Dec 1973

G. R. Steele

Structural Breaks in Linear Regression

Since Gujarati [1, 2], the use of dummy variables as an alternative to the Chow Test for identifying structural breaks in linear regressions has become widely used. It is recognised that these methods give identical F test statistics for the null hypothesis of no structural breaks. However, the writer is unaware of any simple textbook proof of the general case. It is hoped that the following will be useful to teachers of undergraduate courses:

The Chow Test

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

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} X_1 & 0 \\ 0 & X_2 \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} + \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \tag{2}$$

$$\text{i.e. } y = X^* \quad \beta + u$$

where y, u are (T x 1) column vectors; X is a (T x k) matrix; X_1 is a (T_1 x k) matrix; X_2 is a (T_2 x k) matrix; and the remaining vectors are similarly partioned.

Equations (1) and (2) are estimated separately to obtain the respective sums of squares of the residuals, i.e. $\hat{u}'\hat{u}$ and $u^{*\prime}u^*$.

Under H_o: $\beta_1 = \beta_2$

$$\frac{\hat{u}'\hat{u} - u^{*\prime}u^*}{u^{*\prime}u^*} \cdot \frac{T - 2k}{k} \sim F\ (k, T - 2k)$$

The Dummy Variable Approach

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

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} X_1 & 0 \\ X_2 & X_2 \end{pmatrix} \begin{pmatrix} \gamma_1 \\ \gamma_2 \end{pmatrix} + \begin{pmatrix} \eta_1 \\ \eta_2 \end{pmatrix} \tag{3}$$

$$\text{i.e. } y = X^{**} \quad \gamma + \eta$$

When equations (1) and (3) are estimated separately, we obtain respectively $\hat{u}'\hat{u}$ and $\eta'\eta$.

Under H_o: $\gamma_2 = 0$

$$\frac{\hat{u}'\hat{u} - \eta'\eta}{\eta'\eta} \cdot \frac{T - 2k}{k} \sim F\ (k, T - 2k)$$

Clearly the two tests are equivalent if it can be shown that:

$$\eta'\eta = u^{*\prime}u^*$$

Now

$$u^{*\prime}u^* = u'M^*u$$

$$\eta'\eta = u'M^{**}u$$

where

$$M^* = \left[I - X^*(X^{*\prime}X^*)^{-1} X^{*\prime}\right]$$

$$M^{**} = \left[I - X^{**}(X^{**\prime}X^{**})^{-1} X^{**\prime}\right]$$

Thus it is sufficient to show that

$$M^* = M^{**}$$

The proof makes use of the following relationship:

$$\begin{pmatrix} X_1 & 0 \\ 0 & X_2 \end{pmatrix} \begin{pmatrix} I & 0 \\ I & I \end{pmatrix} = \begin{pmatrix} X_1 & 0 \\ X_2 & X_2 \end{pmatrix}$$

i.e. $X^* \quad I^* \quad = \quad X^{**}$

Proof:

$$\begin{aligned} M^{**} &= \left[I - X^{**}(X^{**\prime}X^{**})^{-1} X^{**\prime}\right] \\ &= \left[I - X^*I^*(\{X^*I^*\}'X^*I^*)^{-1} \{X^*I^*\}'\right] \\ &= \left[I - X^*I^*(I^{*\prime}X^{*\prime}X^*I^*)^{-1} I^{*\prime}X^{*\prime}\right] \\ &= \left[I - X^*I^*(I^*)^{-1} (X^*)^{-1} (X^{*\prime})^{-1} (I^{*\prime})^{-1} I^{*\prime}X^{*\prime}\right] \\ &= \left[I - X^*(X^*)^{-1} (X^{*\prime})^{-1} X^{*\prime}\right] \\ &= \left[I - X^*(X^{*\prime}X^*)^{-1} X^{*\prime}\right] \\ &= M^* \end{aligned}$$

Q.E.D.

References

1. Gujarati, Domar, (1970) "Use of Dummy Variables in Testing for Equality between sets of Coefficients in Two Linear Regressions: A Note". The American Statistician Vol.24, No.1, February 1970, pp.50-52

2. Gujarati, Domar, (1970) "Use of Dummy Variables in Testing for Equality between sets of Coefficients in Linearf Regressions: A Generalisation". The American Statistician Vol.24, No.5, December 1970, pp.18-27.