Econometrica, Vol. 68, No. 3 (May, 2000), 715–719

ROBUST WALD TESTS IN SUR SYSTEMS WITH ADDING-UP RESTRICTIONS

By B. Ravikumar, Surajit Ray, and N. Eugene Savin[1]

1. INTRODUCTION

For SUR systems with adding-up restrictions, it is well known that the covariance matrix of disturbances is singular. The usual approach to hypothesis testing in such cases is to construct the relevant test statistics after deleting an equation. A common application of this approach is in the context of complete demand systems where the sum of expenditure shares must equal one. Barten (1969) considered the maximum likelihood estimation of such a system of equations with independent and identical normal disturbance vectors. He proved that the value of the likelihood function, and hence, the maximum likelihood estimates of the parameters are invariant to the equation deleted. This, in turn, implies that the value of the likelihood ratio statistic for testing linear restrictions on the coefficients is invariant to the equation deleted. Similarly, McGuire, Farley, Lucas, and Ring (1968) and Powell (1969) considered the Generalized Least Squares (GLS) estimation of a system of demand equations. Under the assumption that the covariance matrix of the stacked disturbance vector is known, they showed that the GLS estimator and the corresponding quadratic form are invariant to the equation deleted. Estimation and testing have been extended to SUR systems with specific forms of heteroskedasticity and/or autocorrelations; see, for instance, Mandy and Martins-Filho (1993) and Berndt and Savin (1975).

In practice, the likelihood function and/or the covariance matrix of the stacked disturbance vector are usually unknown. Similarly, the functional form of heteroskedasticity and/or autocorrelations is also unknown. In this paper, we consider SUR systems with adding-up restrictions where the same explanatory variables are present in all equations and where heteroskedasticity and/or autocorrelation of unknown forms may be present. For this case, the coefficients are usually estimated by least squares, equation by equation. For testing the typical hypotheses of interest, we show that the robust Wald statistic, i.e., the statistic based on the heteroskedasticity and autocorrelation consistent (HAC) covariance matrix estimator, is invariant to the equation deleted. Our proof of invariance does not rely on parametric assumptions as in Barten (1969) or on knowledge of the covariance matrix as in Powell (1969). Furthermore, the adding-up restrictions we consider are more general than Barten's. As in Powell, the weighted sum of the dependent variables in this paper adds up to one of the explanatory variables, not necessarily a constant. Our proof exploits the properties of generalized inverses and depends only on the existence of first and second moments. It should be noted that even though our robust Wald test is invariant to the equation deleted, it is *not* invariant to nonlinear transformations of the null hypothesis (see Gregory and Veall (1985)).

[1] We gratefully acknowledge the helpful comments of a co-editor and the referees. In particular, we thank an anonymous referee for suggesting a shorter and more elegant proof of our main theorem.

2. SUR MODEL

Consider the following system of equations:

$$y_{it} = \beta_i' x_t + \varepsilon_{it} \qquad (i = 1, 2, \ldots, n;\ t = 1, 2, \ldots, T), \tag{2.1}$$

where y_{it} is a period t observation corresponding to the ith equation, $x_t = (x_{1t}, x_{2t}, \ldots, x_{kt})'$ is a $(k \times 1)$ vector of nonrandom explanatory variables at period t, $\beta_i = (\beta_{i1}, \beta_{i2}, \ldots, \beta_{ik})'$ is a $(k \times 1)$ vector of parameters, and ε_{it} is a period t error corresponding to the ith equation.

The system of equations in (2.1) can be written in the following compact form:

$$y_i = X\beta_i + \varepsilon_i \qquad (i = 1, 2, \ldots, n), \tag{2.2}$$

where y_i is a $(T \times 1)$ vector, X is a $(T \times k)$ matrix of explanatory variables and $\varepsilon_i = (\varepsilon_{i1}, \varepsilon_{i2}, \ldots, \varepsilon_{it})'$ is a $(T \times 1)$ vector of disturbances.

The system of equations in (2.2) can be stacked as

$$y = (I_n \otimes X)\beta + \varepsilon, \tag{2.3}$$

where $y = (y_1, y_2, \ldots, y_n)'$ is an $(nT \times 1)$ column vector of the dependent variables, β is an $(nk \times 1)$ vector of parameters, $\varepsilon = (\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n)'$ is an $(nT \times 1)$ column vector of disturbances, and I_n is an identity matrix of order n. The assumptions for the stacked model can be written as

$$\begin{aligned} &E(\varepsilon) = 0, \\ &E(\varepsilon\varepsilon') = \Sigma. \end{aligned} \tag{2.4}$$

The OLS estimator of β and the variance of this estimator are given by

$$\begin{aligned} &\hat{\beta} = \left(I_n \otimes (X'X)^{-1}X'\right)y, \\ &V(\hat{\beta}) = \left(I_n \otimes (X'X)^{-1}X'\right)\Sigma\left(I_n \otimes X(X'X)^{-1}\right). \end{aligned} \tag{2.5}$$

We estimate $V(\hat{\beta})$ by

$$\hat{V}(\hat{\beta}) = \hat{\Omega} = \sum_{j=-T+1}^{T-1} k\left(\frac{j}{S_T}\right)\hat{\Gamma}(j), \tag{2.6}$$

where

$$\hat{\Gamma}(j) = \begin{cases} \dfrac{1}{T}\sum_{t=j+1}^{T} \hat{\varepsilon}_t \hat{\varepsilon}_{t-j}' \otimes (X'X/T)^{-1} x_t x_{t-j}' (X'X/T)^{-1} & \text{for } j \geq 0, \\ \dfrac{1}{T}\sum_{t=-j+1}^{T} \hat{\varepsilon}_{t+j} \hat{\varepsilon}_t' \otimes (X'X/T)^{-1} x_{t+j} x_t' (X'X/T)^{-1} & \text{for } j < 0, \end{cases}$$

and $k(\cdot)$ is a real-valued kernel, S_T is a band-width parameter, and $\hat{\varepsilon}_t = (\hat{\varepsilon}_{1t}, \hat{\varepsilon}_{2t}, \ldots, \hat{\varepsilon}_{nt})'$ is the estimated disturbance vector at period t. See Andrews (1991) for details.

Adding-Up: The adding-up condition is

$$\omega' y_t = x_{1t} \qquad \forall t, \tag{2.7}$$

where $\omega = (\omega_1, \omega_2, \ldots, \omega_n)'$ is a weight vector such that $\sum_{j=1}^{n} \omega_j = 1$ and $y_t = (y_{1t}, y_{2t}, \ldots, y_{nt})'$. We have assumed without loss of generality that $\omega' y_t$ adds up to the first variable x_{1t}. Equation (2.7) implies the following restrictions on the parameters and disturbances:

$$\text{(2.8)} \qquad \omega'\beta_1 = 1, \qquad \omega'\beta_j = 0 \qquad (j = 2, 3, \ldots, k),$$

and

$$\text{(2.9)} \qquad \omega'\varepsilon_t = 0 \qquad \forall t,$$

where $\beta_j = (\beta_{1j}, \beta_{2j}, \ldots, \beta_{nj})'$ is an $(n \times 1)$ column vector of parameters corresponding to the jth explanatory variable.

Equation (2.9) implies that the ε_{it}'s, $i = 1, 2, \ldots, n$, are linearly dependent at each period t. It can be easily shown that the above adding-up restrictions also hold for LS *estimates* and LS *residuals*, in particular, $\omega'\hat{\varepsilon}_t = 0$, $\forall t$.

3. HYPOTHESIS TESTING

The null hypothesis of interest is

$$\text{(3.1)} \qquad H_0\colon \beta_j = 0,$$

or equivalently H_0: $(I_n \otimes R_j')\beta = 0$, where $I_n \otimes R_j'$ is a selection matrix, and R_j is a k dimensional column vector with 1 in the jth position and zeros elsewhere. The alternative hypothesis is H_1: $\beta_j \neq 0$.

Consequently, the general form of the Wald statistic to test the above hypothesis (3.1) is given by

$$\text{(3.2)} \qquad J = T\hat{\beta}_j'\left[(I_n \otimes R_j')\hat{\Omega}(I_n \otimes R_j)\right]^{-}\hat{\beta}_j,$$

where $[\cdot]^-$ indicates a generalized inverse of $[\cdot]$. Now, we will propose a g-inverse for $[\cdot]$. Let $S_{(k)}$ denote the $(n-1 \times n)$ selector matrix that selects all but the kth element of an n-vector.

PROPOSITION 1: *The matrix* $S_{(k)}'[S_{(k)}(I_n \otimes R_j')\hat{\Omega}(I_n \otimes R_j)S_{(k)}']^{-1}S_{(k)}$ *is a g-inverse for* $(I_n \otimes R_j')\hat{\Omega}(I_n \otimes R_j)$. *Further, it is also a reflexive g-inverse.* (*For details on generalized inverses and reflexive generalized inverses, see Rao and Mitra* (1971, *p.* 14).)

PROOF: If $\omega_i \neq 0$, $i = 1, 2, \ldots, n$, the rows of $S_{(k)}$ are linearly independent and the matrix $S_{(k)}(I_n \otimes R_j')\hat{\Omega}(I_n \otimes R_j)S_{(k)}'$ is invertible. Given invertibility, the first part of Proposition 1 follows immediately from Rao and Mitra (1971, Lemma 2.2.5(c), p. 22). The reflexivity of the g-inverse is trivial. *Q.E.D.*

There is a g-inverse for every k, that is, for every equation deleted. Hence, a question of interest is whether the Wald statistic is invariant to the choice of k. The invariance of the Wald statistic to the equation deleted can be demonstrated by using the following theorem in Rao and Mitra (1971).

THEOREM: *Let A be an $(n \times n)$ matrix and p be an n-vector. Then $p'A^-p$ is invariant for any choice of A^- if $p \in \mathscr{R}(A)$ and $p \in \mathscr{R}(A')$ where $\mathscr{R}(A)$ denotes the range or column space of A.*

To show invariance of the Wald statistic in (3.2), we need to verify $\hat{\beta}_j \in \mathscr{R}(A)$ and $\hat{\beta}_j \in \mathscr{R}(A')$ where $A = (I_n \otimes R_j')\hat{\Omega}(I_n \otimes R_j)$. First, it can be easily seen that A is symmetric. Second, $\omega' A = 0$ implies that rank of A is $(n-1)$. Since A is an $(n \times n)$ matrix with rank $(n-1)$, the columns of A span $(n-1)$ dimensional hyperplanes in $\Re^n$. Further, $\omega' \beta_j = 0$, $j = 2, 3, \ldots, k$, i.e., $\hat{\beta}_j$ lies in an $(n-1)$ dimensional hyperplane in $\Re^n$. Thus, it must be the case that $\hat{\beta}_j$ belongs to the column space of A.

Two variants of the SUR system defined in (2.1) are considered in the remarks below.

REMARK 1: For the classical SUR system, the disturbance vectors are i.i.d. and

$$E(\varepsilon_t \varepsilon_s') = \begin{cases} \Sigma^* & \text{if } t = s, \\ 0 & \text{if } t \neq s, \end{cases}$$

where Σ^* is an $(n \times n)$ contemporaneous covariance matrix. Hence, the Wald statistic (3.2) reduces to

$$J_1 = c^{-1} \hat{\beta}_j^{(k)\prime} (\Sigma^{*(k)})^{-1} \hat{\beta}_j^{(k)},$$

where a superscript "k" denotes that the kth equation is deleted and c is the jth diagonal element of $(X'X)^{-1}$. Under the null hypothesis, J_1 has an asymptotic chi-square distribution with $n-1$ degrees of freedom. In addition, if the disturbance vectors follow a normal distribution, then the exact finite-sample critical values are available; see Anderson (1958).

REMARK 2: Consider the system of equations in (2.1) and suppose that the ε_t's are serially correlated up to lag l and have time-varying covariance matrices (SUR system with heteroskedasticity and autocorrelation):

$$E(\varepsilon_t \varepsilon_s') = \begin{cases} \Sigma_{t,t} & \text{if } t = s, \\ \Sigma_{t,s} & \text{if } |t-s| \leq l, \\ 0 & \text{if } |t-s| > l, \end{cases}$$

where $\Sigma_{t,s}$ is an $(n \times n)$ time dependent lag $(t-s)$ covariance matrix. The Wald statistic for this case can be written

$$J_2 = \hat{\beta}_j^{(k)\prime} \left[\sum_{s=0}^{l} k\left(\frac{s}{S_T}\right) \sum_{t=s}^{T} [q_{t,t-s} \Sigma_{t,t-s}^{(k)} + q_{t-s,t} \Sigma_{t-s,t}^{(k)}] \right]^{-1} \hat{\beta}_j^{(k)},$$

where $q_{t,s}$ is the jth diagonal element of $Q_{t,s} \equiv (X'X)^{-1} x_t x_s' (X'X)^{-1}$. Under the null hypothesis, J_2 has an asymptotic chi-square distribution with $n-1$ degrees of freedom under some regularity conditions. See Hamilton (1994, p. 225).

4. CONCLUDING COMMENTS

We have established the invariance of the robust Wald statistic to the equation deleted in SUR systems with adding-up restrictions where the heteroskedasticity and autocorrelation are of unknown form.

In Ravikumar, Ray, and Savin (1999) we illustrate our results using the Sharpe-Lintner version of the Capital Asset Pricing Model (CAPM). The theory of the CAPM implies that the return on the market portfolio is a weighted sum of returns on the individual assets. Hence, the CAPM is a SUR system with an adding-up restriction. In conventional tests of the CAPM, the null hypothesis is that the vector of intercepts is zero. In our illustration, we use the Wald statistic which takes account of the adding-up restriction.

Department of Economics, University of Iowa, Pappajohn Business Bldg., Iowa City, IA 52242-1000.

Manuscript received February, 1998; final revision received March, 1999.

REFERENCES

ANDERSON, T. W. (1958): *An Introduction to Multivariate Statistical Analysis*. New York: Wiley.

ANDREWS, D. W. K. (1991): "Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation," *Econometrica*, 59, 817–858.

BARTEN, A. P. (1969): "Maximum Likelihood Estimation of a Complete System of Demand Equations," *European Economic Review*, 17–73.

BERNDT, E. R., AND N. E. SAVIN (1975): "Estimation and Hypothesis Testing in Singular Equation Systems with Autoregressive Disturbances," *Econometrica*, 43, 937–957.

GREGORY, A. W., AND M. R. VEALL (1985): "Formulating Wald Tests of Nonlinear Restrictions," *Econometrica*, 53, 1465–1468.

HAMILTON, J. D. (1994): Time Series Analysis. Princeton, NJ: Princeton University Press.

MANDY, D. M., AND C. MARTINS-FILHO (1993): "SUR under Additive Heteroskedasticity: Theory and Share Equation Applications," *Journal of Econometrics*, 58, 315–346.

MCGUIRE, T. W., J. U. FARLEY, R. E. LUCAS, AND L. W. RING (1978): "Estimation and Inference for Linear Models in which Subsets of the Dependent Variable are Constrained," *Journal of the American Statistical Association*, 63, 1201–1213.

POWELL, A. (1969): "Aitken Estimators as a Tool in Allocating Predetermined Aggregates," *Journal of the American Statistical Association*, 64, 913–922.

RAO, C. R., AND S. K. MITRA (1971): *Generalized Inverse of Matrices and Its Applications*. New York: Wiley.

RAVIKUMAR, B., SURAJIT RAY, AND N. E. SAVIN (1999): "CAPM Reconsidered: A Robust Finite Sample Evaluation," Working Paper, Department of Economics, University of Iowa.