

A Characterization of Invariant Tests for Identification in Linear Structural Equations

Giovanni Forchini^{1,2}

Monash University

November 6, 2006

Abstract

We derive the maximal invariant for tests of identification in a linear structural equation model. We find that the maximal invariant has three components that are the basis for existing tests for rank, over-identification and exogeneity.

Introduction

The rapidly growing literature on weak instruments has stressed the importance of reporting tests for possible failures of identification in applied research (e.g. Staiger and Stock (1997)). These have either the form of tests for the rank of some reduced form coefficients (e.g. Cragg and Donald (1993), Cragg and Donald (1997), Robin and Smith (2000), Stock and Yogo (2001)) or tests for over-identifying restrictions (e.g. Sargan (1958), Basmann (1960a), Basmann (1960b), Byron (1974), Hansen (1982)).

In this note, we study tests for identification in a linear structural equation model. The structure of the model suggests invariance arguments that can reduce the testing problem to functions of a maximal invariant (i.e. a few statistics of considerably smaller dimension than the sample size) and simplify the analysis. We first identify the transformations that leave the problem invariant, and then derive a maximal invariant.

¹ Address for correspondence: Giovanni Forchini, Department of Econometrics and Business Statistics, Monash University, Clayton, Victoria 3800, Australia. E-mail: Giovanni.Forchini@BusEco.monash.edu.au

² I thank Don Poskitt and an anonymous referee for useful and encouraging comments.

The maximal invariant has three components, commonly used in tests of rank, over-identifying restrictions and exogeneity.

Main result

We consider a linear structural equation

$$y = Y_2 \beta + Z_1 \gamma + u, \quad (1)$$

$(T \times 1)$ $(T \times n)$ $(n \times 1)$ $(T \times k_1)$ $(k_1 \times 1)$

with corresponding reduced form

$$[y_1, Y_2] = Z_1 [\phi_1, \Phi_2] + Z_2 [\pi_1, \Pi_2] + [v_1, V_2]. \quad (2)$$

$(T \times k_2)$

The dimensions of vectors and matrices are reported in square brackets the first time they are used, unless they are obvious from the context. We assume that $[v_1, V_2] | [Z_1, Z_2] \sim N(0, I_T \otimes \Omega)$, where Ω is an $(n+1 \times n+1)$ matrix of parameters. We are implicitly assuming that the reduced form (2) contains information on (functions) of the parameters $[\phi_1, \Phi_2]$, $[\pi_1, \Pi_2]$ and Ω only. Alternative specifications for the structural equation exist (e.g. Hillier (1990)). However, the one in equation (1) is the most commonly used in applied and theoretical research, and is therefore of particular interest.

Compatibility of the structural equation and the reduced form requires

$$\pi_1 = \Pi_2 \beta \quad (3)$$

$$\phi_1 = \Phi_2 \beta + \gamma \quad (4)$$

$$v_1 = V_2 \beta + u. \quad (5)$$

The vector of coefficients of the endogenous variables β is identified if it is uniquely determined as a map from the set of all joint densities of $[y_1, Y_2]$ conditional on $[Z_1, Z_2]$ specified by the reduced form (2) to the n dimensional Euclidean space. This requires equation (3) to hold and the rank of Π_2 to be n . With no extra knowledge about Φ_2 , the structural parameter γ is identified by (4) provided β is identified (see Phillips (1989) for further discussion).

Joint sufficient statistics for $[\phi_1, \Phi_2]$, $[\pi_1, \Pi_2]$ and Ω are

$$\left[\tilde{\pi}_1, \tilde{\Pi}_2 \right] = \left(Z_2' M_{Z_1} Z_2 \right)^{-1/2} Z_2' M_{Z_1} [y_1, Y_2], \quad (6)$$

$$\left[\tilde{\phi}_1, \tilde{\Phi}_2 \right] = (Z_1' Z_1)^{-1/2} Z_1' [y_1, Y_2] \quad (7)$$

and

$$S = \begin{bmatrix} s_{11} & s_{21}' \\ (1 \times 1) & \\ s_{21} & s_{22} \\ (n \times 1) & (n \times n) \end{bmatrix} = [y_1, Y_2]' M_{[Z_1, Z_2]} [y_1, Y_2], \quad (8)$$

where $P_A = A(A'A)^{-1}A'$ and $M_A = I_T - P_A$ for any full column rank matrix A , and $Z = [Z_1, Z_2]$. These satisfy

$$\left[\tilde{\pi}_1, \tilde{\Pi}_2 \right] \sim N \left(\left(Z_2' M_{Z_1} Z_2 \right)^{1/2} [\pi_1, \Pi_2], I_{k_2} \otimes \Omega \right), \quad (9)$$

$$\left[\tilde{\phi}_1, \tilde{\Phi}_2 \right] \sim N \left((Z_1' Z_1)^{1/2} [\phi_1, \Phi_2] + (Z_1' Z_1)^{-1/2} Z_1' Z_2 \Pi_2 [\beta_2, I_n], I_{k_1} \otimes \Omega \right), \quad (10)$$

and

$$S \sim W_{n+1}(T - k_1 - k_2, \Omega). \quad (11)$$

Moreover, $\left[\tilde{\pi}_1, \tilde{\Pi}_2 \right]$, $\left[\tilde{\phi}_1, \tilde{\Phi}_2 \right]$ and S are independently distributed. Notice that there is a one-to-one relationship between $\left[\tilde{\pi}_1, \tilde{\Pi}_2 \right]$, $\left[\tilde{\phi}_1, \tilde{\Phi}_2 \right]$ and the OLS estimators of the reduced form parameters. Let $\left[\bar{\pi}_1, \bar{\Pi}_2 \right] = \left(Z_2' M_{Z_1} Z_2 \right)^{1/2} [\pi_1, \Pi_2]$. The following result can be easily established.

Lemma 1. (i) The rank of $\bar{\Pi}_2$ is the same as the rank of Π_2 and (ii) $\bar{\pi}_1 = \bar{\Pi}_2 \beta$ if and only if $\pi_1 = \Pi_2 \beta$.

It is natural to consider tests for identification based on some transformations of the sufficient statistics $\left[\tilde{\pi}_1, \tilde{\Pi}_2 \right]$, $\left[\tilde{\phi}_1, \tilde{\Phi}_2 \right]$ and S . Since these induce transformations of the parameters $[\phi_1, \Phi_2]$, $[\pi_1, \Pi_2]$ and Ω we need to make sure that the same relationships hold for the transformed parameters as for the original parameters. For example, if the over-identifying restrictions (3) hold (resp. do not hold) for the original parameters, they must hold (resp. not hold) for the transformed ones. This concept is expressed formally by saying that the testing problem is invariant under some group of

transformations (for a detailed technical discussion see Lehmann (1997) and Muirhead (1982)). In this case, it is sensible to focus only on the class of invariant tests for which the test statistics are functions of $[\tilde{\pi}_1, \tilde{\Pi}_2]$, $[\tilde{\phi}_1, \tilde{\Phi}_2]$ and S that are left unchanged by the transformations of $[\tilde{\pi}_1, \tilde{\Pi}_2]$, $[\tilde{\phi}_1, \tilde{\Phi}_2]$ and S leaving the testing problem invariant.

Proposition 1. *Let A be any $(k_1 \times n + 1)$ matrix, $l_{11} > 0$, L_{22} be an $(n \times n)$ non-singular matrix and l_{21} be any $n \times 1$ vector and E_1 and E_2 be orthogonal matrices of dimension $(k_2 \times k_2)$ and $(k_1 \times k_1)$. Then, the transformations*

$$[\tilde{\pi}_1, \tilde{\Pi}_2] \rightarrow E_1 [\tilde{\pi}_1, \tilde{\Pi}_2] \begin{bmatrix} l_{11} & 0 \\ l_{21} & L_{22} \end{bmatrix} \quad (12)$$

$$[\tilde{\phi}_1, \tilde{\Phi}_2] \rightarrow E_2 [\tilde{\phi}_1, \tilde{\Phi}_2] \begin{bmatrix} l_{11} & 0 \\ l_{21} & L_{22} \end{bmatrix} + A, \quad (13)$$

and

$$S \rightarrow \begin{bmatrix} l_{11} & 0 \\ l_{21} & L_{22} \end{bmatrix}, S \begin{bmatrix} l_{11} & 0 \\ l_{21} & L_{22} \end{bmatrix} \quad (14)$$

do not affect the over-identifying restrictions nor the rank of Π_2 .

The reason for choosing the transformation (12) rather than the more standard one

$$[\tilde{\pi}_1, \tilde{\Pi}_2] \rightarrow E_1 [\tilde{\pi}_1, \tilde{\Pi}_2] L \quad (15)$$

where L is an $(n+1 \times n+1)$ non-singular matrix, is that (15) affects both the rank of Π_2 and (3), and consequently it does not leave the problem invariant.

A statistic that is left unchanged by the transformations (12), (13) and (14) is said to be invariant to them. One can show that there is a statistic, called the maximal invariant, having the properties that all statistics invariant to the transformations of $[\tilde{\pi}_1, \tilde{\Pi}_2]$, $[\tilde{\phi}_1, \tilde{\Phi}_2]$ and S specified in Proposition 1 are functions of the maximal invariant. Formally, the statistic $f([\tilde{\pi}_1, \tilde{\Pi}_2], [\tilde{\phi}_1, \tilde{\Phi}_2], S)$ is a maximal invariant if

$$f([\tilde{\pi}_1, \tilde{\Pi}_2], [\tilde{\phi}_1, \tilde{\Phi}_2], S) = f([\tilde{\pi}_1, \tilde{\Pi}_2]^*, [\tilde{\phi}_1, \tilde{\Phi}_2]^*, S^*)$$

implies that $\left\{ \left[\tilde{\pi}_1, \tilde{\Pi}_2 \right]^*, \left[\tilde{\phi}_1, \tilde{\Phi}_2 \right]^*, S^* \right\}$ can be obtained from $\left\{ \left[\tilde{\pi}_1, \tilde{\Pi}_2 \right], \left[\tilde{\phi}_1, \tilde{\Phi}_2 \right], S \right\}$ using the transformations (12), (13) and (14). Notice that the maximal invariant is not uniquely defined and that any one-to-one function of the maximal invariant is itself a maximal invariant. In our case, we have the following result.

Proposition 2. *The maximal invariant under the transformations (12), (13) and (14) is*

$$\left(\hat{\tau}, \hat{\beta}_{TSLs}^*, \hat{F} \right)$$

where

$$\begin{aligned} \hat{\tau} &= \tilde{\pi}_1' M_{\tilde{\Pi}_2} \tilde{\pi}_1 / s_{11.2} \\ \hat{\beta}_{TSLs}^* &= \left(L_{22}^{-1} \hat{\beta}_{TSLs} - L_{22}' s_{21} \right) / \sqrt{s_{11.2}} \\ \hat{\beta}_{TSLs} &= \left(\tilde{\Pi}_2' \tilde{\Pi}_2 \right)^{-1} \tilde{\Pi}_2' \tilde{\pi}_1 \\ s_{11.2} &= s_{11} - s_{21}' S_{22}^{-1} s_{21} \end{aligned}$$

and L_{22} is a matrix such that $L_{22}' S_{22} L_{22} = I_n$ and $L_{22}' \tilde{\Pi}_2' \tilde{\Pi}_2 L_{22} = \hat{F}$ where \hat{F} is an $(n \times n)$ matrix containing as diagonal elements the eigenvalues of $S_{22}^{-1/2} \tilde{\Pi}_2' \tilde{\Pi}_2 S_{22}^{-1/2}$.

Moreover, $\left(\hat{\tau}, \hat{\beta}_{TSLs}^*, \hat{F} \right)$ depends only on $\left(\tau, \beta^*, F \right)$ where

$$\begin{aligned} \tau &= \bar{\pi}_1' M_{\bar{\Pi}_2} \bar{\pi}_1 / \omega_{11.2} \\ \beta^* &= \left(L_{22}^{-1} \beta - L_{22}' \omega_{21} \right) / \sqrt{\omega_{11.2}} \end{aligned}$$

and L_{22} is now a matrix such that $L_{22}' \Omega_{22} L_{22} = I_n$ and $L_{22}' \bar{\Pi}_2' \bar{\Pi}_2 L_{22} = F$, where F is an $(n \times n)$ matrix containing as diagonal elements the eigenvalues of $\Omega_{22}^{-1/2} \bar{\Pi}_2' \bar{\Pi}_2 \Omega_{22}^{-1/2}$.

Discussion

Proposition 2 identifies all the relevant statistics that may be used to construct tests for identification that are optimal. Since the distribution of the maximal invariant is very complicated, it is difficult to derive optimal tests for identification. It may, however, be possible to use it to prove the admissibility or inadmissibility of existing tests. We hope to be able to report some results in future work.

Here we discuss the relationship between the statistics $(\hat{\tau}, \hat{\beta}_{TSLs}^*, \hat{F})$ and existing tests. For the sake of simplicity, we define

$$\hat{W} = \begin{bmatrix} \hat{w}_{11} & \hat{w}_{21}' \\ \hat{w}_{21} & \hat{W}_{22} \end{bmatrix} = TL' [\tilde{\pi}_1, \tilde{\Pi}_2]' [\tilde{\pi}_1, \tilde{\Pi}_2] L, \quad (16)$$

where L is a block diagonal matrix such that $L'\Omega L = I_{n+1}$. The maximal invariant identified in Proposition 2 is the same as \hat{W} with the right bottom $(n \times n)$ block replaced by its eigenvalues and L replaced by a consistent estimate \hat{L} (that is also a block diagonal matrix).

Under standard assumptions

$$\sqrt{T} \left([\tilde{\pi}_1, \tilde{\Pi}_2] - [\bar{\pi}_1, \bar{\Pi}_2] \right) L = \sqrt{T} \left([\tilde{\pi}_1^*, \tilde{\Pi}_2^*] - [\bar{\pi}_1^*, \bar{\Pi}_2^*] \right) \rightarrow^D N(0, I_{k_2} \otimes I_{n+1}) \quad (17)$$

where $\tilde{\Pi}_2^* = \tilde{\Pi}_2 L_{22}$, $\tilde{\pi}_1^* = (\tilde{\pi}_1 - \tilde{\Pi}_2 \Omega_{22}^{-1} \omega_{21}) / \sqrt{\omega_{11.2}}$, $\bar{\Pi}_2^* = \bar{\Pi}_2 L_{22}$ and $\bar{\pi}_1^* = (\bar{\pi}_1 - \bar{\Pi}_2 \Omega_{22}^{-1} \omega_{21}) / \sqrt{\omega_{11.2}}$. For all $\beta \in \mathbb{R}^n$ let $\beta^* = (L_{22}^{-1} \beta - L_{22}' \omega_{21}) / \sqrt{\omega_{11.2}}$, then it follows from (17) that

$$\sqrt{T} \left(\tilde{\pi}_1^* - \tilde{\Pi}_2^* \beta^* - (\bar{\pi}_1^* - \bar{\Pi}_2^* \beta^*) \right) \rightarrow^D N(0, (1 + \beta^{*'} \beta^*) I_{k_2}). \quad (18)$$

If the over-identifying restrictions and the identification condition hold $\bar{\pi}_1^* - \bar{\Pi}_2^* \beta^* = (\bar{\pi}_1 - \bar{\Pi}_2 \beta) / \sqrt{\omega_{11.2}} = 0$ and β^* is the (unique) vector of partial correlations between the rows of Y_2 and the corresponding component of u . Thus the quantity $\beta^{*'} \beta^* / (1 + \beta^{*'} \beta^*)$ measures the degree of endogeneity (e.g. Phillips (1983)).

A test for exogeneity is a test for $H_0 : \beta^* = 0$ (equivalently $H_0 : \beta = \Omega_{22}^{-1} \omega_{21}$) versus $H_1 : \beta^* \neq 0$ (equivalently $H_0 : \beta \neq \Omega_{22}^{-1} \omega_{21}$). Notice that the maintained hypothesis under both the null and the alternative hypothesis is that the over-identifying restrictions $\bar{\pi}_1 - \bar{\Pi}_2 \beta = 0$ hold and $\bar{\Pi}_2$ has rank n . It is clear from (18) that a test based on $\tilde{\pi}_1^*$ (and thus \hat{w}_{11}) would pick up failures of the null hypothesis of exogeneity as well as failures of the over-identifying restrictions. Thus, one may project (18) onto the space spanned by the columns of $\sqrt{T} \tilde{\Pi}_2^*$ to obtain

$$\frac{\hat{w}_{21}' \hat{W}_{22}^{-1} \hat{w}_{21}}{1 + \beta^{*'} \beta^*} \rightarrow^D \chi^2(n)$$

that can be written as

$$\frac{T \left(\hat{\beta}_{TSLs} - \Omega_{22}^{-1} \omega_{21} \right)' \tilde{\Pi}_2' \tilde{\Pi}_2 \left(\hat{\beta}_{TSLs} - \Omega_{22}^{-1} \omega_{21} \right)}{\omega_{11.2} (1 + \beta^{*'} \beta^*)} = \frac{T \hat{\beta}_{TSLs}^*{}' \tilde{\Pi}_2^*{}' \tilde{\Pi}_2^* \hat{\beta}_{TSLs}^*}{1 + \beta^{*'} \beta^*} \rightarrow^D \chi^2(n),$$

where $\hat{\beta}_{TSLs} = \left(\tilde{\Pi}_2' \tilde{\Pi}_2 \right)^{-1} \tilde{\Pi}_2' \tilde{\pi}_1$ denotes the TSLS estimator of β , and $\hat{\beta}_{TSLs}^* = \left(L_{22}^{-1} \hat{\beta}_{TSLs} - L_{22}' \omega_{21} \right) / \sqrt{\omega_{11.2}}$. Notice that once we replace the unknown quantities with consistent estimators under the null hypothesis, the resulting expression can be interpreted as a Wald test for $\beta - \Omega_{22}^{-1} \omega_{21} = 0$ as well as a test in the sense of Hausman (1978) (since it compares $\hat{\beta}_{TSLs}$ with a consistent estimate $(\hat{\Omega}_{22}^{-1} \hat{\omega}_{21})$ of β under the null hypothesis that $\beta = \Omega_{22}^{-1} \omega_{21}$).

Notice that (18) can be a basis for tests for over-identifying restrictions, too. However, one faces the problem that if the null hypothesis is true, then $\sqrt{T} \left(\tilde{\pi}_1^* - \tilde{\Pi}_2^* \beta^* \right) \rightarrow^D N \left(0, (1 + \beta^{*'} \beta^*) I_{k_2} \right)$ for the “true”, but unknown, β^* . An obvious way to proceed is to project $\sqrt{T} \left(\tilde{\pi}_1 - \tilde{\Pi}_2 \beta^* \right)$ onto the space orthogonal to the one spanned by $\sqrt{T} \tilde{\Pi}_2^*$ and consider the quadratic form

$$\frac{\hat{w}_{11.2}}{1 + \beta^{*'} \beta^*} = \frac{\tilde{\pi}_1^{*'} M_{\tilde{\Pi}_2^*} \tilde{\pi}_1^*}{1 + \beta^{*'} \beta^*} \rightarrow^D \chi^2(k_2 - n),$$

where $\hat{w}_{11.2} = \hat{w}_{11} - \hat{w}_{21}' \hat{W}_{22}^{-1} \hat{w}_{21} = T \hat{\tau}$. This statistic yields the over-identification tests suggested by Sargan (1958), Basman (1960a), Basman (1960b), Byron (1974) and Hansen (1982) for which $\omega_{11.2} (1 + \beta^{*'} \beta^*)$ (i.e. the variance of the structural error) is estimated in different ways. Notice that such test is based on the maintained hypothesis that that Π_2 has rank n .

Functions of the statistic \hat{F} (i.e. the eigenvalues of \hat{W}_{22}) have often been suggested to test the rank of Π_2 (e.g. Anderson and Rubin (1949), Anderson (1951),

Cragg and Donald (1997), Robin and Smith (2000)). Note that \hat{F} is the maximal invariant under the transformations

$$\begin{aligned}\hat{\Pi}_2 &\rightarrow E_1 \hat{\Pi}_2 L_{22} \\ \hat{\Phi}_2 &\rightarrow E_2 \hat{\Phi}_2 L_{22} + A_2 \\ S_{22} &\rightarrow L_{22}' S_{22} L_{22}\end{aligned}$$

of that part of the reduced form (2) involving Y_2 , so that \hat{F} depends only on F irrespective of equation (3). Therefore, it is particularly suited to construct tests and equivariant estimators for the rank of Π_2 .

Since \hat{F} , $\hat{w}_{11.2}$ and $\hat{w}_{21}' \hat{W}_{22}^{-1} \hat{w}_{21}$ are asymptotically independent, it is easy to control the overall size of tests for identification and exogeneity. However, given the maintained hypotheses of the tests, their interpretation has to be based on the logical sequence according to which tests are performed in the following order: test of rank, test of over-identifying restriction, test of exogeneity.

Appendix: Proofs

Proof of Proposition 1

Let

$$L = \begin{bmatrix} l_{11} & 0 \\ l_{21} & L_{22} \end{bmatrix}.$$

The transformations (12), (13) and (14) change the sufficient statistics to

$$E_1 [\tilde{\pi}_1, \tilde{\Pi}_2] L \sim N \left([\bar{\pi}_1, \bar{\Pi}_2]^*, I_{k_2} \otimes L' \Omega L \right),$$

$$E_2 [\tilde{\phi}_1, \tilde{\Phi}_2] L \sim N \left([\bar{\phi}_1, \bar{\Phi}_2]^* + E_2 (Z_1' Z_1)^{-1/2} Z_1' Z_2 \Pi_2 E_1^{-1} [\bar{\pi}_1, \bar{\Pi}_2]^* + A, I_{k_1} \otimes L' \Omega L \right),$$

and

$$L' S L \sim W_{n+1} (T - k_1 - k_2, L' \Omega L),$$

where

$$[\bar{\pi}_1, \bar{\Pi}_2]^* = E_1 [\bar{\pi}_1, \bar{\Pi}_2] \begin{bmatrix} l_{11} & 0 \\ l_{21} & L_{22} \end{bmatrix} = [l_{11} E_1 \bar{\pi}_1 + E_1 \bar{\Pi}_2 l_{21}, E_1 \bar{\Pi}_2 L_{22}], \quad (19)$$

and

$$[\bar{\phi}_1, \bar{\Phi}_2]^* = E_1 [\bar{\phi}_1, \bar{\Phi}_2] \begin{bmatrix} l_{11} & 0 \\ l_{21} & L_{22} \end{bmatrix} + A. \quad (20)$$

Notice that $[\bar{\phi}_1, \bar{\Phi}_2]^*$ and $[\bar{\phi}_1, \bar{\Phi}_2]$ are arbitrary, the rank of $\bar{\Pi}_2^*$ is the same as the rank of $\bar{\Pi}_2$ and $\bar{\pi}_1^* = \bar{\Pi}_2^* \beta^*$ if and only if (3) holds since by definition $\bar{\pi}_1^* = l_{11} E_1 \bar{\pi}_1 + E_1 \bar{\Pi}_2 l_{21} = \bar{\Pi}_2^* \beta^* = E_1 \bar{\Pi}_2 L_{22} \beta^*$, and this can be rewritten as $\bar{\pi}_1 = \bar{\Pi}_2 (L_{22} \beta^* - l_{21}) / l_{11} = \bar{\Pi}_2 \beta$ with $\beta = (L_{22} \beta^* - l_{21}) / l_{11}$.

Proof of Proposition 2

Proof. First, we consider the transformation

$$\{[\tilde{\pi}_1, \tilde{\Pi}_2], [\tilde{\phi}_1, \tilde{\Phi}_2], S\} \rightarrow \{[\tilde{\pi}_1, \tilde{\Pi}_2], [\tilde{\phi}_1, \tilde{\Phi}_2] + A, S\},$$

so that the maximal invariant is $\{[\tilde{\pi}_1, \tilde{\Pi}_2], S\}$ and $[\tilde{\phi}_1, \tilde{\Phi}_2]$ can be dropped. We now verify that the definition of maximal invariant is satisfied: the statistic $\{[\tilde{\pi}_1, \tilde{\Pi}_2], S\} = f([\tilde{\pi}_1, \tilde{\Pi}_2], [\tilde{\phi}_1, \tilde{\Phi}_2], S)$ is invariant and is such that

$$f([\tilde{\pi}_1, \tilde{\Pi}_2], [\tilde{\phi}_1, \tilde{\Phi}_2], S) = f([\tilde{\pi}_1, \tilde{\Pi}_2], [\tilde{\phi}_1, \tilde{\Phi}_2]^*, S)$$

implies that $\{[\tilde{\pi}_1, \tilde{\Pi}_2], [\tilde{\phi}_1, \tilde{\Phi}_2]^*, S\} = \{[\tilde{\pi}_1, \tilde{\Pi}_2], [\tilde{\phi}_1, \tilde{\Phi}_2] + A, S\}$ for a suitable matrix A .

Thus, it is a maximal invariant.

Second, we focus on the transformation (12) and derive the maximal invariant in two steps. (i) Consider $\{[\tilde{\pi}_1, \tilde{\Pi}_2], S\} \rightarrow \{E_1 [\tilde{\pi}_1, \tilde{\Pi}_2], S\}$. The maximal invariant is $\{W, S\}$ where $W = [\tilde{\pi}_1, \tilde{\Pi}_2]' [\tilde{\pi}_1, \tilde{\Pi}_2]$. The fact that the statistic $\{W, S\}$ is a maximal invariant follows from the definition of maximal invariant: it is an invariant function of $\{[\tilde{\pi}_1, \tilde{\Pi}_2], S\}$ and if $\{[\tilde{\pi}_1, \tilde{\Pi}_2]' [\tilde{\pi}_1, \tilde{\Pi}_2], S\} = \{[\tilde{\pi}_1, \tilde{\Pi}_2]^* [\tilde{\pi}_1, \tilde{\Pi}_2]^*, S\}$ then $[\tilde{\pi}_1, \tilde{\Pi}_2]^* = H [\tilde{\pi}_1, \tilde{\Pi}_2]$ for some orthogonal matrix H (e.g. Theorem A9.5 of Muirhead (1982)). (ii) Then, we focus on the transformations

$$\{W, S\} \rightarrow \{L'WL, L'SL\} \quad (21)$$

where

$$L = \begin{bmatrix} l_{11} & 0 \\ l_{21} & L_{22} \end{bmatrix}.$$

The maximal invariant is given by \hat{W} , where $\hat{W} = L'WL$ and L is chosen so that $L'SL = I_{n+1}$. Once again, this follows immediately from the definition of maximal invariant by noting that the product of two block diagonal matrices conformable to L is a block diagonal matrix conformable to L . Now, we just need to choose a convenient matrix L . To do this we partition W , \hat{W} and S conformably to L ,

$$W = \begin{bmatrix} w_{11} & w_{21}' \\ w_{21} & W_{22} \end{bmatrix} \quad \hat{W} = \begin{bmatrix} \hat{w}_{11} & \hat{w}_{21}' \\ \hat{w}_{21} & \hat{W}_{22} \end{bmatrix} \quad S = \begin{bmatrix} s_{11} & s_{21}' \\ s_{21} & S_{22} \end{bmatrix},$$

and select L as follows:

(1) Since $W_{22} \rightarrow L_{22}'W_{22}L_{22}$ and $S_{22} \rightarrow L_{22}'S_{22}L_{22}$, we can choose L_{22} in such a way that $L_{22}'S_{22}L_{22} = I_n$ and $\hat{W}_{22} = L_{22}'W_{22}L_{22} = F$ where F is an $n \times n$ matrix containing the eigenvalues of $S_{22}^{-1/2}W_{22}S_{22}^{-1/2} = S_{22}^{-1/2}\tilde{\Pi}_2'\tilde{\Pi}_2S_{22}^{-1/2}$ (e.g. Theorem A9.9 of Muirhead (1982)).

The matrix L_{22} is uniquely determined up to changes of sign in its first row.

(2) We choose L_{22} as defined above and $l_{11} = 1/\sqrt{s_{11.2}}$ and $l_{21} = -S_{22}^{-1}s_{21}'/\sqrt{s_{11.2}}$ where $s_{11.2} = s_{11} - s_{21}'S_{22}^{-1}s_{21}$, so that $L'SL = I_{n+1}$.

Then we can write

$$\begin{bmatrix} \hat{w}_1 & \hat{W}_2 \end{bmatrix} = \begin{bmatrix} \tilde{\pi}_1 & \tilde{\Pi}_2 \end{bmatrix} L = \begin{bmatrix} (\tilde{\pi}_1 - \tilde{\Pi}_2 S_{22}^{-1} s_{21}) / \sqrt{s_{11.2}} & \tilde{\Pi}_2 L_{22} \end{bmatrix}$$

so that

$$\begin{aligned} \hat{w}_{11} &= \hat{w}_1' \hat{w}_1 \\ &= (\tilde{\pi}_1 - \tilde{\Pi}_2 S_{22}^{-1} s_{21})' (\tilde{\pi}_1 - \tilde{\Pi}_2 S_{22}^{-1} s_{21}) / s_{11.2} \\ &= (\tilde{\pi}_1 - \tilde{\Pi}_2 S_{22}^{-1} s_{21})' P_{\tilde{\Pi}_2} (\tilde{\pi}_1 - \tilde{\Pi}_2 S_{22}^{-1} s_{21}) / s_{11.2} + (\tilde{\pi}_1 - \tilde{\Pi}_2 S_{22}^{-1} s_{21})' M_{\tilde{\Pi}_2} (\tilde{\pi}_1 - \tilde{\Pi}_2 S_{22}^{-1} s_{21}) / s_{11.2} \\ &= (\hat{\beta}_{TSLs} - S_{22}^{-1} s_{21})' \tilde{\Pi}_2' \tilde{\Pi}_2 (\hat{\beta}_{TSLs} - S_{22}^{-1} s_{21}) / s_{11.2} + \tilde{\pi}_1' M_{\tilde{\Pi}_2} \tilde{\pi}_1 / s_{11.2} \\ &= \hat{\beta}_{TSLs}^* ' F \hat{\beta}_{TSLs}^* + \hat{\tau} \end{aligned}$$

and

$$\begin{aligned}
\hat{w}_{21} &= \hat{W}_2' \hat{w}_1 = L_{22}' \tilde{\Pi}_2' (\tilde{\pi}_1 - \tilde{\Pi}_2 S_{22}^{-1} s_{21}) / \sqrt{s_{11.2}} \\
&= L_{22}' \tilde{\Pi}_2' \tilde{\Pi}_2 (\hat{\beta}_{TSLs} - S_{22}^{-1} s_{21}) / \sqrt{s_{11.2}} \\
&= F \hat{\beta}_{TSLs}^*
\end{aligned}$$

where $\hat{\beta}_{TSLs} = (\tilde{\Pi}_2' \tilde{\Pi}_2)^{-1} \tilde{\Pi}_2' \tilde{\pi}_1$ is the TSLS estimator of β and $\hat{\beta}_{TSLs}^* = (L_{22}^{-1} \hat{\beta}_{TSLs} - L_{22}' s_{21}) / \sqrt{s_{11.2}}$.

(3) Finally, the transformation $(\hat{w}_{11}, F \hat{\beta}_{TSLs}^*, F) \rightarrow (\hat{\tau}, \hat{\beta}_{TSLs}^*, F)$ is one-to-one. The rest of Proposition 2 follows from Theorem 6.1.12 of Muirhead (1982).

References

- Anderson, T. W. (1951), "The Asymptotic Distribution of Certain Characteristic Roots and Vectors", *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, J. Neyman, Berkeley, University of California Press, 103-130.
- 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 Statistics* 20, 46-62.
- Basmann, R. L. (1960a), "On Finite Sample Distributions of Generalized Classical Linear Identifiability Test Statistics", *Journal of the American Statistical Association* 55, 650-659.
- Basmann, R. L. (1960b), "On the Asymptotic Distribution of Generalised Linear Estimators", *Econometrica* 28, 97-107.
- Byron, R. P. (1974), "Testing Structural Specification Using the Unrestricted Reduced Form", *Econometrica* 42, 869-883.
- Cragg, J. G. and S. G. Donald (1993), "Testing Identification and Specification in Instrumental Variables Models", *Econometric Theory* 9, 222-240.
- Cragg, J. G. and S. G. Donald (1997), "Inferring the Rank of a Matrix", *Journal of Econometrics* 76, 223-250.
- Hansen, L. P. (1982), "Large Sample Properties of Generalized Method of Moments Estimators", *Econometrica* 40, 1029-1054.
- Hausman, J. A. (1978), "Specification Tests in Econometrics", *Econometrica* 46(6), 1251-1271.
- Hillier, G. H. (1990), "On the Normalization of Structural Equations: Properties of Direction Estimators", *Econometrica* 58, 1181-1194.
- Lehmann, E. L. (1997), *Testing Statistical Hypotheses* New York, Springer-Verlag.
- Muirhead, R. J. (1982), *Aspects of Multivariate Statistical Theory* New York, John Wiley and Sons, Inc.

- Phillips, P. C. B. (1983), "Exact Small Sample Theory in the Simultaneous Equation Model", *Handbook of Econometrics*, Z. Griliches, Amsterdam, North Holland, 449-516.
- Phillips, P. C. B. (1989), "Partially Identified Econometric Models", *Econometric Theory* 5, 181-240.
- Robin, J.-M. and R. J. Smith (2000), "Tests of Rank", *Econometric Theory* 16, 151-176.
- Sargan, J. D. (1958), "The Estimation of Economic Relationships using Instrumental Variables", *Econometrica* 26, 393-415.
- Staiger, D. and J. H. Stock (1997), "Instrumental Variables Regression with Weak Instruments", *Econometrica* 65, 557-586.
- Stock, J. and M. Yogo (2001), "Testing for Weak Instruments in Linear IV regression", *Identification and Inference for Econometric Models: A Festschrift for Tom Rothenberg*, J. Stock, Cambridge, Cambridge University Press.