연세경제연구 제 VI 권, 제 2 호
Yonsei Economic Studies, Vol. VI, No. 2 (September 1999)

A Note on Partial *F*-tests in a Multivariate Linear Regression Model

Suk Bum Yoon

This note demonstrates that partial *F*-tests may result in different test conclusions, when some degree of multicollinearity exists among the variables of two sets of explanatary variables in a multivariate linear regression model.

I. Introduction

This paper investigates differences in powers of partial F-test in a multivariate linear regression model, when the order of the test procedures is switched, in a case where some degree of multicollinearity exists among the explanatary variables. This aspect has often been overlooked when partial F-tests are applied in order to investigate significance of some of explanatary variables.

Suppose we have

$$y = X\beta + u \tag{1}$$

Professor of Economics, Department of Economics, Yonsei University, Seoul 120-749, Korea. A helpful example supporting the conclusion of this paper was provided by Professor Byung Soo Kim of Yonsei University, and Professor Byung Sam Yoo provided with useful comments. However, remaining errors, if any, are authors.

where y is $n \times 1$ explained variables vector, X $n \times (k+1)$ matrix of explanatary variables, β $(k+1) \times 1$ coefficients vector, and u $n \times 1$ error terms with the following properties

$$u \sim N(0, \sigma^2 I) \tag{2}$$

Ordinary least squares estimators of β, which we denote $\hat{\beta}$, are in this case obtained as

$$\hat{\beta} = (X'X)^{-1}X'y \tag{3}$$

with

$$\hat{\beta} \sim N(\beta, \sigma^2(X'X)^{-1}) \tag{4}$$

Ⅱ. Distributions of Estimators

Suppose that we partition Equation (1) as

$$y = X_1\beta_1 + X_2\beta_2 + u \tag{5}$$

where X_1 $n \times r$ matrix, β_1 $r \times 1$ vector, X_2 $n \times (k - r + 1)$ matrix, and β_2 $(k - r + 1) \times 1$vector, respectively.

The equation (3) may accordingly be rewritten as

$$\hat{\beta} = \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix}$$
$$= \begin{bmatrix} X_1'X_1 & X_1'X_2 \\ X_2'X_1 & X_2'X_2 \end{bmatrix}^{-1} \begin{bmatrix} X_1'y \\ X_2'y \end{bmatrix}$$
$$= \begin{bmatrix} W_{11} & W_{12} \\ W_{21} & W_{22} \end{bmatrix} \begin{bmatrix} X_1'y \\ X_2'y \end{bmatrix}$$

$$= \begin{bmatrix} (X_1' M_{22} X_1)^{-1} X_1' M_{22} y \\ (X_2' X_2)^{-1} X_2' y - (X_2' X_2)^{-1} X_2' X_1 (X_1' M_{22} X_1)^{-1} X_1' M_{22} y \end{bmatrix} \tag{6}$$

where

$$W_{11} = (X_1' M_{22} X_1)^{-1}$$

$$W_{12} = -(X_1' M_{22} X_1)^{-1} X_1' X_2 (X_2' X_2)^{-1}$$

$$W_{21} = W_{12}'$$

$$W_{22} = (X_2' X_2)^{-1} + (X_2' X_2)^{-1} X_2' X_1 (X_1' M_{22} X_1)^{-1} X_1' X_2 (X_2' X_2)^{-1}$$

and

$$M_{22} = I - X_2 (X_2' X_2)^{-1} X_2'$$

And we know that the marginal distributions of $\tilde{\beta}_1$ or $\tilde{\beta}_2$[1] may be obtained, under the null hypothesis that remaining explanatory variables do not significantly contribute to y, as

$$\tilde{\beta}_1 \sim N(\beta_1, \sigma^2 (X_1' X_1)^{-1}) \tag{7}$$

or

$$\tilde{\beta}_2 \sim N(\beta_2, \sigma^2 (X_2' X_2)^{-1}) \tag{8}$$

where

$$\tilde{\beta}_1 = (X_1' X_1)^{-1} X_1' y$$

and

$$\tilde{\beta}_2 = (X_2' X_2)^{-1} X_2' y$$

The conditional distribution of β_1 estimators given β_2 estimators, which we denote $\hat{\beta}_1 \mid \tilde{\beta}_2$ are obtained as

---

1) Dhrymes [1].

$$\hat{\beta}_1 \mid \tilde{\beta}_2 \sim N(\beta_1 + X_1' X_2 (X_2' X_2)^{-1} (\tilde{\beta}_2 - \beta_2), \sigma^2 (X_1' M_{22} X_1)^{-1}) \tag{9}$$

or

$$\hat{\beta}_2 \mid \tilde{\beta}_1 \sim N(\beta_2 + X_2' X_1 (X_1' X_1)^{-1} (\tilde{\beta}_1 - \beta_1), \sigma^2 (X_2' M_{11} X_2)^{-1}) \tag{10}$$

Ⅲ. Partial *F*-tests

In order to investigate whether r variables of X_1 significantly explain y, that is whether $\beta_1 = 0$ or not, the following procedure of establishing a ratio of explained sum of squares (ESS) over unexplained sum of squares (USS) is usually formulated.[2]

$$F = \frac{ESS_1 / r}{USS / (n - k - 1)} \tag{11}$$

where

$$ESS_1 = \hat{\beta}' X' X \hat{\beta} - \tilde{\beta}'_2 X_2' X_2 \tilde{\beta}_2 \tag{12}$$

and

$$USS = (y - X\hat{\beta})' (y - X\hat{\beta}) \tag{13}$$

Further expanding $\hat{\beta}' X' X \hat{\beta}$ in the form of equation (5), we have

$$\begin{aligned} \hat{\beta}' X' X \hat{\beta} = {} & y' X_1 (X_1' M_{22} X_1)^{-1} X_1' y - \\ & y' X_2 (X_2' X_2)^{-1} X_2' X_1 (X_1' M_{22_1} X_1)^{-1} X_1' y + \\ & y' X_2 (X_2' X_2)^{-1} X_2' y + \\ & y' X_2 (X_2' X_2)^{-1} X_2' X_1 (X_1' M_{22} X_1)^{-1} X_1' X_2 (X_2' X_2)^{-1} y - \\ & y' X_1 (X_1' M_{22} X_1)^{-1} X_1' X_2 (X_2' X_2)^{-1} X_2' y \\ = {} & y' [\, I - X_2 (X_2' X_2)^{-1} X_2' \,] X_1 (X_1' M_{22} X_1)^{-1} X_1' y - \end{aligned}$$

2) Any typical econometrics textbook may be cited. For example, refer to Johnston [2].

$$y' [I - X_2' (X_2' X_2)^{-1} X_2] X_1 (X_1' M_{22} X_1)^{-1} X_1' X_2 (X_2' X_2)^{-1} X_2' y +$$
$$y' X_2 (X_2' X_2)^{-1} X_2' y$$
$$= y' M_{22} X_1 (X_1' M_{22} X_1)^{-1} X_1' y -$$
$$y' M_{22} X_1 (X_1' M_{22} X_1)^{-1} X_1' X_2 (X_2' X_2)^{-1} X_2' y +$$
$$y' X_2 (X_2' X_2)^{-1} X_2' y$$
$$= y' M_{22} X_1 (X_1' M_{22} X_1)^{-1} [I - X_2 (X_2' X_2)^{-1} X_2'] y +$$
$$y' X_2 (X_2' X_2)^{-1} X_2' y$$
$$= y' M_{22} X_1 (X_1' M_{22} X_1)^{-1} X_1' M_{22} y + y' X_2 (X_2' X_2)^{-1} X_2' y$$
$$= y' [M_{22} X_1 (X_1' M_{22} X_1)^{-1} X_1' M_{22} + X_2 (X_2' X_2)^{-1} X_2'] y$$

Since $\tilde{\beta}_2$ is defined as Equation (8), Equation (12) is reduced to Equation (14) as

$$ESS_1 = y' M_{22} X_1 (X_1' M_{22} X_1)^{-1} X_1' M_{22} y \tag{14}$$

Now, we reverse the order and obtain ESS_2 as Equation (15)

$$ESS_2 = y' M_{11} X_2 (X_2' M_{11} X_2)^{-1} X_2 M_{11} y \tag{15}$$

Suppose further that ESS_i are obtained from marginal distributions of the estimators $\tilde{\beta}_i$ as

$$ESS_{1M} = y' X_1 (X_1' X_1)^{-1} X_1' y \tag{16}$$

and

$$ESS_{2M} = y' X_2 (X_2' X_2)^{-1} X_2' y \tag{17}$$

It is clear that

$$ESS_1 = ESS_{1M}$$

when $X_1' X_2 = 0$, which means that X_1 and X_2 are orthogonal and there dose not exist any collinearity between elements of X_1 and those of X_2. As expected, in this case the conditional distributions are exactly the same with the respective marginal distributions. However, in most cases X_1 and X_2 are not orthogonal, and $ESS_1 \neq ESS_{1M}$ and $ESS_2 \neq ESS_{2M}$ hold.

Ⅳ. Concluding Remarks

In applying a partial F-test in a multivariate linear regression in order to test joint significance of some of explanatary variables, the order of estimating $\hat{\beta}_i$ and $\tilde{\beta}_i$ ($i = 1, 2$) would produce different results of test statistics. In sum, $\frac{ESS_1 / r}{USS / (n-k-1)}$ and $\frac{ESS_2 / (k-r+1)}{USS / (n-k-1)}$ may both fail to support the significance, while $\frac{ESS_{1M} / r}{USS / (n-k-1)}$ and $\frac{ESS_{2M} / (k-r+1)}{USS / (n-k-1)}$ both support the significance or vice versa, when there exists some degree of muticollinearity between X_1 and X_2. Furthermore, because of Equation (9) and (10), the test statistics may be treated as non-central F.

❖ *REFERENCES* ❖

1. Dhrymes, Phoebus J., *Econometrics : Statistical Foundations and Applications*, New York : Harpers and Row, 1970, pp. 15~19.
2. Johnston, J., *Econometic Methods*, Second Edition, New York : McGraw-Hill, 1974.