adjusted, the estimator obtained for $\mathrm{vec}(\alpha_0, \Gamma_{01}, \dots, \Gamma_{0p-1})$ is consistent in probability and $\sqrt{T}$-asymptotically normal. On the other hand the estimator $\hat{\beta}$ of β_0 is such that $T(\hat{\beta} - \beta_0) = O_p(1)$, so we can suppose the long run parameters known without a loss of generality. The resulting residuals will be denoted by $\hat{\epsilon}_t$. The adequacy of the adjusted lag length is usually checked by testing the following pair of hypotheses

$$H_0 : \gamma_m = 0 \quad \text{vs} \quad H_1 : \gamma_m \neq 0,$$

where $\gamma_m = \mathrm{vec}\left\{(E(\epsilon_t\epsilon'_{t-1}), \dots, E(\epsilon_t\epsilon'_{t-m}))\right\}$ for $m \in \{1, \dots\}$. In practice several values of m are considered. If there is no prior information, the adjusted lag length p is increased until the null hypothesis is not rejected.

We first recall the asymptotic behaviors of the LM and the portmanteau test statistics under H_0 $(p = p_0)$. In our cointegrated framework, these results are established in Brüggeman *et al* (2006). Let us define $\tilde{X}_t = (X'_{t-1}\beta_0, \Delta X'_{t-1} \dots, \Delta X'_{t-p+1})'$. We write:

$$\tilde{X}_t = \sum_{i=0}^{\infty} \tilde{\psi}_i \tilde{\epsilon}_{t-i},$$

where $\tilde{\epsilon}_t = \mathbf{1}_p \otimes \epsilon_t$, $\mathbf{1}_p$ is the vector of ones of dimension p, the $\tilde{\psi}_i$'s are $(r_0+d(p-1)) \times (r_0+d(p-1))$ dimensional, and can be obtained from the MA(∞) forms (2.2) and (2.3). Define the residual autocovariances $\hat{\Gamma}(h) := T^{-1} \sum_{t=1}^{T} \hat{\epsilon}_t \hat{\epsilon}'_{t-h}$ and the vector of the m first residual autocovariances $\hat{\gamma}_m = \text{vec}\left\{(\hat{\Gamma}(1), \ldots, \hat{\Gamma}(m))\right\}$. We have

$$T^{\frac{1}{2}} \left(I_m \otimes \hat{\Gamma}^{-\frac{1}{2}}(0) \otimes \hat{\Gamma}^{-\frac{1}{2}}(0) \right) \hat{\gamma}_m \Rightarrow \mathcal{N}\left(0, I_{d^2 m} - K_m K_\infty^{-1} K'_m \otimes I_d\right),$$

where

$$K_m = \begin{pmatrix} (\mathbf{1}'_p \otimes \Sigma_\epsilon^{\frac{1}{2}}) \tilde{\psi}'_0 \\ \vdots \\ (\mathbf{1}'_p \otimes \Sigma_\epsilon^{\frac{1}{2}}) \tilde{\psi}'_{m-1} \end{pmatrix} \tag{2.4}$$

is of dimension $dm \times (r_0 + d(p-1))$ and

$$K_\infty = E\left(\tilde{X}_{t-1} \tilde{X}'_{t-1}\right) = \sum_{i=0}^{\infty} \tilde{\psi}_i (\mathbf{1}_p \otimes \Sigma_\epsilon^{\frac{1}{2}})(\mathbf{1}'_p \otimes \Sigma_\epsilon^{\frac{1}{2}}) \tilde{\psi}'_i. \tag{2.5}$$

The multivariate Box-Pierce portmanteau statistic introduced by Chitturi (1974) for testing H_0 vs H_1 is given by:

$$\begin{aligned} Q_m^{BP} &= T \sum_{h=1}^{m} \text{tr}\left(\hat{\Gamma}'(h) \hat{\Gamma}^{-1}(0) \hat{\Gamma}(h) \hat{\Gamma}^{-1}(0)\right) \\ &= T \hat{\gamma}'_m \left(I_m \otimes \hat{\Gamma}^{-1}(0) \otimes \hat{\Gamma}^{-1}(0)\right) \hat{\gamma}_m. \end{aligned} \tag{2.6}$$

We also consider the Ljung-Box statistic introduced in the VAR framework by Hosking (1980)

$$Q_m^{LB} = T^2 \sum_{h=1}^{m} (T-h)^{-1} \text{tr}\left(\hat{\Gamma}'(h) \hat{\Gamma}^{-1}(0) \hat{\Gamma}(h) \hat{\Gamma}^{-1}(0)\right). \tag{2.7}$$

It is clear that K_∞ can be approximated by $K'_m K_m$ as $m \to \infty$, so that under the null hypothesis the asymptotic distribution of the Q_m^{BP} and Q_m^{LB} statistics can be approximated by a $\chi^2_{d^2(m-p+1)-dr_0}$ for large enough m. Therefore at the asymptotic level α the BP (resp. the LB) test rejects the null hypothesis if $\chi^2_{d^2(m-p+1)-dr_0,1-\alpha} < Q^{BP}_m$ (resp. $\chi^2_{d^2(m-p+1)-dr_0,1-\alpha} < Q^{LB}_m$), where $\chi^2_{d^2(m-p+1)-dr_0,1-\alpha}$ is the $(1-\alpha)th$ quantile of the $\chi^2_{d^2(m-p+1)-dr_0}$ law with $m > p$. However it is well known that m must be chosen carefully to make the portmanteau tests control the type I errors reasonably well, and it is usual to take $m \to \infty$ as $T \to \infty$.

The LM test may be viewed as based on the following model for the errors:

$$\epsilon_t = \sum_{j=1}^{m} B_{0j}\epsilon_{t-j} + e_t, \tag{2.8}$$

where under the null hypothesis $B_{0j} = 0$ for all $j \in \{1, \dots, m\}$. The process (e_t) is iid, with positive definite covariance matrix Σ_e. Let us define $\hat{J} = T^{-1}\sum_{t=1}^{T}\hat{\zeta}_{t-1}\hat{\zeta}'_{t-1} \otimes \hat{\Gamma}(0)^{-1}$, with $\hat{\zeta}_{t-1} = (\tilde{X}'_{t-1}, \hat{\epsilon}'_{t-1}, \dots, \hat{\epsilon}'_{t-m})'$, taking $\hat{\epsilon}_t = 0$ for $t \leq 0$, and $R = (0_{d^2m\times d(r_0+d(p-1))}, I_{d^2m})$ is of dimension $d^2m \times (d^2m + d(r_0 + d(p-1)))$. The nullity of the B_{0i}'s is tested using the LM approach in the context of (2.1), so that we obtain the LM test statistic:

$$Q^{LM}_m = T\hat{\gamma}'_m(I_{dm} \otimes \hat{\Gamma}(0)^{-1})(R\hat{J}^{-1}R')(I_{dm} \otimes \hat{\Gamma}(0)^{-1})\hat{\gamma}_m. \tag{2.9}$$

Under the null hypothesis we have $Q^{LM}_m \Rightarrow \chi^2_{d^2m}$. Then at the asymptotic level α the LM test reject the null hypothesis when $\chi^2_{d^2m,1-\alpha} < Q^{LM}_m$. Note that contrary to the portmanteau tests, the LM test is not based on the approximation of the asymptotic distribution.

In this part the asymptotic power properties of the above described tests is studied. We consider the Bahadur (1960) approach which consists in comparing the ability of the tests to detect a fixed alternative $H_1 : \gamma_m = \varrho \neq 0$ as

$T \to \infty$. For any $x > 0$, define $q_{LM}(x) = -\log P_0(Q_m^{LM} > x)$ where P_0 stands for the limit distribution of Q_m^{LM} under H_0. Let us consider the *(asymptotic) slope* $c_{LM}(\varrho) = 2\lim_{T\to\infty} T^{-1} q_{LM}(Q_m^{LM})$ under the alternative $H_1 : \gamma_m = \varrho \neq 0$ such that the limit exists in probability. Similarly define $c_{BP}(\varrho)$ and $c_{LB}(\varrho)$. We consider the asymptotic relative efficiencies of the test based on Q_m^{LM} with respect to the tests based on Q_m^{BP} and Q_m^{LB}, as the ratios $ARE_{LM,BP}(\varrho) = c_{LM}(\varrho)/c_{BP}(\varrho)$ and $ARE_{LM,LB}(\varrho) = c_{LM}(\varrho)/c_{LB}(\varrho)$. A relative efficiency $ARE_{LM,LB}(\varrho) \geq 1$ suggests that the LM test is better suited to detect H_1 than the LB test because the $p-$values associated with the LM test wane faster or equally faster than the $p-$values obtained using the LB test.

Proposition 1. *Under assumption* **A1***, the relative efficiencies* $ARE_{LM,LB}(\varrho)$ *and* $ARE_{LM,BP}(\varrho)$ *are larger or equal to 1 for every* $\varrho \in \mathbb{R}^{d^2 m}$.

Proof of Proposition 1 We only give the proof for the BP test for conciseness. We have for any $\varrho \neq 0$:

$$T^{-1}Q_m^{BP} = \varrho' \left(I_m \otimes \Sigma_\epsilon^{-1} \otimes \Sigma_\epsilon^{-1}\right) \varrho + o_p(1), \tag{2.10}$$

and

$$\begin{aligned} T^{-1}Q_m^{LM} &= \varrho' \left(I_m \otimes \Sigma_\epsilon^{-\frac{1}{2}} \otimes \Sigma_\epsilon^{-\frac{1}{2}}\right) \left(I_m \otimes \Sigma_\epsilon^{\frac{1}{2}} \otimes \Sigma_\epsilon^{-\frac{1}{2}}\right) \\ &\quad (RJ^{-1}R') \left(I_m \otimes \Sigma_\epsilon^{\frac{1}{2}} \otimes \Sigma_\epsilon^{-\frac{1}{2}}\right) \left(I_m \otimes \Sigma_\epsilon^{-\frac{1}{2}} \otimes \Sigma_\epsilon^{-\frac{1}{2}}\right) \varrho + o_p(1), \end{aligned} \tag{2.11}$$

using the ergodic theorem and the identity $(K \otimes L)(M \otimes N) = (KM) \otimes (LN)$, with $J = E(\zeta_{t-1}\zeta_{t-1}') \otimes \Sigma_\epsilon^{-1}$.

Using the inverse of partitioned matrices and the above identity again we write:

$$\left(I_m \otimes \Sigma_\epsilon^{\frac{1}{2}} \otimes \Sigma_\epsilon^{-\frac{1}{2}}\right)(RJ^{-1}R')\left(I_m \otimes \Sigma_\epsilon^{\frac{1}{2}} \otimes \Sigma_\epsilon^{-\frac{1}{2}}\right) =$$

$$\left(I_m \otimes \Sigma_\epsilon^{\frac{1}{2}} \otimes \Sigma_\epsilon^{-\frac{1}{2}}\right)\left(I_m \otimes \Sigma_\epsilon \otimes \Sigma_\epsilon^{-1} - K_m K_\infty^{-1} K_m' \otimes \Sigma_\epsilon^{-1}\right)^{-1}\left(I_m \otimes \Sigma_\epsilon^{\frac{1}{2}} \otimes \Sigma_\epsilon^{-\frac{1}{2}}\right)$$

$$= \left(I_{d^2m} - \left(I_m \otimes \Sigma_\epsilon^{-\frac{1}{2}}\right)(K_m K_\infty^{-1} K_m')\left(I_m \otimes \Sigma_\epsilon^{-\frac{1}{2}}\right) \otimes I_d\right)^{-1} := \left(I_{d^2m} - \Omega\right)^{-1}, say.$$

Since Σ_e is positive definite, then Σ_ϵ and K_∞ are positive definite. As a consequence Ω is positive semidefinite, and it follows that

$$\delta'\delta - \delta'\left(I_{d^2m} - \Omega\right)\delta \geq 0,$$

where $\delta = \left(I_m \otimes \Sigma_\epsilon^{-\frac{1}{2}} \otimes \Sigma_\epsilon^{-\frac{1}{2}}\right)\varrho$, and

$$\delta'\delta - \delta'\left(I_{d^2m} - \Omega\right)^{-1}\delta \leq 0. \tag{2.12}$$

It is easy to check that $q_{LM}(x) = x/2\{1 + o(1)\}$ for large values of x, since the asymptotic law of the LM test statistic is χ_m^2 with $m \geq 1$. Similarly $q_{BP}(x) := -\log P(\chi^2_{d^2(m-p+1)-dr_0} > x) = x/2\{1 + o(1)\}$. Then from (2.10), (2.11) and (2.12) we have $c_{LM}(\varrho) > c_{BP}(\varrho)$. $\square$

From Proposition 1 it turns out that the LM test is asymptotically more powerful than the portmanteau tests against linear alternatives. On the other hand Hatemi (2004) and Brüggeman *et al* (2006) considered VAR alternatives in their Monte Carlo experiments, and found some power advantages for the portmanteau tests. These empirical results may appear surprising in view of (2.8) and Proposition 1. Such paradoxical observations may be explained by the fact that the asymptotic advantage is particulary noticeable when the J^{-1} matrix has high eigenvalues (see equation (2.12)). In view of (2.4) and (2.5) this is likely to occur when m takes relatively large values as J becomes close to a singular matrix in such a case. Nevertheless from (2.9) the LM test is also likely to lose the control of the type I errors for large m. This is consistent with the fact that in the literature, it is advised to avoid the LM test for large m. It is obvious that a test that has a good type I errors control should be preferred to a more powerful test but with a bad control of the type I error. Finally recall that when m is small, the LM test is found to control the type I errors better than the portmanteau tests based on the $\chi^2_{d^2(m-p)}$ approximation (see Hatemi (2004) and Brüggeman *et al* (2006)). In this case it is likely that J^{-1} has relatively small eigenvalues, in such a way that the LM test is able to display good size results, but also possibly some *finite sample* power disadvantages. In the next section we provide an illustration of the above discussion.

3 Numerical illustrations

In this part we give an illustration of the above discussion. To this aim we simulated $N = 1000$ independent trajectories of the following process

$$\Delta X_t = \Pi_0 X_{t-1} + \Gamma_{01} \Delta X_{t-1} + \epsilon_t,$$

where

$$\Pi_0 = \begin{pmatrix} -0.2 & 0.2 & 0 \\ 0 & -0.2 & 0.2 \\ 0 & 0 & 0 \end{pmatrix}, \quad \Gamma_{01} = \begin{pmatrix} 0.5 & -0.2 & -0.2 \\ -0.2 & 0.5 & -0.2 \\ -0.2 & -0.2 & 0.5 \end{pmatrix}$$

and the errors are iid with $\epsilon_t \sim \mathcal{N}(0, I_d)$. For each iteration the supremum of the eigenvalues of the matrix $R\hat{J}^{-1}R'$ is computed under the null hypothesis. Note that in our case since $\Sigma_\epsilon = I_d$, we have $RJ^{-1}R' = (I_{d^2m} - \Omega)^{-1}$. The medians over the $N = 1000$ iterations are given in Table 1.

It can be seen that for large m the matrix $\hat{J}$ becomes difficult to invert and unstable. In such a case it is well known that the LM test has to be avoided. On the other hand if an alternative is fixed such that $\varrho \approx 0$, we would obtain similar results to those of Table 1. The term giving the theoretical power advantage of the LM seems more limited for small m. In particular from the above section recall that the critical values for the portmanteau tests are smaller than those of the LM test. Following the simulation results of Hatemi (2004) and Brüggeman *et al* (2006), it is likely that these observations lead to some finite sample power advantages for the portmanteau tests.

Table 1: The medians of the sup of the eigenvalues of $R\hat{J}^{-1}R'$.

T	$m=1$	$m=2$	$m=3$	$m=4$	$m=5$	$m=10$
100	83.18	299.26	390.73	480.21	582.27	1532.76
200	103.10	441.52	539.65	585.50	676.91	1520.54
400	110.29	892.64	1066.968	1120.17	1165.65	2076.11

References

Bahadur, R.R. (1960) Stochastic comparison of tests. *Annals of Mathematical Statistics* 31, 276-295.

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

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.

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

GONZALO, J. AND PITARAKIS J.-Y. (1998) Specification via Model Selection in Vector Error Correction Models. *Economics Letters* 60, 321-328.

HATEMI A.J. (2004) Multivariate tests for autocorrelation in the stable and unstable VAR models. *Economic Modelling* 21, 661-683.

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

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

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.

LÜTKEPOHL, H. AND KRÄTZIG (2004) *Applied Time Series Econometrics.* Cambridge University Press, New York.

PFAFF, B. (2008) VAR, SVAR and SVEC Models: Implementation Within R Package vars. *Journal of Statistical Software* 27, 1-25.

Stock, J.H., and Watson, M.W. (1989) Interpreting the evidence on money-income causality. *Journal of Econometrics* 40, 161-181.