Exact tests for contemporaneous correlation of disturbances in seemingly unrelated regressions *

Jean-Marie Dufour †
Université de Montréal

Lynda Khalaf ‡
Université Laval

First version: January 1998
Revised: April 2000
This version: July 2001
Compiled: February 10, 2005, 1:09pm

This paper has been published in the *Journal of Econometrics*, 106 (2002), 143-170.

*The authors thank Bryan Campbell, Judith Giles, Mohamed Taamouti, Allan Wurtz, Victoria Zinde-Walsh, and the Editor Arnold Zellner for several useful comments. This work was supported by the Canada Research Chair Program (Chair in Econometrics, Université de Montréal), the Canadian Network of Centres of Excellence [program on *Mathematics of Information Technology and Complex Systems* (MITACS)], the Canada Council for the Arts (Killam Fellowship), the Natural Sciences and Engineering Research Council of Canada, the Social Sciences and Humanities Research Council of Canada, and the Fonds FCAR (Government of Québec).

†Centre de recherche et développement en économique (C.R.D.E.), Centre interuniversitaire de recherche en analyse des organisations (CIRANO), and Département de sciences économiques, Université de Montréal. Mailing address: Département de sciences économiques, Université de Montréal, C.P. 6128 succursale Centre Ville, Montréal, Québec, Canada H3C 3J7. TEL: (514) 343 2400; FAX: (514) 343 5831; e-mail: jean.marie.dufour@umontreal.ca .
Web page: http://www.fas.umontreal.ca/SCECO/Dufour .

‡GREEN, Université Laval, Pavillon J.-A.-De Sève, St. Foy, Québec, Canada, G1K 7P4. TEL: (418) 656 2131-2409; FAX: (418) 656 7412; email: lynda.khalaf@ecn.ulaval.ca

ABSTRACT

This paper proposes finite-sample procedures for testing the SURE specification in multi-equation regression models, *i.e.* whether the disturbances in different equations are contemporaneously uncorrelated or not. We apply the technique of Monte Carlo (MC) tests [Dwass (1957), Barnard (1963)] to obtain exact tests based on standard LR and LM zero correlation tests. We also suggest a MC quasi-LR (QLR) test based on feasible generalized least squares (FGLS). We show that the latter statistics are pivotal under the null, which provides the justification for applying MC tests. Furthermore, we extend the exact independence test proposed by Harvey and Phillips (1982) to the multi-equation framework. Specifically, we introduce several induced tests based on a set of simultaneous Harvey/Phillips-type tests and suggest a simulation-based solution to the associated combination problem. The properties of the proposed tests are studied in a Monte Carlo experiment which shows that standard asymptotic tests exhibit important size distortions, while MC tests achieve complete size control and display good power. Moreover, MC-QLR tests performed best in terms of power, a result of interest from the point of view of simulation-based tests. The power of the MC induced tests improves appreciably in comparison to standard Bonferroni tests and in certain cases outperform the likelihood-based MC tests. The tests are applied to data used by Fischer (1993) to analyze the macroeconomic determinants of growth.

Key words: seemingly unrelated regressions; SURE system; multivariate linear regression; contemporaneous correlation; exact test; finite-sample test; Monte Carlo test; bootstrap; induced test; LM test; likelihood ratio test; specification test; macroeconomics; growth.

Journal of Economic Literature classification: C3; C12; C15; C30; C33; C52.

RÉSUMÉ

Cet article propose des procédures exactes pour tester la spécification SURE (régressions empilées) dans le contexte des régressions linéaires multivariées, *i.e.* si les perturbations des différentes équations sont corrélées ou non. Nous appliquons la technique des tests de Monte Carlo (MC) [Dwass (1957), Barnard (1963)] pour obtenir des tests d'indépendance exacts fondés sur les critères du quotient de vraisemblance (LR) et du multiplicateur de Lagrange (LM). Nous suggérons aussi un critère du type quasi-quotient de vraisemblance (QLR) dérivé sur base des moindres carrés généralisés réalisables (FGLS). Nous démontrons que ces statistiques sont libres de paramètres de nuisance sous l'hypothèse nulle, ce qui justifie l'application des tests de Monte Carlo. Par ailleurs, nous généralisons le test exact proposé par Harvey et Phillips (1982) au contexte des équations multiples. En particulier, nous proposons plusieurs tests induits basés sur des tests de type Harvey-Phillips et nous suggérons une technique basée sur des simulations afin de résoudre le problème de combinaison de tests. Nous évaluons les propriétés des tests que nous proposons dans le cadre d'une étude de Monte Carlo. Nos résultats montrent que les tests asymptotiques usuels présentent de sérieuses distorsions de niveau, alors que les tests de MC contrôlent parfaitement le niveau et ont une bonne puissance. De plus, les tests QLR se comportent bien du point de vue de la puissance; ce résultat est intéressant vu que les tests (multivariés) que nous proposons sont basés sur des simulations. La puissance des tests de MC induits augmente sensiblement par rapport aux tests fondés sur l'inégalité de Bonferroni tests et, dans certains cas, dépasse la puissance des tests de MC fondés sur la vraisemblance. Nous appliquons les tests sur des données utilisées par Fischer (1993) pour analyser des modèles de croissance.

Mots clés: régressions empilées; système SURE; test d'indépendance; régression linéaire multivariée; corrélation contemporaine; test exact; test à distance finie; test de Monte Carlo; bootstrap; test induit; test LM; quotient de vraisemblance; test de spécification; mavroéconomie; croissance.

Classification du Journal of Economic Literature classification: C3; C12; C15; C30; C33; C52.

Contents

List of Definitions, Propositions and Theorems

List of Tables

1. Introduction

Multi-equation models which use both cross-section and time series data are common in econometric studies. These include, in particular, the seemingly unrelated regressions (SURE) model introduced by Zellner (1962). The SURE specification is expressed as a set of linear regressions where the disturbances in the different equations are correlated. The non-diagonality of the error covariance matrix usually entails that individual equation estimates are sub-optimal; hence, generalized least squares (GLS) estimation which exploits the correlations across equations may improve inference. However, the implementation of GLS requires estimating the error covariance from the data. Further the cross-equation dependence must be taken into account when testing cross-equation parameter restrictions. As it is well known, feasible generalized least squares (FGLS) estimators need not be more efficient than ordinary least squares (OLS); see Srivastava and Giles (1987, Chapter 2). Indeed, the closer the error covariance comes to being spherical, the more likely it is that OLS estimates will be superior. This has extensively been discussed in the SURE literature; see, for example, Zellner (1962, 1963), Mehta and Swamy (1976), Kmenta and Gilbert (1968), Revankar (1974, 1976), Kunitomo (1977), Kariya (1981*a*), and Srivastava and Dwivedi (1979). In this sense, choosing between GLS and OLS estimation in the SURE model corresponds to the problem of testing for sphericity of the error covariance matrix.

This paper studies and proposes finite-sample tests for independence against contemporaneous correlation of disturbances in a SURE model. Independence tests in multivariate models have been discussed in both the econometric and statistical literatures. In particular, Breusch and Pagan (1980) derived a Lagrange multiplier (LM) test for the diagonality of the error covariance matrix. Kariya (1981*c*) derived locally best invariant tests in a two-equation framework. Shiba and Tsurumi (1988) proposed Wald, likelihood ratio (LR), LM and Bayesian tests for the hypothesis that the error covariance is block-diagonal. Related results are also available in Kariya (1981*b*), Kariya, Fujikoshi and Krishnaiah (1984) and Cameron and Trivedi (1993). Except for one special case, these test procedures are only justified by asymptotic arguments. The exception is Harvey and Phillips (1982, Section 3) who proposed exact independence tests between the errors of an equation and those of the other equations of the system. These tests (which we will denote EFT) involve conventional F statistics for testing whether the (estimated) residuals added to each equation have zero coefficients. EFT tests may be applied in the context of general diagonality tests; for example, one may assess in turn whether the disturbances in each equation are independent of the disturbances in all other equations. Such a sequence of tests however raises the problem of taking into account the dependence between multiple tests, a problem not solved by Harvey and Phillips (1982).

A major problem in the SURE context comes from the fact that relevant null distributions are either difficult to derive or too complicated for practical use. This is true even in the case of identical regressor matrices. Hence the applicable procedures rely heavily on asymptotic approximations whose accuracy can be quite poor. This is evident from the Monte Carlo results reported in Harvey and Phillips (1982) and Shiba and Tsurumi (1988), among others. In any case, it is widely acknowledged by now that standard multivariate LR-based asymptotic tests are unreliable in finite samples, in the sense that test sizes deviate from the nominal significance levels; see Dufour and Khalaf (1998) for related simulation evidence.

In this paper, we reemphasize this fact and propose to use the technique of Monte Carlo (MC) tests [Dwass (1957), Barnard (1963)] in order to obtain provably exact procedures. We apply the MC test technique to: (i) the standard likelihood ratio (LR) and Lagrange multiplier (LM) criteria, and (ii) OLS and FGLS-based quasi-LR (QLR) statistics. We also introduce several induced tests based on a set of simultaneous Harvey/Phillips-type tests and suggest a simulation-based solution to the associated combination problem. The critical regions of conventional induced tests are usually computed using probability inequalities (*e.g.,* the well know Boole-Bonferroni inequality) which yields conservative p-values whenever non independent tests are combined [see, for example, Savin (1984), Folks (1984), Dufour (1990) and Dufour and Torrès (1998)]. Here, we propose to construct the induced tests such that size-correct p-values can be readily obtained by simulation.

The first step towards an exact test procedure involves deriving nuisance-parameter-free null distributions. In the context of standard independence tests, invariance results are known given two univariate or multivariate regression equations [Kariya (1981*b*, 1981*c*), Kariya et al. (1984)]. The problem of nuisance parameters is yet unresolved in models involving more than two regression equations. Here, we show that the LR, LM and QLR independence test statistics are pivotal under the null, for multi-equation SURE systems. Though the proof of this result is not complex, it does not appear to be known in the literature. Of course, existing work in this area has typically focused on deriving p-values analytically. By contrast, the approach taken in this article does not require extracting exact distributions; the technique of MC tests allows one to obtain provably exact randomized tests in finite samples using very small numbers of MC replications of the original test statistic under the null hypothesis. In the present context, this technique can easily be applied whenever the distribution of the errors is continuous and specified up to an unknown covariance matrix (or linear transformation). Note this distribution does not have to be Gaussian. For further references regarding MC tests, see Dufour (1995), Dufour and Kiviet (1996, 1998), Kiviet and Dufour (1997), Dufour, Farhat, Gardiol and Khalaf (1998), and Dufour and Khalaf (2001). We investigate the size and power of suggested tests in a Monte Carlo study. The results show that, while the asymptotic LR and LM tests seriously overreject, the MC versions of these tests achieve perfect size control and have good power. The power of the MC induced tests improves appreciably in comparison to the standard Bonferroni tests and in several cases outperform the corresponding MC-LR and LM tests.

The outline of this study is as follows. In Section 2, we present the model and the estimators used, while the test statistics are described in Section 3. In Section 4, we show that the proposed test statistics have nuisance-parameter free distributions under the null hypothesis and describe how exact MC tests can be implemented. In Section 5, we report the simulation results. In Section 6, we apply the tests to data used by Fischer (1993) to analyze the macroeconomic determinants of growth. We conclude in Section 7.

2. Framework

Consider the seemingly unrelated regression model

$$Y_i = X_i\beta_i + u_i\,, \quad i = 1,\ \ldots\ ,\ p\,, \tag{2.1}$$

where Y_i is a vector of n observations on a dependent variable, X_i a full-column rank $n \times k_i$ matrix of regressors, β_i a vector of k_i unknown coefficients, and $u_i = (u_{1i},\, u_{2i},\, \ldots\, ,\, u_{ni})'$ a $n \times 1$ vector of random disturbances. When $X_i = X_j,\ i,\ j = 1,\ \ldots\ ,\ p$, we have a *multivariate linear regression* (MLR) model; see Anderson (1984, chapters 8 and 13), Berndt and Savin (1977), and Kariya (1985). The system (2.1) may be rewritten in the stacked form

$$y = X\beta + u \tag{2.2}$$

where

$$y = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_p \end{bmatrix},\ X = \begin{bmatrix} X_1 & 0 & \cdots & 0 \\ 0 & X_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & X_p \end{bmatrix},\ u = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_p \end{bmatrix},\ \beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{bmatrix}, \tag{2.3}$$

so that X is a $(np) \times k$ matrix, y and u each have dimension $(np) \times 1$ and β has dimension $k \times 1$, with $k = \sum_{i=1}^{p} k_i$. Set

$$U = \begin{bmatrix} u_1 & u_2 & \cdots & u_p \end{bmatrix} = \begin{bmatrix} U'_{1\cdot} \\ U'_{2\cdot} \\ \vdots \\ U'_{n\cdot} \end{bmatrix} \tag{2.4}$$

where $U_{t\cdot} = (u_{t1},\, u_{t2},\, \ldots\, ,\, u_{tp})'$ is the disturbance vector for the t-th observation. In the sequel, we shall also use, when required, some or all of the following assumptions and notations:

$$U_{t\cdot} = JW_t\ , \quad t = 1,\, \ldots\, ,\, n\, , \tag{2.5}$$

where J is a fixed lower triangular $p \times p$ matrix such that

$$\Sigma \equiv JJ' = \left[\sigma_{ij}\right]_{i,j=1,\, \ldots\, ,p} \text{ is nonsingular,} \tag{2.6}$$

where we set $\sigma_i \equiv \sigma_{ii}^{1/2}$, $i = 1,\, \ldots\, ,\, p$;

$$\begin{gathered} W_1,\, \ldots\, ,\, W_n \text{ are } p \times 1 \text{ random vectors} \\ \text{whose joint distribution is completely specified;} \end{gathered} \tag{2.7}$$

$$u \text{ is independent of } X\, . \tag{2.8}$$

Assumption (2.8) is a strict exogeneity assumption, which clearly holds when X is fixed. The assumptions (2.5) - (2.7) mean that the disturbance distribution is completely specified up an unknown linear transformation that can modify the scaling and dependence properties of the disturbances in different equations. Note (2.5) - (2.7) do not necessarily entail that Σ is the covariance matrix of

$U_{t\cdot}$. However, if we make the additional assumption that

$$\begin{gathered} W_1, \ldots , W_n \text{ are uncorrelated with} \\ E(W_t) = 0 \,, \quad E(W_t W_t') = I_p \,, \quad t = 1, \ldots , n \,, \end{gathered} \tag{2.9}$$

or, more restrictively,

$$W_1, \ldots , W_n \overset{i.i.d.}{\sim} N[0,\, I_p] \,, \tag{2.10}$$

we have:

$$E\big(U_{t\cdot}\big) = 0 \,, \quad E\big(U_{t\cdot} U_{t\cdot}'\big) = \Sigma \,, \quad t = 1, \ldots , n \,, \tag{2.11}$$

$$E(u) = 0 \,,\; E(u_i u_j') = \sigma_{ij} I_n \,, \quad i \,,\, j = 1 \,, \ldots \,, p \,, \tag{2.12}$$

and

$$E(uu') = \Sigma \otimes I_p \,. \tag{2.13}$$

The coefficients of the regression equations can be estimated by several methods among which the most well known are: (i) ordinary least squares (OLS) applied to each equation, (ii) two-step feasible generalized least squares (FGLS), (iii) iterative FGLS (IFGLS), and (iv) maximum likelihood (ML) assuming u follows a multinormal distribution. The OLS estimator of β is

$$\hat{\beta}_{OLS} = (\hat{\beta}_1' \,, \ldots \,, \hat{\beta}_p')' , \quad \hat{\beta}_i = (X_i' X_i)^{-1} X_i' Y_i \,, \quad i = 1 \,, \ldots \,, p \,. \tag{2.14}$$

An associated estimate $\hat{\Sigma}$ for Σ can be obtained from the OLS residuals:

$$\hat{u}_i = Y_i - X_i \hat{\beta}_i = M(X_i) u_i \,, \quad M(X_i) = I_n - X_i (X_i' X_i)^{-1} X_i' \,, \quad i = 1 \,, \ldots \,, p \,. \tag{2.15}$$

The two-step FGLS estimate based on any consistent estimate S of Σ, is given by

$$\tilde{\beta}_{FGLS} = \left[X'(S^{-1} \otimes I_n) X \right]^{-1} X'(S^{-1} \otimes I_n) y \,. \tag{2.16}$$

If the disturbances are normally distributed, we have the log-likelihood function

$$\mathcal{L} = -\frac{np}{2} \ln(2\pi) - \frac{n}{2} \ln(|\Sigma|) - \frac{1}{2} (y - X\beta)'(\Sigma^{-1} \otimes I_n)(y - X\beta) \,. \tag{2.17}$$

The corresponding maximum likelihood (ML) estimators $\tilde{\beta}$ and $\tilde{\Sigma}$ of β and Σ satisfy the following normal equations:

$$X'(\tilde{\Sigma}^{-1} \otimes I_n) X \tilde{\beta} = X'(\tilde{\Sigma}^{-1} \otimes I_n) y \,, \quad \tilde{\Sigma} = \frac{1}{n} \tilde{U}' \tilde{U} = \left[\tilde{\sigma}_{ij} \right]_{i,j=1,\ldots,p} \tag{2.18}$$

where $\tilde{\beta} = (\tilde{\beta}_1' \,, \ldots \,, \tilde{\beta}_p')'$ and

$$\tilde{U} = [\tilde{u}_1 \,, \ldots \,, \tilde{u}_p] \,,\; \tilde{u}_i = Y_i - X_i \tilde{\beta}_i \,,\; \tilde{\sigma}_{ij} = \tilde{u}_i' \tilde{u}_j / n \,. \tag{2.19}$$

Of course, the estimators in (2.18) are well defined provided the matrix $\tilde{\Sigma}$ has full column rank, an assumption we shall make in the sequel.

Iterative procedures are typically applied to obtain the ML estimates. Suppose $\tilde{\Sigma}^{(0)}$ is an initial estimate of Σ. Using (2.18), we can solve for a first GLS estimate of β,

$$\tilde{\beta}^{(0)} = \left[X'(\tilde{\Sigma}^{(0)} \otimes I_n)^{-1} X\right]^{-1} X'(\tilde{\Sigma}^{(0)} \otimes I_n)^{-1} y \, , \tag{2.20}$$

from which a new estimate of u may be obtained:

$$\tilde{u}^{(1)} = y - X\tilde{\beta}^{(0)} \, . \tag{2.21}$$

This residual leads to further estimators $\tilde{\Sigma}^{(1)}$ and $\tilde{\beta}^{(1)}$ of Σ and β. Pursuing this iterative process, we see that the estimators at the h-th iteration take the form:

$$\tilde{\beta}^{(h)} = \left[X'(\tilde{\Sigma}^{(h)} \otimes I_n)^{-1} X\right]^{-1} X'(\tilde{\Sigma}^{(h)} \otimes I_n)^{-1} y \, , \tag{2.22}$$

$$\tilde{\Sigma}^{(h)} = \frac{1}{n}\tilde{U}^{(h)\prime}\tilde{U}^{(h)} = \left[\tilde{\sigma}_{ij}^{(h)}\right]_{i,j=1,\ldots,p} \, , \tag{2.23}$$

$h = 1, 2, \ldots$, where

$$\tilde{U}^{(h)} = [\tilde{u}_1^{(h)} \, , \, \ldots \, , \, \tilde{u}_p^{(h)}] \, , \; \tilde{u}_i^{(h)} = Y_i - X_i\tilde{\beta}_i^{(h-1)} \, , \; \tilde{\sigma}_{ij}^{(h)} = \tilde{u}_i^{(h)\prime}\tilde{u}_j^{(h)}/n \, . \tag{2.24}$$

Under standard assumptions, iterating this procedure to convergence yields the ML estimates [see Oberhofer and Kmenta (1974)]. For a more general discussion of the properties of such partially iterated estimators, the reader may consult Robinson (1988).

3. Test statistics for cross-equation disturbance correlation

3.1. Likelihood-based tests

Given the setup described above, we consider the problem of testing the hypothesis H_0 that Σ is diagonal. For any vector $d = (d_1 \, , \, \ldots \, , \, d_N)'$, let us denote $D_N(d_i)$ the diagonal matrix whose diagonal elements are $d_1 \, , \, \ldots \, , \, d_N$:

$$D_N(d_i) = diag(d_1 \, , \, \ldots \, , \, d_N) \, . \tag{3.1}$$

Then H_0 may be expressed as

$$H_0 : \Sigma = D_p(\sigma_i^2) \, . \tag{3.2}$$

Since J is lower triangular, it is easy to see that: $\Sigma = D_p(\sigma_i^2)$ if and only $J = D_p(\sigma_i)$. Thus, under H_0, $u_{ti} = \sigma_i W_{ti}$, $i = 1 \, , \, \ldots \, , \, p$, where $W_t = (W_{t1}, W_{t2}, \ldots \, , \, W_{tp})'$. If (2.9) holds, H_0 is equivalent to the absence of contemporaneous correlation between the components of $U_{t\cdot}$. If the components of W_t are independent, H_0 is equivalent to the independence between the components of u_t; when $W_1, \ldots \, , \, W_n$ are independent, the latter condition entails that the disturbance vectors $u_1, \ldots \, , \, u_p$ are independent.

In the sequel, we will frequently refer to the standardized disturbances

$$w = \left(w_1', \ldots, w_p'\right)', \quad \text{where } w_i = (1/\sigma_i)u_i, \; i = 1, \ldots, p. \tag{3.3}$$

Under the assumptions (2.5) - (2.7), the vector w has a completely specified distribution if H_0 holds.

Let us now consider the case where, in addition to (2.5) - (2.7), we make the normality assumption (2.10). Then the disturbance vectors $U_{t\cdot} = JW_t$, $t = 1, \ldots, n$, are $i.i.d.$ $N[0, \Sigma]$ where $\Sigma = JJ'$ and we have the log-likelihood function (2.17). In this case, the LR and LM statistics for testing H_0 take relatively simple forms. The LR statistic is $\xi_{LR} = n \ln(\tilde{\Lambda})$ where

$$\tilde{\Lambda} = |D_p(\hat{\sigma}_i^2)| / |\tilde{\Sigma}|, \tag{3.4}$$

while the LM criterion is

$$\xi_{LM} = n \sum_{i=2}^{p} \sum_{j=1}^{i-1} r_{ij}^2 \tag{3.5}$$

where $r_{ij} = \hat{u}_i' \hat{u}_j / [(\hat{u}_i' \hat{u}_i)(\hat{u}_j' \hat{u}_j)]^{1/2}$. Under standard regularity conditions, both ξ_{LR} and ξ_{LM} follow a $\chi^2\big(p(p-1)/2\big)$ distribution asymptotically under H_0 [see Breusch and Pagan (1980)].

In the sequel, we shall also consider quasi-LR statistics $\xi_{LR}^{(h)} = n \ln(\tilde{\Lambda}^{(h)})$ where $\tilde{\Sigma}^{(h)}$ is used instead of the unrestricted ML estimator $\tilde{\Sigma}$:

$$\tilde{\Lambda}^{(h)} = |D_p(\hat{\sigma}_i^2)| / |\tilde{\Sigma}^{(h)}|. \tag{3.6}$$

Since unrestricted ML estimators of the SURE model parameters are usually obtained through iterative numerical methods, such QLR statistics are easier to compute than the fully-iterated LR statistic.

3.2. Induced Harvey-Phillips tests

A finite-sample exact independence test was developed by Harvey and Phillips (1980). Their procedure is applicable under the assumptions (2.5) - (2.10) to test a null hypothesis of the form

$$H_{01} : \Sigma = \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \Sigma_{11} \end{bmatrix} \tag{3.7}$$

where Σ_{11} is a $(p-1) \times (p-1)$ matrix. Specifically, they propose the following statistic:

$$EFT = \frac{\hat{u}_1' \hat{V}_1 \left(\hat{V}_1' M_1 \hat{V}_1\right)^{-1} \hat{V}_1' \hat{u}_1 / (p-1)}{\hat{u}_1' \left[I - \hat{V}_1 \left(\hat{V}_1' M_1 \hat{V}_1\right)^{-1} \hat{V}_1'\right] \hat{u}_1 / (n - k_1 - p + 1)}, \tag{3.8}$$

where $\hat{V}_1 = [\hat{u}_2, \ldots, \hat{u}_p]$, which follows an F distribution with $(p-1,\ n-k_1-p+1)$ degrees of freedom under H_{01}. The EFT statistic can be obtained as the usual F-statistic for testing whether the coefficients on $\hat{V}_1$ are zero in the regression of Y_1 on X_1 and $\hat{V}_1$.

More generally, we can consider any particular disturbance vector u_i (or equation) from the p

regressions in (2.1) and test in this way whether u_i is independent of $V_{K_{(i)}} \equiv \left[u_j\right]_{j\in K_{(i)}}$, where $K_{(i)}$ is some non-empty subset of $\{j : 1 \leq j \leq p, j \neq i\}$. This can be done by estimating an extended regression of the form

$$Y_i = X_i\beta_i + \sum_{j\in K_{(i)}} \widehat{u}_j\gamma_{ij} + u_i \tag{3.9}$$

and testing the hypothesis $H_0[K_{(i)}] : \gamma_{ij} = 0$ for $j \in K_{(i)}$. Under the null hypothesis H_0 of independence [see (3.2)], the corresponding F-statistic

$$F_i[K_{(i)}] = \frac{\left(\widehat{u}_i'\widehat{u}_i - SS(K_{(i)})\right)/\overline{p}_i}{SS(K_{(i)})/(n - k_i - \overline{p}_i)} \tag{3.10}$$

follows an $F(\overline{p}_i\,,\, n - k_i - \overline{p}_i)$, where $\overline{p}_i$ is the number of elements in $K_{(i)}$ and $SS(K_{(i)})$ is the unrestricted residual sum of squares from regression (3.9).

As things stand, the latter procedures only test the independence of one disturbance vector u_i with respect to the other disturbance vectors. It is straightforward to see that the test of H_0 based on $F_i[K_{(i)}]$ can only detect correlations between u_i and the other disturbances. In order to test H_0 against all possible covariance matrices Σ, we need a different procedure. A simple way to do this, which still exploits the Harvey-Phillips procedure, consists in using induced tests that combine several tests of the form $F_i[K_{(i)}]$. Here we shall consider two methods for combining tests.

Denote $G_F[x|\nu_1, \nu_2]$ the survival function of the Fisher distribution with (ν_1, ν_2) degrees of freedom; *i.e.*, if F is a random variable that follows an $F(\nu_1, \nu_2)$ distribution, we have $G_F[x|\nu_1, \nu_2] = P[F \geq x]$. We consider the test statistics

$$EFT_i \equiv F_i[K_i]\ ,\ \text{where } K_i \equiv \{j : 1 \leq j \leq n\,,\, j \neq i\}\,,\ i = 1, \ldots, p\,, \tag{3.11}$$

each of which tests whether u_i is independent of all the other disturbance vectors. The p-value associated with EFT_i is:

$$pv_i[K_i] = G_F[EFT_i\,|\,p - 1\,,\, n - k_i - p + 1] \tag{3.12}$$

which follows a uniform distribution on the interval $[0, 1]$. The level-α F-test based on EFT_i is equivalent (with probability 1) to rejecting the null hypothesis when $pv_i[K_i] \leq \alpha$, or equivalently when $1 - pv_i[K_i] \geq 1 - \alpha$.

A difficulty we meet here consists in controlling the overall level of a procedure based on several separate tests. A simple way to do this consists in running each one of the p tests $F_i[K_i]$ at level α_i, so that $\sum_{i=1}^{p}\alpha_i = \alpha$, and rejecting H_0 when at least one of the p separate tests rejects the null hypothesis; for example, we may take $\alpha_i = \alpha/p$, $i = 1, \ldots, p$. By the Boole-Bonferroni inequality, this ensures that the probability of rejecting H_0 is not greater than α (although it could be smaller). When $\alpha_i = \alpha/p$, this procedure is equivalent to rejecting H_0 when $pv_{\min} \leq \alpha/p$, where

$$pv_{\min} \equiv \min\{pv_i[K_i] : i = 1, \ldots, p\} \tag{3.13}$$

is the minimum of the p-values.

Note that using the minimum of several p-values as a test statistic was originally proposed by Tippett (1931) and Wilkinson (1951), in the case of independent test statistics. The independence condition does not however hold here for the EFT_i statistics, hence the necessity of taking into account the dependence. Because it is conservative, the Boole-Bonferroni bound may lead to a power loss with respect to a procedure that avoids the use of a bound. In the next section, we will see that the conservative property of the Bonferroni-based $pv_{\min}$ procedure can be corrected by using the technique of Monte Carlo tests. In other words, we consider the procedure that rejects H_0 when $pv_{\min}$, as defined by (3.12) and (3.13), is small, and we shall show that its size can be controlled by using the Monte Carlo test technique.

A second fairly natural way of "aggregating" separate tests consists in rejecting H_0 when the product

$$pv_{\times} = \prod_{i=1}^{p} pv_i[K_i] \tag{3.14}$$

is small. Such a procedure was originally suggested by Fisher (1932) and Pearson (1933), again for independent test statistics. As for the $pv_{\min}$ procedure, we will see that the size of such a test based on $pv_{\times}$ can be controlled by Monte Carlo techniques, even if the individual p-values $pv_i[K_i]$ are not independent.

For convenience reasons, we shall implement both these tests by taking the test criteria:

$$F_{\min} = 1 - pv_{\min} \;, \tag{3.15}$$

$$F_{\times} = 1 - \; pv_{\times} \;, \tag{3.16}$$

each one of which rejects H_0 when it is large.

We also considered a "sequential" approach in which we test the sequence of hypotheses

$$H_{0i} : u_i \text{ is independent of } u_{i+1}, \, \dots \, , u_p \tag{3.17}$$

for $i = 1, \, ..., \, p-1,$ using Harvey-Phillips tests based on regressions of the form

$$Y_i = X_i \beta_i + \sum_{j=i+1}^{p} \widehat{u}_j \gamma_{ij} + u_i \;, \tag{3.18}$$

$i = 1 \, , \, \dots \, , \, p-1$. Clearly H_0 is equivalent to the conjunction of the $p-1$ hypotheses $H_{0i}\, , i = 1 \, , \, ..., \, p-1,$ so that we should reject H_0 when at least one of these tests is significant. This yields the $p-1$ test statistics $F_i[\{i+1, \, \dots \, , \, p\}]$, $i = 1 \, , \, ..., \, p-1$ for which it is easy to see that $F_i[\{i+1, \, \dots \, , \, p\}] \sim F(p-i\, , \, n-k_i-p+i)$ under H_0. The problem then consists again in controlling the overall level of this combined procedure. Since it is not clear the test statistics are independent, one way to achieve this control consists in using again the Boole-Bonferroni inequality. For this, we test H_{0i} at level α_i, where $\sum_{i=1}^{p} \alpha_i = \alpha$, and reject H_0 when one of the tests is significant. In a sequential context, a standard way of doing this consists in considering geometrically declining levels, such as

$$\alpha_1 = \alpha/2\,,\ \alpha_2 = \alpha/(2^2)\,,\ \dots\,,\ \alpha_{p-2} = \alpha/(2^{p-2})\,,\ \alpha_{p-1} = \alpha/(2^{p-2})\,; \tag{3.19}$$

see Anderson (1971, Chapter 4) and Lehmann (1957). Here we shall consider the bound procedure based on (3.19), as well as tests on the minimum and the product of the p separate p-values associated with the test statistics $F_i[\{i+1, \dots, p\}]$:

$$FS_{\min} = 1 - \min\{pv_i[\{i+1, \dots, p\}] : i = 1, \dots, p-1\}\ , \tag{3.20}$$

$$FS_{\times} = 1 - \prod_{i=1}^{p-1} pv_i[\{i+1, \dots, p\}]\ . \tag{3.21}$$

Again the levels of the two latter procedures will be controlled through the Monte Carlo test technique.

For further discussion of multiple test procedures, the reader may consult Miller (1981), Folks (1984), Savin (1984), Dufour (1989, 1990), Westfall and Young (1993), Gouriéroux and Monfort (1995, Chapter 19), and Dufour and Torrès (1998, 2000).

4. Finite-sample theory

We proceed next to examine the finite-sample distributions of the above defined LM, LR and QLR test criteria. In particular, we show that the associated null distributions are free of nuisance parameters. To do this, we will first demonstrate in the three following propositions that all the statistics considered are functions of the standardized disturbances w_i , $i = 1, \dots, p$. Interestingly, these properties hold under very weak distributional assumptions on u and X.

Proposition 4.1 STANDARDIZED REPRESENTATION OF LM AND HARVEY-PHILLIPS STATISTICS. *Under the assumptions and notations* (2.1) *to* (2.6), *the* LM *statistic defined in* (3.5) *can be written in the form*

$$\xi_{LM} = n \sum_{i=2}^{p} \sum_{j=1}^{i-1} \overline{r}_{ij}^2 \tag{4.1}$$

where $\overline{r}_{ij} = \widehat{w}_i'\widehat{w}_j/[(\widehat{w}_i'\widehat{w}_i)(\widehat{w}_j'\widehat{w}_j)]^{1/2}$, $\widehat{w}_i = \widehat{u}_i/\sigma_i = M(X_i)w_i$ *and* $w_i = (1/\sigma_i)u_i$, *while each statistic* $F_i[K_{(i)}]$ *defined in* (3.10) *is identical to the* F*-statistic* $\overline{F}_i[K_{(i)}]$ *for testing* $H_0^* : \gamma_{ij}^* = 0$ *for* $j \in K_{(i)}$ *in the regression*

$$Y_i^* = X_i\beta_i^* + \sum_{j \in K_{(i)}} \widehat{w}_j \gamma_{ij}^* + w_i \tag{4.2}$$

where $Y_i^* = (1/\sigma_i)y_i$.

PROOF. The result for the LM statistic follows on observing that

$$r_{ij} = \frac{\hat{u}_i'\hat{u}_j}{[(\hat{u}_i'\hat{u}_i)(\hat{u}_j'\hat{u}_j)]^{1/2}} = \frac{\widehat{w}_i'\widehat{w}_j}{[(\hat{w}_i'\hat{w}_i)\,(\hat{w}_j'\hat{w}_j)]^{1/2}} = \overline{r}_{ij}\,.$$

For $F_i[K_{(i)}]$, we note that

$$\hat{u}_i'\hat{u}_i = u_i'M(X_i)u_i = \sigma_i^2 w_i'M(X_i)w_i = \sigma_i^2\hat{w}_i'\hat{w}_i\ ,\ SS[K_{(i)}] = \sigma_i^2 SS_i^*$$

where $\widehat{w}_i'\widehat{w}_i$ and SS_i^* are the restricted and unrestricted residual sum of squares from the linear regression

$$Y_i^* = X_i\beta_i^* + \sum_{j\in K_{(i)}}\widehat{w}_j\gamma_{ij}^* + w_i\,.$$

We then see that

$$\begin{aligned} F_i[K_{(i)}] &= \frac{\big(\widehat{u}_i'\widehat{u}_i - SS[K_{(i)}]\big)/\overline{p}_i}{SS[K_{(i)}]/(n-k_i-\overline{p}_i)} = \frac{\big(\sigma_i^2\hat{w}_i'\hat{w}_i - \sigma_i^2 SS_i^*\big)/\overline{p}_i}{\sigma_i^2 SS_i^*/(n-k_i-\overline{p}_i)} \\ &= \frac{\big(\hat{w}_i'\hat{w}_i - SS_i^*\big)/\overline{p}_i}{SS_i^*/(n-k_i-\overline{p}_i)} = \overline{F}_i[K_{(i)}]\,. \end{aligned}$$

□

Proposition 4.2 STANDARDIZED REPRESENTATION OF THE LR STATISTIC. *Under the assumptions and notations of Proposition* **4.1**, *suppose the matrix* $\tilde{\Sigma}$ *defined in* (2.18) *has full column rank. Then the LR-based statistic* $\tilde{\Lambda}$ *defined in* (3.4) *can be written in the form*

$$\tilde{\Lambda} = \frac{\prod_{i=1}^{p}[w_i'M(X_i)w_i/n]}{|\widetilde{\Sigma}_*|} \tag{4.3}$$

where $\widetilde{\Sigma}_*$ *is the ML estimator of* Σ *obtained by maximizing the Gaussian log-likelihood*

$$\mathcal{L}_* = -\frac{np}{2}\ln(2\pi) - \frac{n}{2}\ln(|\Sigma|) - \frac{1}{2}(w-X\beta)'(\Sigma^{-1}\otimes I_n)(w-X\beta) \tag{4.4}$$

where $w = (w_1', w_2', \ldots, w_p')'$.

PROOF. From (3.4) we can write

$$\tilde{\Lambda} = \frac{|D_p(\sigma_i^{-1})|\,|D_p(\hat{\sigma}_i^2)|\,|D_p(\sigma_i^{-1})|}{|D_p(\sigma_i^{-1})|\,|\tilde{\Sigma}|\,|D_p(\sigma_i^{-1})|} = \frac{\prod_{i=1}^{p}\hat{\sigma}_i^2/\sigma_i^2}{|D_p(\sigma_i^{-1})\tilde{\Sigma}D_p(\sigma_i^{-1})|} \tag{4.5}$$

where $\hat{\sigma}_i^2/\sigma_i^2 = \widehat{w}_i'\widehat{w}_i/n = w_i'M(X_i)w_i/n$. Further, it is easy to see that the Gaussian log-likelihood (2.17) is invariant under data transformations of the form $y_* = vec[\ Y_{1*}\ \ Y_{2*}\ \ \cdots\ \ Y_{p*}\ ]$ with

$$Y_{i*} = c_i(Y_i + X_i\delta_i)\,,\ i = 1,\,\ldots\,,\,p\,, \tag{4.6}$$

where c_i is an arbitrary non-zero constant and δ_i an arbitrary $k_i \times 1$ vector ($i = 1, \ldots, p$). In other words, if the log-likelihood function of y is given by (2.17), the likelihood of y_* has the same form with β_i replaced by $\beta_{i*} = c_i(\beta_i + \delta_i)$ and Σ replaced by $\Sigma_* = D_p(c_i)\Sigma D_p(c_i)$. In particular, if we take $\delta_i = -\beta_i$ and $c_i = 1/\sigma_i$, we get $Y_{i*} = (1/\sigma_i)u_i = w_i$ with $\mathcal{L}_*$ as the corresponding log-likelihood function. Consequently, by the equivariance of maximum likelihood estimators [see Dagenais and Dufour (1991)], we have $\widetilde{\Sigma}_* = D_p(\sigma_i^{-1})\tilde{\Sigma}D_p(\sigma_i^{-1})$, from which (4.3) follows. □

Proposition 4.3 STANDARDIZED REPRESENTATION OF QLR STATISTICS. *Under the assumptions and notations of Proposition* **4.1**, *let $\tilde{\Sigma}^{(0)}$ be an initial positive definite estimator of Σ, and suppose the matrices $\widetilde{\Sigma}^{(h)}$, $h = 1, \ldots, H$, defined in* (2.23) *have full column rank. Then, the approximate LR statistics $\tilde{\Lambda}^{(H)}$ defined by* (3.6) *can be written in the form*

$$\tilde{\Lambda}^{(H)} = \frac{\prod_{i=1}^{p}[w_i'M(X_i)w_i/n]}{|\tilde{\Sigma}_*^{(H)}|} \tag{4.7}$$

where $\tilde{\Sigma}_*^{(H)}$ is the estimate of Σ obtained through the formulas:

$$\tilde{\beta}_*^{(h)} = \left[X'(\tilde{\Sigma}_*^{(h)} \otimes I_n)^{-1}X\right]^{-1}X'(\tilde{\Sigma}_*^{(h)} \otimes I_n)^{-1}w\ , \tag{4.8}$$

$$\begin{aligned}\tilde{\Sigma}_*^{(h)} &= D_p(\sigma_i^{-1})\tilde{\Sigma}^{(0)}D_p(\sigma_i^{-1}) \quad &, \textit{for } h = 0\,,\\ &= \tfrac{1}{n}\tilde{U}_*^{(h)\prime}\tilde{U}_*^{(h)} &, \textit{for } h \geq 1\,,\end{aligned} \tag{4.9}$$

$h = 0, 1, \ldots, H$, *where* $\tilde{U}_*^{(h)},\ h \geq 1$, *obeys the recursion*

$$\tilde{U}_*^{(h)} = [\tilde{u}_{1*}^{(h)}\ ,\ \ldots\ ,\ \tilde{u}_{p*}^{(h)}]\,,\ \tilde{u}_{i*}^{(h)} = w_i - X_i\tilde{\beta}_{i*}^{(h-1)}\ ,\ i = 1,\,\ldots\,,\,p\,. \tag{4.10}$$

PROOF. From the definition (3.6), we can write, for $h \geq 0$,

$$\tilde{\Lambda}^{(h)} = \frac{|D_p(\sigma_i^{-1})|\,|D_p(\hat{\sigma}_i^2)|\,|D_p(\sigma_i^{-1})|}{|D_p(\sigma_i^{-1})|\,|\widetilde{\Sigma}^{(h)}|\,|D_p(\sigma_i^{-1})|} = \frac{|D_p(\hat{\sigma}_i^2/\sigma_i^2)|}{|\tilde{\Sigma}_*^{(h)}|} = \frac{\prod_{i=1}^{p}[w_i'M(X_i)w_i/n]}{|\tilde{\Sigma}_*^{(h)}|} \tag{4.11}$$

where

$$\tilde{\Sigma}_*^{(h)} \equiv D_p(\sigma_i^{-1})\widetilde{\Sigma}^{(h)}D_p(\sigma_i^{-1})\,. \tag{4.12}$$

For $h = 0$, the result holds trivially. For $h \geq 1$, we have:

$$\tilde{\Sigma}_*^{(h)} = \frac{1}{n}\tilde{U}_*^{(h)\prime}\tilde{U}_*^{(h)} = \big[\tilde{u}_{i*}^{(h)\prime}\tilde{u}_{j*}^{(h)}/n\big]_{i,j=1,\,\dots\,,p}\,, \tag{4.13}$$

$$\tilde{U}_*^{(h)} = \tilde{U}^{(h)}D_p(\sigma_i^{-1}) = [\tilde{u}_1^{(h)}\,,\,\dots\,,\,\tilde{u}_p^{(h)}]D_p(\sigma_i^{-1}) = [\tilde{u}_{1*}^{(h)}\,,\,\dots\,,\,\tilde{u}_{p*}^{(h)}]\,, \tag{4.14}$$

$$\tilde{u}_{i*}^{(h)} \equiv (1/\sigma_i)\tilde{u}_i^{(h)} = (1/\sigma_i)[Y_i - X_i\tilde{\beta}_i^{(h-1)}]\,, \quad i = 1,\,\dots\,,\,p\,. \tag{4.15}$$

Putting (4.15) in vector form, we see that

$$\tilde{u}_*^{(h)} \equiv vec[\tilde{u}_{1*}^{(h)}\,,\,\dots\,,\,\tilde{u}_{p*}^{(h)}] = (D \otimes I_n)\tilde{u}^{(h)} = \overline{D}_n\big(y - X\tilde{\beta}^{(h-1)}\big) \tag{4.16}$$

where $D \equiv D_p(\sigma_i^{-1})$ and $\overline{D}_n \equiv D \otimes I_n$.

Now, for $h \geq 0$, the feasible GLS estimator $\tilde{\beta}^{(h)}$ minimizes the quadratic form

$$\widetilde{S}(\beta) = (y - X\beta)'\big(\widetilde{\Sigma}^{(h)} \otimes I_n\big)^{-1}(y - X\beta)$$

with respect to β. Since

$$\begin{aligned}\widetilde{S}(\beta) &= (y - X\beta)'(D \otimes I_n)(D^{-1} \otimes I_n)\big(\widetilde{\Sigma}^{(h)} \otimes I_n\big)^{-1}(D^{-1} \otimes I_n)(D \otimes I_n)(y - X\beta)\\ &= [(D \otimes I_n)(y - X\beta)]'\big[(D\widetilde{\Sigma}^{(h)}D) \otimes I_n\big]^{-1}[(D \otimes I_n)(y - X\beta)]\,,\end{aligned}$$

this entails that

$$\tilde{\beta}^{(h)} = \big[\big(\overline{D}_nX\big)'\big(\tilde{\Sigma}_*^{(h)} \otimes I_n\big)^{-1}\big(\overline{D}_nX\big)\big]^{-1}\big(\overline{D}_nX\big)'\big(\tilde{\Sigma}_*^{(h)} \otimes I_n\big)^{-1}\overline{D}_ny\,.$$

Further, on noting that

$$w \equiv (w_1',\, w_2',\,\dots\,,\, w_p')' = (D \otimes I_n)u = \overline{D}_nu$$

and

$$\overline{D}_nX = [D_p(\sigma_i^{-1}) \otimes I_n]X = \begin{bmatrix} \sigma_1^{-1}X_1 & 0 & \cdots & 0 \\ 0 & \sigma_2^{-1}X_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_p^{-1}X_p \end{bmatrix} = XD_p(\sigma_i^{-1}I_{k_i}) = X\Delta$$

where $\Delta \equiv D_p(\sigma_i^{-1}I_{k_i})$ is a non-singular matrix, we see that

$$\begin{aligned}\tilde{\beta}^{(h)} &= \big[\big(X\Delta\big)'\big(\tilde{\Sigma}_*^{(h)} \otimes I_n\big)^{-1}\big(X\Delta\big)\big]^{-1}\big(X\Delta\big)'\big(\tilde{\Sigma}_*^{(h)} \otimes I_n\big)^{-1}\overline{D}_ny\\ &= \beta + \Delta^{-1}\big[X'\big(\tilde{\Sigma}_*^{(h)} \otimes I_n\big)^{-1}X\big]^{-1}X'\big(\tilde{\Sigma}_*^{(h)} \otimes I_n\big)^{-1}\overline{D}_nu\\ &= \beta + \Delta^{-1}\tilde{\beta}_*^{(h)}\end{aligned}$$

where

$$\tilde{\beta}_*^{(h)} = \left[X'(\tilde{\Sigma}_*^{(h)} \otimes I_n)^{-1}X\right]^{-1} X'(\tilde{\Sigma}_*^{(h)} \otimes I_n)^{-1} w\,.$$

For $h \geq 1$, we then see that:

$$\begin{aligned}\overline{D}_n \tilde{u}^{(h)} &= \overline{D}_n(y - X\tilde{\beta}^{(h-1)}) = \overline{D}_n(X\beta + u - X\beta - X\Delta^{-1}\tilde{\beta}_*^{(h-1)}) \\ &= w - X\tilde{\beta}_*^{(h-1)}\end{aligned}$$

hence

$$\tilde{u}_{i*}^{(h)} \equiv (1/\sigma_i)\tilde{u}_i^{(h)} = w_i - X_i\tilde{\beta}_{i*}^{(h-1)}\,,\ i = 1,\, \ldots\, ,\, p.$$

This completes the proof of the proposition. □

Propositions **4.1** and **4.2** show that the distributions of the LM, Harvey-Phillips and LR statistics only depend on the distributions of X and w, irrespective whether the null hypothesis H_0 holds or not. This property also carries to procedures based on combining several of these test statistics, such as the induced Harvey-Phillips tests proposed in Section 3.2. In particular, under the strict exogeneity assumption (2.8), this means that the conditional distributions (given X) of these test statistics only depend on the distribution of w (and the known value of X). If we further assume that the joint distribution of $W_1,\, \ldots\, ,\, W_n$ is completely specified [assumption (2.7)], then under H_0 the distribution of w does not involve any unknown parameter, and similarly for the LM, Harvey-Phillips and LR statistics. For the QLR statistics, the same properties will hold provided we assume that $\tilde{\Sigma}_*^{(0)} \equiv D_p(\sigma_i^{-1})\tilde{\Sigma}^{(0)}D_p(\sigma_i^{-1})$ can be rewritten as a function of X and w. In particular, this will be the case if the initial value $\tilde{\Sigma}^{(0)}$ is obtained from the least squares residuals from the p separate regressions in (2.1), *i.e.* if

$$\tilde{\Sigma}^{(0)} = \frac{1}{n}\widehat{U}'\widehat{U}\,,\ \widehat{U} = [\widehat{u}_1,\, \ldots\, ,\, \widehat{u}_p]\,,\ \widehat{u}_i = M(X_i)Y_i\,,\ i = 1,\, \ldots\, ,\, p. \tag{4.17}$$

We can thus state the following proposition.

Proposition 4.4 PIVOTAL PROPERTY OF TESTS FOR CROSS-EQUATION CORRELATION. *Under the assumptions and notations* (2.1) *to* (2.8)*, the* LM *statistic, the LR-based statistic* $\tilde{\Lambda}$ *and all the statistics of the form* $F_i[K_{(i)}]$*, where* $K_{(i)}$ *is some* (*non-empty*) *subset of* $\{j : 1 \leq j \leq p\,,\, j \neq i\}$*, follow a joint distribution* (*conditional on* X) *that does not depend on any unknown parameter under the null hypothesis* $H_0 : \Sigma = D_p(\sigma_i^2)$ *. If furthermore*

$$D_p(\sigma_i^{-1})\tilde{\Sigma}^{(0)}D_p(\sigma_i^{-1}) = H(X,\, w) \tag{4.18}$$

where $H(X,\, w)$ *is a known function of* X *and* w*, the same property holds for the QLR statistics* $\tilde{\Lambda}^{(h)}$, $h \geq 0$.

It is of interest to note here that the pivotal property for the LR statistics $\tilde{\Lambda}$ could also be obtained by using the invariance results for generalized regressions models given by Breusch (1980).

However this would not simplify our proof and would not yield the explicit representation provided by Proposition **4.2**. As we will see below, the latter can be useful for implementing MC tests.

The fact that the LM, Harvey-Phillips, LR and QLR statistics have nuisance-parameter-free null distributions entails that MC tests can be applied here to obtain a finite-sample version of the corresponding tests. Such tests can be implemented as follows. Consider a test statistic T for H_0 with a continuous nuisance-parameter-free null distribution, suppose H_0 is rejected when T is large $\big[$*i.e.*, when $T \geq c(\alpha)$, where $P[T \geq c(\alpha)] = \alpha$ under $H_0\big]$, and denote by $G(x) = P[T \geq x]$ its survival function under the null hypothesis. Let T_0 be the test statistic computed from the observed data. Then the associated critical region of size α may be expressed as $G(T_0) \leq \alpha$. By Monte Carlo methods, generate N independent realizations $T_1 , \ldots , T_N$ of T under H_0. Now compute the randomized "p-value" $\hat{p}_N(T_0)$, where

$$\hat{p}_N(x) = \frac{N\hat{G}_N(x) + 1}{N + 1} , \tag{4.19}$$

$$\hat{G}_N(x) = \frac{1}{N}\sum_{i=1}^{N} I_{[0,\infty)}(T_i - x), \;\; I_A(x) = \begin{cases} 1, & \text{if } x \in A \\ 0, & \text{if } x \notin A \end{cases} .$$

Then we can show that

$$P[\hat{p}_N(T_0) \leq \alpha] = \frac{I[\alpha(N+1)]}{N+1} ; \tag{4.20}$$

see Dufour and Kiviet (1998). In particular, if we choose N so that $\alpha(N + 1)$ is an integer (*e.g.*, for $\alpha = 0.05$, we can take $N = 19$, 39, 99, etc.), we have:

$$P[\hat{p}_N(T_0) \leq \alpha] = \alpha . \tag{4.21}$$

In other words, the randomized critical region $\hat{p}_N(T_0) \leq \alpha$ has the same level as the critical region $G(T_0) \leq \alpha$. This procedure is of course valid when the error vectors $U_{t\cdot}$ are i.i.d. normal [Assumption (2.10)], but also under parametric distributional assumptions when J is the only unknown parameter in the distribution of $U_{t\cdot}$, $t = 1, \ldots , n$.

MC tests can be interpreted as parametric bootstrap methods applied to statistics whose null distribution does not depend on nuisance parameters, with however the central additional observation that the randomization allows one to exactly control the size of the test for a given (possibly small) number of MC simulations. For further discussion of Monte Carlo tests (including its relation with the bootstrap), see Dufour (1995), Dufour and Kiviet (1996), Kiviet and Dufour (1997), Dufour et al. (1998), and Dufour and Khalaf (2001). On the bootstrap, the reader may consult Hall (1992), Efron and Tibshirani (1993), Jeong and Maddala (1993), Vinod (1993), Shao and Tu (1995), and Horowitz (1997).

5. Simulation experiments

In order to assess the performance of the various procedures discussed above, we conducted a set of Monte Carlo experiments for a five-equation model ($p = 5$) with $n = 25$ observations. To assess test

Table 1. Covariance matrices used in the Monte Carlo experiments

	.0007773	6.616e-06	-1.082e-05	.0003573	-.0001443
	.0024550	.0001923	-.0010390	-.0006195	
Σ_1	.0002950	1.747e-05	.0002829		
	.0007560	.0004105			
	.0006790				
	.0007773	1.654e-06	-1.353e-06	3.969e-05	-1.804e-05
	.0024550	2.405e-05	-.0001737	-7.732e-05	
Σ_2	.0002800	2.427e-05	5.417e-05		
	.0001276	2.495e-05			
	4.863e-05				
	.0007773	3.308e-06	-3.607e-06	8.931e-05	-3.608e-05
	.0024550	9.618e-05	-.0003471	-.0001238	
Σ_3	.0002836	3.804e-05	.0001051		
	.0001800	7.966e-05			
	.0001029				
	.0007773	8.271e-07	-1.803e-06	.0001786	-2.062e-05
	.0024550	2.138e-05	-.0002083	-.0002061	
Σ_4	.0002800	1.513e-05	3.485e-05		
	.0001707	2.421e-05			
	5.630e-05				

size, we also considered $n = 50, 100$. In each experiment, the design matrices X_i , $i = 1, \ldots , p$, include a constant term and equal numbers of regressors ($k_i = k$, $i = 1, \ldots , p$). The values of k considered are $k = 5, 6, \ldots , 15$. The variables in each matrix X_i were generated using a multivariate normal distribution and kept constant over all replications. The disturbances were generated from multivariate normal distributions. Several choices for the error covariance were considered and are listed in Table 1. The Σ_1 matrix as well as the regression coefficients used were taken from the empirical example discussed in Section 6.[1] The other matrices were obtained by dividing certain elements of the Cholesky decomposition of Σ_1 by appropriate constants to decrease the covariance terms. Of course, the parameters under the null were obtained by setting the non-diagonal elements of Σ_1 to zero. The numbers of trials for the MC tests were set to 19 and 99 ($N = 19, 99$). The number of overall replications was 1000. All experiments were performed with Gauss 386iVM (version 3.2.13). The results are presented in Tables 2 and 3.

Our main findings can be summarized as follows.

1. The asymptotic tests (Asy.) consistently overreject. Indeed, we can see that the empirical sizes can be substantially larger than the nominal 5%. This is in accordance with well documented results on LR-based multivariate tests. On the other hand, our conclusions with

[1]The statistics studied are all invariant to the values of the regression coefficients.

Table 2. Empirical sizes of LM and quasi-LR independence tests

$p=5$	$n=25$				$n=50$				$n=100$			
k	QLR_{OLS}		LM		QLR_{OLS}		LM		QLR_{OLS}		LM	
	Asy.	MC	Asy.	MC	Asy.	MC	Asy.	MC	Asy.	MC	Asy.	MC
5	.193	.040	.105	.045	.115	.057	.081	.057	.070	.040	.062	.037
6	.198	.046	.122	.052	.115	.055	.082	.050	.071	.046	.054	.036
7	.307	.050	.172	.057	.137	.061	.108	.057	.069	.050	.054	.037
8	.322	.048	.200	.054	.150	.057	.106	.050	.080	.048	.069	.045
9	.413	.049	.263	.052	.158	.048	.107	.046	.087	.049	.073	.038
10	.478	.055	.336	.058	.184	.050	.139	.052	.091	.055	.071	.040
11	.536	.038	.353	.049	.190	.054	.146	.056	.092	.038	.076	.036
12	.601	.040	.432	.045	.210	.048	.150	.049	.096	.040	.079	.041
13	.650	.057	.505	.043	.230	.047	.179	.040	.109	.057	.088	.037
14	.725	.059	.577	.051	.236	.042	.185	.048	.115	.059	.095	.036
15	.816	.052	.684	.064	.271	.045	.213	.055	.120	.052	.109	.047

Table 3. Empirical rejections of various independence tests

$n=25$	Σ_0 (H_0)		Σ_1		Σ_2		Σ_3		Σ_4	
	Asy.	MC	MC		MC		MC		MC	
MC replications	-	19	19	99	19	99	19	99	19	99
LM	.105	.045	.998	1.0	.911	.954	.704	.794	.444	.500
QLR_{OLS}	.193	.040	1.0	1.0	.947	.971	.744	.820	.438	.494
QLR_{GLS}	.260	.040	1.0	1.0	.959	.979	.750	.825*	.429	.504
LR	.267	.047	1.0	1.0	.961	.980	.746	.824	.428	.494
$F_{\min}$	-	.043	1.0	1.0	.925	.965	.632	.693	.360	.409
$F_{\times}$	-	.052	1.0	1.0	.944	.980	.714	.784	.382	.438
$FS_{\min}$	-	.049	1.0	1.0	.846	.912	.562	.653	.368	.399
$FS_{\times}$		.052	1.0	1.0	.963	.984*	.721	.799	.490	.562*
	Bonferroni Harvey-Phillips type tests									
$F_{\min}$	.034		1.0		.963		.665		.356	
$FS_{\min}$	.049		1.0		.896		.687		.316	

respect to the LM test are not in agreement with the available Monte Carlo evidence, in which LM independence test was found to work well. This was due to the fact that small numbers of equations were studied in the earlier literature. Here we find that it does not always work well in larger systems. In contrast, the MC versions of the tests achieve perfect size control, i.e. the probability of type I error is equal to the nominal size of 5%.

2. The size corrected tests perform quite well. The power of all four MC tests are comparable to each other, although the LR-type tests exhibit better power. The EFT test shows relatively lower power, as would be expected.

3. Iterating SURE estimators to convergence is clearly not worthwhile, in the sense of improving the power of the associated LR test. In fact, in some cases, iterations resulted in slight power losses. Furthermore, our results give very favorable support to the OLS-based QLR test. This issue is particularly pertinent in the context of simulation-based tests.

4. The MC induced tests based on the Harvey-Phillips statistics perform very well over all the parameter values considered. As expected, the Tippet/Wilkinson-type MC induced tests perform better than their Bonferroni counterparts. The power of the Fisher/Pearson-type induced tests is generally higher than the power of the Tippet/Wilkinson-type ones. Further, the sequential variants of the induced tests perform better than the non-sequential ones. Indeed, in two cases over three, the sequential Fisher/Pearson-type induced test $(FS_{\times})$ exhibits the best power among all the tests considered.

6. Application to growth equations

For illustrative purposes, we studied data previously analyzed by Fischer (1993) which contains several series of macroeconomic aggregates observed yearly for a large panel of countries. The dependent variables of interest are four growth indicators: GDP growth, capital accumulation, productivity growth (measured by Solow residuals), and labor force growth. The following determinants of growth are considered: the inflation rate, the ratio of budget surplus to GDP, the terms of trade, and the black market premium on the exchange rate. Fischer focuses on explaining the determinants of growth. The econometric specification consists of an unbalanced panel model, which assumes contemporaneously uncorrelated disturbances. Here, we shall test the latter specification. Attention is restricted to the multiple regressions (17), (23), (29) and (35) in Fischer (1993), which include all four explanatory variables. The choice of countries was motivated by the availability of observations on all included variables. We consider:

A) the South-American region (1973-1987): 1) Mexico, 2) Argentina, 3) Chile, 4) Colombia, 5) Ecuador, and 6) Paraguay;

B) the African region (1977-88): 1) Ghana, 2) Côte d'Ivoire, 3) Kenya, 4) Malawi, 5) Morocco, and 6) Zambia;

C) the Asian region (1978-87): 1) Korea, 2) Pakistan, 3) Thailand, 4) India, and 5) Indonesia.

Then, for each region, we considered four SURE different systems corresponding to each one of the four growth indicators considered (where $t = 1, \ldots, n$, $i = 1, \ldots, p$) :

$$\Delta GDP_{ti} = \beta_0^G + \beta_{1i}^G INFLAT_{ti} + \beta_{2i}^G TRMTRD_{ti} + \beta_{3i}^G SRPLS_{ti} + \beta_{4i}^G EXCM_{ti} + u_{ti}^G \,;$$

$$\Delta CPTL_{ti} = \beta_0^K + \beta_{1i}^K INFLAT_{ti} + \beta_{2i}^K TRMTRD_{ti} + \beta_{3i}^K SRPLS_{ti} + \beta_{4i}^K EXCM_{ti} + u_{ti}^K \,;$$

$$\Delta PRDCT_{ti} = \beta_0^P + \beta_{1i}^P INFLAT_{ti} + \beta_{2i}^P TRMTRD_{ti} + \beta_{3i}^P SRPLS_{ti} + \beta_{4i}^P EXCM_{ti} + u_{ti}^P \,;$$

$$\Delta LABOR_{ti} = \beta_0^L + \beta_{1i}^L INFLAT_{ti} + \beta_{2i}^L TRMTRD_{ti} + \beta_{3i}^L SRPLS_{ti} + \beta_{4i}^L EXCM_{ti} + u_{ti}^L \,.$$

Here, for each country i and each year t, ΔGDP_{ti}, $\Delta CPTL_{ti}$, $\Delta PRDCT_{ti}, \Delta LABOR_{ti}$, and $EXCM_{ti}$ represent respectively GDP growth, capital accumulation, productivity growth, and labor force growth. The explanatory variables are: inflation $(INFLAT_{ti})$, terms of trade $(TRMTRD_{ti})$, the ratio of budget surplus to GDP $(SRPLS_{ti})$, and the black market premium on the exchange rate $(EXCM_{ti})$. Overall, we consider 12 different SURE systems with either 6 equations (South America, Africa) or 5 equations (Asia), each system corresponding to a region and one of the four growth indicators. Countries are numbered inside each region as indicated in the list presented at beginning of this section (this ordering correspond to the World Bank database that we used).

We will now test whether the disturbances inside each one of these SURE systems are contemporaneously correlated, using a Gaussian distributional assumption. The assumption that the disturbances are not correlated across countries is important to justify pooling the data as done by Fischer (1993). In each case, we applied LM, LR and QLR tests, as well as Harvey/Phillips-type induced tests. The MC tests are based on $N = 999$ replications of the statistics considered. The QLR tests are based on two-step feasible GLS estimators, using OLS residuals to estimate the disturbance covariance matrix. For completeness, we also report the individual Harvey-Phillips tests (based on the statistics $F_i(K_i)$ and $F_i[\{i+1, \ldots, p\}]$ defined in Section 3.2) which are combined by the MC induced tests. Note that in the case of the sequential tests, the ordering of the countries may affect the outcome of the test; here, we present results based on the ordering given above. The results are presented (as p-values) in Tables 4 to 7. The MC test results which are significant at the 10% level are highlighted with one star (*), while those which are significant at the 5% level are highlighted with two stars (**). In view of the simulation evidence of Section 5, we shall stress the conclusions provided by the MC LR-based and FS tests. Asymptotic p-values (Asy.) are only reported for comparison sake.

For GDP growth (Table 4), no test is significant (at the 10% level) in the case of the South-American countries. For Africa, the MC LR-type tests are significant at the 10% level (but not 5%), but the $FS_{\min}$ induced test is significant at the 5% level. On looking at the individual sequential Harvey-Phillips tests, it appears this may be due to correlations between the disturbances in the Malawi equation and those for Morocco and/or Zambia. Turning to the Asian region, while the LR-based tests are not significant again, we nevertheless observe that the $F_{\min}$, $FS_{\min}$ and $FS_{\times}$ are significant at the 10% level. In this case, the Harvey-Phillips sequential tests suggest that there may be dependence between Korea and the other countries. For all regions, it is of interest to observe that the asymptotic approximations and the MC procedure yield very different p-values for the LR-based statistics, which may lead to quite different conclusions. This observation also applies to the

Table 4. GDP growth SURE systems: independence tests

	South America $p=6$		Africa $p=6$		Asia $p=5$	
$F_1(K_1)$	.7927		.1613		.0215**	
$F_2(K_2)$	.7906		.2882		.6068	
$F_3(K_3)$	.2669		.1308		.7127	
$F_4(K_4)$	.9901		.0516*		.5453	
$F_5(K_5)$	.8503		.9571		.3771	
$F_6(K_6)$	.8253		.2652		-	
$F_1(\{2, \dots, p\})$	.7927		.1613		.0215**	
$F_2(\{3, \dots, p\})$	.7470		.4964		.3113	
$F_3(\{4, \dots, p\})$	.8810		.9137		.4277	
$F_4(\{5, \dots, p\})$	.8647		.0055**		.3873	
$F_5(\{p\})$	.9290		.6005		-	
	Asy.	MC	Asy.	MC	Asy.	MC
LM	.9425	.977	.0466	.081*	.4384	.611
QLR_{OLS}	.9242	.981	.0100	.062*	.0872	.470
LR	.4374	.978	.0000	.082*	.0000	.412
$F_{\min}$	-	.742	-	.224	-	.094*
$F_{\times}$	-	.917	-	.130	-	.258
$FS_{\min}$	-	1.0	-	.025**	-	.085*
$FS_{\times}$	-	1.0	-	.072*	-	.096*

Table 5. Capital growth SURE systems: independence tests

	South America $p = 6$		Africa $p = 6$		Asia $p = 5$	
$F_1(K_1)$	.1592		.2503		.3399	
$F_2(K_2)$	.3514		.2035		.2200	
$F_3(K_3)$	.0561*		.7065		.1373	
$F_4(K_4)$	.0333**		.0589*		.2255	
$F_5(K_5)$	.2288		.0143**		.0357	
$F_6(K_6)$	.3509		.8389		-	
$F_1(\{2, \ldots, p\})$	.1592		.2503		.3399	
$F_2(\{3, \ldots, p\})$	.2249		.0854*		.6323	
$F_3(\{4, \ldots, p\})$	.0111*		.3004		.0069*	
$F_4(\{5, \ldots, p\})$	.7156		.1422		.6355	
$F_5(\{p\})$	.7679		.7581		-	
	Asy.	MC	Asy.	MC	Asy.	MC
LM	.0350	.061*	.1023	.026**	.2449	.367
QLR_{OLS}	.0058	.096*	.0049	.132	.0000	.002**
LR	.0000	.053*	.0000	.249	.0000	.001**
$F_{\min}$	-	.167	-	.063*	-	.137
$F_{\times}$	-	.075*	-	.080*	-	.056*
$FS_{\min}$	-	.055*	-	.359	-	.027**
$FS_{\times}$	-	.086*	-	.141	-	.091*

Table 6. Productivity growth SURE systems: independence tests

	South America $p=6$		Africa $p=6$		Asia $p=5$	
$F_1(K_1)$	.9765		.1312		.5003	
$F_2(K_2)$	.9162		.1909		.8182	
$F_3(K_3)$	.5362		.2965		.4958	
$F_4(K_4)$	.9976		.0242**		.1246	
$F_5(K_5)$	.9430		.8209		.0918	
$F_6(K_6)$	.7528		.2454		-	
$F_1(\{2, \ldots, p\})$	.9765		.1312		.5003	
$F_2(\{3, \ldots, p\})$	.8294		.3912		.5421	
$F_3(\{4, \ldots, p\})$	.6037		.3738		.8683	
$F_4(\{5, \ldots, p\})$	.9442		.0519*		.2284	
$F_5(\{p\})$	.6962		.8069		-	
	Asy.	MC	Asy.	MC	Asy.	MC
LM	.9913	.998	.0356	.061*	.5070	.698
QLR_{OLS}	.9891	.997	.0012	.074*	.0658	.415
LR	.7929	.998	.0000	.016**	.0000	.266
$F_{\min}$	-	.943	-	.111	-	.337
$F_{\times}$	-	.979	-	.093*	-	.282
$FS_{\min}$	-	.988	-	.212	-	.636
$FS_{\times}$	-	.990	-	.152	-	.664

Table 7. Labor force growth SURE systems: independence tests

	South America $p = 6$		Africa $p = 6$		Asia $p = 5$	
$F_1(K_1)$	.4051		.0734*		.0153*	
$F_2(K_2)$	.4051		.2266		.684	
$F_3(K_3)$	.1976		.0397**		.0957*	
$F_4(K_4)$	.0189**		.0153**		.8999	
$F_5(K_5)$	.0288**		.3571		.4434	
$F_6(K_6)$	.1011		.1761		-	
$F_1(\{2, \ldots, p\})$	.4051		.0734*		.0153*	
$F_2(\{3, \ldots, p\})$	.2594		.1201		.5698	
$F_3(\{4, \ldots, p\})$	.5535		.0085**		.2187	
$F_4(\{5, \ldots, p\})$	.0580*		.0897*		.2661	
$F_5(\{p\})$	.0799*		.4900		-	
	Asy.	MC	Asy.	MC	Asy.	MC
LM	.0686	.103	.0007	.004**	.3040	.457
QLR_{OLS}	.0108	.126	.0000	.006**	.0063	.140
LR	.0000	.020*	.0000	.004**	.0000	.194
$F_{\min}$	-	.102	-	.065*	-	.062*
$F_{\times}$	-	.045**	-	.021**	-	.128
$FS_{\min}$	-	.272	-	.033**	-	.062*
$FS_{\times}$	-	.092*	-	.008**	-	.050**

results for the other growth indicators discussed below.

For capital growth (Table 5), the MC LR and $FS_{\min}$ tests are strongly significant for Asia and close to being significant at the 5% level for South America. The same tests do not come out significant at usual levels for Africa, although the LM, $F_{\min}$ and $F_{\times}$ also provide indications of dependence in this case too. The Harvey-Phillips individual tests suggest there is dependence between the disturbances in the equation for Chile and those for Colombia, Ecuador and Paraguay; in Africa, the dependence appears to be between Thailand, India and Indonesia.

In the case of productivity growth (Table 6), we see no evidence of cross-equation correlation for both South America and Asia, but some for Africa. In the latter case, the MC-LR statistic is strongly significant, and to a lesser extent the quasi-LR and $F_{\times}$ tests. On looking at the individual sequential Harvey-Phillips tests, it seems this may be due to correlation between Morocco and Zambia.

For labor force growth (Table 7), we see strong evidence of cross-equation correlation in the cases of South America and Africa. For Asia, the LR-based tests are not significant at the 10% level, but induced Harvey-Phillips tests are significant at the 5% level (or close to it).

Overall, these results provide several examples where asymptotic p-values grossly overstate test significance. Despite this fact, using more reliable finite-sample methods, we also found quite significant evidence of contemporaneous correlation between the disturbances in several of the equations considered, a feature that should be taken into account when analyzing these data. Of course, it is beyond the scope of the present paper to perform a complete reanalysis of the Fischer (1993) data.

7. Conclusion

In this paper, we have proposed simulation-based procedures to derive exact p-values for standard LR and LM independence tests in the context of SURE models. We have also proposed alternative OLS and IFGLS-based QLR criteria. In multi-equation models, conventional independence tests only have an asymptotic justification. The reason for the lack of popularity of finite sample procedures is clearly the intractable nature of available distributional results. Here, we have considered an alternative and considerably more straightforward approach to independence tests. We have shown that LR and LM statistics are pivotal under the null, which implies that exact critical values can be obtained easily by MC techniques.

The feasibility of the approach suggested was illustrated through both a simulation experiment and an empirical application. The results show that asymptotic tests are indeed highly unreliable; in contrast, MC tests achieve size control and have good power. We emphasize that OLS-based MC QLR tests performed extremely well. This aspect is important particularly in larger systems, since test procedures based on iterative estimators are typically more expensive from the point of view of MC tests. The MC induced tests turned out to have surprisingly good power. Since the MC test procedure yields size-correct significance points, this approach seems very promising in the context of non-independent simultaneous tests.

References

Anderson, T. W. (1971), *The Statistical Analysis of Time Series*, John Wiley & Sons, New York.

Anderson, T. W. (1984), *An Introduction to Multivariate Statistical Analysis*, second edn, John Wiley & Sons, New York.

Barnard, G. A. (1963), 'Comment on 'The spectral analysis of point processes' by M. S. Bartlett', *Journal of the Royal Statistical Society, Series B* **25**, 294.

Berndt, E. R. and Savin, N. E. (1977), 'Conflict among criteria for testing hypotheses in the multivariate linear regression model', *Econometrica* **45**, 1263–1277.

Breusch, T. S. (1980), 'Useful invariance results for generalized regression models', *Journal of Econometrics* **13**, 327–340.

Breusch, T. S. and Pagan, A. R. (1980), 'The Lagrange multiplier test and its applications to model specification in econometrics', *Review of Economic Studies* **47**, 239–254.

Cameron, A. C. and Trivedi, P. K. (1993), 'Tests of independence in parametric models with applications and illustrations', *Journal of Business and Economic Statistics* **11**, 29–43.

Dagenais, M. G. and Dufour, J.-M. (1991), 'Invariance, nonlinear models and asymptotic tests', *Econometrica* **59**, 1601–1615.

Dufour, J.-M. (1989), 'Nonlinear hypotheses, inequality restrictions, and non-nested hypotheses: Exact simultaneous tests in linear regressions', *Econometrica* **57**, 335–355.

Dufour, J.-M. (1990), 'Exact tests and confidence sets in linear regressions with autocorrelated errors', *Econometrica* **58**, 475–494.

Dufour, J.-M. (1995), Monte Carlo tests with nuisance parameters: A general approach to finite-sample inference and nonstandard asymptotics in econometrics, Technical report, C.R.D.E., Université de Montréal.

Dufour, J.-M., Farhat, A., Gardiol, L. and Khalaf, L. (1998), 'Simulation-based finite sample normality tests in linear regressions', *The Econometrics Journal* **1**, 154–173.

Dufour, J.-M. and Khalaf, L. (1998), Simulation based finite and large sample tests in multivariate regressions, Technical report, C.R.D.E., Université de Montréal. 27 pages. forthcoming in the Journal of Econometrics.

Dufour, J.-M. and Khalaf, L. (2001), Monte Carlo test methods in econometrics, *in* B. Baltagi, ed., 'Companion to Theoretical Econometrics', Blackwell Companions to Contemporary Economics, Basil Blackwell, Oxford, U.K., chapter 23, pp. 494–519.

Dufour, J.-M. and Kiviet, J. F. (1996), 'Exact tests for structural change in first-order dynamic models', *Journal of Econometrics* **70**, 39–68.

Dufour, J.-M. and Kiviet, J. F. (1998), 'Exact inference methods for first-order autoregressive distributed lag models', *Econometrica* **66**, 79–104.

Dufour, J.-M. and Torrès, O. (1998), Union-intersection and sample-split methods in econometrics with applications to SURE and MA models, *in* A. Ullah and D. E. A. Giles, eds, 'Handbook of Applied Economic Statistics', Marcel Dekker, New York, pp. 465–505.

Dufour, J.-M. and Torrès, O. (2000), 'Markovian processes, two-sided autoregressions and exact inference for stationary and nonstationary autoregressive processes', *Journal of Econometrics* **99**, 255–289.

Dwass, M. (1957), 'Modified randomization tests for nonparametric hypotheses', *Annals of Mathematical Statistics* **28**, 181–187.

Efron, B. and Tibshirani, R. J. (1993), *An Introduction to the Bootstrap*, Vol. 57 of *Monographs on Statistics and Applied Probability*, Chapman & Hall, New York.

Fischer, S. (1993), 'The role of macroeconomic factors in growth', *Journal of Monetary Economics* **32**, 485–512.

Fisher, R. A. (1932), *Statistical Methods for Research Workers*, Oliver and Boyd, Edinburgh.

Folks, J. L. (1984), Combination of independent tests, *in* P. R. Krishnaiah and P. K. Sen, eds, 'Handbook of Statistics 4: Nonparametric Methods', North-Holland, Amsterdam, pp. 113–121.

Gouriéroux, C. and Monfort, A. (1995), *Statistics and Econometric Models, Volumes One and Two*, Cambridge University Press, Cambridge, U.K.

Hall, P. (1992), *The Bootstrap and Edgeworth Expansion*, Springer-Verlag, New York.

Harvey, A. C. and Phillips, G. D. A. (1980), 'Testing for serial correlation in simultaneous equation models', *Econometrica* **48**, 747–759.

Harvey, A. C. and Phillips, G. D. A. (1982), 'Testing for contemporaneous correlation of disturbances in systems of regression equations', *Bulletin of Economic Research* **34(2)**, 79–81.

Horowitz, J. L. (1997), Bootstrap methods in econometrics: Theory and numerical performance, *in* D. Kreps and K. W. Wallis, eds, 'Advances in Economics and Econometrics: Theory and Applications. Seventh World Congress. 3 Volumes', Cambridge University Press, Cambridge, U.K., pp. 188–222.

Jeong, J. and Maddala, G. S. (1993), A perspective on application of bootstrap methods in econometrics, *in* (Maddala, Rao and Vinod 1993), pp. 573–610.

Kariya, T. (1981*a*), 'Bounds for the covariance matrices of Zellner's estimator in the SUR model and 2SAE in a heteroscedastic model', *Journal of the American Statistical Association* **76**, 975–979.

Kariya, T. (1981*b*), 'Robustness of multivariate tests', *The Annals of Statistics* **6**, 1267–1275.

Kariya, T. (1981*c*), 'Tests for the independence between two seemingly unrelated regression equations', *The Annals of Statistics* **9(2)**, 381–390.

Kariya, T. (1985), *Testing in the Multivariate General Linear Model*, number 22 *in* 'Economic Research Series, The Institute of Economic Research, Hitotsubashi University, Japan', Kinokuniya Company Ltd., Tokyo.

Kariya, T., Fujikoshi, Y. and Krishnaiah, P. R. (1984), 'Tests for independence between two multivariate regression equations with different design matrices', *Journal of Multivariate Analysis* **15**, 383–407.

Kiviet, J. F. and Dufour, J.-M. (1997), 'Exact tests in single equation autoregressive distributed lag models', *Journal of Econometrics* **80**, 325–353.

Kmenta, J. and Gilbert, R. F. (1968), 'Small sample properties of alternative estimators of seemingly unrelated regressions', *Journal of the American Statistical Association* **63**, 1180–1200.

Kunitomo, N. (1977), 'A note on the efficiency of Zellner's estimator for the case of two seemingly unrelated regression', *Economic Studies Quarterly* **28**, 73–77.

Lehmann, E. L. (1957), 'A theory of some multiple decision problems. II', *Annals of Mathematical Statistics* **28**, 547–572.

Maddala, G. S., Rao, C. R. and Vinod, H. D., eds (1993), *Handbook of Statistics 11: Econometrics*, North-Holland, Amsterdam.

Mehta, J. S. and Swamy, P. A. V. B. (1976), 'Further evidence on the relative efficiencies of Zellner's seemingly unrelated regressions estimator', *Journal of the American Statistical Association* **71**, 634–639.

Miller, Jr., R. G. (1981), *Simultaneous Statistical Inference*, second edn, Springer-Verlag, New York.

Oberhofer, W. and Kmenta, J. (1974), 'A general procedure for obtaining maximum likelihood estimates in generalized regression models', *Econometrica* **42**, 579–590.

Pearson, K. (1933), 'On a method of determining whether a sample of size n supposed to have been drawn from a parent population', *Biometrika* **25**, 379–410.

Revankar, N. S. (1974), 'Some finite sample results in the context of two seemingly unrelated regressions', *Journal of the American Statistical Association* **68**, 187–190.

Revankar, N. S. (1976), 'Use of restricted residuals in SUR systems: Some finite sample results', *Journal of the American Statistical Association* **71**, 183–188.

Robinson, P. M. (1988), 'The stochastic difference between econometric statistics', *Econometrica* **56**(3), 531–548.

Savin, N. E. (1984), Multiple hypothesis testing, *in* Z. Griliches and M. D. Intrilligator, eds, 'Handbook of Econometrics, Volume 2', North-Holland, Amsterdam, chapter 14, pp. 827–879.

Shao, S. and Tu, D. (1995), *The Jackknife and Bootstrap*, Springer-Verlag, New York.

Shiba, T. and Tsurumi, H. (1988), 'Bayesian and non-Bayesian tests of independence in seemingly unrelated regressions', *International Economic Review* **29**, 377–389.

Srivastava, V. K. and Dwivedi, T. D. (1979), 'Estimation of seemingly unrelated regression equations: A brief survey', *Journal of Econometrics* **10**, 15–32.

Srivastava, V. K. and Giles, D. E. (1987), *Seemingly Unrelated Regression Equations Models. Estimation and inference*, Marcel Dekker, New York.

Tippett, L. H. (1931), *The Methods of Statistics*, Williams and Norgate, London.

Vinod, H. D. (1993), Bootstrap methods: Applications in econometrics, *in* (Maddala et al. 1993), pp. 629–661.

Westfall, P. H. and Young, S. S. (1993), *Resampling-Based Multiple Testing: Examples and Methods for p-Value Adjustment*, John Wiley & Sons, New York.

Wilkinson, B. (1951), 'A statistical consideration in psychological research', *Psychology Bulletin* **48**, 156–158.

Zellner, A. (1962), 'An efficient method for estimating seemingly unrelated regressions and tests for aggregate bias', *Journal of the American Statistical Association* **57**, 348–368.

Zellner, A. (1963), 'Estimators for seemingly unrelated regressions: Some exact finite sample results', *Journal of the American Statistical Association* **58**, 977–992.