Finite-sample simulation-based tests in seemingly unrelated regressions [*]

Jean-Marie Dufour [†] and Lynda Khalaf [‡]

November 1998
Revised: February 2001
Compiled: February 10, 2005

This paper has been published in the *Oxford Bulletin of Economics and Statistics*, 65 (2003), 891-906.

[*] This work was supported by 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). This paper is a revised version of the earlier working paper Dufour and Khalaf (1998*a*). The authors thank Emanuela Cardia, Marcel Dagenais, John Galbraith, Eric Ghysels, James McKinnon, Christophe Muller, Olivier Torrès and Michael Veall for useful comments. Earlier versions of this paper were presented at the North American Meetings of the Econometric Society, the Third International Conference on Computing and Finance (Hoover Institution, Stanford University), the Annual Meetings of the American Statistical Association, the Canadian Econometric Study Group, the Canadian Economic Association, and at Ohio State University (Economics).

[†] C.R.D.E, CIRANO and Département de sciences économiques, Université de Montréal. Mailing address: C.R.D.E, 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. e-mail: lynda.khalaf@ecn.ulaval.ca

ABSTRACT

In this paper, we propose finite and large sample likelihood based test procedures for possibly nonlinear hypotheses on the coefficients of SURE systems. Two complementary approaches are described. First, we propose an exact Monte Carlo bounds test based on the standard likelihood ratio criterion. Second, we consider alternative Monte Carlo tests which can be run whenever the bounds are not conclusive. These include, in particular, quasi-likelihood ratio criteria based on non-maximum-likelihood estimators. Illustrative Monte Carlo experiments show that: (i) the bounds are sufficiently tight to yield conclusive results in a large proportion of cases, and (ii) the randomized procedures correct all the usual size distortions in such contexts. The procedures proposed are finally applied to test restrictions on a factor demand model.

Key words: multivariate linear regression; seemingly unrelated regressions; Monte Carlo test; bounds test; nonlinear hypothesis; finite-sample test; exact test; bootstrap; factor demand; cost function.

Journal of Economic Literature classification: C3; C12; C33; C15 O4; O5.

Contents

List of Definitions, Propositions and Theorems

List of Tables

1. Introduction

In this paper, we study the problem of testing general, possibly nonlinear constraints on the coefficients of the seemingly unrelated regressions (SURE) model introduced by Zellner (1962). The SURE model may be cast as a system of regression equations with contemporaneously correlated disturbances, where the regressors may differ across equations. For a detailed review, the reader may consult Srivastava and Giles (1987).

In connection with the SURE model, very few analytical finite-sample results are available. A rare exception is provided by Harvey and Phillips (1982, Section 3) who derived independence tests between the disturbances of an equation and those of the other equations of a SURE model. The tests involve conventional F-statistics and are based on the residuals obtained from regressing each dependent variable on all the independent variables of the system. Of course this problem is a very special one. In a different vein, Phillips (1985) derived the exact distribution of a two-stage SURE estimator using a fractional matrix calculus. However, the analytical expressions obtained are very complex and, more importantly, involve unknown nuisance parameters, namely the elements of the error covariance matrix. The latter fact makes the application of Phillips' distributional results to practical hypothesis testing problematic.

Asymptotic Wald, Lagrange multiplier and likelihood ratio tests are available and commonly employed in empirical applications of the SURE model; see, for example, Breusch (1979) or Srivastava and Giles (1987). It has been shown however that, in finite samples, the asymptotic criteria are seriously biased towards overrejection, with the problem getting worse as the number of equations grows relative to the sample size; see, for example, Laitinen (1978), Meisner (1979), Bera, Byron and Jarque (1981), Theil and Fiebig (1985), and Dufour and Khalaf (1998*b*). Attempts to improve standard asymptotic tests include, in particular: (i) Bartlett-type corrections, and (ii) bootstrap and simulation-based methods. See, for example, Rocke (1989), Rayner (1990), Rilstone and Veall (1996), Theil, Shonkwiler and Taylor (1985), Theil, Taylor and Shonkwiler (1986), Taylor, Shonkwiler and Theil (1986), and Theil and Fiebig (1985).

Further results relevant to the SURE model can be found in the statistics and econometrics literature on multivariate linear regressions (MLR). These are relevant because the MLR model can be viewed as a special case of the SURE model where the regressor matrices for the different equations are identical. For reviews and further references on exact and asymptotic inference in MLR models, the reader may consult Rao (1973, Chapter 8), Anderson (1984, chapters 8 and 13), Kariya (1985), Stewart (1997) and Dufour and Khalaf (1998*b*). In particular, besides showing the inadequacy of various size-correction procedures (including Bartlett corrections) through simulation, we derived in Dufour and Khalaf (1998*b*) exact bounds on the null distribution of LR test statistics for possibly non-linear hypotheses on regression coefficients in MLR models. Even though computing analytically these bounds may be difficult, they can easily be evaluated by simulation and implemented as finite-sample *bounds Monte Carlo tests.* The implications for hypothesis testing are two-fold. First, the finite-sample bounds on the LR criterion easily yield conservative tests, for both linear and non-linear hypotheses. Second, Monte Carlo test methods can lead to tests with correct levels. This is related to the fact that LR statistics are *pivotal* or *boundedly pivotal* for quite general hypotheses in the MLR model [see the discussion in Dufour (1997) on boundedly pivotal statistics].

In this paper, we extend the results presented in Dufour and Khalaf (1998*b*) to the case of SURE systems, under both gaussian and non-gaussian disturbance distributions. Indeed the model considered here is an extension of the standard gaussian SURE model that allows for both gaussian and non-gaussian disturbance distribution, as long the latter is specified up an unknown linear transformation (or contemporaneous covariance matrix). In particular, we discuss two approaches that can be applied on their own or sequentially, namely: (i) a bounds procedure, and (ii) Monte Carlo tests. Practical implementation of both techniques is simple. To obtain the bounds, we exploit the fact that the SURE specification can be viewed as a special case of a properly chosen MLR model constrained by regressor exclusion restrictions on the different equations.

To be more specific, we give at this point a preliminary discussion of the proposed conservative bound, which can be viewed as an extension of an approach described earlier in Dufour (1989) and Dufour and Kiviet (1998). First, we reconsider the testing problem within the framework of an appropriate MLR model, namely the MLR setup of which the model on hand is a restricted form. As pointed out above, this setup allows for gaussian and non-gaussian error distributions, provided the latter can be simulated. Second, we introduce, in the relevant MLR framework, a "uniform linear (UL) hypothesis" [Berndt and Savin (1977)] which is a special case of the set of restrictions specified by the null hypothesis. The intuition behind this suggestion follows from the fact that exact nuisance-parameter free critical values for the LR criterion are available when the null is UL within a MLR. Indeed, it turns out that the LR criterion for testing the suggested UL hypothesis conveniently bounds the LR statistic for testing the general constraints.

In addition, we propose alternative Monte Carlo (MC) tests [see Dwass (1957), Barnard (1963), Jöckel (1986) or Dufour (1998)] that can be run whenever the bounds tests are not conclusive. We consider: (i) an asymptotically valid procedure that may be interpreted as a parametric bootstrap, and (ii) a method which is exact for any sample size, following Dufour (1998). Further, in situations where maximum likelihood (ML) methods may be computationally expensive, we introduce LR-type test criteria based on non-ML estimators. In particular, we consider two-stage statistics or estimators at any step of the process by which the likelihood is maximized iteratively. We emphasize that parametric bootstrap and bounds tests should be viewed as complementary rather than alternative procedures.

The paper is organized as follows. In Section 2, we present the model studied and define the test statistics which will be considered. In Section 3, we describe the proposed bounds and Monte Carlo est procedures. Simulation results are reported in Section 4. Section 5 illustrates the procedures proposed by applying them to test restrictions on a factor demand model. We conclude in Section 6.

2. Framework

We consider here a $p-$equation SURE system of the form:

$$Y_j = X_j\beta_j + u_j \ , \ j = 1, \ \dots \ , \ p \ , \tag{2.1}$$

where Y_j is a vector of n observations on a dependent variable, X_j is a full-column rank $n \times k_j$ matrix of regressors, $\beta_j = (\beta_{0j},\, \beta_{1j},\, \ldots\, ,\, \beta_{k_j-1,j})'$ is a vector of k_j unknown coefficients, and $u_j = (u_{1j},\, u_{2j},\, \ldots\, , u_{nj})'$ is a $n \times 1$ vector of random disturbances. The system (2.1) can be rewritten in the stacked form

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

where

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

so X is a $(np) \times k$ matrix, y and u each have dimension $(np) \times 1$ and β has dimension $k \times 1$, with $k = \sum_{j=1}^{p} k_j$. Let us also set:

$$U = \left[\; u_1\, ,\, u_2\, ,\, \ldots\, ,\, u_p \;\right] = \begin{bmatrix} U'_{1\cdot} \\ U'_{2\cdot} \\ \vdots \\ U'_{n\cdot} \end{bmatrix} \tag{2.4}$$

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

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

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

and

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

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 to an unknown linear transformation that can modify the scaling and dependence properties of the disturbances in the different equations. Note (2.5) - (2.7) do not necessarily entail that Σ is the covariance matrix of $U_{t\cdot}$, because the distribution of $W_1,\, \ldots\, ,\, W_n$ is not restricted (*e.g.*, it may not have finite second moments). 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 the stronger assumption

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

we have $E\big(U_{t\cdot}U_{t\cdot}'\big) = \Sigma\,,\ t = 1, \ldots , n$, and

$$E(uu') = \Sigma \otimes I_p . \tag{2.11}$$

Assumption (2.10) yields the gaussian SURE model. For further reference, we shall write

$$W = [W_1, \ldots , W_n]' = UJ^{-1}. \tag{2.12}$$

In this paper, we consider the problem of testing general hypotheses of the form:

$$H_0 : A\beta \in \Delta_0 \tag{2.13}$$

where A is a full row-rank $v_0 \times k$ matrix and Δ_0 is a non-empty subset of $\mathbb{R}^{v_0}$.

For our subsequent arguments, it will be important to spell out the relation between SURE and MLR models. The MLR model may be defined as a SURE model where the regressors in all the equations are the same $(X_1 = X_2 = \cdots = X_p)$. Conversely, a SURE model can be viewed as a restricted MLR system. To be more specific, for each $1 \le j \le p$ in the context of (2.1), let $\overline{X}_j$ be any matrix such that the columns of $[X_j, \overline{X}_j]$ are linearly independent and span the same space as the columns of the matrix $[X_1,\ X_2,\ \ldots ,\ X_p]$. In most practical situations, $\overline{X}_j$ will simply contain the regressors from the matrices $X_k,\ k \neq j$, which are excluded from the *j*-th equation. Further, let X_* be any full-column rank $n \times k_*$ matrix whose columns span the same space as those of $[X_1,\ X_2,\ \ldots ,\ X_p]$, i.e.

$$\text{sp}(X_*) = \text{sp}([X_1,\ X_2,\ \ldots ,\ X_p])\,,\ \det(X_*'X_*) \neq 0\,,\ \text{rank}(X_*) = k_*\,, \tag{2.14}$$

where, for any matrix Z, sp(Z) represents the vector space spanned by the columns of Z. Then, for each j, we can find matrices S_j and $\overline{S}_j$ of dimensions $k_* \times k_j$ and $k_* \times (k_* - k_j)$ respectively such that $X_*S_j = X_j,\ X_*\overline{S}_j = \overline{X}_j$ and the matrix $T_j = [S_j,\ \overline{S}_j]$ is invertible. Consequently, (2.1) may be rewritten as

$$Y_j = X_*S_j\beta_j + X_*\overline{S}_j\overline{\beta}_j + u_j = X_*\beta_{*j} + u_j\ ,\ j = 1, \ldots , p\,, \tag{2.15}$$

where $\beta_{*j} = S_j\beta_j + \overline{S}_j\overline{\beta}_j = T_j(\beta_j',\ \overline{\beta}_j')'$, with the restrictions

$$\overline{\beta}_j = 0,\ j = 1, \ldots , p\,. \tag{2.16}$$

The latter restrictions may also be expressed in implicit form on $\beta_{*j},\ j = 1, \ldots , p$, as:

$$M(S_j)\beta_{*j} = 0,\ j = 1, \ldots , p\,, \tag{2.17}$$

where $M(S_j) = [I_{k_*} - S_j(S_j'S_j)^{-1}S_j']$. Clearly (2.15) - (2.16) define a constrained MLR model.

On relaxing the SURE restrictions (2.17), this MLR model can be put in the stacked form:

$$y = (I_p \otimes X_*)\beta_* + u = \overline{X}_* \beta_* + u \tag{2.18}$$

where $\overline{X}_* = I_p \otimes X_*$ and $\beta_* = \left(\beta'_{*1},\, \beta'_{*2},\, \ldots\,,\, \beta'_{*p}\right)'$, or equivalently,

$$Y = X_* B_* + U \tag{2.19}$$

where $Y = [Y_1,\, Y_2,\, \ldots\,,\, Y_p]$ and $B_* = [\beta_{*1},\, \beta_{*2},\, \ldots\,,\, \beta_{*p}]$. We shall call the unrestricted MLR model (2.18) [or (2.19)] an *embedding* MLR model for the SURE model (2.1). It is clear any hypothesis on β can be expressed equivalently in terms of β_* within the corresponding embedding MLR model.

In this paper, we shall emphasize LR-type tests of H_0 derived under the gaussian distributional assumptions (2.5) - (2.10). In this case, the log-likelihood function associated with the SURE model (2.2) has the form:

$$\mathcal{L}(\beta,\, \Sigma) = -\frac{np}{2}\ln(2\pi) - \frac{n}{2}\ln(|\Sigma|) - \frac{1}{2}(y - X\beta)'(I_n \otimes \Sigma)^{-1}(y - X\beta)\,. \tag{2.20}$$

Then, provided the relevant optima do exist and are unique, the "unconstrained" maximized value of $\mathcal{L}(\beta,\, \Sigma)$ can be written

$$L(H_{\mathcal{S}}) = \sup\{\mathcal{L}(\beta,\, \Sigma) : \beta \in \mathbb{R}^k \text{ and } \Sigma \text{ is p.d.}\} = -\frac{np}{2}[\ln(2\pi) + 1] - \frac{n}{2}\ln\left(|\widehat{\Sigma}_{\mathcal{S}}|\right)\ , \tag{2.21}$$

while its constrained maximized value subject to H_0 is

$$L(H_0) = \sup\{\mathcal{L}(\beta,\, \Sigma) : A\beta \in \Delta_0 \text{ and } \Sigma \text{ is p.d.}\} = -\frac{np}{2}[\ln(2\pi) + 1] - \frac{n}{2}\ln\left(|\widehat{\Sigma}_0|\right) \tag{2.22}$$

where $\widehat{\Sigma}_0$ and $\widehat{\Sigma}_{\mathcal{S}}$ are the restricted and unrestricted ML estimates of Σ, assuming Σ is positive definite (p.d.). Thus, the gaussian LR statistic for testing H_0 against the unrestricted SURE model [or, equivalently, against the embedding MLR model (2.18) with the restrictions (2.17)] is given by:

$$LR(H_0) = 2[L(H_{\mathcal{S}}) - L(H_0)] = n\ln(\Lambda_{\mathcal{S}})\ , \quad \Lambda_{\mathcal{S}} = |\widehat{\Sigma}_0|/|\widehat{\Sigma}_{\mathcal{S}}|\,. \tag{2.23}$$

Note the exclusion SURE restrictions are imposed under both the null and the alternative hypotheses. In the statistics literature, $\Lambda_{\mathcal{S}}^{-1}$ is known as the Wilks criterion.

Similarly, the log-likelihood function associated with the MLR model (2.18), taken jointly with the gaussian distributional assumptions (2.5) - (2.10), is:

$$\begin{aligned} \mathcal{L}_*(\beta_*,\, \Sigma) &= -\frac{np}{2}\ln(2\pi) - \frac{n}{2}\ln(|\Sigma|) - \frac{1}{2}(y - \overline{X}_*\beta_*)'(I_n \otimes \Sigma)^{-1}(y - \overline{X}_*\beta_*) \\ &= -\frac{np}{2}\ln(2\pi) - \frac{n}{2}\ln(|\Sigma|) - \frac{1}{2}\mathrm{tr}[\Sigma(Y - X_*B_*)(Y - X_*B_*)'] \end{aligned} \tag{2.24}$$

where $\beta_* = vec(B_*)$. It is clear

$$\mathcal{L}_*(\beta_*,\, \Sigma) = \mathcal{L}(\beta,\, \Sigma)\,, \text{ when } \overline{\beta}_j = 0,\; j = 1,\, \ldots\,,\, p\,. \tag{2.25}$$

Let

$$\widehat{B}_* = (X_*'X_*)^{-1}X_*'Y\;, \quad \widehat{U} = Y - X_*\widehat{B}_* = M(X_*)U\;, \quad \widehat{\Sigma}_{\mathcal{M}} = \frac{1}{n}\widehat{U}'\widehat{U}\;, \tag{2.26}$$

where $M(X_*) \equiv I_n - X_*(X_*'X_*)X_*'$. Then, provided $\widehat{U}$ has full column rank (which requires $n \geq k_* + p$), $\mathcal{L}_*(\beta_*,\, \Sigma)$ attains a unique (unconstrained) finite maximum at $B_* = \widehat{B}_*$ and $\Sigma = \widehat{\Sigma}_{\mathcal{M}}$ [see Anderson (1984, Chapter 3)], yielding the maximal value:

$$\begin{aligned} L(H_{\mathcal{M}}) &= \sup\{\mathcal{L}_*(\beta_*,\, \Sigma) : \beta_* \in \mathbb{R}^{k_*} \text{ and } \Sigma \text{ is p.d.}\} \\ &= -\frac{np}{2}[\ln(2\pi) + 1] - \frac{n}{2}\ln\left(|\widehat{\Sigma}_{\mathcal{M}}|\right)\,. \end{aligned} \tag{2.27}$$

Thus $\widehat{B}_*$ and $\widehat{\Sigma}_{\mathcal{M}}$ are the ML estimators of the parameters of the unrestricted embedding MLR model associated with the SURE model (2.2). Given the assumptions (2.5) - (2.10), a necessary and sufficient condition for $\widehat{U}$ to have full column rank with probability one is:

$$P\big[\text{rank}\,[M(X_*)W] = p\,\big] = 1\,.$$

The latter will hold, for example, if $\text{rank}(X_*) = k_* \leq n - p$ and $vec(W)$ follows an absolutely continuous distribution on $\mathbb{R}^{np}$. In view of the relation between the SURE model and an embedding MLR model, we shall also consider the LR statistic for testing H_0 against the completely unrestricted MLR model [(2.18) without the SURE restrictions (2.17)]:

$$LR_{\mathcal{M}}(H_0) = 2[L(H_{\mathcal{M}}) - L(H_0)] = n\ln(\Lambda_{\mathcal{M}})\;, \quad \Lambda_{\mathcal{M}} = |\widehat{\Sigma}_0|/|\widehat{\Sigma}_{\mathcal{M}}|\;. \tag{2.28}$$

3. Test procedures

We will now show how one can obtain finite-sample LR-based tests in the context of SURE models as defined above. For that purpose, we shall exploit special features of so-called *uniform linear* restrictions for which LR test statistics have nuisance-parameter null distributions in the context of MLR models (which entails they are pivotal statistics under the null hypothesis). In the MLR case (where $X_1 = \cdots = X_p \equiv X_*$ and $k_1 = \cdots = k_p \equiv k_*$), the hypothesis H_0 in (2.13) is uniform linear if can be expressed in the form

$$H_{\mathcal{UL}} : \left(C' \otimes R\right)\beta = vec(D_0) \tag{3.1}$$

or, equivalently,

$$H_{\mathcal{UL}} : RBC = D_0 \tag{3.2}$$

where B is the $k_* \times p$ matrix such that $\beta = vec(B)$, R is a known $r \times k_*$ matrix of rank r, C is a known $p \times c$ matrix of rank c, and D_0 is a known $r \times c$ matrix. In Dufour and Khalaf (1998*b*), it is shown that the null distribution of the gaussian LR statistic [derived under the assumptions (2.5) - (2.10)] for testing $H_{\mathcal{UL}}$ (against the unrestricted MLR model) does not involve any nuisance parameter under the weaker assumptions (2.1) - (2.8) _ which allow for non-normal disturbances _ and may easily be simulated. In particular, the parameters of the covariance matrix $\Sigma = JJ'$ do not appear in the distribution. Beyond this specific hypothesis class, it is well known that the LR statistic is not pivotal, even if the null hypothesis is linear. For further discussion of uniform linear hypotheses in MLR models, the reader may consult Berndt and Savin (1977), Stewart (1997) and Dufour and Khalaf (1998*b*).

Let us now turn to the SURE model. In the context of the embedding MLR model (2.18), the null hypothesis H_0 in (2.13) is equivalent to the conjunction of $A\beta \in \Delta_0$ with the SURE restrictions (2.16):

$$H_0^* : AF\beta_* \in \Delta_0 \text{ and } M(S_j)\beta_{*j} = 0,\ j = 1,\ \ldots\ ,\ p\ , \tag{3.3}$$

where β_*, S_j, $M(S_j)$ and k_* are defined as in (2.14) - (2.18), and

$$F = \begin{bmatrix} (S_1'S_1)^{-1}S_1' & 0 & \cdots & 0 \\ 0 & (S_2'S_2)^{-1}S_2' & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & (S_p'S_p)^{-1}S_p' \end{bmatrix} . \tag{3.4}$$

It is clear we can find a full row-rank matrix A_* of dimension $v_{0*} \times (pk_*)$ such that H_0^* can be reexpressed in terms of β_* according to a form similar to H_0 in (2.13):

$$H_0^* : A_*\beta_* \in \Delta_{0*} \tag{3.5}$$

where $v_{0*} \geq v_0$ and Δ_{0*} is a non-empty subset of $\mathbb{R}^{v_{0*}}$.

We now state our main result on the distribution of LR statistics in SURE models.

Theorem 3.1 BOUND ON LR STATISTICS IN SURE MODELS. *Suppose the assumptions* (2.1) – (2.8) *hold, with*

$$P\big[rank\,[M(X_*)W] = p\big] = 1 \tag{3.6}$$

where X_* is defined as in (2.14)*, and let $H_{\mathcal{UL}}^* : RB_*C = D_0$ be a set of uniform linear restrictions on* (2.19) *such that $H_{\mathcal{UL}}^*$ entails H_0^*, where the matrices R, B_*, C, D_0 and the hypothesis H_0^* are defined as in* (3.2) *and* (3.3)*. Then the following inequalities hold:*

$$\Lambda_{\mathcal{S}} \leq \Lambda_{\mathcal{M}} \leq \Lambda_{\mathcal{UL}} \tag{3.7}$$

where $\Lambda_{\mathcal{S}}$ and $\Lambda_{\mathcal{M}}$ are defined as in (2.23) *and* (2.28)*, $\Lambda_{\mathcal{UL}} = |\widehat{\Sigma}_{\mathcal{UL}}|/|\widehat{\Sigma}_{\mathcal{M}}|$ and $\widehat{\Sigma}_{\mathcal{UL}}$ is a ML estimator of Σ obtained under the uniform linear restrictions $H_{\mathcal{UL}}^*$. Furthermore, under $H_{\mathcal{UL}}^*$, the distribution of $\Lambda_{\mathcal{UL}}$ (conditional on X) does not depend on the unknown parameter matrices B and Σ nor on the values of the constants in D_0.*

PROOF. The proof is based on observing that the hypotheses involved in the definitions of the statistics $\Lambda_{\mathcal{S}}$, $\Lambda_{\mathcal{M}}$ and $\Lambda_{\mathcal{UL}}$ can be viewed as special cases of the embedding MLR model (2.19). First we note that, under the assumptions (2.1)-(2.8) and (3.6), the log-likelihood $\mathcal{L}_*(\beta_*, \Sigma)$ has (with probability one) a unique maximum given by $L(H_{\mathcal{M}})$ in (2.27). This entails that the supremum of $\mathcal{L}_*(\beta_*, \Sigma)$ under any set of restrictions on β_* must be finite. In particular, the supremum (with respect to β_* and Σ) of $\mathcal{L}_*(\beta_*, \Sigma)$ under $H^*_{\mathcal{UL}}$ can be written

$$\begin{aligned} L(H^*_{\mathcal{UL}}) &= \sup\{\mathcal{L}_*(\beta_*, \Sigma) : RB_*C = D_0 \text{ and } \Sigma \text{ is p.d.}\} \\ &= -\frac{np}{2}[\ln(2\pi) + 1] - \frac{n}{2}\ln\left(|\widehat{\Sigma}_{\mathcal{UL}}|\right) \end{aligned} \tag{3.8}$$

where $\widehat{\Sigma}_{\mathcal{UL}}$ is the ML estimator of Σ under $H^*_{\mathcal{UL}}$. Thus the gaussian LR statistic for testing $H^*_{\mathcal{UL}}$ against the embedding MLR model [i.e., (2.18) jointly with (2.5)-(2.10)] is:

$$LR_{\mathcal{M}}(H^*_{\mathcal{UL}}) = 2[L(H_{\mathcal{M}}) - L(H^*_{\mathcal{UL}})] = n\ln(\Lambda_{\mathcal{UL}}) \,, \quad \Lambda_{\mathcal{UL}} = |\widehat{\Sigma}_{\mathcal{UL}}|/|\widehat{\Sigma}_{\mathcal{M}}| \,. \tag{3.9}$$

From Theorem 3.1 in Dufour and Khalaf (1998*b*), the exact distribution of $\Lambda_{\mathcal{UL}}$ under $H^*_{\mathcal{UL}}$ only depends on the distribution of W and the known matrices X, R and C _ but not on D_0 nor on the otherwise unknown parameters in B_* and Σ.

Now, by the definition of the embedding MLR model, (2.25) and the equivalence between H_0 and H_0^*, we see that $L(H_{\mathcal{S}})$ and $L(H_0)$ in (2.21)-(2.22) can also be expressed in terms of $\mathcal{L}_*(\beta_*, \Sigma)$:

$$L(H_{\mathcal{S}}) = \sup\{\mathcal{L}_*(\beta_*, \Sigma) : M(S_j)\beta_{*j} = 0,\ j = 1, \ldots, p, \text{ and } \Sigma \text{ is p.d.}\} \,, \tag{3.10}$$

$$L(H_0) = \sup\{\mathcal{L}_*(\beta_*, \Sigma) : \beta_* \text{ satisfies } H_0^* \text{ and } \Sigma \text{ is p.d.}\} \,. \tag{3.11}$$

Since $H^*_{\mathcal{UL}}$ entails H_0^*, which in turn is a restricted form of the SURE model $(H_{\mathcal{S}})$, and since $H_{\mathcal{S}}$ can be obtained by imposing linear restrictions on the embedding MLR model $(H_{\mathcal{M}})$, it follows that

$$L(H^*_{\mathcal{UL}}) \leq L(H_0) \leq L(H_{\mathcal{S}}) \leq L(H_{\mathcal{M}}) \,, \tag{3.12}$$

hence

$$L(H_{\mathcal{S}}) - L(H_0) \leq L(H_{\mathcal{M}}) - L(H_0) \leq L(H_{\mathcal{M}}) - L(H^*_{\mathcal{UL}}) \tag{3.13}$$

and

$$\Lambda_{\mathcal{S}} \leq \Lambda_{\mathcal{M}} \leq \Lambda_{\mathcal{UL}} \,. \tag{3.14}$$

This completes the proof of the theorem. □

It follows from the latter theorem that

$$P[\Lambda_{\mathcal{S}} \geq \lambda_{\mathcal{UL}}(\alpha)] \leq P[\Lambda_{\mathcal{M}} \geq \lambda_{\mathcal{UL}}(\alpha)] \leq \alpha \tag{3.15}$$

under H_0, where $\lambda_{\mathcal{UL}}(\alpha)$ is determined such that $P[\Lambda_{\mathcal{UL}} \geq \lambda_{\mathcal{UL}}(\alpha)] = \alpha$ (or, at least, $P[\Lambda_{\mathcal{UL}} \geq \lambda_{\mathcal{UL}}(\alpha)] \leq \alpha$) and $0 \leq \alpha \leq 1$. It is important to note here that the inequality (3.7) holds for *any* set of uniform linear restrictions which entails H_0. In particular, on taking $R = I_{k_*}$, $C = I$ and $D_0 = B_*$ (the true value of B_*), it is clear such a set of restrictions does always exist, although other choices (for R, C and D_0) may be available in view of the form of H_0 and lead to a tighter bound. Further, the distribution of $\Lambda_{\mathcal{UL}}$ (when $RB_*C = D_0$) only depends on X, R and C, but not on D_0, so one can use any possible value of D_0 (such as $D_0 = 0$) in order to compute or simulate this distribution.

Using well known results from Anderson (1984) and Rao (1973) for gaussian MLR models, it is also possible to show that the bounding statistic $\Lambda_{\mathcal{UL}}$ is distributed like the product of the inverse of p beta variables with degrees of freedom which depend only on r, c, p and k_* [see Dufour (1997) and Dufour and Khalaf (1998*b*)]. The latter result is, however, hardly useful for practical applications of the proposed bound. Hence we do not restate our conclusions here for this specific gaussian case: it is more convenient to derive $\lambda_{\mathcal{UL}}(\alpha)$ by simulation as shown below and using Theorem **3.1**, under any distributional assumptions that satisfy (2.5) including the normal case. Finally, we note that the same bound applies to both criteria $\Lambda_{\mathcal{S}}$ and $\Lambda_{\mathcal{M}}$. Since $\Lambda_{\mathcal{M}} \geq \Lambda_{\mathcal{S}}$, it will thus be preferable to apply the bound to $\Lambda_{\mathcal{U}}$ rather than $\Lambda_{\mathcal{S}}$, since this will yield a more powerful test.

Theorem **3.1** has further implications for LR-based hypothesis tests. The fact that the null distribution of the LR statistic can be bounded (in a non trivial way) by a statistic whose distribution can be simulated fairly easily entails that MC test techniques may be used to obtain finite-sample *p*-values based on the LR-based statistics when the bounds test is not conclusive. In earlier work, we have discussed in detail how such procedures can be implemented; see Dufour (1998), Dufour, Farhat, Gardiol and Khalaf (1998), Dufour and Kiviet (1996, 1998), and Dufour and Khalaf (1998*b*, 2001). These include techniques for the construction of: (i) a (parametric) bootstrap-type *p*-value which we denote a *local* Monte Carlo *p*-value (LMC) to account for the fact that the underlying simulation routine is implemented given a specific nuisance parameter estimate, and (ii) an exact randomized *p*-value which corresponds to the largest MC *p*-value over the relevant nuisance parameter space; conformably, we call the latter a *maximized* Monte Carlo (MMC) *p*-value [Dufour (1998)]. Both procedures are summarize below. Although the LMC *p*-value is only valid asymptotically, non-rejections are conclusive from a finite-sample perspective, in the following sense. Indeed, for all $0 \leq \alpha \leq 1$, if the LMC *p*-value exceeds α, we can be sure that the maximum p-value also exceeds α. We emphasize the fact that the MMC test can be implemented in complementarity with the above defined bounds tests. Indeed, if the BMC test rejects the null then the MMC test is certainly significant. For a more detailed discussion of the justification and implementation of such simulation-based procedures, we refer the reader to the papers just cited.

To illustrate how the above results may be used in the context of a SURE model, we will now discuss an illustrative example.

Example 3.2 THREE-EQUATION SURE MODEL. In the SURE model (2.1), with gaussian errors and $p = 3$, $k_i = 2$, $X_i = [\iota_n,\ x_i]$ where ι_n denotes a vector of n 1's, consider the problem of testing

$$H_0 : \beta_{11} = \beta_{22} = \beta_{33} \ . \tag{3.16}$$

We suppose also that the matrix $X_* = [\iota_n,\ x_1,\ x_2,\ x_3]$ has full column rank $k_* = 4 \leq n - 3$. Then, it is easy to see that this problem is equivalent to testing

$$H_0^* : b_{11} = b_{22} = b_{33} \text{ and } b_{12} = b_{13} = b_{21} = b_{23} = b_{31} = b_{32} = 0$$

in the framework of the MLR model

$$Y = X_* B_* + U \tag{3.17}$$

with $Y = [Y_1,\ Y_2\ ,\ Y_3]$, $U = [U_1,\ U_2,\ U_3]$, $B_* = [b_1,\ b_2,\ b_3]$, and $b_j = (b_{0j},\ b_{1j}, b_{2j}, b_{3j})$, $j = 1,\ 2,\ 3$. In order to use the above results on the conservative bound, we need to construct a set of UL restrictions on the coefficients of the later MLR model that satisfy the hypothesis in question. It is easy to see that constraints setting the values of the coefficients $b_{ij},\ i,\ j = 1\ , \ldots ,\ 3$ to specific values (9 restrictions on 12 coefficients) meet this purpose:

$$H_{\mathcal{UL}}^* : \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} B_* = D_0\,. \tag{3.18}$$

All that remains is to calculate the LR, as defined in (2.23), and use the critical value associated with the uniform linear restriction (3.18).

A bounds MC test may then be applied as described in Dufour and Khalaf (1998*b*) for testing general possibly non-linear restrictions in MLR models. The procedure can be described as follows for the example just considered. For further reference, and conformably with Dufour and Khalaf (1998*b*), we call the latter bounds test a BMC test.

1. Denote $\Lambda^{(0)}$ the observed test statistic, which could be the statistic $\Lambda_{\mathcal{S}}$ in (2.23) or $\Lambda_{\mathcal{M}}$ in (2.28).

2. By Monte Carlo methods, draw N simulated samples (conditional on the right-hand-side regressors) from model (2.18) in $H_{\mathcal{UL}}^*$. For instance, in the above example **3.2**, one way to do this is to draw the simulated samples from the base model (3.17) with parameters set to their constrained (imposing H_0^*) SURE estimates on which the additional restrictions underlying $H_{\mathcal{UL}}^*$ have been imposed (simply setting to fixed values the relevant coefficients of B_*). To be more specific, let us denote by $\widetilde{b}_{ij}$ the SURE estimate of β_{ij}, $0 \leq i \leq 3$, $1 \leq j \leq 3$, several of which should be zero to account for the SURE exclusion restrictions). Then, if we choose to draw from model (3.17) with coefficients $\widetilde{b}_{ij}, i = 0,\ \ldots\ ,\ 3,\ j = 1\ , \ldots ,\ 3$, and the conformable covariance matrix estimate, then $H_{\mathcal{UL}}^*$ could be of the form: $b_{ij} = \widetilde{b}_{ij}$, $i,\ j = 1\ , \ldots ,\ 3$.[1] Furthermore, any error distribution that satisfies (2.5)-(2.8) may be considered.

[1] It is also possible _ and perhaps more efficient _ to rewrite the bounding LR statistic as a pivotal quantity, using the results in Dufour and Khalaf (1998*b*). The procedure just presented exploits the pivotal property of the statistic *implicitly*. Yet it is quite intuitive and relates to the familiar parametric bootstrap.

3. From each simulated sample, compute the bounding statistic $\Lambda_{\mathcal{UL}}$ which corresponds to the LR-based statistic associated with $H^*_{\mathcal{UL}}$: this yields $\Lambda^{(h)}_{\mathcal{UL}},\ h = 1\,,\ \ldots\ ,\ N$. As emphasized above, it is important to make sure that the regression coefficients selected to generate the Monte Carlo drawings correspond to the restrictions implied by $H^*_{\mathcal{UL}}$. In other words, the simulated values of the bounding statistic should satisfy the null hypothesis.

4. Compute the simulated $p-$value $\widehat{p}_N(\Lambda^{(0)})$ where

$$\widehat{p}_N(x) = \left\{ 1 + \sum_{h=1}^{N} I[\Lambda^{(h)}_{\mathcal{UL}} - x] \right\} /(N+1)\ , \tag{3.19}$$

   $I[z] = 1$ if $z \geq 0$ and $I[z] = 0$ if $z < 0$.

5. The procedure rejects at level α if $\widehat{p}_N(\Lambda^{(0)}) \leq \alpha$.

Using the same arguments as in Dufour (1998), Dufour and Kiviet (1996, 1998)and Dufour and Khalaf (1998*b*), it is easy to see that

$$P\left[\widehat{p}_N(\Lambda^{(0)}) \leq \alpha\right] \leq \alpha \text{ under } H_0\,,$$

so the critical region $\widehat{p}_N(\Lambda^{(0)}) \leq \alpha$ has level α. If the above procedure is implemented replacing $\Lambda^{(h)}_{\mathcal{UL}},\ j = 1\,,\ \ldots\ ,\ N$ with $\widehat{\Lambda}^{(h)},\ h = 1\,,\ \ldots\ ,\ N$, which refer to realized values of the LR criterion associated with H^*_0 and the simulated samples, then (3.19) yields a parametric bootstrap or an LMC *p*-value. In this case, the *p*-value in question depends on the choice of the intervening nuisance parameters. The MMC p-value corresponds to the largest MC *p*-value overall nuisance parameters compatible with the null hypothesis. A global optimizing algorithm is required to maximize the MC p-value. In this paper, we used simulated annealing [see Goffe, Ferrier and Rogers (1994)]. The reader may consult Dufour (1998), Dufour et al. (1998), Dufour and Kiviet (1996, 1998), and Dufour and Khalaf (2001) for further discussion of MC tests in econometrics.

4. Simulation study

In this section, we present simulation results illustrating the performance of the above proposed procedures along with the one of more traditional asymptotically justified methods. In particular, we examine the performance of BMC and MMC tests in SURE contexts. In the linear case, we also consider LMC tests based on standard Wald-type criteria and several alternative statistics justified on the basis of computational cost as opposed to those relying on full maximum likelihood estimation.

4.1. Design

We studied two gaussian SURE designs, similar to Example **3.2**. In the first one (D1), we considered the problem of testing linear cross-equation constraints, while in the second one (D2), we studied a nonlinear constraint.

D1. **SURE system, cross-equation constraints**
Model (2.1); $k_j = 2\ ,\ j = 1\ ,\ \ldots\ ,\ p\,;\ p = 3, 5\ ;\ n = 25\ ;$
$H_0 : \beta_{jj} = \beta_{11}\ ,\ j = 2\ ,\ \ldots\ ,\ p\ .$

D2. **SURE system, non-linear constraints**
Model (2.1); $k_j = 3\ ,\ j = 1\ ,\ \ldots\ ,\ p\,;\ p = 7\ ;\ n = 25\,;$
the regressor which corresponds to β_{1j} is common to all equations;
$H_0 : \beta_{1j} = \gamma\beta_{2j}\ ,\ j = 1\ ,\ \ldots\ ,\ p\ ,\ \gamma$ unknown.

For each model, a constant regressor was included and the other regressors were independently drawn (once) from a normal distribution; the errors were independently generated as *i.i.d.* $N(0, \Sigma)$ with $\Sigma = JJ'$ and the elements of J drawn (once) from a normal distribution. The coefficients are reported in Table 1.

The statistics examined for D1 include the relevant LR criteria defined by (2.23) and (2.28), as well as three other types of statistics: (1) quasi-LR statistics based on incompletely maximized likelihood functions; (2) test statistics similar to those suggested by Theil et al. (1985); (3) a number of Wald-type criteria. To be more precise, the latter are defined as follows.

1. The quasi-LR (QLR) statistics are:

$$QLR_{(l)} = n \ln(\Lambda_{(l)})\ ,\ \Lambda_{(l)} = |\widetilde{\Sigma}_{0(l)}|/|\widetilde{\Sigma}_{(l)}|\ , \tag{4.1}$$

   where $\widetilde{\Sigma}_{0(l)}$ and $\widetilde{\Sigma}_{(l)}$ denote the constrained and unconstrained iterative estimators of Σ and the subscript l refers to the number of iterations involved. Though we did not analytically establish the asymptotic distribution of the latter criteria, we assessed their asymptotic significance using the χ^2 reference distribution for the usual LR statistic. We append the subscript LMC to the notation for the QLR test to refer to the corresponding LMC test.

2. The test statistics suggested by Theil et al. (1985) may be interpreted as unscaled Wald-type statistics, whose level is controlled by a Monte Carlo (or bootstrap) method. We consider these here mainly for historical reasons, because they are really the first simulation-based test procedures proposed in the SURE setup. For the model with three equations, we considered:

$$\begin{aligned} \mu_{31} &= |\widehat{\beta}_{11} - \widehat{\beta}_{22}| + |\widehat{\beta}_{22} - \widehat{\beta}_{33}| \quad , \\ \mu_{32} &= |\widehat{\beta}_{11} - \widehat{\beta}_{33}| + |\widehat{\beta}_{22} - \widehat{\beta}_{33}| \quad , \\ \mu_{33} &= |\widehat{\beta}_{11} - \widehat{\beta}_{22}| + |\widehat{\beta}_{11} - \widehat{\beta}_{33}| \quad . \end{aligned}$$

   In the five-equation case, the following were selected among many possible choices:

$$\begin{aligned} \mu_{51} &= |\widehat{\beta}_{11} - \widehat{\beta}_{22}| + |\widehat{\beta}_{22} - \widehat{\beta}_{33}| + |\widehat{\beta}_{33} - \widehat{\beta}_{44}| + |\widehat{\beta}_{44} - \widehat{\beta}_{55}| \quad , \\ \mu_{52} &= |\widehat{\beta}_{22} - \widehat{\beta}_{33}| + |\widehat{\beta}_{33} - \widehat{\beta}_{44}| + |\widehat{\beta}_{44} - \widehat{\beta}_{55}| + |\widehat{\beta}_{55} - \widehat{\beta}_{11}| \quad , \\ \mu_{53} &= |\widehat{\beta}_{33} - \widehat{\beta}_{44}| + |\widehat{\beta}_{44} - \widehat{\beta}_{55}| + |\widehat{\beta}_{55} - \widehat{\beta}_{11}| + |\widehat{\beta}_{11} - \widehat{\beta}_{22}| \quad , \\ \mu_{54} &= |\widehat{\beta}_{44} - \widehat{\beta}_{55}| + |\widehat{\beta}_{55} - \widehat{\beta}_{11}| + |\widehat{\beta}_{11} - \widehat{\beta}_{22}| + |\widehat{\beta}_{22} - \widehat{\beta}_{33}| \quad , \\ \mu_{55} &= |\widehat{\beta}_{55} - \widehat{\beta}_{11}| + |\widehat{\beta}_{11} - \widehat{\beta}_{22}| + |\widehat{\beta}_{22} - \widehat{\beta}_{33}| + |\widehat{\beta}_{33} - \widehat{\beta}_{44}| \quad . \end{aligned}$$

For the purpose of this experiment, we used for $\widehat{\beta}$ the gaussian ML estimator of β.

3. The Wald-type criteria are based on feasible generalized least squares (GLS) parameter estimates. Specifically, we considered the statistic suggested in Srivastava and Giles (1987, Chapter 10) for an hypothesis of the form $A\beta - r = 0$:

$$\mathcal{W} = \left(\frac{v_1}{v_0}\right) \frac{\left(A\widehat{\beta} - r\right)' \left[A\left(X' \left(S^{-1} \otimes I_n\right) X\right)^{-1} A'\right]^{-1} \left(A\widehat{\beta} - r\right)}{(y - X\widehat{\beta})' \left(S^{-1} \otimes I_n\right) (y - X\widehat{\beta})} \tag{4.2}$$

where A is a $\nu_0 \times k$ full-row rank fixed matrix, $v_1 = np - k$, $k = \sum_{j=1}^{p} k_j$, while S and $\widehat{\beta}$ are feasible generalized least squares parameter estimates.[2] Under the null hypothesis (and standard regularity conditions), $v_0\mathcal{W}$ has a $\chi^2(v_0)$ asymptotic distribution. Theil (1971, Chapter 6) suggests that the $F(v_0, v_1)$ distribution better captures the finite-sample distribution of the statistics. Yet this claim is not supported by either analytical or simulation evidence. Maximum likelihood estimators may also be substituted for $\widehat{\beta}$ and S in the formulae for the Wald criterion. Here we have considered both the standard feasible estimator of β [using the estimate of S based on least squares residuals applied to each one of the regressions in (2.1)] as well as the ML estimators (iterated to convergence) of β and Σ. For further reference, the GLS and ML based $\mathcal{W}$ tests will be denoted $\mathcal{W}[GLS,\, j]$ and $\mathcal{W}[ML,\, j]$ respectively, where $j \in \{\chi^2\,,\, F\}$ indicates whether the critical value was obtained from the χ^2 asymptotic distribution or from the $F(v_0, v_1)$ distribution. The local MC (parametric bootstrap) counterparts will be denoted $\mathcal{W}[GLS,\, LMC]$ and $\mathcal{W}[ML,\, LMC]$.

In both D1 and D2, we computed empirical frequencies of type I errors, based on a nominal size of 5% and 1000 replications. In D1, the powers of the tests were investigated by simulating the model with the same parameter values except for β_{11} respectively.[3] The LMC and BMC tests were applied with 19 and 99 replications. Because of the computational cost involved, the MMC test was only applied with $p = 3$ and $N = 19$. The BMC test was performed based on the bounding statistic as described in Example **3.2**. For each test statistic, the LMC randomized procedure was based on simulations that use a restricted estimator similar to the estimator(s) involved in the corresponding test statistic: a restricted ML (or quasi-ML) estimator for LR or Wald-type tests based on ML (quasi-ML) estimators, restricted feasible GLS estimators for tests based on GLS estimators. All the experiments were conducted using Gauss-386i VM version 3.1.

4.2. Results and discussion

The results of the limited size-study in D2 reveal the following: the observed empirical frequency of type I errors for the LR statistic was 12.5% whereas the one of the bounds test (2.6%) satisfied the 5% level constraint. The results of experiment D1 are summarized in Tables 2 to 5. The subscripts asy, BMC, LMC and MMC which appear in these tables refer respectively to the standard asymptotic tests, MC bounds tests, local MC tests (parametric bootstrap), and maximized MC

[2] The statistic $\mathcal{W}$ above corresponds to the z statistic in equation (10.11) of Srivastava and Giles (1987, Chapter 10).

[3] For the purpose of power comparisons, the asymptotic tests were size corrected using an independent simulation.

tests. $LR[asy]$, $LR[BMC]$, $LR[LMC]$ and $LR[MMC]$ refer to the corresponding LR tests, and similarly to $LR_{\mathcal{M}}$ and QLR. Our results show the following.

1. The asymptotic criteria have an upward bias in size; as can be seen in Table 2, rejection of the null is repeatedly many times larger than what it should be. The bias clearly worsens in the 5 equation example (5EQ). Across the cases examined, the Wald-type statistics have larger sizes when based on their asymptotic χ^2 critical values. Although the F approximation seems to correct the problem in the 3EQ model, it clearly fails to do so in the 5EQ case. The non-linear LR test examined in D2 is also over-sized.

2. The BMC test was found to be well behaved. Power gains are possible in other test problems where a tighter critical bound is available. Indeed, we have observed reasonable power even if we have experimented with the worst scenario, in the sense that bounding test statistics correspond to a null hypothesis which fixes the values of all regression coefficients (except the intercept). Furthermore, we found that the BMC and the MMC tests based on LR_M yield equivalent decisions for all cases examined; the MMC test based on LR performs marginally better. This illustrates the value of the conservative bounds test as a tool to be used in conjunction with LMC test methods and not necessarily as an alternative to those methods. As emphasized earlier, the bounds procedure is computationally inexpensive and exact. In addition, whenever the bounds test rejects, inference may be made without further appeal to randomized tests.

3. There is no indication of overrejection for the LMC tests considered. While the critical values used, conditional on the particular choice of consistent estimator for the error covariance matrix, are only asymptotically justified, the procedure was remarkably effective in correcting the bias. Whether this conclusion would carry to larger systems remains an open question. In this regard, note that available simulation evidence on the SURE model, specifically the experiment in Rocke (1989) on *large* systems is limited to three equations at best.

4. While they did exhibit adequate sizes, the statistics inspired by Theil et al. (1985) did not fare well in terms of power. For the 3EQ model, the performance was dramatically poor for μ_{32} and μ_{33} but less so in the case of μ_{31}. Even then, as compared to the randomized LR, the performance is less than satisfactory.

5. The LMC tests performed noticeably well in terms of power in all instances, even when the number of replications was as low as 19. It is worth noting however that simulation evidence does not favor the randomized usual LR tests over those based on $\Lambda_{(l)}$ typically involving fewer iterations, although we are uncertain as to the asymptotic equivalence of both procedures. This observation has an important bearing on empirical practice. The simplicity of the method based on $\Lambda_{(l)}$ has much to recommend it for larger models in which statistics requiring full MLE may be quite expensive to randomize.

5. Empirical illustration

In this section, we present an empirical application that illustrates the results presented in this paper. We consider testing restrictions on the parameters of a generalized Leontief cost function. We used the data from Berndt and Wood (1975) and the factor demand system from Berndt (1991, pp. 460-462). The model imposes constant returns to scale and linear homogeneity in prices, and includes

TABLE 1. COEFFICIENT VALUES USED IN THE SIMULATION EXPERIMENTS

D1.	$\beta_{(3EQ)} = (1.2\,,\ 0.1\,,\ 0.8\,,\ 0.1\,,\ -1.1\,,\ 0.1)'$ $\beta_{(5EQ)} = (1.2\,,\ 0.1\ ,\ 0.8\,,\ 0.1\,,\ -1.1\,,\ 0.1\,,\ 1.9\,,\ 0.1\,,\ -0.2\,,\ 0.1)'$
D2.	$\gamma = .009$ and $\beta_{0j},\ \beta_{2j},\ \ j = 1\ ,\ ...\ ,\ p$, drawn (once) as $i.i.d.\ N(0, .16)$

TABLE 2. EMPIRICAL LEVELS OF VARIOUS TESTS: EXPERIMENT D1

Asymptotic tests			MC tests					
Test	3EQ	5EQ	Test	3EQ	5EQ	Test	3EQ	5EQ
$\mathcal{W}[GLS,\ \chi^2]$	.061	.130	$\mathcal{W}[GLS_{LMC}]$	.049	.047	μ_{31}	.058	-
$\mathcal{W}[ML,\ \chi^2]$	.124	.254	$\mathcal{W}[ML_{LMC}]$	.047	.049	μ_{32}	.051	-
$\mathcal{W}[GLS,\ F]$	.052	.121	$LR[LMC]$	.047	.043	μ_{33}	.055	-
$\mathcal{W}[ML,\ F]$	.111	.242	$LR[MMC]$	.038		μ_{51}	-	.027
$LR[asy]$	.094	.143	$LR_{\mathcal{M}}[MMC]$	.036		μ_{52}	-	.026
$QLR_{(0)}$	.068	.077	$LR[BMC]$	.036	.029	μ_{53}	-	.025
$QLR_{(1)}$	.088	.131	$QLR_{(0)}[LMC]$	.045	.052	μ_{54}	-	.011
$QLR_{(2)}$	.094	.143	$QLR_{(1)}[LMC]$	.048	.052	μ_{55}	-	.025
			$QLR_{(2)}[LMC]$	.047	.044			

TABLE 3. POWER OF THE BOUNDS TESTS: EXPERIMENT D1

$H_0 : \beta_{11} = .1$

		3 equations					5 equations				
N	β_{11}	.3	.5	.7	.9	1	.3	.5	.7	.9	1.0
19	p_1	.065	.383	.791	.963	.987	.082	.416	.792	.958	.995
	p_2	.171	.324	.171	.034	.013	.249	.497	.207	.042	.005
	p_3	.030	.021	.008	0.00	0.00	.038	.011	0.00	0.00	0.00
	p_4	.734	.272	.030	.003	0.00	.631	.076	.001	0.00	0.00
19	p_1	.077	.434	.858	.986	.999	.075	.474	.877	.990	1.0
	p_2	.204	.372	.127	.014	.001	.256	.439	.122	.010	0.00
	p_3	.022	.007	.003	0.00	0.00	.035	.010	0.00	0.00	0.00
	p_4	.697	.187	.012	0.00	0.00	.634	.077	.001	0.00	0.00

Note: p_1 is the empirical probability that $LR[LMC]$ and $LR[BMC]$ reject, p_2 measures the probability that $LR[BMC]$ fails to reject and $LR[LMC]$ rejects, p_3 measures the probability that $LR[BMC]$ rejects and $LR[LMC]$ fails to reject and p_4 is the empirical probability that both tests fail to reject.

TABLE 4. POWER OF VARIOUS TESTS: EXPERIMENT D1, 3 EQUATIONS

$H_0 : \beta_{11} = .1$

	19 replications					99 replications				
β_{11}	.3	.5	.7	.9	1.0	.3	.5	.7	.9	1.0
$\mathcal{W}[GLS,\, \chi^2]$	.192	.647	.939	.993	.999	.192	.647	.939	.993	.999
$\mathcal{W}[ML,\, \chi^2]$	.264	.787	.984	1.0	1.0	.264	.787	.984	1.0	1.0
$LR[asy]$	.281	.806	.985	1.0	1.0	.281	.806	.985	1.0	1.0
$\mathcal{W}[GLS,\, LMC]$	.185	.579	.884	.974	.986	.202	.640	.934	.990	.998
$\mathcal{W}[ML,\, LMC]$	.225	.704	.958	.997	1.00	.260	.774	.985	1.00	1.00
$LR[LMC]$	.236	.707	.962	.997	1.00	.262	.779	.985	1.00	1.00
$QLR_{(0)}$	.227	.689	.950	.993	.988	.256	.762	.977	.997	.999
$QLR_{(1)}$	.238	.709	.961	.997	1.00	.259	.776	.986	1.00	1.00
$QLR_{(2)}$	.236	.707	.962	.997	1.00	.262	.776	.985	1.00	1.00
$LR_M[MMC]$	.095	.404	.799	.963	.987	.099	.441	.861	.986	.999
$LR[MMC]$	.054	.388	.804	.978	.993	-	-	-	-	-
$LR[BMC]$	.095	.404	.799	.963	.987	.099	.441	.861	.986	.999
μ_{31}	.076	.108	.148	.216	.259	.064	.108	.165	.219	.268
μ_{32}	.197	.552	.869	.974	.992	.210	.641	.935	.995	.998
μ_{33}	.093	.183	.307	.432	.489	.088	.184	.328	.503	.601

TABLE 5. POWER OF VARIOUS TESTS: EXPERIMENT D1, 5 EQUATIONS

$H_0 : \beta_{11} = .1$

	19 replications					99 replications				
β_{11}	.3	.5	.7	.9	1.1	.3	.5	.7	.9	1.1
$\mathcal{W}[GLS,\, \chi^2]$	.200	.703	.961	.994	.999	.200	.703	.961	.994	.999
$\mathcal{W}[ML,\, \chi^2]$	.317	.918	1.0	1.0	1.0	.317	.918	1.0	1.0	1.0
LR_{asy}	.331	.913	.999	1.0	1.0	.331	.913	.999	1.0	1.0
$\mathcal{W}[GLS,\, LMC]$	.162	.619	.918	.982	.998	.186	.684	.946	.990	.999
$\mathcal{W}[ML,\, LMC]$	.265	.832	.991	.999	1.00	.297	.903	1.0	1.00	1.00
$LR[LMC]$	.286	.841	.999	.999	1.00	.328	.908	.998	1.00	1.00
$QLR_{(0)}$	.265	.806	.971	.998	1.00	.316	.864	.983	.999	1.00
$QLR_{(1)}$	.290	.849	.988	.998	1.00	.334	.900	.997	1.00	1.00
$QLR_{(2)}$	.287	.842	.991	.999	1.00	.331	.908	.997	1.00	1.00
$LR[BMC]$	.120	.427	.792	.958	.995	.110	.484	.877	.990	1.00
μ_{51}	.029	.034	.038	.041	.048	.032	.036	.039	.041	.044
μ_{52}	.031	.036	.039	.042	.045	.031	.034	.038	.040	.041
μ_{53}	.042	.085	.154	.258	.359	.035	.077	.152	.241	.397
μ_{54}	.023	.071	.159	.289	.456	.025	.067	.175	.302	.512
μ_{55}	.031	.050	.071	.118	.170	.033	.056	.092	.128	.180

four inputs: capital (K), labor (L), energy (E) and non-energy intermediate materials (M). If we denote the output by Y and the input prices P_j, j = K, L, E, M, the stochastic cost minimizing input-output KLEM equations are:

$$\begin{aligned}
\mathtt{K/Y} &= d_{\mathtt{KK}} + d_{\mathtt{KL}}\,(P_{\mathtt{L}}/P_{\mathtt{K}})^{1/2} + d_{\mathtt{KE}}\,(P_{\mathtt{E}}/P_{\mathtt{K}})^{1/2} + d_{\mathtt{KM}}\,(P_{\mathtt{M}}/P_{\mathtt{K}})^{1/2} + e_{\mathtt{K}}\ , && (5.1)\\
\mathtt{L/Y} &= d_{\mathtt{LL}} + d_{\mathtt{LK}}\,(P_{\mathtt{K}}/P_{\mathtt{L}})^{1/2} + d_{\mathtt{LE}}\,(P_{\mathtt{E}}/P_{\mathtt{L}})^{1/2} + d_{\mathtt{LM}}\,(P_{\mathtt{M}}/P_{\mathtt{L}})^{1/2} + e_{\mathtt{L}}\ , && (5.2)\\
\mathtt{E/Y} &= d_{\mathtt{EE}} + d_{\mathtt{EK}}\,(P_{\mathtt{K}}/P_{\mathtt{E}})^{1/2} + d_{\mathtt{EL}}\,(P_{\mathtt{L}}/P_{\mathtt{E}})^{1/2} + d_{\mathtt{EM}}\,(P_{\mathtt{M}}/P_{\mathtt{E}})^{1/2} + e_{\mathtt{E}}\ , && (5.3)\\
\mathtt{M/Y} &= d_{\mathtt{MM}} + d_{\mathtt{MK}}\,(P_{\mathtt{K}}/P_{\mathtt{M}})^{1/2} + d_{\mathtt{ML}}\,(P_{\mathtt{L}}/P_{\mathtt{M}})^{1/2} + d_{\mathtt{ME}}\,(P_{\mathtt{E}}/P_{\mathtt{M}})^{1/2} + e_{\mathtt{M}}\ , && (5.4)
\end{aligned}$$

where the error terms $e_{\mathtt{K}},\ e_{\mathtt{L}},\ e_{\mathtt{E}},\ e_{\mathtt{M}}$ satisfy the distributional assumptions (2.10). We focus on testing the symmetry restrictions entailed by microeconomic theory, *i.e.*

$$H_{01} : \begin{cases} d_{\mathtt{KL}} = d_{\mathtt{LK}}, & d_{\mathtt{KM}} = d_{\mathtt{MK}} \\ d_{\mathtt{KE}} = d_{\mathtt{EK}}, & d_{\mathtt{LM}} = d_{\mathtt{ML}} \\ d_{\mathtt{LE}} = d_{\mathtt{EL}}, & d_{\mathtt{EM}} = d_{\mathtt{ME}} \end{cases}$$

as well as a subset of these constraints

$$H_{02} : \quad d_{\mathtt{EM}} = d_{\mathtt{ME}}, \quad d_{\mathtt{KM}} = d_{\mathtt{MK}} \quad .$$

Conforming with the procedures described above, we reconsider the testing problem in the context of the MLR model of which the KLEM system is a restricted form. The individual equations of the latter model include the 32 price ratios $(P_i/P_j)^{1/2}$, i, j = K, L, E, M as regressors. The unrestricted MLE SURE estimates using the data provided in Berndt (1991) on the manufacturing sector of the U.S. economy over the period 1947-71 are given below (with asymptotic standard errors in parentheses):

$$\begin{aligned}
\mathtt{K/Y} &= \underset{(.0143)}{.0263} + \underset{(.0088)}{.0036}\,(P_{\mathtt{L}}/P_{\mathtt{K}})^{1/2} + \underset{(.0301)}{.0649}\,(P_{\mathtt{E}}/P_{\mathtt{K}})^{1/2} - \underset{(.0426)}{.0443}\,(P_{\mathtt{M}}/P_{\mathtt{K}})^{1/2} + \widehat{e}_{\mathtt{K}}\ , && (5.5)\\
\mathtt{L/Y} &= -\underset{(.0157)}{.0719} + \underset{(.0245)}{.0517}\,(P_{\mathtt{K}}/P_{\mathtt{L}})^{1/2} + \underset{(.0476)}{.2200}\,(P_{\mathtt{E}}/P_{\mathtt{L}})^{1/2} + \underset{(.0676)}{.0264}\,(P_{\mathtt{M}}/P_{\mathtt{L}})^{1/2} + \widehat{e}_{\mathtt{L}}\ , && (5.6)\\
\mathtt{E/Y} &= \underset{(.0183)}{.0403} - \underset{(.0088)}{.0111}\,(P_{\mathtt{K}}/P_{\mathtt{E}})^{1/2} - \underset{(.0053)}{.0048}\,(P_{\mathtt{L}}/P_{\mathtt{E}})^{1/2} + \underset{(.0259)}{.0150}\,(P_{\mathtt{M}}/P_{\mathtt{E}})^{1/2} + \widehat{e}_{\mathtt{E}}\ , && (5.7)\\
\mathtt{M/Y} &= \underset{(.1214)}{.7401} - \underset{(.0420)}{.0542}\,(P_{\mathtt{K}}/P_{\mathtt{M}})^{1/2} - \underset{(.0258)}{.1374}\,(P_{\mathtt{L}}/P_{\mathtt{M}})^{1/2} + \underset{(.0855)}{.0399}\,(P_{\mathtt{E}}/P_{\mathtt{M}})^{1/2} + \widehat{e}_{\mathtt{M}}\,. && (5.8)
\end{aligned}$$

For both hypotheses, we computed the FGLS and ML-based Wald statistics (4.2), the LR and $LR_{\mathcal{M}}$ criteria as defined in (2.23) and (2.28) and the QLR statistics (4.1). In the case of the Wald and QLR test, we obtained the asymptotic χ^2 and LMC p-values using 19 and 99 simulated samples. The exact BMC and MMC p-values were also obtained for the LR criteria. The bounding statistic $LR_{\mathcal{UL}} = n\ln(\Lambda_{\mathcal{UL}})$ corresponds to the UL hypothesis that sets all the coefficients of the MLR model (except the intercepts) to specific values. As stated in Section 3, the BMC procedure based on LR_M yields tighter bounds [see inequality (3.14)]. Our results are summarized in Table 6.

From these results, we see that the symmetry hypothesis H_{01} is rejected using all asymptotic

TABLE 6A: GENERALIZED LEONTIEF FACTOR DEMANDS:

CROSS-EQUATION SYMMETRY TESTS

$H_{01}: \begin{cases} d_{\mathtt{KL}} = d_{\mathtt{LK}},\ d_{\mathtt{KM}} = d_{\mathtt{MK}},\ d_{\mathtt{KE}} = d_{\mathtt{EK}}, \\ d_{\mathtt{LM}} = d_{\mathtt{ML}},\ d_{\mathtt{LE}} = d_{\mathtt{EL}},\ d_{\mathtt{EM}} = d_{\mathtt{ME}} \end{cases}$								
		LR_M	LR	$QLR_{(0)}$	$QLR_{(1)}$	$QLR_{(2)}$	$Wald_{GLS}$	$Wald_{ML}$
Statistic		176.582	74.159	75.545	75.140	74.911	239.597	238.777
Asymptotic p-value		.000	.000	.000	.000	.000	.000	.000
Reps	MC p-value							
19	BMC	.05	.70	-	-	-	-	-
99		.01	.67	-	-	-	-	-
19	MMC	.05	05	-	-	-	-	-
99		.01	.01	-	-	-	-	-
19	LMC	.05	.05	.05	.05	.05	.05	.05
99		.01	.01	.01	.01	.01	.01	.01

Note: Under H_{01}, $LR_{\mathcal{M}} \overset{asy}{\sim} \chi^2(42)$ while the other statistics have asymptotic $\chi^2(6)$ distributions. The $LR_{\mathcal{M}}$ statistic tests the symmetry restrictions (6 constraints) jointly with the SURE exclusion restrictions (36 constraints) _ a total of 42 restrictions _ against the unrestricted MLR model. *Reps* stands for *replications*.

TABLE 6B: GENERALIZED LEONTIEF FACTOR DEMANDS:

PARTIAL CROSS-EQUATION SYMMETRY TESTS PARTIAL

$H_{02}: \quad d_{\mathtt{EM}} = d_{\mathtt{ME}}, \quad d_{\mathtt{KM}} = d_{\mathtt{MK}}$								
		$LR_{\mathcal{M}}$	LR	$QLR_{(0)}$	$QLR_{(1)}$	$QLR_{(2)}$	$Wald_{GLS}$	$Wald_{ML}$
Statistic		102.574	.15179	.15180	.15179	.15179	.1283	.1279
Asymptotic p-value		.000	.927	.927	.927	.927	.937	.938
Reps	MC p-value							
19	BMC	.15	1.0	-	-	-	-	-
99		.17	1.0	-	-	-	-	-
19	MMC	.05	1.0	-	-	-	-	-
99		.05	1.0	-	-	-	-	-
19	LMC	.05	.90	90	90	90	90	90
99		.04	.94	.94	.94	.94	.94	.94

Note: Under H_{02}, $LR_{\mathcal{M}} \overset{asy}{\sim} \chi^2(38)$ while the other statistics have asymptotic $\chi^2(2)$ distributions. $LR_{\mathcal{M}}$ tests a subset of symmetry restrictions (2 constraints) jointly with the SURE exclusion restrictions (36 constraints) _ 38 restrictions in all _ against the unrestricted MLR model.

and exact tests. In the case of H_{02}, all tests against the unconstrained SURE specification are not significant. However, the asymptotic χ^2 and LMC tests $LR_{\mathcal{M}}$ are significant at the 5% level. Although the bounds p-value is larger than 0.05, the MMC test is significant at the 5% level, even with 19 simulated samples. It is worth noting that the QLR and the LR LMC tests yield equivalent decisions for both testing problems. Moreover, all MC tests based on 19 and 99 replications also yield similar decisions.

6. Conclusion

In this paper we have extended the MLR-based LR test procedure to the SURE framework. We have combined the bounds and Monte Carlo test approaches to provide p-values for test statistics that can yield provably exact tests in finite samples even for nonlinear hypothesis as well as more reliable large sample tests. The feasibility of the test strategy was also illustrated with an extensive Monte Carlo experiment and an empirical application. We have found that standard asymptotic tests exhibit serious errors in level, particularly in larger systems. In contrast, the various tests we have proposed displayed excellent size and power properties.

References

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.

Bera, A. K., Byron, R. P. and Jarque, C. M. (1981), 'Further evidence on asymptotic tests for homogeneity and symmetry in large demand systems', *Economics Letters* **8**, 101–105.

Berndt, E. R. (1991), *The Practice of Econometrics: Classic and Contemporary*, Addison-Wesley, Reading (MA).

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

Berndt, E. R. and Wood, D. O. (1975), 'Technology, prices and the derived demand for energy', *Review of Economics and Statistics* **53**, 259–268.

Breusch, T. S. (1979), 'Conflict among criteria for testing hypotheses: Extensions and comments', *Econometrica* **47**, 203–207.

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. (1997), 'Some impossibility theorems in econometrics, with applications to structural and dynamic models', *Econometrica* **65**, 1365–1389.

Dufour, J.-M. (1998), 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*a*), Simulation based finite and large sample inference methods in multivariate regressions and seemingly unrelated regressions, Technical report, C.R.D.E., Université de Montréal. 36 pages.

Dufour, J.-M. and Khalaf, L. (1998*b*), 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.

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

Goffe, W. L., Ferrier, G. D. and Rogers, J. (1994), 'Global optimization of statistical functions with simulated annealing', *Journal of Econometrics* **60**, 65–99.

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.

Jöckel, K.-H. (1986), 'Finite sample properties and asymptotic efficiency of Monte Carlo tests', *The Annals of Statistics* **14**, 336–347.

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.

Laitinen, K. (1978), 'Why is demand homogeneity so often rejected?', *Economics Letters* **1**, 187–191.

Meisner, J. F. (1979), 'The sad fate of the asymptotic Slutsky symmetry test for large systems', *Economics Letters* **2**, 231–233.

Phillips, P. C. B. (1985), 'The exact distribution of the SUR estimator', *Econometrica* **53**, 745– 756.

Rao, C. R. (1973), *Linear Statistical Inference and its Applications*, second edn, John Wiley & Sons, New York.

Rayner, R. K. (1990), 'Bootstrapping p-values and power in the First-Order autoregression: A Monte Carlo study', *Journal of Business and Economic Statistics* **8**, 251–263.

Rilstone, P. and Veall, M. (1996), 'Using bootstrapped confidence intervals for improved inferences with seemingly unrelated regression equations', *Econometric Theory* **12**, 569–580.

Rocke, D. M. (1989), 'Bootstrap Bartlett adjustment in seemingly unrelated regressions', *Journal of the American Statistical Association* **84**, 598–601.

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

Stewart, K. G. (1997), 'Exact testing in multivariate regression', *Econometric Reviews* **16**, 321–352.

Taylor, T. G., Shonkwiler, J. S. and Theil, H. (1986), 'Monte Carlo and bootstrap testing of demand homogeneity', *Economics Letters* **20**, 55–57.

Theil, H. (1971), *Principles of Econometrics*, John Wiley & Sons, New York.

Theil, H. and Fiebig, D. G. (1985), Small sample and large equation systems, *in* E. J. Hannan, P. R. Krishnaiah and M. M. Rao, eds, 'Handbook of Statistics 5: Time Series in the Time Domain', North-Holland, Amsterdam, pp. 451–480.

Theil, H., Shonkwiler, J. S. and Taylor, T. G. (1985), 'A Monte Carlo test of Slutsky symmetry', *Economics Letters* **19**, 331–332.

Theil, H., Taylor, T. and Shonkwiler, J. S. (1986), Monte Carlo testing in systems of equations, *in* D. J. Slottje and G. F. Rhodes Jr., eds, 'Advances in Econometrics, Volume 5, Innovations in Quantitative Economics: Essays in Honor of Robert L. Basmann', Vol. 5, JAI Press, Greenwich, Conn., pp. 227–239.

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.