ESSAYS IN HONOUR OF FABIO CANOVA

ADVANCES IN ECONOMETRICS

Series Editors: Thomas B. Fomby, R. Carter Hill, Ivan Jeliazkov, Juan Carlos Escanciano, Eric Hillebrand, Daniel L. Millimet, Rodney Strachan

Recent Volumes:

Volume 25	Nonparametric Econometric Methods – Edited by Qi Li and Jeffrey S. Racine
Volume 26	Maximum Simulated Likelihood Methods and Applications – Edited by R. Carter Hill and William Greene
Volume 27A	Missing Data Methods: Cross-sectional Methods and Applications – Edited by David M. Drukker
Volume 27B	Missing Data Methods: Time-series Methods and Applications – Edited by David M. Drukker
Volume 28	DSGE Models in Macroeconomics: Estimation, Evaluation and New Developments – Edited by Nathan Balke, Fabio Canova, Fabio Milani and Mark Wynne
Volume 29	Essays in Honor of Jerry Hausman – Edited by Badi H. Baltagi, Whitney Newey, Hal White and R. Carter Hill
Volume 30	30th Anniversary Edition – Edited by Dek Terrell and Daniel Millmet
Volume 31	Structural Econometric Models – Edited by Eugene Choo and Matthew Shum
Volume 32	VAR Models in Macroeconomics — New Developments and Applications: Essays in Honor of Christopher A. Sims – Edited by Thomas B. Fomby, Lutz Kilian and Anthony Murphy
Volume 33	Essays in Honor of Peter C. B. Phillips – Edited by Thomas B. Fomby, Yoosoon Chang and Joon Y. Park
Volume 34	Bayesian Model Comparison – Edited by Ivan Jeliazkov and Dale J. Poirier
Volume 35	Dynamic Factor Models – Edited by Eric Hillebrand and Siem Jan Koopman
Volume 36	Essays in Honor of Aman Ullah – Edited by Gloria Gonzalez-Rivera, R. Carter Hill and Tae-Hwy Lee
Volume 37	Spatial Econometrics – Edited by Badi H. Baltagi, James P. LeSage and R. Kelley Pace
Volume 38	Regression Discontinuity Designs: Theory and Applications – Edited by Matias D. Cattaneo and Juan Carlos Escanciano
Volume 39	The Econometrics of Complex Survey Data: Theory and Applications – Edited by Kim P. Huynh, David T. Jacho-Chávez and Guatam Tripathi
Volume 40A	Topics in Identification, Limited Dependent Variables, Partial Observability, Experimentation, and Flexible Modelling Part A – Edited by Ivan Jeliazkov and Justin L. Tobias
Volume 40B	Topics in Identification, Limited Dependent Variables, Partial Observability, Experimentation, and Flexible Modelling Part B – Edited by Ivan Jeliazkov and Justin L. Tobias
Volume 41	Essays in Honor of Cheng Hsiao – Edited by Tong Li, M. Hashem Pesaran and Dek Terrell
Volume 42	The Econometrics of Networks – Edited by Áureo de Paula, Elie Tamer and Marcel-Cristian Voia
Volume 43A	Essays in Honor of M. Hashem Pesaran – Edited by Alexander Chudik, Cheng Hsiao and Allan Timmermann
Volume 43B	Essays in Honor of M. Hashem Pesaran – Edited by Alexander Chudik, Cheng Hsiao and Allan Timmermann
Volume 44A	Essays in Honour of Fabio Canova – Edited by Juan J. Dolado, Luca Gambetti and Christian Matthes

ADVANCES IN ECONOMETRICS VOLUME 44B

ESSAYS IN HONOUR OF FABIO CANOVA

EDITED BY

JUAN J. DOLADO
Universidad Carlos III de Madrid, Spain

LUCA GAMBETTI
Universitat Autónoma de Barcelona, Spain
Universitá di Torino, Italy

And

CHRISTIAN MATTHES
Indiana University, USA

emerald PUBLISHING

United Kingdom – North America – Japan
India – Malaysia – China

Emerald Publishing Limited
Howard House, Wagon Lane, Bingley BD16 1WA, UK

First edition 2022

Reprints and permissions service
Contact: permissions@emeraldinsight.com

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

ISBN: 978-1-80382-832-9 (Print)
ISBN: 978-1-80382-831-2 (Online)
ISBN: 978-1-80382-833-6 (Epub)

ISSN: 0731-9053 (Series)

CONTENTS

This page intentionally left blank

CHAPTER 1

TESTS FOR RANDOM COEFFICIENT VARIATION IN VECTOR AUTOREGRESSIVE MODELS

Dante Amengual[a], Gabriele Fiorentini[b] and Enrique Sentana[c]

[a]CEMFI, Madrid, Spain
[b]Università di Firenze and RCEA, Firenze, Italy
[c]CEMFI, Madrid, Spain

ABSTRACT

The authors propose the information matrix test to assess the constancy of mean and variance parameters in vector autoregressions (VAR). They additively decompose it into several orthogonal components: conditional heteroskedasticity and asymmetry of the innovations, and their unconditional skewness and kurtosis. Their Monte Carlo simulations explore both its finite size properties and its power against i.i.d. coefficients, persistent but stationary ones, and regime switching. Their procedures detect variation in the autoregressive coefficients and residual covariance matrix of a VAR for the US GDP growth rate and the statistical discrepancy, but they fail to detect any covariation between those two sets of coefficients.

Keywords: Gross domestic product; gross domestic income; Hessian matrix; information matrix test; outer product of the score

JEL classifications: C32; C52; E01

Essays in Honour of Fabio Canova
Advances in Econometrics, Volume 44B, 1–35

ISSN: 0731-9053/doi:10.1108/S0731-90532022000044B001

1. INTRODUCTION

Following the path-breaking work by Sims (1980), vector autoregressions (VARs) have become one of the most commonly employed tools by empirical macroeconomists in academia, central banks and research departments of financial institutions. Their ability to capture dynamic relationships between multiple variables makes them particularly apt for short- and medium-run economic analysis, either on their own or in combination with structural macroeconomic models (see Canova, 2007, for a textbook treatment). In fact, VARs have become the forecasting benchmark to beat, thereby replacing the univariate ARIMA models in vogue in the 1970s and 1980s.

Unlike those univariate models, though, VARs are hardly ever subject to a battery of specification tests. Part of the reason is that specification testing does not fit well with the Bayesian approach to inference predominant among macroeconometricians. But the scarcity of specification tests for those models also plays an important role. The purpose of this chapter is precisely to apply the information matrix test of White (1982) to VARs.

Our choice of specification test is far from random. The neglected heterogeneity interpretation of the information matrix test in Chesher (1984) provides a very relevant justification in macroeconomic applications, in which changes in the structure of the economy are a first-order concern (see e.g. Perron, 1989). There is, in fact, a long tradition of autoregressive models with time-varying parameters, which are sometimes called Random Coefficient Autoregressions (RCAs) in the time series literature (see e.g. Nicholls & Quinn, 1982, for an earlier treatment and Regis, Serra, & van den Heuvel, 2021, for a recent survey). Moreover, in recent years the macroeconometric literature has paid considerable attention to models in which not only the parameters governing the conditional mean change over time, but also the parameters corresponding to the variances and covariances of the innovations may also time-vary (see e.g. Primiceri, 2005 or D'Agostino, Gambetti, & Giannone, 2013).

Several tests for constant versus random coefficients in autoregressive models already exist in the literature, mostly in the univariate case (see e.g. Akharif & Hallin, 2003; Chen, Wang, Li, & Huang, 2020; Horváth & Trapani, 2019; Lee, 1998). Some of them use a likelihood framework, but they tend to focus on the classical triad of Wald, likelihood ratio and Lagrange multiplier (LM) tests, which have a somewhat non-standard distribution under the null. There is also a huge literature on structural break tests, as well as on testing for recurrent regime switches (see, respectively, Hansen, 2001 and Carrasco, Hu, & Ploberger, 2014, and the references therein).

Our information matrix test, though, has two main advantages: (i) unlike many of those tests, the test statistic has an asymptotic chi-squared distributions under the null when the autoregressive process is covariance stationary; and (ii) it can be additively decomposed into four easily interpretable orthogonal components: (a) a test for conditional heteroskedasticity of the innovations; (b) a test for conditional asymmetry of those innovations; (c) a test for their unconditional skewness; and (d) a test for their unconditional kurtosis. These four tests can be combined into other easily interpretable tests. For example, the sum of (b) and (c) assesses the null hypothesis of zero covariance between the mean and variance parameters. Similarly, the sum of (c) and (d) checks the multivariate normality of the innovations.

An additional advantage of the information matrix test is that it can capture multiple types of deviations from constant coefficients even though it is not specifically designed for them. For that reason, we conduct an extensive Monte Carlo exercise in which we study the power of the test against three types of random coefficient variation: *i.i.d.* coefficients, as in RCAs, persistent but stationary coefficients, and finally regime switching models. In all cases, we calibrate the designs so that the unconditional variances of the coefficients is the same across these three alternatives, and compute critical values under the null using the bootstrap.

Finally, we apply our procedures to test the parameter constancy of a VAR in an important empirical context. Specifically, we study the dynamic relationship between the equally weighted average of the growth rates of the expenditure and income measures of US gross domestic product (GDP) produced by the Bureau of Economic Analysis (BEA), and the statistical discrepancy, which is the difference between the (log) levels of those two measurements. Thus, we follow Almuzara, Amengual, and Sentana (2019) and Almuzara, Fiorentini, and Sentana (2022) in imposing cointegration between the gross domestic expenditure (GDE) and gross domestic income (GDI) measures. The empirical results that we obtain with our proposed information matrix tests confirm time variation in both the autoregressive coefficients and the residual covariance matrix of the innovations, but they fail to detect any covariance between those two groups of coefficients.

The rest of this chapter is organized as follows. In Section 2, we derive the information matrix test for a multivariate linear regression model, while in Section 3 we specialize it to VARs. We then present the results of our simulation experiments on its size and power properties in Section 4, and apply it to US GDE and GDI in Section 5. Our conclusions appear in Section 6, followed by appendices that contain proofs together with some additional material.

2. INFORMATION MATRIX TESTS OF THE MULTIVARIATE REGRESSION MODEL

Consider the following multivariate normal regression model:

$$\mathbf{y}_t = \mathbf{B}\mathbf{x}_t + \boldsymbol{\Omega}^{1/2}\varepsilon_t^*, \tag{1}$$

where the vector of dependent variables $\mathbf{y}_t$ is $N \times 1$, the vector of regressors $\mathbf{x}_t$, which often includes a constant, is $M \times 1$, the matrix of regression coefficients $\mathbf{B}$ is $N \times M$, the residual covariance matrix $\boldsymbol{\Omega}$ is $N \times N$ symmetric and positive definite, and the vector of standardized innovations ε_t^* follows a spherical normal distribution conditional on the regressors and the past values of the observed variables. Thus, the conditional mean vector and covariance matrix of $\mathbf{y}_t$ will be $\boldsymbol{\mu}_t(\boldsymbol{\theta})=\mathbf{B}\mathbf{x}_t$ and $\boldsymbol{\Sigma}_t(\boldsymbol{\theta})=\boldsymbol{\Omega}$, respectively, where $\boldsymbol{\theta}=(\mathbf{b}',\boldsymbol{\omega}')'$, $\mathbf{b}=\text{vec}(\mathbf{B})$ and $\boldsymbol{\omega}=\text{vec}h(\boldsymbol{\Omega})$.

Given these assumptions, the contribution from a single observation to the log-likelihood function is

$$-\frac{N}{2}\ln(2\pi)-\frac{1}{2}\ln|\boldsymbol{\Omega}|-\frac{1}{2}(\mathbf{y}_t-\mathbf{B}\mathbf{x}_t)'\boldsymbol{\Omega}^{-1}(\mathbf{y}_t-\mathbf{B}\mathbf{x}_t)=-\frac{N}{2}\ln(2\pi)-\frac{1}{2}\ln|\boldsymbol{\Omega}|-\frac{1}{2}\varsigma_t(\boldsymbol{\theta}),$$

where $\varsigma_t(\boldsymbol{\theta})=\boldsymbol{\varepsilon}_t^{*'}(\boldsymbol{\theta})\boldsymbol{\varepsilon}_t^*(\boldsymbol{\theta})$, $\boldsymbol{\varepsilon}_t^*(\boldsymbol{\theta})=\boldsymbol{\Omega}^{-1/2}\boldsymbol{\varepsilon}_t(\boldsymbol{\theta})$ and $\boldsymbol{\varepsilon}_t(\boldsymbol{\theta})=\mathbf{y}_t-\mathbf{B}\mathbf{x}_t$.

The maximum likelihood estimators of the model parameters are known in closed form without the need to conduct any numerical optimization. Specifically,

$$\hat{\mathbf{B}}=\left(\sum_{t=1}^{T}\mathbf{y}_t\mathbf{x}_t'\right)\left(\sum_{t=1}^{T}\mathbf{x}_t\mathbf{x}_t'\right)^{-1}$$

and

$$\hat{\boldsymbol{\Omega}}=\frac{1}{T}\left[\sum_{t=1}^{T}(\mathbf{y}_t-\hat{\mathbf{B}}\mathbf{x}_t)(\mathbf{y}_t-\hat{\mathbf{B}}\mathbf{x}_t)'\right].$$

Nevertheless, we need expressions for the score and Hessian matrix to be able to derive the information matrix test.

To compute the score, we first differentiate $\boldsymbol{\mu}_t(\boldsymbol{\theta})$ and $\boldsymbol{\Sigma}_t(\boldsymbol{\theta})$ with respect to the $q=MN+N(N+1)/2$ model parameters in $\boldsymbol{\theta}$. Specifically, the first derivatives are given by,

$$\frac{\partial\boldsymbol{\mu}_t(\boldsymbol{\theta})}{\partial\mathbf{b}'}=\mathbf{x}_t'\otimes\mathbf{I}_N$$
$$\frac{\partial\text{vec}[\boldsymbol{\Sigma}_t(\boldsymbol{\theta})]}{\partial\boldsymbol{\omega}'}=\mathbf{D}_N,$$

where $\mathbf{D}_N$ is the duplication matrix of order N (see Magnus & Neudecker, 2019). Thus, the log-likelihood score is:

$$\mathbf{s}_t(\boldsymbol{\theta})=\mathbf{Z}_{lt}(\boldsymbol{\theta})\boldsymbol{\varepsilon}_t^*(\boldsymbol{\theta})+\mathbf{Z}_{st}(\boldsymbol{\theta})\text{vec}[\boldsymbol{\varepsilon}_t^*(\boldsymbol{\theta})\boldsymbol{\varepsilon}_t^{*'}(\boldsymbol{\theta})-\mathbf{I}_N],$$

where

$$\mathbf{Z}_{lt}(\boldsymbol{\theta})=\begin{bmatrix}\mathbf{x}_t\otimes\boldsymbol{\Omega}^{-1/2'}\\ \mathbf{0}\end{bmatrix},$$

$$\mathbf{Z}_{st}(\boldsymbol{\theta})=\begin{bmatrix}\mathbf{0}\\ \frac{1}{2}\mathbf{D}_N'(\boldsymbol{\Omega}^{-1/2'}\otimes\boldsymbol{\Omega}^{-1/2'})\end{bmatrix}.$$

As a result, the scores will be,

$$\begin{aligned} \mathbf{s}_{\mathbf{b}t}(\boldsymbol{\theta}) &= [\mathbf{x}_t \otimes \boldsymbol{\Omega}^{-1/2'} \boldsymbol{\varepsilon}_t^*(\boldsymbol{\theta})] = [\mathbf{x}_t \otimes \boldsymbol{\Omega}^{-1}(\mathbf{y}_t - \mathbf{B}\mathbf{x}_t)] \\ &= \mathrm{vec}[\boldsymbol{\Omega}^{-1}(\mathbf{y}_t - \mathbf{B}\mathbf{x}_t)\mathbf{x}_t'] \end{aligned} \tag{2}$$

and

$$\begin{aligned} \mathbf{s}_{\omega t}(\boldsymbol{\theta}) &= \frac{1}{2}\mathbf{D}_N'(\boldsymbol{\Omega}^{-1/2'} \otimes \boldsymbol{\Omega}^{-1/2'})\mathrm{vec}[\boldsymbol{\varepsilon}_t^*(\boldsymbol{\theta})\boldsymbol{\varepsilon}_t^{*'}(\boldsymbol{\theta}) - \mathbf{I}_N] \\ &= \frac{1}{2}\mathbf{D}_N'\mathrm{vec}[\boldsymbol{\Omega}^{-1}(\mathbf{y}_t - \mathbf{B}\mathbf{x}_t)(\mathbf{y}_t - \mathbf{B}\mathbf{x}_t)'\boldsymbol{\Omega}^{-1} - \boldsymbol{\Omega}^{-1}]. \end{aligned} \tag{3}$$

Consequently, the outer product of the scores will be,

$$\begin{aligned} \mathbf{s}_{\mathbf{b}t}(\boldsymbol{\theta})\mathbf{s}_{\mathbf{b}t}'(\boldsymbol{\theta}) &= [\mathbf{x}_t\mathbf{x}_t' \otimes \boldsymbol{\Omega}^{-1/2'}\boldsymbol{\varepsilon}_t^*(\boldsymbol{\theta})\boldsymbol{\varepsilon}_t^{*'}(\boldsymbol{\theta})\boldsymbol{\Omega}^{-1/2}] \\ &= [\mathbf{x}_t\mathbf{x}_t' \otimes \boldsymbol{\Omega}^{-1}(\mathbf{y}_t - \mathbf{B}\mathbf{x}_t)(\mathbf{y}_t - \mathbf{B}\mathbf{x}_t)'\boldsymbol{\Omega}^{-1}] \end{aligned}$$

$$\begin{aligned} \mathbf{s}_{\omega t}(\boldsymbol{\theta})\mathbf{s}_{\mathbf{b}t}'(\boldsymbol{\theta}) &= \frac{1}{2}\mathbf{D}_N'(\boldsymbol{\Omega}^{-1/2'} \otimes \boldsymbol{\Omega}^{-1/2'})\mathrm{vec}[\boldsymbol{\varepsilon}_t^*(\boldsymbol{\theta})\boldsymbol{\varepsilon}_t^{*'}(\boldsymbol{\theta}) - \mathbf{I}_N][\mathbf{x}_t' \otimes \boldsymbol{\varepsilon}_t^{*'}(\boldsymbol{\theta})\boldsymbol{\Omega}^{-1/2}] \\ &= \frac{1}{2}\mathbf{D}_N'\mathrm{vec}[\boldsymbol{\Omega}^{-1}(\mathbf{y}_t - \mathbf{B}\mathbf{x}_t)(\mathbf{y}_t - \mathbf{B}\mathbf{x}_t)'\boldsymbol{\Omega}^{-1} - \boldsymbol{\Omega}^{-1}][\mathbf{x}_t' \otimes (\mathbf{y}_t - \mathbf{B}\mathbf{x}_t)'\boldsymbol{\Omega}^{-1}] \end{aligned}$$

and

$$\begin{aligned} \mathbf{s}_{\omega t}(\boldsymbol{\theta})\mathbf{s}_{\omega t}'(\boldsymbol{\theta}) &= \frac{1}{4}\mathbf{D}_N'(\boldsymbol{\Omega}^{-1/2'} \otimes \boldsymbol{\Omega}^{-1/2'})\mathrm{vec}[\boldsymbol{\varepsilon}_t^*(\boldsymbol{\theta})\boldsymbol{\varepsilon}_t^{*'}(\boldsymbol{\theta}) - \mathbf{I}_N] \\ &\times \mathrm{vec}'[\boldsymbol{\varepsilon}_t^*(\boldsymbol{\theta})\boldsymbol{\varepsilon}_t^{*'}(\boldsymbol{\theta}) - \mathbf{I}_N](\boldsymbol{\Omega}^{-1/2'} \otimes \boldsymbol{\Omega}^{-1/2'})\mathbf{D}_N \\ &= \frac{1}{4}\mathbf{D}_N'\mathrm{vec}[\boldsymbol{\Omega}^{-1}(\mathbf{y}_t - \mathbf{B}\mathbf{x}_t)(\mathbf{y}_t - \mathbf{B}\mathbf{x}_t)'\boldsymbol{\Omega}^{-1} - \boldsymbol{\Omega}^{-1}] \\ &\times \mathrm{vec}'[\boldsymbol{\Omega}^{-1}(\mathbf{y}_t - \mathbf{B}\mathbf{x}_t)(\mathbf{y}_t - \mathbf{B}\mathbf{x}_t)'\boldsymbol{\Omega}^{-1} - \boldsymbol{\Omega}^{-1}]\mathbf{D}_N. \end{aligned}$$

To compute the Hessian, it is convenient to use the general expressions for elliptical distributions in Supplementary Appendix C of Fiorentini and Sentana (2021), namely,

$$\mathbf{h}_{\theta\theta t}(\phi) = \frac{\partial^2 d_t(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}\,\partial\boldsymbol{\theta}'} + \frac{\partial^2 g[\varsigma_t(\boldsymbol{\theta}),\boldsymbol{\eta}]}{(\partial\varsigma)^2}\frac{\partial\varsigma_t(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\frac{\partial\varsigma_t(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}'} + \frac{\partial g[\varsigma_t(\boldsymbol{\theta}),\boldsymbol{\eta}]}{\partial\varsigma}\frac{\partial^2\varsigma_t(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}\,\partial\boldsymbol{\theta}'},$$

where

$$\partial^2 d_t(\boldsymbol{\theta})/\partial\boldsymbol{\theta}\,\partial\boldsymbol{\theta}' = 2\mathbf{Z}_{st}(\boldsymbol{\theta})\mathbf{Z}_{st}'(\boldsymbol{\theta}) - \frac{1}{2}\left\{\mathrm{vec}'\left[\boldsymbol{\Sigma}_t^{-1}(\boldsymbol{\theta})\right]\otimes\mathbf{I}_q\right\}\partial\mathrm{vec}\left\{\partial\mathrm{vec}'\left[\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\right]/\partial\boldsymbol{\theta}\right\}/\partial\boldsymbol{\theta}'$$

and

$$\begin{aligned}\partial^2 \varsigma_t(\boldsymbol{\theta})/\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}' &= 2\mathbf{Z}_{lt}(\boldsymbol{\theta})\mathbf{Z}_{lt}'(\boldsymbol{\theta}) + 8\mathbf{Z}_{st}(\boldsymbol{\theta})[\mathbf{I}_N \otimes \boldsymbol{\varepsilon}_t^*(\boldsymbol{\theta})\boldsymbol{\varepsilon}_t^{*'}(\boldsymbol{\theta})]\mathbf{Z}_{st}'(\theta)\\ &+4\mathbf{Z}_{lt}(\boldsymbol{\theta})[\boldsymbol{\varepsilon}_t^{*'}(\boldsymbol{\theta})\otimes\mathbf{I}_N]\mathbf{Z}_{st}'(\boldsymbol{\theta}) + 4\mathbf{Z}_{st}(\boldsymbol{\theta})[\boldsymbol{\varepsilon}_t^{*}(\boldsymbol{\theta})\otimes\mathbf{I}_N]\mathbf{Z}_{lt}'(\boldsymbol{\theta})\\ &-2[\boldsymbol{\varepsilon}_t^{*'}(\boldsymbol{\theta})\boldsymbol{\Sigma}_t^{-\frac{1}{2}'}(\boldsymbol{\theta})\otimes\mathbf{I}_q]\partial\mathrm{vec}[\partial\boldsymbol{\mu}_t'(\boldsymbol{\theta})/\partial\boldsymbol{\theta}]\partial\boldsymbol{\theta}'\\ &-\{\mathrm{vec}'[\boldsymbol{\Sigma}_t^{-\frac{1}{2}}(\boldsymbol{\theta})\boldsymbol{\varepsilon}_t^*(\boldsymbol{\theta})\boldsymbol{\varepsilon}_t^{*'}(\boldsymbol{\theta})\boldsymbol{\Sigma}_t^{-\frac{1}{2}'}(\boldsymbol{\theta})]\otimes\mathbf{I}_q\}\partial\mathrm{vec}\{\partial\mathrm{vec}'[\boldsymbol{\Sigma}_t(\boldsymbol{\theta})]/\partial\boldsymbol{\theta}\}/\partial\boldsymbol{\theta}'.\end{aligned}$$

In the case of model (1), $d_t(\boldsymbol{\theta}) = -\frac{1}{2}\ln|\boldsymbol{\Omega}|$ and

$$\partial^2 d_t(\boldsymbol{\theta})/\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}' = \frac{1}{2}\begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_N'(\boldsymbol{\Omega}^{-1}\otimes\boldsymbol{\Omega}^{-1})\mathbf{D}_N \end{bmatrix}.$$

Similarly, we have that $g[\varsigma_t(\boldsymbol{\theta}),\boldsymbol{\eta}] = -\frac{1}{2}\varsigma_t(\theta)$ under normality, so that $\partial g[\varsigma_t(\boldsymbol{\theta}),\boldsymbol{\eta}]/\partial\varsigma = -\frac{1}{2}$ and $\partial^2 g[\varsigma_t(\boldsymbol{\theta}),\boldsymbol{\eta}]/(\partial\varsigma)^2 = 0$. Finally,

$$\partial^2 \varsigma_t(\boldsymbol{\theta})/\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}' = 2\begin{pmatrix} \mathbf{x}_t\mathbf{x}_t'\otimes\boldsymbol{\Omega}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}$$

$$+2\begin{Bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_N'(\boldsymbol{\Omega}^{-1/2'}\otimes\boldsymbol{\Omega}^{-1/2'})[\mathbf{I}_N\otimes\boldsymbol{\varepsilon}_t^*(\boldsymbol{\theta})\boldsymbol{\varepsilon}_t^{*'}(\boldsymbol{\theta})](\boldsymbol{\Omega}^{-1/2}\otimes\boldsymbol{\Omega}^{-1/2})\mathbf{D}_N \end{Bmatrix}$$

$$+2\begin{Bmatrix} \mathbf{0} & (\mathbf{x}_t\otimes\boldsymbol{\Omega}^{-1/2'})[\boldsymbol{\varepsilon}_t^{*'}(\boldsymbol{\theta})\otimes\mathbf{I}_N](\boldsymbol{\Omega}^{-1/2}\otimes\boldsymbol{\Omega}^{-1/2})\mathbf{D}_N \\ \mathbf{0} & \mathbf{0} \end{Bmatrix}$$

$$+2\begin{Bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{D}_N'(\boldsymbol{\Omega}^{-1/2'}\otimes\boldsymbol{\Omega}^{-1/2'})[\boldsymbol{\varepsilon}_t^*(\theta)\otimes\mathbf{I}_N](\mathbf{x}_t'\otimes\boldsymbol{\Omega}^{-1/2}) & \mathbf{0} \end{Bmatrix}$$

$$= 2\begin{Bmatrix} (\mathbf{x}_t\mathbf{x}_t'\otimes\boldsymbol{\Omega}^{-1}) & [\mathbf{x}_t(\mathbf{y}_t-\mathbf{B}\mathbf{x}_t)'\boldsymbol{\Omega}^{-1}\otimes\boldsymbol{\Omega}^{-1}]\mathbf{D}_N \\ \mathbf{D}_N'[\boldsymbol{\Omega}^{-1}(\mathbf{y}_t-\mathbf{B}\mathbf{x}_t)\mathbf{x}_t'\otimes\boldsymbol{\Omega}^{-1}] & \mathbf{D}_N'[\boldsymbol{\Omega}^{-1}\otimes\boldsymbol{\Omega}^{-1}(\mathbf{y}_t-\mathbf{B}\mathbf{x}_t)(\mathbf{y}_t-\mathbf{B}\mathbf{x}_t)'\boldsymbol{\Omega}^{-1}]\mathbf{D}_N \end{Bmatrix}$$

where we have exploited the fact that the second derivatives of the conditional mean and covariance functions with respect to the model parameters are all 0.

Therefore, we can write the Hessian matrix as

$$-\begin{Bmatrix} (\mathbf{x}_t\mathbf{x}_t'\otimes\boldsymbol{\Omega}^{-1}) & [\mathbf{x}_t(\mathbf{y}_t-\mathbf{B}\mathbf{x}_t)'\boldsymbol{\Omega}^{-1}\otimes\boldsymbol{\Omega}^{-1}]\mathbf{D}_N \\ \mathbf{D}_N'[\boldsymbol{\Omega}^{-1}(\mathbf{y}_t-\mathbf{B}\mathbf{x}_t)\mathbf{x}_t'\otimes\boldsymbol{\Omega}^{-1}] & \mathbf{D}_N'\{\boldsymbol{\Omega}^{-1}\otimes[\boldsymbol{\Omega}^{-1}(\mathbf{y}_t-\mathbf{B}\mathbf{x}_t)(\mathbf{y}_t-\mathbf{B}\mathbf{x}_t)'\boldsymbol{\Omega}^{-1}-\frac{1}{2}\boldsymbol{\Omega}^{-1}]\}\mathbf{D}_N \end{Bmatrix}.$$

The sum of the outer product of the score and the Hessian yields the following three terms:

$$\mathbf{bb}:\ [\mathbf{x}_t\mathbf{x}_t' \otimes \mathbf{\Omega}^{-1}(\mathbf{y}_t - \mathbf{B}\mathbf{x}_t)(\mathbf{y}_t - \mathbf{B}\mathbf{x}_t)'\mathbf{\Omega}^{-1}] - (\mathbf{x}_t\mathbf{x}_t' \otimes \mathbf{\Omega}^{-1}), \tag{4}$$

$$\begin{aligned}\boldsymbol{\omega}\mathbf{b}:\ &\frac{1}{2}\mathbf{D}_N'\mathrm{vec}[\mathbf{\Omega}^{-1}(\mathbf{y}_t - \mathbf{B}\mathbf{x}_t)(\mathbf{y}_t - \mathbf{B}\mathbf{x}_t)'\mathbf{\Omega}^{-1} - \mathbf{\Omega}^{-1}][\mathbf{x}_t' \otimes (\mathbf{y}_t - \mathbf{B}\mathbf{x}_t)'\mathbf{\Omega}^{-1}]\\ &-\mathbf{D}_N'[\mathbf{\Omega}^{-1}(\mathbf{y}_t - \mathbf{B}\mathbf{x}_t)\mathbf{x}_t' \otimes \mathbf{\Omega}^{-1}],\end{aligned} \tag{5}$$

and

$$\begin{aligned}\boldsymbol{\omega}\boldsymbol{\omega}:\ &\frac{1}{4}\mathbf{D}_N'\mathrm{vec}[\mathbf{\Omega}^{-1}(\mathbf{y}_t - \mathbf{B}\mathbf{x}_t)(\mathbf{y}_t - \mathbf{B}\mathbf{x}_t)'\mathbf{\Omega}^{-1} - \mathbf{\Omega}^{-1}]\\ &\times\mathrm{vec}'[\mathbf{\Omega}^{-1}(\mathbf{y}_t - \mathbf{B}\mathbf{x}_t)(\mathbf{y}_t - \mathbf{B}\mathbf{x}_t)'\mathbf{\Omega}^{-1} - \mathbf{\Omega}^{-1}]\mathbf{D}_N\\ &-\mathbf{D}_N'\{\mathbf{\Omega}^{-1} \otimes [\mathbf{\Omega}^{-1}(\mathbf{y}_t - \mathbf{B}\mathbf{x}_t)(\mathbf{y}_t - \mathbf{B}\mathbf{x}_t)'\mathbf{\Omega}^{-1} - \frac{1}{2}\mathbf{\Omega}^{-1}]\}\mathbf{D}_N.\end{aligned} \tag{6}$$

When $x_t = 1$, these formulas coincide with those in Amengual, Fiorentini, and Sentana (2021), who re-write the $\boldsymbol{\omega}\mathbf{b}$ and $\boldsymbol{\omega}\boldsymbol{\omega}$ expressions in terms of multivariate Hermite polynomials of orders 3 and 4, respectively.[1] We can generalize their results for any $\mathbf{x}_t$ as follows. As in Barndorff-Nielsen and Petersen (1979), define the (centred) multivariate Hermite polynomials of ε of order $k = k_1 + \ldots + k_N \geq 0$ as

$$H_{\underset{1\ldots1\ldots N\ldots N}{k_1\quad k_N}}(\varepsilon;\mathbf{\Delta})\cdot e^{-\frac{1}{2}\varepsilon'\mathbf{\Delta}\varepsilon} = (-1)^k \frac{\partial^k}{(\partial\varepsilon_1)^{k_1}\ldots(\partial\varepsilon_N)^{k_N}}\left(e^{-\frac{1}{2}\varepsilon'\mathbf{\Delta}\varepsilon}\right), \tag{7}$$

where $\mathbf{\Delta} = \mathbf{\Omega}^{-1}$ is the inverse covariance matrix of ε. As is well known, when model (1) is correctly specified: (i) the expected value of any multivariate Hermite polynomial of positive degree k conditional on the regressors and the past values of the observed variables is 0; and (ii) the conditional and unconditional covariance matrices of those polynomials coincide.

Let

$$\mathbf{H}_k(\varepsilon;\mathbf{\Delta}) = \begin{bmatrix} H_{k,0,\cdots,0}(\varepsilon;\mathbf{\Delta}) \\ H_{k-1,1,\cdots,0}(\varepsilon;\mathbf{\Delta}) \\ \vdots \\ H_{0,\cdots,0,k}(\varepsilon;\mathbf{\Delta}) \end{bmatrix}$$

denote the $\binom{N+k-1}{k}\times 1$ vector that contains all the non-redundant multivariate Hermite polynomials of order k, which we will simply denote by $\mathbf{H}_k(\varepsilon_t^*)$ for the special case of $\mathbf{\Delta} = \mathbf{I}_N$.

Similarly, let

$$\mathbf{m}_{ht}(\boldsymbol{\theta}) = \mathbf{H}_2[\varepsilon_t^*(\boldsymbol{\theta})] \otimes \mathrm{vec}h(\mathbf{x}_t\mathbf{x}_t'), \tag{8}$$

$$\mathbf{m}_{at}(\boldsymbol{\theta}) = \mathbf{H}_3[\varepsilon_t^*(\boldsymbol{\theta})] \otimes \mathbf{x}_t, \tag{9}$$

$$\mathbf{m}_{kt}(\boldsymbol{\theta}) = \mathbf{H}_4[\varepsilon_t^*(\boldsymbol{\theta})], \tag{10}$$

which effectively span (4), (5) and (6), respectively. Finally, let

$$\bar{\mathbf{m}}_{lT}(\boldsymbol{\theta}) = \frac{1}{T}\sum_{t=1}^{T}\mathbf{m}_{lt}(\boldsymbol{\theta}) \text{ for } l = h, a, k$$

We can then state our main result:

Proposition 1. Assume $\mathbf{x}_t$ has finite fourth moments. Then, the information matrix test that compares the outer product of the score with the Hessian of the multivariate regression model (1) evaluated at the Gaussian maximum likelihood estimators $\hat{\boldsymbol{\theta}}_T = (\hat{\mathbf{b}}_T', \hat{\boldsymbol{\omega}}_T')'$ is asymptotically equivalent under the null hypothesis of correct specification to the sum of the following three moment tests:

$$h_{hT} = T \cdot \bar{\mathbf{m}}_{hT}'(\hat{\boldsymbol{\theta}}_T)\hat{V}^{+}[\mathbf{m}_{ht}(\hat{\boldsymbol{\theta}}_T)]\bar{\mathbf{m}}_{hT}(\hat{\boldsymbol{\theta}}_T), \tag{11}$$

$$h_{aT} = T \cdot \bar{\mathbf{m}}_{aT}'(\hat{\boldsymbol{\theta}}_T)\hat{V}^{-1}[\mathbf{m}_{at}(\hat{\boldsymbol{\theta}}_T)]\bar{\mathbf{m}}_{aT}(\hat{\boldsymbol{\theta}}_T), \tag{12}$$

and

$$h_{kT} = T \cdot \bar{\mathbf{m}}_{kT}'(\hat{\boldsymbol{\theta}}_T)\hat{V}^{-1}[\mathbf{m}_{kt}(\hat{\boldsymbol{\theta}}_T)]\bar{\mathbf{m}}_{kT}(\hat{\boldsymbol{\theta}}_T), \tag{13}$$

where $+$ denotes the Moore-Penrose generalized inverse,

$$\lim_{T\to\infty} V[\sqrt{T}\bar{\mathbf{m}}_{hT}(\hat{\boldsymbol{\theta}}_T)] = V\{\mathbf{H}_2[\varepsilon_t^*(\boldsymbol{\theta})]\} \otimes V\left[\mathrm{vec}h(\mathbf{x}_t\mathbf{x}_t')\right], \tag{14}$$

$$\lim_{T\to\infty} V[\sqrt{T}\bar{\mathbf{m}}_{aT}(\hat{\boldsymbol{\theta}}_T)] = V\{\mathbf{H}_3[\varepsilon_t^*(\boldsymbol{\theta})]\} \otimes E(\mathbf{x}_t\mathbf{x}_t'), \text{ and} \tag{15}$$

$$\lim_{T\to\infty} V[\sqrt{T}\bar{\mathbf{m}}_{kT}(\hat{\boldsymbol{\theta}}_T)] = V\{\mathbf{H}_4[\varepsilon_t^*(\boldsymbol{\theta})]\},$$

which converge in distribution to three independent chi-squared random variables whose degrees of freedom are $\binom{N+1}{2}\mathrm{rank}\{V[\mathrm{vec}h(\mathbf{x}_t\mathbf{x}_t')]\}$, $\binom{N+2}{3}M$ and $\binom{N+3}{4}$, respectively.

Note that if $\mathbf{x}_t$ contains either a constant or a set of dummy variables that linearly span a constant term, then $V[\mathrm{vec}h(\mathbf{x}_t\mathbf{x}_t')]$ will be singular with nullity 1, which explains the generalized inverse in (11). Given that the diagonal covariance matrices of $\mathbf{H}_k(\varepsilon_t^*)$ for $k = 2, 3, 4$ do not depend on any unknown quantities