Simulation-Based Finite-Sample Inference in Simultaneous Equations

Jean-Marie Dufour*
Université de Montréal

Lynda Khalaf†
Université Laval

December 8, 2003

Centre Interuniversitaire de recherche en économie quantitative (CIREQ), *Centre interuniversitaire de recherche en analyse des organisations* (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.

†Département d'économique and *Groupe de Recherche en économie de l'énergie de l'environement et des ressources naturelles* (GREEN), Université Laval, and *Centre Interuniversitaire de recherche en économie quantitative* (CIREQ), Université de Montréal. Mailing address: 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

In simultaneous equation (SE) contexts, nuisance parameter, weak instruments and identification problems severely complicate exact and asymptotic tests (except for very specific hypotheses). In this paper, we propose exact likelihood based tests for possibly nonlinear hypotheses on the coefficients of SE systems. We discuss a number of bounds tests and Monte Carlo simulation based tests. The latter involves maximizing a randomized p-value function over the relevant nuisance parameter space which is done numerically by using a simulated annealing algorithm. We consider limited and full information models. We extend, to non-Gaussian contexts, the bound given in Dufour (Econometrica, 1997) on the null distribution of the LR criterion, associated with possibly non-linear- hypotheses on the coefficients of one Gaussian structural equation. We also propose a tighter bound which will hold: (i) for the limited information (LI) Gaussian hypothesis considered in Dufour (1997) and for more general, possibly cross-equation restrictions in a non-Gaussian multi-equation SE system. For the specific hypothesis which sets the value of the full vector of endogenous variables coefficients in a limited information framework, we extend the Anderson-Rubin test to the non-Gaussian framework. We also show that Wang and Zivot's (Econometrica, 1998) asymptotic bounds-test may be seen as an asymptotic version of the bound we propose here. In addition, we introduce a multi-equation Anderson-Rubin-type test. Illustrative Monte Carlo experiments show that: (i) bootstrapping standard instrumental variable (IV) based criteria fails to achieve size control, especially (but not exclusively) under near non-identification conditions, and (ii) the tests based on IV estimates do not appear to be boundedly pivotal and so no size-correction may be feasible. By contrast, likelihood ratio based tests work well in the experiments performed.

Contents

List of Assumptions, Propositions and Theorems

List of Tables

1. Introduction

Hypotheses tests in simultaneous equation (SE) models are among the most enduring problems in econometrics. With few exceptions, the distributions of standard test statistics are known only asymptotically due to feedback from the dependent variables to the explanatory variables. Indeed, exact procedures have been proposed only for a few highly special cases. Early in the development of econometric theory relating to the SE model, Haavelmo (1947) constructed exact confidence regions for OLS reduced form parameter estimates and corresponding structural parameter estimates. Bartlett (1948) and Anderson and Rubin (1949, (AR)) proposed exact F-tests for specific classes of hypotheses in the context of a structural equation along with corresponding confidence sets; see also Maddala (1974). Promising extensions of the AR test have recently been discussed in Dufour and Jasiak (2001), Dufour and Taamouti (2003*c*, 2003*b*, 2003*a*) and Dufour (2003). Some exact specification tests have also been suggested for SE. In particular, Durbin (1957) proposed a bounds test against serial correlation in SE and Harvey and Phillips (1980, 1981*a*, 1981*b*, 1989) have suggested tests against serial correlation, heteroskedasticity and structural change in a single structural equation; these tests are based on residuals from a regression of the estimated endogenous part of an equation on all exogenous variables. An exact F-test involving reduced form residuals was proposed by Dufour (1987, Section 3), for the hypothesis of independence between the full vector of stochastic explanatory variables and the disturbance term of a structural equation.[1] Beside these exceptions, available and routinely applied inference procedures in SE are asymptotic. In particular, instrumental variable (IV) methods are the most widely used in empirical practice.

The finite sample distributions of standard estimators and test statistics have received attention early on in this literature. Initial studies (for surveys, see Phillips (1983) and Taylor (1983)) have revealed that: (i) exact distributions are highly complex; (ii) nuisance parameter problems severely hinder the development of exact tests (except for very specific hypotheses); (ii) asymptotic distributions may provide poor approximations in several cases. However, the severity of these findings and their implications on applied work were not recognized until the recent research on *near-identification* or *weak instruments*. Published papers dealing with such problems include: Nelson and Startz (1990*a*), Nelson and Startz (1990*b*), Buse (1992), Choi and Phillips (1992), Maddala and Jeong (1992), Angrist and Krueger (1994), McManus, Nankervis and Savin (1994), Bound, Jaeger and Baker (1995), Cragg and Donald (1996), Hall, Rudebusch and Wilcox (1996), Dufour (1997), Shea (1997), Staiger and Stock (1997), Wang and Zivot (1998), Zivot, Startz and Nelson (1998), Stock and Wright (2000), Dufour and Jasiak (2001), Hahn and Hausman (2002, 2003), Kleibergen (2002), Moreira (2003*a*, 2003*b*), Stock, Wright and Yogo (2002), Kleibergen and Zivot (2003), Perron (2003), Wright (2003); several recent working papers are also cited in Dufour (2003) and Stock et al. (2002). Studies on weak instruments convincingly demonstrate that standard asymptotic procedures (i.e. procedure which *impose identification away* without correcting for local-almost-identification (LAU)) are fundamentally flowed and lead to serious overrejections; these problems are not *small* sample related and occur with fairly large sample sizes, since they are caused by asymptotics failures. In particular Dufour (1997) shows that usual t-type tests, based on common IV estimators, have significance levels that may deviate arbitrarily from their nominal levels since

[1]This procedure generalizes earlier tests suggested by Wu (1973) and Hausman (1978)

it is not possible to bound their null distributions.

To circumvent weak-instruments related difficulties, the above cited recent work on SE has focused on three main directions (see the surveys of Dufour (2003) and Stock et al. (2002)): (i) refinements in asymptotic analysis which include the local-to-zero or local-to-unity frameworks (e.g. Staiger and Stock (1997), Wang and Zivot (1998)), (ii) asymptotic approximations which hold whether instruments are weak or not (e.g. Kleibergen (2002), Moreira (2003*b*)), and (iii) new finite-sample-justified procedures based on proper pivots, i.e. statistics whose null distributions are either nuisance parameter free or bounded by nuisance parameter free distribution [i.e. are *boundedly pivotal*], (e.g. Dufour (1997), Dufour and Jasiak (2001), Dufour and Khalaf (2002), Dufour and Taamouti (2003*c*, 2003*b*, 2003*a*)). So far, provably exact procedures are still in short supply, and typically require normal errors.

With the declining cost of computing, a natural alternative to traditional inference are simulation-based methods such as bootstrapping; for reviews, see Efron (1982), Efron and Tibshirani (1993), Hall (1992), Jeong and Maddala (1993), Vinod (1993), Shao and Tu (1995), Li and Maddala (1996). These surveys suggest that bootstrapping can provide more reliable inference for many problems. In connection with the SE model, examples in which the bootstrap outperforms conventional asymptotics include: Freedman and Peters (1984*a*), Green, Hahn and Rocke (1987), Hu, Lau, Fung and Ulveling (1986), Korajczyk (1985), Dagget and Freedman (1985), and Moreira and Rothenberg (2003). Others however, find that the method leads to little improvement, *e.g.* Freedman and Peters (1984*b*), Park (1985) and Beran and Srivastava (1985), Moreira and Rothenberg (2003). Clearly, there appears to be a conflict in the conclusions regarding the effectiveness of the bootstrap in SE contexts.[2]

This paper addresses these issues and develops alternative simulation based test procedures in limited and full information SE models. The tests we propose are motivated by finite sample arguments. We focus on likelihood ratio (LR) based statistics. This choice is motivated by the propositions in Dufour (1997) pertaining to LR's *boundedly pivotal* characteristic, *i.e.* the fact that LR admits nuisance-parameter-free bounds. Our contributions can be classified in five categories.

First, we extend, to *non-Gaussian* contexts, the bound given in Dufour (1997, (Theorem 5.1)) on the null distribution of the LR criterion, associated with possibly non-linear- hypotheses on the coefficients of *one* Gaussian structural equation.[3] We also propose a tighter bound which will hold: (i) for the limited information (LI) Gaussian hypothesis considered in Dufour (1997, (Theorem 5.1)) (*i.e.* in the context of the LR statistic based on limited information maximum likelihood (LIML) estimation), and (ii) for more general, possibly cross-equation restrictions in a non-Gaussian multi-equation SE system. Formally, we show that Dufour (1997)'s result may be obtained as a special - although non-optimal - case of our proposed bound. To do this, we use the results of Dufour and Khalaf (2002) on hypotheses tests in multivariate linear regression (MLR) models.[4]

[2]In fact, it is well known that bootstrapping may fail to achieve size control when the asymptotic distribution of the underlying test statistic involves nuisance parameters [see Athreya (1987), Basawa, Mallik, McCormick, Reeves and Taylor (1991) and Sriram (1994), and Dufour (2002).

[3]SE LR tests often involve non-linear hypotheses implied by the structure; in connection, see Bekker and Dijkstra (1990) or Byron (1974)

[4]The relationship between the MLR and the SE model is readily seen: when all the predetermined variables of a SE system are strictly exogenous, the reduced form is equivalent to a (restricted) MLR system.

Second, for the specific hypothesis which sets the value of the full vector of endogenous variables coefficients in a LI framework, we show that Wang and Zivot (1998)'s asymptotic bounds-test may be seen as an asymptotic version of the bound we propose here. We use this result to extend the validity of Wang and Zivot (1998)'s bound to the case of general linear hypotheses on structural coefficients. To do this, we show that our general bound on the LIML is based on an AR-type bounding pivotal statistic.

Third, we extend the AR-test to the non-Gaussian framework. Specifically, we show analytically that the proof of its pivotality in finite samples *does not require normal errors*. This is achieved by re-writing the AR statistic as an LR-type criterion (based on the LI reduced form). To date, available exact AR-type tests require normality assumptions. In this regard, our results are noteworthy.

Fourth, our re-interpretation of the AR-test allows to re-write Kleibergen (2002)'s test as a approximate generalized AR-test (see Dufour (2003) and Dufour and Taamouti (2003*c*, 2003*b*, 2003*a*)) obtained with a specific instrument substitution choice. Specifically, we prove analytically that Kleibergen (2002)'s test can be obtained as an F-test for the exclusion of a specific instrument matrix, based on a constrained estimate of the coefficient of the excluded regressors in the first stage regression. To do this, we use the expression provided in Dufour (2003, Section 6.3 (d)) as well as known results from the MLR literature (Berndt and Savin (1977), Dufour and Khalaf (2002)).

Fifth, we propose a multi-equation Anderson-Rubin-type test which also admits a pivotal bound based on the results of Dufour and Khalaf (2003) relating to SURE models. In view of the renewed interest in the Anderson-Rubin test (see Dufour (1997), Dufour and Jasiak (2001), Staiger and Stock (1997), Wang and Zivot (1998) and Dufour and Taamouti(2003*c*, 2003*b*, 2003*a*)), extensions to a systems context may prove useful.

It is important, at this stage, to emphasize that the distributional theory which underlies all the above procedures holds whether identification constraints are imposed or not. Consequently, identification problems are resolved without the need to introduce non-standard, e.g. local-to-zero, asymptotics. Furthermore, although exactness is obtained under parametric assumptions (which are duly defined in the paper), normality is not strictly required.

Sixth, this paper makes several contributions relevant to simulation-based tests. Indeed, the null distribution of all statistics considered may be quite complex, particularly in non-Gaussian contexts. In view of this, we propose, following Dufour and Khalaf (2002), to apply the Monte Carlo (MC) test procedure [Dwass (1957), Barnard (1963), Dufour (2002)] to obtain simulation based exact p-values. MC test procedures may be viewed as parametric bootstrap tests applied to statistics whose null distribution does not involve nuisance parameters, with however a fundamental additional observation: the associated randomized test procedure can easily be performed to control test size exactly, for a given number of replications.

Here, recall that we consider two types of statistics, the pivotal ones (our extensions of the AR test), and the boundedly pivotal ones (general LR-LIML and multi-equation AR test). The MC test method easily yields exact p-values given pivotal statistics; to avoid confusion in what follows, we will refer to MC tests based on exact pivots as *pivotal MC tests* (PMC). Boundedly pivotal statistics are approached through two MC test procedures. First, we consider the *bounds-MC technique* (BMC) (Dufour (2002), Dufour and Khalaf (2002)). This methods differs from the PMC one in the fact that the null distribution of the bounding statistics (which is pivotal by construction) is considered. Second, we apply the *maximized MC method* (MMC) (Dufour (2002)); this method requires; (i) defining a p-value function which gives a bootstrap-type MC p-value conditional on relevant nuisance parameters, (ii) maximizing the latter function (using global maximization algorithms) over these nuisance parameters.[5] The latter method may be viewed as a numerical search for the optimal bound.

It is clear that such a search may be computationally expensive. So we propose to combine the BMC with an MMC test, which can be run whenever the bounds test is not significant. To understand this strategy, recall that the BMC test is exact in the sense that rejections (at level α) are conclusive. Furthermore, we show that the MMC algorithm may be written in a way to include a standard parametric bootstrap as a first step. Possibly expensive iterations - to obtain the maximal MC p-value in question which underlies the MMC test - may thus be saved if the bootstrap p-value exceeds α.

To illustrate the performance of these tests particularly given identification issues, we run a small-scale simulation experiment. Our main findings are: (i) MC methods based on randomization procedures where unknown parameters are replaced by estimators do not achieve size control, and (ii) MMC p-values for IV-based test are always one; in other words, it is does not appear possible to find a non trivial bound on the rejection probabilities, so that standard asymptotic and bootstrap procedures are deemed to fail when applied to such statistics. In contrast, LR-based MMC tests allow one to control the level of the procedure.

The paper is organized as follows. Section 2 develops the notation and definitions. In Section 3 we discuss pivotal statistics in full and sub-systems; general hypotheses are considered in Section 4. The MC test procedures applied to pivotal and general hypotheses are presented in 5. Simulation results are reported in Section 6 and Section 7 concludes the paper.

2. Framework

We consider a system of p simultaneous equations of the form

$$YB + X\mathit{\Gamma} = U, \tag{2.1}$$

where $Y = [y_1 \, , \, \dots \, , \, y_p]$ is an $n \times p$ matrix of observations on p endogenous variables, X is an $n \times k$ matrix of fixed (or strictly exogenous) variables and $U = [u_1 \, , \, \dots \, , \, u_p] = [U_1 \, , \, \dots \, , \, U_n]'$ is a matrix of random disturbances. The coefficient matrix B is assumed to be invertible. The equations in (2.1) give the *structural form* of the model. Post-multiplying both sides by B^{-1} leads to the *reduced form*

$$Y = X\mathit{\Pi} + V, \;\; \mathit{\Pi} = -\mathit{\Gamma}B^{-1}, \;\; \pi = vec(\mathit{\Pi}), \tag{2.2}$$

where $V = [v_1 \, , \, \dots \, , \, v_p] = [V_1 \, , \, \dots \, , \, V_n]'$ is the matrix of reduced form disturbances. Further, we suppose the rows of U satisfy the following distributional assumptions:

$$U_t \sim JW_t, \;\; t = 1 \, , \, \dots \, , \, n, \tag{2.3}$$

[5]MMC *p*-values are computed using a simulated annealing (SA) optimization algorithm; see Corana, Marchesi, Martini and Ridella (1987) or Goffe, Ferrier and Rogers (1994).

where the vector $w = vec(W_1 , \ldots , W_n)$ has a known distribution and J is an unknown nonsingular matrix; for further reference, let $W = [W_1 , \ldots , W_n]'$ where (2.3) implies that

$$W = U(J^{-1})'. \tag{2.4}$$

When $Var(W_t) = I_p$, $var(U_t) = JJ' \equiv \Omega$ and $var(V_t) = (B^{-1})'\Omega B^{-1} = (B^{-1})'JJ'B^{-1} \equiv \Sigma$. Of course, condition (2.3) will be satisfied when

$$W_t \sim N(0, I_p), \;\; t = 1 , \ldots , n. \tag{2.5}$$

A key feature of SE models is the imposition of identification conditions on the structural coefficients. Usually, these conditions are formulated in terms of zero restrictions on B and Γ. In addition, a normalization constraint is imposed which is usually achieved by setting the diagonal elements of B equal to one. We can rewrite model (2.1), given exclusion and normalization restrictions as

$$y_i = Y_i\beta_i + X_{1i}\gamma_{1i} + u_i, \;\; i = 1 , \ldots , p, \tag{2.6}$$

where Y_i and X_{1i} are $n \times m_i$ and $n \times k_i$ matrices which respectively contain the observations on the included endogenous and exogenous variables of the model. Many problems are also formulated in terms of limited-information (LI) models such as

$$\begin{aligned} y_i &= Y_i\beta_i + X_{1i}\gamma_{1i} + u_i = Z_i\delta_i + u_i, \\ Y_i &= X_{1i}\Pi_{1i} + X_{2i}\Pi_{2i} + V_i, \end{aligned} \tag{2.7}$$

where $Z_i = [Y_i, X_{1i}]$, $\delta_i = (\beta_i', \gamma_{1i}')'$ and X_{2i} refers to the excluded exogenous variables. The associated LI reduced form is

$$\begin{bmatrix} y_i & Y_i \end{bmatrix} = X\Pi_i + \begin{bmatrix} v_i & V_i \end{bmatrix}, \; \Pi_i = \begin{bmatrix} \pi_{1i} & \Pi_{1i} \\ \pi_{2i} & \Pi_{2i} \end{bmatrix}, \; X = \begin{bmatrix} X_{1i} & X_{2i} \end{bmatrix} \tag{2.8}$$

$$\pi_{1i} = \Pi_{1i}\beta_i + \gamma_{1i}, \;\; \pi_{2i} = \Pi_{2i}\beta_i, \tag{2.9}$$

which lead to the necessary and sufficient condition for identification

$$rank(\Pi_{2i}) = m_i. \tag{2.10}$$

Our LI-analogue of (2.3) can be stated as follows. Let V_{it} refer to the tth row of V_i, then the rows of $\begin{bmatrix} u_i & V_i \end{bmatrix}$ satisfy the following distributional assumptions:

$$\begin{pmatrix} u_{it} & V_{it}' \end{pmatrix} \sim J_iW_t^i, \;\; t = 1 , \ldots , n, \tag{2.11}$$

where $vec(W_1^i , \ldots , W_n^i)$ has a known distribution and J_i is an unknown non-singular matrix. When $Var(W_t^i) = I_{m_i+1}$,

$$var \begin{pmatrix} u_{it} & V_{it}' \end{pmatrix} = J_iJ_i' \equiv \Omega_i. \tag{2.12}$$

For further reference, let $W^i = [W_1^i \,,\, ... \,,\, W_n^i]'$ where (2.11) implies that

$$W^i = \left[\begin{array}{cc} u_i & V_i \end{array}\right] (J_i^{-1})'. \tag{2.13}$$

In this context, LIML corresponds to maximizing, imposing (2.9), the likelihood function

$$\mathcal{L}(y_i, Y_i | X_{1i}, X_{2i}) = -\frac{n(m+1)}{2} \ln(2\pi) - \frac{n}{2} \ln |\Sigma_i| - \frac{1}{2} tr \Sigma_i^{-1} D_i' D_i, \tag{2.14}$$

where $D_i = \left[\begin{array}{cc} y_i & Y_i \end{array}\right] - X\Pi_i$ and Σ_i denotes the relevant reduced form error covariance. Numerical maximization may be considered, yet it is well know that an equivalent solution obtains through an eigenvalue/eigenvector problem based on the following determinantal equation

$$\left| \left[\begin{array}{cc} y_i & Y_i \end{array}\right]' M_{1i} \left[\begin{array}{cc} y_i & Y_i \end{array}\right] - \lambda_i \left[\begin{array}{cc} y_i & Y_i \end{array}\right]' M \left[\begin{array}{cc} y_i & Y_i \end{array}\right] \right| = 0 \tag{2.15}$$

where $M = I - X(X'X)^{-1}X'$, $M_{1i} = I - X_{1i}(X_{1i}'X_{1i})^{-1}X_{1i}'$ and λ_i refers to the eigen value in question. Indeed, it can be shown (see, for example Davidson and MacKinnon (1993, Chapter 18), Wang and Zivot (1998)) that the estimator of β is $\widetilde{\beta}_i = \underset{\beta_i}{\text{ARGMIN}} \{\lambda(\beta_i)\}$

$$\lambda(\beta_i) = \frac{[y_i - Y_i\beta_i]' M_{1i} [y_i - Y_i\beta_i]}{[y_i - Y_i\beta_i]' M_{1i} (I - M_{1i}X_{2i}(X_{2i}'M_{1i}X_2)^{-1}X_{2i}'M_{1i}) M_{1i} [y_i - Y_i\beta_i]}. \tag{2.16}$$

Formally, the LIML estimator of β_i and γ_{1i} is

$$\widetilde{\delta}_i = \left[\begin{array}{c} \widetilde{\beta}_i \\ \widetilde{\gamma}_{1i} \end{array}\right] = \left[\begin{array}{cc} Y_i'Y_i - \widetilde{\lambda}_i Y_i' M Y_i & Y_i'X \\ X'Y_i & X'X \end{array}\right]^{-1} \left[\begin{array}{c} Y_i' - \widetilde{\lambda}_i Y_i' M \\ X_i' \end{array}\right] y_i \tag{2.17}$$

where $\widetilde{\lambda}_i$ is the smallest root of (2.15), which corresponds to $\lambda(\widetilde{\beta}_i)$ [where $\lambda(\beta_i)$ is given by (2.16)]. Correspondingly, expressions for the reduced form parameter estimates obtain as follows (see Theil (1971), appendix B):

$$\left[\begin{array}{cc} \widetilde{\pi}_{1i} & \widetilde{\Pi}_{1i} \end{array}\right] = (X_{1i}'X_{1i})^{-1}X_{1i}' \left(\left[\begin{array}{cc} y_i & Y_i \end{array}\right] - X_{2i} \left[\begin{array}{cc} \widetilde{\pi}_{2i} & \widetilde{\Pi}_{2i} \end{array}\right] \right) \tag{2.18}$$

$$\begin{aligned} \left[\begin{array}{cc} \widetilde{\pi}_{2i} & \widetilde{\Pi}_{2i} \end{array}\right] &= (X_{2i}'M_{1i}X_{2i})^{-1}X_{2i}'M_{1i} \left[\begin{array}{cc} y_i & Y_i \end{array}\right] \\ &\quad - \frac{(X_{2i}'M_{1i}X_{2i})^{-1}X_{2i}'M_{1i} \left[\begin{array}{cc} y_i & Y_i \end{array}\right]}{\left[\begin{array}{c} 1 \\ -\widetilde{\beta}_i \end{array}\right]' \widetilde{\Sigma}_i \left[\begin{array}{c} 1 \\ -\widetilde{\beta}_i \end{array}\right]} \left[\begin{array}{c} 1 \\ -\widetilde{\beta}_i \end{array}\right] \left[\begin{array}{c} 1 \\ -\widetilde{\beta}_i \end{array}\right]' \widetilde{\Sigma}_i \end{aligned} \tag{2.19}$$

$$\widetilde{\Sigma}_i = \frac{\left[\begin{array}{cc} y_i & Y_i \end{array}\right]' M \left[\begin{array}{cc} y_i & Y_i \end{array}\right]}{n} + \frac{(\widetilde{\lambda} - 1)}{n} \tag{2.20}$$

$$\times \frac{\left[\begin{array}{cc} y_i & Y_i \end{array}\right]' M \left[\begin{array}{cc} y_i & Y_i \end{array}\right] \left[\begin{array}{c} 1 \\ -\widetilde{\beta}_i \end{array}\right] \left(\left[\begin{array}{cc} y_i & Y_i \end{array}\right]' M \left[\begin{array}{cc} y_i & Y_i \end{array}\right] \left[\begin{array}{c} 1 \\ -\widetilde{\beta}_i \end{array}\right] \right)'}{\left[\begin{array}{c} 1 \\ -\widetilde{\beta}_i \end{array}\right]' \left[\begin{array}{cc} y_i & Y_i \end{array}\right]' M \left[\begin{array}{cc} y_i & Y_i \end{array}\right] \left[\begin{array}{c} 1 \\ -\widetilde{\beta}_i \end{array}\right]}.$$

The derivations of Theil (1971) also imply that $\left|\widetilde{\Sigma}_i\right|$ satisfies

$$\left|\widetilde{\Sigma}_i\right| = \widetilde{\lambda}_i \left|' \left[\begin{array}{cc} y_i & Y_i \end{array}\right]' M \left[\begin{array}{cc} y_i & Y_i \end{array}\right]\right|. \tag{2.21}$$

For hypotheses of the form $R_i \delta_i = r_i$ on the coefficients of (2.7), where R_i is a known $q_i \times m_i$ matrix of rank q_i and r_i is known, Wald statistics are routinely applied and take the form

$$\begin{aligned} \tau_w &= \frac{1}{s^2}(r_i - R_i \widehat{\hat{\delta}}_i)' - [R_i'(Z_i P_i (P_i' P_i)^{-1} P_i' Z_i)^{-1} R_i]\, (r_i - R_i \widehat{\hat{\delta}}_i), \\ s^2 &= \frac{1}{n}(y_i - Z_i \widehat{\hat{\delta}}_i)'(y_i - Z_i \widehat{\hat{\delta}}_i)' \end{aligned} \tag{2.22}$$

where $\widehat{\hat{\delta}}_i$ is a consistent asymptotically normal estimator such as (2.17) or the 2SLS

$$\hat{\delta}_i = [Z_i' P_i (P_i' P_i)^{-1} P_i' Z_i]^{-1} Z_i' P_i (P_i' P_i)^{-1} P_i' y_i, \;\; P_i = \left[\begin{array}{cc} X & X(X'X)^{-1} X' Y_i \end{array}\right].$$

Imposing identification, the asymptotic null distribution of τ_w is $\chi^2(q)$. For an asymptotic theory conformable with under-identification, see Staiger and Stock (1997).

3. Pivotal Statistics in systems and subsystems

The recent literature on SE models has underscored the importance of proper pivots. This section characterizes pivotal statistics in possibly non-Gaussian systems and subsystems, which include the case of one single structural equation (the LI case). We first consider the LI context, since it is a fundamental one, and because it may be used to explicate our multi-equation approach.

3.1. Non-Gaussian extensions of the Anderson-Rubin test

In the context of the LI model (2.7), consider hypotheses of the form:

$$H_{AR} : \beta_i = \beta_i^0, \tag{3.1}$$

where β_i^0 is a known vector. Let $y_i^0 = y_i - Y_i \beta_i^0$; then (3.1) may be tested in the context of the transformed structural system

$$y_i^0 = Y_i(\beta_i - \beta_i^0) + X_{1i}\gamma_{1i} + u_i, \tag{3.2}$$

$$Y_i = X_{1i} \Pi_{1i} + X_{2i} \Pi_{2i} + V_i, \tag{3.3}$$

with reduced form

$$\begin{aligned} \left[\begin{array}{cc} y_i^0 & Y_i \end{array}\right] &= \left[\begin{array}{cc} X_{1i} & X_{2i} \end{array}\right] \varPi_i + \left[\begin{array}{cc} u_i + V(\beta_i - \beta_i^0) & V_i \end{array}\right], \\ \pi_{1i} &= \varPi_{1i}(\beta_i - \beta_i^0) + \gamma_{1i}, \quad \pi_{2i} = \varPi_{2i}(\beta_i - \beta_i^0). \end{aligned}$$

Let $O_{(s,j)}$ denotes a zero $s \times j$ matrix. In this context, (3.1) corresponds to testing

$$\left[O_{(k-k_i,\, k_i)},\ I_{(k-k_i)}\right] \varPi_i C_i = 0, \qquad C_i = \left[\begin{array}{c} 1 \\ O_{(m_i,\, 1)} \end{array}\right]. \tag{3.4}$$

To simplify the presentation, note that since the hypothesis concerns solely the element of β_i, the test may be recast in the context of:

$$\begin{aligned} M_{1i} \left[\begin{array}{cc} y_i^0 & Y_i \end{array}\right] C_i &= M_{1i} X_{2i} \varPi_{AR} + M_{1i} \left[\begin{array}{cc} u_i + V_i(\beta_i - \beta_i^0) & V_i \end{array}\right] C_i \\ \varPi_{AR} &= \left[\begin{array}{cc} \pi_{2i} & \varPi_{2i} \end{array}\right] C_i \end{aligned}$$

with null hypothesis $\varPi_{AR} = 0$. The QLR statistic in this case takes the form (see Dufour and Khalaf (2002)) where $P_{M_{1i}X_{2i}} = I - M_{1i}X_{2i}(X_{2i}'M_{1i}X_{2i})^{-1}X_{2i}'M_{1i}$

$$\frac{|\hat{\varSigma}_{AR}^0|}{|\hat{\varSigma}_{AR}|} = \frac{C_i' \left[\begin{array}{cc} y_i^0 & Y_i \end{array}\right]' M_{1i} \left[\begin{array}{cc} y_i^0 & Y_i \end{array}\right] C_i}{C_i' \left[\begin{array}{cc} y_i^0 & Y_i \end{array}\right]' M_{1i} P_{M_{1i}X_{2i}} M_{1i} \left[\begin{array}{cc} y_i^0 & Y_i \end{array}\right] C_i} = \frac{y_i^{0\prime} M_{1i} y_i^0}{y_i^{0\prime} M_{1i} P_{M_{1i}X_{2i}} M_{1i} y_i^0}$$

which is a monotonic transformation of the Anderson-Rubin statistic.

Theorem 3.1 DISTRIBUTION OF THE AR TEST STATISTIC. *In the context of the LI model* (2.7)*, consider the problem of testing* (3.1)

$$H_{AR} : \beta_i = \beta_i^0$$

imposing (2.11) *where the first row of J_i has zeros everywhere except for the first element. Let*

$$\varLambda_{AR} = \frac{[y_i - Y_i\beta_{0i}]' M_{1i} [y_i - Y_i\beta_{0i}]}{[y_i - Y_i\beta_{0i}]' M_{1i} P_{M_{1i}X_{2i}} M_{1i} [y_i - Y_i\beta_{0i}]} \tag{3.5}$$

be the associated Anderson-Rubin statistic. Then under the null hypothesis

$$P[\varLambda_{AR} \geq x] = P\left[\frac{|w_i' M_{1i} w_i|}{|w_i' M_{1i} P_{M_{1i}X_{2i}} M_{1i} w_i|} \geq x\right], \ \forall x,$$

where $w_i = \left(\begin{array}{cccc} w_1^i & w_2^i & \ldots & w_n^i \end{array}\right)'$ gives the first column of W^i as defined in (2.11)-(2.13).

PROOF. Under the null hypothesis,

$$\frac{|\hat{\varSigma}_{AR}^0|}{|\hat{\varSigma}_{AR}|} = \frac{u_i' M_{1i} u_i}{u_i' M_{1i} P_{M_{1i}X_{2i}} M_{1i} u_i}.$$

Given assumption (2.11), $u_i = \begin{bmatrix} u_i & V_i \end{bmatrix} C_i = W^i J_i' C_i$. When the first row of J_i in (2.11) has zeros everywhere, except for the first element which equals $\sigma \neq 0$, then $J_i' C_i = \sigma C_i$ and $W^i J_i' C_i = \sigma w_i = \sigma W^i C_i$, so

$$\frac{|\hat{\Sigma}_{AR}^0|}{|\hat{\Sigma}_{AR}|} = \frac{\sigma C_i' W^{i\prime} M_{1i} W^i C_i \sigma}{\sigma C_i' W^{i\prime} M_{1i} P_{M_{1i}X_{2i}} M_{1i} M_{1i} W^i C_i \sigma} = \frac{C_i' W^{i\prime} M_{1i} W^i C_i}{C_i' W^{i\prime} M_{1i} P_{M_{1i}X_{2i}} M_{1i} W^i C_i}. \tag{3.6}$$

Then the result obtains on observing that $w_i = W^i C_i$. ■

The latter result means that an exact test can be carried out in non-normal context without the need to specify the distribution of the full W^i matrix. If normality is further imposed, then it is straightforward to see (see also Dufour and Khalaf (2002)) that

$$[\Lambda_{AR} - 1]\frac{n-k}{k-k_i} \sim F(k - k_i, n - k).$$

As usual, the AR procedure can be adapted to test hypotheses on γ_{1i} (in addition to constraints on β_i). It is clear that our results will apply to this case as well. So consider now the problem of testing

$$H_{ARX} : \beta_i = \beta_i^0, \;\; \gamma_{11i} = \gamma_{11i}^0 \tag{3.7}$$

where $\gamma_{1i} = (\gamma_{11i}', \gamma_{12i}')$, γ_{11i} is $k_{1i} \times 1$, and $X_{1i} = \begin{bmatrix} X_{11i} & X_{12i} \end{bmatrix}$ is decomposed conformably. The associated Anderson-Rubin statistic

$$\begin{aligned} \Lambda_{ARX} &= \frac{\left[y_i - Y_i\beta_{0i} - X_{11i}\gamma_{11i}^0\right]' M_{12i} \left[y_i - Y_i\beta_{0i} - X_{11i}\gamma_{11i}^0\right]}{\left[y_i - Y_i\beta_{0i} - X_{11i}\gamma_{11i}^0\right]' M_{12i} P_{M_{12i}X_{22i}} M_{12i} \left[y_i - Y_i\beta_{0i} - X_{11i}\gamma_{11i}^0\right]} \\ M_{12i} &= I - X_{12i}(X_{12i}'X_{12i})^{-1}X_{12i}', \;\; X_{22i} = \begin{bmatrix} X_{11i} & X_{2i} \end{bmatrix} \\ P_{M_{12i}X_{22i}} &= I - M_{12i}X_{22i}(X_{22i}'M_{12i}X_{22i})^{-1}X_{22i}'M_{12i}. \end{aligned}$$

Then following the same arguments as in Theorem **3.1**, we can show that under the null hypothesis

$$P[\Lambda_{ARX} \geq x] = P\left[\frac{|w_i' M_{12i} w_i|}{|w_i' M_{12i} P_{M_{12i}X_{22i}} M_{12i} w_i|} \geq x\right], \forall x, \tag{3.8}$$

and if normality is further imposed,

$$[\Lambda_{ARX} - 1]\frac{n-k}{k-k_i-k_{1i}} \sim F(k - k_i - k_{1i}, n - k). \tag{3.9}$$

Finally, consider the hypothesis analyzed in Dufour and Jasiak (2001, Section 4):

$$H_{ARQX} : \beta_i = \beta_i^0, \;\; Q_{1i}\gamma_{1i} = \nu_0 \tag{3.10}$$

where Q_{1i} is a $q_{1i} \times k_i$ matrix where $q_{1i} = rank(Q_{1i})$; Q_{1i} can be treated as submatrix of an invertible $k_i \times k_i$ matrix $Q_i = \begin{bmatrix} Q'_{1i} & Q'_{2i} \end{bmatrix}'$ so that

$$Q_i \gamma_{1i} = \begin{bmatrix} Q_{1i}\gamma_{11i} \\ Q_{2i}\gamma_{21i} \end{bmatrix} = \begin{bmatrix} \nu_{1i} \\ \nu_{2i} \end{bmatrix}.$$

Let $X_{Q_i} = X_{1i}Q_i^{-1} = \begin{bmatrix} X_{Q_{1i}} & X_{Q_{2i}} \end{bmatrix}$ where $X_{Q_{1i}}$ and $X_{Q_{2i}}$ are $T \times q_{1i}$ and $T \times (k_i - q_{1i})$ matrices, so the LI equation can be re-written as

$$y_i = Y_i\beta_i + X_{Q_{1i}}\nu_{1i} + X_{Q_{2i}}\nu_{2i} + u_i,$$

in which case testing H_{ARQX} amounts to assessing $\beta_i = \beta_i^0, \ \nu_{1i} = \nu_0$. The associated Anderson-Rubin statistic

$$\begin{aligned} \Lambda_{ARQX} &= \frac{[y_i - Y_i\beta_{0i} - X_{Q_{1i}}\nu_0]' \, M_{Q_{2i}} \, [y_i - Y_i\beta_{0i} - X_{Q_{1i}}\nu_0]}{[y_i - Y_i\beta_{0i} - X_{Q_{1i}}\nu_0]' \, M_{Q_{2i}} P_{M_{Q_{2i}}X_{22i}} M_{Q_{2i}} \, [y_i - Y_i\beta_{0i} - X_{Q_{1i}}\nu_0]} \\ M_{Q_{2i}} &= I - X_{Q_{2i}}(X'_{Q_{2i}}X_{Q_{2i}})^{-1}X'_{Q_{2i}}, \ \ X_{22i} = \begin{bmatrix} X_{Q_{1i}} & X_{2i} \end{bmatrix} \\ P_{M_{Q_{2i}}X_{22i}} &= I - M_{Q_{2i}}X_{22i}(X'_{22i}M_{Q_{2i}}X_{22i})^{-1}X'_{22i}M_{Q_{2i}}. \end{aligned}$$

The same arguments underlying (3.8) yield

$$P[\Lambda_{ARQX} \geq x] = P\left[\frac{|w_i' M_{Q_{2i}} w_i|}{\left|w_i' M_{Q_{2i}} P_{M_{Q_{2i}}X_{22i}} M_{Q_{2i}} w_i\right|} \geq x\right], \forall x, \tag{3.11}$$

and imposing normality

$$P[[\Lambda_{ARQX} - 1]\frac{n-k}{k - k_i - q_{1i}} \geq x] = P\left[F(k - k_i - q_{1i}, n-k) \geq x\right]. \tag{3.12}$$

It is also easy to show, using the same arguments as in the above Theorems, that all the roots of the determinantal equation

$$\begin{aligned} \left|y_i^{0\prime} M_{Q_{2i}} y_i^0 - \mu \, y_i^{0\prime} M_{Q_{2i}} P_{M_{Q_{2i}}X_{22i}} M_{Q_{2i}} y_i^0\right| &= 0 \\ [y_i - Y_i\beta_{0i} - X_{Q_{1i}}\nu_0] &= y_i^0 \end{aligned}$$

are pivotal under the null hypothesis, which lead to alternative statistics, such as the Lawley-Hotelling trace criterion, the Bartlett-Nanda-Pillai trace criterion and the maximum Root criterion.[6]

To conclude this section, it is useful to consider the test proposed by Kleibergen (2002) in the context of (3.1). Dufour (2003) shows that the latter test corresponds to an AR-type test applied with a specific instrument choice (denoted Z_K). Specifically, equations 83-86 from Dufour (2003)

[6] For references, see Rao (1973, Chapter 8) or Anderson (1984, chapters 8 and 13) and Dufour and Khalaf (2002).

rewritten in terms of the transformed model

$$M_{1i}\begin{bmatrix} y_i & Y_i \end{bmatrix} = M_{1i}X_{2i}\begin{bmatrix} \pi_{2i} & \Pi_{2i} \end{bmatrix} + M_{1i}\begin{bmatrix} u_i & V_i \end{bmatrix} \tag{3.13}$$

lead to the instrument

$$\begin{aligned} Z_K &= X_{2i}\overline{\Pi}_{2i}, \quad \overline{\Pi}_{2i} = \widehat{\Pi}_{2i} - \widehat{\pi}_{2i}(\widetilde{\beta}_i)\frac{S_{\varepsilon V}(\beta_i^0)}{S_{\varepsilon\varepsilon}(\beta_i^0)} \qquad (3.14)\\ \widehat{\Pi}_{2i} &= (X_{2i}'M_{1i}X_{2i})^{-1}X_{2i}'M_{1i}Y_i, \quad \widehat{\pi}_2(\beta_i^0) = (X_{2i}'M_{1i}X_{2i})^{-1}X_{2i}'M_{1i}\left[y_i - Y_i\beta_i^0\right]\\ S_{\varepsilon V}\left(\beta_i^0\right) &= \frac{1}{T-k}\left[y_i - Y_i\beta_i^0\right]' MY_i, \quad S_{\varepsilon\varepsilon}(\beta_i^0) = \frac{1}{T-k}\left[y_i - Y_i\beta_i^0\right]' M\left[y_i - Y_i\beta_i^0\right]. \end{aligned}$$

Here we argue that the later expression is a constrained OLS estimator of Π_{2i}, imposing the LIML structure. Expressions for constrained OLS estimates of (3.13) can be derived using the formulae from the general theory on MLR imposing uniform linear hypotheses (see Berndt and Savin (1977, equations 5 and 6) and Dufour and Khalaf (2002)). In is context, the AR null hypothesis takes the form (in the notation of Berndt and Savin (1977)) $F\begin{bmatrix} \pi_{2i} & \Pi_{2i} \end{bmatrix} G = E$,where $F = I_{k_2}$, $E = 0$ and $G = \left(1, -\beta_i^{0\prime}\right)'$. Then applying equation (5) from Berndt and Savin (1977) which we reproduce here for convenience (where P_0 and P give the formula for the constrained and unconstrained estimators of $\begin{bmatrix} \pi_{2i} & \Pi_{2i} \end{bmatrix}$ in (3.13))

$$\begin{aligned} P_0 &= P - \left(\widetilde{X}'\widetilde{X}\right)^{-1} F'\left[F\left(\widetilde{X}'\widetilde{X}\right)^{-1} F'\right]^{-1}(FPG - E)\left[G'SG\right]^{-1} G'S,\\ P &= \left(\widetilde{X}'\widetilde{X}\right)^{-1}\widetilde{X}\widetilde{y}, \quad S = (\widetilde{y} - \widetilde{X}P_0)'(\widetilde{y} - \widetilde{X}P_0),\\ \widetilde{y} &= M_{1i}\begin{bmatrix} y_i & Y_i \end{bmatrix}, \quad \widetilde{X} = M_{1i}X_{2i}, \end{aligned}$$

yields the following expression for the constrained QMLE estimates:[7]

$$\begin{aligned} \begin{bmatrix} \widehat{\pi}_{2i}^0 & \widehat{\Pi}_{2i}^0 \end{bmatrix} &= (X_{2i}'M_{1i}X_{2i})^{-1}X_{2i}'M_{1i}\begin{bmatrix} y_i & Y_i \end{bmatrix}\\ &\quad - \frac{(X_{2i}'M_{1i}X_{2i})^{-1}X_{2i}'M_{1i}\begin{bmatrix} y_i & Y_i \end{bmatrix}}{\begin{bmatrix} 1 \\ -\beta_i^0 \end{bmatrix}'\widehat{\Sigma}_i\begin{bmatrix} 1 \\ -\beta_i^0 \end{bmatrix}}\begin{bmatrix} 1 \\ -\beta_i^0 \end{bmatrix}\begin{bmatrix} 1 \\ -\beta_i^0 \end{bmatrix}'\widehat{\Sigma}_i \end{aligned}$$

or alternatively

$$\begin{aligned} \begin{bmatrix} \widehat{\pi}_2^0 & \widehat{\Pi}_2^0 \end{bmatrix} &= (X_{2i}'M_{1i}X_{2i})^{-1}X_{2i}'M_{1i}\begin{bmatrix} y_i & Y_i \end{bmatrix}\\ &\quad -(X_{2i}'M_{1i}X_{2i})^{-1}X_{2i}'M_{1i}\left[y_i - Y_i\beta_i^0\right]\frac{\left[y_i - Y_i\beta_i^0\right]' M\begin{bmatrix} y_i & Y_i \end{bmatrix}}{\left[y_i - Y_i\beta_i^0\right]' M\left[y_i - Y_i\beta_i^0\right]}. \end{aligned}$$

[7] A similar expression for the constrained LIML estimator of Π_{2i} obtains, replacing β_i^o by $\widetilde{\beta}_i$; see (2.19).

Post-multiplying the latter expression by $\begin{bmatrix} O(1,m) \\ I_m \end{bmatrix}$ leads to the estimator

$$\widehat{\Pi}_{2i}^0 = \widehat{\Pi}_{2i} - (X_{2i}'M_{1i}X_{2i})^{-1}X_{2i}'M_{1i}\left[y_i - Y_i\beta_i^0\right]\frac{\left[y_i - Y_i\beta_i^0\right]' MY_i}{\left[y_i - Y_i\beta_i^0\right]' M\left[y_i - Y_i\beta_i^0\right]}.$$

which is exactly equal to $\overline{\Pi}_{2i}$ as defined in (3.14). Recall that Dufour (2003) has shown that Wang and Zivot (1998)'s LM_{GMM} test obtains as an AR-type test with instrument $X_{2i}\widehat{\Pi}_{2i}$. We thus see that Kleibergen (2002) is highly related to the latter, since it is obtained in a similar way, replacing the unconstrained OLS estimator of Π_{2i} by a constrained OLS estimator which imposes the structure. As mentioned in Dufour (2003), these tests are affected by the fact that instruments are not independent from the error term u_i, and thus are not pivotal in finite samples.

3.2. Multi-equation non-Gaussian extensions of the Anderson-Rubin test

The results of the previous section provide the basis for extending the AR procedure to multi-equation contexts. Consider a subset of the p-equation system (2.6),

$$y_i = Y_i\beta_i + X_{1i}\gamma_{1i} + u_i, \;\; i = 1\;, \;...\;, \;m, \tag{3.16}$$

where $m \leq p$. In this context, consider the problem of testing,

$$H_{MAR} : \beta_i = \beta_i^0, \; i = 1, ..., m. \tag{3.17}$$

Typically, when equations in (3.16) are viewed as a system, the first stage in an IV-type procedure consist in regressing each left-hand side endogenous variable on all the exogenous variables of the full sub-system. Conformably, let Z_2 refer to the set of exogenous variables that are excluded from all m equations, so that $X = \begin{bmatrix} Z_1 & Z_2 \end{bmatrix}$; then the first stage regression corresponds to

$$\overline{Y} = Z_1\Pi_1 + Z_2\Pi_2 + \overline{V}, \tag{3.18}$$

where $\overline{Y}$ includes all the distinct right-hand-side endogenous variables and the error term $\overline{V}$ is defined conformably; suppose that $\overline{Y}$ is $T \times \overline{m}$ and Z_1 is $T \times \overline{k}$. By definition, postmultiplying $\overline{Y}$ by a selection matrix (of zeros and ones) gives Y_i, which allows to decompose (3.18) as follows:

$$Y_i = Z_1\Pi_{1i} + Z_2\Pi_{2i} + V_i, \; i = 1, ..., m,$$

where V_i includes the relevant columns of $\overline{V}$, and Π_{1i} and Π_{2i} are the relevant sub-matrices of Π_1 and Π_2. Transform the system setting $y_i^0 = y_i - Y_i\beta_i^0, \; i = 1, ..., m$, as follows:

$$y_i^0 = Y_i(\beta_i - \beta_i^0) + Z_1\gamma_i + u_i \tag{3.19}$$

$$Y_i = Z_1\Pi_{1i} + Z_2\Pi_{2i} + V_i, \tag{3.20}$$

where γ_i may include zeros so that $Z_1\gamma_i = X_{1i}\gamma_{1i}$. This leads to the reduced form

$$\begin{aligned}\left[\begin{array}{ccc} y_1^0 & \dots & y_m^0 \quad \overline{Y}\end{array}\right] &= \left[\begin{array}{cc} Z_1 & Z_2\end{array}\right]\left[\begin{array}{cccc}\pi_{11} & \dots & \pi_{1m} & \Pi_1 \\ \pi_{21} & \dots & \pi_{2m} & \Pi_2\end{array}\right] \\ &\quad + \left[\begin{array}{cccc} u_1 + V_1(\beta_1-\beta_1^0) & \dots & u_m + V_m(\beta_m - \beta_m^0) & \overline{V}\end{array}\right], \\ \pi_{1i} &= \Pi_{1i}(\beta_i - \beta_i^0) + \gamma_i,\ \pi_{2i} = \Pi_{2i}(\beta_i - \beta_i^0),\end{aligned}$$

in which case (3.17) corresponds to testing:

$$\left[O_{(k-\overline{k},\,\overline{k})},\ I_{(k-\overline{k})}\right]\left[\begin{array}{cccc}\pi_{11} & \dots & \pi_{1m} & \Pi_1 \\ \pi_{21} & \dots & \pi_{2m} & \Pi_2\end{array}\right]\overline{C} = 0, \qquad \overline{C} = \left[\begin{array}{c} I_m \\ O_{(\overline{m},\, m)}\end{array}\right]. \tag{3.21}$$

These constraints do not consider the exclusions implied by the zeros in γ_i. Let $M_{Z_1} = I - Z_1(Z_1'Z_1)^{-1}Z_1'$ and $\overline{\Pi}_{AR} = \left[\begin{array}{cccc}\pi_{21} & \dots & \pi_{2m} & \Pi_2\end{array}\right]\overline{C}$. Then the test amounts to assessing $\overline{\Pi}_{AR} = 0$ in the context of:

$$\begin{aligned} M_{Z_1}\left[\begin{array}{cccc} y_1^0 & \dots & y_m^0 & \overline{Y}\end{array}\right]\overline{C} &= M_{Z_1}Z_2\overline{\Pi}_{AR} \\ &\quad + M_{Z_1}\left[\begin{array}{cccc} u_1 + V_1(\beta_1-\beta_1^0) & \dots & u_m + V_m(\beta_m-\beta_m^0) & \overline{V}\end{array}\right]\overline{C}\end{aligned}$$

Let $P_{M_{Z_1}Z_2} = I - M_{Z_1}Z_2(Z_2'M_{Z_1}X_2)^{-1}Z_2'M_{Z_1}$, then the LR statistic to test $\overline{\Pi}_{AR} = 0$ is

$$\begin{aligned}\Lambda_{MAR} &= \frac{\left|\overline{C}'\left[\begin{array}{cccc} y_1^0 & \dots & y_m^0 & \overline{Y}\end{array}\right]' M_{Z_1}\left[\begin{array}{cccc} y_1^0 & \dots & y_m^0 & \overline{Y}\end{array}\right]\overline{C}\right|}{\left|\overline{C}'\left[\begin{array}{cccc} y_1^0 & \dots & y_m^0 & \overline{Y}\end{array}\right]' M_{Z_1}P_{M_{Z_1}Z_2}M_{Z_1}\left[\begin{array}{cccc} y_1^0 & \dots & y_m^0 & \overline{Y}\end{array}\right]\overline{C}\right|}, \\ &= \frac{\left|\left[\begin{array}{ccc} y_1^0 & \dots & y_m^0\end{array}\right]' M_{Z_1}\left[\begin{array}{ccc} y_1^0 & \dots & y_m^0\end{array}\right]\right|}{\left|\left[\begin{array}{ccc} y_1^0 & \dots & y_m^0\end{array}\right]' M_{Z_1}P_{M_{Z_1}Z_2}M_{Z_1}\left[\begin{array}{ccc} y_1^0 & \dots & y_m^0\end{array}\right]\right|}.\end{aligned}$$

Theorem 3.2 DISTRIBUTION OF THE AR MULTIVARIATE TEST. *In the context of the subsystem* (3.16) *of the SE model* (2.1), *consider the problem of testing* (3.17)

$$H_{MAR} : \beta_i = \beta_i^0,\ i = 1, ..., m,$$

where, without loss of generality, the m-equations under test are the first m equations of the system so that

$$\left[\begin{array}{ccc} u_1 & \dots & u_m\end{array}\right] = UC,\ C = \left[\begin{array}{c} I_m \\ O_{(p-m,\, m)}\end{array}\right]$$

where U satisfies (2.3)*, with*

$$J = \left[\begin{array}{cc} J_{11} & 0 \\ J_{21} & J_{22}\end{array}\right],\ J_{11} : m \times m,\ J_{11} \text{ is nonsingular} \tag{3.22}$$

and W is partitioned conformably as follows

$$\begin{aligned} W &= \begin{bmatrix} W_1 & W_2 \end{bmatrix}, \; W_i = [W_{i1}, \ldots, W_{iT}]', \; i = 1, 2, \\ W_t &= \begin{bmatrix} W_{1t} \\ W_{2t} \end{bmatrix}, \; W_{1t} : m \times 1 . \end{aligned} \tag{3.23}$$

Let

$$\Lambda_{MAR} = \frac{\left| \begin{bmatrix} y_1^0 & \ldots & y_m^0 \end{bmatrix}' M_{Z_1} \begin{bmatrix} y_1^0 & \ldots & y_m^0 \end{bmatrix} \right|}{\left| \begin{bmatrix} y_1^0 & \ldots & y_m^0 \end{bmatrix}' M_{Z_1} P_{M_{Z_1} Z_2} M_{Z_1} \begin{bmatrix} y_1^0 & \ldots & y_m^0 \end{bmatrix} \right|}$$

be the associated multivariate Anderson-Rubin statistic. Then under the null hypothesis

$$P[\Lambda_{MAR} \geq x] = P\left[\frac{|W_1' M_{Z_1} W_1|}{\left| W_1' M_{Z_1} P_{M_{Z_1} Z_2} M_{Z_1} W_1 \right|} \geq x \right], \; \forall x.$$

PROOF. Under the null hypothesis,

$$\begin{aligned} \Lambda_{MAR} &= \frac{\left| \begin{bmatrix} u_1 & \ldots & u_m \end{bmatrix}' M_{Z_1} \begin{bmatrix} u_1 & \ldots & u_m \end{bmatrix} \right|}{\left| \begin{bmatrix} u_1 & \ldots & u_m \end{bmatrix}' M_{Z_1} P_{M_{Z_1} Z_2} M_{Z_1} \begin{bmatrix} u_1 & \ldots & u_m \end{bmatrix} \right|}, \\ \begin{bmatrix} u_1 & \ldots & u_m \end{bmatrix} &= UC = WJ'C = W \begin{bmatrix} J_{11}' \\ 0 \end{bmatrix} = W_1 J_{11}'. \end{aligned}$$

Substituting $W_1 J_{11}'$ for UC in Λ_{MAR} leads to

$$\begin{aligned} \Lambda_{MAR} &= \frac{\left| J_{11} W_1' M_{Z_1} W_1 J_{11}' \right|}{\left| J_{11} W_1' M_{Z_1} P_{M_{Z_1} Z_2} M_{Z_1} W_1 J_{11}' \right|} = \frac{|J_{11}| \, |W_1' M_{Z_1} W_1| \left| J_{11}' \right|}{|J_{11}| \left| W_1' M_{Z_1} P_{M_{Z_1} Z_2} M_{Z_1} W_1 \right| \left| J_{11}' \right|} \\ &= \frac{|W_1' M_{Z_1} W_1|}{\left| W_1' M_{Z_1} P_{M_{Z_1} Z_2} M_{Z_1} W_1 \right|}. \end{aligned}$$

This completes the proof.■

In this case as well, it easy to show, using the same arguments as in the above Theorem, that all the roots of the determinantal equation

$$\begin{aligned} \left| \mathcal{Y}_i^{0\prime} M_{Z_1} \mathcal{Y}_i^0 - \mu \, \mathcal{Y}_i^{0\prime} M_{Z_1} P_{M_{Z_1} Z_2} M_{Z_1} \mathcal{Y}_i^0 \right| &= 0 \\ \begin{bmatrix} y_1^0 & \ldots & y_m^0 \end{bmatrix} &= \mathcal{Y}_i^0 \end{aligned}$$

are pivotal under the null hypothesis, which lead to alternative statistics. The case where $m = p$ deserves a special attention, and leads to the full system approach.

3.3. Pivots in full systems

In the context of (2.1) with (2.3), consider testing $H_B : B = B_0$; recall that B includes normalization and exclusion restrictions (since all endogenous variables do not appear in all equations). These constraints may be tested by assessing the exclusion restrictions in the regression of YB_0 on X. Indeed, if we examine the reduced form (2.1), we see that H_B implies that the coefficient of

$$YB_0 = X\Pi B_0 + VB_0$$

should reflect the exclusion (identifying) restrictions in Γ. Typically, these exclusions are of the SURE type (i.e. they do not affect the coefficient of the same regressor for all equations), yet its is possible to obtain a pivot if we focus on assessing the exclusion of the common instruments.[8] This hypothesis takes the following form:

$$Q\Pi B_0 C = 0 \tag{3.24}$$

where Q and C are full-row rank and full column rank selection matrices.[9] Without loss of generality, suppose that

$$C = \begin{bmatrix} I_c \\ 0 \end{bmatrix}, \; J = \begin{bmatrix} J_{11} & 0 \\ J_{21} & J_{22} \end{bmatrix}, \; J_{11} \text{ is } c \times c, \text{ nonsingular,} \tag{3.25}$$

$$W_t = \begin{bmatrix} W_{1t} \\ W_{2t} \end{bmatrix}, \; W_{1t} : c \times 1\,. \tag{3.26}$$

Then

$$UC = WJ'C = W \begin{bmatrix} J_{11}' \\ 0 \end{bmatrix} = W_1 J_{11}', \;\; W_i = [W_{i1}, \ldots, W_{iT}]', \; i = 1, 2. \tag{3.27}$$

The LR statistic to test the latter hypothesis is:

$$\Lambda_B = \frac{\left|C'B_0'(Y - X\widehat{\Pi}_0)'(Y - X\widehat{\Pi}_0)B_0C\right|}{\left|C'B_0'(Y - X\widehat{\Pi})'(Y - X\widehat{\Pi})B_0C\right|}$$

where $\widehat{\Pi}_0$ and $\widehat{\Pi}$ are the constrained and unconstrained OLS estimates in the regression of YB_0C on X. Let $M = I - X(X'X)^{-1}X'$, $M_0 = M + X(X'X)^{-1}Q'[Q(X'X)^{-1}Q']^{-1}Q(X'X)^{-1}X'$. Then under the null hypothesis,

$$\Lambda_B = \frac{|C'B_0'V'M_0VB_0C|}{|C'B_0'V'MVB_0C|} = \frac{\left|C'B_0'\left(B_0^{-1}\right)'U'M_0UB_0^{-1}B_0C\right|}{\left|C'B_0'\left(B_0^{-1}\right)'U'MUB_0^{-1}B_0C\right|} = \frac{|C'U'M_0UC|}{|C'U'MUC|} \tag{3.28}$$

[8]If no common instruments are available, then exact bounds tests of the implied SURE constraints can be considered as in Dufour and Khalaf (2003).

[9]The matrix C allows to select-out the equations of the system that will not be subject to exclusion tests, *e.g.* the equations which in the first place did not include endogenous regressors.

$$= \frac{\left|J_{11}W_1'M_0W_1J_{11}'\right|}{\left|J_{11}W_1'MW_1J_{11}'\right|} = \frac{|J_{11}|\,|W_1'M_0W_1|\,\left|J_{11}'\right|}{|J_{11}|\,|W_1'MW_1|\,\left|J_{11}'\right|} = \frac{|W_1'M_0W_1|}{|W_1'MW_1|}. \tag{3.29}$$

No assumption on the distribution W_2 is required and the matrix J in (2.3) only needs to be block triangular. It is worth noting that hypotheses which test further common constraints on Γ in addition to fixing $B = B_0$ can be accommodated in the same way, by adjusting Q and C and allowing a non-zero matrix of known constants on the right hand side of (3.24). Pivots can also be obtained for such hypotheses, as is demonstrated in the following Theorem.

Theorem 3.3 CHARACTERIZATION OF PIVOTAL STATISTICS. *In the context of the SE model* (2.1) *consider the hypothesis which when written in terms of the reduced form* (2.2) *takes the form*

$$H_{ULB} : Q\Pi B_0 C = D \tag{3.30}$$

where Q is a $q \times k$ known matrix with rank q, D is known,

$$C = \begin{bmatrix} C_{11} \\ 0 \end{bmatrix}, \; C_{11} \textit{ is } c \times c \textit{ nonsingular,}$$

U satisfies (2.3) *with*

$$J = \begin{bmatrix} J_{11} & 0 \\ J_{21} & J_{22} \end{bmatrix}, \; J_{11} \textit{ is } c \times c\,, \textit{ nonsingular,} \tag{3.31}$$

and W_t is partitioned conformably

$$W_t = \begin{bmatrix} W_{1t} \\ W_{2t} \end{bmatrix}, \; W_{1t} : c \times 1\,.$$

Let

$$\Lambda_{ULB} = \frac{\left|C'B_0'(Y - X\widehat{\Pi}_0)'(Y - X\widehat{\Pi}_0)B_0C\right|}{\left|C'B_0'(Y - X\widehat{\Pi})'(Y - X\widehat{\Pi})B_0C\right|}$$

denote the LR statistic for testing the latter restrictions where $\widehat{\Pi}_0$ and $\widehat{\Pi}$ are the constrained and unconstrained OLS estimates in the regression of YB_0C on X. Then under the null hypothesis

$$P[\Lambda_{ULB} \geq x] = P\left[\frac{|W_1'M_0W_1|}{|W_1'MW_1|} \geq x\right], \; \forall x,$$

where $M = I - X(X'X)^{-1}X'$, $M_0 = M + X(X'X)^{-1}Q'[Q(X'X)^{-1}Q']^{-1}Q(X'X)^{-1}X'$ and $W_i = [W_{i1}\,, \ldots\,, W_{iT}]'\,,\; i = 1, 2.$

PROOF. Following (2.21)-(3.29), we see that under the null hypothesis,

$$\Lambda_{ULB} = \frac{|C'U'M_0UC|}{|C'U'MUC|}, \quad UC = WJ'C = W\begin{bmatrix} J_{11}'C_{11} \\ 0 \end{bmatrix} = W_1\, J_{11}'C_{11}$$

where $J_{11}'C_{11}$ is nonsingular. So

$$\Lambda_{ULB} = \frac{\left|C_{11}'J_{11}W_1'M_0W_1\, J_{11}'C_{11}\right|}{\left|C_{11}'J_{11}W_1'MW_1\, J_{11}'C_{11}\right|} = \frac{|C_{11}'J_{11}|\,|W_1'M_0W_1\,|\,\left|J_{11}'C_{11}\right|}{|C_{11}'J_{11}|\,|W_1'MW_1\,|\,|J_{11}'C_{11}|} = \frac{|W_1'M_0W_1|}{|W_1'MW_1|}.$$

This completes the proof.■

The same arguments as in the above Theorem show that all the roots of the determinantal equation

$$\left|C'U'M_0UC - \mu\, C'U'MUC\right| = 0$$

are also pivotal under the null hypothesis. The above derivations show that pivotal statistics can be obtained for all hypotheses of the form (3.30); these constraints are Uniform Linear; see Dufour and Khalaf (2002) and Berndt and Savin (1977). Here we show that pivots obtain when the coefficients of the left-hand side endogenous variables of the equations subject to test are all fixed. Indeed, since the error term of the reduced form equals UB^{-1}, the framework differs from Dufour and Khalaf (2002): invariance to J obtains when B is fixed (to allow the decomposition in (3.28)). One exception is noteworthy, and is stated in the following Theorem.

Theorem 3.4 PIVOTAL STATISTICS: A SPECIAL CASE. *Consider the MLR model* (2.1) *with* (2.3) *and the hypothesis which when written in terms of* (2.2) *takes the form*

$$H_{UL} : Q\Pi C = D \tag{3.32}$$

where C is an invertible $p \times p$ matrix, Q is a $q \times k$ known matrix with rank q and D is known. Let

$$\Lambda_{UL} = \frac{\left|C'(Y - X\widehat{\Pi}_0)'(Y - X\widehat{\Pi}_0)C\right|}{\left|C'(Y - X\widehat{\Pi})'(Y - X\widehat{\Pi})C\right|}$$

be the LR statistic for testing the latter restrictions, where $\widehat{\Pi}_0$ and $\widehat{\Pi}$ are the constrained and unconstrained OLS estimates in the regression of YC on X. Then under the null hypothesis

$$P[\Lambda_{UL} \geq x] = P\left[\frac{|W'M_0W|}{|W'MW|} \geq x\right], \ \forall x,$$

where $M = I - X(X'X)^{-1}X'$, $M_0 = M + X(X'X)^{-1}Q'[Q(X'X)^{-1}Q']^{-1}Q(X'X)^{-1}X'$ and W is as defined in (2.3).

PROOF. Under the null hypothesis,

$$\begin{aligned}\Lambda_{UL} &= \frac{|C'V'M_0VC|}{|C'V'MVC|} = \frac{\left|C'\left(B^{-1}\right)'U'M_0UB^{-1}C\right|}{\left|C'\left(B^{-1}\right)'U'MUB^{-1}C\right|} \\ &= \frac{\left|C'\left(B^{-1}\right)'\right|\left|U'M_0U\right|\left|B^{-1}C\right|}{\left|C'\left(B^{-1}\right)'\right|\left|U'MU\right|\left|B^{-1}C\right|} = \frac{|U'M_0U|}{|U'MU|} = \frac{|JW'M_0WJ'|}{|JW'MWJ'|} \\ &= \frac{|J|\,|W'M_0W|\,|J'|}{|J|\,|W'MW|\,|J'|} = \frac{|W'M_0W|}{|W'MW|}.\end{aligned}$$

This completes the proof. An example of the latter case in the LI context includes the problem where Π_{2i} is tested in addition to β_i.■

We emphasize again that the above results do not require the normality assumption. Eventually, when the normality hypothesis (2.5) holds, the distribution of the bounding statistic for special cases of Q and C is well known (see Rao (1973, chapter 8), Anderson (1984, chapters 8 and 13) and the appendix of Dufour and Khalaf (2002)) and involves the product of p independent *beta* variables with degrees of freedom that depend on the sample size, the number of restrictions and the number of parameters involved in these restrictions. For example, when $C = I_p$,

$$P[\Lambda_{NL}^{-1} \geq x] = P\left[\mathcal{L} \geq x\right],\ \forall x, \tag{3.33}$$

where $\mathcal{L}$ *i*s distributed like the product of p independent beta variables with parameters $(\frac{1}{2}(n-k-p+i)\,,\,\frac{q}{2})\,,\ i=1,\ \ldots\,,\ p$. When $c = 1$,

$$[\Lambda_{UL} - 1]\frac{n-k}{q} \sim F(q, n-k). \tag{3.34}$$

4. General Hypotheses tests on structural coefficients

In this section, we consider hypotheses for which pivots are not available. These hypotheses may be linear or non-linear, and may be approached from a full or sub-system approach. We first consider the full system case which will lead to useful results for the single equation problem.

4.1. The full system approach

Consider the problem of testing arbitrary restrictions on the structural parameters of model (2.1), under (2.3), which when expressed in terms of the reduced form coefficients, take the form

$$H_{NL} : R\pi \in \Delta_0, \tag{4.1}$$

where R is $(r \times kp)$ of rank r and Δ_0 is a non-empty subset of $\Re^r$. This characterization of the hypothesis allows for nonlinear as well as inequality constraints. The Gaussian QLR criterion to test H_{NL} is $n\ln(\Lambda_{NL})$,where

$$\Lambda_{NL} = \frac{|\hat{\Sigma}^{NL}|}{|\hat{\Sigma}|}\,, \tag{4.2}$$

with $\hat{\Sigma}^{NL}$ and $\hat{\Sigma}$ being the restricted and unrestricted ML estimators of Σ; in the statistics literature, Λ_{NL}^{-1} corresponds to Wilks' criterion. The discussion in the previous section does not lead to pivotal statistics for these hypotheses, yet we will show that Λ_{NL} is boundedly pivotal, in the sense of Dufour (1997), *i.e.* its null distribution can be bounded by a pivotal quantity; see Dufour and Khalaf (2002). To do this, we first observe that the general hypothesis (4.1) always admits as a special case, some hypothesis for which a pivot exists; indeed, the case where all the coefficients of the reduced form equation are restricted provides a trivial case which always satisfies our purpose. To relate our results with Dufour (1997), consider this special case

$$H_L : \Pi = D, \tag{4.3}$$

which obtains as in (3.32) with the further restriction that $Q = I_k$. Clearly, $H_L \subseteq H_{NL}$. In general, its is also possible to find a hypothesis of the form (3.30) which is special case of H_{NL}. Let $H_{ULB} \subseteq H_{NL}$ denote the hypothesis of the latter form which obtains from H_{NL} with the least number of restrictions.

Theorem 4.1 BOUNDELDY PIVOTAL STATISTICS. *Consider the MLR model* (2.1) *and let* Λ_{NL} *be the statistic defined by* (4.2) *for testing restrictions which, when written in terms of the reduced form* (2.2), *take the form* (4.1). *Further, consider restrictions of the form* (3.30) $H_{NL} : Q\Pi B_0 C = D$ *where* Q *is a* $q \times k$ *known matrix with rank* q, D *is known,*

$$C = \begin{bmatrix} C_{11} \\ 0 \end{bmatrix}, \; C_{11} \textit{ is } c \times c \textit{ (nonsingular)}$$

and Q, B_0 *and* C_{11} *are chosen such that* $H_{UL} \subseteq H_{NL}$. *Then under the null hypothesis imposing* (2.3) *with*

$$J = \begin{bmatrix} J_{11} & 0 \\ J_{21} & J_{22} \end{bmatrix}, \; J_{11} \textit{ is } c \times c \textit{ (nonsingular)}$$

and $W_t = \begin{bmatrix} W_{1t} \\ W_{2t} \end{bmatrix}, \; W_{1t} : p_1 \times 1,$

$$P[\Lambda_{NL} \geq x] \leq P\left[\frac{|W_1' M_0 W_1|}{|W_1' M W_1|}\right], \; \forall x,$$

where $M = I - X(X'X)^{-1}X'$, $M_0 = M + X(X'X)^{-1}Q'[Q(X'X)^{-1}Q']^{-1}Q(X'X)^{-1}X'$ *and* $W_i = [W_{i1}, \ldots, W_{iT}]'$, $i = 1, 2$.

PROOF. Let Λ_{ULB} be the reciprocal of Wilks' criterion for testing H_{ULB}. Since by construction $H_{ULB} \subseteq H_{NL}$, and since both Λ_{UL} and Λ_{NL} use the URF as the unconstrained hypothesis, then it is straightforward to see that $\Lambda_{NL} \leq \Lambda_{ULB}$. The null distribution of Λ_{ULB} was established in Theorem **3.3**, which leads to above bound.■

If we consider the bound associated with (3.32), and we further impose normality, then using (3.34) leads to the results of Dufour (1997).

Theorem 4.2 BOUNDELDY PIVOTAL STATISTICS: A SPECIAL CASE. *Consider the MLR model* (2.1) *and let* Λ_{NL} *be the statistic defined by* (4.2) *for testing restrictions which, when written in terms of the reduced form* (2.2)*, take the form* (4.1)*. Then under the null hypothesis imposing* (2.3) *and normal errors*

$$P[\Lambda_{NL}^{-1} \geq x] \leq P\left[\mathcal{L} \geq x\right], \; \forall x,$$

where $\mathcal{L}$ *is distributed like the product of* p *independent beta variables with parameters* $(\frac{1}{2}(n-k-p+i)\,,\,\frac{k}{2})\,,\; i=1,\,\ldots\,,\,p.$

PROOF. Consider restrictions of the form (4.3) $H_L : \Pi = D$, and let Λ_L be the reciprocal of Wilks' criterion for testing H_L. Following the arguments of Theorem **4.1**, we see that $\Lambda_{NL} \leq \Lambda_L$. The null distribution of Λ_L^{-1} obtains as a special case of (3.33) with $q = k$, which leads to above beta-based bound.■

Since Dufour (1997)'s bound was formally stated in the context of a LI model, let us turn the LI context.

4.2. The LI context

Let us first consider the case of the LIML LR statistic associated with $H_{AR} : \beta_i = \beta_i^0$, in the context of the LI model (2.7). Wang and Zivot (1998) have shown that this statistic is a monotonic transformation of

$$\Lambda_{LIML} = \lambda(\beta_i^0) - \lambda(\widetilde{\beta}_i)$$

where $\lambda(\beta_i)$ is defined in (2.16) and $\widetilde{\beta}_i$ is the LIML estimate of β defined in (2.17). Recall that $\lambda(\widetilde{\beta}_i) = \min_{\beta_i}\{\lambda(\beta_i)\}$ and $\lambda(\beta_i^0) = \Lambda_{AR}$ as defined in (3.5). It is thus easy to see that $\Lambda_{LIML} \leq \Lambda_{AR}$, so under the null hypothesis, using Theorem **3.1**, we have:

$$P[\Lambda_{LIML} \geq x] \leq P\left[\frac{|w_i' M_{1i} w_i|}{|w_i' M_{1i} P_{M_{1i}X_{2i}} M_{1i} w_i|} \geq x\right], \; \forall x, \tag{4.4}$$

where $w_i = \begin{pmatrix} w_1^i & w_2^i & \ldots & w_n^i \end{pmatrix}'$ gives the first column of W^i as defined in (2.11)-(2.13). If the normality hypothesis is further imposed, then

$$P\left[[\Lambda_{LIML} - 1]\frac{n-k}{k-k_i} \geq x\right] \leq P\left[F(k-k_i, n-k) \geq x\right], \; \forall x.$$

Whereas $n[\ln(\Lambda_{LIML})]$ has a $\chi^2(m_i)$ asymptotic distribution only under identification assumptions, $n[\ln(\Lambda_{AR})]$ is asymptotically distributed as $\chi^2(k-k_i)$ whether the rank condition holds or not. The above inequality implies that the asymptotic distribution of the LR-LIML statistic is thus bounded by a $\chi^2(k-k_i)$ distribution independently of the conditions for identification. This result was derived under *local-to-zero* asymptotics in Wang and Zivot (1998). Our result also shows that using the LR-LIML in this context will lead to power losses compared to the AR criterion.

Consider the problem of testing arbitrary restrictions on the parameters of model (2.7), under (2.11), which when expressed in terms of the reduced form (2.8), take the form

$$H_{NL} : R\pi_i \in \Delta_0, \tag{4.5}$$

where R is $(r \times km_i)$ of rank r and Δ_0 is a non-empty subset of $\Re^r$ and $\pi_i = vec(\Pi_i)$. The Gaussian QLR criterion to test H_{NL} is $n\ln(\Lambda_{NL})$,where

$$\Lambda_{NL} = \frac{|\hat{\Sigma}_i^{NL}|}{|\hat{\Sigma}_i|}\ , \tag{4.6}$$

with $\hat{\Sigma}_i^{NL}$ and $\hat{\Sigma}_i$ being the restricted and unrestricted ML estimators of Σ_i; note that the denominator is completely unconstrained, i.e. does not reflect the LIML exclusion restrictions. As in the full system approach, we first observe that the general hypothesis (4.5) always admits as a special case, some hypothesis for which a pivot exists; indeed, the case where all the coefficients of the LI reduced form equation are restricted

$$H_L : \Pi_i = D, \tag{4.7}$$

provides such a trivial example: clearly, $H_L \subseteq H_{NL}$.

Theorem 4.3 BOUNDELDY PIVOTAL LI STATISTICS: A SPECIAL CASE. *Consider the MLR model* (2.7)-(2.8) *and let* Λ_{NL} *be the statistic defined by* (4.6) *for testing restrictions which, when written in terms of the reduced form* (2.8)*, take the form* (4.5)*. Then under the null hypothesis imposing* (2.11) *and normal errors*

$$P[\Lambda_{NL} \geq x] \leq P\left[\frac{|W^{i\prime}W^i|}{|W^{i\prime}MW^i|} \geq x\right],\ \forall x,$$

where $M = I - X(X'X)^{-1}X'$*, and* W^i *is as defined in* (2.3); *imposing normal errors we further obtain that*

$$P[\Lambda_{NL}^{-1} \geq x] \leq P\left[\mathcal{L} \geq x\right],\ \forall x,$$

where $\mathcal{L}$ *is distributed like the product of* m_i+1 *independent beta variables with parameters* $(\frac{1}{2}(n - k - (m_i+1) + i)\ ,\ \frac{k}{2})\ ,\ i = 1,\ \ldots\ ,\ m_i + 1.$

PROOF. Let Λ_L be the reciprocal of Wilks' criterion for testing H_L applied to the LI context. Following the arguments of Theorem **4.1**, we see that $\Lambda_{NL} \leq \Lambda_L$. The null distribution of Λ_L obtains as in **3.4**, applied to the LI context. The normal case also derives from (3.33) with $q = k$ and $p = m_i$ which leads to above beta-based bound.■

The normal case is exactly the same result obtained in Dufour (1997). Following the reasoning explicated for our full system approach, tighter bounds can be obtained by a proper choice of the linear hypothesis which is a special case of (4.5). As an illustration, let us consider the important special case where restrictions in (4.5) only affect δ_i, the coefficients of the structural equation. In this case, it is possible to find a linear hypothesis of the form (3.10) which is a special case of the hypothesis under test. Then using the same arguments underlying Theorem **4.1** and the distributional result (3.11) yields the following bounds test procedure.

Theorem 4.4 BOUNDEDLY PIVOTAL LI STATISTICS. *Consider the problem of testing arbitrary restrictions on the structural parameters of model* (2.7) *under* (2.11) *of the form*

$$H_{NLS} : R\delta_i \in \varDelta_0, \tag{4.8}$$

where R is $r \times (m_i + k_i)$ of rank r and $\varDelta_0$ is a non-empty subset of $\Re^r$. The Gaussian QLR criterion to test H_{NLS} is $n\ln(\varLambda_{NLS})$,where

$$\varLambda_{NLS} = \frac{|\hat{\varSigma}_i^{NLS}|}{|\hat{\varSigma}_i|}\ , \tag{4.9}$$

with $\hat{\varSigma}_i^{NLS}$ and $\hat{\varSigma}_i$ being the restricted and unrestricted ML estimators of $\varSigma_i$. Consider a hypothesis of the form (3.10) *which is a special case of* (4.8)

$$H_{ARQX} : \beta_i = \beta_i^0,\ \ Q_{1i}\gamma_{1i} = \nu_0 \subseteq H_{NL}, \tag{4.10}$$

where Q_{1i} is a $q_{1i} \times k_i$ matrix with $q_{1i} = rank(Q_{1i})$; Q_{1i} can be treated as submatrix of an invertible $k_i \times k_i$ matrix $Q_i = \begin{bmatrix} Q_{1i}' & Q_{2i}' \end{bmatrix}'$ so that

$$Q_i\gamma_{1i} = \begin{bmatrix} Q_{1i}\gamma_{11i} \\ Q_{2i}\gamma_{21i} \end{bmatrix} = \begin{bmatrix} \nu_{1i} \\ \nu_{2i} \end{bmatrix}.$$

Let $X_{Q_i} = X_{1i}Q_i^{-1} = \begin{bmatrix} X_{Q_{1i}} & X_{Q_{2i}} \end{bmatrix}$ where $X_{Q_{1i}}$ and $X_{Q_{2i}}$ are $T \times q_{1i}$ and $T \times (k_i - q_{1i})$ matrices. Then imposing (2.11) *where the first row of J_i has zeros everywhere except for the first element,*

$$P[\varLambda_{NLS} \geq x] \leq P\left[\frac{|w_i'M_{Q_{2i}}w_i|}{\left|w_i'M_{Q_{2i}}P_{M_{Q_{2i}}X_{22i}}M_{Q_{2i}}w_i\right|} \geq x\right],\ \forall x,$$

$$\begin{aligned} M_{Q_{2i}} &= I - X_{Q_{2i}}(X_{Q_{2i}}'X_{Q_{2i}})^{-1}X_{Q_{2i}}',\ \ X_{22i} = \begin{bmatrix} X_{Q_{1i}} & X_{2i} \end{bmatrix} \\ P_{M_{Q_{2i}}X_{22i}} &= I - M_{Q_{2i}}X_{22i}(X_{22i}'M_{Q_{2i}}X_{22i})^{-1}X_{22i}'M_{Q_{2i}}. \end{aligned}$$

where $w_i = \begin{pmatrix} w_1^i & w_2^i & \ldots & w_n^i \end{pmatrix}'$ gives the first column of W^i as defined in (2.11). *Imposing normality, we further obtain*

$$P\left[\varLambda_{NLS} - 1]\frac{n-k}{k-k_i-q_{1i}} \geq x\right] \leq P\left[F(k-k_i-q_{1i}, n-k) \geq x\right].$$

Note that the LR statistics considered use an unconstrained MLR as the alternative hypothesis.

An alternative statistic which considers the exclusion constraints can also be considered and will admit the same bound; see e.g. (4.4); indeed, by construction, the LIML-constrained statistic is larger than its unconstrained-alternative counterpart. However, this also means that bounds-tests should be based on the latter.

5. Simulation based pivotal and bounds tests

As is evident from the above results, the exact distributional results we have derived typically involve non-standard distributions, even in some Gaussian based contexts. However, they can be easily obtained using the MC method; see Dufour (2002), Dufour and Khalaf (2002). In the following, we describe the methodology in full and LI systems. To facilitate the presentation, in what follows: (i) S denotes the statistic considered, (ii) $\mathcal{W}$ refers to W in (2.3) or W^i in (2.11), (ii) X denotes the exogenous variables used for the test including instruments, and (iv) the number of MC draws N is obtained so that $\alpha(N+1)$ is an integer, where $0 < \alpha < 1$ is the level of the test.

Let us first consider the case where S is pivotal, i.e. $S = S(\mathcal{W}, X)$, where $S(\mathcal{W}, X)$ refers to the pivotal expression of S under the null hypothesis, as in Theorems **3.1 - 3.4**. Let $S^{(0)}$ denote the test statistic calculated from the observed sample; generate N of replications $S^{(1)}, \ldots, S^{(N)}$ of S which satisfy the null hypothesis, using draws from the null distribution of $\mathcal{W}$ and $S(\mathcal{W}, X)$. Compute $\hat{p}_N[S] \equiv p_N(S^{(0)}\,;\, S)$, where

$$p_N(x\,;\, S) \equiv \frac{NG_N(x\,;\, S)+1}{N+1}\,, G_N(x\,;\, S) \equiv \frac{1}{N}\sum_{i=1}^{N} s\big(S^{(i)} - x\big)\,, \tag{5.1}$$

and $s(x) = 1$ if $x \geq 0$, and $s(x) = 0$ if $x < 0$. In other words, $p_N(S^{(0)}\,;\, S) = [N\widehat{G}_N(S^{(0)}) + 1]/(N+1)$ where $N\widehat{G}_N(S^{(0)})$ is the number of simulated values which are greater than or equal to $S^{(0)}$. The MC critical region is $p_N(S^{(0)}\,;\, S) \leq \alpha$, where, under the null hypothesis, $\mathsf{P}\big[p_N(S^{(0)}\,;\, S) \leq \alpha\big] = \alpha$; see Dufour (2002). To avoid confusion, we refer to p-values based on the latter method as *Pivotal MC* (PMC) p-values.

If S is nuisance parameter dependant but boundedly pivotal, let $\overline{S}(\mathcal{W}, X)$ refer to the pivotal expression of the relevant bound under the null hypothesis, as in Theorems **4.1 - 4.3**. The associated MC procedure applies as in the PMC case, where $S^{(1)}, \ldots, S^{(N)}$ are obtained using $\overline{S}(\mathcal{W}, X)$; here, (5.1) leads to a level correct MC p-value which we denote *Bounds MC* (BMC) p-value, such that $\mathsf{P}\big[p_N(S^{(0)}\,;\, S) \leq \alpha\big] \leq \alpha$; see Dufour (2002) and Dufour and Khalaf (2002).

When S depends on nuisance parameters (say θ), a MC p-value, conditional on θ which we will denote $\widehat{p}_N(S|\theta)$ may be obtained as follows. Let $S^{(0)}$ denote the test statistic calculated from the observed sample; generate N of replications $S^{(1)}, \ldots, S^{(N)}$ of S given θ, using draws from the simulated model under the null hypothesis. Applying (5.1) yields a conditional MC p-value $\widehat{p}_N\,[S|\theta]$. The (standard) parametric bootstrap (denoted *Local MC* (LMC)) corresponds to the case where a consistent estimate of θ (compatible with the null hypothesis), say $\widehat{\theta}$, is used in the latter procedure. The MMC method involves maximizing $\widehat{p}_N\,[S|\theta]$ over all values of θ compatible with the null hypothesis, which provides a numerical search for the tightest bound available.

It is evident that for all $0 \leq \alpha \leq 1$ and $\forall\widehat{\theta}$, if the LMC p-value exceeds α, then the MMC p-value will also exceed α. This means that non-rejections in the context of LMC tests may be interpreted ”exactly”, with reference to the MMC test. Furthermore, if the BMC p-value is less than α, then we can be sure that the MMC p-value is also less than α. Since the BMC procedure is numericall less expensive than MMC, we recommend the following sequential procedure (with level α). Obtain a BMC p-value first and reject the null hypothesis if the BMC p-value is $\leq \alpha$. If not, obtain an LMC p-value using the constrained QMLE of θ. If the LMC p-value exceeds α, then conclude the test is not significant. Otherwise, run an MMC algorithm.

6. A Simulation study

This section reports an investigation, by simulation, of the performance of the various proposed test procedures. We focus on the LI examples. In each case, we also study 2SLS-based Wald tests, which are routinely computed in empirical practice. The asymptotic and MC test versions of the latter tests are considered. Since a bound is not available for these tests, we focus on the LMC and MMC tests. Each experiment was based on 1000 replications. We use Simulated Annealing to obtain the maximal p-values. The MC tests are applied with 99 replications.

The experiments are based on the LI model (2.7). We consider three endogenous variables ($p_i = 3$ and $m_i = 2$) and $k = 3,\ 4,\ 5$ and 6 exogenous variables. In all cases, the structural equation includes only one exogenous variable, the constant regressor. In the following tables, $d = (k-1) - (p-1)$ refers to the *degree of over-identification.* The restrictions tested are of the form precisely, we consider in turn: hypotheses which set the full vector of endogenous variables coefficients i.e. of the the form: (3.1), and hypotheses which set a subset of endogenous variables coefficients of the the form:

$$\beta_{1i} = \beta_{1i}^{0}, \tag{6.1}$$

where $\beta_i = (\beta_{1i}', \beta_{2i}')'$ and β_{1i} is $m_{1i} \times 1$, with $m_{1i} = 1$. The sample sizes are set to $n = 25$, 50, 100. The exogenous regressors are independently drawn from the normal distribution, with means zero and unit variances. These were drawn only once. The errors were generated according to a multinormal distribution with mean zero and covariance

$$\Sigma_i = \begin{bmatrix} 1 & .95 & -.95 \\ .95 & 1 & -1.91 \\ -.95 & -1.91 & 12 \end{bmatrix} \tag{6.2}$$

The other coefficients were

$$\gamma_{1i} = 1,\ \beta_i = (10, -1.5)',\ \Pi_{1i} = (1.5, 2)',\ \Pi_{2i} = \begin{bmatrix} \tilde{\Pi} \\ O_{(k-3,2)} \end{bmatrix}, \tag{6.3}$$

The identification problem becomes mores serious as the determinant of $\Pi_2'\Pi_2$ gets closer to zero. In view of this, we consider various choices for $\tilde{\Pi}$:

$$\tilde{\Pi}_{(1)} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix},\ \tilde{\Pi}_{(2)} = \begin{bmatrix} 2 & 1.999 \\ 1.999 & 2 \end{bmatrix},$$

$$\tilde{\Pi}_{(3)} = \begin{bmatrix} .5 & .499 \\ .499 & .5 \end{bmatrix}, \quad \tilde{\Pi}_{(4)} = \begin{bmatrix} .01 & .009 \\ .009 & .01 \end{bmatrix}.$$

We examine LR statistics which use an unconstrained MLR as the alternative hypothesis and their counterpart which considers the LIML exclusion constraints. For convenience and clarity, the former are denoted LR_{OLS} and the latter LR_{LIML}. We also consider Wald statistics of the form (2.22) based on LIML and 2SLS estimators and denote these statistics Wald_{LIML} and Wald_{2SLS} respectively. We report the probability of Type I error for the standard asymptotic χ^2 test, and the LMC, MMC and BMC based procedures. The subscripts asy, LMC, MMC and BMC which appear in the subsequent Tables are used to identify these procedures respectively. In the case of the statistic LR_{OLS} under (3.1), the local MC test is denoted PMC to account for the fact that the test is exact since the statistic is pivotal. We have also examined the generalized Wang and Zivot (1998) asymptotic bounds tests to which we refer as BND_z. We perform a power study by varying the value of β_1 away from the null value of 10 and given $\tilde{\Pi}_{(1)}$, for the tests which size was adequate.

To generate the simulated samples in the LMC case, we consider the restricted LIML estimates of the parameters that are not specified by the null, except for the Wald_{2SLS} statistic. In this case, we use restricted 2SLS estimates for the structural equation and OLS based estimates for reduced form equations which complement the system. From these estimates, sum-of-squared-residuals are constructed which yield the usual estimate covariance estimate. Furthermore, to ensure the complementarity of the MMC and the bounds procedures, the exact bounds are obtained by simulation (we do not use the F distribution). Tables 1-5 summarize our findings. Our results show the following.

1. Identification problems severely distort the sizes of standard asymptotic tests. While the evidence of size distortions is notable even in identified models, the problem is far more severe in near-unidentified situations. The results for the Wald test are especially striking: empirical sizes exceeding 80 and 90% were observed! More importantly, increasing the sample size does not correct the problem. This result substantiates so-called "weak instruments" effects. The asymptotic LR behaves more smoothly in the sense that size distortions are not as severe; still some form of size correction is most certainly called for.

2. The performance of the standard bootstrap is disappointing. In general, the empirical sizes of LMC tests exceed 5% in most instances, even in identified models. In particular, bootstrap Wald tests fail completely in near-unidentified conditions.

3. Whether the rank condition for identification is imposed or not, more serious size distortions are observed in over-identified systems. This holds true for asymptotic and bootstrap procedures. While the problems associated with the Wald tests conform to general expectations, it is worth noting that the traditional bootstrap does not completely correct the size of LR tests.

4. *In all cases, the Wald tests maximal randomized p-values are always one.* This meant that under the null and the alternative, MMC empirical rejections were always zero (this result, for space considerations, is not reported in the Tables).

5. The bounds tests and the MMC tests achieve size control in all cases. The strategy of resorting to MMC when the bounds test is not conclusive would certainly pay off, for the critical bound

is easier to compute. However, it is worth noting that although the MMC are thought to be computationally burdensome, the SA maximization routine was observed to converge quite rapidly irrespective of the number of intervening nuisance parameters.

6. The LIML-LMC performs generally better than the generalized Wang and Zivot (1998) asymptotic bounds tests. Observe however that the LMC test is not exactly size correct, whereas Wang and Zivot (1998)'s tests sizes were not observed to exceed 5%. In situations were size was adequate, the LMC test showed superior power.

7. The performance of the Wald-LIML LMC test may seem acceptable, although the above remark in the case of the MMC p-value also holds in this case. As expected, power losses with respect to the LR test are noted. It is worth noting that since constrained and unconstrained MLE is done analytically, there seems to be arguments in favor of a Wald test if a LIML approach is considered.

The above findings mean that 2SLS-based tests are inappropriate in the weak instrument case and cannot be corrected by bootstrapping. Much more reliable tests will be obtained by applying the proposed LR-based procedures. The usual arguments on computational inconveniences should not be overemphasized. With the increasing availability of more powerful computers and improved software packages, there is less incentive to prefer a procedure on the grounds of execution ease.

7. Conclusion

The serious inadequacy of standard asymptotic tests in finite samples is widely observed in the SE context. Here, we have proposed alternative, simulation-based procedures and demonstrated their feasibility in an extensive Monte Carlo experiment. Particular attention was given to the identification problem. By exploiting MC methods and using these in combination with bounds procedures, we have constructed provably exact tests for arbitrary, possibly nonlinear hypotheses on the systems coefficients. We have also investigated the ability of the conventional bootstrap to provide more reliable inference in finite samples. The simulation results show that the latter fails when the simulated statistic is IV-based. In the case of the LR criteria, although the bootstrap did reduce the error in level, it did not achieve size control. In contrast, MMC LR tests perfectly controlled levels. The exact randomized procedures are computer intensive; however, with modern computer facilities, computational costs are no longer a hindrance.

References

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

Anderson, T. W. and Rubin, H. (1949), 'Estimation of the parameters of a single equation in a complete system of stochastic equations', *Annals of Mathematical Statistics* **20**, 46–63.

Table 1. Empirical P(Type I error): Testing a subset of endogenous variables coefficients, LR tests.

			LR_{LIML}					LR_{OLS}				
d	$\widetilde{\Pi}$	n	Asy	LMC	MMC	BMC	BD_Z	Asy	LMC	MMC	BMC	BD_Z
1	$\widetilde{\Pi}_{(1)}$	25	7.5	5.3	0.8	0.8	1.3	6.1	3.8	2.3	2.3	3.2
		50	7.9	5.3	0.4	0.2	0.1	4.9	5.2	1.7	1.7	2.2
		100	6.3	5.1	0.8	0.8	0.7	5.8	4.9	2.4	2.5	2.4
2		25	10.2	5.9	0.4	0.4	0.8	5.9	5.3	1.1	1.1	1.6
		50	8.9	5.7	0.8	0.7	0.4	5.8	4.7	2.6	2.6	2.9
		100	6.4	4.5	0.3	0.2	0.4	5.2	5.0	1.6	1.6	2.0
3		25	14.9	6.8	0.6	0.6	0.9	8.2	4.5	2.3	2.3	5.2
		50	9.8	5.0	0.2	0.2	0.1	6.3	3.9	1.9	1.9	3.1
		100	7.4	5.1	0.2	0.1	0.0	4.8	4.5	1.7	1.7	2.7
1	$\widetilde{\Pi}_{(2)}$	25	14.2	7.1	1.7	1.6	3.0	7.5	4.9	2.3	2.1	4.3
		50	12.7	5.6	1.1	1.1	1.2	5.2	4.5	1.6	1.6	2.0
		100	12.0	6.1	1.5	1.5	1.6	6.1	5.5	2.0	1.9	2.5
2		25	20.0	7.8	1.2	1.1	2.4	7.0	4.4	2.6	2.6	4.2
		50	17.0	6.7	1.8	1.5	1.7	6.4	4.4	2.6	2.6	2.9
		100	15.6	6.1	0.9	0.9	0.9	5.1	4.5	1.6	1.6	2.1
3		25	22.3	8.9	1.4	1.0	2.4	8.8	5.7	3.3	3.3	5.9
		50	23.6	8.6	0.9	0.8	1.4	7.1	4.4	2.5	2.5	4.1
		100	21.0	6.4	1.2	1.0	1.1	5.4	4.1	2.1	2.1	2.2
1	$\widetilde{\Pi}_{(3)}$	25	2.4	2.0	0.2	0.2	0.1	1.5	1.7	0.3	0.3	0.8
		50	3.4	3.9	0.4	0.4	0.4	1.6	2.6	0.5	0.5	0.6
		100	6.0	5.4	0.5	0.5	0.5	2.8	3.7	7.0	7.0	0.8
2		25	4.3	3.9	0.0	0.0	0.0	1.5	1.2	0.3	0.3	0.2
		50	6.6	5.2	0.3	0.3	0.0	1.6	2.4	0.7	0.7	0.6
		100	7.2	5.4	0.1	0.1	0.2	1.8	2.4	0.4	0.4	0.5
3		25	6.4	3.8	0.0	0.0	0.0	2.2	1.4	0.6	0.6	0.1
		50	6.3	4.6	0.1	0.0	0.1	1.0	0.9	0.1	0.1	0.3
		100	10.9	7.1	0.1	0.1	0.2	1.4	2.1	0.7	0.7	0.7
1	$\widetilde{\Pi}_{(4)}$	25	2.1	1.3	0.1	0.1	0.1	0.6	1.0	0.3	0.3	0.4
		50	1.8	1.3	0.0	0.0	0.0	0.2	0.1	0.0	0.0	0.0
		100	2.5	1.6	0.0	0.0	0.0	0.1	0.1	0.0	0.0	0.0
2		25	5.4	1.8	0.0	0.0	0.0	0.5	0.8	0.3	0.2	0.2
		50	0.4	1.1	0.0	0.0	0.0	0.4	0.6	0.0	0.0	0.0
		100	1.7	1.7	0.0	0.0	0.0	0.1	0.5	0.0	0.0	0.0
3		25	9.0	2.6	1.0	1.0	0.1	1.4	1.0	0.3	0.3	0.8
		50	5.8	2.0	0.0	0.0	0.0	0.7	0.8	0.0	0.0	0.2
		100	0.4	1.3	0.1	0.1	0.0	0.4	0.5	0.2	0.2	0.1

Table 2. Empirical P(Type I error): Testing a subset of endogenous variables coefficients, Wald tests.

			Wald - 2SLS		Wald - LIML					Wald - 2SLS		Wald - LIML	
d	$\widetilde{\Pi}$	n	Asy	LMC	Asy	LMC	d	$\widetilde{\Pi}$	n	Asy	LMC	Asy	LMC
1	$\widetilde{\Pi}_{(1)}$	25	8.6	5.8	8.3	3.9	1	$\widetilde{\Pi}_{(3)}$	25	10.9	5.8	6.0	2.0
		50	6.4	5.9	6.2	5.1			50	7.2	5.6	4.8	2.2
		100	5.4	4.9	5.5	4.9			100	6.8	5.2	5.9	2.9
2		25	11.0	6.8	9.9	4.3	2		25	17.7	11.6	10.5	2.7
		50	8.0	5.8	8.5	5.1			50	13.3	7.4	6.7	2.4
		100	7.6	5.9	7.2	4.7			100	11.0	6.8	8.3	3.1
3		25	14.2	8.5	14.3	4.9	3		25	22.6	10.2	10.2	2.4
		50	10.4	6.0	10.9	4.7			50	18.3	10.5	10.4	3.4
		100	8.1	6.1	7.4	5.0			100	14.3	7.0	6.3	2.7
1	$\widetilde{\Pi}_{(2)}$	25	8.2	5.3	8.6	3.3	1	$\widetilde{\Pi}_{(3)}$	25	88.9	57.9	75.1	0.4
		50	4.6	4.9	5.2	3.0			50	84.9	49.6	66.8	0.7
		100	4.2	4.3	5.1	4.0			100	85.0	44.8	68.0	0.6
2		25	12.6	5.9	13.9	3.1	2		25	85.0	44.8	79.7	0.1
		50	8.3	5.1	10.4	3.8			50	55.5	21.0	76.9	0.5
		100	7.6	3.7	11.7	3.5			100	95.3	58.7	74.3	0.6
3		25	14.7	7.3	18.7	4.1	3		25	99.3	82.3	84.4	1.0
		50	13.4	7.9	18.8	4.5			50	98.9	76.4	81.6	0.6
		100	11.6	5.1	17.1	3.7			100	98.9	70.0	77.8	0.5

Table 3. Empirical P(Type I error): Testing the full vector of endogenous variables coefficients.

			Wald_{2SLS}		LR_{LIML}				LR_{OLS}		Wald_{LIML}		AR
d	$\widetilde{\Pi}$	n	Asy	LMC	Asy	LMC	MMC	BD_Z	Asy	LMC	Asy	MC	
1	$\widetilde{\Pi}_{(1)}$	25	9.7	5.1	10.9	5.5	3.1	5.2	8.9	5.3	9.2	3.4	4.8
		50	7.1	5.1	6.8	4.4	2.1	3.5	6.1	4.	6.7	4.1	4.7
		100	6.5	4.8	6.6	4.7	2.2	2.4	6.3	4.3	6.3	4.7	5.3
2		25	11.4	6.2	13.3	6.5	1.6	3.5	8.6	5.0	12.1	4.0	4.6
		50	9.5	5.6	10.1	6.8	2.3	2.5	6.9	5.9	8.9	5.0	4.9
		100	8.2	5.9	6.2	4.1	0.8	1.2	5.2	4.2	7.9	5.6	4.2
3		25	14.8	7.2	16.0	7.5	1.4	2.6	11.4	6.3	15.5	5.0	4.4
		50	11.8	5.4	10.2	4.8	1.2	1.7	7.5	5.2	13.0	4.2	5.6
		100	8.4	6.4	7.4	5.2	0.6	0.2	5.0	4.7	8.0	5.9	4.3
1	$\widetilde{\Pi}_{(2)}$	25	8.1	5.0	12.9	5.4	3.8	6.9	8.9	5.3	7.7	2.7	4.8
		50	4.9	3.3	9.7	5.7	3.4	4.3	6.1	4.6	4.4	1.9	4.7
		100	4.4	4.0	11.1	5.5	3.6	4.8	13.3	6.3	4.0	4.1	5.3
2		25	12.8	6.5	18.1	6.6	2.4	4.7	8.6	5.0	11.8	4.1	4.6
		50	9.9	5.2	15.6	7.2	3.8	3.6	6.9	5.9	9.0	3.6	4.9
		100	6.5	4.0	13.2	5.7	2.7	2.5	5.2	4.2	6.0	3.2	4.2
3		25	14.9	6.9	20.7	7.3	2.3	4.1	11.4	6.3	14.8	3.3	4.4
		50	12.1	5.7	20.8	7.3	2.4	3.7	7.5	5.2	14.2	3.6	5.6
		100	9.2	5.0	17.3	6.4	2.2	2.6	5.0	4.7	11.2	3.1	4.3
1	$\widetilde{\Pi}_{(3)}$	25	11.9	6.4	12.8	5.4	3.7	6.7	8.5	5.3	8.9	2.7	4.8
		50	6.5	5.2	9.7	5.8	3.4	4.5	6.1	4.6	4.8	3.1	4.7
		100	5.6	4.4	11.1	5.5	3.6	4.8	6.3	4.3	4.1	4.2	5.3
2		25	18.9	10.3	18.0	6.6	2.4	4.7	8.6	5.0	14.2	3.3	4.6
		50	12.1	6.2	15.7	7.3	3.8	3.6	6.9	5.9	10.2	2.6	4.9
		100	9.4	5.0	13.2	5.7	2.7	2.5	5.2	4.2	7.2	2.8	4.2
3		25	23.0	10.2	20.9	7.2	2.4	4.1	11.4	6.3	16.8	3.5	4.4
		50	18.5	8.2	20.9	7.1	2.5	3.7	7.5	5.2	15.8	3.4	5.6
		100	12.4	6.1	17.2	6.4	2.2	2.6	5.0	4.7	12.2	3.6	4.3
1	$\widetilde{\Pi}_{(4)}$	25	92.5	72.6	14.3	6.1	4.9	7.6	8.9	5.3	79.0	3.8	4.8
		50	91.1	66.5	10.9	6.0	4.1	4.9	6.1	4.6	73.1	3.9	4.7
		100	90.2	61.3	11.3	5.1	3.7	5.0	6.3	4.3	7.11	3.2	5.3
2		25	98.9	85.3	21.8	6.4	3.1	5.8	8.6	5.0	82.3	2.8	4.6
		50	98.4	79.4	18.1	6.1	4.4	4.6	6.9	4.6	73.1	3.9	4.9
		100	97.5	71.5	14.7	5.4	3.1	2.9	5.2	4.2	76.9	3.2	4.2
3		25	99.6	90.7	26.5	7.7	3.1	5.3	11.4	6.3	84.9	2.5	4.4
		50	99.3	87.2	23.6	6.5	3.0	5.3	7.5	5.2	82.2	3.8	5.6
		100	99.1	81.9	20.7	6.2	2.8	3.0	5.0	4.7	78.5	2.7	4.3

Table 4. Power: Testing the full vector of endogenous variables coefficients

$H_0 : \beta_{11} = 10$			LR_{LIML}				LR_{OLS}	$Wald_{LIML}$	AR
Sample Size	d	β_{11}	LMC	MMC	BMC	BD_z	PMC	LMC	
25	1	10.1	12.2	8.0	8.0	11.2	11.4	13.8	10.4
		10.2	30.2	22.8	22.8	31.7	27.4	32.6	29.0
		10.3	55.0	44.4	44.3	55.8	50.4	51.1	51.5
		10.5	88.6	80.8	80.8	89.6	84.8	73.8	86.6
		11	99.9	99.4	99.4	99.9	99.6	83.6	99.7
	2	10.1	14.6	5.5	5.4	9.3	10.5	12.6	10.7
		10.2	35.9	19.0	18.8	29.0	29.4	29.9	29.7
		10.3	59.6	39.2	38.8	51.5	48.8	48.9	51.0
		10.5	91.5	77.0	76.8	87.6	84.5	69.8	86.6
		11	1.0	99.2	99.2	99.8	99.5	80.1	99.6
	3	10.1	15.0	4.7	4.2	8.2	9.9	12.6	10.5
		10.2	35.9	14.2	13.7	23.1	24.6	30.9	25.7
		10.3	61.4	32.6	30.7	46.9	46.5	49.7	48.1
		10.5	93.1	73.8	71.3	86.2	84.1	72.3	85.1
		11	1.0	99.1	99.0	99.7	99.6	81.9	1.0
$H_0 : \beta_{11} = 10$			LR_{LIML}				LR_{OLS}	$Wald_{LIML}$	AR
Sample Size	d	β_{11}	LMC	MMC	BMC	BD_z	PMC	LMC	
50	1	10.1	22.4	15.6	15.6	19.0	19.3	25.9	20.5
		10.2	66.9	54.0	54.0	59.2	60.2	62.9	60.9
		10.3	93.2	88.4	88.4	92.0	90.7	84.8	92.8
		11.0	1.0	1.0	1.0	1.0	1.0	96.4	1.0
	2	10.1	24.3	11.9	11.8	15.5	20.1	24.1	20.4
		10.2	67.6	46.4	45.8	53.1	57.5	59.0	59.8
		10.3	93.2	83.8	83.6	88.6	89.1	80.7	89.9
		11.0	1.0	1.0	1.0	1.0	1.0	94.8	1.0
	3	10.1	22.8	7.7	6.9	9.7	16.9	22.9	17.1
		10.2	61.8	31.8	30.7	38.5	46.2	54.5	48.8
		10.3	90.1	68.7	67.2	74.5	79.4	78.8	81.8
		11.0	1.0	1.0	1.0	1.0	1.0	95.3	1.0
$H_0 : \beta_{11} = 10$			LR_{LIML}				LR_{OLS}	$Wald_{LIML}$	AR
Sample Size	d	β_{11}	LMC	MMC	BMC	BD_z	PMC	LMC	
100	1	10.1	41.4	31.6	31.6	33.9	37.9	44.3	45.9
		10.2	93.8	87.0	87.0	89.0	90.3	89.9	91.4
		10.5	1.0	1.0	1.0	1.0	1.0	99.1	1.0
	2	10.1	40.7	19.4	18.9	23.2	31.4	41.6	33.6
		10.2	95.1	77.0	76.6	81.0	84.3	88.6	87.1
		10.3	99.6	98.4	98.4	98.7	98.8	98.9	99.1
		10.5	1.0	1.0	1.0	1.0	1.0	99.6	1.0
	3	10.1	38.3	15.6	13.7	15.3	27.3	40.6	27.6
		10.2	89.0	70.2	70.1	71.6	82.0	88.0	83.0
		10.5	1.0	1.0	1.0	1.0	1.0	99.3	1.0

Table 5. Power: Testing a subset of endogenous variables coefficients

$H_0 : \beta_{11} = 10$			LR_{LIML}				LR_{OLS}			
Sample Size	d	β_{11}	LMC	MMC	BMC	BD_z	LMC	MMC	BMC	BD_z
25	1	10.3	15.3	7.2	3.5	6.1	12.7	8.0	5.6	9.1
		10.5	18.8	10.3	5.8	8.9	16.2	11.2	8.4	12.4
		11.0	21.6	10.8	7.8	11.1	18.4	13.2	10.5	14.8
	2	10.3	13.6	6.2	2.8	5.3	10.6	6.5	5.5	10.1
		10.5	15.7	7.9	4.6	7.5	13.0	9.0	7.4	12.9
		11.0	19.4	10.1	5.6	9.6	15.3	11.6	9.0	16.0
	3	10.3	15.1	5.4	2.0	3.7	8.3	4.8	4.8	10.1
		10.5	17.7	7.3	2.5	5.2	11.4	7.7	6.6	12.8
		11.0	22.4	8.7	3.5	7.6	13.8	10.4	8.7	16.3
$H_0 : \beta_{11} = 10$			LR_{LIML}				LR_{OLS}			
Sample Size	d	β_{11}	LMC	MMC	BMC	BD_z	LMC	MMC	BMC	BD_z
50	1	10.1	11.0	4.9	2.5	2.6	8.8	5.9	4.0	5.2
		10.3	28.8	18.4	10.5	12.8	24.7	20.3	15.6	17.7
		10.5	39.1	27.6	17.0	19.5	33.3	28.1	21.7	5.6
		11.0	48.2	35.5	24.0	27.5	42.7	36.6	29.0	33.2
	2	10.1	10.3	3.5	1.2	1.4	6.9	4.1	3.3	4.6
		10.3	23.4	14.0	5.4	8.2	17.9	14.1	9.6	13.9
		10.5	30.3	19.5	10.0	14.0	25.3	20.7	15.9	19.8
		11.0	37.3	22.9	15.6	18.2	31.4	27.1	21.3	26.7
	3	10.1	11.9	3.4	0.7	1.1	7.1	4.5	3.4	5.7
		10.3	28.8	12.7	4.8	6.3	19.2	14.9	12.0	17.4
		10.5	37.5	20.2	8.9	12.3	26.9	21.5	17.7	24.1
		11.0	45.3	27.5	13.6	18.3	35.6	29.2	25.3	32.8
$H_0 : \beta_{11} = 10$			LR_{LIML}				LR_{OLS}			
Sample Size	d	β_{11}	LMC	MMC	BMC	BD_z	LMC	MMC	BMC	BD_z
100	1	10.1	16.6	10.4	4.4	4.7	14.6	10.7	6.7	7.9
		10.2	38.9	25.9	15.6	16.8	32.6	26.7	21.5	23.5
		10.3	54.6	44.3	26.0	28.1	47.7	39.4	32.1	34.6
		10.5	69.9	58.9	42.2	45.6	63.5	58.1	49.1	52.5
		11.0	80.4	68.6	56.1	60.8	76.3	72.5	62.5	66.7
	2	10.1	19.0	12.6	3.3	3.8	12.9	9.3	7.5	7.9
		10.2	42.1	271	10.7	13.0	29.7	25.4	19.9	22.3
		10.3	58.6	42.7	21.5	24.8	45.0	40.2	33.6	37.4
		10.5	70.9	59.6	38.6	43.0	62.5	58.3	50.8	54.6
		11.0	82.0	70.7	53.2	58.0	75.6	71.1	65.1	69.0
	3	10.1	18.2	8.4	1.7	2.2	11.0	8.4	6.7	7.4
		10.2	40.6	20.8	7.7	8.7	27.1	22.0	19.2	21.7
		10.3	55.8	34.1	15.3	17.7	40.6	36.0	30.6	34.5
		10.5	72.8	49.5	28.4	31.8	59.1	53.1	46.9	51.7
		11.0	81.8	64.6	44.4	48.7	74.0	69.9	64.0	68.7

Angrist, J. D. and Krueger, A. B. (1994), Split sample instrumental variables, Technical Working Paper 150, N.B.E.R., Cambridge, MA.

Athreya, K. B. (1987), 'Bootstrap of the mean in the infinite variance case', *The Annals of Statistics* **15**, 724–731.

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.

Bartlett, M. S. (1948), 'A note on the statistical estimation of supply and demand relations from time series', *Econometrica* **16**, 323–329.

Basawa, I. V., Mallik, A. K., McCormick, W. P., Reeves, J. H. and Taylor, R. L. (1991), 'Bootstrapping unstable first-order autoregressive processes', *The Annals of Statistics* **19**, 1098–1101.

Bekker, P. A. and Dijkstra, T. K. (1990), 'On the nature and number of the constraints on the reduced form as implied by the structural form', *Econometrica* **58**, 507–514.

Beran, R. and Srivastava, M. (1985), 'Bootstrap tests and confidence regions for functions of a covariance matrix', *The Annals of Statistics* **13**, 95–115.

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

Bound, J., Jaeger, D. A. and Baker, R. M. (1995), 'Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak', *Journal of the American Statistical Association* **90**, 443–450.

Buse, A. (1992), 'The bias of instrumental variables estimators', *Econometrica* **60**, 173–180.

Byron, R. P. (1974), 'Testing structural specification using the unrestricted reduced form', *Econometrica* **42**, 869–883.

Choi, I. and Phillips, P. C. B. (1992), 'Asymptotic and finite sample distribution theory for IV estimators and tests in partially identified structural equations', *Journal of Econometrics* **51**, 113–150.

Corana, A., Marchesi, M., Martini, C. and Ridella, S. (1987), 'Minimizing multimodal functions of continuous variables with the 'simulated annealing' algorithm', *ACM Transactions on Mathematical Software* **13**, 262–280.

Cragg, J. G. and Donald, S. G. (1996), Testing overidentifying restrictions in unidentified models, Discussion paper, Department of Economics, University of British Columbia and Boston University, Boston.

Dagget, R. S. and Freedman, D. A. (1985), Econometrics and the law: A case study in the proof of antitrust damages, *in* L. Le Cam and R. A. Olshen, eds, 'Berkeley Conference in Honor of Jerzy Neyman and Jack Kiefer', Wadsworth, Belmont, California.

Davidson, R. and MacKinnon, J. G. (1993), *Estimation and Inference in Econometrics*, Oxford University Press, New York.

Dufour, J.-M. (1987), Linear Wald methods for inference on covariances and weak exogeneity tests in structural equations, *in* I. B. MacNeill and G. J. Umphrey, eds, 'Advances in the Statistical Sciences: Festschrift in Honour of Professor V.M. Joshi's 70th Birthday. Volume III, Time Series and Econometric Modelling', D. Reidel, Dordrecht, The Netherlands, pp. 317–338.

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

Dufour, J.-M. (2002), 'Monte Carlo tests with nuisance parameters: A general approach to finite-sample inference and nonstandard asymptotics in econometrics', *Journal of Econometrics* **forthcoming**.

Dufour, J.-M. (2003), 'Identification, weak instruments and statistical inference in econometrics', *Canadian Journal of Economics* **36**, forthcoming.

Dufour, J.-M. and Jasiak, J. (2001), 'Finite sample limited information inference methods for structural equations and models with generated regressors', *International Economic Review* **42**, 815–843.

Dufour, J.-M. and Khalaf, L. (2002), 'Simulation based finite and large sample tests in multivariate regressions', *Journal of Econometrics* **111**(2), 303–322.

Dufour, J.-M. and Khalaf, L. (2003), Finite sample tests in seemingly unrelated regressions, *in* D. E. A. Giles, ed., 'Computer-Aided Econometrics', Marcel Dekker, New York, chapter 2, pp. 11–35.

Dufour, J.-M. and Taamouti, M. (2003*a*), On methods for selecting instruments, Technical report, C.R.D.E., Université de Montréal.

Dufour, J.-M. and Taamouti, M. (2003*b*), Point-optimal instruments and generalized Anderson-Rubin procedures for nonlinear models, Technical report, C.R.D.E., Université de Montréal.

Dufour, J.-M. and Taamouti, M. (2003*c*), Projection-based statistical inference in linear structural models with possibly weak instruments, Technical report, C.R.D.E., Université de Montréal.

Durbin, J. (1957), 'Testing for serial correlation in systems of simultaneous regression equations', *Biometrika* **44**, 370–377.

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

Efron, B. (1982), *The Jacknife, the Bootstrap and Other Resampling Plans*, CBS-NSF Regional Conference Series in Applied Mathematics, Monograph No. 38, Society for Industrial and Applied Mathematics, Philadelphia, PA.

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.

Freedman, D. A. and Peters, S. C. (1984*a*), 'Bootstraping a regression equation: Some empirical results', *Journal of the American Statistical Association* **79**, 97–106.

Freedman, D. A. and Peters, S. C. (1984*b*), 'Bootstraping an econometric model: Some empirical results', *Journal of Business and Economic Statistics* **2**, 150–158.

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

Green, R., Hahn, W. and Rocke, D. (1987), 'Standard errors for elasticities: A comparison of bootstrap and asymptotic standard errors', *Journal of Business and Economic Statistics* **5**, 145–149.

Haavelmo, T. (1947), 'Methods of measuring the marginal propensity to consume', *Journal of the American Statistical Association* **42**, 105–122.

Hahn, J. and Hausman, J. (2002), Weak instruments: Diagnosis and cures in empirical econometrics, Technical report, Department of Economics, Massachusetts Institute of Technology, Cambridge, Massachusetts.

Hahn, J. and Hausman, J. (2003), IV estimation with valid and invalid instruments, Technical report, Department of Economics, Massachusetts Institute of Technology, Cambridge, Massachusetts.

Hall, A. R., Rudebusch, G. D. and Wilcox, D. W. (1996), 'Judging instrument relevance in instrumental variables estimation', *International Economic Review* **37**, 283–298.

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. (1981*a*), 'Testing for heteroscedasticity in simultaneous equation models', *Journal of Econometrics* **15**, 311–340.

Harvey, A. C. and Phillips, G. D. A. (1981*b*), 'Testing for serial correlation in simultaneous equation models: Some further results', *Journal of Econometrics* **17**, 99–105.

Harvey, A. C. and Phillips, G. D. A. (1989), Testing for structural change in simultaneous equation models, *in* P. Hackl, ed., 'Statistical Analysis and Forecasting of Economic Structural Change', Springer-Verlag, Berlin, pp. 25–36.

Hausman, J. (1978), 'Specification tests in econometrics', *Econometrica* **46**, 1251–1272.

Hu, Y. S., Lau, N., Fung, H. and Ulveling, E. F. (1986), 'Monte Carlo studies on the effectiveness of the bootstrap bias reduction methods on 2SLS estimates', *Economics Letters* **20**, 233–239.

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.

Kleibergen, F. (2002), 'Pivotal statistics for testing structural parameters in instrumental variables regressions', *Econometrica* **70**(5), 1781–1804.

Kleibergen, F. and Zivot, E. (2003), 'Bayesian and classical approaches to instrumental variable regression', *Journal of Econometrics* **114**(1), 29–72.

Korajczyk, R. (1985), 'The pricing of forward contracts for foreign exchange', *Journal of Political Economy* **93**, 346–368.

Li, H. and Maddala, G. S. (1996), 'Bootstrapping time series models', *Econometric Reviews* **15**, 115–158.

Maddala, G. S. (1974), 'Some small sample evidence on tests of significance in simultaneous equations models', *Econometrica* **42**, 841–851.

Maddala, G. S. and Jeong, J. (1992), 'On the exact small sample distribution of the instrumental variable estimator', *Econometrica* **60**, 181–183.

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

McManus, D. A., Nankervis, J. C. and Savin, N. E. (1994), 'Multiple optima and asymptotic approximations in the partial adjustment model', *Journal of Econometrics* **62**, 91–128.

Moreira, M. J. (2003*a*), 'A conditional likelihood ratio test for structural models', *Econometrica* **71**(4), 1027–1048.

Moreira, M. J. (2003*b*), A general theory of hypothesis testing in the simultaneous equations model, Technical report, Department of Economics, Harvard University, Cambridge, Massachusetts.

Moreira, M. and Rothenberg, T. (2003), Bootstrap inference for structural parameters when identication is weak, Technical report, Department of Economics, Harvard University, Cambridge, Massachusetts.

Nelson, C. R. and Startz, R. (1990*a*), 'The distribution of the instrumental variable estimator and its t-ratio when the instrument is a poor one', *Journal of Business* **63**, 125–140.

Nelson, C. R. and Startz, R. (1990*b*), 'Some further results on the exact small properties of the instrumental variable estimator', *Econometrica* **58**, 967–976.

Park, S. (1985), Bootstraping 2SLS of a dynamic econometric model: Some empirical results, Working paper, Carleton University.

Perron, B. (2003), 'Semiparametric weak instrument regressions with an application to the risk return tradeoff', *Review of Economics and Statistics* **85**(2), 424–443.

Phillips, P. C. B. (1983), Exact small sample theory in the simultaneous equations model, *in* Z. Griliches and M. D. Intrilligator, eds, 'Handbook of Econometrics, Volume 1', North-Holland, Amsterdam, chapter 8, pp. 449–516.

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

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

Shea, J. (1997), 'Instrument relevance in multivariate linear models: A simple measure', *Review of Economics and Statistics* **LXXIX**, 348–352.

Sriram, T. N. (1994), 'Invalidity of bootstrap for critical branching processes with immigration', *The Annals of Statistics* **22**, 1013–1023.

Staiger, D. and Stock, J. H. (1997), 'Instrumental variables regression with weak instruments', *Econometrica* **65**, 557–586.

Stock, J. H. and Wright, J. H. (2000), 'GMM with weak identification', *Econometrica* **68**, 1097–1126.

Stock, J. H., Wright, J. H. and Yogo, M. (2002), 'A survey of weak instruments and weak identification in generalized method of moments', *Journal of Business and Economic Statistics* **20**(4), 518–529.

Taylor, W. E. (1983), 'On the relevance of finite sample distribution theory', *Econometric Reviews* **2**, 1–39.

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

Wang, J. and Zivot, E. (1998), 'Inference on structural parameters in instrumental variables regression with weak instruments', *Econometrica* **66**, 1389–1404.

Wright, J. H. (2003), 'Detecting lack of identification in GMM', *Econometric Theory* **19**(2), 322–330.

Wu, D.-M. (1973), 'Alternative tests of independence between stochastic regressors and disturbances', *Econometrica* **41**, 733–750.

Zivot, E., Startz, R. and Nelson, C. R. (1998), 'Valid confidence intervals and inference in the presence of weak instruments', *International Economic Review* **39**, 1119–1144.