**Identification du jury**

Université de Montréal

Faculté des études supérieures

Cette thèse intitulée:

Simulation Based Finite and Large Sample Inference Methods

in Seemingly Unrelated Regressions

and Simultaneous Equations

Présentée par

Lynda Khalaf

a été évaluée par un jury composé des personnes suivantes:

Thèse acceptée le:.............................................................................................

Université de Montréal

Simulation Based Finite and Large Sample Inference Methods

in Seemingly Unrelated Regressions

and Simultaneous Equations

par

Lynda Khalaf

Département de sciences économiques

Faculté des arts et sciences

Thèse présentée à la Faculté des études supérieures

en vue de l'obtention du grade de

Philosophiæ Doctor (Ph.D.)

en sciences économiques

Novembre, 1996

# Sommaire

Dans cette thèse, nous proposons des tests d'hypothèses fondés sur des simulations applicables à des systèmes d'équations linéaires multiples. Nous considérons successivement les régressions empilées, les équations apparamment non-reliées (le modèle SURE) et les systèmes d'équations simultanées. Dans ces modèles, les méthodes de test sont pour la plupart limitées aux procédures asymptotiques dont la performance sur des échantillons finis peut être arbitrairement mauvaise. En effet, les problèmes associés aux tests conventionels dans le contexte multivarié sont maintenant bien connus. Des études théoriques et pratiques montrent qu'il y a de sérieux problèmes de distorsion de niveau provenant d'une part du nombre élevé de paramètres dans les grands systèmes et d'autre part des conditions d'identifiabilité.

Les essais que nous présentons ici proposent des méthodes alternatives de test basées sur des simulations, visant à résoudre les problèmes de contrôle de niveau sur des échantillons finis. L'approche utilisée pour obtenir ces tests requiert l'exécution de trois techniques complémentaires: un test conservateur basé sur des bornes pivotales, un test simulé libéral s'apparentant au bootstrap, et (lorsque requis) un test de Monte Carlo exact randomisés. Les développements récents de l'informatique permettent maintenant d'exécuter ces procédures à faible coût malgré les difficultés liées au grand nombre de paramètres de nuisance.

Le premier essai est axé sur le modèle de référence: le modèle linéaire multivarié. Nous considérons en premier lieu le cas des hypothèses uniformes linéaires et nous généralisons au cadre non-gaussien des résultats classiques pour les modèles d'analyse de variance multivariée. Afin de traiter les hypothèses linéaires ou non-linéaires plus générales, nous dérivons une borne pivotale sur la distribution du critère du quotient de vraisemblance.

Nous en déduisons la validité de la procédure de tests décrits çi-haut. Ensuite, nous généralisons les résultats au contexte des régressions apparemment non-reliées. Des expériences de Monte Carlo sont effectuées pour illustrer les principaux résultats obtenus. Nos simulations confirment le manque de fiabilité des procédures asymptotiques usuelles et des ajustements de type Bartlett. En revanche, les tests que nous proposons contrôlent parfaitement le niveau et ont une excellente puissance. Nous illustrons les tests proposés sur les données de Fischer (1993) portant sur un modèle de détermination de la croissance économique.

Dans le deuxième essai, nous considérons l'hypothèse de diagonalité de la matrice de covariance des erreurs dans le contexte SURE. Nous démontrons que les statistiques du quotient de vraisemblance et du multiplicateur de Lagrange sont pivotales. Par conséquent, on peut facilement calculer des points critiques exacts par la méthode de Monte Carlo. Nous suggérons également un test de Monte Carlo exact basé sur un critère de quasi-maximum de vraisemblance (QLR) calculé à partir des estimés des moindres carrés ordinaires. Nos simulations montrent la supériorité des tests QLR simulés, un résultat très intéressant du point vue pratique. Nous illustrons l'utilité des diverses procédures de test par des applications au modèle de croissance spécifié dans le premier chapitre.

Dans le chapitre trois, nous nous plaçons dans le contexte des équations simultanées et considérons des tests portant sur les paramètres structuraux. En premier lieu, nous exprimons l'hypothèse nulle en fonction des paramètres de la forme réduite. Nous nous retrouvons ainsi dans une situation où nous pouvons mettre en œuvre les diverses stratégies de test proposées dans le chapitre premier. En particulier, nous dérivons une borne pivotale sur la distribution du critère LR pour des contraintes croisées générales et des hypothèses linéaires concernant une équation structurelle. Nous présentons aussi une généralisation aux cadre multivarié du test proposé par Anderson et Rubin (Annals of Mathematical statistics, 1949). Nous évaluons la performance des diverses statistiques de test en présence de

problème d'indentification. Nos simulations montrent que la statistique de Wald usuelle est fondamentalement erronnée et ne peut être améliorée par Monte Carlo. Par contre, les résultats concernant la statistique LR illustrent le manque de fiabilité du bootstrap usuel et la validité du test de Monte Carlo maximisé.

# Résumé

Nous étudions dans cette thèse divers tests d'hypothèses possiblement non-linéaires, concernant les paramètres d'un modèle linéaire à équations multiples. En particulier, nous considérons: (1) les modèles linéaires multivariés (MLR), (2) les systèmes d'équations apparemment non-reliées (SURE), et (3) les systèmes d'équations simultanées (SE). Rappelons que le système SURE se réduit à un modèle linéaire multivarié avec des contraintes d'exclusion sur les coefficients de régression. Par ailleurs, le modèle à équations simultanées écrit sous forme réduite peut être vu comme un système de régressions empilées.

Ces problèmes de test sont évidemment classiques; de fait, les méthodes de test usuelles sont applicables mais leur principales propriétés sont asymptotiques. Alors que des tests exacts existent pour un nombre de cas particuliers intéressants, ce fait est souvent négligé par les analystes qui appliquent généralement des procédures asymptotiques. A cet égard, nous commençons par un survol rapide des résultats exacts disponibles pour des hypothèses spécifiques dans le cas des trois modèles considérés. Nous nous intéressons particulièrement au cas où l'hypothèse s'écrit sous forme de contraintes *uniformes linéaires* sur la matrice des coefficients de régression. Dans ce dernier cas, nous généralisons au cadre non-gaussien des résultats distributionnels classiques apparus dans la litérature statistique portant sur l'analyse de variance multiple [voir Anderson (1984, chapitres 8 et 13) et Rao (1974, chapitre 8)].

Bien que les hypothèses uniformes linéaires sont fréquentes dans la pratique économétrique [voir Stewart (1995)], il est clair qu'elles sont fort restrictives. Pour des hypothèses générales linéaires ou non-linéaires, les distributions exactes des diverses statistiques de test usuelles dépendent de paramètres de nuisance inconnus. Le problème fondamental est alors de calculer les points critiques pertinents tels que l'on puisse contrôler

le niveau des tests sur des échantillons finis. Dans un contexte multivarié, c'est rarement le cas. En effet, de nombreuses expériences de simulation suggèrent que les erreurs de type I associées aux tests standards peuvent considérablement dépasser le niveau affiché. A ce sujet, nous citons dans ce texte un ensemble de résultats portant sur la performance des tests multivariés, apparus dans la litérature sur: (1) l'homogénéité et la symétrie dans les systèmes d'équations de demande, (2) l'efficacité des marchés financiers, (4) les tests d'indépendance, et (4) les régressions instrumentales fondées sur des instruments faibles. Nos simulations permettent aussi d'illustrer le manque de fiabilité des tests habituellement employés en pratique.

L'importance des distorsions de niveau dans le contexte des équations simultanées a été soulignée récemment par un résultat de Dufour (1996) qui montre que les statistiques fondées sur des estimateurs instrumentaux ne peuvent être bornées et par conséquent, les niveaux des tests correspondants peuvent dévier arbitrairement des niveaux nominaux. De plus, les techniques de correction de niveau usuelles (*e.g.*, le bootstrap) ne peuvent résoudre ce problème. Un test valide devrait être fondé sur une statistique *uniformément bornable, i.e.* dont la distribution sous l'hypothèse nulle admet une borne qui ne dépend pas de paramètres de nuisance [voir Lehmann (1986, chapitre 3)]. Dans les cas qui nous concernent, nous démontrons que la statistique du quotient de vraisemblance (LR) satisfait cette dernière propriété. Pour ce faire, nous dérivons une borne pivotale sur la distribution du critère LR. Cette borne est construite à partir de la statistique de test portant sur des contraintes uniformes linéaires qui constituent un cas particulier de l'hypothèse à tester. Nous généralisons ainsi au cadre multivarié le résultat de Dufour (1989) pour le modèle linéaire à une équation. Il est important de noter que: (1) les résultats démontrés permettent des erreurs non-gaussiennes, et (2) la statistique bornante est facile à simuler ce qui rend les bornes faciles à utiliser. Par conséquent, nous observons que les méthodes fondées sur une borne pertinente combinées à des techniques de Monte Carlo peuvent aisément fournir des

tests valides basés sur la vraisemblance.

En effet, étant donné les progrès de l'informatique récents, les méthodes de correction de niveau basées sur des simulations (*e.g.*, le bootstrap) apparaisent de façon naturelle dans le contexte multivarié. Sur cette question, nous passons en revue les divers résultats portant sur l'application du bootstrap dans les modèles à équations multiples. Nous constatons que la technique du bootstrap usuelle n'offre pas de solution satisfaisante au problème du contrôle de niveau des tests sur des échantillons finis. Pour cette raison, nous proposons des méthodes différentes basées sur les techniques de Monte Carlo introduites par Dwass (1957) et Barnard (1957). Dufour (1995) fournit une discussion détaillée des tests de Monte Carlo en présence de paramètres de nuisance. En effet, ces méthodes sont applicables en présence de paramètres de nuisance, mais il est nécessaire de considérer des critères de test uniformément bornables. Les procédures de Monte Carlo s'apparentent au bootstrap. Cependant, contrairement à ce dernier, les tests de Monte Carlo tiennent compte explicitement du nombre de replications aboutissant ainsi à un test exact randomisé. En effet, le nombre de réplications peut être assez petit (*e.g.*, un minimum de 19 replications est requis pour un test de niveau 5%). Par ailleurs, des points critiques libéraux sont aisément déterminés à partir des statistiques de test simulées évaluées en fonction d'estimateurs convergents des paramètres de nuisance. En pratique, l'application des tests de Monte Carlo exacts conduit à maximiser les *p*-values sur l'espace des paramètres de nuisance pertinents. Afin de mener numériquement la maximisation, nous exploitons l'algorithme d'optimisation globale "*simulated annealing*" proposé par Corona et al (1987) et Goffe et Ferrier (1994). Nous déduisons alors une procédure en plusieurs étapes qui peut être appliquée dans l'ordre suivant:

(1) effectuer le test à borne; si ce dernier rejette, la procédure conclut en faveur d'un rejet de l'hypothèse nulle;

(2) sinon, on obtient le point critique simulé libéral; si l'hypothèse est acceptée, la procédure accepte l'hypothèse nulle;

(3) en poursuivant la démarche, on peut par la suite calculer la $p$-value maximale et décider de rejetter l'hypothèse nulle si cette dernière est inférieure au seuil de signification désiré. Nous appliquons cette stratégie à divers problèmes de tests dans le contexte des trois modèles multivariés spécifiés plus-haut.

Dans le premier chapitre, nous considérons le modèle MLR de référence. Nous examinons d'abord des hypothèses uniformes linéaires pour lesquelles nous présentons des statistiques pivotales qui fournissent des tests exacts. Nous en déduisons le résultat distributionnel de base portant sur le critère LR qui permettra de dériver les bornes pivotales pour des hypothèses générales linéaires ou non-linéaires. Nous vérifions par Monte Carlo la validité des procédures proposées. De plus, nous démontrons que la correction de Barlett proposée par Attfield (1995) ne corrige pas le niveau du test LR dans le contexte des grands systèmes. Par la suite, nous considérons les contraintes générales sur les coefficients des modèles MLR et SURE. Nous évaluons la taille et la puissance des tests proposés au moyen d'expériences de Monte Carlo. En particulier, nous constatons que: (1) les tests de Monte Carlo libéraux fondés sur des estimateurs convergents des paramètres de nuisance corrigent le niveau, et (2) la borne est concluante dans une grande proportion de cas. Nous illustrons les tests proposés à l'aide de données déjà utilisées par Fischer (1993) portant sur un modèle de détermination de la croissance économique.

Dans le chapitre deux, nous nous intéressons toujours au modèle SURE et considèrons l'hypothèse de diagonalité de la matrice de covariance des erreurs. Pour ce problème, nous démontrons que la statistique LR et la statistique du multiplicateur de Lagrange (LM) sont pivotales quelque soit le nombre d'équations constituant le système. Ce résultat n'a été établi à cette date que dans le cas de deux équations [voir Kariya (1981a,b)]. Un test exact peut donc être effectué par la méthode de Monte Carlo appliquée aux critères LR ou LM. Nous suggérons également un test de Monte Carlo exact basé sur un critère de quasi-maximum de vraisemblance (QLR) calculé à partir des estimés des moindres carrés

ordinaires (MCO). Par ailleurs, nous généralisons au cadre de plusieurs équations un test d'indépendance exact proposé par Harvey et Phillips (1982). Nous comparons entre elles les méthodes de tests proposées. Nos résultats indiquent que les tests QLR simulés utilisant les estimateurs MCO sont généralement meilleurs que les autres tests disponibles. Ainsi, sur le plan du calcul il est fort peu utile d'obtenir le maximum de vraisemblance. Nous illustrons l'utilité des diverses procédures de test par des applications au modèle de croissance spécifié dans le chapitre précédent.

Dans le chapitre trois, nous nous plaçons dans le contexte des équations simultanées et considérons des tests portant sur les paramètres structuraux. Afin de développer un cadre d'analyse, nous exprimons l'hypothèse nulle en fonction des paramètres de la forme réduite. On peut noter ici que des hypothèses linéaires sur les paramètres structuraux impliquent des restrictions non-linéaires sur les coefficents de la forme réduite. Une fois de plus, nous nous retrouvons dans une situation où nous pouvons utiliser les divers résultats distributionels établis dans le chapitre premier. En particulier, nous soulignons l'existence d'une borne pivotale sur la distribution du critère LR. Nous explicitons les caratéristiques de cette borne en premier lieu pour des hypothèses genérales et ensuite pour des hypothèses spécifiques intéressantes, notamment des contraintes linéaires sur les coefficients d'une équation structurelle. Nous présentons aussi un test de type système qui peut être interprété comme une généralisation à un cadre multivarié du test proposé par Anderson et Rubin (1949). A cet égard, soulignons que la performance des tests Anderson-Rubin a été étudiée dans plusieurs contextes par Dufour et Jasiak (1996). Nous évaluons la performance des diverses statistiques de test en présence de problème d'indentification. En particulier, nous vérifions par des expériences de simulation que la statistique de Wald usuelle sur-rejette considérablement et ne peut être améliorée par Monte Carlo. De fait, nous démontrons numériquement que le maximum de la $p$-value associée au test de Wald sur l'espace des paramètres de nuisance est toujours un, de sorte qu'un point critique utile ne peut être obtenu. De même, nous constatons

que la statistique LR ne peut être corrigée par un bootstrap usuel, alors que la procédure de Monte Carlo maximisée contrôle parfaitement le niveau du test. Ces propriétés tiennent même si les paramètres impliqués sont presque non-identifiables.

Finalement, le dernier chapitre conclut le rapport et présente une discussion critique de nos résultats.

## Bibliographie

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

Anderson, T.W. and H. Rubin (1949). *Estimation of the parameters of a single equation in a complete system of stochastic equations*, Annals of Math. Stat. 20, 46-63.

Attfield, C.L.F. (1995). *A Bartlett adjustment to the likelihood ratio test for a system of equations*, Journal of Econometrics 66, 207-223.

Barnard, G.A. (1963). Comment, Journal of the Royal Statistical Society B, 25, 294.

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

Dufour, J.M. (1989). *Nonlinear Hypotheses, Inequality Restrictions and Non-Nested Hypotheses: Exact Simultaneous Tests in Linear Regressions*, Econometrica, 57 (2), 335-355.

Dufour, J.M. (1996). *Some impossibility theorems in econometrics with applications to structural and dynamic models*, Econometrica, forthcoming.

Dufour, J.M. (1995). *Monte Carlo tests in the presence of nuisance parameters, with econometric applications*, working paper, Centre de recherche et développement en

économique (C.R.D.E.) and Département de sciences économiques, Université de Montréal.

Dufour, J.M. and J. Jasiak (1996). *Finite sample inference methods for simultaneous equations and models with unobserved and generated regressors*, working paper, Centre de recherche et développement en économique (C.R.D.E.) and Département de sciences économiques, Université de Montréal.

Dwass, M. (1957). *Modified Randomization tests for Non-parametric Hypotheses*, Annals f Mathematical Statistics, 28, 181-187.

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

Harvey, A.C. and G.D.A. Phillips (1989). *Testing for structural change in simultaneous equation models*, in: Hackl, Peter ed., *Statistical analysis and forecasting of economic structural change*, Springer-Verlag, 25-36.

Kariya, T. (1981a). *Tests for the independence between two seemingly unrelated regression equations*, The Annals of Statistics 9 (2), 381-390.

Kariya, T. (1981b). *Robustness of multivariate tests*, The Annals of Statistics 6, 1267-1275.

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

Stewart, K.G. (1995). *Exact testing in multivariate regression*, working paper, department of economics, University of Victoria.

# Table des matières

# Liste des tableaux

# Chapitre 2.

# Chapitre 3.

# Chapter 1

# Simulation Based
# Finite and Large Sample Inference Methods
# in Multiple Equation Regression Models

Abstract

In the context of multivariate regression (MLR) and seemingly unrelated regression (SURE) models it is well known that commonly employed asymptotic test criteria are seriously biased towards overrejection. In this paper, we propose finite and large sample likelihood based test procedures for possibly non-linear hypotheses on the coefficients of MLR and SURE systems. Two complementary approaches are described. First, we derive general nuisance-parameter free bounds on the distribution of standard likelihood ratio criteria. Even though it may be difficult to compute these bounds analytically, they can be easily obtained by simulation. Second, we propose a number of Monte Carlo tests which can be run whenever the bounds are not conclusive and develop an extension of the bootstrap method to statistics whose asymptotic distributions involve nuisance parameters. They include, in particular, quasi-likelihood ratio criteria based on non-maximum-likelihood estimators. Illustrative Monte Carlo experiments show that: (i) the bound provides conclusive results in a large proportion of cases, and (ii) the randomized procedures correct all the usual size distortions in such contexts. We also present an extension of standard tests of uniform linear hypotheses in MLR contexts to non-Gaussian error distributions; in fact, the normality assumption is not necessary for most of the results we obtain. Practical implementations of both procedures is relatively straightforward. The methods proposed are applied to data used in Fischer (1993) to analyze the macroeconomic determinants of growth.

## 1. Introduction

Testing the validity of restrictions on the coefficients of a multivariate linear regression (MLR) model is a common issue which arises in statistics and econometrics. Extensive discussion of this problem can be found in the statistics literature on multivariate analysis of variance (MANOVA) and the econometric literature on seemingly unrelated regressions (SURE). The MLR model can be viewed as a special case of the SURE model where the regressor matrices for the different equations are the same. Conversely, the SURE specification may be viewed as a special case of the MLR model constrained by various exclusion restrictions on the different equations.

In the MLR framework, several finite sample procedures have been proposed for testing linear restrictions. These include, in particular, tests based on the likelihood ratio (LR) criterion [Wilks (1932), Bartlett (1947)], the Lawley-Hotelling (LH) trace criterion [Lawley (1938), Bartlett (1939), Hotelling (1947, 1951)], the Bartlett-Nanda-Pillai (BNP) trace criterion [Bartlett (1939), Nanda (1950), Pillai (1955)] and the Maximum Root (MR) criterion [Roy (1953)]. The literature concerning the moments, Laplace transforms and exact densities of these statistics is vast; see, for example, Rao (1973, chapter 8), Anderson (1984, chapters 8 and 13) and Kariya (1985). Yet the use of these methods is limited to very specific problems: tests of *uniform mixed linear* hypotheses [Bernt and Savin (1977)]. Examples of *uniform mixed linear* constraints include: (i) the case where the same transformations of the regression coefficients are set to given values, within or across equations, and (2) the hypothesis that a single parameter equals zero. Further, in most instances, exact distributional results are difficult to exploit and approximate distributions are suggested. Thus far less restrictive testing problems have not apparently been considered from a finite sample perspective, with perhaps the notable exception of the Hashimoto and Ohtani's (1990) exact test for general linear restrictions. This procedure is similar to Jayatissa's (1977) test for equality of regression coefficients in two linear regressions with unequal error variances.

However, the authors recognize that as with Jayatissa's procedure, the test involves complicated computations and has low power. Further, the test relies on a non-unique decomposition of the OLS residuals. These observations suggest that the test may be of limited practical interest.

In connection with the SURE model, the standard literature on hypothesis tests is asymptotic; see, for example, Srivastava and Giles (1987). Very few analytical finite sample results are available. A rare exception is provided by Harvey and Phillips (1982, Section 3) who derived independence tests between the disturbances of an equation and those of the other equations of a SURE model. The tests involve conventional F-statistics and are based on the residuals obtained from regressing each dependent variable on all independent variables of the system. Of course this problem is a very specific one. In a different vein, Phillips (1985) derived the exact distribution of a two-stage SURE estimator using a fractional matrix calculus. However, the analytical expressions obtained are very complex and, more importantly, involve unknown nuisance parameters, namely the elements of the error covariance matrix. The latter fact makes the application of Phillips' distributional results to practical hypothesis testing problems difficult.

Asymptotic Wald, Lagrange multiplier and likelihood ratio tests are available and commonly employed in econometric applications of the MLR model; see for example, Berndt and Savin (1977), Evans and Savin (1982), Breusch (1979), Gouriéroux, Monfort and Renault (1993, 1995) or Stewart (1995a,b). It has been shown however that in finite samples, these asymptotic criteria are seriously biased towards overrejection when the number of equations relative to the sample size is moderate to large. Well known examples include Laitinen (1978), Meisner (1979), Bera et al (1981) and Theil and Fiebig (1985) in the context of homogeneity and symmetry testing in demand systems. Further evidence is reported in relation to multivariate tests of the CAPM; see for example Stambaugh (1982), Jobson and Korkie (1982), Amsler and Schmidt (1985) and Mackinlay (1987). These and other references

are discussed in Stewart (1995b).

It is clear that standard asymptotic approximations are quite unsatisfactory in this context. Attempts to improve those include in particular (i) Bartlett-type corrections, and (ii) simulation-based methods. Basically, Bartlett corrections involve rescaling the test statistic by a suitable constant obtained such that the mean of the scaled statistic equals that of the approximating distribution to a given order [Bartlett (1937), Lawley (1956), Rothenberg (1984), Barndorff-Nielsen and Blaesid (1986)]. Formulae explicitly directed towards systems of equations are given in Attfield (1995). Overall, the correction factors require cumulants and joint cumulants of first and second order derivatives of the log likelihood function, and cannot, outside a small class of problems, be implemented as easily as might be expected. Furthermore, simulation studies [e.g. Ohtani and Toyoda (1985), Frydenberg and Jensen (1989), Hollas (1991), Rocke (1989), Wong (1989, 1991) and Gonzalo and Pitarkis (1994)] suggest that in many instances Bartlett adjustments are not as effective as expected. A more simple correction factor is proposed by Italianer (1985) yet the procedure is rather heuristic and has little theoretical background.

In connection with simulation-based tests, the bootstrap method [see Hall (1988), Beran (1988) or Hall and Horowitz(1996)] was suggested to obtain size corrected critical points. Jeong and Maddala (1993) and Vinod (1993) provide a comprehensive survey of econometric applications of the bootsrap; regarding MLR models, several Monte Carlo and/or empirical studies are discussed most of which report the efficacy of the procedure [e.g. Williams (1986), Rocke (1989), Rayner (1990a, 1990b), Eakin et al. (1990), Affleck-Graves et al. (1990), Martin (1990), Atkinson et al (1992) and Rilstone et al (1993)]. Although long recognized as a proper alternative to standard asymptotic theory, the bootstrap has only asymptotic justification when the null distributions of the test statistics involve nuisance parameters, hence the finite sample validity of resulting inference remains to be established. This point should be born in mind while interpreting results on the usefulness of the

bootstrap. In a different vein, several randomized tests have been suggested in the MLR literature for particular test problems and are referred to under the name of Monte Carlo tests, examples being Theil, Shonkwiler and Taylor (1985), Theil, Taylor and Shonkwiler (1986), Taylor, Shonkwiler and Theil (1986) and Theil and Fiebig (1985). However, these authors do not supply neither asymptotic nor finite sample theoretical results.

In this paper, we propose finite sample likelihood based tests for possibly non-linear hypotheses on the coefficients of seemingly unrelated regressions. We discuss two approaches that can be applied on their own or sequentially, namely: (i) a conservative bounds test, and (ii) Monte Carlo tests. Practical implementation of both procedures is basically straightforward. The methods we propose are best motivated by the propositions in Dufour (1996) relating to likelihood based inference in MLR settings: using an argument similar to the one in Dufour (1989) for a univariate regression, it is shown that LR statistics have null distributions which are boundedly pivotal, i.e. which admit nuisance-parameter-free bounds. Even though it may be difficult to compute analytically these bounds, they can easily be obtained by simulation. Here, we explicit and apply this result in the context of MLR and SURE systems. The implications for hypothesis testing are two-fold. First, the finite sample bounds on the LR criterion easily yield conservative tests. Second, bootstrap techniques can lead to tests with correct levels.

To be more specific, we give at this point a preliminary discussion of the proposed conservative bound with regards to SURE systems. First, we reconsider the testing problem within the framework of an appropriate MLR model, namely the MLR setup of which the model on hand is a restricted form. Secondly, we introduce, in the relevant MLR framework, a uniform linear hypothesis that is a special case of the general restrictions in the null. The intuition behind this suggestion follows from the fact that exact nuisance-parameter free critical values for the LR criterion are available when the null is uniform linear within a MLR. Indeed, it turns out that the LR criterion for testing the suggested uniform linear

hypothesis conveniently bounds the LR statistic for testing the general constraints.

In addition, we propose alternative Monte Carlo tests [see Dwass (1957), Barnard (1963), Jöckel (1986) or Dufour (1995)] that can be run whenever the bounds tests are not conclusive. We consider (i) an asymptotically valid procedure that may be interpreted as a parametric bootstrap, and (ii) a method which is exact for any sample size, following Dufour (1995). While the normality assumption underlies the motivation for the statistics we consider, this is not necessary for most of the results obtained. In fact, we discuss an extension of standard tests of uniform linear hypotheses in MLR contexts to non-Gaussian distributions. Further, in situations where maximum likelihood (ML) methods may be computationaly expensive, we introduce LR-type test criteria based on non-ML estimators. In particular, we consider two-stage statistics or estimators at any step of the process by which the likelihood is maximized iteratively. We emphasize that Monte Carlo and bounds tests should be viewed as complementary rather than alternative procedures.

The paper is organized as follows. Section 2 develops the notation and definitions. Section 3 discusses the known distributional results pertaining to the test criteria in the context of the MLR model and provides an extension of standard tests of uniform linear hypothesis to non-gaussian distributions. Section 4 presents test statistics for general linear hypotheses with respect to the MLR model and establishes bounds on the significance points for these statistics. We also discuss how to apply the results to non-linear and inequality restrictions. The generalization to the SURE is discussed in Section 5. Simulation results are reported in Section 6. Section 7 provides empirical illustrations of the various tests and Section 8 concludes.

## 2. The general framework

In this Section we introduce the models and notations to be used in the paper. The first model we consider is the MLR model. Then, we focus our attention on the SURE model,

which can be viewed as a special case of the MLR model obtained by imposing different exclusion restrictions on the different equations of a MLR model.

## 2.1. The multivariate linear regression model

The MLR model can be expressed as follows:

$$(2.1) \quad Y = XB + U,$$

where $Y = [Y_1, \dots, Y_p]$ is an $(n,p)$ matrix of observations on $p$ dependent variables, $X$ is an $(n,k)$ full-column rank matrix of fixed regressors, $B = [\beta_1, \dots, \beta_p]$ is a $(k,p)$ matrix of unknown coefficients and $U = [U_1, \dots, U_p] = [\tilde{U}_1, \dots, \tilde{U}_n]'$ is an $(n,p)$ matrix of random disturbances with covariance matrix $\Sigma$ where $\det(\Sigma) \neq 0$. We also assume that the rows $\tilde{U}_i'$, $i = 1, \dots, n$, of $U$ satisfy the following distributional assumptions:

$$(2.2a) \quad \tilde{U}_i = JW_i, \ i = 1, \dots, n,$$

where the vector $w = vec(W_1, \dots, W_n)$ has a known distribution and $J$ is an unknown, non-singular matrix; for further reference, let $W = [W_1, \dots, W_n]' = UG'$, where $G = J^{-1}$. In particular, this condition will be satisfied when

$$(2.2b) \quad W_i \sim N(0, I_p), \ i = 1, \dots, n,$$

in which case the covariance matrix of $\tilde{U}_i$ is $JJ' = (G'G)^{-1}$. An alternative representation of the model is

$$(2.3) \quad Y_i = X\beta_i + U_i, \ i = 1, \dots, p,$$

where $Y_i$ is a vector of $n$ observations on a dependent variable. By writing (2.2) as

$$(2.4) \quad \begin{bmatrix} Y_1 \\ Y_2 \\ \cdot \\ Y_p \end{bmatrix} = \begin{bmatrix} X & 0 & . & 0 \\ 0 & X & . & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & . & X \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \cdot \\ \beta_p \end{bmatrix} + \begin{bmatrix} U_1 \\ U_2 \\ \cdot \\ U_p \end{bmatrix},$$

the model may be expressed in compact form

(2.5)     $y = (I_p \otimes X) \, b + u$ ,

where   $y = vec(Y)$ , $b = vec(B)$ , and $u = vec(U)$ . The least squares estimate of $B$ is

(2.6)     $\hat{B} = (X'X)^{-1} X'Y$ ,

and the corresponding residual matrix is

(2.7)     $\hat{U} = Y - X\hat{B} = M \, Y = M \, U$ ,

where   $M = I - X \, (X'X)^{-1} X'$ . In this model, it is well known that under (2.2b) the maximum likelihood (ML) estimators of the parameters reduce to $\hat{B}$ and $\hat{\Sigma} = \hat{U}'\hat{U} / n$ . Thus the maximum of the likelihood function over the unrestricted parameter space is

(2.8)     $\displaystyle \max_{B,\Sigma} L = - \frac{np}{2} \ln(2\pi) - \frac{n}{2} \ln(|\hat{\Sigma}|) - \frac{np}{2}$ .

To derive the distributions of the relevant test statistics, we shall exploit the following decomposition of the sum of squared errors matrix $\hat{U}'\hat{U}$ :

(2.9)     $\hat{U}'\hat{U} = U' M \, U = G^{-1} \, (UG')' \, M \, (UG') \, (G^{-1})' = G^{-1} \, W' \, M \, W \, (G^{-1})'$ ,

where the matrix $W$ has a distribution which does not involve nuisance parameters.

## 2.2 The seemingly unrelated regression model

Let us now consider the following $p$ equations regression model:

(2.10)     $Y_i = X_i \beta_i + U_i$ , $\quad i = 1 , \dots , p$ ,

where $X_i$ is a $(n,k_i)$ full-column rank matrix of fixed regressors and $U_1, U_2, \dots, U_p$ satisfy the same distributional assumptions as in (2.2). This model is known as the SURE model. Let

$$(2.11) \quad y = \begin{bmatrix} Y_1 \\ Y_2 \\ \cdot \\ Y_p \end{bmatrix} , \quad X^{\bullet} = \begin{bmatrix} X_1 & 0 & . & 0 \\ 0 & X_2 & . & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & . & X_p \end{bmatrix} , \quad \beta^{\bullet} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \cdot \\ \beta_p \end{bmatrix} , \quad u = \begin{bmatrix} U_1 \\ U_2 \\ \cdot \\ U_p \end{bmatrix} .$$

Then an alternative compact representation of the model is

(2.12) $\quad y = X^*\beta + u$ .

In this case, $\beta$ is $(k^*,1)$ with $k^* = \sum_{i=1}^{p} k_i$ . The likelihood function associated with (2.12) is:

(2.13) $\quad L_1 = -\dfrac{np}{2} \ln(2\pi) - \dfrac{n}{2} \ln(|\Sigma|) - \dfrac{1}{2} (y - X^*\beta)'(\Sigma^{-1} \otimes I_n)(y - X^*\beta)$ ,

which is usually maximized using iterative numerical procedures. To develop finite sample tests for the SURE model, we will find useful to explicate the relation between SURE and MLR models. Let $\bar{X}_i$ , $i = 1, \dots , p$ , denote the matrix of observations on the explanatory variables excluded from the $i$-th equation. Further, define $Z$ as any full column rank $(n, k)$ matrix which spans the same space as $[X_1, X_2, \dots , X_p]$ , and $J_i$ , $\bar{J}_i$ selection matrices such that

(2.14) $\quad ZJ_i = X_i$ , $\quad Z\bar{J}_i = \bar{X}_i$ , $\quad i = 1 , \dots , p$ .

Then the SURE model (2.10) may be written as

(2.15) $\quad Y_i = ZJ_i\beta_i + Z\bar{J}_i\bar{\beta}_i + U_i$ , $\quad i = 1 , \dots , p$ ,

with the restriction

(2.16) $\quad \bar{\beta} = 0$ ,

where $\bar{\beta}_i$ is the vector of coefficients which do not appear in the equation for $Y_i$ and

(2.17) $\quad \bar{\beta} = [ \bar{\beta}_1' , \bar{\beta}_2' , \dots , \bar{\beta}_p' ]'$ .

To simplify notation, we will rewrite the unrestricted MLR (2.15) as

(2.18) $\quad Y_i = Z\beta_i + U_i$ , $i = 1, \dots , p$ ,

where $\beta_i$ is a $(k, 1)$ vector which includes all the elements of $\beta_i$ and $\bar{\beta}_i$ (though possibly not in the same order). The latter may also be written in the more compact form

(2.19) $\quad y = (I_p \otimes Z) \beta + u$ ,

where $\beta$ is the $(pk, 1)$ vector

(2.20) $\quad \beta = [ \beta_1' , \beta_2' , \dots , \beta_p' ]'$ .

## 3. Hypothesis testing in the Multivariate Linear Regression model

In this Section we shall review known finite sample distributional results pertaining to various criteria for testing a general linear hypothesis in the context of the MLR model (2.1), and provide some extensions that will allow analogous tests to be performed in a large set of models with non-gaussian errors. Finite sample procedures are available only for the case where the constraints take the special uniform linear (UL) form $H_0: RBC = D$ , where $R$ is a known $(r,k)$ matrix of rank $r \le k$, $C$ is a known $(p,c)$ matrix of rank $c \le p$, and $D$ is a known $(r,c)$ matrix. We will first study the problem $H_{01}: R\beta_i = \delta_i$ , $i = 1$ , ... , $p$ that corresponds to $C = I_p$. The exposition of this special case will be simplified and likewise, $H_{01}$ has traditionally been the subject of multivariate analysis of variance; see Rao (1973, chapter 8), Anderson (1984, chapter 8) or Kariya (1985). In this context, the most commonly used criteria are:

(a) the likelihood ratio (LR) or the Wilks $(\mathcal{L})$ criteria [Wilks (1932), Bartlett (1947)];

(b) the Lawley-Hotelling (LH) trace criterion [Lawley (1938), Bartlett (1939), Hotelling (1947, 1951)];

(c) the Bartlett-Nanda-Pillai (BNP) criterion [Bartlett (1939), Nanda (1950), Pillai (1955)];

(d) the Maximum Root (MR) criterion [Roy (1953)].

All these test criteria are functions of the roots $m_1$ , $m_2$ , ... , $m_p$ of the equation

(3.1) $\quad |\hat{U}'\hat{U} - m \, \hat{U}_0'\hat{U}_0| = 0$

where $\hat{U}_0'\hat{U}_0$ and $\hat{U}'\hat{U}$ are respectively the constrained and unconstrained sum of squared error matrices. For convenience, the roots are reordered so that $m_1 \ge \ldots \ge m_p$ . In particular, we have:

(3.2a) LR $= -n \ln(\mathcal{L})$ , $\mathcal{L} = \prod_{i=1}^{p} m_i = \dfrac{|\hat{U}'\hat{U}|}{|\hat{U}_0'\hat{U}_0|}$ ;

(3.2b) LH $= \sum_{i=1}^{p} \dfrac{1-m_i}{m_i}$ ;

(3.2c) BNP $= \sum_{i=1}^{p} (1 - m_i)$ ;

(3.2d) MR $= \underset{1 \le i \le p}{Maximum} \quad \dfrac{1 - m_i}{m_i}$ .

The hypothesis is rejected when LR is suitably small or when LH, BNP or MR are suitably large. Note that the criteria LH and BNP can be interpreted as Wald and Lagrange multiplier test statistics, respectively. For details of the relationship, see Berndt and Savin (1977), Breusch (1979) or Stewart (1995a).

In Section 2, we saw that $\hat{U}'\hat{U}$ can be expressed as $\hat{U}'\hat{U} = G^{-1}W'MW' (G^{-1})'$ which depends on $\Sigma$ only through $G$. Similarly, $\hat{U}_0'\hat{U}_0$ can be expressed as

(3.3) $\quad \hat{U}_0'\hat{U}_0 = G^{-1} W' M_0 W' (G^{-1})'$ ,

where $M_0 = I - X(X'X)^{-1} ( X'X - R'(R(X'X)^{-1}R')^{-1}R ) (X'X)^{-1}X'$. These observations yield the following basic result which allows one to derive finite sample tests based on the above criteria.

___

*Theorem 3.1*    *Under (2.1), (2.2a) and* $H_{01}$ *, the vector* $(m_1, m_2, ..., m_p)'$ *of the roots of (3.1) is distributed like the vector of the corresponding roots of*

*(3.4)* $\quad |W'MW - mW'M_0W| = 0$ ,

*where M is defined as in (2.7),* $M_0$ *as in (3.3),* $W$ *is defined in (2.2a), and the roots are put in descending order in both cases* .

___

PROOF: Let $M_0$ be defined as in (3.3) in the context of the constrained model. Then

$\hat{U}'\hat{U} = G^{-1} W' M W (G^{-1})'$ ,

$\hat{U}_0'\hat{U}_0 = G^{-1} W' M_0 W (G^{-1})'$ .

Consequently, the determinantal equation (3.1) can be expressed as:

$|G^{-1} W' M W (G^{-1})' - m G^{-1} W' M_0 W (G^{-1})'| = 0$ ,

hence $\quad |G^{-1}||W'\ M\ W\ -\ m\ W'\ M_0\ W||(G^{-1})'| \ =\ 0 \quad$ and

(3.5) $\quad |W'\ M\ W\ -\ m\ W'\ M_0\ W| \ =\ 0$ .

Since the rows of $W$ are $iid \sim N\ (\ 0,\ I_p\ )$ , the roots of equation (3.5) have distributions which does not involve $\Sigma$. Q.E.D

The above result implies that the joint distribution of $(m_1,\ ...\ ,\ m_p)$ does not depend on the regression coefficients nor does it involve nuisance parameters. Hence the test criteria obtained as functions of the roots are pivotal under the null and yield exact inference given assumption (2.2a). Although Theorem 3.1 is not explicitly stated by Anderson (1984) or Rao (1973), it is implicit in their demonstrations. Since an explicit proof of Theorem 3.1 is not apparently available, we supply one in the Appendix. On the basis of the above, the distribution of the Wilks $\wp$ criterion can be readily established.

*Corollary 3.2*    *Under (2.1), (2.2a) and* $H_{01}$ , *Wilks'* $\wp$ *statistic for testing* $H_{01}$ *is distributed like the product of the roots of* $|W'MW\ -\ mW'M_0W| \ =\ 0$ , *where M is defined as in (2.4),* $M_0$ *as in (3.3) and* $W$ *is defined in (2.2a).*

It may be useful, for simulation purposes, to restate Theorem (3.2) as follows.

*Corollary 3.3*    *Under (2.1), (2.2a) and* $H_{0l}$, *Wilks'* $\wp$ *statistic for testing* $H_{0l}$ *is distributed like* $|\ W'MW\ |\ /\ |\ W'M_0W\ |$ , *where M is defined as in (2.4),* $M_0$ *as in (3.3), and* $W$ *is defined in (2.2a).*

Note that the above characterization of the exact distribution do not require the normality assumption. Eventually, when the normality hypothesis (2.2b) holds, the distribution of the Wilks criterion, as stated in Theorem 3.4 below, is well known [Anderson (1984)].

_Theorem 3.4_      _Under (2.1), (2.2a) and $H_{01}$, Wilks' $\Lambda$ statistic for testing $H_{01}$ is distributed_

_like the product of p independent beta variables with parameters ( $(n - r_X - p + i)/2$, $r/2$ ),_

_$i = 1, ..., p$, where $r_X$ is the rank of the regressor matrix and r is the rank of the matrix R._

For non-Gaussian errors [_i.e._ when $W_t$ follows a known distribution which differs

from the $N(0, I_p)$ distribution], the null distribution of Wilks' statistic cannot be assessed

analytically. However, the above results can be used to obtain randomized or Monte Carlo

tests that are applicable given the general distributional assumption (2.2a). Such procedures

were originally suggested by Dwass (1957) and Barnard (1963). In the following, we briefly

outline the methodology involved as it applies to the present context; for a more detailed

discussion, see Dufour (1995), Dufour and Kiviet (1994, 1996) and Kiviet and Dufour

(1996).

Let $T_1$ denote the observed test statistic $T$, where $T$ is the adopted test criterion, for

instance LR, as defined in (3.2a). By Monte Carlo methods and for a given number $N$ of

replications, generate $T_j$, $j = 1, ..., N$ independent realizations of the statistic in

question, under the null, given a consistent estimator of the error covariance matrix.

Specifically, we consider several choices for the error covariance estimates, based on

restricted and unrestricted OLS coefficient estimates. While the level of the test is controlled

irrespective of the number of replications, the statistic typically performs better in terms of

power the larger the number of replications. Define $R_j(N)$ as the rank

of $T_j$ when $T_1, ..., T_j, ..., T_N$ are ranked in non-decreasing order $(j = 1, ..., N)$ ·

For $0 < \alpha < 1$, choose $C_N(\alpha)$ to be a positive real number such that

(3.6)      $C_N(\alpha) = 1 - ( I[N\alpha] / N ) + ( 1 / N )$ ,

where $I[x]$ is the largest integer less than or equal to $x$. Then the test's critical region

corresponds to

(3.7)      $R_1(N) / N \geq C_N(\alpha)$ ·

The latter critical value is exact given the assumptions of Theorems 3.1.

We now turn to the uniform mixed linear hypotheses $H_0$. A reparametrization of model (2.1) establishes the main distributional results for this problem. Indeed, it is worth noting that the maximum likelihood estimators [MLE] subject to $H_0$ may simply be obtained by maximizing the likelihood associated with:

$$(3.8) \qquad Y_c = XB_c + U_c \text{ ,}$$

where $Y_c = YC$, $B_c = BC$ and $U_c = UC$ with covariance $C'\Sigma C$ , subject to $RB_c = D$ . The resulting Wilks test statistic will satisfy the assumptions of Theorem 3.1-3.4. Hence the finite sample results established above also hold in the general uniform mixed linear case.

For certain values of $r$ and $c$ and normal errors, the null distribution of the Wilks criterion reduces to the $F$ distribution. For instances, if $\min(r, c) \leq 2$ , then

$$(3.9) \qquad \frac{1 - \mathcal{L}^{1/\tau}}{\mathcal{L}^{1/\tau}} \frac{\rho\tau - 2\lambda}{rc} \sim F(rc, \rho\tau - 2\lambda) \text{ ,}$$

where

$$\rho = \frac{n-k-(r-c+1)}{2}, \quad \lambda = \frac{rc-2}{4}, \quad \tau = \begin{cases} (r^2c^2-4)/(r^2+c^2-5), & \text{if } r^2+c^2-5 > 0 \\ 1, & \text{otherwise} \end{cases}.$$

Further, the special case $r = 1$ leads to the *Hotelling* $T^2$ criterion which is a monotone function of $\mathcal{L}$. If $r > 2$ and $c > 2$, then the distributional result (3.9) holds asymptotically [Rao 1973, chapter 8]. Stewart (1995b) provides an extensive discussion of these special $F$ tests.

In Section 6.1, we report simulations on Monte Carlo tests based on the above finite sample theory. For a proof of Theorem 3.4 and a review of asymptotic results pertaining to the criteria (3.2a-3.2d), the reader may consult Anderson (1984, chapter 8) or Rao (1971, chapter 8]. Finally, recall that not all linear hypotheses can be expressed as in $H_0$; we discuss other linear hypotheses in the following Section.

## 4. General linear hypotheses in the multivariate linear model

In this Section, we introduce a preliminary result relating to general hypothesis tests in MLR contexts. The issues we raise are of interest for their own sake and have a crucial bearing on hypothesis testing in the SURE framework. We consider the general case of $q^*$ independent restrictions on the coefficients of model (2.5) of the form

$$(4.1) \qquad H_{01} : R^*b \in \Delta_0 ,$$

where $Rank(R^*) = q^*$ , $\Delta_0$ is a non-empty subset of $\mathbb{R}^{q^*}$ . This characterization of the hypothesis includes linear restrictions, both within and across equations and allows for non-linear or inequality constraints. We adopt the LR testing procedure. Recognising the difficulty of obtaining exact critical values, we derive exact bounds on the null distribution of the LR statistic. Throughout this Section, we suppose that $\Sigma$ is not subject to restrictions, other than being in the class of positive definite symmetric matrices. We center our attention on the statistic

$$(4.2) \qquad \Lambda^* = \frac{|\hat{\Sigma}_{01}|}{|\hat{\Sigma}|} ,$$

where $\hat{\Sigma}_{01}$ , $\hat{\Sigma}$ maximize the log-likelihood function associated with (2.5) imposing and ignoring the restrictions in $H_{01}$ . The hypothesis is rejected when $\Lambda^*$ is suitably large. The LR statistic for testing $H_{01}$ is $n\ln(\Lambda^*)$ . As extensively discussed earlier, standard tests carried out on the basis of the above statistic will only have asymptotic validity. Indeed, it may be shown that the null distribution of $\Lambda^*$ depends on the error covariance matrix $\Sigma$ [Breusch (1980)]. Here we emphasize that $\Sigma$ is the only intervening nuisance parameter; as demonstrated in Breusch (1980), the null distributions of LR, Lagrange multiplier and Wald statistics in generalized linear models are invariant with respect to regression coefficients. Consequently, we proceed to derive finite sample, nuisance-parameter free bounds on the null distribution of $\Lambda^*$ . To do this, we shall extend the methodology proposed in Dufour (1989)

in the context of single equation linear models.

Consider the MLR model (2.5) and let $L(H_{12})$ denote the unrestricted maximum of the associated likelihood function. In the Gaussian model, $L(H_{12})$ is expressed by (2.8). Further, suppose we can find another set of restrictions of the form $\bar{R}BC = D$, that are UL in the notation of Section 3 and may be obtained as a special case of the restrictions in $H_{01}$. Then a hypothesis involving such UL restrictions would be nested within $H_{01}$. Formally, we have $H_{02} \subseteq H_{01}$ , where

(4.3)  $\quad H_{02} : \bar{R}BC = D$ ,  $\quad H_{01} : R^{*}b \in \Delta_{0}$ .

Now define $L(H_{0i})$ , $i = 1,2$ , to be the maximum of the log-likelihood function under $H_{0i}$. Given the normality assumption (2.2b), $L(H_{0i})$ , $i = 1,2$ can be expressed as

(4.4)  $\qquad L(H_{0i}) = -\dfrac{np}{2} \ln(2\pi) - \dfrac{n}{2} \ln(|\hat{\Sigma}_{0i}|) - \dfrac{np}{2}$ ,  $i = 1,2$ ,

provided $\hat{\Sigma}_{02}$ maximizes the likelihood under $H_{02}$ . Then it is straightforward to see that

(4.5)  $\qquad L(H_{02}) \leq L(H_{01}) \leq L(H_{12})$ .

Applying (4.4) to (4.5) gives

(4.6)  $\qquad \Lambda^{*} \leq \Lambda^{*}_{c}$ ,

where

(4.7)  $\qquad \Lambda^{*}_{c} = \dfrac{|\hat{\Sigma}_{02}|}{|\hat{\Sigma}|}$ .

In addition, the null distribution of $\Lambda^{*}_{c}$ may be obtained in finite samples following the results in Section 3. Most importantly, the null distribution does not depend on the error covariance matrix. Thus, it can be used to obtain critical values for $\Lambda^{*}$ . Indeed, (4.7) implies that, under the null,

(4.8)  $\qquad P [\Lambda^{*} \geq x ] \leq P [ \Lambda^{*}_{c} \geq x ]$ ,  $\forall x$ ,

where $P [ \Lambda^{*}_{c} \geq x ]$ does not depend on the nuisance parameter $\Sigma$. Under (2.2b) the null

distribution of $\Lambda_c^*$ involves the product of $p$ independent *beta* variables with degrees of freedom that depend on data, parameter and restriction counts and thus may easily be obtained by simulation. Let $\Psi_\alpha(\,.\,)$ be such that

(4.9) $\qquad P\,[\ \Psi(v_1\,,\,v_2\,,\,v_3)\geq \Psi_\alpha(v_1\,,\,v_2\,,\,v_3)\ ]\,=\,\alpha\ \ ,\ \ 0\leq\alpha\leq 1\ ,$

where $\Psi(v_1\,,\,v_2\,,\,v_3)$ is distributed like the product of the inverse of $v_2$ independent *beta* variables with parameters $(\ (v_1\,-\,v_2\,+\,i)/2\ ,\ v_3/2\ )\ ,\ i\,=\,1\ ,\,...\,,\,v_2$ . Then (4.8) may be rewritten as

(4.10) $\qquad P\,[\ \Lambda^*\geq\Psi_\alpha(n-k\,,\,p\,,\,\tilde{q})\ ]\leq\alpha\ \ ,\ \ 0\leq\alpha\leq 1\ ,$

where $\tilde{q}\,=\,min(r,\,c)\ \ ,\ \ r\,=\,rank(\tilde{R})\ ,\ c\,=\,rank(C)$ . Consequently, the critical value $Q_\alpha$ defined by

(4.11) $\qquad Q_\alpha\,=\,\Psi_\alpha(n-k\,,\,p\,,\,\tilde{q})\ ,$

is conservative at level $\alpha$. Of course, one should seek the smallest critical bound possible. This would mean expressing $\tilde{R}$ so that $\tilde{q}$ is as small as possible. We proceed next to state our main conclusion for the normal model.

*Theorem 4.1*　　　　*Consider the MLR model (2.5) under (2.2b). Let* $\Lambda^*$ *be the statistic defined by (4.2) for testing* $R^*b\,\in\,\Delta_0$ *where* $R^*$ *is a* $(q^*\,,\,k)$ *full column rank matrix and* $\Delta_0$ *is a non-empty subset of* $\mathbf{R}^{q^*}$ *. Further, consider restrictions of the form* $\tilde{R}BC\,=\,D$ *that satisfy* $R^*b\,\in\,\Delta_0$ *. Then, under the null, for all* $0\leq\alpha\leq 1$ *,* $P\,[\ \Lambda^*\geq\Psi_\alpha(n-k\,,\,p\,,\,\tilde{q})\ ]\leq\alpha$ *, where* $\tilde{q}\,=\,min(r,c)\ \ ,\ \ r\,=\,rank(\tilde{R})\ ,\ c\,=\,rank(C)$ *and* $\Psi_\alpha(\,.\,)$ *is defined by (4.9).*

At this point, it is worth noting that normality, *i.e.* hypothesis (2.2b) by no way constitutes a necessary assumption in this case. Indeed, the critical values of the bounding statistic may still be determined by simulation under the general assumption (2.2a). Inequality (4.7) results from the properties of maximum likelihood estimation irrespective of the

likelihood density function. For the purpose of generality, we restate next our main result for model (2.5) given the distributional assumption (2.2a).

_Theorem 4.2_        _Consider the MLR model (2.5) under (2.2a). Let_ $\Lambda^*$ _be the Wilks statistic defined by (4.2) for testing_ $R^*b \in \Delta_0$ _where_ $R^*$ _is a_ $(q^*,k)$ _full column rank matrix and_ $\Delta_0$ _is a non-empty subset of_ $\mathbb{R}^{q^*}$. _Further, consider restrictions of the form_ $\bar{R}BC = D$ _that satisfy_ $R^*b \in \Delta_0$ _with_ $\bar{q} = min(r,c)$ , $r = rank(\bar{R})$ , $c = rank(C)$. _Let_ $\Lambda_c^*$ _be the Wilks criterion for testing the latter restrictions. Then under the null,_

$$P [ \Lambda^* \geq \lambda_c^*(\alpha) ] \leq \alpha \quad , \quad for\ all\ 0 \leq \alpha \leq 1 \ , where\ \lambda_c^*(\alpha)\ is\ determined\ such\ that$$

$$P [ \Lambda_c^* \geq \lambda_c^*(\alpha) ] = \alpha \ .$$

Clearly, the above results hold when the hypothesis is linear of the form $R^*b = \delta_0$ . In addition, the fact that the null distribution of the LR statistic can be bounded (in a non trivial way) as in Theorem 4.1 and Corollary 4.2 implies that simulation based techniques may be used to obtain valid inference based on the statistic in (4.2) when the bounds test is not conclusive. Dufour (1995) presents the large and finite sample theory underlying Monte Carlo tests in the presence of nuisance parameters. The methodology involved is basically as described in the previous Section. However, since the null distribution of $\Lambda^*$ is not generally free of nuisance parameters, the critical region defined in (3.6) is not provably exact. To obtain an exact critical region, the p-value associated with (3.6) ought to be maximized with respect to the elements of the error covariance matrix $\Sigma$. Whenever the last step involving the numerical optimization of the randomized p-value is not carried out, the method just outlined is closely related to a parametric bootstrap. Indeed, as demonstrated in Dufour (1995), the critical region defined by (3.6) has the correct level under asymptotic considerations whenever the asymptotic distribution of the randomized statistic depends continuously on the intervening nuisance parameters. This property clearly holds in this case

since the LR statistic is asymptotically pivotal. Furthermore, we must observe that the critical region defined in (3.6) may be viewed as exact in a liberal sense, *i.e.* if the test based on the latter critical region fails to reject, we can be sure that the exact test involving the maximum $p$-value is not significant at level $\alpha$. We emphasize the fact that the proposed randomized test can be implemented in complementarity with the above defined bounds tests. If the conservative test rejects the null then the LR test is most certainly significant. These issues will be taken up further in Section 7, in the framework of a Monte Carlo experiment. We next employ a well known example from the finance literature to illustrate how the above results may be used.

A fundamental problem in financial economics involves testing the mean-variance efficiency of a candidate benchmark portfolio. Let $R_{jt}$, $R_{kt}$, $j = 1, ..., N$, $k = 1, ..., K$ be security returns for period $t$, $t = 1, ..., T$. The hypothesis of interest is that some portfolio of the $K$ security subset is efficient with respect to the total set of $N+K$ securities. If it is assumed that a riskless asset $R_F$ exists, then efficiency can be tested based on the following multivariate regression

$$(4.12) \quad r_{jt} = \alpha_j + \sum_{k=1}^{K} \beta_{jk} \, r_{kt} + e_{jt} \quad , \quad j = 1, ..., N \quad , \quad t = 1, ..., T \, ,$$

where $r_{jt} = R_{jt} - R_{Ft}$, $r_{kt} = R_{kt} - R_{Ft}$. The hypothesis of efficiency implies that the intercepts $\alpha_j$ are jointly equal to zero, *i.e.*

$$(4.13) \quad \alpha_j = 0 \, , \, j = 1, ..., N \, .$$

A well known example of (1.1) is the *capital asset pricing model* (CAPM)

$$(4.14) \quad R_{jt} - R_{Ft} = \alpha_j + \beta_j \, (R_{Mt} - R_{Ft}) + e_{jt} \quad , \quad j = 1, ..., N \quad , \quad t = 1, ..., T \, ,$$

where $R_{Mt}$ are the returns on the market benchmark. Gibbons, Ross and Shanken [GRS] (1989) show that a transformation of the LR criterion to test (4.13) has an exact $F$ distribution given normality of asset returns. Mackinlay (1987) proposes a similar statistic in

the context of a single beta CAPM. Specifically, GRS suggest the following test statistic:

$$(4.15) \quad Q = \frac{T \hat{\alpha}' \hat{S}^{-1} \hat{\alpha}}{1 + \bar{r}' \hat{\Delta}^{-1} \bar{r}} \quad ,$$

where $\hat{\alpha}$ is the vector of intercept OLS estimates, $\hat{S} = \hat{\Sigma} \, T/(T-K-1)$ is the unbiased estimator of $\Sigma$, $\bar{r} = (\bar{r}_1, \ldots, \bar{r}_K)'$ is the vector of time series means for $r_t = (r_{1t}, \ldots, r_{Kt})'$, and $\hat{\Delta}$ is the sample covariance matrix for $r_t$. Under (4.13), $Q$ has the *Hotelling* $T^2(N, T-K-1)$ distribution or alternatively,

$$(4.16) \quad \frac{Q (T-K-N)}{N (T-K-1)} \sim F (N, T-K-N) \quad .$$

Let $\Lambda_c$ denote the statistic defined by (4.2) in this context. It can be shown [see, for example Stewart (1995b)] that $\Lambda_c$ is related to the GRS criterion as follows:

$$(4.17) \quad \Lambda_c - 1 = Q / (T-K-1) \quad .$$

The econometric analysis is more complicated when the zero beta intercept is unknown and must be inferred using the return data [see, for example Gibbons (1982)]. In this case, the excess-return MLR becomes

$$(4.18) \quad R_{jt} - \gamma = \alpha_j + \sum_{k=1}^{K} \beta_{jk} (R_{kt} - \gamma) + e_{jt} \quad , \quad j = 1, \ldots, N \quad , \quad t = 1, \ldots, T \quad ,$$

where $\gamma$ is the unknown zero-beta intercept. Under the hypothesis of mean-variance efficiency, there exists a scalar $\gamma$ such that the vector of $\alpha_j$ is equal to zero. This implies the following non-linear constraint on the raw-return MLR

$$(4.19) \quad R_{jt} = a_j + \sum_{k=1}^{K} \beta_{jk} R_{kt} + e_{jt} \quad , \quad j = 1, \ldots, N \quad , \quad t = 1, \ldots, T \quad ,$$

$$(4.20) \qquad a_j = \gamma \left( 1 - \sum_{k=1}^{K} \beta_{jk} \right) \, , \, j = 1 \, , \, \dots \, , \, N \, \cdot$$

Suppose that we wish to test (4.20) and let $\Lambda^*$ denote the associated statistic from (4.2). Exact tests for this specific problem have been studied by Shanken (1986) and more recently by Stewart (1995b). In what follows, we show that the exact procedures in question may be obtained as an application of our general methodology.

Shanken (1986) employs the statistic $Q(\hat{\gamma})$ , where, in the context of (1.8)

$$(4.21) \qquad Q(\gamma) = \frac{T \, \hat{a}'(\gamma) \, \hat{S}^{-1} \, \hat{a}(\gamma)}{1 + (\bar{R} - \gamma \iota_K)' \, \hat{\Delta}^{-1} \, (\bar{R} - \gamma \iota_K)} \, ,$$

where $\hat{\gamma} = ARGMIN_{\gamma} Q(\gamma)$ , $\hat{a}(\gamma) = \hat{a} - \gamma(\iota_N - \hat{\beta}\iota_K)$ , $\hat{a}$ is the vector of intercept estimates, $\hat{\beta}$ is the $(N,K)$ matrix of OLS *beta* estimates, $\hat{\Sigma}$ is the unbiased estimate of $\Sigma$, $\bar{R} = (\bar{R}_1 , \dots , \bar{R}_K)'$ is the vector of time series means for $R_t = (R_{1t} , \dots , R_{Kt})'$ , $\hat{\Delta}$ is the sample covariance matrix for $R_t$ and $\iota_J$ denotes a vector of $J$ 1's. Shanken shows that (i) the LR statistic for testing (4.20) is a transformation of $Q(\hat{\gamma})$ , (ii) $\hat{\gamma}$ is the constrained maximum likelihood estimator of $\gamma$ , and (iii) the null distribution of $Q(\hat{\gamma})$ may be bounded by the *Hotelling* $T^2(N, T-K-1)$ distribution. Turning to our proposed bound on the statistic $\Lambda^*$ , we suggest to consider the statistic $\Lambda_c$ associated with the special case of (4.20) where $\gamma$ is any known constant. By (4.16), (4.17) and using (4.5), this naturally leads to the use of conservative critical points involving the $F$ $(N, T-K-N)$ distribution. This is the same result obtained by Shanken (1986) and Stewart (1995b).

## 5. Hypothesis testing in the Seemingly Unrelated Regressions model

This section considers testing hypotheses about the parameters of the SURE model. Indeed, the results for the MLR model furnish interesting applications for systems inference in SURE model. We begin by deriving, along the lines of the previous Section, a conservative bound

on the distribution of the LR statistic for testing arbitrary hypotheses. We then deal with general linear hypotheses. Bounds tests and the Monte Carlo approach are combined to develop exact inference techniques.

Consider first the problem of testing in the context of model (2.12) the general hypothesis

$$(5.1) \qquad H_0 : C^* \beta \in \Delta_0^* ,$$

where $rank(C^*) = v_0^*$ , $\Delta_0^*$ is a non-empty subset of $\mathbb{R}^{v_0^*}$ . In terms of the MLR model (2.19) which includes (2.12) as a special case, the restrictions in $H_0$ may be stated as

$$(5.2) \qquad H_{01} : C\tilde{\beta} \in \Delta_0 ,$$

where $rank(C) = v_0$ , $\Delta_0$ is a non-empty subset of $\mathbb{R}^{v_0}$ and $C$ is expressed so that it incorporates the SURE restrictions (2.16). The usual LR tests of $H_{01}$ are based on the restricted and unrestricted estimators of the error covariance matrix. Let us denote by $\hat{\Sigma}_{01}$ and $\hat{\Sigma}$ the maximum likelihood estimators of $\Sigma$ associated with (2.19) respectively imposing and ignoring the restrictions $H_{01}$ . Further, let $\hat{\Sigma}_{11}$ be the maximum likelihood estimators of $\Sigma$ under the SURE restrictions (2.16). In this case, the usual LR statistic is a monotonic transformation of

$$(5.3) \qquad \Lambda = \frac{|\hat{\Sigma}_{01}|}{|\hat{\Sigma}_{11}|} .$$

For the purpose of deriving the conservative bound, we introduce another LR based statistic, namely

$$(5.4) \qquad \Lambda^* = \frac{|\hat{\Sigma}_{01}|}{|\hat{\Sigma}|} .$$

As it stands, testing $H_{01}$ based on the LR criterion (5.4) is exactly the type of problem discussed in Section 4. Here, one simply needs to consider a UL hypothesis, hereafter denoted by $H_{02}$ such that $H_{02} \subseteq H_{01}$ . The associated LR statistic is

$$(5.5) \qquad \Lambda_c^* = \frac{|\hat{\Sigma}_{02}|}{|\hat{\Sigma}|} \, ,$$

provided $\hat{\Sigma}_{02}$ maximizes the log likelihood under $H_{02}$. As established in Section 3, the exact null distribution of $\Lambda_c^*$ is nuisance-parameter-free and may be easily simulated. Conformably with the notation in Section 4, let $L(H_{0i})$, $i = 1,2$ be the maximum of the likelihood function under $H_{0i}$. Further, let $L(H_{1i})$, $i = 1,2$ refer to the maximum of the log likelihood under (2.16) and the unrestricted maximum, respectively. Hence, the following inequality holds under the general distributional assumption (2.2a):

$$(5.6) \qquad L(H_{02}) \leq L(H_{01}) \leq L(H_{11}) \leq L(H_{12}) \, .$$

Consequently, it is straightforward to show that

$$(5.7) \qquad \Lambda \leq \Lambda^* \leq \Lambda_c^* \, .$$

The critical bound may be accordingly obtained from the null distribution of $\Lambda_c^*$ as described in Section 4. To facilitate the analysis, we shall, in the following, provide an illustrative example.

_Example 5.1_    In the three equations SURE model with Gaussian errors

$$(5.8) \qquad Y_1 = \beta_{10} + \beta_{11} X_1 + U_1$$

$$Y_2 = \beta_{20} + \beta_{22} X_2 + U_2$$

$$Y_3 = \beta_{30} + \beta_{33} X_3 + U_3 \, ,$$

consider testing $H_0 : \beta_{11} = \beta_{22} = \beta_{33}$. In terms of the corresponding MLR model

$$(5.9) \qquad Y_1 = \beta_{10} + \beta_{11} X_1 + \beta_{12} X_2 + \beta_{13} X_{13} + U_1$$

$$Y_2 = \beta_{20} + \beta_{21} X_1 + \beta_{22} X_2 + \beta_{23} X_3 + U_2$$

$$Y_3 = \beta_{30} + \beta_{31} X_1 + \beta_{32} X_2 + \beta_{33} X_3 + U_3 \, ,$$

the problem implies testing the simultaneous hypothesis $H_0 : \beta_{11} = \beta_{22} = \beta_{33}$ and $\beta_{12} = \beta_{13} = \beta_{21} = \beta_{23} = \beta_{31} = \beta_{32} = 0$. In order to use the above results on the conservative bound, we need to construct a set of UL restrictions in the sense of Section 3

that satisfy the hypothesis in question. It is easy to see that the constraints setting the coefficients $\beta_{ij}$, $i$, $j = 1$, ... , 3 to specific values do serve the purpose. All that remains is to calculate Wilks' statistic conforming with (5.3) and use the critical value defined by (4.11) as a conservative cut-off point.

Having presented our basic result in terms of arbitrary hypotheses, let us now take up the special case of linear constraints in the context of the SURE model (2.12). We wish to treat linear restrictions both within and across equation *i.e.* restrictions of the form

$$(5.10) \quad C^{*}\beta = \delta_{0}^{*} \ ,$$

where $rank(C^{*}) = v_{0}^{*}$ . The MLR model (2.19) reduces to (2.12) when the restrictions (2.16) are imposed. Accordingly, we often refer to the equivalent problem of testing, in the MLR setup (2.19), a set of independent linear constraints of the form

$$(5.11) \quad C\beta = \delta_{0} \ ,$$

where $rank(C) = v_{0}$ and $C$ is formulated so that it incorporates (2.16). Several relevant test criteria have been suggested in the SURE literature [see, for example, Srivastava and Giles (1987, chapter 10)]. Among those, we cite the LR statistic as expressed in (5.3) and the following well-known Wald-type statistic, defined in the notation of (2.12) and (5.10) as:

$$(5.12) \quad \tau = \left(\frac{v_{1}}{v_{0}^{*}}\right) \frac{\left(C^{*}\hat{\beta} - \delta_{0}^{*}\right)' \left[C^{*}\left(X^{*'}\left(S^{-1} \otimes I_{n}\right)X^{*}\right)^{-1}C^{*'}\right]^{-1} \left(C^{*}\hat{\beta} - \delta_{0}^{*}\right)}{\left(y - X^{*}\hat{\beta}\right)' \left(S^{-1} \otimes I_{n}\right) \left(y - X^{*}\hat{\beta}\right)}$$

where $v_{1} = np - k^{*}$, $S$ is a consistent estimator of the error covariance matrix and $\hat{\beta}$ is the feasible generalized least squares estimate (FGLS) of $\beta$ :

$$(5.13) \quad \hat{\beta} = \left[X^{*'}\left(S^{-1}\otimes I_{n}\right)X^{*}\right]^{-1}X^{*'}\left(S^{-1}\otimes I_{n}\right)y \ .$$

Under the null, $v_{0}^{*}\tau$ has an asymptotic distribution that is chi-square with $v_{0}^{*}$ degrees of freedom. Theil (1971, chapter 6) advises that the $F$ distribution better captures the finite sample properties of the statistic in (5.12). Yet this claim is not supported by neither analytical nor simulation evidence. Although there are many possible choices for $S$, two ways

of constructing consistent estimates are common, each of which relies on residuals obtained beginning with OLS. The first approach involves so-called "restricted residuals" [Srivastava and Giles (1987, chapter 2)], *i.e.* the least-squares residuals for model (2.19) restricted by the SURE constraints (2.16). Clearly, these correspond to the OLS residuals from the regression (2.12). With this replacement for $S$, $\hat{\beta}$ coincides with the two-step Zellner (1962) estimator. Iterative FGLS estimators may be substituted for $\hat{\beta}$ **and** $S$ in the formulae for the Wald criterion. Incidentally, ML estimation may be implemented through iterating to desired convergence the FGLS estimators cited above. In relation, see Srivastava and Giles (1987, chapter 5). Hereafter, we denote the statistic in (5.12) based on the two-step parameter estimates $\tau_F$ , to which we refer by *two-step Wald* statistic. The criterion (5.12) obtained with iterative estimates we denote $\tau_T$ and call it the *iterative Wald* statistic.

We also introduce another class of LR-type statistics justified on the basis of computational cost as opposed to those relying on full ML estimation. More specifically, let $\hat{\Sigma}_{ii}^{(j)}$ , $i = 0$ , $1$ denote the iterative estimators of $\Sigma$ imposing (5.11) and (2.16) respectively, at the $j$-th step of the iteration process. Define the statistic

$$(5.14) \qquad \Lambda_{(j)} = \frac{|\hat{\Sigma}_{01}^{(j)}|}{|\hat{\Sigma}_{11}^{(j)}|} \ ,$$

where the superscript $j$ refers to the number of iterations involved in the derivation of the error covariance estimates and may assume any value ranging from zero to the maximum required for convergence.

Proceeding as outlined in the previous Section, randomized critical points may be obtained for all the above suggested test criteria. It is evident that statistics resting on estimators necessitating fewer iterations have a broader scope of applicability in the context of Monte Carlo methods. We next present the basic result underlying the asymptotic validity of the bootstrap method for LR type statistics based on consistent restricted and unrestricted

parameter estimates that do not necessarily maximise the likelihood. Whereas the null distribution of the LR criterion is asymptotically pivotal, the limiting null distribution of such statistics may depend on nuisance parameters. As demonstrated in Dufour (1995), it is sufficient that the limiting distribution satisfy the following property.

*Proposition 5.1*        *Consider the MLR model (2.19). Let* $\Lambda_{(i)}$ *be the LR-type statistic defined by (5.14) for testing* $C^{*\prime}\beta = \delta_0$ *where* $rank(C^*) = v_0^*$ *and* $C^*$ *incorporates the SURE restrictions (2.16) against the SURE specification. Then the asymptotic null distribution of the statistic depends continuously on the error covariance matrix.*

The latter proposition derives from the following well known result: the above defined two-step and iterative estimators have the same asymptotic distribution as the MLE; see, for example Srivastava et al. (1987, chapters 5 and 10). Then, under standard regularity conditions, the quasi-LR statistics follow usual limiting chi-square distributions.

## 6. A simulation study

This Section reports an investigation, by simulation, of the performance of the various proposed statistics in the context of SURE and SEM models. All the experiments were conducted using Gauss-386i VM version 3.1 and each was based on 1000 replications.

### 6.1 MLR model with uniform linear hypothesis

To give an idea of the value of randomized tests, a Monte Carlo simulation experiment was conducted for the MLR model with the same restrictions on each equation. The assumed model was

$$
(6.1) \qquad Y_{ij} = \beta_{j0} + \sum_{k=1}^{m} \beta_{jk} x_{ik} + u_{ij} \, , \quad i = 1 \, , \, ... \, , \, n \quad , \quad j = 1 \, , \, ... \, , \, m \, ,
$$

where $n$ refers to the sample size and $m$ to the number of equations in the system. The restrictions tested were of the form

$$(6.2) \qquad \sum_{k=1}^{m} \beta_{jk} = 0 \ , \quad j = 1 \ , \ ... \ , m \ \cdot$$

The model illustrates the problem of homogeneity testing in demand systems as noted in Section 1. Attfield (1995) considers the model for $m = 2$ regarding the Bartlett adjusted LR test. Here, we reexamined the example to provide evidence that the randomized procedures do control test sizes in instances where the Bartlett correction fails to reduce the bias, particularly in larger systems. Indeed, the experiment was carried out for 5, 7 and 8 equations models. The sample sizes were set to 20, 25, 40, 50 and 100. The matrices of fixed regressors were independently drawn from the normal distribution with means (65, 35, 45, 45, 35, 55, 45, 50) and standard deviations (.5, .6, .2, .3, .6, .4, .3, .2). The errors were generated as $NID(0,\Sigma)$ , with $\Sigma = GG'$ and $G$ randomly drawn from the normal distribution. The regression coefficients are reported in Table 1a. The LR and adjusted LR statistics were calculated and denoted $\Lambda$ and $\Lambda/c$ where the Bartlett correction $c$ was derived following Attfield (1995, Section 3.3) as:

$$(6.3) \quad c = \qquad 1 + 7/n \ , \quad \text{for} \ \ m = 5$$
$$1 + 9/n \ , \quad \text{for} \ \ m = 7$$
$$1 + 10/n \ , \quad \text{for} \ \ m = 8 \ \cdot$$

For the LR and adjusted LR criteria, the percentage of the 1000 replications greater than the 5% critical chi-square values were evaluated. Further, the finite sample critical points were obtained by simulation following the lines of Section 3 and the percentage of replications greater than the randomized cut-off points were calculated. As emphasized in Section 3, the null distribution of the LR statistic is nuisance-parameter free in this case hence the critical points obtained by simulation are exact. The results are summarized in Table 1b.

The experiment has three main conclusions. First, the results on the LR test conform to the well documented fact in this context namely of severe overrejection. Indeed, for $m = 8$ and $n = 20, 25$ the empirical sizes were observed to be 75.3 and 47.6% respectively. Second, the results on the adjusted LR imply that the Bartlett correction, though providing some improvement, does not correct the bias in larger systems. For $m = 8$ and $n = 20, 25$ the empirical sizes after the correction remained at 49.3 and 24.0%. In contrast, we finally show that the randomized testing procedure corrects the levels in all cases examined.

We have also conducted another experiment to investigate the accuracy of the asymptotic $F$ test (3.9) where $r > 2$ and $c > 2$. Several choices for the number of equations ($p$), the number of regressors ($k$), $r$ and $c$ were considered. The regressors and the error covariance matrices were selected as described above. In all cases the regression coefficients and the matrices $R$ and $C$ were drawn at random. The results are reported in Table 8.c. We observe that the asymptotic Chi-square approximation is extremely poor; the asymptotic F test performs relatively better but size correction is still needed. The Monte Carlo test achieves size control.

### 6.2 MLR model with cross-equation restrictions

Considering the same MLR model as in (6.1) with $m = 3, 5$, we also studied tests of the following hypothesis:

(6.4) $\quad \beta_{jj} = \beta_{11}$ , $j = 2, \dots , m$ and $\beta_{jk} = 0$ , $j \neq k$ , $j , k = 1 , \dots , m$ .

The coefficients for this example are presented in Table 2a. For each Monte Carlo trial, the LR and the conservative bound were calculated as in (4.2) and (4.8), and the observed significance was computed using the asymptotic distribution and the randomized critical region (3.6) as outlined in Section 4, for a nominal significance level of 5%. In addition, the power of the tests was investigated by simulating the model under alternative values for the regression parameter of the first equation, namely $\beta_{11}$ , in both systems. For the purpose

of power comparisons, the sizes of the asymptotic tests were *locally corrected, i.e.* an independent simulation was conducted for the same parameter choices as the initial experiment to determine empirical 5% cut-off points. The randomized tests were applied with 20 and 100 replications. To generate the independent realizations of the randomized criteria, we experimented with several consistent estimates of the error covariance matrix relying on restricted and unrestricted OLS coefficient estimates. Since the results were insensitive to choice of consistent estimator for $\Sigma$, we report only the results based on restricted estimates. Tables 2b to 2d summarize our findings. An apparent implication is that the conservative bound provides conclusive results in a large proportion of cases. Further, the parametric bootstrap provides substantial improvement for inferences over the conventional asymptotic techniques. Indeed, the randomized procedure corrects the test size with no substantial power loss. Increasing the number of equations does not have a great effect on the relative performance of the methods proposed.

*6.3 Monte Carlo evidence: the SURE model*

Two Gaussian SURE models, modelled after Example 4.1, were used for the study. Systems involving three equations and five equations were considered to which we will refer as the 3EQ and the 5EQ models, respectively. Each equation includes an intercept term and one fixed regressor. The regressors were independently drawn from the normal distribution with means (65, 35, 45, 45, 35) and standard deviations (.5, .6, .2, .3, .6). The sample size was set to 25. The error covariance matrices were also randomly drawn as in Section 6.1. The regression parameters were $(1.2, .1, .8, .1, -1.1, .1)'$ for the 3EQ model and $(1.2, .1, .8, .1, -1.1, .1, 1.9, .1, -.2, .1)'$ for the 5EQ case. Clearly, the results are invariant to the true values of the regression coefficients. The restrictions were as in $H_0$, Example 4.1, i.e. involve testing the equality of the equations' regression coefficients, apart from the intercepts. We experimented with the following

statistics: $\Lambda$ , $\Lambda^*$ , $\tau_F$ , $\tau_T$ and $\Lambda_{(j)}$ , $j = 0$ , $1$ , $3$ as defined by (5.3), (5.4), (5.12) and (5.14) respectively. The conservative bound was evaluated based on the statistic in (5.5). Further, criteria inspired by those suggested in Theil et al (1985) were also studied. For the 3EQ, we considered:

(6.5)
$$\mu_{31} = |\hat{\beta}_{11} - \hat{\beta}_{22}| + |\hat{\beta}_{22} - \hat{\beta}_{33}|$$
$$\mu_{32} = |\hat{\beta}_{11} - \hat{\beta}_{33}| + |\hat{\beta}_{22} - \hat{\beta}_{33}|$$
$$\mu_{33} = |\hat{\beta}_{11} - \hat{\beta}_{22}| + |\hat{\beta}_{11} - \hat{\beta}_{33}| \ .$$

In the 5EQ case, the following were selected among the many possible choices

(6.6)
$$\mu_{51} = |\hat{\beta}_{11} - \hat{\beta}_{22}| + |\hat{\beta}_{22} - \hat{\beta}_{33}| + |\hat{\beta}_{33} - \hat{\beta}_{44}| + |\hat{\beta}_{44} - \hat{\beta}_{55}|$$
$$\mu_{52} = |\hat{\beta}_{22} - \hat{\beta}_{33}| + |\hat{\beta}_{33} - \hat{\beta}_{44}| + |\hat{\beta}_{44} - \hat{\beta}_{55}| + |\hat{\beta}_{55} - \hat{\beta}_{11}|$$
$$\mu_{53} = |\hat{\beta}_{33} - \hat{\beta}_{44}| + |\hat{\beta}_{44} - \hat{\beta}_{55}| + |\hat{\beta}_{55} - \hat{\beta}_{11}| + |\hat{\beta}_{11} - \hat{\beta}_{22}|$$
$$\mu_{54} = |\hat{\beta}_{44} - \hat{\beta}_{55}| + |\hat{\beta}_{55} - \hat{\beta}_{11}| + |\hat{\beta}_{11} - \hat{\beta}_{22}| + |\hat{\beta}_{22} - \hat{\beta}_{33}|$$
$$\mu_{55} = |\hat{\beta}_{55} - \hat{\beta}_{11}| + |\hat{\beta}_{11} - \hat{\beta}_{22}| + |\hat{\beta}_{22} - \hat{\beta}_{33}| + |\hat{\beta}_{33} - \hat{\beta}_{44}| \ .$$

For each trial, the various statistics were calculated and the observed significance was computed using the asymptotic distribution and the randomized cut-off points (3.6), for a nominal significance level of 5%. Though we did not analytically establish the asymptotic distribution of the criteria $\Lambda_{(j)}$ , we assessed their asymptotic significance using the chi-square reference distribution for the usual LR statistic. In addition, the power of the various tests was investigated by simulating the model under alternative values for the regression parameter of the first equation, namely $\beta_{11}$ , in both 3EQ and 5EQ systems. For the purpose of power comparisons, the sizes of the asymptotic tests were *locally corrected* as explained in Section 7.2. The randomized tests were applied with 20 and 100 replications. Several consistent estimates of the error covariance matrix relying on restricted and unrestricted OLS and two-step GLS coefficient estimates were considered. As in the general MLR case, the results were insensitive to choice of consistent estimator for $\Sigma$, hence we report only the results based on restricted GLS estimates in Table 3 to Table 7b. Although the Monte Carlo

experiments are conditional on the selected design and the values of the covariance matrices, our results show the following:

(1)    The asymptotic criteria have an upward bias in size; as can be seen in Table 3, rejection of the null is repeatedly many times more than what it should be. The bias clearly worsens in the 5EQ example. Across the cases examined, the Wald-type statistics give larger sizes when referred to the chi-square distribution. Although the $F$ approximation seems to correct the problem in the 3EQ model, it clearly fails to do so in the 5EQ case.

(2)    The conservative statistic was found to be well behaved. Power gains are possible in other test problems where a tighter critical bound is available. Indeed, we have observed reasonable power even if we have experimented with the worst scenario, in the sense that bounding test statistics correspond to a null hypothesis which fixes the values of all regression coefficients (except the intercept). This illustrates the value of the conservative test as a tool to be used in conjunction with Monte Carlo methods and not necessarily as an alternative to those methods. As emphasized earlier, the bounds procedure is computationaly inexpensive and exact. In addition, whenever the bounds test rejects, inference may be made without further appeal to randomized tests.

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

(4)    The Monte Carlo tests performed noticeably well in terms of power in all instances, even when the number of replications was as low as 20. We emphasize that the size-corrected asymptotic tests are *unavailable* in practical testing situations since the *local*

correction it entails requires that $\Sigma$ be known. The statistic with the best power properties across the alternatives examined was the randomized LR.

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

(6)    Simulation evidence does not favour the randomized usual LR tests over those based on $\Lambda_{(i)}$ typically involving fewer iterations, although we are uncertain as to the asymptotic equivalence of both procedures. This observation has an important bearing on empirical practice. The simplicity of the method based on $\Lambda_{(i)}$ has much to recommend it for larger models in which statistics requiring full MLE may be quite expensive to randomize.

## 6.4 Non-linear hypotheses

To conclude this section, we report the following studies that treat non-linear hypotheses. First, an experiment was modeled after the MLR system (4.19) under the null hypothesis (4.20) with $N = 40$, $T = 60$ and $K = 1$. Random errors were generated as $NID(0,\Sigma)$, with $\Sigma = GG'$ and $G$ drawn from the normal distribution. The coefficients were also drawn at random from the normal distribution with mean zero and variance .16; the parameter $\gamma$ was set to .009. To derive the LR statistic, the constrained MLE was numerically computed according to Shanken (1986). Empirical rejections were calculated using the asymptotic $\chi^2(N-1)$ distribution, the Monte Carlo cut-off points and the $\mathcal{F}(N, T-K-N)$ bound. For a nominal level of 5%, the observed size of the asymptotic test was 89.5%. As was the case with linear hypotheses, the asymptotic tests severely overreject. In contrast, the MC and bounds test had the correct sizes, .047 and .038 respectively.

Secondly, another experiment was set up based on the following SURE system

(6.7)     $Y_{it} = \delta_i + \alpha_i X_t^M + \beta_i X_{it},$   $i = 1, \ldots, N,$   $t = 1, \ldots, T$ ,

where $N = 7$ and $T = 25$. The regressor denoted $X^M$ is common to all equations. The null hypothesis was

(6.8)     $H_0$: $\alpha_i = \gamma \beta_i,$   $i = 1, \ldots, N,$   $\gamma$ unknown.

The parameters and regressors and random errors were generated as just described. Denote the stacked coefficient vector ß. For this experiment, we consider a quasi-LR (QLR) test derived as follows. The unconstrained estimate of ß, say $\hat{\beta}$ , is the two-step Zellner FGLS estimator. The constrained estimate of $\gamma$ is evaluated so as to minimize the following quadratic form

(6.9)     $\hat{\gamma} = ARGMIN$ $[w(\gamma)],$   $w(\gamma) = (R_{(\gamma)}\hat{\beta})'[R_{(\gamma)} COV(\hat{\beta}) R_{(\gamma)}]^{-1}(R_{(\gamma)}\hat{\beta})$ ,

where $R_{(\gamma)} = (I_{(7)} \otimes [0\ 1\ -\gamma])$ . The restricted and unrestricted residuals are obtained conformably and yield a QLR statistic. We have also examined a Wald-type criterion computed as $w(\hat{\gamma})$ . The conventional $\chi^2(N-1)$ asymptotic critical value was adopted. The observed empirical sizes of the Wald and QLR statistics were 32% and 24.6% respectively, whereas the levels of the Monte Carlo Wald and QLR statistic (7.7% and 2.6% respectively) are adequate.

## 7. Applications

In this Section, we present an empirical application that illustrates the results presented in this paper. Fischer (1993) presents panel regressions to examine the effect of macroeconomic factors on growth, specifically GDP growth, the rate of capital accumulation, productivity growth (measured by the Solow residuals) and the labor force growth. Four determinants of growth are specified: the inflation rate (INFLT), the ratio of budget surplus to GDP (SRPLS), the terms of trade (TTRD) and the black market premium on foreign exchange (EXPM). Regional dummies are included but observed not to affect conclusions. A random coefficient

unbalanced panel regression model is fitted using a data set on a large sample of countries. Here, we center our attention on the multiple regressions (17), (23) and (29) that include all four explanatory variables, i.e. we consider the sample of countries on which data is available on all variables. We estimate the system of growth equations when we cannot rule out the possibility that random disturbances are contemporaneously correlated across countries, within three geographical regions[1]. We also test the equality of coefficients across countries, a maintained hypothesis given the panel regressions. We apply and compare both the results of the usual asymptotic tests and the Monte Carlo techniques discussed in this paper. Furthermore, we provide the empirical rejection percentages associated with the asymptotic statistic, using the simulation experiment required to perform the Monte Carlo tests. Since we consider a balanced data set for the SURE systems, we actually use a smaller sample than in Fischer's analysis; for illustrative purposes, we reestimate the latter regressions based on our subsample. We report two Monte Carlo $p$-values, depending on the parameter values adopted for the simulation. As was emphasized earlier, consistent estimates that satisfy the null in question are needed. It is natural for the problem on hand to chose the restricted SURE estimates, imposing the null overall the estimated parameters. We also use SURE estimates constraining only the coefficients that are explicitly restricted in the null. The $p$-values thus generated we denote respectively MC1 and MC2. The findings are summarized in Tables 9-17.

First of all, it is noteworthy that, in accordance with the findings in Section 6, the empirical rejections of asymptotic tests by far exceed the 5% nominal levels, and in many cases are as high as 60%. Note that asymptotic bias is more serious for tests based on MLE. Correspondingly, we observe that the exact tests typically reverse the results of standard asymptotic tests in that they do not yield the strong rejections that asymptotic tests do.

---

[1] Africa (Ghana, Ivory Coast, Kenya, Malawi, Morocco, Zambia: 1977-88), South America (Mexico, Argentina, Chili, Colombia, Ecuador, Paraguay: 1973-88) and Asia (India, Indonesia, Korea, Pakistan, Thailand: 1978-87).

Further, in many instances where asymptotic tests show very strong significance, the $p$-values calculated on the basis of the Monte Carlo tests are greater than $\alpha$, so that the exact tests do not reject.[2] This issue proves to be particularly important when spurious significance supports coefficients whose signs contradict macroeconomic expectations. Indeed, this is a situation in which asymptotic arguments are very misleading, even if the asymptotic $p$-values are substantially less than $\alpha$. Second, we find that asymptotic $t$-tests based on two-step estimators do not lead to the same decisions as those based on MLE estimates. In contrast, no conflict arises if significance is assessed using the MC method. Indeed, this is exactly what the simulation experiment would lead one to expect.

*7.1 Parameter significance[3]*

| GDP | INFLT | | SRPLS | | TTRD | | EXPM | |
|---|---|---|---|---|---|---|---|---|
| | GLS | MLE | GLS | MLE | GLS | MLE | GLS | MLE |
| AFRICA | .025 | .021 | .185* | .113 | -.018 | .021 | -.004 | .004 |
| S-AMER. | **-.059*** | **-.085*** | .253* | .274* | **.107*** | **.101*** | -.051* | -.034* |
| ASIA | -.205* | -.445* | .646 | .694* | .126* | .074* | .038 | .074 |

Asymptotic 5% significance is indicated by the subscript "*"; the bold characters imply the coefficients are significant at 5% using MC tests

The result of the asymptotic tests associated with GLS estimates is that in the South American countries, all macroeconomic factors decisively affect GDP growth: inflation and exchange premiums adversely affect growth while higher budget surplus are associated with higher growth, adverse changes in terms of trade reduce GDP growth. The Panel regression estimates lead to the same conclusions, but the numerical value of all SURE estimates except

---

[2] The Monte Carlo tests reported may be interpreted as liberal exact tests, in the sense that failure to reject on the part of any of the criteria is compelling.

[3] In the following discussion, unless differently stated, the significance level is set a 5%.

the budget surplus coefficient are higher. Relying on Monte Carlo tests, we find that the budget surplus and the exchange premium lose their significance. In the African countries, all SURE estimates are numerically lower than the panel regression estimates; the coefficients on the budget surplus and exchange premium have the expected sign and the asymptotic $t$-statistic is significant only in the case of the budget surplus variable. However the exact tests indicate that none of the coefficients are significantly different from zero. In the asian countries, all coefficients except the coefficient on inflation are numerically higher than Fischer's panel regressions; all coefficients except the coefficient on the black market exchange premium have the anticipated sign. However, the latter coefficient and the budget surplus coefficient are not significantly different from zero. Using exact test, we find that all coefficients are not significantly different from zero.

We now turn to MLE-based tests for the GDP growth regressions. We first observe that for the South-American region, iterating the SURE estimates has no effect on coefficient signs nor significance; the coefficients on inflation and the budget surplus are numerically larger. Regarding Africa, we find that maximum likelihood estimation reverses the signs of coefficients on the terms of trade and the exchange premium; all estimates are however not significantly different from zero. Iteration preserves the signs in the regressions on Asian countries; the coefficients are numerically larger with the exception of the terms of trade factor. The asymptotic $t$-test is significant for all parameter estimates except the coefficient on the exchange premium; yet the Monte Carlo $p$-values are larger than 5%. It is of interest to note that the alternative asymptotic tests lead to conflicting decisions regarding the significance of the coefficient on the budget surplus in the context of Africa. In contrast, the Monte Carlo tests produce similar results whether the underlying estimates are the two-step GLS or are iterated to convergence. In summary, it appears that only inflation and the terms of trade affect GDP growth, and the effect is only noted in the South American countries. Next, we examine the capital accumulation regressions.

| Capital | INFLT | | SRPLS | | TTRD | | EXPM | |
|---------|-------|-----|-------|-----|------|-----|------|-----|
| | GLS | MLE | GLS | MLE | GLS | MLE | GLS | MLE |
| AFRICA | .004 | .039* | -.034 | .119* | -.018 | -.016* | -.074* | -.014* |
| S-AMER. | **-.041*** | **-.058*** | -.060* | -.092* | .014* | .014* | **-.052*** | **-.071*** |
| ASIA | .094* | -.072* | .091 | .149* | .017 | .004 | .007 | -.004 |

Asymptotic 5% significance is indicated by the subscript "*"; the bold characters imply the coefficients are significant at 5% using MC tests

In the South American countries, asymptotic tests based on GLS estimates imply that, in accordance with the results relating to panel regressions, inflation and black market exchange premiums decisively reduce the growth rate of capital; the numerical value of all coefficients except the coefficient on terms of trade are larger in the SURE regressions. The coefficient on the budget surplus variable is negative, contrary to what is typically presumed. However, while appearing strongly significant using standard $t$-tests, the latter coefficient and the coefficient on the terms of trade variable are not significant, based on Monte Carlo tests. In the African countries, the SURE estimates are numerically larger than the panel regression estimates, except for the case of inflation; all estimates except the exchange premium coefficient do not have the expected sign. The asymptotic $t$-statistic is significant only in the case of the exchange premium variable; however the exact tests indicate that none of the coefficients are significantly different from zero. In the Asian countries, all coefficients are numerically larger than the Panel estimates. The asymptotic $p$-values suggest that inflation significantly affects capital accumulation; yet the coefficient on inflation is (surprisingly) positive. The latter coefficient loses its significance in the exact test.

We see again that maximum likelihood produces the same results in terms of parameter signs and significance in the South-American context; all coefficients except the coefficient on the terms of trade are numerically larger. However, for the African region, we observe that the maximum likelihood and the two-step estimators of the coefficient on the budget surplus have opposite signs; although the sign reversal agrees with macro-economic

expectations, the associated Monte Carlo $t$-statistic is not significant. We also find that all MLE-based asymptotic tests are significant. Hence, the usual asymptotic arguments would imply that inflation and adverse changes in the terms of trade are favourable to the growth of capital. Interpreting the Monte Carlo tests, the latter effects are found to be insignificant. With respect to Asia, we see that MLE have the anticipated signs, which was not the case given GLS estimation. The coefficients on inflation and the budget surplus are asymptotically different from zero but the Monte Carlo tests are not significant. As was noted earlier, the alternative asymptotic tests associated with either GLS or MLE lead to conflicting decisions regarding the significance of the coefficient on the budget surplus in the context of Asia and all parameter estimates except the coefficient on the exchange premium with respect to Africa. In contrast, all Monte Carlo tests yield the same inference.

To recapitulate, Monte Carlo tests imply that only inflation and the exchange premium affect the growth of capital, and the effect is only observed in the South American countries. Next, we analyze the productivity growth SURE equations.

| Productivity | INFLT | | SRPLS | | TTRD | | EXPM | |
|---|---|---|---|---|---|---|---|---|
| | GLS | MLE | GLS | MLE | GLS | MLE | GLS | MLE |
| AFRICA | .025 | .080* | .286* | .092 | -.024 | .056* | -.004 | -.014* |
| S-AMER. | -.036* | -.052* | .317* | .357* | **.103*** | **.093*** | -.031 | -.009* |
| ASIA | -.263* | -.845* | .537* | -.010 | .103* | -.058* | .003 | .268 |

Asymptotic 5% significance is indicated by the subscript "*"; the bold characters imply the coefficients are significant at 5% using MC tests

In the South American case, asymptotic $t$-tests lead to the same conclusions as in Fischer (1993): the coefficient on the black market exchange premium is not significantly different from zero, inflation is significantly negatively correlated with the rate of productivity growth, increases in the budget surplus and favourable changes in terms of trade are associated with higher productivity growth; the SURE estimates are numerically larger

than the panel estimates. On the basis of the Monte Carlo tests, we find that the coefficient on the budget surplus loses its strong significance and the coefficient on inflation is no longer significant. In the Asian countries, all SURE estimates except the coefficient of the exchange premium are numerically higher than the panel estimates and the signs conform to expectations. All macroeconomic factors except the exchange premium variable appear significant, interpreting asymptotic $p$-values. The exact tests indicate that theses coefficients are not significantly different from zero. Turning to the African countries, we observe that all SURE estimates except the budget surplus coefficient are numerically lower than the panel estimates and not significant; but the coefficients on the budget surplus loses its significance in the exact tests. Again, it appears that iteration has no effect on parameter signs and significance in the South-American context. The coefficients on inflation and the budget deficit are numerically larger than the two-step estimates. The MLE of the coefficient of the terms of trade in the context of Africa has the expected sign; however, iteration to convergence produces negative estimates for the coefficients on the budget surplus and the terms of trade for the Asian region. Relying on asymptotic GLS or MLE-based tests produces conflicting inference regarding all the coefficient in the African case and all except the coefficient on the terms of trade for the Asian countries. The Monte Carlo MLE-based significance tests fail to reject except for the coefficient on the terms of trade in the context of South-America, as was the case with the GLS-based tests.

In summary, it appears that only the terms of trade affect productivity growth, and the effect is only noted in the South American countries.

*7.2 Testing Equality of regression coefficients*

The asymptotic tests strongly suggest that unconstrained SURE regressions are mostly called for: the Wald tests are significant for all the coefficients in all growth equation relating to Africa; the same is true of South America, with the exception of the coefficients on the

surplus variable in the GDP and productivity growth equations, and the coefficient on the terms of trade in the capital accumulation equation. Relatively fewer rejections are noted in the Asian countries: the test is significant for the terms of trade variable in the GDP growth equation, for the inflation and the exchange premium variable in the capital accumulation regression and for all the coefficients except the coefficient on the budget surplus variable in the productivity growth equation.

Turning to the Monte Carlo tests, we observe that in some cases, there is conflict among the alternative tests. For instance, consider, in the context of the GDP and capital growth regressions the hypothesis that inflation coefficients are equal across African countries; the Wald statistic rejects the null (the $p$-value is 4%) while the quasi-LR accepts. If we instead test the hypothesis that the coefficients on the budget surplus variable are equal across the African countries in the context of the productivity growth regressions, we find that the quasi-LR criterion suggests rejection at 4% but the Wald statistic is not significant. The same occurs when testing the hypothesis that the coefficients of the inflation, budget surplus and terms of trade variables are equal across the capital accumulation equations for Asia. It is natural, in cases where conflict among criteria arises, to assess the comparative power of the tests involved; in our case however, the simulation experiments suggest that no one test has a definitive power advantage. For the purpose of resolving the conflict, we reject the null at 5% if at least one of the two alternative tests is significant at 2.5%. Upon applying the latter rule, we find statistical evidence to reject the null in the context of the capital accumulation regressions accumulation: the test is significant for the coefficients of inflation in all regions and the coefficients on the budget surplus and the terms of trade in the Asian countries.

To conclude, we estimate the capital growth regressions relaxing the hypothesis of equality of coefficients when required. In the following, we briefly summarize the implications of Monte Carlo tests on parameter significance. The results are reported in Table

17. Within the African countries, the coefficient on inflation is significantly different from zero in Cote d'Ivoire, Malawi and Morocco. The coefficient has the anticipated sign only in Malawi. In the South American regressions, the coefficient on inflation is significant for Mexico, argentina and Ecuador and has the expected sign. As for Asia, the coefficient on inflation is significant in the case of Korea but is positive.

## 8. Conclusion

In this paper we have shown that the LR test on the coefficients of the MLR model is boundedly pivotal under the null. The bounds we have derived under general, possibly non-linear hypotheses are finite sample exact and may easily be obtained by simulation. In view of this, we have combined the approach of bounds test and Monte Carlo tests to provide $p$-values for tests statistics that are more accurate than those based on asymptotic approximations. The basic results were stated in terms of arbitrary hypotheses in MLR contexts. We have also focused on special cases, namely uniform and general linear hypotheses and have extended the methodology to the SURE framework. We have reported the results of an extensive Monte Carlo experiment that covered uniform linear, cross-equation and non-linear restrictions in MLR and SURE models. The feasibility of the test strategy was also illustrated with an empirical application. We have found that standard asymptotic tests exhibit serious errors in level, particularly in larger systems; usual size correction techniques (*e.g.* the Bartlett adjustment) may fail. in contrast, the various tests we have proposed displayed excellent size and power properties.

Table 1a. Coefficients for the MLR simulation experiment, the uniform linear hypothesis

| 5 EQ | $\beta_1$ | $(-2,.1,.9,-.1,.3,1.8)'$ |
|------|-----------|--------------------------|
|      | $\beta_2$ | $(.8,.8,.1,.5,-.5,2.1)'$ |
|      | $\beta_3$ | $(-1.1,0,-.5,2,1,.5)'$ |
|      | $\beta_4$ | $(1.9,.7,.3,.3,1.8,-.1)'$ |
|      | $\beta_5$ | $(-.2,1.5,.9,-.2,.7,.1)'$ |
| 7 EQ | $\beta_1$ | $(-2,.1,.9,-.1,.3,1.8,.2,.4)'$ |
|      | $\beta_2$ | $(.8,.8,.1,.5,-.5,2.1,.1,.5)'$ |
|      | $\beta_3$ | $(-1.1,0,-.5,2,1,.5,.1,.5)'$ |
|      | $\beta_4$ | $(1.9,.7,.3,.3,1.8,-.1,0,.6)'$ |
|      | $\beta_5$ | $(-.2,1.5,.9,-.2,.7,.1,.2,.4)'$ |
|      | $\beta_6$ | $(-1.8,2,.1,.2,.3,.4,1.6,-1)'$ |
|      | $\beta_7$ | $(1,.1,-1.5,2,.1,1.5,-.1,1.5)'$ |
| 8 EQ | $\beta_1$ | $(-2,.1,.9,-.1,.3,0,1.8,.2,.4)'$ |
|      | $\beta_2$ | $(.8,.8,.1,.5,0,-.5,2.1,.1,.5)'$ |
|      | $\beta_3$ | $(-1.1,0,-.5,2,1,0,.5,.1,.5)'$ |
|      | $\beta_4$ | $(1.9,.7,0,.3,.3,1.8,-.1,0,.6)'$ |
|      | $\beta_5$ | $(-.2,1.5,.9,0,-.2,.7,.1,.2,.4)'$ |
|      | $\beta_6$ | $(-1.8,2,.1,.2,0,.3,.4,1.6,-1)'$ |
|      | $\beta_7$ | $(1,.1,-1.5,2,.1,1.5,0,-.1,1.5)'$ |
|      | $\beta_8$ | $(-.5,-2.5,1.5,2.5,-1.5,1.5,.6,1.5,0)'$ |

Table 1b. Empirical levels of various test criteria: the MLR model with uniform linear restrictions

| Sample Size | 5 EQ | | | 7 EQ | | | 8 EQ | | |
|-------------|------|-----------------|-----------|------|-----------------|-----------|------|-------------|-----------|
|             | $LR$ | $LR_{Bartlett}$ | $LR_{MC}$ | $LR$ | $LR_{Bartlett}$ | $LR_{MC}$ | $LR$ | $R_{Bartlett}$ | $LR_{MC}$ |
| 20  | .299 | .136 | .051 | .563 | .322 | .047 | .753 | .493 | .046 |
| 25  | .190 | .093 | .049 | .383 | .182 | .036 | .476 | .240 | .042 |
| 40  | .130 | .079 | .056 | .189 | .088 | .051 | .223 | .117 | .051 |
| 50  | .095 | .058 | .055 | .155 | .078 | .050 | .190 | .088 | .053 |
| 100 | .071 | .054 | .042 | .071 | .050 | .041 | .094 | .065 | .049 |

Note:LR, $LR_{Bartlett}$ and $LR_{MC}$ refer to the LR, the Bartlett-adjusted and the randomized LR criteria. respectively.

Table 2a. Coefficients for the MLR simulation experiment with cross-equation restrictions

| 5 EQ | $\beta_1$ | $(1.2,.1,0,0)'$ |
| | $\beta_2$ | $(.8,0,.1,0)'$ |
| | $\beta_3$ | $(-1.1,0,0,.1)'$ |
| 7 EQ | $\beta_1$ | $(1.2,.1,0,0,0,0)'$ |
| | $\beta_2$ | $(.8,0,.1,0,0,0)'$ |
| | $\beta_3$ | $(-1.1,0,0,.1,0,0)'$ |
| | $\beta_4$ | $(1.9,0,0,0,.1,0)'$ |
| | $\beta_5$ | $(-.2,0,0,0,0,.1)'$ |

Table 2b. Empirical levels of various LR-based tests of cross-equation restrictions in the MLR model

| | 3 EQ | 5 EQ |
|---|---|---|
| LR (Asymptotic cut-off point) | .122 | .310 |
| LR (Randomized cut-off point) | .055 | .044 |
| LR (Bound) | .036 | .029 |

Note: for definitions, refer to (4.2) and (4.7).

Table 2c. Powers of Monte Carlo tests of cross-equation restrictions in the MLR model with three equations.

| | 20 replications | | | | | 100 replications | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | $\beta_{11}=.3$ | $\beta_{11}=.5$ | $\beta_{11}=.7$ | $\beta_{11}=.9$ | $\beta_{11}=1$ | $\beta_{11}=.3$ | $\beta_{11}=.5$ | $\beta_{11}=.7$ | $\beta_{11}=.9$ | $\beta_{11}=1$ |
| $LR_{AST}$ | .140 | .522 | .918 | .995 | 1.00 | .140 | .522 | .918 | .995 | 1.00 |
| $LR_{MC}$ | .137 | .468 | .849 | .987 | .991 | .135 | .539 | .912 | .995 | 1.00 |
| $LR_{BND}$ | .095 | .404 | .799 | .963 | .987 | .099 | .441 | .861 | .986 | .999 |

Note: For definitions, see (4.2) and (4.7); $LR_{AST}$ refers to the infeasible locally corrected asymptotic test.

Table 2d. Powers of Monte Carlo tests of cross-equation restrictions in the MLR model with five equations.

| | 20 Replications | | | | | 100 Replications | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | $\beta_{11}=.3$ | $\beta_{11}=.5$ | $\beta_{11}=.7$ | $\beta_{11}=.9$ | $\beta_{11}=1$ | $\beta_{11}=.1$ | $\beta_{11}=.5$ | $\beta_{11}=.7$ | $\beta_{11}=.9$ | $\beta_{11}=1$ |
| $LR_{AST}$ | .128 | .515 | .904 | .995 | 1.00 | .128 | .515 | .904 | .995 | 1.0 |
| $LR_{MC}$ | .138 | .467 | .837 | .967 | 1.00 | .137 | .537 | .904 | .994 | 1.0 |
| $LR_{BOUND}$ | .120 | .427 | .792 | .958 | .995 | .110 | .484 | .877 | .990 | 1.0 |

Note: For definitions, see (4.2) and (4.7); $LR_{AST}$ refers to the infeasible locally corrected asymptotic test.

Table 3. Empirical levels of various asymptotic criteria: the SURE model

|  | 3 EQ | 5 EQ |
|---|---|---|
| $Wald_{FGLS}$ (Chi-square) | .061 | .130 |
| $Wald_{FGLS}$ (F) | .052 | .121 |
| $Wald_{MLE}$ (Chi-square) | .124 | .254 |
| $Wald_{MLE}$ (F) | .111 | .242 |
| $LR$ | .094 | .143 |
| $QLR_{(0)}$ (0 Iterations) | .068 | .077 |
| $QLR_{(1)}$ (1 Iteration) | .088 | .131 |
| $QLR_{(2)}$ (2 Iterations) | .094 | .143 |

Note: For definitions, see (5.3), (5.4), (5.12) and (5.14).

Table 4a. Empirical levels of various Monte Carlo and bounds test: the SURE model

|  | 3 EQ | 5 EQ |
|---|---|---|
| $Wald_{FGLS}$ | .049 | .047 |
| $Wald_{MLE}$ | .047 | .049 |
| $LR_{MC}$ | .047 | .043 |
| $QLR_{(0)}$ (0 Iterations) | .045 | .052 |
| $QLR_{(1)}$ (1 Iteration) | .048 | .052 |
| $QLR_{(2)}$ (2 Iterations) | .047 | .044 |
| $LR_{BOUND}$ | .036 | .029 |
| $\mu_{31}$ | .058 | - |
| $\mu_{32}$ | .051 | - |
| $\mu_{33}$ | .055 | - |
| $\mu_{51}$ | - | .027 |
| $\mu_{52}$ | - | .026 |
| $\mu_{53}$ | - | .025 |
| $\mu_{54}$ | - | .011 |
| $\mu^{55}$ | - | .025 |

Note: For definitions, see (5.3), (5.4), (5.12), (5.14), (6.5) and (6.6).

Table 5. Powers of size corrected asymptotic tests for the SURE model

| | 3 EQ | | | | | 5EQ | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | $\beta_{11}=.3$ | $\beta_{11}=.5$ | $\beta_{11}=.7$ | $\beta_{11}=.9$ | $\beta_{11}=1$ | $\beta_{11}=.3$ | $\beta_{11}=.5$ | $\beta_{11}=.7$ | $\beta_{11}=.9$ | $\beta_{11}=1.1$ |
| $W_{FGLS}$ | .192 | .647 | .939 | .993 | .999 | .200 | .703 | .961 | .994 | .999 |
| $W_{MLE}$ | .264 | .787 | .984 | 1.00 | 1.00 | .317 | .918 | 1.0 | 1.0 | 1.0 |
| $LR_{ASY}$ | .281 | .806 | .985 | 1.00 | 1.00 | .331 | .913 | .999 | 1.0 | 1.0 |

Note: For definitions, see (5.3), (5.4), (5.12), (5.14), (6.5) and (6.6).

Table 6a. Powers of the bounds tests for the SURE model with three equations

| | 20 replications | | | | | 100 replications | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | $\beta_{11}=.3$ | $\beta_{11}=.5$ | $\beta_{11}=.7$ | $\beta_{11}=.9$ | $\beta_{11}=1$ | $\beta_{11}=.3$ | $\beta_{11}=.5$ | $\beta_{11}=.7$ | $\beta_{11}=.9$ | $\beta_{11}=1.0$ |
| $p_1$ | .065 | .383 | .791 | .963 | .987 | .077 | .434 | .858 | .986 | .999 |
| $p_2$ | .171 | .324 | .171 | .034 | .013 | .204 | .372 | .127 | .014 | .001 |
| $p_3$ | .030 | .021 | .008 | 0.00 | 0.00 | .022 | .007 | .003 | 0.00 | 0.00 |
| $p_4$ | .734 | .272 | .030 | .003 | 0.00 | .697 | .187 | .012 | 0.00 | 0.00 |

Note: $p_1$ refers to the empirical probability that the optimal and conservative tests reject, $p_4$ refers to the probability that the tests fail to reject, $p_2$ measures the probability that the optimal test rejects and the conservative test fails to reject and $p_3$ measures the probability that the conservative test rejects and the optimal test fails to reject. The null hypothesis corresponds to $\beta_{11}=.1$.

Table 6b. Powers of Monte Carlo tests for the SURE model with three equations

| | 20 replications | | | | | 100 replications | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | $\beta_{11}=.3$ | $\beta_{11}=.5$ | $\beta_{11}=.7$ | $\beta_{11}=.9$ | $\beta_{11}=1.0$ | $\beta_{11}=.3$ | $\beta_{11}=.5$ | $\beta_{11}=.7$ | $\beta_{11}=.9$ | $\beta_{11}=1.0$ |
| $W_{FGLS}$ | .185 | .579 | .884 | .974 | .986 | .202 | .640 | .934 | .990 | .998 |
| $W_{MLE}$ | .225 | .704 | .958 | .997 | 1.00 | .260 | .774 | .985 | 1.00 | 1.00 |
| $LR_{MC}$ | .236 | .707 | .962 | .997 | 1.00 | .262 | .779 | .985 | 1.00 | 1.00 |
| $QLR_{(0)}$ | .227 | .689 | .950 | .993 | .988 | .256 | .762 | .977 | .997 | .999 |
| $QLR_{(1)}$ | .238 | .709 | .961 | .997 | 1.00 | .259 | .776 | .986 | 1.00 | 1.00 |
| $QLR_{(2)}$ | .236 | .707 | .962 | .997 | 1.00 | .262 | .776 | .985 | 1.00 | 1.00 |
| $LR_{BOUND}$ | .095 | .404 | .799 | .963 | .987 | .099 | .441 | .861 | .986 | .999 |
| $\mu_{31}$ | .076 | .108 | .148 | .216 | .259 | .064 | .108 | .165 | .219 | .268 |
| $\mu_{32}$ | .197 | .552 | .869 | .974 | .992 | .210 | .641 | .935 | .995 | .998 |
| $\mu_{33}$ | .093 | .183 | .307 | .432 | .489 | .088 | .184 | .328 | .503 | .601 |

Note: For definitions, see (5.3), (5.4), (5.12), (5.14) and (6.5).

Table 7a. Powers of Monte Carlo tests for the SURE model with five equations

| | 20 replications | | | | | 100 replications | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | $\beta_{11}=.3$ | $\beta_{11}=.5$ | $\beta_{11}=.7$ | $\beta_{11}=.9$ | $\beta_{11}=1.1$ | $\beta_{11}=.3$ | $\beta_{11}=.5$ | $\beta_{11}=.7$ | $\beta_{11}=.9$ | $\beta_{11}=1.1$ |
| $W_{FGLS}$ | .162 | .619 | .918 | .982 | .998 | .186 | .684 | .946 | .990 | .999 |
| $W_{MLE}$ | .265 | .832 | .991 | .999 | 1.00 | .297 | .903 | 1.0 | 1.00 | 1.00 |
| $LR_{MC}$ | .286 | .841 | .999 | .999 | 1.00 | .328 | .908 | .998 | 1.00 | 1.00 |
| $QLR_{(0)}$ | .265 | .806 | .971 | .998 | 1.00 | .316 | .864 | .983 | .999 | 1.00 |
| $QLR_{(1)}$ | .290 | .849 | .988 | .998 | 1.00 | .334 | .900 | .997 | 1.00 | 1.00 |
| $QLR_{(2)}$ | .287 | .842 | .991 | .999 | 1.00 | .331 | .908 | .997 | 1.00 | 1.00 |
| $LR_{BOUND}$ | .120 | .427 | .792 | .958 | .995 | .110 | .484 | .877 | .990 | 1.00 |
| $\mu_{51}$ | .029 | .034 | .038 | .041 | .048 | .032 | .036 | .039 | .041 | .044 |
| $\mu_{52}$ | .031 | .036 | .039 | .042 | .045 | .031 | .034 | .038 | .040 | .041 |
| $\mu_{53}$ | .042 | .085 | .154 | .258 | .359 | .035 | .077 | .152 | .241 | .397 |
| $\mu_{54}$ | .023 | .071 | .159 | .289 | .456 | .025 | .067 | .175 | .302 | .512 |
| $\mu_{55}$ | .031 | .050 | .071 | .118 | .170 | .033 | .056 | .092 | .128 | .180 |

Note: For definitions, see (5.3), (5.4), (5.12), (5.14) and (6.6).

Table 7.b Powers of bounds tests for the SURE model with five equations

| | 20 replications | | | | | 100 replications | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | $\beta_{11}=.3$ | $\beta_{11}=.5$ | $\beta_{11}=.7$ | $\beta_{11}=.9$ | $\beta_{11}=1.1$ | $\beta_{11}=.3$ | $\beta_{11}=.5$ | $\beta_{11}=.7$ | $\beta_{11}=.9$ | $\beta_{11}=1.1$ |
| $p_1$ | .082 | .416 | .792 | .958 | .995 | .075 | .474 | .877 | .990 | 1.00 |
| $p_2$ | .249 | .497 | .207 | .042 | .005 | .256 | .439 | .122 | .010 | 0.00 |
| $p_3$ | .038 | .011 | 0.00 | 0.00 | 0.00 | .035 | .010 | 0.00 | 0.00 | 0.00 |
| $p_4$ | .631 | .076 | .001 | 0.00 | 0.00 | .634 | .077 | .001 | 0.00 | 0.00 |

Note: for definitions, refer to Table 6.a.

Table 8.a: MLR model, NON-LINEAR hypothesis: Asymptotic, Monte Carlo and bounds tests

| | |
|---|---|
| *LR* (Asymptotic cut-off point) | .857 |
| *LR* (Randomized cut-off point) | **.037** |
| *LR* (Bound) | **.047** |

Table 8.b: SURE model, NON-LINEAR hypothesis: Asymptotic and Monte Carlo tests

| | Quasi-LR | Wald |
|---|---|---|
| Asymptotic cut-off point | .320 | .246 |
| Monte Carlo cut-off point | .024 | .077 |

Table 8.c: MLR model, Rao's asymptotic F test

| p | k | r | c | $LR_{asy}$ | $F_{asy}$ | $LR_{MC}$ |
|---|---|---|---|---|---|---|
| 13 | 12 | 12 | 13 | 1.00 | .108 | .052 |
| 13 | 12 | 12 | 13 | 1.00 | .198 | .047 |
| 11 | 12 | 12 | 11 | 1.00 | .096 | .054 |
| 12 | 12 | 12 | 12 | 1.00 | .114 | .048 |
| 12 | 13 | 13 | 12 | 1.00 | .225 | .038 |

Note: for definitions, refer to (3.2) and (3.8).

Table 9.a: GDP GROWTH
(GHANA, IVORY COAST, KENYA, MALAWI, MOROCCO, ZAMBIA, 1977-88)

| Country | variable | Coef. | t-ratio | p-Asy | p-MC1 | %R-MC1 | P-MC2 | %R-MC2 |
|---------|----------|-------|---------|-------|-------|--------|-------|--------|
| Ghana | INFLT | -.036 | -1.153 | .2867 | .51 | .23 | .52 | .27 |
| | SRPLS | -.046 | -1.399 | .2045 | .43 | .30 | .39 | .28 |
| | TTRD | -.018 | -.464 | .6566 | .77 | .24 | .76 | .23 |
| | EXPM | -.032 | -2.417 | .0463 | .20 | .28 | .19 | .29 |
| Cote D'Ivoire | INFLT | .388 | 3.093 | .0175 | .08 | .19 | .08 | .20 |
| | SRPLS | .626 | 3.936 | .0056 | .06 | .25 | .07 | .28 |
| | TTRD | -.331 | -4.847 | .0019 | .03 | .33 | .02 | .35 |
| | EXPM | .453 | 2.942 | .0217 | .10 | .18 | .11 | .23 |
| Kenya | INFLT | -.211 | -1.504 | .1762 | .32 | .23 | .39 | .26 |
| | SRPLS | .682 | 1.950 | .0922 | .22 | .21 | .22 | .21 |
| | TTRD | -.020 | -.425 | .6836 | .75 | .25 | .74 | .24 |
| | EXPM | -.057 | -.909 | .3938 | .48 | .20 | .50 | .23 |
| Malawi | INFLT | -.402 | -6.728 | .0001 | .01 | .18 | .01 | .29 |
| | SRPLS | .669 | 8.896 | .0003 | .01 | .20 | .01 | .29 |
| | TTRD | -.043 | -1.270 | .2446 | .51 | .34 | .54 | .33 |
| | EXPM | -.004 | -.175 | .8660 | .89 | .27 | .88 | .27 |
| Morocco | INFLT | .623 | 1.690 | .1348 | .44 | .38 | .47 | .38 |
| | SRPLS | .941 | 2.290 | .0558 | .26 | .33 | .23 | .33 |
| | TTRD | -.169 | -.718 | .4959 | .65 | .25 | .67 | .25 |
| | EXPM | .099 | .512 | .6246 | .83 | .25 | .82 | .25 |
| Zambia | INFLT | .248 | -1.727 | .2647 | .44 | .24 | .42 | .24 |
| | SRPLS | -.100 | 1.212 | .7792 | .88 | .21 | .91 | .21 |
| | TTRD | -.099 | -1.272 | .2440 | .45 | .30 | .48 | .30 |
| | EXPM | -.002 | -.062 | .9522 | .98 | .28 | .94 | .28 |
| All | INFLT | .025 | .893 | .372 | .70 | .32 | .70 | .32 |
| | SRPLS | .185 | 2.01 | .044 | .23 | .23 | .24 | .24 |
| | TTRD | -.018 | -.66 | .508 | .76 | .31 | .76 | .31 |
| | EXPM | -.004 | -.61 | .55 | .77 | .35 | .75 | .36 |

Table 9.b: CAPITAL GROWTH
(GHANA, IVORY COAST, KENYA, MALAWI, MOROCCO, ZAMBIA, 1977-88)

| Country | variable | Coef. | t-ratio | p-ASY | P-MC1 | %R-MC1 | P-MC2 | %R-MC2 |
|---------|----------|-------|---------|-------|-------|--------|-------|--------|
| Ghana | INFLT | -.005 | -.236 | .8204 | .90 | .24 | .89 | .23 |
| | SRPLS | .179 | 1.203 | .2681 | .48 | .28 | .46 | .30 |
| | TTRD | .030 | 1.492 | .1795 | .34 | .28 | .36 | .27 |
| | EXPM | -.001 | -.154 | .8820 | .95 | .31 | .96 | .31 |
| Cote D'Ivoire | INFLT | 1.095 | 13.108 | .0000 | **.01** | .16 | **.01** | .24 |
| | SRPLS | .202 | 1.953 | .0918 | .29 | .28 | .29 | .28 |
| | TTRD | -.163 | -3.612 | .0086 | .08 | .31 | .09 | .31 |
| | EXPM | .416 | 3.952 | .0055 | **.02** | .24 | **.02** | .20 |
| Kenya | INFLT | .141 | 1.066 | .3220 | .55 | .24 | .53 | .26 |
| | SRPLS | -.424 | -1.194 | .2715 | .48 | .21 | .52 | .22 |
| | TTRD | -.023 | -.472 | .6510 | .72 | .25 | .70 | .25 |
| | EXPM | -.084 | -1.168 | .2812 | .49 | .21 | .47 | .21 |
| Malawi | INFLT | -.191 | -2.159 | .0677 | .19 | .20 | .24 | .26 |
| | SRPLS | .339 | 3.328 | .0126 | **.05** | .23 | .07 | .31 |
| | TTRD | -.088 | -2.039 | .0809 | .26 | .27 | .28 | .30 |
| | EXPM | .145 | 4.737 | .0021 | **.03** | .24 | .01 | .27 |
| Morocco | INFLT | .286 | 1.748 | .1240 | .34 | .33 | .46 | .37 |
| | SRPLS | -.750 | -4.096 | .0046 | **.05** | .34 | **.05** | .32 |
| | TTRD | .092 | .923 | .3865 | .59 | .26 | .58 | .27 |
| | EXPM | -.218 | -2.508 | .0405 | .18 | .28 | .20 | .32 |
| Zambia | INFLT | -.027 | -3.319 | .0128 | .11 | .23 | .11 | .24 |
| | SRPLS | -.078 | -1.467 | .1859 | .33 | .21 | .45 | .28 |
| | TTRD | -.067 | -2.216 | .0623 | .13 | .15 | .20 | .27 |
| | EXPM | .008 | 1.719 | .1293 | .32 | .27 | .32 | .28 |
| All | INFLT | .004 | .248 | .804 | .90 | .23 | .91 | .23 |
| | SRPLS | -.034 | -.792 | .43 | .69 | .37 | .73 | .33 |
| | TTRD | -.018 | -1.87 | .061 | .42 | .41 | .43 | .38 |
| | EXPM | -.074 | -2.2 | .026 | .21 | .25 | .24 | .26 |

Table 9.c: PRODUCTIVITY GROWTH
(GHANA, IVORY COAST, KENYA, MALAWI, MOROCCO, ZAMBIA, 1977-88)

| Country | variable | Coef. | t-ratio | p-ASY | p-MC1 | %R-MC1 | p-MC2 | %R-MC2 |
|---|---|---|---|---|---|---|---|---|
| Ghana | INFLT | -.054 | -1.602 | .1531 | .36 | .26 | .37 | .29 |
| | SRPLS | -.958 | -2.920 | .0223 | .16 | .28 | .16 | .25 |
| | TTRD | -.028 | -.722 | .4939 | .67 | .19 | .67 | .21 |
| | EXPM | -.041 | -3.109 | .0171 | .12 | .27 | .14 | .30 |
| Cote D'Ivoire | INFLT | -.039 | -.347 | .7385 | .79 | .18 | .77 | .18 |
| | SRPLS | .555 | 3.926 | .0057 | .07 | .25 | .06 | .30 |
| | TTRD | -.262 | -4.310 | .0035 | .05 | .37 | .04 | .34 |
| | EXPM | .247 | 1.748 | .1176 | .22 | .18 | .28 | .23 |
| Kenya | INFLT | -.187 | -1.798 | .1152 | .26 | .24 | .30 | .27 |
| | SRPLS | .659 | 2.491 | .0416 | .09 | .23 | .13 | .21 |
| | TTRD | .008 | .209 | .8404 | .85 | .25 | .85 | .25 |
| | EXPM | -.084 | -1.696 | .1337 | .29 | .21 | .28 | .20 |
| Malawi | INFLT | -.316 | -4.236 | .0039 | .02 | .21 | .03 | .33 |
| | SRPLS | .523 | 5.356 | .0011 | .02 | .25 | .02 | .30 |
| | TTRD | .001 | .019 | .9856 | .99 | .30 | 1.0 | .25 |
| | EXPM | -.092 | -3.109 | .0171 | .10 | .19 | .10 | .26 |
| Morocco | INFLT | .577 | 1.904 | .0986 | .29 | .28 | .35 | .31 |
| | SRPLS | 1.357 | 3.997 | .00052 | .03 | .33 | .03 | .30 |
| | TTRD | -.198 | -1.038 | .3340 | .56 | .25 | .52 | .28 |
| | EXPM | .168 | 1.064 | .3225 | .60 | .27 | .60 | .27 |
| Zambia | INFLT | .303 | 1.397 | .2051 | .39 | .28 | .39 | .29 |
| | SRPLS | .024 | -.064 | .9505 | .97 | .30 | .99 | .30 |
| | TTRD | -.138 | -1.697 | .1335 | .35 | .29 | .32 | .25 |
| | EXPM | -.00004 | -.001 | .9992 | 1.0 | .32 | 1.0 | .31 |
| All | INFLT | .025 | .713 | .476 | .79 | .30 | .79 | .30 |
| | SRPLS | .286 | 3.14 | .0017 | .08 | .22 | .07 | .28 |
| | TTRD | -.024 | -.91 | .363 | .72 | .30 | .71 | .29 |
| | EXPM | -.004 | -.39 | .693 | .78 | .28 | .81 | .27 |

Table 10.a: GDP GROWTH
(MEXICO, ARGENTINA, CHILI, COLOMBIA, ECUADOR, PARAGUAY, 1973-88)

| Country | variable | Coef. | t-ratio | p-ASY | p-MC1 | %R-MC1 | p-MC2 | %R-MC2 |
|---|---|---|---|---|---|---|---|---|
| Mexico | INFLT | -.075 | -.1745 | .1117 | .31 | .23 | .40 | .30 |
| | SRPLS | .331 | 2.019 | .0711 | .17 | .16 | .17 | .18 |
| | TTRD | .209 | 4.948 | .0006 | .03 | .19 | .03 | .20 |
| | EXPM | -.153 | -1.403 | .1909 | .43 | .24 | .43 | .28 |
| Argentina | INFLT | -.011 | -.539 | .6014 | .70 | .23 | .70 | .22 |
| | SRPLS | .334 | 1.422 | .1855 | .30 | .15 | .32 | .14 |
| | TTRD | .109 | 1.391 | .1945 | .41 | .24 | .43 | .24 |
| | EXPM | .018 | .849 | .4158 | .61 | .26 | .64 | .26 |
| Chili | INFLT | -.043 | -1.807 | .1009 | .33 | .29 | .34 | .28 |
| | SRPLS | .386 | 2.072 | .0650 | .28 | .30 | .29 | .29 |
| | TTRD | .244 | 4.375 | .0014 | .04 | .25 | .05 | .20 |
| | EXPM | -.631 | -3.907 | .0029 | .05 | .29 | .05 | .26 |
| Colombia | INFLT | .257 | -1.429 | .1835 | .44 | .32 | .46 | .34 |
| | SRPLS | .164 | 1 | .3411 | .52 | .18 | .52 | .19 |
| | TTRD | .020 | .537 | .6028 | .71 | .29 | .73 | .28 |
| | EXPM | -.191 | -1.987 | .0750 | .25 | .26 | .23 | .23 |
| Ecuador | INFLT | -.427 | -3.730 | .0039 | .03 | .22 | .05 | .24 |
| | SRPLS | .532 | 2.483 | .0324 | .15 | .25 | .17 | .26 |
| | TTRD | .082 | 1.927 | .0828 | .23 | .20 | .25 | .24 |
| | EXPM | -.071 | -1.857 | .0930 | .22 | .18 | .21 | .19 |
| Paraguay | INFLT | .137 | 1.057 | .3155 | .45 | .20 | .47 | .22 |
| | SRPLS | .797 | 1.716 | .1170 | .21 | .13 | .20 | .16 |
| | TTRD | .123 | 3.927 | .0028 | .03 | .19 | .03 | .19 |
| | EXPM | -.121 | -3.501 | .0057 | .05 | .17 | .03 | .19 |
| All | INFLT | -.059 | -4.83 | .000 | .01 | .24 | .02 | .22 |
| | SRPLS | .253 | 2.85 | .004 | .13 | .32 | .14 | .34 |
| | TTRD | .107 | 5.53 | .000 | .01 | .24 | .01 | .25 |
| | EXPM | -.051 | -2.9 | .003 | .12 | .25 | .10 | .23 |

Table 10.b: CAPITAL GROWTH
(MEXICO, ARGENTINA, CHILI, COLOMBIA, ECUADOR, PARAGUAY, 1973-88)

| Country | variable | Coef. | t-ratio | p-ASY | p-MC1 | %R-MC1 | p-MC2 | %R-MC2 |
|---|---|---|---|---|---|---|---|---|
| Mexico | INFLT | -.111 | -6.630 | .0001 | **.01** | .20 | **.01** | .23 |
| | SRPLS | -.212 | -3.177 | .0099 | .06 | .17 | .06 | .20 |
| | TTRD | .021 | 1.275 | .2311 | .34 | .19 | .37 | .20 |
| | EXPM | -.034 | -.799 | .4428 | .61 | .16 | .61 | .17 |
| Argentina | INFLT | -.031 | -2.472 | .0330 | .17 | .23 | .15 | .24 |
| | SRPLS | .200 | 1.378 | .1982 | .26 | .10 | .33 | .13 |
| | TTRD | .076 | 1.591 | .1427 | .32 | .24 | .32 | .23 |
| | EXPM | -.010 | -.737 | .4778 | .59 | .17 | .61 | .17 |
| Chili | INFLT | -.010 | -1.276 | .2309 | .41 | .20 | .44 | .22 |
| | SRPLS | .019 | .317 | .7576 | .83 | .21 | .82 | .20 |
| | TTRD | -.024 | -1.264 | .2350 | .45 | .27 | .43 | .23 |
| | EXPM | -.087 | -1.454 | .1766 | .34 | .21 | .38 | .20 |
| Colombia | INFLT | -.089 | -2.338 | .0415 | .22 | .27 | .29 | .30 |
| | SRPLS | -.124 | -3.363 | .0072 | **.02** | .14 | **.02** | .15 |
| | TTRD | .013 | 1.676 | .1247 | .22 | .19 | .22 | .19 |
| | EXPM | -.072 | -3.337 | .0075 | **.03** | .10 | **.05** | .11 |
| Ecuador | INFLT | -.214 | -4.132 | .0020 | **.02** | .20 | **.05** | .24 |
| | SRPLS | -.051 | -.524 | .6120 | .80 | .18 | .80 | .18 |
| | TTRD | .029 | 1.557 | .1505 | .27 | .18 | .29 | .19 |
| | EXPM | -.040 | -2.229 | .0499 | .14 | .19 | .15 | .19 |
| Paraguay | INFLT | -.330 | -4.305 | .0016 | **.02** | .23 | **.01** | .24 |
| | SRPLS | .932 | 3.408 | .0067 | **.03** | .17 | **.04** | .17 |
| | TTRD | .006 | .324 | .7528 | .87 | .17 | .88 | .20 |
| | EXPM | -.091 | -3.918 | .0029 | **.03** | .20 | **.05** | .22 |
| All | INFLT | -.041 | -10.4 | .000 | **.01** | .22 | **.01** | .21 |
| | SRPLS | -.060 | -2.6 | .010 | .15 | .32 | .16 | .34 |
| | TTRD | .014 | 3.08 | .002 | .17 | .28 | .17 | .29 |
| | EXPM | -.052 | -9.54 | .000 | **.01** | .27 | **.01** | .22 |

Table 10.c: PRODUCTIVITY GROWTH
(MEXICO, ARGENTINA, CHILI, COLOMBIA, ECUADOR, PARAGUAY, 1973-88)

| Country | variable | Coef. | t-ratio | p-ASY | p-MC1 | %R-MC1 | p-MC2 | %R-MC2 |
|---|---|---|---|---|---|---|---|---|
| Mexico | INFLT | -.037 | -.809 | .4374 | .62 | .25 | .63 | .31 |
| | SRPLS | .387 | 2.224 | .0503 | .12 | .17 | .13 | .15 |
| | TTRD | .194 | 4.237 | .0017 | .03 | .21 | .03 | .22 |
| | EXPM | -.082 | -.706 | .4961 | .68 | .24 | .68 | .27 |
| Argentina | INFLT | -.002 | -.099 | .9228 | .96 | .29 | .96 | .29 |
| | SRPLS | .325 | 1.297 | .2238 | .38 | .17 | .38 | .21 |
| | TTRD | .059 | .063 | .5038 | .65 | .26 | .64 | .24 |
| | EXPM | .013 | .589 | .5689 | .68 | .24 | .68 | .24 |
| Chili | INFLT | -.040 | -1.631 | .1340 | .37 | .28 | .36 | .30 |
| | SRPLS | .346 | 1.812 | .1000 | .34 | .32 | .34 | .28 |
| | TTRD | .251 | 4.401 | .0013 | .04 | .21 | .05 | .24 |
| | EXPM | -.623 | -3.799 | .0035 | .06 | .27 | .05 | .25 |
| Colombia | INFLT | -.224 | -1.227 | .2480 | .47 | .33 | .48 | .31 |
| | SRPLS | .252 | 1.520 | .1594 | .35 | .16 | .30 | .22 |
| | TTRD | .015 | .408 | .6922 | .81 | .32 | .82 | .32 |
| | EXPM | -.143 | -1.475 | .1709 | .40 | .24 | .41 | .23 |
| Ecuador | INFLT | -.342 | -3.093 | .0114 | .10 | .24 | .09 | .25 |
| | SRPLS | .604 | 2.898 | .0159 | .11 | .25 | .12 | .25 |
| | TTRD | .063 | 1.516 | .1605 | .32 | .21 | .34 | .22 |
| | EXPM | -.057 | -1.541 | .1544 | .27 | .18 | .31 | .16 |
| Paraguay | INFLT | .263 | 2.219 | .0508 | .16 | .18 | .18 | .22 |
| | SRPLS | .291 | .692 | .5050 | .55 | .16 | .55 | .15 |
| | TTRD | .116 | 4.040 | .0024 | .01 | .21 | .02 | .18 |
| | EXPM | -.080 | -2.560 | .0284 | .11 | .19 | .13 | .19 |
| All | INFLT | -.036 | -3.1 | .002 | .12 | .23 | .09 | .24 |
| | SRPLS | .317 | 3.8 | .0001 | .07 | .33 | .06 | .35 |
| | TTRD | .103 | 5.67 | .000 | .01 | .22 | .01 | .24 |
| | EXPM | -.031 | -1.8 | .061 | .28 | .26 | .27 | .25 |

Table 11.a: GDP GROWTH
(INDIA, INDONESIA, KOREA, PAKISTAN, THAILAND, 1978-87)

| Country | variable | Coef. | t-ratio | p-ASY | p-MC1 | %R-MC1 | p-MC2 | %R-MC2 |
|---|---|---|---|---|---|---|---|---|
| India | INFLT | .103 | .323 | .7595 | .86 | .23 | .86 | .23 |
| | SRPLS | -.377 | -.602 | .5733 | .70 | .31 | .69 | .33 |
| | TTRD | .077 | .817 | .4511 | .68 | .38 | .67 | .41 |
| | EXPM | .012 | .062 | .9526 | .97 | .37 | .97 | .37 |
| Idonesia | INFLT | -1.276 | -.2038 | .0971 | .31 | .33 | .32 | .37 |
| | SRPLS | .577 | .597 | .5766 | .76 | .39 | .76 | .41 |
| | TTRD | .308 | 1.956 | .1079 | .33 | .34 | .36 | .37 |
| | EXPM | -.373 | -1.163 | .2974 | .59 | .36 | .58 | .36 |
| Korea | INFLT | -.004 | -.025 | .9809 | .99 | .28 | .99 | .28 |
| | SRPLS | .670 | 1.254 | .1880 | .39 | .30 | .49 | .35 |
| | TTRD | .837 | 5.100 | .0038 | .04 | .31 | .05 | .35 |
| | EXPM | .038 | .508 | .6330 | .80 | .32 | .85 | .32 |
| Pakistan | INFLT | -.173 | -.284 | .7877 | .92 | .30 | .90 | .28 |
| | SRPLS | 1.231 | 1.185 | .2894 | .50 | .26 | .49 | .29 |
| | TTRD | .016 | .192 | .8554 | .94 | .21 | .91 | .23 |
| | EXPM | -.075 | -.868 | .4252 | .59 | .25 | .63 | .26 |
| Thailand | INFLT | .225 | 1.417 | .2156 | .42 | .33 | .47 | .30 |
| | SRPLS | .031 | .023 | .9823 | .98 | .29 | .98 | .30 |
| | TTRD | .169 | 1.494 | .1954 | .39 | .25 | .35 | .25 |
| | EXPM | .333 | 1.303 | .2492 | .45 | .26 | .48 | .26 |
| All | INFLT | -.205 | -2.45 | .014 | .16 | .28 | .16 | .24 |
| | SRPLS | .646 | 2.66 | .077 | .18 | .33 | .23 | .39 |
| | TTRD | .126 | 3.63 | .0003 | .13 | .36 | .13 | .35 |
| | EXPM | .038 | .77 | .4420 | .70 | .35 | .71 | .35 |

Table 11.b: CAPITAL GROWTH
(INDIA, INDONESIA, KOREA, PAKISTAN, THAILAND, 1978-87)

| Country | variable | Coef. | t-ratio | p-ASY | p-MC1 | %R-MC1 | p-MC2 | %R-MC2 |
|---|---|---|---|---|---|---|---|---|
| India | INFLT | -.004 | -.096 | .9273 | .92 | .27 | .93 | .25 |
| | SRPLS | -.042 | -.563 | .5976 | .71 | .72 | .73 | .27 |
| | TTRD | .008 | .728 | .4995 | .72 | .72 | .73 | .30 |
| | EXPM | -.003 | -.153 | .8847 | .94 | .92 | .93 | .32 |
| Indonesia | INFLT | .086 | .414 | .6961 | .82 | .85 | .86 | .33 |
| | SRPLS | .002 | .009 | .9929 | 1.0 | .99 | 1.0 | .35 |
| | TTRD | .062 | 1.386 | .2243 | .36 | .44 | .45 | .31 |
| | EXPM | -.249 | -2.364 | .0645 | .24 | .26 | .27 | .33 |
| Korea | INFLT | .152 | 1.640 | .1620 | .39 | .36 | .37 | .36 |
| | SRPLS | .590 | 1.892 | .1171 | .29 | .32 | .33 | .32 |
| | TTRD | -.013 | -.123 | .9066 | .89 | .94 | .95 | .22 |
| | EXPM | .146 | 2.664 | .0447 | .21 | .21 | .22 | .32 |
| Pakistan | INFLT | -.133 | -1.733 | .1437 | .31 | .35 | .36 | .25 |
| | SRPLS | .147 | 1.238 | .2706 | .47 | .45 | .46 | .26 |
| | TTRD | .001 | .144 | .8911 | .92 | .91 | .92 | .25 |
| | EXPM | -.022 | -2.163 | .0829 | .29 | .29 | .30 | .33 |
| Thailand | INFLT | .177 | 2.392 | .0622 | .22 | .23 | .24 | .29 |
| | SRPLS | .325 | 1.907 | .1148 | .34 | .23 | .24 | .21 |
| | TTRD | .002 | .549 | .6064 | .78 | .76 | .77 | .30 |
| | EXPM | -.088 | -1.360 | .2318 | .34 | .36 | .37 | .25 |
| All | INFLT | .094 | 2.46 | .0138 | .29 | .34 | .27 | .30 |
| | SRPLS | .091 | 1.09 | .276 | .56 | .30 | .60 | .35 |
| | TTRD | .017 | 1.31 | .1915 | .51 | .38 | .55 | .39 |
| | EXPM | .0007 | .034 | .9725 | .97 | .37 | .98 | .38 |

Table 11.c: PRODUCTIVITY GROWTH
(INDIA, INDONESIA, KOREA, PAKISTAN, THAILAND, 1978-87)

| Country | variable | Coef. | t-ratio | p-ASY | p-MC1 | %R-MC1 | p-MC2 | %R-MC2 |
|---|---|---|---|---|---|---|---|---|
| India | INFLT | .031 | .113 | .9143 | .99 | .23 | .99 | .23 |
| | SRPLS | -.479 | -.843 | .4376 | .64 | .27 | .59 | .31 |
| | TTRD | .080 | .935 | .3926 | .67 | .38 | .67 | .39 |
| | EXPM | -.059 | -.339 | .7487 | .87 | .31 | .89 | .32 |
| Indonesia | INFLT | -1.056 | -1.835 | .1259 | .40 | .38 | .40 | .32 |
| | SRPLS | .966 | 1.113 | .3163 | .55 | .38 | .56 | .39 |
| | TTRD | .150 | 1.040 | .3462 | .55 | .34 | .54 | .33 |
| | EXPM | -.380 | -1.275 | .2585 | .55 | .37 | .57 | .37 |
| Korea | INFLT | -.184 | -1.121 | .3131 | .55 | .28 | .55 | .36 |
| | SRPLS | .322 | .640 | .5503 | .70 | .36 | .77 | .35 |
| | TTRD | .069 | 3.792 | .0127 | .11 | .36 | .09 | .31 |
| | EXPM | .031 | .385 | .7159 | .83 | .31 | .82 | .30 |
| Pakistan | INFLT | -.131 | -.919 | .8581 | .97 | .26 | .95 | .26 |
| | SRPLS | .784 | .684 | .5246 | .71 | .30 | .69 | .35 |
| | TTRD | -.119 | -1.121 | .3132 | .59 | .31 | .59 | .35 |
| | EXPM | -.127 | -1.140 | .3060 | .47 | .29 | .54 | .32 |
| Thailand | INFLT | .144 | 1.149 | .3025 | .48 | .25 | .53 | .24 |
| | SRPLS | .145 | .283 | .7889 | .88 | .19 | .86 | .28 |
| | TTRD | .100 | 1.075 | .3316 | .57 | .32 | .54 | .29 |
| | EXPM | .606 | 2.240 | .0519 | .13 | .25 | .21 | .20 |
| All | INFLT | -.263 | -3.3 | .001 | .08 | .25 | .06 | .27 |
| | SRPLS | .537 | 2.13 | .033 | .21 | .25 | .27 | .32 |
| | TTRD | .103 | 2.54 | .011 | .24 | .37 | .24 | .37 |
| | EXPM | .033 | .624 | .533 | .79 | .27 | .79 | .27 |

Table 12: MLE (GHANA, IVORY COAST, KENYA, MALAWI, MOROCCO, ZAMBIA, 1977-88)

| | variable | Coef. | t-ratio | p-Asy | p-MC1 | %R-MC1 | P-MC2 | %R-MC2 |
|---|---|---|---|---|---|---|---|---|
| GDP Growth | INFLT | .021 | .944 | .345 | .81 | .66 | .82 | .65 |
| | SRPLS | .113 | 1.41 | .158 | .81 | .69 | .76 | .67 |
| | TTRD | .0212 | .942 | .345 | .85 | .73 | .86 | .73 |
| | EXPM | .004 | .78 | .438 | .85 | .70 | .87 | .70 |
| Capital Growth | INFLT | .039 | 12.34 | .000 | .09 | .67 | .15 | .59 |
| | SRPLS | .119 | 16.45 | .000 | .09 | .79 | .10 | .62 |
| | TTRD | -.016 | -10.7 | .000 | .17 | .65 | .15 | .65 |
| | EXPM | -.014 | -19.2 | .000 | .09 | .72 | .07 | .68 |
| Productivity Growth | INFLT | .080 | 3.40 | .0007 | .79 | .61 | .44 | .68 |
| | SRPLS | .092 | 1.09 | .288 | .77 | .63 | .77 | .64 |
| | TTRD | .056 | 2.57 | .010 | .57 | .66 | .62 | .72 |
| | EXPM | .014 | 2.47 | .013 | .58 | .61 | .55 | .63 |

Table 13: MLE (MEXICO, ARGENTINA, CHILI, COLOMBIA, ECUADOR, PARAGUAY: 1973-88)

| | variable | Coef. | t-ratio | p-ASY | p-MC1 | %R-MC1 | p-MC2 | %R-MC2 |
|---|---|---|---|---|---|---|---|---|
| GDP Growth | INFLT | -.085 | -9.5 | .000 | .03 | .41 | .02 | .41 |
| | SRPLS | .274 | 4.09 | .000 | .04 | .47 | .18 | .46 |
| | TTRD | .101 | 6.2 | .000 | .04 | .40 | .03 | .42 |
| | EXPM | -.034 | -2.5 | .012 | .36 | .52 | .34 | .48 |
| Capital Growth | INFLT | -.058 | -18.1 | .000 | .01 | .38 | .01 | .43 |
| | SRPLS | -.092 | -5.32 | .000 | .08 | .41 | .09 | .46 |
| | TTRD | .014 | 4.54 | .000 | .09 | .37 | .08 | .38 |
| | EXPM | .071 | -14.3 | .000 | .02 | .34 | .01 | .37 |
| productivity Growth | INFLT | -.052 | -5.34 | .000 | .08 | .40 | .07 | .46 |
| | SRPLS | .357 | 5.02 | .000 | .10 | .38 | .12 | .50 |
| | TTRD | .093 | 5.7 | .000 | .03 | .37 | .04 | .40 |
| | EXPM | -.009 | -.61 | .543 | .84 | .47 | .85 | .45 |

Table 14: MLE (INDIA, INDONESIA, KOREA, PAKISTAN, THAILAND: 1978-87)

| | variable | Coef. | t-ratio | p-ASY | p-MC1 | %R-MC1 | p-MC2 | %R-MC2 |
|---|---|---|---|---|---|---|---|---|
| GDP Growth | INFLT | -.445 | -5.62 | .000 | .35 | .70 | .45 | .70 |
| | SRPLS | .694 | 2.99 | .014 | .52 | .70 | .60 | .76 |
| | TTRD | .074 | 2.3 | .037 | .73 | .72 | .68 | .72 |
| | EXPM | .074 | 1.8 | .090 | .67 | .66 | .72 | .71 |
| Capital Growth | INFLT | -.072 | -3.95 | .000 | .58 | .72 | .57 | .70 |
| | SRPLS | .149 | 6.254 | .000 | .43 | .74 | .44 | .73 |
| | TTRD | .004 | 1.52 | .150 | .82 | .76 | .80 | .86 |
| | EXPM | -.004 | -1.1 | .290 | .85 | .70 | .88 | .77 |
| Productivity Growth | INFLT | -.845 | -19.9 | .000 | .14 | .63 | .15 | .69 |
| | SRPLS | -.010 | -.09 | .93 | .97 | .75 | .96 | .76 |
| | TTRD | -.058 | -3.05 | .009 | .65 | .73 | .67 | .75 |
| | EXPM | .268 | 11.56 | .000 | .19 | .66 | .25 | .72 |

Table 15. Panel regression estimates

| | GDP GROWTH | | | | CAPITAL ACCUMULATION | | | |
|---|---|---|---|---|---|---|---|---|
| | INFLT | SRPLS | TTRD | EXPM | INFLT | SRPLS | TTRD | EXPM |
| Coef. | -.033 | .273 | .074 | -.016 | -.028 | .025 | .015 | -.020 |
| t-Stat | -3.41 | 4.03 | 3.62 | -2.24 | -2.95 | .40 | .96 | -3.05 |
| p-val | .001 | .000 | .000 | .026 | .003 | .693 | .338 | .003 |

Table 15 cont. Panel regression estimates

| | PRODUCTIVITY GROWTH | | | |
|---|---|---|---|---|
| | INFLT | SRPLS | TTRD | EXPM |
| Coef. | -.025 | .275 | .064 | -.011 |
| t-Stat | -1.97 | 3.26 | 3.34 | -1.29 |
| p-val | .051 | .001 | .001 | .1 |

Table 16. Testing equality of SURE coefficients

| | INFLT | SRPLS | TTRD | EXPM | INFLT | SRPLS | TTRD | EXPM | INFLT | SRPLS | TTRD | EXPM |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | GDP: AFRICA | | | | Capital: AFRICA | | | | Productivity: AFRICA | | | |
| Wald | 51.2 | 17.4 | 18.6 | 10.6 | 198 | 35.9 | 15.6 | 48.3 | 15.3 | 32.6 | 7.16 | 8.36 |
| $P_{ART}$ | .000 | .004 | .002 | .059 | .000 | .000 | .000 | .000 | .009 | .000 | .209 | .137 |
| $P_{WC1}$ | .05 | .44 | .37 | .55 | .01 | .17 | .39 | .05 | .35 | .13 | .60 | .55 |
| $\%R_{WC1}$ | .50 | .58 | .63 | .53 | .58 | .60 | .55 | .57 | .45 | .59 | .47 | .41 |
| $P_{WC2}$ | .05 | .38 | .44 | .62 | .02 | .10 | .39 | .09 | .55 | .19 | .71 | .67 |
| $\%R_{WC2}$ | .67 | .59 | .63 | .57 | .48 | .58 | .57 | .59 | .67 | .68 | .48 | .53 |
| QLR | 32.3 | 12.4 | 9.69 | 5.16 | 29.2 | 24.0 | 15.8 | 21.2 | 35.4 | 54.3 | -27 | -14 |
| $P_{WC1}$ | .14 | .47 | .58 | .60 | .12 | .21 | .38 | .29 | .06 | .03 | .95 | .86 |
| $P_{WC2}$ | .15 | .40 | .56 | .67 | .14 | .20 | .33 | .27 | .13 | .04 | .98 | .93 |
| | GDP: South America | | | | Capital: South America | | | | Productivity: South America | | | |
| Wald | 19.8 | 3.05 | 16.5 | 33.3 | 47.9 | 25.8 | 6.41 | 13.5 | 1.56 | 1.81 | 11.6 | 39.7 |
| $P_{ART}$ | .001 | .693 | .005 | .000 | .000 | .000 | .268 | .019 | .002 | .874 | .041 | .000 |
| $P_{WC1}$ | .40 | .94 | .38 | .09 | .01 | .05 | .72 | .27 | .41 | .97 | .41 | .02 |
| $\%R_{WC1}$ | .57 | .46 | .54 | .54 | .52 | .38 | .43 | .38 | .59 | .52 | .47 | .43 |
| $P_{WC2}$ | .34 | .96 | .36 | .09 | .02 | .05 | .78 | .31 | .35 | .99 | .44 | .03 |
| $\%R_{WC2}$ | .61 | .47 | .53 | .54 | .43 | .38 | .49 | .39 | .63 | .53 | .47 | .58 |
| QLR | 11.5 | 2.01 | 5.51 | 9.72 | 28.6 | 26.7 | 6.13 | 3.43 | 10.5 | .65 | 3.98 | 31.6 |
| $P_{WC1}$ | .55 | .86 | .68 | .47 | .07 | .08 | .63 | .58 | .61 | .93 | .47 | .04 |
| $P_{WC2}$ | .47 | .84 | .70 | .56 | .03 | .06 | .65 | .64 | .47 | .92 | .50 | .05 |
| | GDP: ASIA | | | | Capital: Asia | | | | Productivity: Asia | | | |
| Wald | 6.30 | 2.59 | 28.6 | 3.93 | 9.97 | 5.93 | 4.16 | 15.7 | 6.58 | 2.18 | 10.3 | 4.44 |
| $P_{ART}$ | .178 | .630 | .000 | .416 | .041 | .205 | .385 | .003 | .016 | .703 | .036 | .350 |
| $P_{WC1}$ | .79 | .92 | .19 | .92 | .60 | .77 | .89 | .41 | .79 | .96 | .65 | .91 |
| $\%R_{WC1}$ | .64 | .60 | .68 | .66 | .63 | .65 | .67 | .64 | .68 | .61 | .65 | .67 |
| $P_{WC2}$ | .83 | .93 | .26 | .92 | .62 | .78 | .93 | .45 | .82 | .95 | .62 | .93 |
| $\%R_{WC2}$ | .62 | .65 | .73 | .71 | .61 | .62 | .62 | .68 | .66 | .64 | .62 | .68 |
| QLR | 19.7 | .94 | 26.8 | 8.25 | 58.7 | 63.0 | 69.1 | 47.1 | 20.2 | -3.0 | 71.6 | 47.6 |
| $P_{WC1}$ | .29 | .70 | .20 | .59 | .03 | .01 | .01 | .07 | .30 | .84 | .19 | .47 |
| $P_{WC2}$ | .39 | .78 | .24 | .59 | .02 | .02 | .02 | .10 | .38 | .88 | .19 | .49 |

Table 17a. CAPITAL GROWTH: UNCONSTRAINED ESTIMATES
(GHANA, IVORY COAST, KENYA, MALAWI, MOROCCO, ZAMBIA: 1977-88)

| Country | variable | Coef. | t-ratio | p-ASY | P-MC1 | %R-MC1 | P-MC2 | %R-MC2 |
|---------|----------|-------|---------|-------|-------|--------|-------|--------|
| Ghana | INFLT | -.012 | -.883 | .3870 | .58 | .23 | .63 | .28 |
|  | SRPLS | .009 | .413 | .6839 | .77 | .28 | .79 | .30 |
|  | TTRD | -.020 | -4.959 | .0001 | .07 | .33 | .07 | .34 |
|  | EXPM | -.008 | -4.289 | .0003 | .09 | .33 | .09 | .36 |
| Cote D'Ivoire | INFLT | .885 | 7.711 | .000 | .01 | .21 | .01 | .19 |
|  | SRPLS | .009 | .413 | .6839 | .77 | .28 | .79 | .30 |
|  | TTRD | -.020 | -4.959 | .0001 | .07 | .33 | .07 | .34 |
|  | EXPM | -.008 | -4.289 | .0003 | .09 | .33 | .09 | .36 |
| Kenya | INFLT | .041 | .448 | .6584 | .79 | .20 | .79 | .21 |
|  | SRPLS | .009 | .413 | .6839 | .77 | .28 | .79 | .30 |
|  | TTRD | -.020 | -4.959 | .0001 | .07 | .33 | .07 | .34 |
|  | EXPM | -.008 | -4.289 | .0003 | .09 | .33 | .09 | .36 |
| Malawi | INFLT | -.360 | -6.05 | .000 | .02 | .19 | .01 | .24 |
|  | SRPLS | .009 | .413 | .6839 | .77 | .28 | .79 | .30 |
|  | TTRD | -.020 | -4.959 | .0001 | .07 | .33 | .07 | .34 |
|  | EXPM | -.008 | -4.289 | .0003 | .09 | .33 | .09 | .36 |
| Morocco | INFLT | .592 | 4.341 | .0003 | .05 | .30 | .02 | .31 |
|  | SRPLS | .009 | .413 | .6839 | .77 | .28 | .79 | .33 |
|  | TTRD | -.020 | -4.959 | .0001 | .07 | .33 | .07 | .34 |
|  | EXPM | -.008 | -4.289 | .0003 | .09 | .33 | .09 | .36 |
| Zambia | INFLT | -.018 | -1.365 | .1862 | .34 | .16 | .45 | .20 |
|  | SRPLS | .009 | .413 | .6839 | .77 | .28 | .79 | .33 |
|  | TTRD | -.020 | -4.959 | .0001 | .07 | .33 | .07 | .34 |
|  | EXPM | -.008 | -4.289 | .0003 | .09 | .33 | .09 | .36 |

Table 17b. CAPITAL GROWTH: UNCONSTRAINED ESTIMATES
(MEXICO, ARGENTINA, CHILI, COLOMBIA, ECUADOR, PARAGUAY: 1973-88)

| Country | variable | Coef. | t-ratio | p-ASY | p-MC1 | %R-MC1 | p-MC2 | %R-MC2 |
|---|---|---|---|---|---|---|---|---|
| Mexico | INFLT | -.1039 | -11.015 | .0000 | .01 | .20 | .01 | .16 |
| | SRPLS | -.043 | -1.430 | .1652 | .40 | .26 | .36 | .25 |
| | TTRD | .009 | 1.254 | .2214 | .55 | .38 | .54 | .36 |
| | EXPM | -.029 | -2.927 | .0072 | .03 | .17 | .04 | .20 |
| Argentina | INFLT | -.0408 | -3.058 | .0053 | .06 | .23 | .06 | .18 |
| | SRPLS | -.043 | -1.430 | .1652 | .40 | .26 | .36 | .25 |
| | TTRD | .009 | 1.254 | .2214 | .55 | .38 | .54 | .36 |
| | EXPM | -.029 | -2.927 | .0072 | .03 | .17 | .04 | .20 |
| Chili | INFLT | -.011 | -2.047 | .0513 | .41 | .20 | .21 | .21 |
| | SRPLS | -.043 | -1.430 | .1652 | .40 | .26 | .36 | .25 |
| | TTRD | .009 | 1.254 | .2214 | .55 | .38 | .54 | .36 |
| | EXPM | -.029 | -2.927 | .0072 | .03 | .17 | .04 | .20 |
| Colombia | INFLT | -.059 | -1.364 | .1846 | .44 | .33 | .42 | .30 |
| | SRPLS | -.043 | -1.430 | .1652 | .40 | .26 | .36 | .25 |
| | TTRD | .009 | 1.254 | .2214 | .55 | .38 | .54 | .36 |
| | EXPM | -.029 | -2.927 | .0072 | .03 | .17 | .04 | .20 |
| Ecuador | INFLT | -.186 | -4.112 | .0084 | .02 | .19 | .01 | .17 |
| | SRPLS | -.043 | -1.430 | .1652 | .40 | .26 | .36 | .25 |
| | TTRD | .009 | 1.254 | .2214 | .55 | .38 | .54 | .36 |
| | EXPM | -.029 | -2.927 | .0072 | .03 | .17 | .04 | .20 |
| Paraguay | INFLT | -.2203 | -1.975 | .0594 | .21 | .20 | .24 | .23 |
| | SRPLS | -.043 | -1.430 | .1652 | .40 | .26 | .36 | .25 |
| | TTRD | .009 | 1.254 | .2214 | .55 | .38 | .54 | .36 |
| | EXPM | -.029 | -2.927 | .0072 | .03 | .17 | .04 | .20 |

Table 17c. CAPITAL GROWTH: UNCONSTRAINED ESTIMATES
(INDIA, INDONESIA, KOREA, PAKISTAN, THAILAND: 1978-87)

| Country | variable | Coef. | t-ratio | p-ASY | p-MC1 | %R-MC1 | p-MC2 | %R-MC2 |
|---------|----------|-------|---------|-------|-------|--------|-------|--------|
| India | INFLT | -.0012 | -.0380 | .9705 | .99 | .18 | .96 | .17 |
| | SRPLS | -.0458 | -.772 | .4599 | .58 | .14 | .57 | .19 |
| | TTRD | .0092 | 1.318 | .2199 | .74 | .98 | .62 | .32 |
| | EXPM | -.0140 | -1.524 | .1617 | .41 | .31 | .37 | .30 |
| Indonesia | INFLT | .3923 | 1.788 | .1073 | .40 | .34 | .43 | .35 |
| | SRPLS | -.1631 | .521 | .6148 | .81 | .33 | .79 | .36 |
| | TTRD | .0484 | .900 | .3917 | .62 | .27 | .61 | .26 |
| | EXPM | -.0140 | -1.524 | .1617 | .41 | .31 | .37 | .30 |
| Korea | INFLT | .3254 | 3.343 | .0086 | **.01** | .23 | **.01** | .29 |
| | SRPLS | 1.051 | 2.941 | .0165 | .14 | .36 | .13 | .36 |
| | TTRD | .2029 | 1.799 | .1050 | .35 | .30 | .29 | .26 |
| | EXPM | -.0140 | -1.524 | .1617 | .41 | .31 | .37 | .30 |
| Pakistan | INFLT | -.1763 | -2.357 | .0428 | .13 | .19 | .18 | .24 |
| | SRPLS | .2201 | 1.744 | .1152 | .17 | .13 | .35 | .27 |
| | TTRD | .001 | .120 | .9070 | .92 | .13 | .93 | .13 |
| | EXPM | -.0140 | -1.524 | .1617 | .41 | .31 | .37 | .30 |
| Thailand | INFLT | .1455 | 2.134 | .0616 | .24 | .27 | .26 | .26 |
| | SRPLS | .3676 | 1.884 | .0922 | .34 | .36 | .24 | .44 |
| | TTRD | -.0287 | -.802 | .4429 | .69 | .26 | .73 | .28 |
| | EXPM | -.0140 | -1.524 | .1617 | .41 | .31 | .37 | .30 |

# BIBLIOGRAPHY

Affleck-Graves, J. and B. McDonald (1990). *Multivariate tests of asset pricing: the comparative power of alternative statistics*, Journal of Financial and Quantitative Analysis 25, 163-185.

Amsler, C.E. and P. Schmidt (1985). *A Monte Carlo investigation of the accuracy of multivariate CAPM tests*, Journal of Financial Economics 14, 359-375.

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

Atkinson, S.E. and P.W. Wilson (1992). *The Bias of bootstrapped versus conventional standard errors in the general linear and SUR models*, Econometric Theory 8, 258-275.

Attfield, C.L.F. (1995). *A Bartlett adjustment to the likelihood ratio test for a system of equations*, Journal of Econometrics 66, 207-223.

Barnard, G.A. (1963). Comment, Journal of the Royal Statistical Society B, 25, 294.

Barndorff-Nielsen, O.E. and P. Blaesild (1986). *A note on the calculation of Bartlett adjustments*, J. R. statist. Soc. B 48, 353-358.

Bartlett, M.S. (1937). *Properties of sufficiency and statistical tests*, Proc. R. Soc. A 160, 268-282.

Bartlett, M.S. (1947). *Multivariate Analysis*, Journal of the Royal Statistical Society Suppl., 9, 176-197.

Bartlett, M.S. (1939). *A Note on Tests of Significance in Multivariate Analysis*, Proceedings of the Cambridge Philosophical Society, 35, 180-185.

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

Beran, R. (1988). *Prepivoting test statistics: a bootstrap view of asymptotic refinements*,

Journal of the American Statistical association 83 (403), 687-697.

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

Breusch, T.S. (1980). *Useful invariance results for generalized regression models*, Journal of Econometrics 13, 327-340.

Breusch, T.S. (1979). *Conflict among criteria for testing hypotheses: extensions and comments*, Econometrica 47 (1), 203-207.

Dufour, J.M. (1996). *Some impossibility theorems in econometrics with applications to structural and dynamic models*, Econometrica (forthcoming).

Dufour, J.M. (1995). *Monte Carlo tests in the presence of nuisance parameters, with econometric applications*, working paper, Centre de recherche et développement en économique (C.R.D.E.) and Département de sciences économiques, Université de Montréal.

Dufour, J.M. (1989). *Nonlinear Hypotheses, Inequality Restrictions and Non-Nested Hypotheses: Exact Simultaneous Tests in Linear Regressions*, Econometrica, 57 (2), 335-355.

Dufour, J.M. and J.F. Kiviet (1994). *Exact inference methods for first-order autoregressive distributed lag models*, working paper, Centre de recherche et développement en économique (C.R.D.E.) and Département de sciences économiques, Université de Montréal and Tinbergen Institute and Faculty of economics and econometrics, University of Amsterdam.

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

Dwass, M. (1957). *Modified Randomization tests for Non-parametric Hypotheses*, Annals of Mathematical Statistics, 28, 181-187.

Eakin, K., McMillen D.P. and M.J. Buono (1990). *Constructing confidence intervals using*

*bootstrap: an application to a multiproduct cost function*, Review of Economics and Statistics 72, 339-344.

Evans, G.B.A. and N.E. Savin (1982). *Conflict among the criteria revisited; the W, LR and LM tests*, Econometrica 50 (3), 737-748.

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

Frydenberg, M. and J.L. Jensen (1989). *Is the 'improved likelihood' really improved in the discrete case?*, Biometrika 76 (4), 655-661.

Gibbons, M.R., Ross, S.A. and J. Shanken (1989). *A test of the efficiency of a given portfolio*, Econometrica 57, 1121-1152.

Gonzalo, J. and J.Y. Pitarakis (1994). Cointegration Analysis in large systems, working paper, Boston university, department of economics.

Gouriéroux, C., Monfort A. and Renault E. (1993). *Tests sur le noyau, l'image et le rang de la matrice des coefficients d'un modèle linéaire multivarié*, Annales d'Économie et Statistiques 32, 81-111.

Gouriéroux, C., Monfort A. and Renault E. (1995). *Inference in factor models*, in: Maddala G.S., Phillips C.B. and T.N. Srinivasan eds., *Advances in econometrics and quantitative economics*, Blackwell, Oxford, U.K. and Cambridge, USA, 311-353.

Hall, P. (1988). *Theoretical comparison of the bootstrap confidence intervals*, The annals of statistics 16 (3), 927- 953.

Hall P. and J. Horowitz (1996). *Bootstrap critical values for tests based on generalized-method-of-moments estimators*, Econometrica 64, 891-916.

Harvey, A.C. and G.D.A. Phillips (1982). *Testing for contemporaneous correlation of disturbances in systems of regression equations*, Bulletin of Economic Research, 79-91.

Hashimoto, N. and K. Ohtani (1990). *An exact test for linear restrictions in Seemingly*

*Unrelated Regressions with the same regressors*, Economics Letters 32, 243-346.

Hollas, E.A. (1991). *Performance of Bartlett adjustment for certain likelihood ratio tests*, Communications in Statistics-Simulation 20 (2&3), 449-462.

Hotelling H. (1947). *Multivariate quality control illustrated by the air testing of sample bomb sights*, Techniques of Statistical analysis, Ch, II New York: McGraw-Hill.

Hotelling H. (1951). *A generalized T-Test and measure of multivariate dispersion*, in *Proceedings of the second Berkeley Symposium*, Berkeley: University of California Press, 23-42

Italianer, A. (1985). *A small-sample correction for the likelihood ratio test*, Economic Letters 19, 315-317.

Jayatissa, W.A. (1977). *Tests of equality between sets of coefficients in two linear regressions when disturbance variances are unequal*, Econometrica 45, 1291-1292.

Jeong, J. and G.S. Maddala (1993). *A perfective on application of bootstrap methods in econometrics*, in: G.S. Maddala, C.R. Rao and H.D. Vinod, eds., *Handbook of Statistics* 11, Elsevier Science Publishers B.V., 573-610.

Jobson, J.D. and B. Korkie (1982), *Potential performance and tests of portfolio efficiency*, Journal of Financial Economics 10, 433-466.

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

Kariya T. (1985). Testing in the multivariate general linear model. Economic research series No. 22, The Institute of Economic Research, Hitotsubashi University, Japan.

Kiviet, J.F. and J.M. Dufour (1996). *Exact tests in single equation autoregressive distributed lag models*, Journal of Econometrics (forthcoming).

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

Lawley, D.N. (1938). *A generalization of Fisher's Z-test*, BiometriKa 30, 180-187.

Lawley, D.N. (1956). *A general method for approximating to the distribution of the likelihood ratio statistic*, BiometriKa 43, 295-303.

MacKinlay, A. C. (1987). *On multivariate tests of the capital asset pricing model*, Journal of Financial Economics 18, 341-372.

Martin, M.A. (1990). *On bootstrap iteration for coverage correction in confidence intervals*, Journal of the American Statistical association 85 (412),1105-1118.

Mackinlay, A.C. (1987). *On multivariate tests of the CAPM*, Journal of Financial economics 18, 341-371.

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

Nanda, D.N. (1950). *Distribution of the Sum of Roots of a Determinantal Equation Under a Certain Condition*, Annals of Mathematical Statistics, 21, 432-439.

Ohtani, K. and T. Toyoda (1985). *A Monte Carlo study of the Wald, LM and LR tests in a heteroscedastic linear model*, Communications in Statistics-Simulation and Computation 14 (3), 735-746.

Phillips, P.C.B. (1985). *The Exact Distribution of the SUR estimator*, Econometrica, 53, 745-756.

Pillai, K.C.S. (1955). *Some New Test Criteria in Multivariate Analysis*, Annals of Mathematical Statistics, 43, 117-121.

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

Rayner, R.K. (1990a). *Bartlett's correction and the bootstrap in normal linear regression models*, Economics Letters 33, 255-258.

Rayner, R.K. (1990b). *Bootstrap tests for the generalized least squares regression models*, Economics Letters 34, 261-265.

Rilstone, P. and M. Veall (1993). *Using bootstrapped confidence intervals for improved*

*inferences with seemingly unrelated regression equations,* working paper.

Rocke, D. M. (1989). *Bootstrap Bartlett adjustment in Seemingly Unrelated Regressions,* Journal of the American Statistical association 84 (406), 598-601.

Rothenberg, T. J. (1984). *Hypothesis testing in linear models when the error covariance matrix is nonscalar,* Econometrica 52, 827-842.

Roy, S.N. (1953). *On a Heuristic Method of Test Construction and its Use on Multivariate Analysis,* Annals of mathematical Statistics, 24, 220-238.

Shanken, J. (1986). *Testing portfolio efficiency when the zero-beta rate is unknown: a note,* Journal of finance 41, 269-276.

Shanken, J. (1985). *Multivariate tests of the zero-beta CAPM,* Journal of Financial Econ. 14, 325-348.

Srivastava, V.K. and D.E. Giles (1987). *Seemingly Unrelated Regression Equation Models,* New York: Marcel Deckker.

Stambaugh, R.F. (1982). *On the exclusion of assets from tests of the two-parameter model,* a sensitivity analysis, Journal of Financial Economics 10, 237-268.

Stewart, K.G. (1995a). *The functional equivalence of the W, LR and LM statistics,* Economic Letters 49, 109-112.

Stewart, K.G. (1995b). *Exact testing in multivariate regression,* working paper, department of economics, University of Victoria.

Taylor, T.G., Shonkwiler, J.S. and H. Theil (1986). *Monte Carlo and bootstrap testing of demand homogeneity,* Economics Letters 20, 55-57.

Theil, H. and D.G. Fiebig (1985). *Small sample and large equation systems,* Handbook of Statistics (V), 451-480.

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

Theil, H., Taylor, T.G. and J.S. Shonkwiler (1986). *Monte Carlo testing in systems of*

*equations*, Advances in Econometrics, vol.5, 227-239.

Vinod, H.D. (1993). *Bootstrap methods: applications in econometrics*, in: G.S. Maddala, C.R. Rao and H.D. Vinod, eds., *Handbook of Statistics* 11, Elsevier Science Publishers B.V., 692-661.

Wilks S.S (1932). *Certain Generalisations in the Analysis of Variance*, Biometrika 24, 471-494.

Williams, M. A. (1986). *An economic application of bootstrap statistical methods: Addyston pipe revisited*, Amer. Econom. 30, 52-58.

Wong, M.Y. (1989). *Likelihood estimation of a simple linear regression model when both variables have error*, Biometrika 76 (1), 141-148.

Wong, M.Y. (1991). *Bartlett adjustment to the likelihood ratio statistic for testing several slopes*, Biometrika 78 (1), 221-224.

Zellner, A. (1962). *An efficient method of estimating Seemingly Unrelated Regressions and tests for aggregation bias*, Journal of the American Statistical Association 57, 348-368.

# Chapter 2

# Monte Carlo tests for contemporaneous correlation of disturbances in multi-equation regression models

Abstract

This paper proposes finite sample procedures for testing the SURE specification in multi-equation regression models. We apply the technique of Monte Carlo (MC) tests [Dwass (1957), Barnard (1963)] to obtain exact tests based on standard LR and LM zero correlation tests. We also suggest a MC quasi-LR (QLR) test based on feasible generalized least squares (FGLS). We show that the latter statistics are pivotal under the null, which provides a basic motivation for applying randomized tests. Furthermore, we extend the exact independence test proposed by Harvey and Phillips (1982) to the multi-equation framework. The properties of the proposed tests are studied in a Monte Carlo experiment which shows that standard asymptotic LR and LM tests exhibit important size distortions. By contrast, MC tests achieve complete size control and display good power. Moreover, QLR MC tests performed best in terms of power, a result of interest from the point of view of simulation-based tests. The tests are applied to data used by Fischer (1993) to analyze the macroeconomic determinants of growth.

## 1. Introduction

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

This paper studies and proposes finite sample tests for independence against contemporaneous correlation of disturbances in a multi-equation SURE model. We use for that purpose the likelihood ratio (LR) and Lagrange multiplier (LM) test criteria; we also introduce quasi-LR (QLR) statistics based on FGLS estimates. These statistics have rather complicated null distributions. So to obtain finite sample tests, we shall exploit the technique of Monte Carlo (MC) tests [see Dwass (1957), Barnard (1963), Birnbaum (1974) and Dufour (1995)] which allows one to obtain provably exact randomized tests in finite samples using very small numbers of MC replications of the original test statistic under the null hypothesis.

Independence tests in multivariate models have been discussed in both the econometric and statistical literature. Breusch and Pagan (1980) provide an LM test for the diagonality of the error covariance matrix. Kariya (1981a) derives locally best invariant tests in a two-equation framework. Shiba and Tsurumi (1988) proposed Wald, LR, LM and Bayesian tests for the hypothesis that the error covariance is block-diagonal. Related results are also available in Kariya (1981b), Kariya et al (1984), and Cameron and Trivedi (1993). Except for one special case, these test procedures are only justified by asymptotic arguments. The exception is Harvey and Phillips (1982, section 3) who proposed exact independence tests between the errors of an equation and those of the other equations of the system. These tests (called EFT) involve conventional $F$ statistics for testing whether the (estimated) residuals added to each equation have zero coefficients. EFT tests may be applied in the context of general diagonality tests; for example, one may assess in turn whether the disturbances in each equation are independent of the disturbances in all other equations. Such a sequence of tests however raises the problem of taking into account the dependence between multiple tests, a problem not solved by Harvey and Phillips (1982).

A major problem in the multi-equation context comes from the fact that relevant null distributions are either difficult to derive or too complicated to use in practice. This is true even in the case of identical regressor matrices. Hence the applicable procedures rely heavily on asymptotic approximations whose accuracy can be quite poor. This is evident from the Monte Carlo results reported in Harvey and Phillips (1982) and Shiba and Tsurumi (1988), among others. In any case, it is widely acknowledged by now that standard multivariate LR-based asymptotic tests are unreliable in finite samples, in the sense that test sizes deviate from the nominal significance levels; see Dufour and Khalaf (1996a, b) for related simulation evidence. On the other hand, most reported studies suggest that LM independence tests have correct sizes. However, for the multi-equation cases examined here, LM asymptotic tests also exhibit size distortions.

The first step towards a finite sample exact test procedure involves deriving nuisance-parameter-free null distributions. In the context of independence tests, invariance results are known given two univariate or multivariate regression equations [Kariya (1981a, b), Kariya et al. (1984)]. The problem of nuisance parameters is yet unresolved in models involving more than two regression equations. Here, we show that the LR, LM and QLR independence test statistics are pivotal under the null, for multi-equation SURE systems. Though the proof of this result is not complex, it does not appear to be known in the literature. Of course, existing work in this area has typically focused on deriving $p$-values analytically. By contrast, the approach taken in this article does not require extracting exact distributions; we obtain provably exact $p$-values by using the technique of MC tests originally proposed by Dwass (1957), Barnard (1963). Note also that the normality assumption is not a prerequisite of the tests proposed here. For further references regarding MC tests, see Dufour and Kiviet (1994), Edgington (1980), Foutz (1980), Jöckel (1986) and Mariott (1979). Monte Carlo evidence on the performance of the various tests is also presented. The results show that while the asymptotic LR and LM tests seriously overreject, the MC versions of these tests achieve perfect size control and have good power.

The outline of this study is as follows. In Section 2, we present the framework and the independence test criteria; we next describe the MC test procedures. In Section 3, we report the Monte Carlo results. In Section 4, we apply the tests to data used by Fischer (1993) to analyze the macroeconomic determinants of growth.

## 2. Framework and test statistics

### 2.1 Model and standard estimators

Consider the seemingly unrelated regression model

$$(2.1) \quad Y_i = X_i \beta_i + u_i , \ i = 1 , \dots , p ,$$

where $Y_i$ is a vector of $n$ observations on a dependent variable, $X_i$ a $(n, k_i)$ matrix of

observations on $k_i$ explanatory variables, $\beta_i$ a vector of $k_i$ coefficients, and $u_i$ a vector of $n$ random disturbances. When $X_i = X_j$ , $i , j = 1 , \ldots , p,$ we have a multivariate linear regression model (MLR). The system may be rewritten in the form

(2.2)    $y = X\beta + u$ ,

where

$$
y = \begin{bmatrix} Y_1 \\ Y_2 \\ \cdot \\ Y_p \end{bmatrix}, \quad
X = \begin{bmatrix} X_1 & \cdot & \cdot & \cdot \\ \cdot & X_2 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & X_p \end{bmatrix}, \quad
u = \begin{bmatrix} u_1 \\ u_2 \\ \cdot \\ u_p \end{bmatrix}, \quad
\beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \cdot \\ \beta_p \end{bmatrix},
$$

so that $X$ has dimension $(nk, k)$, $y$ and $u$ each have dimension $(np, 1)$ and $\beta$ has dimension $(k,1)$, with $k = \sum_{i=1}^{p} k_i$ . We also suppose that

(2.3)    $E(u) = 0$ , $E(u_i \mu_j') = \sigma_{ij} I_n$ , $i , j = 1 , \ldots , p$ ,

hence $E(uu') = \Sigma \otimes I_p$ , where $\Sigma = [ \sigma_{ij} ]$ ; we also set $\sigma_{ii} = \sigma_i^2$ . The coefficients of the regression equations can be estimated by several methods of which the most well known are: (i) OLS applied to each equation, (i) two-step FGLS, (iii) iterative FGLS (IFGLS), and (iv) maximum likelihood (ML). Denote the OLS estimator by

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

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

(2.5)    $\hat{u}_i = Y_i - X_i\hat{\beta}_i = M_i\mu_i$ , $M_i = I_n - X_i(X_i'X_i)^{-1}X_i$ , $i = 1 , \ldots , p$ .

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

(2.6)    $\hat{\beta}_{FGLS} = [X'(S^{-1}\otimes I_n)X]^{-1}X'(S^{-1}\otimes I_n)y$ .

If the disturbances are normally distributed, then the log-likelihood is

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

The ML estimators satisfy the following normal equations

(2.8) $\quad X'(\tilde{\Sigma}^{-1} \otimes I_n)X\tilde{\beta} = X'(\tilde{\Sigma}^{-1} \otimes I_n)y$ , $\quad \tilde{\Sigma} = (\tilde{U}'\tilde{U})/n$ ,

where $\tilde{U} = [\tilde{u}_1, \dots, \tilde{u}_p]$, $\tilde{u}_i = Y_i - X_i\tilde{\beta}_i$, $i = 1, \dots, p$, and $\tilde{\beta} = (\tilde{\beta}_1', \dots, \tilde{\beta}_p')'$ . Iterative procedures are typically applied to obtain the ML estimates. Suppose $\tilde{\Sigma}^{(0)}$ is an initial estimate of $\Sigma$ . Using (2.8), we can solve for a first estimate of $\beta$

(2.9) $\quad \tilde{\beta}^{(0)} = \left(X'(\tilde{\Sigma}^{(0)} \otimes I_n)^{-1}X\right)^{-1} X'(\tilde{\Sigma}^{(0)} \otimes I_n)^{-1}y$ ,

from which an estimate of $u$ may be obtained

(2.10) $\quad \tilde{u}^{(1)} = y - X\tilde{\beta}^{(0)}$ .

This residual leads to an estimator $\tilde{\Sigma}^{(1)}$ of $\Sigma$ and another estimate $\tilde{\beta}^{(1)}$. Thus, the estimators at the $j$th iteration take the form

(2.11) $\quad \tilde{\beta}^{(j)} = [X'(\tilde{\Sigma}^{(j)} \otimes I_n)^{-1}X]^{-1} X'(\tilde{\Sigma}^{(j)} \otimes I_n)^{-1}y$ , $\quad \tilde{\Sigma}^{(j)} = (\tilde{U}^{(j)'}\tilde{U}^{(j)})/n$ ,

where $\tilde{U}^{(j)}$ is as in (2.8) , with $\tilde{u}_i^{(j)} = y_i - X_i\tilde{\beta}_i^{(j-1)}$ . Iterating this procedure to convergence yields ML estimates [Oberhofer and Kmenta (1974)]. With this notation, $\tilde{\beta}^{(0)}$ corresponds to $\tilde{\beta}_{FGLS}$ , while $\tilde{\beta}^{(J)}$ and $\tilde{\Sigma}^{(J)}$ denote the estimates achieved after $J$ iterations.

## 2.2 Independence test statistics

Given the setup described above, consider the problem of testing $H_0$, the hypothesis that $\Sigma$ is diagonal. Some notation we use throughout is first listed. Let $D_N(d_i)$ represent a diagonal matrix of dimension $N$, with $(d_1, \dots, d_N)$ along the diagonal. $H_0$ may then be expressed as

(2.12) $\quad H_0 : \Sigma = D_p(\sigma_i^2)$ , for some vector $(\sigma_1, \dots, \sigma_p)'$ .

Further, we will frequently refer to the standardized disturbances and the standardized residuals which we denote respectively

(2.13) $\quad w = (w_1', \dots, w_p')'$ , $\quad w_i = u_i/\sigma_i$ , $i = 1, \dots, p$ ,

(2.14) $\quad \tilde{w}^{(j)} = (\tilde{w}_1^{(j)'}, \dots, \tilde{w}_p^{(j)'})'$ , $\quad \tilde{w}_i^{(j)} = \tilde{u}_i^{(j)}/\sigma_i$ , $i = 1, \dots, p$, $j = 0, \dots, J$ .

In addition, let

(2.15)  $\tilde{\Omega}^{(j)} = \left[ \left( \tilde{w}_l^{(j)} \right)' \left( \tilde{w}_m^{(j)} \right) \right]_{l,m=1,\ldots,p}$ .

We also use an alternative notation in the case of OLS estimators. Specifically, let $\hat{w}_l$, $\hat{w}$ and $\hat{\Omega}$ correspond to $\tilde{w}_l^{(0)}$, $\tilde{w}^{(0)}$ and $\tilde{\Omega}^{(0)}$. Note the vector $w$ has a known distribution if the null is true.

The LR statistic for testing $H_0$ is a monotonic transformation of

(2.16)  $\tilde{\Lambda} = \dfrac{|D_p(\hat{\sigma}_i)|}{|\tilde{\Sigma}|}$ .

The asymptotic null distribution of $n \ln(\tilde{\Lambda})$ is $\chi^2 \left( \dfrac{p(p-1)}{2} \right)$ . We also consider

(2.17)  $\tilde{\Lambda}^{(j)} = \dfrac{|D_p(\hat{\sigma}_i)|}{|\tilde{\Sigma}^{(j)}|}$ , $j = 0, 1, \ldots, J$ .

Confirming with the notation set above, $\tilde{\Lambda}^{(0)}$ refers to the QLR statistic based on OLS estimates and $\tilde{\Lambda}^{(J)}$ corresponds to the LR criterion. The LM criterion is

(2.18)  $LM = n \displaystyle\sum_{i=2}^{p} \sum_{j=1}^{i-1} r_{ij}^2$ , $r_{ij} = \dfrac{\hat{u}_i' \hat{u}_j}{(\hat{u}_i' \hat{u}_i)^{1/2} (\hat{u}_j' \hat{u}_j)^{1/2}}$ ,

and has the same limiting null distribution as the LR statistic [see Breusch and Pagan (1980)]. On the other hand, a finite sample exact independence test was developed by Harvey and Phillips (1980); their procedure is applicable where the null hypothesis has the form

(2.19)  $H_{01} : \Sigma = \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \Sigma_{11} \end{bmatrix}$ .

Specifically, they propose the following statistic:

$$(2.20) \quad EFT = \frac{\hat{u}_1' \hat{V}_1 (\hat{V}_1' M_1 \hat{V}_1)^{-1} \hat{V}_1' \hat{u}_1}{\hat{u}_1' [I - \hat{V}_1 (\hat{V}_1' M_1 \hat{V}_1)^{-1} \hat{V}_1'] \hat{u}_1} \frac{n - k_1 - p + 1}{p - 1} ,$$

$$\hat{V}_1 = (\hat{u}_2 , \dots , \hat{u}_p) , \quad M_1 = I - X_1 (X_1' X_1)^{-1} X_1 ,$$

which follows an $F$ distribution with $(p-1 , n-k_1-p+1)$ degrees of freedom under $H_{01}$. The EFT statistic may be obtained as the usual $F$ statistic for testing whether the coefficients on $\hat{V}_1$ are zero in the regression of $y_1$ on $X_1$ and $\hat{V}_1$. As it stands, the EFT test is restricted to hypotheses of the form (2.19); in fact, the problem of testing (2.12) is not addressed at all in Harvey and Phillips (1982). Here, we suggest an extension to the multi-equation framework. Specifically, we propose to examine, in turn, whether $u_1$ is independent of $(u_2 , \dots , u_p)$, or $u_2$ is independent of $(u_3 , \dots , u_p)$ and so on so forth; a maximum of $(p-1)$ tests would be involved. Critical points from the relevant $F$-distribution may be obtained by using Boole-Bonferroni inequality. For example, the significance levels may be set to $\alpha_1 = \alpha/2$, $\alpha_2 = \alpha/(2^2)$, $\dots$ , $\alpha_{p-2} = \alpha/(2^{p-2})$, $\alpha_{p-1} = \alpha/(2^{p-2})$. Harvey and Phillips report that individual EFT tests have good power. Whether the proposed sequential EFT tests would fair well is an open question. It appears likely that the problems associated with sequential inference will become worse as the dimension of the error covariance matrix increases.

## 2.3 Exact distributional results

We proceed next to examine the finite sample distributions of the above defined LM, LR and QLR criteria. In particular, we show that the associated null distributions are nuisance parameter free. To do this, we first rewrite the statistics in terms of the vector of standardized residuals. Indeed, by simple manipulations we find that

$$(2.21) \quad LM = n \sum_{i=2}^{p} \sum_{j=1}^{i-1} r_{ij}^2 , \quad r_{ij} = \frac{\hat{w}_i' \hat{w}_j}{(\hat{w}_i' \hat{w}_i)^{1/2} (\hat{w}_j' \hat{w}_j)^{1/2}} ,$$

$$(2.22) \qquad \Lambda^{(j)} = \frac{\left|D_p(\sigma_i^{-1})\right| \left|D_p(\hat{\sigma}_i^2)\right| \left|D_p(\sigma_i^{-1})\right|}{\left|D_p(\sigma_i^{-1})\right| \left|\hat{\Sigma}^{(j)}\right| \left|D_p(\sigma_i^{-1})\right|} = \frac{\prod\limits_{i=1}^{p} \hat{w}_i'\hat{w}_i}{\left|\hat{\Omega}^{(j)}\right|} \quad , \quad 0 < j < J \ .$$

Under the null, the distribution of $\hat{w}_i = M_i \, w_i \ , \quad i = 1 \, , \, \dots \, , p$ is nuisance parameter free. It follows that the null distribution of the LM statistic is pivotal. Further, this would also imply that $\bar{\Lambda}^{(0)}$ has a pivotal null distribution.

We next examine the IFGLS standardized residual vector $\tilde{w}^{(j)}, \quad j = 1 \, , \, \dots \, , J$ . Let $Y$ refer to the stacked vector $\left((y_1/\sigma_1)' \, , \, \dots \, , (y_p/\sigma_p)'\right)'$ . Then $\beta^{(0)}$ may be rewritten as:

$$(2.23) \qquad \beta^{(0)} = \left[X'\left(\Sigma_0^{1/2}\hat{\Omega}^{(0)}\Sigma_0^{1/2}\otimes I_n\right)^{-1}X\right]^{-1}X'\left(\Sigma_0^{1/2}\hat{\Omega}^{(0)}\Sigma_0^{1/2}\otimes I_n\right)^{-1}y$$

$$= \left[X'(\Sigma_0^{-1/2}\otimes I_n)(\hat{\Omega}^{(0)}\otimes I_n)^{-1}(\Sigma_0^{-1/2}\otimes I_n)X\right]^{-1}X'(\Sigma_0^{-1/2}\otimes I_n)(\hat{\Omega}^{(0)}\otimes I_n)^{-1}(\Sigma_0^{-1}\otimes I_n)y$$

$$= D_p(\sigma_i I_k)\left[X'\left(\hat{\Omega}^{(0)}\otimes I_n\right)^{-1}X\right]^{-1}X'\left(\hat{\Omega}^{(0)}\otimes I_n\right)^{-1}Y \ .$$

It follows that

$$(2.24) \qquad \tilde{w}^{(1)} = \bar{M}_{(1)} \, Y = \bar{M}_{(1)} \, w \ , \quad \bar{M}_{(1)} = I_{nk} - X\left[X'(\hat{\Omega}^{(0)}\otimes I_n)^{-1}X\right]^{-1}X'(\hat{\Omega}^{(0)}\otimes I_n)^{-1} \ .$$

The same holds true if we consider further iterations. Formally, we have

$$(2.25) \qquad \tilde{w}^{(j)} = \bar{M}_{(j)}w \ , \quad j = 1 \, , \, \dots \, , J \ ,$$

where $\bar{M}_{(j)} = I_{nk} - X\left(X'(\hat{\Omega}^{(j-1)}\otimes I_n)^{-1}X\right)^{-1}X'(\hat{\Omega}^{(j-1)}\otimes I_n)^{-1}$ . The implications of this is that iterations beyond the first round give rise to pivotal statistics under the null provided $\hat{\Omega}^{(0)}$ is pivotal (which is typically the case).

## 2.4 Exact Monte Carlo p-values

The fact that the LM, LR and QLR statistics have nuisance parameter free null distributions entails that MC tests are easily applicable. In the following, we summarize the basic methodology involved, as it applies to the present context; for a more elaborate discussion of MC tests, see Dufour (1995).

Denote $T_0$ the observed test statistic $T$; the associated critical region of size $\alpha$ may be expressed as $G(T_0) \leq \alpha$ where $G(.)$ is the "$p$-value" function. By Monte Carlo methods, generate $N$ independent realizations $T_N = (T_1 \, , \, \dots \, , T_N)'$. Now rank $T_s$, $s = 0, 1, \dots, N$, in

non-decreasing order and obtain $\hat{p}_N(T_0)$ , where

$$(2.26) \quad \hat{p}_N(x) = \frac{N\hat{G}_N(x) + 1}{N + 1} \ , \quad \hat{G}_N(x) = \frac{1}{N} \sum_{i=1}^{N} I_{[0,\infty)}(T_i - x) \ , \quad I_A(x) = \begin{matrix} 1, \text{ if } x \in A \\ 0, \text{ if } x \in A \end{matrix}$$

The randomized critical region $\hat{p}_N \leq \alpha$ has the same level as the critical region $G(T_0) \leq \alpha$ . Indeed, it is shown in Dufour (1995) that

$$(2.27) \quad P \ [\hat{p}_N(T_0) \leq \alpha] \leq \frac{I[\alpha(N + 1)]}{N + 1} \ ,$$

where $I(z)$ is the largest integer less than or equal to z.

## 3. Monte Carlo experiments

In order to assess the performance of the various procedures discussed above, a set of Monte Carlo experiments were conducted for a five equation model ($p = 5$) with five explanatory variables including a constant term per equation. Three regressors are common to all equations. Hence, nine distinct regressors were used in the experiments and were generated using a multivariate normal distribution. The set of regressors were kept constant for all replications. The disturbances were generated from multivariate normal distributions. Several choices for the error covariance were considered and are listed in Table 1. The matrix labelled $\Sigma_1$ as well as the regression coefficients are from the empirical example we discuss in the next section. The other matrices were obtained by dividing certain elements of the Cholesky decomposition of $\Sigma_1$ by appropriate constants to decrease the covariance terms. Of course, the parameters under the null were obtained by setting the non-diagonal elements of $\Sigma_1$ to zero. The number of trials for the MC tests was set to 19 and 99 (N = 19, 99). The number of overall replications was 1000. All experiments were performed with Gauss 386iVM, version 3.2.13. The results presented in Table 2 consider for four tests:

LM, $\hat{\Lambda}$, $\bar{\Lambda}^{(1)}$, $\bar{\Lambda}$ and EFT. In what follows, we report our main findings.

(1) The asymptotic tests consistently overreject. Indeed, the empirical sizes are substantially larger than the nominal 5%. This is in accordance with well documented results on LR-based multivariate tests. On the other hand, our conclusions with respect to the LM test are not in agreement with available Monte Carlo evidence. The LM independence test was as yet found to work well. Here we show that it does not always work well in larger systems. In contrast, all MC tests achieve size control.

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

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

## 4. Empirical illustration

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

availability of observations on all included variables. We consider:

(i) the South-American region: Mexico, Argentina, Chili, Colombia, Ecuador and Paraguay (1973-1987)

(ii) the African region: Ghana, Côte D'Ivoire, Kenya, Malawi, Morocco and Zambia (1977-88)

(iii) the Asian region: Korea, Pakistan, Thailand, India, Indonesia (1978-87).

To be more specific, we consider, in each region, the following four SURE systems:

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

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

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

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

where $\Delta GDP_{it}$, $\Delta CPTL_{it}$, $\Delta PRDCT_{it}$, $\Delta LABOR_{it}$, $INFLAT_{it}$, $TRMTRD_{it}$, $SRPLS_{it}$ and $EXCM_{it}$ refer respectively to GDP growth, capital accumulation, productivity growth, labor force growth, inflation, terms of trade, the ratio of budget surplus to GDP and the black market premium on the exchange rate, in country $i$ and year $t$.

In all cases, we compute LM, QLR and LR tests, relying on OLS and two-step Zellner estimates respectively. We also calculate the generalized Harvey-Phillips test (EFT) described in Section 2.2. We use the following notation. The statistic $EFT_i$ tests whether the disturbances of the $i$th equation (associated with the $i$th country) are independent of the errors of the subsequent equations. Of course, the ordering of countries may affect the outcome of the test, given the rule devised to deal with the associated multiple tests. For practical purposes, we have considered the countries in alphabetical order within every region. We obtained the following results.

*GDP growth*

The hypothesis of independence was consistently rejected by all tests in the case of Africa: the asymptotic tests are strongly significant, whereas the MC tests are significant at 10%. In contrast all tests were not significant (at 10%) for the South-American countries. Turning to the Asian region, we observe that the EFT test rejects at 5%; the LM asymptotic and Monte Carlo tests are not significant (at 10%); the QLR asymptotic test is significant at 10%; however, the MC QLR is not significant (the $p$-value = .47). Similar evidence is noted for the LR test: while the asymptotic test is strongly significant, the MC tests does not reject.

*Capital growth*

All tests reject independence for the South-american region: the asymptotic LM test rejects at 5%, the other asymptotic tests are very strongly significant. Yet the MC and the EFT test are significant only at 10%. In the case of Africa, it is worth mentioning that although the asymptotic LM test is not significant at 10%, the MC LM test has a $p$-value of .026; the EFT and MC LR-based tests do not reject (p-values are greater than 10%), although the asymptotic counterpart show very small $p$-values. Finally, in the case of the Asian region, both asymptotic and MC tests strongly reject independence; the EFT test is only significant at 10%. However, both conventional and MC LM tests are not significant.

*Productivity growth*

All tests fail to reject independence for the South-american region. The same holds true for MC tests relating to the Asian region, except that the asymptotic LR criterion is very strongly significant. In connection with Africa, all except the EFT tests are significant. As was often the case, the asymptotic tests have much smaller $p$-values; the OLS-based tests reject only at 10%.

*Labor force growth*

The independence hypothesis is consistently rejected at 5% in the African countries; the EFT test is the exception and is only significant at 10%. In contrast, the only significant exact test in the case of South america that is significant is the LR Monte Carlo test; the OLS-based MC test and the EFT test have significance levels that exceed 10%, while all asymptotic tests are significant. With regards to Asia, we note the familiar evidence regarding LR-based tests: asymptotic is strongly significant while the MC tests fail to reject at levels higher than 10%. The EFT test is significant at 5%.

## 5. Conclusion

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

The feasibility of the approach suggested was illustrated through both a simulation experiment and an empirical application. The results show that asymptotic tests are indeed highly unreliable; in contrast, MC tests achieve size control and have good power. We emphasize that OLS-based MC QLR tests performed extremely well. This aspect is important particularly in larger systems, since test procedures based on iterative estimators are typically more expensive from the point of view of MC tests.

Table 1. Empirical sizes of LM and Quasi-LR independence tests

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

Table 2. Empirical rejections of various independence tests (n=25)

| TEST | $\Sigma_0$ ($H_0$) MC reps=20 | | $\Sigma_1$ | | $\Sigma_2$ | | $\Sigma_3$ | | $\Sigma_4$ | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | MC reps | | MC reps | | MC reps | | MC reps | |
| | ASY | MC | 20 | 100 | 20 | 100 | 20 | 100 | 20 | 100 |
| LM | .105 | .045 | .998 | 1.0 | .911 | .954 | .704 | .794 | .444 | .500 |
| $QLR_{OLS}$ | .193 | .040 | 1.0 | 1.0 | .947 | .971 | .744 | .820 | .438 | .494 |
| $QLR_{GLS}$ | .260 | .040 | 1.0 | 1.0 | .959 | .979 | .750 | .825 | .429 | .504 |
| LR | .267 | .047 | 1.0 | 1.0 | .961 | .980 | .746 | .824 | .428 | .494 |
| EFT | .049 | | 1.0 | | .896 | | .687 | | .316 | |

Table 3. Independence tests: GDP growth SURE systems

|  |  | South America | Africa | Asia |
|---|---|---|---|---|
| EFT | $EFT_1$ | .7927 | .1613 | .0215** |
|  | $EFT_2$ | .7470 | .4964 | .3113 |
|  | $EFT_3$ | .8810 | .9137 | .4277 |
|  | $EFT_4$ | .8647 | .0055* | .3873 |
|  | $EFT_5$ | .9290 | .6005 | - |
| LM | ASY | .9425 | .0466 | .4384 |
|  | MC | .977 | .081* | .611 |
| $QLR_{OLS}$ | ASY | .9242 | .010 | .0872 |
|  | MC | .981 | .062* | .470 |
| LR | ASY | .4374 | .0000 | .000 |
|  | MC | .978 | .082* | .412 |

Table 4. Independence tests: Capital growth SURE systems

|  |  | South America | Africa | Asia |
|---|---|---|---|---|
| EFT | $EFT_1$ | .1592 | .2503 | .3399 |
|  | $EFT_2$ | .2249 | .0854 | .6323 |
|  | $EFT_3$ | .0111* | .3004 | .0069* |
|  | $EFT_4$ | .7156 | .1422 | .6355 |
|  | $EFT_5$ | .7679 | .7581 | - |
| LM | ASY | .0350 | .1023 | .2449 |
|  | MC | .061* | .026* | .367 |
| $QLR_{OLS}$ | ASY | .0058 | .0049 | .0000 |
|  | MC | .096* | .132 | .002** |
| LR | ASY | .0000 | .0000 | .0000 |
|  | MC | .053* | .249 | .001** |

Table 5. Independence tests: Productivity growth SURE systems

| | | South America | Africa | Asia |
|---|---|---|---|---|
| EFT | $EFT_1$ | .9765 | .1312 | .5003 |
| | $EFT_2$ | .8294 | .3912 | .5421 |
| | $EFT_3$ | .6037 | .3738 | .8683 |
| | $EFT_4$ | .9442 | .0519 | .2284 |
| | $EFT_5$ | .6962 | .8069 | - |
| LM | ASY | .9913 | .0356 | .5070 |
| | MC | .998 | .061* | .698 |
| $QLR_{OLS}$ | ASY | .9891 | .0012 | .0658 |
| | MC | .997 | .074* | .415 |
| LR | ASY | .7929 | .0000 | .0000 |
| | MC | .998 | .016** | .266 |

Table 6. Independence tests: Labor force growth SURE systems

| | | South America | Africa | Asia |
|---|---|---|---|---|
| EFT | $EFT_1$ | .4051 | .0734 | .0153** |
| | $EFT_2$ | .2594 | .1201 | .5698 |
| | $EFT_3$ | .5535 | .0085* | .2187 |
| | $EFT_4$ | .0580 | .0897 | .2661 |
| | $EFT_5$ | .0799 | .4900 | - |
| LM | ASY | .0686 | .0007 | .3040 |
| | MC | .103 | .004** | .457 |
| $QLR_{OLS}$ | ASY | .0108 | .0000 | .0063 |
| | MC | .126 | .006** | .140 |
| LR | ASY | .0000 | .0000 | .0000 |
| | MC | .020** | .004** | .194 |

Table 7. The covariance matrices (non-redundant elements)
used in the Monte carlo experiment

$$\Sigma_1$$

0.0007773 6.616e-06 -1.082e-05  0.0003573 -0.0001443
0.0024550 0.0001923 -0.0010390 -0.0006195
0.0002950 1.747e-05  0.0002829
0.0007560 0.0004105
0.0006790

$$\Sigma_2$$

0.0007773 1.654e-06 -1.353e-06  3.969e-05 -1.804e-05
0.0024550 2.405e-05 -0.0001737 -7.732e-05
0.0002800 2.427e-05  5.417e-05
0.0001276 2.495e-05
4.863e-05

$$\Sigma_3$$

0.0007773 3.308e-06 -3.607e-06  8.931e-05 -3.608e-05
0.0024550 9.618e-05 -0.0003471 -0.0001238
0.0002836 3.804e-05  0.0001051
0.0001800 7.966e-05
0.0001029

$$\Sigma_4$$

0.0007773 8.271e-07 -1.803e-06  0.0001786 -2.062e-05
0.0024550 2.138e-05 -0.0002083 -0.0002061
0.0002800 1.513e-05  3.485e-05
0.0001707 2.421e-05
5.630e-05

# Bibliography

Barnard, G.A. (1963). *The spectral analysis of point processes*, Discussion on M.S.

Birnbaum, Z.W. (1974). *Computers and unconventional test-statistics*, in F. Proschan and R.J. Serfling (eds.), *Reliability and Biometry*, SIAM, Philadelphia, 441-458.

Breusch, T.S. and A.R. Pagan (1980). *The Lagrange multiplier test and its applications to model specification in econometrics*, Review of Economic Studies 47, 239-254.

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

Dufour, J.M. (1995). *Monte Carlo tests in the presence of nuisance parameters with econometric applications*, working paper, Centre de Recherche et Developpement en Economique CRDE and Département de Sciences Economique, université de Montréal.

Dufour J.-M. and J.F. Kiviet (1994). *Exact inference methods in first order autoregressive distributed lag models*, Discussion paper C.R.D.E., Université de Montréal, and Faculty of Economics and Econometrics, University of Amsterdam.

Dufour, J.M. and L. Khalaf (1996a). *Simulation based finite and large sample inference methods in seemingly unrelated regressions*, working paper, Centre de Recherche et Developpement en Economique CRDE and Département de Sciences Economique, université de Montréal.

Dufour, J.M. and L. Khalaf (1996b). *Simulation based finite and large sample inference methods in simultaneous equations*, working paper, Centre de Recherche et Developpement en Economique CRDE and Département de Sciences Economique, université de Montréal.

Dwass, M. (1957). *Modified randomization tests for non-parametric hypotheses*, Annals of Mathematical Statistics 28, 181-187.

Edgington, E.S. (1980). *Randomization Tests*, Marcel Dekker, New York.

Fisher, S. (1993). *The role of Macroeconomic factors in growth*, Journal of monetary econ., 485-512.

Foutz, R.V. (1980). *A method for constructing exact tests from test statistics that have unknown null distributions*, Journal of Statistical Computation and Simulation 10, 187-193.

Harvey, A.C. and D.A. Phillips (1982). *Testing for contemporaneous correlation of disturbances in systems of regression equations*, Bulletin of Economic Research 34, 79-91.

Jockel K.H. (1986). *Finite sample properties and asymptotic efficiency of Monte carlo tests*, Annals of Statistics 14, 336-347.

Kariya, T. (1981a). *Tests for the independence between two seemingly unrelated regression equations*, The Annals of Statistics 9 (2), 381-390.

Kariya, T. (1981b). *Robustness of multivariate tests*, The Annals of Statistics 6, 1267-1275.

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

Kariya T., Fujikoshi Y. and P.R. Krishnaiah (1984). *Tests for independence between two multivariate regression equations with different design matrices*, Journal of Multivariate Analysis 15, 383-407.

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

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

Marriott, F.H.C. (1979). *Barnard's Monte Carlo tests: how many simulations*, Applied Statistics 28, 75-77.

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

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

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

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

Shiba T. and H. Tsurumi (1988). *Bayesian and non-bayesian tests of independence in seemingly unrelated regressions*, International Economic Review 29, 377-389.

Srivastava , V.K. and Giles D.E. (1987). *Seemingly unrelated regressions equations*, New York: Marcel Deckker

Srivastava , V.K. and T.D. Dwivedi (1979). *Estimation of seemingly unrelated regression equations: a brief survey*, Journal of Econometrics 10, 15-32.

Zellner, A. (1962). *An efficient method of estimating seemingly unrelated regression and tests for aggregation bias*, Journal of the American Statistical Association 57, 348-368.

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

# Chapter 3

# Simulation Based Finite and Large Sample Inference Methods in Simultaneous Equations

Abstract

In the context of multivariate regression (MLR) and simultaneous equations (SE), it is well known that commonly employed asymptotic test criteria are seriously biased towards overrejection. In this paper, we propose finite and large sample likelihood based test procedures for possibly nonlinear hypotheses on the coefficients of SE systems. We discuss a number of Monte Carlo tests and develop an extension of the bootstrap method to statistics whose asymptotic distributions involve nuisance parameters. The latter involves maximizing a randomized $p$-value function over the relevant nuisance parameter space. This is done numerically by using a simulated annealing algorithm. Illustrative Monte Carlo experiments show that: (i) bootstraping 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, LR tests work well in the experiments performed.

## 1. Introduction

Econometricians are often confronted with technical difficulties arising from simultaneity when testing parameter restrictions in systems of equations. With few exceptions, the distributions of standard test statistics are known only asymptotically due to feedback from the dependent variables to the explanatory variables. There will obviously be approximation errors when the asymptotic results are applied to samples of moderate size as is frequently the case in simultaneous equations (SE) applications. Although long recognized as a serious issue in statistical inference, finite sample validity has not received the attention it deserves in such contexts. Indeed, tests of parameter significance have almost invariably been based on asymptotic procedures.

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) proposed exact F-tests for specific classes of hypothesis in the context of a structural equation along with corresponding confidence sets. For a different (although related) problem, Maddala (1974) and Dufour and Jasiak (1996) have described finite sample single-equation procedures which can be viewed as extensions of the latter procedures. Some exact specification tests have also been suggested for SE. In particular, Durbin (1957) proposed a bounds test against serial correlation in SE and, more recently, Harvey and Phillips (1980, 1981a, 1981b, 1989) have suggested tests against serial correlation, heteroscedasticity and structural change in a single structural equation. In both cases, the 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. This procedure generalizes earlier tests suggested by Wu (1973, $T_2$ statistic) and Hausman (1978, eq. 2.23). From a different standpoint, the finite sample distributions of commonly used estimators and test statistics have also received attention in the literature. For a review of the main findings in this area,

the reader may consult Phillips (1983) and Taylor (1983). It is clear from these results that the exact distributions in most cases depend on nuisance parameters. Except for special hypotheses, no work seems presently available that resolves the problem of nuisance parameters in finite samples.

Because of the computational complexity of maximum likelihood methods in SE models, statistical inference has generally been based on instrumental variable (IV) methods. The problems associated with asymptotically valid tests in IV regressions are discussed in Dufour (1996). In particular, it is shown that usual t-type tests, based on common IV estimators, such as two-stage least squares, have significance levels that may deviate arbitrarily from their nominal levels since it is not possible to bound the null distributions of the relevant test statistics to obtain valid inference. This results from identification concerns and is related to the so-called "weak instruments" problem; see, *e.g.* Nelson and Startz (1990a, 1990b), Buse (1992), Maddala and Jeong (1992), Angrist and Krueger (1994), Staiger and Stock (1994), Bound, Jaeger and Baker (1995), Hall, Rudebusch and Wilcox (1996), Cragg and Donald (1996), and Wang and Zivot (1996). For further results relevant to the issue of non-identification, see also Sargan (1983), Phillips (1984, 1985, 1989), Hillier (1990), Choi and Phillips (1992), McManus, Nankervis and Savin (1994)].

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), and Hall and Horowitz (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: Friedman and Peters (1984a), Green et al (1987), Hu et al (1986), Korajczyck (1985) and Dagget and Friedman (1985). Others however, find that the method leads to little improvement, *e.g.* Friedman and Peters (1984b), Park (1985) and Beran and Srivastava (1985). Clearly, there appears to be a conflict in the conclusions regarding the effectiveness of the bootstrap in SE contexts. 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 et

al. (1991), Sriram (1994) and Davidson and MacKinnon (1996)].

This paper addresses these issues and considers alternative Monte Carlo (MC) procedures [Dwass (1957), Barnard (1963), Birnbaum (1974)] for statistical inference in the SE model. In Dufour (1995) the literature on MC tests is reviewed and extensions to nuisance-parameter-dependent statistics are discussed. A randomized procedure termed *maximized Monte Carlo* (MMC) method is specifically proposed which yields provably exact tests, provided the underlying statistics are *boundedly pivotal, i.e.* admit nuisance-parameter-free bounds. Here we consider analogous procedures based on a likelihood framework. Indeed, we show that randomization cannot improve the performance of IV-based tests; given the severity of the problem in the presence of identification difficulties, the case is made here for LR tests rather than IV-based tests. Note that MC tests are closely related to parametric bootstrap tests [see Efron and Tibshirani (1993, Chapter 16), Hall (1992), Hall and Horowitz (1996)], with however a fundamental difference: in MC tests, the number of replications used is explicitly taken into account and can be very small, so that theoretically exact randomized tests can be obtained. For further references regarding MC tests, see Dufour and Kiviet (1994), Edgington (1980), Foutz (1980), Jöckel (1986) and Marriott (1979).

The practical application of LR-based randomized tests is, however, subject to an important consideration: reduced-form tests often involve non-linear hypotheses implied by the structure; in connection, see Bekker and Djikstra (1990) or Byron (1974). Systems tests for nonlinear hypothesis are examined in Dufour and Khalaf (1996), primarily in the context of the multivariate linear regression (MLR) model. The approach used to obtain the tests involves the application of two techniques: bounds tests [similar to those suggested in Dufour (1989, 1990)] and (when required) randomized tests. 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. Here we extend the results in Dufour and Khalaf (1996) to the SE context. After showing how a relevant exact bound can be derived, we use the latter to obtain a conservative test. Next, MMC tests are proposed that can be run whenever the bounds are not conclusive. A multi-equation Anderson-Rubin-type test is also proposed.

We undertake to explore specifically the identification issue in the context of a small simulation experiment. We provide Monte Carlo evidence showing that standard bootstrap methods are unsatisfactory when Wald or quasi-LR statistics are based on IV estimators. More precisely, 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. MMC $p$-values are computed using a simulated annealing (SA) optimization algorithm; for a description of the latter, see Corona et al. (1987) or Goffe and Ferrier (1994).

The paper is organized as follows. Section 2 develops the notation and definitions. Section 3 reviews distributional results from the MLR model, in so far as they are relevant to SE. Section 4 presents test statistics for general hypotheses in the SE framework. Linear hypotheses in the single-equation set-up are considered as a special case. Simulation results are reported in Section 5. Section 6 concludes the paper.

## 2. Framework

We consider a system of $p$ simultaneous equations of the form

(2.1)     $YB + X\Gamma = U$ ,

where $Y = [y_1, \ldots, y_p]$ is an $(n,p)$ matrix of observations on $p$ endogenous variables, $X$ is an $(n,k)$ matrix of fixed (or strictly exogenous) variables and $U = [u_1, \ldots, u_p] = [U_1, \ldots, U_n]'$ 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. Premultiplying both sides by $B^{-1}$ leads to the *reduced form*

(2.2)     $Y = X\Pi + V$ ,   $\Pi = -\Gamma B^{-1}$ ,

or equivalently

(2.3)     $y = (I_p \otimes X) \pi + v$ ,

where $y = vec(Y)$, $\pi = vec(\Pi)$, $v = vec(V)$ and $V = [v_1, \ldots, v_p] = [V_1, \ldots, V_n]'$ is

the matrix of reduced form disturbances. Further, we suppose the rows $U_1'$, ... , $U_n'$ of $U$ satisfy the following distributional assumptions:

(2.4a)　　$U_t = Jw_t$, $t = 1$, ... , $n$,

where the vector $w = vec(w_1, ... , w_n)$ has a known distribution and $J$ is an unknown nonsingular matrix. In particular, this condition will be satisfied when

(2.4b)　　$w_t \sim N(0,I_p)$, $t = 1$, ... , $n$.

More generally, when $U_t$ has finite second moments, its covariance matrix will be $Var(U_t) = JJ' \equiv \Omega$ and the covariance matrix of $V_t$ will be $Var(V_t) = (B^{-1})'\Omega B^{-1} \equiv \Sigma$. Note that the system's unrestricted reduced form (URF) is an MLR model.

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 $\Gamma$. In addition, a normalization constraint is imposed on the endogenous variables coefficients; this 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

(2.5)　　$y_i = Y_i\beta_i + X_{1i}\gamma_{1i} + u_i$, $i = 1, ... , p$,

where $Y_i$ and $X_{1i}$ are $(n, m_i)$ and $(n, k_i)$ matrices which respectively contain the observations on the included endogenous and exogenous variables of the model. If more than $m_i$ variables are excluded from the $i$-th equation, this equation is said to be *over-identified.*

Many problems are formulated in terms of limited-information (LI) models, comprised by a particular structural equation and the reduced form associated with the included right-hand side endogenous variables such as

(2.6)　　$y_i = Y_i\beta_i + X_{1i}\gamma_{1i} + u_i = Z_i\delta_i + u_i$,

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

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

(2.7)　　$[y_i \ Y_i] = X\Pi_i + [v_i \ V_i]$, $\Pi_i = \begin{bmatrix} \pi_{1i} & \Pi_{1i} \\ \pi_{2i} & \Pi_{2i} \end{bmatrix}$, $\pi_{1i} = \Pi_{1i}\beta_i + \gamma_{1i}$, $\pi_{2i} = \Pi_{2i}\beta_i$.

The necessary and sufficient condition for identification follows from the

relation $\pi_{2i} = \Pi_{2i}\beta_i$ : $\beta_i$ is recoverable if and only if

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


## 3. Hypothesis tests in the multivariate linear regression

In this section, we review some important distributional results pertaining to LR test criteria in reduced form contexts. We consider general restrictions on $q^*$ independent linear transformations of $\pi$, the coefficient vector of model (2.3):

(3.1)    $H_{01} : R\pi \in \Delta_0$ ,

where $R$ is a $(q^*, kp)$ matrix such that $rank(R) = q^*$ and $\Delta_0$ is a non-empty subset of $\mathbf{R}^{q^*}$ . This characterization of the hypothesis includes linear restrictions, both within and across equations, and allows for nonlinear as well as inequality constraints. Here, the theory and test procedures detailed in Dufour and Khalaf (1996) are directly applicable. The following exposition will focus on the main issues that are useful for our present purpose.

The LR criterion to test $H_{0l}$ is $n\ln(\Lambda)$, where

(3.2)    $\Lambda = |\hat{\Sigma}_{01}|/|\hat{\Sigma}|$ ,

with $\hat{\Sigma}_{01}$ and $\hat{\Sigma}$ being the restricted and unrestricted ML estimators of $\Sigma$ in (2.3). In the statistics literature, $\Lambda^{-1}$ is often called the Wilks criterion. As argued in Dufour and Khalaf (1996), the null distribution of $\Lambda$ depends on nuisance parameters, yet it is boundedly pivotal. To see the point, consider restrictions of the form

(3.3)    $H_{02}: Q\Pi C = D$ ,

such that $H_{02} \subset H_{01}$ , where $Q$ is a $(q, k)$ matrix of rank $q$ and $C$ is a $(p, c)$ matrix of rank $c$. Linear restrictions that decompose into the latter specific form are called *uniform linear* (UL) in the MLR literature. Let $\Lambda^c(q, c)$ be the reciprocal of the Wilks criterion for testing the latter restrictions. Then the distribution of $\Lambda$ is bounded by the distribution of $\Lambda^c(q, c)$ . Specifically, in Dufour and Khalaf (1996) it is shown that:

(i) the null distribution of the LR statistic for uniform linear hypothesis involves no nuisance parameters and may easily be obtained by simulation;

(ii) under the null, $P[\Lambda \geq \lambda_\alpha^c(q, c)] \leq \alpha$ for all $0 < \alpha < 1$, where $\lambda_\alpha^c(q, c)$ is determined such that $P[\Lambda^c(q, c) \geq \lambda_\alpha^c(q, c)] = \alpha$.

The underlying distributional conditions, namely of the (2.4a) form, are appreciably less restrictive than those of traditional multivariate analysis of variance which require normal errors. Results that correspond to these can be derived in a similar way for the SURE model. The exact critical bound can be obtained by rewriting the test problem in terms of the MLR model of which the SURE system under consideration is a restricted form; see Dufour and Khalaf (1996, section 5) for details.

The fact that the null distribution of the LR statistic can be bounded (in a non trivial way) entails that simulation based techniques (such as Monte Carlo tests) may be effective in this context. In Dufour (1995) the finite and large sample theory underlying Monte Carlo tests in the presence of nuisance parameters is discussed. The methodology involved may be summarized as follows.

Let $T_0$ denote the observed test statistic $T$. Suppose its null distribution depends on the unknown parameter $\theta$. Using Monte Carlo methods, generate $i.i.d$ realizations $T_1$ , ... , $T_N$ of $T_0$ under the null, given $\theta_n$ a consistent estimator of the intervening nuisance parameters, and a specified number $N$ of replications. Rank $T_j$ , $j = 1$ , ... , $N$ in non-decreasing order and obtain $\hat{p}_N(T_0)$ where

$$(3.4) \qquad \hat{p}_N(x) = \frac{N\hat{G}_N(x) + 1}{N + 1} \quad , \quad \hat{G}_N(x) = \frac{1}{N}\sum_{i=1}^{N} I_{[0,\infty)}(T_i - x) \quad , \quad I_A(x) = \begin{matrix} 1 \ , \ if \ x \in A \\ 0 \ , \ if \ x \notin A \end{matrix}$$

Then the test's critical region corresponds to

$$(3.5) \qquad \hat{p}_N(T_0) \leq \alpha \quad , \quad 0 \leq \alpha \leq 1 \quad \cdot$$

Dufour (1995) gives general conditions under which the latter critical region has the correct level asymptotically, $i.e.$ in order to have

$$(3.6) \qquad \lim_{n \to \infty} \left\{ P\left[ \hat{p}_N(T_0|\theta_n) \leq \alpha \right] - P\left[ \hat{p}_N(T_0|\theta) \leq \alpha \right] \right\} = 0 \quad ,$$

for $0 \leq \alpha \leq 1$ and $\plim_{n \to \infty}(\theta_n) = \theta$.

The method just outlined is closely related to a parametric bootstrap, with however a

fundamental difference. In MC tests, the number of replications used is explicitly taken into account and can be very small, so that theoretically exact randomized tests can be obtained; the fact that the procedure is randomized plays a crucial role in determining the size of the test. Here, to obtain an exact critical region, the $p$-value associated with (3.5) ought to be maximized with respect to the elements of the intervening nuisance parameters. Specifically, we have:

$$(3.7) \qquad P\left( \underset{\theta \in M_0}{sup} \left[ \hat{p}_N(T_0|\theta) \right] \leq \alpha \right) \leq \frac{I[\alpha(N + 1)]}{N + 1} , 0 \leq \alpha \leq 1 ,$$

where $I[x]$ is the largest integer less than or equal to $x$ and $M_0$ refers to the nuisance parameter space under the null.

In practical applications of exact MC tests, a global optimization procedure is needed to obtain the maximal randomized $p$-value in (3.7). One such procedure, originally proposed by Corona et al. (1987) and later modified by Goffe et al. (1994) is the simulated annealing (SA) algorithm. SA starts from an initial point, say $\theta_n$ , and sweeps the parameter space at random. An *uphill* step is always accepted while a *downhill* step may be accepted: the decision is made using the Metropolis criterion. The direction of all moves is determined by probabilistic criteria. As it progresses, SA constantly adjusts the step length so that downhill moves are less and less likely to be accepted. In this manner, the algorithm escapes local optima and gradually converges towards the most probable area for optimizing. SA is robust with respect to non-quadratic and even non-continuous surfaces and typically escapes local optima. The procedure is known not to depend on starting values. Most importantly, SA readily handles problems involving a fairly large number of parameters. These procedures are applied in the context of the Monte Carlo experiment reported in Section 5.

## 4. Hypothesis tests in the simultaneous equation model

This section discusses tests on structural parameters in SE models. We first take up LR tests of arbitrary hypotheses based on reduced forms. We then focus on specific procedures such as (i) tests of linear constraints on the coefficients of a single structural equation, and (iii) a generalization of the Anderson-Rubin test to the multi-equation context. Exact simulation-

based bounds on the null distribution of LR statistics are first derived; in an important special case, we also obtain conservative bounds based on the Fisher distribution. Alternative randomized LR tests are next considered. For completeness, we also discuss Wald and quasi-LR IV-based tests. However, a case is made here for LR rather than IV-based criteria.

### 4.1 General hypotheses on structural coefficients based on reduced forms

Consider the problem of testing arbitrary restrictions on the structural parameters of the possibly over-identified model (2.1). Given the transformation that takes the structural system into its reduced form (2.3), namely $\Pi = -\Gamma B^{-1}$, the constraints in question imply nonlinear restrictions on the reduced form parameters. In general, the induced restrictions on $\Pi$ may be expressed as

$$(4.1) \qquad H_{04} : R\pi \in \Delta_0 \ ,$$

where $\pi = vec(\Pi)$, $r = rank(R)$ and $\Delta_0$ is a non-empy subset of $\mathbf{R}^r$.

As it stands, the problem of testing the constraints on structural parameters using the reduced form is exactly the kind of problems discussed in Section 3. Indeed, after expressing the structural restrictions as nonlinear reduced form restrictions, the multivariate procedures given previously are applicable even in case of identification problems. A sound test strategy would be to perform the bounds tests first and, on failure to reject, to apply randomized tests. The methodology can be summarized as follows:

(1) consider the URF and derive the relevant reduced-form restrictions including the exclusion and normalization restrictions implied by the structure;

(2) compute the ratio of the determinant of the maximum likelihood estimates (MLE) of $\Sigma$, imposing and ignoring the restrictions considered;

(3) use the bound $\lambda_\alpha^c(q, c)$ defined in Section 3 as a conservative cut-off point;

(4) whenever the bound is not conclusive, use the MC $p$-value underlying (3.7).

The tests may be applied in a full information, sub-system or single-equation set-ups. Finally, a word about the single-equation tests. As pointed out by Pagan (1979), the LI model involves a triangular system in which case MLE can be derived applying iterated SURE (ITSURE) techniques; on this issue, see also Lahiri and Schmidt (1978). From the practical

point of view, ITSURE have a computational advantage over general nonlinear constrained ML algorithms.

## 4.2 General linear hypotheses in the limited information context

To illustrate how the above results may be used, we consider here the problem of testing linear restrictions in a LI framework. For exposition simplicity, we shall restrict attention to hypotheses that set several structural coefficients to specific values. More precisely, we consider in turn hypotheses of the form:

$$(4.2) \qquad H_0: \beta_i = \beta_i^0 \, ,$$

$$(4.3) \qquad H_0: \beta_{1i} = \beta_{1i}^0 \, ,$$

where $\beta_i = (\beta_{1i}', \beta_{2i}')'$ and $\beta_{1i}$ is $(m_{1i}, 1)$, and

$$(4.4) \qquad H_0: \beta_{1i} = \beta_{1i}^0 \, , \quad \gamma_{12i} = \gamma_{12i}^0 \, ,$$

where $\gamma_{1i} = (\gamma_{11i}', \gamma_{12i}')'$ and $\gamma_{12i}$ is $(k_{2i}, 1)$.

When the model is identified, (4.2) corresponds to the following restrictions

$$(4.5) \qquad \Pi_{2i}\beta_i^0 = \pi_{2i} \, ,$$

or equivalently,

$$(4.6) \qquad \left[ O_{(k-k_i, \, k_i)} \, , \, I_{(k-k_i)} \right] \begin{bmatrix} \pi_{1i} & \Pi_{1i} \\ \pi_{2i} & \Pi_{2i} \end{bmatrix} \begin{bmatrix} 1 \\ -\beta_i^0 \end{bmatrix} = 0 \, ,$$

where $O_{(s, \, j)}$ denotes a zero $(s,j)$ matrix. Let $\hat{\Sigma}_0$ and $\hat{\Sigma}$ be the error covariance LIML estimates imposing and ignoring (4.5), where the latter corresponds to the unrestricted reduced form. Further, let $\hat{\Sigma}_{LI}$ denote the LIML error covariance estimate imposing the exclusion restrictions implied by the structure. Conformably with the notation set above, define

$$(4.7) \qquad \Lambda_{LI} = |\hat{\Sigma}_0| \, / \, |\hat{\Sigma}_{LI}| \, ,$$

$$(4.8) \qquad \Lambda^c(k-k_i, \, 1) = |\hat{\Sigma}_0| \, / \, |\hat{\Sigma}| \, .$$

Following the arguments of Section 3, we see that the distribution of $\Lambda_{LI}$ is bounded by the distribution of $\Lambda^c(k-k_i, \, 1)$.

The LI LR statistic $(LR_{LI})$ may be obtained as $n\ln(\Lambda_{LI})$. Whereas $n[\ln(\Lambda_{LI})]$ has a $\chi^2(m_i)$ asymptotic distribution only under identification assumptions, $n[\ln(\Lambda^c(k-k_i, \, 1))]$

is asymptotically distributed as $\chi^2(k-k_i)$ whether the rank condition holds or not. The asymptotic distribution of the LI LR 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 Nelson et al (1996) and Wang and Zivot (1996) for the special case where $\beta_i$ is scalar. Furthermore, exact bounds based on the $F(k-k_i, n-k)$ distribution may also be derived for this problem if the normality assumption (2.4b) holds. Indeed, as pointed out in Stewart (1995),

$$(4.9) \qquad q[\Lambda^c(q,1) - 1] / (n-k) \sim F(q, n-k) \; ,$$

where $\Lambda^c(q, 1)$ is the reciprocal of the Wilks statistics for testing uniform linear restrictions of the form $Q\Pi C = D$ , $rank(Q) = q$ , $rank(C) = c = 1$ .

The important thing to note regarding the latter bound is that it relates to the well known Anderson-Rubin (AR) statistic. Bartlett (1948) and Anderson and Rubin (1949) suggested an exact test that can be applied only if the null takes the (4.2) form. The idea behind the test is quite simple. Define $y_i^* = y_i - Y_i\beta_i^0$ . Under the null, the model can be written as $y_i^* = X_{1i}\gamma_{1i} + u_i$ . On the other hand, if the hypothesis is not true, $y_i^*$ will be a linear function of all the exogenous variables. Thus, the null may be assessed by the $F$ test that the coefficient of the "excluded" regressors is zero in the regression of $y_i^*$ on all the exogenous variables. It is straightforward to show using the results on UL hypotheses in Dufour and Khalaf (1996) and Stewart (1995) that the AR statistic associated with $H_0$: $\beta_i = \beta_i^0$ corresponds to a monotonic transformation of the LI LR criterion for testing the UL hypothesis $\Pi_{2i}\beta_i^0 = \pi_{2i}$ against an unrestricted alternative.

Let us now consider the hypothesis (4.3). On partitioning $\Pi_{1i} = [\Pi_{11i}, \Pi_{12i}]$ and $\Pi_{2i} = [\Pi_{21i}, \Pi_{22i}]$ conformably with $\beta_i = (\beta_{1i}', \beta_{2i}')'$ the corresponding reduced form restrictions may be expressed as

$$(4.10) \qquad \left[ O_{(k-k_i, \, k_i)} \, , \, I_{(k-k_i)} \right] \begin{bmatrix} \pi_{1i} & \Pi_{11i} & \Pi_{12i} \\ \pi_{2i} & \Pi_{21i} & \Pi_{22i} \end{bmatrix} \begin{bmatrix} 1 \\ -\beta_{1i}^0 \\ -\beta_{2i} \end{bmatrix} = 0 \; .$$

Let $\Lambda_{LI}$ be the reciprocal of the Wilks statistic for testing (4.10) against the restrictions implied by the structure. The nonlinearities in connection with (4.10) stem from the fact

that $\beta_{2i}$ is unknown. However, the special case of (4.10) that corresponds to specific (unknown) values of $\beta_{2i}$ takes the UL form. Let $\Lambda_a^c(k-k_i$ , 1) the reciprocal of the Wilks statistic for testing these UL restrictions against an unrestricted alternative. Then conservative bounds for $\Lambda_U$ can be obtained form the statistic $\Lambda^c(k-k_i$ , 1) or the $F(k-k_i$ , $n-k)$ when applicable. Similar results hold under (4.4). In this case, the implied reduced form constraints are

$$(4.11) \qquad \left[ O_{(k-(k_i-k_{2i}) \, , \, k_i-k_{2i})} \, , \, I_{(k-(k_i-k_{2i}))} \right] \begin{bmatrix} \pi_{1i} & \Pi_{11i} & \Pi_{12i} \\ \pi_{2i} & \Pi_{21i} & \Pi_{22i} \end{bmatrix} \begin{bmatrix} 1 \\ -\beta_{1i}^0 \\ -\beta_{2i} \end{bmatrix} = \begin{bmatrix} 0 \\ \gamma_{12i}^0 \\ 0 \end{bmatrix} .$$

Thus, conservative bounds for the associated $\Lambda_U$ can be obtained from the statistic $\Lambda^c((k-(k_i-k_{2i}))$ , 1) corresponding to the special case of (4.11) where $\beta_{2i}$ is known, or the $F((k-(k_i-k_{2i}))$ , $n-k)$ when applicable, as previously shown.

In closing this section, we review the well-known IV analogue of the Wald test for the hypothesis $R\delta_i = r$, where $rank(R) = q$. Such tests have usually been derived for cases in which the maximum likelihood estimation is thought, on grounds of computational cost, to be impractical. For instance, consider the two-stage least squares (2SLS) estimator

$$(4.12) \qquad \hat{\delta}_i = [Z_i'P(P'P)^{-1}P'Z_i]^{-1} \, Z_i'P(P'P)^{-1}P'y_i \, ,$$

where $P$ is the following matrix of instruments $P = [X_i \, , \, X(X'X)^{-1}X'Y_i]$. Application of the Wald principle yields the following criterion

$$(4.13) \qquad \tau_w = \frac{1}{s^2} \, ( \, r \, - \, R\hat{\delta}_i \, )' \, [ \, R' \, (Z_i'P(P'P)^{-1}P'Z_i)^{-1} \, R \, ]^{-1} \, ( \, r \, - \, R\hat{\delta}_i \, ) \, ,$$

where

$$(4.14) \qquad s^2 = \frac{1}{n} \, ( \, y_i \, - \, Z_i\hat{\delta}_i \, )' \, ( \, y_i \, - \, Z_i \, \hat{\delta}_i)' \, .$$

Under usual regularity conditions and imposing identification, $\tau_w$ is distributed like a $\chi^2(q)$ variable. Despite the widespread recognition of the need for caution in the application of IV-based tests, the standard econometric software packages typically implement IV-based Wald tests. In particular, the $t$-tests on individual parameters are routinely computed in the context of 2SLS or 3SLS procedures. Unfortunately, the Monte Carlo experiments we report in Section 5 confirm that IV-based tests realize computational savings at the risk of very poor performance.

*4.3 A multi-equation Anderson-Rubin type test*

The AR test has recently received renewed interest. See, for example, Revankar and Mallela (1972), Maddala (1974), Staiger and Stock (1994), Dufour and Jasiak (1996), Nelson et al (1996) and Wang and Zivot (1996). However, as it stands, the AR test ignores any restrictions relating to equations other than the $i$th. Here, we discuss an extension to the multiple equation framework.

Consider, in the context of (2.5) hypotheses of the form

(4.15)  $H_0: \beta_i = \beta_i^0$ , $i = 1, ..., p$ .

Now define $Y^* = [\ y_1^*, ... , y_p^*\ ]$ , where $y_i^* = y_i - Y_i\beta_i^0$ , $i = 1, ..., p$ . Under the null, the system of equations corresponds to the SURE model

(4.16)  $y_i^* = X_{1i}\gamma_{1i} + u_i$ , $i = 1, ..., p$ ,

whereas under the alternative the relevant specification is the MLR model including all the exogenous variables. Thus the problem reduces to testing the underlying SURE exclusion restrictions. Since the test involves the coefficients of different regressors within a MLR model, an exact critical value is not available. Nevertheless, the tests described in Section 3 are applicable and lead to valid inference. In addition, the test can be readily extended to accommodate additional constraints on the exogenous variables coefficients. Maddala (1974) treats the single equation case. Specifically, consider hypothesis of the form

(4.17)  $\beta_i = \beta_i^0$ , $\gamma_{11i} = \gamma_{11i}^0$ ,

where $\gamma_{11i}$ is a subset of $\gamma_{1i}$ . Partition the matrix $X_{1i}$ accordingly and let

(4.18)  $y_i^{**} = y_i - Y_i\beta_i^0 - X_{11i}\gamma_{11i}$ , $i = 1, ..., p$ .

Then the restricted model becomes the SURE system

(4.19)  $y_i^{**} = X_{12i}\gamma_{12i} + u_i$ , $i = 1, ..., p$ ,

and the test may be carried out as above. Note that the tests are also applicable in a sub-system framework. However, as with the single equation AR test, the requirement is that all structural coefficients pertaining to the right-hand-side endogenous variables be specified under the null.

## 5. A simulation study

This section reports an investigation, by simulation, of the performance of the various proposed test procedures. All the experiments were conducted using Gauss-386i VM version 3.2 and each was based on 1000 replications.

### 5.1 Monte Carlo design

The experiments were based on the LI model (2.6). The system involves three endogenous variables ($p = 3$) and $k = 3,4,5$ and 6 exogenous variables. The different models are denoted (1) to (4) corresponding to $k = 3$ to 6. In all cases, the structural equation includes only one exogenous variable, the constant regressor. To simplify the exposition, we henceforth drop the subscript $i$ when referring to the equation under consideration. The restrictions tested were of the form

$$(5.1) \qquad H_0: \beta_1 = \beta_1^0 \ ,$$

where $\beta = (\beta_1 , \beta_2)'$. The sample sizes were set to $n = 25$ and $n = 100$ (the latter in certain cases only). The exogenous regressors were 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

$$(5.2) \qquad \Sigma = \begin{bmatrix} 1 & .95 & -.95 \\ .95 & 1 & -1.91 \\ -.95 & -1.91 & 12 \end{bmatrix} .$$

The other coefficients were

$$(5.3) \qquad \gamma_1 = 1, \ \beta_1 = 10, \ \beta_2 = -1.5, \ \Pi_1 = (1.5 , 2)', \ \Pi_2 = \begin{bmatrix} \bar{\Pi} \\ O_{(k-3, 2)} \end{bmatrix} ,$$

with

$$(5.4a) \qquad \bar{\Pi} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} , \ \text{or} \qquad (5.4b) \qquad \bar{\Pi} = \begin{bmatrix} 2 & 1.999 \\ 1.999 & 2 \end{bmatrix} ,$$

and $O_{(s,j)}$ is a zero $(s,j)$ matrix. The identification problem becomes more serious as the determinant of $\Pi_2'\Pi_2$ gets closer to zero. In view of this, we also consider:

$$(5.5a) \ \bar{\Pi} = \begin{bmatrix} .5 & .499 \\ .499 & .5 \end{bmatrix} , \quad (5.5b) \ \bar{\Pi} = \begin{bmatrix} .4 & .399 \\ .399 & .4 \end{bmatrix} , \quad (5.5c) \ \bar{\Pi} = \begin{bmatrix} .3 & .299 \\ .299 & .3 \end{bmatrix} ,$$

$$(5.5d) \quad \tilde{\Pi} = \begin{bmatrix} .2 & .199 \\ .199 & .2 \end{bmatrix}, \quad (5.5e) \quad \tilde{\Pi} = \begin{bmatrix} .1 & .099 \\ .099 & .1 \end{bmatrix}, \quad (5.5f) \quad \tilde{\Pi} = \begin{bmatrix} .01 & .009 \\ .009 & .01 \end{bmatrix}.$$

The Wald and LR statistics were calculated as defined in Section 4 and denoted $W$ and $LR$. In each case, we report the probability of Type I error for asymptotic (ASY), Monte Carlo (MC), maximized Monte Carlo (MMC) and bounds (BND) procedures. The standard asymptotic $\chi^2(1)$ critical value was adopted. The MC and MMC critical points were obtained by simulation following the lines of Section 3; the Tsionas (1995) SA algorithm was implemented to obtain the maximal p-values. The randomized tests were applied with $N = 19$ replications. Recall that MC $p$-values require consistent estimates of the intervening nuisance parameters. Here we use restricted 2SLS estimates in the case of the Wald test and the LIML estimates for the LR tests. Tables 1-3 summarize our findings.

## 5.2 *Results*

Although the Monte Carlo experiments were conditional on the selected design, 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 a standard bootstrap is disappointing. For both LR and Wald criteria, the empirical sizes of MC 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, the failure of the traditional bootstrap LR test is worth emphasizing.

(4) In all cases, the Wald tests maximal randomized $p$-values are always one. Another Monte Carlo experiment (not reported here) confirms similar results in the context of a quasi-LR statistic based on derived 2SLS reduced form estimates.

(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.[1]

The above findings mean that 2SLS-based tests are inappropriate in the weak instrument case and cannot be corrected by bootstraping. 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.

## 6. 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 completely 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, MC LR tests perfectly controlled levels. The exact

---

[1] The average recorded time-to-convergence using a Pentium 135 was approximately 3 minutes for all models reported in Tables 1 and 2.

randomized procedures are computer intensive; however, with modern computer facilities, computational costs are longer a hinderance.

Table 1. Empirical sizes of various asymptotic and Monte Carlo tests

| | | $\bar{\Pi} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$ | | $\bar{\Pi} = \begin{bmatrix} 2 & 1.999 \\ 1.999 & 2 \end{bmatrix}$ | |
|---|---|---|---|---|---|
| | | LR | WALD | LR | WALD |
| Model (1) [n = 25] | ASY | .057 | .048 | .055 | .023 |
| | MC | .049 | .044 | .060 | .023 |
| | MMC | .016 | .000* | .025 | .000 |
| | BND | .006 | - | .011 | - |
| Model (2) [n = 25] | ASY | .075 | .086 | .142 | .082 |
| | MC | .053 | .058 | .071 | .053 |
| | MMC | .017 | .000* | .027 | .000* |
| | BND | .007 | - | .013 | - |
| Model (3) [n = 25] | ASY | .102 | .110 | .200 | .126 |
| | MC | .059 | .068 | .079 | .059 |
| | MMC | .006 | .000* | .015 | .000* |
| | BND | .003 | - | .013 | - |
| Model (4) [n = 25] | ASY | .149 | .142 | .223 | .147 |
| | MC | .068 | .085 | .090 | .073 |
| | MMC | .012 | .000* | .024 | .000* |
| | BND | .001 | - | .011 | - |
| Model (4) [n = 100] | ASY | .097 | .081 | .210 | .116 |
| | MC | .053 | .061 | .064 | .051 |
| | MMC | .001 | .000* | .020 | .000* |
| | BND | .000 | - | .009 | |

Note: The subscript (*) indicates that the MMC p-value is *one* throughout the 1000 replications.

Table 2. Empirical sizes of various asymptotic and Monte Carlo tests
Near-unidentified conditions (n = 25)

| | | Model (2) | | Model (4) | |
|---|---|---|---|---|---|
| | | LR | WALD | LR | WALD |
| $\bar{\Pi} = \begin{bmatrix} .5 & .499 \\ .499 & .5 \end{bmatrix}$ | ASY | .024 | .109 | .064 | .226 |
| | MC | .021 | .058 | .038 | .102 |
| | MMC | .002 | .000* | .001 | .000* |
| | BND | .001 | - | .000 | - |
| $\bar{\Pi} = \begin{bmatrix} .4 & .399 \\ .399 & .4 \end{bmatrix}$ | ASY | .020 | .124 | .051 | .264 |
| | MC | .014 | .066 | .030 | .115 |
| | MMC | .002 | .000* | .001 | .000* |
| | BND | .000 | - | .000 | - |
| $\bar{\Pi} = \begin{bmatrix} .3 & .299 \\ .299 & .3 \end{bmatrix}$ | ASY | .017 | .156 | .041 | .345 |
| | MC | .014 | .080 | .026 | .140 |
| | MMC | .002 | .000* | .000 | .000* |
| | BND | .000 | - | .000 | - |
| $\bar{\Pi} = \begin{bmatrix} .2 & .199 \\ .199 & .2 \end{bmatrix}$ | ASY | .009 | .245 | .031 | .532 |
| | MC | .010 | .105 | .020 | .210 |
| | MMC | .001 | .000* | .000 | .000* |
| | BND | .000 | - | .000 | - |
| $\bar{\Pi} = \begin{bmatrix} .1 & .099 \\ .099 & .1 \end{bmatrix}$ | ASY | .006 | .507 | .032 | .823 |
| | MC | .012 | .212 | .019 | .444 |
| | MMC | .000 | .000* | .000 | .000* |
| | BND | .000 | - | .000 | - |
| $\bar{\Pi} = \begin{bmatrix} .01 & .009 \\ .009 & .01 \end{bmatrix}$ | ASY | .021 | .889 | .090 | .993 |
| | MC | .013 | .579 | .026 | .823 |
| | MMC | .001 | .000* | .004 | .000* |
| | BND | .000 | - | .000 | - |

Note: The subscript (*) indicates that the MMC p-value is *one* throughout the 1000 replications.

Table 3. Empirical sizes of Wald-type tests
Near-unidentified conditions (n = 100)

| Model (4) | Wald test, n=100 | | Model (4) | Wald test, n=100 | |
|---|---|---|---|---|---|
| $\bar{\Pi} = \begin{bmatrix} .5 & .499 \\ .499 & .5 \end{bmatrix}$ | ASY | .143 | $\bar{\Pi} = \begin{bmatrix} .1 & .099 \\ .099 & .1 \end{bmatrix}$ | ASY | .227 |
| | MC | .070 | | MC | .106 |
| | MMC | .000* | | MMC | .000* |
| $\bar{\Pi} = \begin{bmatrix} .4 & .399 \\ .399 & .4 \end{bmatrix}$ | ASY | .162 | $\bar{\Pi} = \begin{bmatrix} .1 & .099 \\ .099 & .1 \end{bmatrix}$ | ASY | .581 |
| | MC | .074 | | MC | .212 |
| | MMC | .000* | | MMC | .000* |
| $\bar{\Pi} = \begin{bmatrix} .3 & .299 \\ .299 & .3 \end{bmatrix}$ | ASY | .191 | $\bar{\Pi} = \begin{bmatrix} .01 & .009 \\ .009 & .01 \end{bmatrix}$ | ASY | .989 |
| | MC | .083 | | MC | .700 |
| | MMC | .000* | | MMC | .000* |

Note: The subscript (*) indicates that the MMC p-value is *one* throughout the 1000 replications.

# Bibiliography

Anderson, T.W. and H. Rubin (1949). *Estimation of the parameters of a single equation in a complete system of stochastic equations*, Annals of Mathematical Stat. 20, 46-63.

Angrist, J.D. and A.B. Krueger (1994). *Split sample instrumental variables*, Technical working Paper No. 150, NBER, Cambridge MA.

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

Barnard, G.A. (1963). *The spectral analysis of point processes*, Discussion on M.S. Bartlett, Journal of the Royal Statistical Society 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., A.K. Mallik, W.P. McCormick, W.P. Reeves and R.L. Reeves (1991), *Bootstrapping unstable first-order autoregressive processes*, Annals of Statistics 19, 1098-1101.

Bekker P.A. and T.K. Djikastra (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 M. Srivastava (1985). *Bootsrap tests and confidence regions for functions of a covariance matrix*, Annals of statistics 13, 95-115.

Birnbaum, Z.W. (1974). *Computers and unconventional test-statistics*, in F. Proschan and R.J. Serfling (eds.), *Reliability and Biometry*, SIAM, Philadelphia, 441-458.

Bound, J., D.A. Jaeger and R.M. Baker (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 biais of instrumental variables estimators*, Econometrica 60, 173-180.

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

Corona, A., Marchesi M., Martini C. and S. Ridella (1987). *Minimizing multimodal functions*

*of continuous variables with the 'simulated annealing algorithm'*, ACM Transactions on Mathematical Software 13, 262-280.

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

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

Dagget, R.S. and D.A. Freedman (1985). *Econometrics and the law: a case study in the proof of antitrust damages*, in: L.m. Le Cam and R.A. Olshen, eds, Proc. Berkeley Conf. in Honor of Jerzy Neyman and Jack Kiefer, 1, Wadsworth, Belmont, CA.

Davidson, R. and J.G. MacKinnon (1996). *The size distortion of Bootstrap tests*, Discussion paper, department of Economics, Queen's University.

Dufour, J.M. (1996). *Some impossibility theorems in econometrics with applications to instrumental variables*, dynamic models and cointegration, cahier 3695, Centre de recherche et développement en économique (C.R.D.E.) and Département de sciences économiques, Université de Montréal.

Dufour, J.M. (1987). *Linear Wald methods for inference on covariance matrices and weak exogeneity tests in structural equations*, in *Time Series and Econometric Modelling*, edited by I.B. MacNeil and G.J. Umphrey, D. Reidel Publishing Company (Dordrecht, Holland),1987, 317-338.

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

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

Dufour, J.M. (1995). *Monte Carlo tests in the presence of nuisance parameters, with econometric applications*, working paper, Centre de recherche et développement en économique (C.R.D.E.) and Département de sciences économiques, Université de Montréal.

Dufour, J.M. (1996). *Some impossibility theorems in econometrics with applications to structural and dynamic models*, Econometrica, forthcoming.

Dufour, J.M. and J. Jasiak (1996). *Finite sample inference methods for simultaneous equations and models with unobserved and generated regressors*, working paper, Centre de recherche et développement en économique (C.R.D.E.) and Département de sciences économiques, Université de Montréal.

Dufour J.-M. and J.F. Kiviet (1994). *Exact inference methods in first order autoregressive distributed lag models*, Discussion paper C.R.D.E., Université de Montréal, and Faculty of Economics and Econometrics, University of Amsterdam.

Dufour, J.M. and L. Khalaf (1996). *Simulation based finite and large sample inference methods in seemingly unrelated regressions*, working paper, Centre de recherche et développement en économique (C.R.D.E.) and Département de sciences économiques, 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 non-parametric hypotheses*, Annals of Mathematical Statistics 28, 181-187.

Edgington, E.S. (1980). *Randomization Tests*, Marcel Dekker, New York.

Efron, B. (1982). *The jackknife, the bootstrap and other resampling plans*. CBS-NSF Regional conference series in applied mathematics, Monograph No. 38, Philadelphia, PA: Society for Industrial and Applied Mathematics.

Efron, B. and R.J. Tibshirani (1993). *An introduction to the bootstrap*, New York: Capman and Hall.

Foutz, R.V. (1980). *A method for constructing exact tests from test statistics that have unknown null distributions*, Journal of Statistical Computation and Simulation 10, 187-193.

Freedman, D.A. and S.C. Peters (1984a). *Bootstraping a regression equation: some empirical results*, Journal of the American Statistical Association, 79 (385), 97-106.

Freedman, D.A. and S.C. Peters (1984b). *Bootstraping an econometric model: some empirical*

*results*, Journal of Business and Economics Statistics 2 (2), 150-158.

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

Green, R., Hahn, W. and D. Rocke (1987). *Standard errors for elasticities: a comparison of bootsrap and asymptotic standard errors*, Journal of Business and economics Statistics 5, 145-149.

Haavelmo, T. (1947). *Methods of measuring the marginal propensity to consume*, Journal of the American Statistical Association 42, 105-122.

Hall, P. (1992). *The bootstrap and Edgeworth expansion*, New York: Springer-Verlag.

Hall A.R., G.D. Rudebusch, and D.W. Wilcox (1996). *Judging instrument relevance in instrumental variables estimation*,

Hall, P. and J. Horowitz (1996). *Bootstrap critical values for tests based on generalized-method-of-moments estimators*, Econometrica 64, 891-916.

Harvey, A.C. and G.D.A. Phillips (1989). *Testing for structural change in simultaneous equation models*, in: Hackl, Peter ed., *Statistical analysis and forecasting of economic structural change*, Springer-Verlag, 25-36.

Harvey, A.C. and G.D.A. Phillips (1981a). *Testing for heteroscedasticity in simultaneous equation models*, Journal of Econometrics 15, 311-340.

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

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

Hauseman, J. (1978). *Specification tests in econometrics*, Econometrica 46, 1251-1272.

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

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

Jeong, J. and G.S. Maddala (1993). *A perspective on application of bootstrap methods in*

*econometrics*, in: G.S. Maddala, C.R. Rao and H.D. Vinod, eds, *Handbook of statistics*, 11, Elsevier Science publishers B.V., 573-610.

Korajczyk, R. (1985). *The pricing of forward contracts for foreign exchange*, Journal of Politic. Econom. 93, 346-368.

Lahiri K. and P. Schmidt (1978). *On the estimation of triangular structural systems*, Econometrica 46, 1217-1221.

Maddala, G.S. (1974). *Small sample evidence on tests of significance in simultaneous equations models*, Econometrica 42 (5), 841-851.

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

Marriott, F.H.C. (1979). *Barnard's Monte Carlo tests: how many simulations*, Applied Statistics 28, 75-77.

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

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

Nelson, C.R. and R. Startz (1990b). *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., R. Startz and E. Zivot (1996). *Valid confidence intervals and inference in the presence of weak instruments*, Discussion paper, department of Economics, University of Washington.

Pagan A. (1979). *Some consequences of viewing LIML as iterated Aitken estimator*, Economic Letters 3, 369-372.

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

Phillips, P.C.B. (1983). *Exact small sample theory in the simultaneous equations model*, Handbook of Econometrics (I), 451-561.

Phillips, P.C.B. (1984). *The exact distribution of LIML:I*, International Economic Review 25, 249-261.

Phillips, P.C.B. (1985). *The exact distribution of LIML:II*, International Economic Review 26, 21-36.

Phillips, P.C.B. (1989). *Partially identified econometric models*, Econometrica 39, 577-591.

Revankar, N. and P. Mallela (1972). *The power of an F-test in the context of a structural equation*, Econometrica 16, 323-329.

Sargan J.D. (1983). *Identification and lack of identification*, Econometrica 51, 1605-1633.

Sriram, T.N. (1994). *Invalidity of Bootstrap for critical branching processes with immigration*, Annals of Statistics 22, 1013-1023.

Staiger, D. and J.H. Stock (1994). *Instrumental Variables regression with weak instruments*, Technical working Paper No. 151, NBER, Cambridge, MA.

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: G.S. Maddala, C.R. Rao and H.D. Vinod, eds, *Handbook of statistics*, 11, Elsevier Science publishers B.V., 629-661.

Wang J. and E. Zivot (1996). *Inference on a structural parameters in instrumental variable regression with weakly correlated instruments*, Discussion paper, department of Economics, University of Washington.

Wu, D.M. (1973). *Alternative tests of independence between stochastic regressors and disturbances: finite sample results*, Econometrica 42, 529-546.

# Conclusion générale

Les méthodes d'inférence basées sur des simulations ont connu ces dernières années un développement important. Ceci est dû en grande partie aux progrès actuels de l'informatique. Les stratégies de tests simulés décrites dans les chapitres précédents ont pour objectif de contrôler les erreurs de première espèce sur des échantillons finis. Ce problème revêt une importance considérable dans le cadre multivarié car les tests couramment utilisés sont fondés sur des arguments asymptotiques et sont généralement sujets à des distorsions de niveau importantes. De plus, les distributions exactes des statistiques intervenant dans les procédures usuelles sont compliquées et dépendent de paramètres inconnus. Mais avec l'application des tests de Monte Carlo, ces problèmes sont pratiquement résolus. Or ces méthodes seront effectives lorsqu'on a spécifié un critère de test dont la distribution sous l'hypothèse nulle admet une borne qui ne dépend pas de paramètres de nuisance. La contribution fondamentale de ce rapport est de démontrer cette propriété dans le cas du critère LR sous des hypothèses générales, linéaires ou non-linéaires portant sur les coefficients d'un modèle linéaire multivarié. Nous en déduisons (i) un test à bornes, et (ii) la validité des procédures Monte Carlo fondées sur la vraisemblance. Sur le plan de la facilité de calcul, le test à bornes est particulièrement utile; nous suggérons donc une stratégie de test par étapes qui consiste à recourir aux procédures de Monte Carlo lorsque la borne n'est pas concluante.

Les résultats de nombreuses expériences suggèrent que les procédures proposées contrôlent effectivement le niveau et ont une bonne puissance. Par ailleurs, nous vérifions qu'il suffit d'un faible nombre de simulations pour obtenir des résultats très satisfaisants. Ce type de propriété rend les méthodes de tests simulés appliquées au quotient de vraissemblance fort prometteuses. En revanche, nous démontrons pour le cas des équations simulatnées, que l'application mécanique de tests utilisant des variables instrumentales (*e.g.* les statistiques de

*Student* fournies par les logiciels usuels) peut entraîner de sérieuses erreurs. Ce résultat provient des problèmes qui se rattachent à l'identification. Nos simulations permettent aussi de souligner le manque de fiabilité du bootstrap dans le contexte des équations simultanées.

La popularité des tests usuels basés sur des régressions instrumentales pourrait être attribuée au fait que les procédures alternatives fondées sur la vraisemblance sont plus lourdes à appliquer. Certes, les tests du quotient de vraisemblance simulé que nous proposons dans cette thèse sont intensives en temps de calcul. Mais avec les progrès rapides de l'informatique, l'exécution de ces méthodes est considérablement simplifiée. D'ailleurs, tout compte fait, nous soulignons que les coûts de calcul ne sont que le prix à payer pour une inférence valide.