A NOTE ON THE DERIVATION OF THEIL'S BLUS RESIDUALS

Gregory C. Chow

ECONOMETRIC RESEARCH PROGRAM
Research Memorandum No. 176
March 1975

The research described in this paper was supported by NSF Grant GS43747X. Paper typed by Alice Furth.

Econometric Research Program
PRINCETON UNIVERSITY
207 Dickinson Hall
Princeton, New Jersey

A NOTE ON THE DERIVATION OF THEIL'S BLUS RESIDUALS

Gregory C. Chow*

In two papers the results of which are summarized in his _Principles of Econometrics_ [1, Chapter 5], H. Theil has proposed a best linear unbiased (BLU) estimator of the residuals in a standard linear regression which has a scalar (S) covariance matrix (of the form $\sigma^2 I$). Theil's proof in the form of seven Lemma's [1, pp. 209-213] and other related theorems appears complicated and can be further motivated. This note provides a simple constructive proof of the BLUS residuals.

To set up the problem, consider n observations of a regression model with k explanatory variables

(1) $$y = X\beta + \varepsilon$$

where $E\varepsilon\varepsilon' = I\sigma^2$. For a linear estimator Cy to be unbiased, we have

(2) $$E(Cy) = EC(X\beta + \varepsilon) = CX\beta = 0 ,$$

which implies

(3) $$CX = 0 .$$

*I would like to thank Donald T.Sant and T. James Trussell for useful comments and to acknowledge financial support from the National Science Foundation (NSF Grant GS43747X).

It follows from (3) [1, p. 207, Theorem 5.5] that the linear estimator can be written in three alternative forms

(4) $$Cy = C\varepsilon = Ce$$

where e is the vector of residuals by the method of least squares

(5) $$e = [I - X(X'X)^{-1}X']\varepsilon \ .$$

In order for $Cy = C\varepsilon$ to have a scalar covariance matrix

(6) $$E(C\varepsilon\varepsilon'C') = CC'\sigma^2 = I\sigma^2 \ ,$$

we require

(7) $$CC' = I \ .$$

The restriction (3), with X having k columns, is a set of k linear restrictions on the columns of C. Therefore, the rank of C cannot exceed $n-k$. We thus let C be an $(n-k) \times n$ matrix and Cy be an estimator of only $n-k$ components of ε. Accordingly Theil partitions model (1) as

(8) $$\begin{bmatrix} y_o \\ y_1 \end{bmatrix} = \begin{bmatrix} X_o \\ X_1 \end{bmatrix} \beta + \begin{bmatrix} \varepsilon_o \\ \varepsilon_1 \end{bmatrix} = \begin{bmatrix} X_o \\ X_1 \end{bmatrix} b + \begin{bmatrix} e_o \\ e_1 \end{bmatrix}$$

with X_o being a nonsingular $k \times k$ matrix and X_1 being an $(n-k) \times k$ matrix. The linear estimator can be written as

(9) $$Ce = C_o e_o + C_1 e_1 ,$$

C_1 being $(n-k)\times(n-k)$. Using $X_o' e_o + X_1' e_1 = 0$, we have $e_o = -(X_1 X_o^{-1})' e_1 = -Z' e_1$, where $Z = X_1 X_o^{-1}$. Further using $C_o X_o + C_1 X_1 = 0$, we have $C_o = -C_1 X_1 X_o^{-1} = -C_1 Z$. (9) thus becomes

(10) $$Ce = (-C_1 Z)(-Z' e_1) + C_1 e_1 = C_1 [I + ZZ'] e_1 .$$

The constraint (7) for a scalar covariance matrix yields

(11) $$C_o C_o' + C_1 C_1' = C_1 ZZ' C_1' + C_1 C_1' = C_1 [I + ZZ'] C_1' = I .$$

The linear estimator is alternatively written as

(12) $$C_o \varepsilon_o + C_1 \varepsilon_1 = -C_1 Z \varepsilon_o + C_1 \varepsilon_1 .$$

The vector of errors in using (12) to estimate ε_1 has a covariance matrix

(13) $$\mathrm{Cov}[-C_1 Z \varepsilon_o + (C_1 - I)\varepsilon_1] = \sigma^2 [C_1 (I + ZZ') C_1' + I - C_1 - C_1'] .$$

To find the matrix C_1 in the estimator (10) which minimizes the trace of (13), following Theil's definition of being "best," subject to the constraint (11) for a scalar covariance matrix, I would propose to form the Lagrangean expression

(14) $$L = tr[C_1(I+ZZ')C_1' + I - C_1 - C_1'] - tr\, M[C_1(I+ZZ')C_1' - I]$$

where M is a symmetric $(n-k)\times(n-k)$ matrix of Lagrange multipliers, and differentiate with respect to C_1. Using the differentiation rule $\partial tr(AB)/\partial tr(BA)/\partial A = B'$, we have

(15) $$\frac{\partial L}{\partial C_1} = 2C_1(I+ZZ') - 2I - 2MC_1(I+ZZ') = 0$$

To solve equations (15) and (11) for the two unknowns C_1 and M, we post-multiply (15) by C_1' and use (11) to obtain

(16) $$M = I - C_1' .$$

Substituting (16) for M in (15), one gets

(17) $$C_1'C_1(I+ZZ') = I .$$

Since C_1 is symmetric because M in (16) is symmetric, (17) implies $C_1^2 = (I+ZZ')^{-1}$ or

(18) $$C_1 = (I+ZZ')^{-1/2}$$

which is the solution. Theil has written the solution in the form of $C_1 = PDP'$ where P is a matrix consisting of the characteristics vectors of $(I+ZZ')$ corresponding to the roots d_i^{-2}, D having diagonal elements d_i and $(I+ZZ') = PD^{-2}P'$. The reader may consult Theil [1] for further treatment of this topic.

REFERENCE

[1] Theil, H: Principles of Econometrics. New York, John Wiley & Sons, Inc., 1971.