

## **TESTING A HYPOTHESIS ABOUT A STRUCTURAL COEFFICIENT IN A SIMULTANEOUS EQUATION MODEL FOR KNOWN COVARIANCE MATRIX**

THEODORE W. ANDERSON

*Department of Statistics and Department of Economics, Stanford University*  
*twa@stanford.edu*

NAOTO KUNITOMO

*Graduate School of Economics, University of Tokyo*  
*kunitomo@e.u-tokyo.ac.jp*

### **ABSTRACT**

There is now considerable econometrics literature on estimating the coefficients of one equation in a simultaneous equation model. Recently, econometricians are entertaining the problem of testing a hypothesis concerning the coefficients of a single equation. In this paper we derive the likelihood ratio test for a vector of coefficients in the case when the error covariance is known.

#### ***Key words***

Error covariance matrix; likelihood ratio test; simultaneous equation model; testing hypothesis.

### **RESUMEN**

Existe considerable literatura en econometría sobre la estimación de los coeficientes de una ecuación individual en un modelo de ecuaciones simultáneas. Recientemente, algunos economistas se están ocupando del problema de la prueba de una hipótesis concerniente a los coeficientes de una ecuación individual. En este artículo deducimos una prueba de razón de verosimilitud para un vector de coeficientes en el caso en que la matriz de covarianzas de errores es conocida.

#### ***Palabras clave***

Matriz de covarianzas de errores; prueba de razón de verosimilitud; modelo de ecuaciones simultáneas; prueba de hipótesis.

## 1. Introduction

There is a considerable econometrics literature about *estimating* coefficients of one equation in a simultaneous equation model (Anderson and Rubin (1949), for example). More recently, econometricians have been studying the problem of *testing* a hypothesis concerning the coefficients of a single equation. See Andrews, Moreira and Stock (2006), for example. In a previous study Anderson and Kunitomo (2007) have derived the likelihood ratio test of a vector of coefficients when it is assumed that the single equation is identified and the covariance is estimated. In the present paper we derive the likelihood ratio test in the case of the error covariance matrix known.

## 2. The model

Let  $\mathbf{Y}$  be a  $T \times G$  matrix of endogenous or dependent variables; let  $\mathbf{Z}$  be a  $T \times K$  matrix of exogenous or independent variables; and let  $\mathbf{V}$  be a  $T \times G$  matrix of unobservable errors or shocks, which are statistically independent of  $\mathbf{Z}$ . The rows of  $\mathbf{V}$  are assumed independent, the  $t$ -th row of which has mean  $\mathcal{E}(\mathbf{v}_t) = \mathbf{0}$  and covariance  $\mathcal{E}(\mathbf{v}_t \mathbf{v}'_t) = \boldsymbol{\Omega}$ . (The matrix  $\boldsymbol{\Omega}$  is assumed known.) The "reduced form" of the model is

$$\mathbf{Y} = \mathbf{Z}\boldsymbol{\Pi} + \mathbf{V}, \quad (1)$$

where  $\boldsymbol{\Pi}$  is a  $K \times G$  matrix of parameters. Statisticians will recognize (1) as a multivariate regression.

A single *structural* equation of interest may be written as

$$\mathbf{Y}\boldsymbol{\beta} = \mathbf{Z}\boldsymbol{\gamma} + \mathbf{u}, \quad (2)$$

where  $\mathbf{u} = \mathbf{V}\boldsymbol{\beta}$ . To distinguish this equation from other linear combinations of  $\mathbf{Y}$  and  $\mathbf{Z}$ , some "identifying" assumptions concerning  $\boldsymbol{\beta}$  and  $\boldsymbol{\gamma}$  are made; in particular some components of  $\boldsymbol{\beta}$  and  $\boldsymbol{\gamma}$  may be assumed  $\mathbf{0}$ . The relation between (1) and (2) implies

$$\boldsymbol{\Pi}\boldsymbol{\beta} = \mathbf{u}. \quad (3)$$

The equation (1) is said to be *identified* by 0's if given  $\boldsymbol{\Pi}$  (3) has a unique solution for  $(\boldsymbol{\beta}, \boldsymbol{\gamma})$  except a constant of proportionality.

In this paper first we pay special attention to the case that  $\boldsymbol{\gamma} = \mathbf{0}$ . Then the relation (3) is

$$\boldsymbol{\Pi}\boldsymbol{\beta} = \mathbf{0}. \quad (4)$$

Then  $\boldsymbol{\beta}$  is identified if  $\boldsymbol{\Pi}$  has rank  $G - 1$ .

We consider the null hypothesis  $H_0 : \boldsymbol{\beta} = \boldsymbol{\beta}_0$ , where  $\boldsymbol{\beta}_0$  is a specified vector. This null hypothesis is equivalent to the hypothesis that  $\boldsymbol{\Pi}\boldsymbol{\beta}_0 = \mathbf{0}$ .

The unrestricted regression coefficient  $\Pi$  can be estimated by the sample regression

$$\mathbf{P} = (\mathbf{Z}' \mathbf{Z})^{-1} \mathbf{Z}' \mathbf{Y}. \quad (5)$$

When  $\Omega$  is known and  $\mathbf{V}$  is normal, the matrix  $\mathbf{P}$  constitutes a sufficient set of statistics. The density of  $\mathbf{P}$  for  $\Omega = \mathbf{I}_G$  is

$$L(\Pi) = (2\pi)^{-\frac{1}{2}TG} \exp \left\{ -\frac{1}{2} \text{tr}(\mathbf{P} - \Pi)' \mathbf{A}(\mathbf{P} - \Pi) \right\}, \quad (6)$$

where

$$\mathbf{A} = \mathbf{Z}' \mathbf{Z}.$$

The likelihood maximized with respect to  $\Pi$  is

$$L_{H_2} = (2\pi)^{-\frac{1}{2}TG},$$

where  $H_2$  denotes the model with no restriction on  $\Pi$ .

The model and the inference problem have some invariance properties. If  $\mathbf{Z}$  is replaced by  $\tilde{\mathbf{Z}} = \mathbf{Z}\mathbf{C}$  and  $\Pi$  is replaced by  $\tilde{\Pi} = \mathbf{C}^{-1}\Pi$ , then the model (1) is unchanged since  $\Pi\beta = \mathbf{0}$  is transformed to  $\tilde{\Pi}\beta = \mathbf{0}$ . The null hypothesis is unaffected. Note that  $\mathbf{C}$  is an arbitrary nonsingular matrix.

If  $\mathbf{Y}$  is replaced by  $\mathbf{Y}\Phi = \mathbf{Y}^*$ ,  $\mathbf{V}$  is replaced by  $\mathbf{V}\Phi = \mathbf{V}^*$  ( $\mathbf{v}_t^* = \Phi' \mathbf{v}_t$ ),  $\Pi$  is replaced by  $\Pi\Phi = \Pi^*$ ,  $\beta$  is replaced by  $\Phi^{-1}\beta = \beta^*$ , and  $\beta_0$  is replaced by  $\Phi^{-1}\beta_0 = \beta_0^*$ , the model is unchanged with  $\Omega$  replaced by  $\Phi'\Omega\Phi = \Omega^*$ . When  $\Omega = \Omega^* = \mathbf{I}_G$  is required,  $\Phi'\Phi = \mathbf{I}_G$ , that is,  $\Phi$  is an orthogonal matrix.

### 3. Likelihood Ratio Tests

We denote the model with  $\Pi$  having rank  $G - 1$  as  $H_1$ . We shall derive the likelihood ratio criterion to test  $H_0$  when  $H_1$  is assumed, namely  $L_{H_0}/L_{H_1}$ , as the ratio of the criterion to test  $H_0$  vs  $H_2$  and the criterion to test  $H_1$  vs  $H_2$ .

$$\frac{L_{H_0}}{L_{H_1}} = \frac{L_{H_0}/L_{H_2}}{L_{H_1}/L_{H_2}}.$$

The hypothesis that  $\Pi\beta_0 = \mathbf{0}$  can be considered as the hypothesis that  $\Pi$  has rank  $G - 1$  and  $\Pi\beta_0 = \mathbf{0}$ . When  $G = 2$ ,  $\Pi$  has rank 1, that is, can be written as  $\mu\gamma'$  for the K-vector  $\mu$  and the 2 vector  $\gamma$ , and that  $\gamma'\beta_0 = 0$ .

**Lemma:** If  $G = 2$  and  $\Pi = \mu\gamma'$ , then the likelihood function of  $\mathbf{P}$  maximized with respect to  $\mu$  is

$$\max_{\mu} L(\mu\gamma') = (2\pi)^{-T} \exp \left\{ -\frac{1}{2} \text{tr} \mathbf{P}' \mathbf{A} \mathbf{P} + \frac{1}{2} \gamma' \mathbf{P}' \mathbf{A} \mathbf{P} \gamma \right\}.$$

**Proof:**

$$\begin{aligned} L(\boldsymbol{\mu}\boldsymbol{\gamma}') &= (2\pi)^{-T} \exp \left\{ -\frac{1}{2} \text{tr}(\mathbf{P} - \boldsymbol{\mu}\boldsymbol{\gamma}')' \mathbf{A} (\mathbf{P} - \boldsymbol{\mu}\boldsymbol{\gamma}') \right\} \\ &= (2\pi)^{-T} \exp \left\{ -\frac{1}{2} \text{tr} [\mathbf{P}\mathbf{A}\mathbf{P}' - \boldsymbol{\gamma}\boldsymbol{\mu}' \mathbf{A}\mathbf{P} - \mathbf{P}'\mathbf{A}\boldsymbol{\mu}\boldsymbol{\gamma}' + \boldsymbol{\gamma}\boldsymbol{\mu}' \mathbf{A}\boldsymbol{\mu}\boldsymbol{\gamma}'] \right\} \end{aligned}$$

is maximized at  $\boldsymbol{\mu} = \mathbf{P}\boldsymbol{\gamma}/\boldsymbol{\gamma}'\boldsymbol{\gamma}$ . Q.E.D.

To eliminate the indeterminacy of a scale factor in  $\boldsymbol{\beta}$  we require

$$\boldsymbol{\beta}'\boldsymbol{\beta} = 1.$$

An alternative normalization is

$$\boldsymbol{\beta}'\boldsymbol{\Psi}\boldsymbol{\beta} = 1,$$

where  $\boldsymbol{\Psi}$  is a known positive semi-definite matrix. To specify  $\boldsymbol{\beta}' = (1, \boldsymbol{\beta}_2)$ , for example, we take

$$\boldsymbol{\Psi} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

When  $\boldsymbol{\Pi}$  is  $K \times 2$ , then  $\boldsymbol{\Pi}\boldsymbol{\beta} = \mathbf{0}$  implies

$$\boldsymbol{\beta}'\boldsymbol{\gamma} = 0.$$

Let  $\boldsymbol{\gamma}_0$  be a vector such that  $\boldsymbol{\beta}_0'\boldsymbol{\gamma}_0 = 0$ . Then

$$\begin{aligned} L_{H_0} &= (2\pi)^{-T} \exp \left\{ -\frac{1}{2} \text{tr} \mathbf{P}\mathbf{A}\mathbf{P}' + \frac{1}{2} \boldsymbol{\gamma}_0'\mathbf{P}'\mathbf{A}\mathbf{P}\boldsymbol{\gamma}_0 \right\} \\ &= (2\pi)^{-T} \exp \left\{ -\frac{1}{2} \boldsymbol{\beta}_0'\mathbf{P}'\mathbf{A}\mathbf{P}\boldsymbol{\beta}_0 \right\} \end{aligned}$$

since

$$\text{tr} \mathbf{P}\mathbf{A}\mathbf{P}' = \boldsymbol{\beta}'\mathbf{P}'\mathbf{A}\mathbf{P}\boldsymbol{\beta} + \boldsymbol{\gamma}'\mathbf{P}'\mathbf{A}\mathbf{P}\boldsymbol{\gamma}.$$

The likelihood ratio criterion for testing  $H_0$  against  $H_2$  is

$$\frac{L_{H_0}}{L_{H_2}} = \exp \left\{ -\frac{1}{2} \boldsymbol{\beta}_0'\mathbf{P}'\mathbf{A}\mathbf{P}\boldsymbol{\beta}_0 \right\}.$$

When the null hypothesis  $H_0$  is true,

$$-2 \log \frac{L_{H_0}}{L_{H_2}} = \boldsymbol{\beta}_0'\mathbf{P}'\mathbf{A}\mathbf{P}\boldsymbol{\beta}_0$$

has a  $\chi^2$ -distribution with  $K$  degrees of freedom since  $\mathbf{P}\boldsymbol{\beta}_0$  has the distribution  $N(\mathbf{0}, \sigma^2 \mathbf{A}^{-1})$ .

Now consider

$$H_1 : \text{rank } \boldsymbol{\Pi} = G - 1,$$

that is,  $\boldsymbol{\Pi}$  is written as  $\boldsymbol{\Pi} = \boldsymbol{\mu}\boldsymbol{\gamma}'$ . The likelihood of  $\mathbf{P}$  maximized subject to  $\boldsymbol{\Pi} = \boldsymbol{\mu}\boldsymbol{\gamma}'$  is

$$\begin{aligned} L_{H_1} &= \max_{\boldsymbol{\gamma}} \max_{\boldsymbol{\mu}} L(\boldsymbol{\mu}\boldsymbol{\gamma}') \\ &= (2\pi)^{-T} \exp \left\{ -\frac{1}{2} \text{tr} \mathbf{P} \mathbf{A} \mathbf{P}' + \frac{1}{2} \max_{\boldsymbol{\gamma}} \frac{\boldsymbol{\gamma}' \mathbf{P}' \mathbf{A} \mathbf{P} \boldsymbol{\gamma}}{\boldsymbol{\gamma}' \boldsymbol{\gamma}} \right\} \\ &= (2\pi)^{-T} \exp \left\{ -\frac{1}{2} \min_{\boldsymbol{\beta}} \frac{\boldsymbol{\beta}' \mathbf{P}' \mathbf{A} \mathbf{P} \boldsymbol{\beta}}{\boldsymbol{\beta}' \boldsymbol{\beta}} \right\}. \end{aligned}$$

The likelihood ratio criterion for testing  $H_1$  against  $H_2$  is

$$\begin{aligned} \frac{L_{H_1}}{L_{H_2}} &= \exp \left\{ -\frac{1}{2} \min_{\boldsymbol{\beta}} \frac{\boldsymbol{\beta}' \mathbf{P}' \mathbf{A} \mathbf{P} \boldsymbol{\beta}}{\boldsymbol{\beta}' \boldsymbol{\beta}} \right\} \\ &= \exp \left\{ -\frac{1}{2} \frac{\widehat{\boldsymbol{\beta}}' \mathbf{P}' \mathbf{A} \mathbf{P} \widehat{\boldsymbol{\beta}}}{\widehat{\boldsymbol{\beta}}' \widehat{\boldsymbol{\beta}}} \right\}, \end{aligned}$$

where  $\widehat{\boldsymbol{\beta}}$  is the maximum likelihood estimator of  $\boldsymbol{\beta}$  for  $H_1$ .

When the null hypothesis  $H_1$  is true,

$$\begin{aligned} -2 \log \frac{L_{H_1}}{L_{H_2}} &= \min_{\boldsymbol{\beta}} \frac{\boldsymbol{\beta}' \mathbf{P}' \mathbf{A} \mathbf{P} \boldsymbol{\beta}}{\boldsymbol{\beta}' \boldsymbol{\beta}} \\ &= \frac{\widehat{\boldsymbol{\beta}}' \mathbf{P}' \mathbf{A} \mathbf{P} \widehat{\boldsymbol{\beta}}}{\widehat{\boldsymbol{\beta}}' \widehat{\boldsymbol{\beta}}} \end{aligned}$$



has a limiting  $\chi^2$ -distribution with 1 degree of freedom (Anderson and Rubin, 1950). Note that  $\min_{\boldsymbol{b}} (\boldsymbol{b}' \mathbf{P}' \mathbf{A} \mathbf{P} \boldsymbol{b} / \boldsymbol{b}' \boldsymbol{b})$  is the smaller root of

$$|\mathbf{P}' \mathbf{A} \mathbf{P} - d \mathbf{I}_2| = 0.$$

The likelihood ratio criterion for testing  $H_0$  against  $H_1$  is

$$\begin{aligned}\frac{L_{H_0}}{L_{H_1}} &= \frac{L_{H_0}/L_{H_2}}{L_{H_1}/L_{H_2}} \\ &= \exp \left\{ \frac{1}{2} \beta'_0 \mathbf{P}' \mathbf{A} \mathbf{P} \beta_0 + \frac{1}{2} \min_{\beta} \frac{\beta' \mathbf{P}' \mathbf{A} \mathbf{P} \beta}{\beta' \beta} \right\} \\ &= \exp \left\{ \frac{1}{2} \left[ \frac{\hat{\beta}' \mathbf{P}' \mathbf{A} \mathbf{P} \hat{\beta}}{\hat{\beta}' \hat{\beta}} - \frac{\beta'_0 \mathbf{P}' \mathbf{A} \mathbf{P} \beta_0}{\beta'_0 \beta_0} \right] \right\}.\end{aligned}$$

Equivalently, the likelihood ratio criterion for testing  $H_0$  against  $H_1$  is

$$-2 \log \frac{L_{H_0}}{L_{H_1}} = \frac{\beta'_0 \mathbf{P}' \mathbf{A} \mathbf{P} \beta_0}{\beta'_0 \beta_0} - \frac{\hat{\beta}' \mathbf{P}' \mathbf{A} \mathbf{P} \hat{\beta}}{\hat{\beta}' \hat{\beta}}. \quad (7)$$

Note that the criterion does not depend on the normalization of  $\beta_0$  and of  $\hat{\beta}$ . The criterion (7) has been proposed by Moreira (2003). He called (7) the conditional likelihood ratio.

**Corollary:** The likelihood ratio criterion (7) is invariant with respect to the groups of transformations

$$\tilde{\mathbf{Z}} = \mathbf{ZC}, \tilde{\boldsymbol{\Pi}} = \mathbf{C}^{-1} \boldsymbol{\Pi}$$

and

$$\mathbf{Y}^* = \mathbf{Y}\Phi, \mathbf{P}^* = \mathbf{P}\Phi, \beta^* = \Phi'\beta, \beta_0^* = \Phi'\beta_0, \quad (8)$$

where  $\Phi\Phi' = \mathbf{I}$ . When  $\Omega = \sigma^2\mathbf{I}$ , the transformation (8) is appropriate. A normalization of  $\beta'\Psi\beta = 1$  is transformed to

$$\beta^{*\prime} \Phi' \Psi \Phi \beta^* = 1.$$

When  $G = 2$  and  $\Omega = \mathbf{I}$ , the condition  $\beta'\beta = 1$  implies that  $\beta$  can be written as

$$\beta = \begin{bmatrix} \cos \theta \\ \sin \theta \end{bmatrix} \quad (9)$$

and  $\gamma$  can be written as

$$\gamma = \text{constant} \times \begin{bmatrix} -\sin \theta \\ \cos \theta \end{bmatrix}.$$

The term  $\max_{\boldsymbol{\mu}} L(\boldsymbol{\mu}\boldsymbol{\gamma}')$  is independent of the constant which can be taken as 1. The null hypothesis  $\boldsymbol{\beta} = \boldsymbol{\beta}_0$  where

$$\boldsymbol{\beta}_0 = \begin{bmatrix} \cos \theta_0 \\ \sin \theta_0 \end{bmatrix}$$

implies that the null hypothesis  $\boldsymbol{\beta} = \boldsymbol{\beta}_0$  is equivalent to  $\boldsymbol{\theta} = \boldsymbol{\theta}_0$  and

$$\boldsymbol{\gamma} = \boldsymbol{\gamma}_0 = \begin{bmatrix} -\sin \theta_0 \\ \cos \theta_0 \end{bmatrix}.$$

Anderson and Kunitomo (2007) derived the likelihood ratio criterion for testing  $H_0$  against  $H_1$  when  $\boldsymbol{\Omega}$  is unknown and must be estimated. In that case let

$$\mathbf{H} = \mathbf{Y}'\mathbf{Y} - \mathbf{P}'\mathbf{A}\mathbf{P}.$$

The maximum likelihood estimators of  $\boldsymbol{\Pi}$  and  $\boldsymbol{\Omega}$  are  $\mathbf{P}$  and  $(1/T)\mathbf{H}$  (under  $H_2$ ). Then

$$-2 \log \frac{L_{H_0}^*}{L_{H_1}^*} = \log \left[ 1 + \frac{\boldsymbol{\beta}_0' \mathbf{P}' \mathbf{A} \mathbf{P} \boldsymbol{\beta}_0}{\boldsymbol{\beta}_0' \mathbf{H} \boldsymbol{\beta}_0} \right] - \log \left[ 1 + \min_{\mathbf{b}} \frac{\mathbf{b}' \mathbf{P}' \mathbf{A} \mathbf{P} \mathbf{b}}{\mathbf{b}' \mathbf{H} \mathbf{b}} \right]$$

is the likelihood ratio criterion for testing  $H_0$  against  $H_1$ . It might be observed that Anderson and Rubin (1949) derived the likelihood ratio criterion for testing  $H_1$  against  $H_2$  explicitly and the LR criterion for testing  $H_0$  against  $H_2$  implicitly (the converse of a confidence region), but did not derive the LR criterion for testing  $H_0$  against  $H_1$ .

When the null hypothesis is true,  $\mathbf{P}\boldsymbol{\beta}_0 \xrightarrow{p} 0$ ,  $(1/T)\mathbf{H} \xrightarrow{p} \boldsymbol{\Omega}$  as  $T \rightarrow \infty$ . Assume  $(1/T)\mathbf{A} \rightarrow \mathbf{A}_0$ . Then expansion of  $\log(1+x) = x + o_p(1)$  for  $x \rightarrow 0$  yields

$$-2 \log \frac{L_{H_0}^*}{L_{H_1}^*} = \frac{\boldsymbol{\beta}_0' \mathbf{P}' \mathbf{A} \mathbf{P} \boldsymbol{\beta}_0}{\boldsymbol{\beta}_0' \mathbf{H} \boldsymbol{\beta}_0} - \min_{\mathbf{b}} \frac{\mathbf{b}' \mathbf{P}' \mathbf{A} \mathbf{P} \mathbf{b}}{\mathbf{b}' \mathbf{H} \mathbf{b}} + o_p(1).$$

The limiting distribution of  $-2 \log[L_{H_0}/L_{H_1}]$  is the same as the limiting distribution  $-2 \log[L_{H_0}^*/L_{H_1}^*]$ . Under the null hypothesis the limiting distribution is  $\chi^2$  with 1 degree of freedom.

#### 4. More General Models

Suppose that  $\mathbf{Y}, \mathbf{Z}, \boldsymbol{\beta}$  and  $\boldsymbol{\gamma}$  in (2) are

$$\mathbf{Y} = (\mathbf{Y}_1, \mathbf{Y}_2), \mathbf{Z} = (\mathbf{Z}_1, \mathbf{Z}_2), \boldsymbol{\beta} = \begin{bmatrix} \boldsymbol{\beta}_1 \\ \mathbf{0} \end{bmatrix}, \boldsymbol{\gamma} = \begin{bmatrix} \boldsymbol{\gamma}_1 \\ \mathbf{0} \end{bmatrix},$$

where  $\mathbf{Y}_1$  and  $\mathbf{Y}_2$  have  $G_1$  and  $G_2$  columns, respectively, and  $\mathbf{Z}_1$  and  $\mathbf{Z}_2$  have  $K_1$  and  $K_2$  columns, respectively. Then (2) is

$$\mathbf{Y}_1\boldsymbol{\beta}_1 = \mathbf{Z}_1\boldsymbol{\gamma}_1 + \mathbf{u}. \quad (10)$$

Let

$$\boldsymbol{\Pi} = \begin{bmatrix} \boldsymbol{\Pi}_{11} & \boldsymbol{\Pi}_{12} \\ \boldsymbol{\Pi}_{21} & \boldsymbol{\Pi}_{22} \end{bmatrix}, \quad \boldsymbol{\Omega} = \begin{bmatrix} \boldsymbol{\Omega}_{11} & \boldsymbol{\Omega}_{12} \\ \boldsymbol{\Omega}_{21} & \boldsymbol{\Omega}_{22} \end{bmatrix}.$$

Then (2) and (10) imply

$$\begin{bmatrix} \boldsymbol{\Pi}_{11}\boldsymbol{\beta}_1 \\ \boldsymbol{\Pi}_{21}\boldsymbol{\beta}_1 \end{bmatrix} = \begin{bmatrix} \boldsymbol{\gamma}_1 \\ \mathbf{0} \end{bmatrix} \quad (11)$$

and  $\mathbf{u} = \mathbf{V}_1\boldsymbol{\beta}_1$ , where  $\mathbf{V} = (\mathbf{V}_1, \mathbf{V}_2)$ . The second subvector of (11) corresponds to (4).

It will be convenient to transform the model so that the two sets of exogenous variables are orthogonal. Let

$$\mathbf{A} = \mathbf{Z}'\mathbf{Z} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix} = \begin{bmatrix} \mathbf{Z}'_1\mathbf{Z}_1 & \mathbf{Z}'_1\mathbf{Z}_2 \\ \mathbf{Z}'_2\mathbf{Z}_1 & \mathbf{Z}'_2\mathbf{Z}_2 \end{bmatrix}$$

and

$$\mathbf{Z}_{2.1} = \mathbf{Z}_2 - \mathbf{Z}_1\mathbf{A}_{11}^{-1}\mathbf{A}_{12}, \quad \mathbf{A}_{22.1} = \mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12}.$$

Define  $(\boldsymbol{\Pi}_{11}^*, \boldsymbol{\Pi}_{12}^*) = (\mathbf{I}_{K_1}, \mathbf{A}_{11}^{-1}\mathbf{A}_{12})\boldsymbol{\Pi}$ . Then

$$\mathbf{Z}\boldsymbol{\Pi} = (\mathbf{Z}_1, \mathbf{Z}_{2.1}) \begin{bmatrix} \boldsymbol{\Pi}_{11}\boldsymbol{\beta}_1 \\ \boldsymbol{\Pi}_{21}\boldsymbol{\beta}_1 \end{bmatrix} = (\mathbf{Z}_1, \mathbf{Z}_{2.1})\boldsymbol{\Pi}^*,$$

by defining  $\boldsymbol{\Pi}^*$ . Define also

$$\begin{aligned} \mathbf{A}^* &= \begin{bmatrix} \mathbf{Z}'_1 \\ \mathbf{Z}'_{2.1} \end{bmatrix} [\mathbf{Z}_1, \mathbf{Z}_{2.1}] = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{O} \\ \mathbf{O} & \mathbf{A}_{22.1} \end{bmatrix}, \\ \mathbf{P}^* &= (\mathbf{A}^*)^{-1} \begin{bmatrix} \mathbf{Z}'_1 \\ \mathbf{Z}'_{2.1} \end{bmatrix} \mathbf{Y} = \begin{bmatrix} \mathbf{A}_{11}^{-1}\mathbf{Z}'_1\mathbf{Y} \\ \mathbf{A}_{22.1}^{-1}\mathbf{Z}'_{2.1}\mathbf{Y} \end{bmatrix} = \begin{bmatrix} \mathbf{P}_1^* \\ \mathbf{P}_2^* \end{bmatrix}. \end{aligned}$$

The density of  $\mathbf{P}^*$  is defined from the density of  $\mathbf{P}$  given (5) as

$$\begin{aligned} L(\boldsymbol{\Pi}^*) &= (2\pi)^{-\frac{1}{2}TG_1} \exp \left\{ -\frac{1}{2} \text{tr}(\mathbf{P}_1^* - \boldsymbol{\Pi}_1^*)' \mathbf{A}_{11} (\mathbf{P}_1^* - \boldsymbol{\Pi}_1^*) \right\} \\ &\quad \times (2\pi)^{-\frac{1}{2}TG_2} \exp \left\{ -\frac{1}{2} \text{tr}(\mathbf{P}_2^* - \boldsymbol{\Pi}_2^*)' \mathbf{A}_{22.1} (\mathbf{P}_2^* - \boldsymbol{\Pi}_2^*) \right\}, \end{aligned}$$

where  $\boldsymbol{\Pi}_1^* = (\boldsymbol{\Pi}_{11}^*, \boldsymbol{\Pi}_{12}^*)$  and  $\boldsymbol{\Pi}_2^* = (\boldsymbol{\Pi}_{21}, \boldsymbol{\Pi}_{22})$ . The maximum of the first part of  $L(\boldsymbol{\Pi}^*)$  occurs at  $\boldsymbol{\Pi}_1^* = \mathbf{P}_1^*$  and is  $(2\pi)^{-\frac{1}{2}TG_1}$ . The maximum of the second part with respect to  $\boldsymbol{\Pi}_{22}$  is

$$L(\boldsymbol{\Pi}_{21}) = (2\pi)^{-\frac{1}{2}TG_2} \exp \left\{ -\frac{1}{2} \text{tr}(\mathbf{P}_{21} - \boldsymbol{\Pi}_{21})' \mathbf{A}_{22,1} (\mathbf{P}_{21} - \boldsymbol{\Pi}_{21}) \right\}.$$

This corresponds to (5).

## 5. Distribution of Test Criterion

The matrix

$$\mathbf{G} = \mathbf{P}' \mathbf{A} \mathbf{P}$$

has the noncentral Wishart distribution with  $T-K$  degrees of freedom, covariance matrix  $\boldsymbol{\Omega} = \mathbf{I}_G$ , and the noncentrality parameter

$$\boldsymbol{\Gamma} = \boldsymbol{\Pi}' \mathbf{A} \boldsymbol{\Pi}$$

when  $\mathbf{P}$  has the density (6). The noncentral Wishart distribution was derived by Anderson and Girshick (1944) with correction (1964) for  $\boldsymbol{\Pi}$  of rank 1. The criterion (7) can be written

$$-2 \log \frac{L_{H_0}}{L_{H_1}} = \frac{\boldsymbol{\beta}_0' \mathbf{G} \boldsymbol{\beta}_0}{\boldsymbol{\beta}_0' \boldsymbol{\beta}_0} - \min_{\mathbf{b}} \frac{\mathbf{b}' \mathbf{G} \mathbf{b}}{\mathbf{b}' \mathbf{b}}. \quad (12)$$

Since the criterion is invariant with respect to orthogonal transformations (9), we can write (12) for  $\boldsymbol{\beta}_0 = (1, 0)'$  and

$$\boldsymbol{\Pi} = \boldsymbol{\mu} \boldsymbol{\gamma}' = \boldsymbol{\mu}(0, 1) = (0, \boldsymbol{\mu})$$

and

$$\min_{\mathbf{b}} \frac{\mathbf{b}' \mathbf{G} \mathbf{b}}{\mathbf{b}' \mathbf{b}} = d_1,$$

the smaller root of

$$0 = |\mathbf{G} - d \mathbf{I}_2| = d^2 - (g_{11} + g_{22})d + g_{11}g_{22} - g_{12}^2.$$

Thus (12) is

$$g_{11} - d_1 = \frac{1}{2} \left[ g_{11} - g_{22} + \sqrt{(g_{11} - g_{22})^2 + 4g_{12}^2} \right]. \quad (13)$$

The distribution of (13) can be found from the noncentral Wishart distribution. See Anderson, Kunitomo, and Matsushita (2009).

## References

- ANDERSON, T.W. and KUNITOMO, N. (2007). "On Likelihood Ratio Tests of Structural Coefficients : Anderson-Rubin (1949) revisited," Discussion Paper CIRJE-F-499, Graduate School of Economics, University of Tokyo (<http://www.e.u-tokyo.ac.jp/cirje/research/dp/2007>).
- ANDERSON, T.W., KUNITOMO, N., and MATSUSHITA, Y. (2009). "The Limited Information Maximum Likelihood Estimator as an Angle," Discussion Paper CIRJE-F-619, Graduate School of Economics, University of Tokyo (<http://www.e.u-tokyo.ac.jp/cirje/research/dp/2009>).
- ANDERSON, T.W. and GIRSCHICK, M.A. (1944). "Some extensions of the Wishart distribution," *Annals of Mathematical Statistics*. **15**: 345–357. (Corrections, 1964, **35**: 923–925).
- ANDERSON, T.W. and RUBIN, H. (1949). "Estimation of the Parameters of a Single Equation in a Complete System of Stochastic Equations," *Annals of Mathematical Statistics*. **20**: 46–63.
- ANDERSON, T.W. and RUBIN, H. (1950). "The Asymptotic Properties of Estimates of the Parameters of a Single Equation in a Complete System of Stochastic Equation," *Annals of Mathematical Statistics*. **21**: 570–582.
- ANDREWS, D.W.H., MOREIRA, M.J., and STOCK, J.H. (2006). "Optimal Two-Sided Invariant Similar Tests for Instrumental Variables Regression," *Econometrica*. **74**: 715–752.
- MOREIRA, M. (2003). "A Conditional Likelihood Ratio Test for Structural Models," *Econometrica*. **71**: 1027–1048.

Received June 2009

Revised August 2009