Econ 203B: Single Equation Models

The F-Statistic

Michael Powell
Department of Economics, UCLA
February 11th, 2006

For all that follows, assume that all the assumptions of the CNLR model hold. The purpose of these notes is to provide rigorous links among the three characterizations of the F statistic that show up frequently in econometrics:

$$\begin{aligned} F_0 &= \frac{1}{p}\left(\Gamma\hat{\beta}-\gamma_0\right)'\left(\hat{\sigma}^2\Gamma\left(X'X\right)^{-1}\Gamma'\right)^{-1}\left(\Gamma\hat{\beta}-\gamma_0\right) \\ F_0 &= \frac{\left(\hat{\varepsilon}_R'\hat{\varepsilon}_R-\hat{\varepsilon}_{UR}'\hat{\varepsilon}_{UR}\right)/p}{\hat{\varepsilon}_{UR}'\hat{\varepsilon}_{UR}/\left(n-k\right)} \\ F_0 &= \frac{\left(R_{UR}^2-R_R^2\right)/p}{\left(1-R_{UR}^2\right)/\left(n-k\right)} \end{aligned}$$

Definition 1 *Let $X\sim\chi^2\left(n_1\right)$, $Y\sim\chi^2\left(n_2\right)$ be independent random variables. Then we say that the random variable $F\equiv\frac{X/n_1}{Y/n_2}$ has an* ***F distribution*** *with numerator degrees of freedom n_1 and denominator degrees of freedom n_2. That is, $F\sim F\left(n_1,n_2\right)$.*

Proposition 2 *Let $\hat{\beta}$ be the OLS estimator for the CNLR model and let $\hat{\sigma}^2=\frac{\hat{\varepsilon}'\hat{\varepsilon}}{n-k}$. Then, $F_0\equiv\frac{1}{p}\left(\Gamma\hat{\beta}-\Gamma\beta\right)'\left(\hat{\sigma}^2\Gamma\left(X'X\right)^{-1}\Gamma'\right)^{-1}\left(\Gamma\hat{\beta}-\Gamma\beta\right)\sim F\left(p,n-k\right)$.*

Proof. Recall that

$$W=\left(\Gamma\hat{\beta}-\Gamma\beta\right)'\left(\sigma^2\Gamma\left(X'X\right)^{-1}\Gamma'\right)^{-1}\left(\Gamma\hat{\beta}-\Gamma\beta\right)\sim\chi^2\left(p\right)$$

And

$$\frac{\hat{\varepsilon}'\hat{\varepsilon}}{\sigma^2}\sim\chi^2\left(n-k\right)$$

This gives us

$$\begin{aligned} F &= \frac{1}{p}\left(\Gamma\hat{\beta}-\Gamma\beta\right)'\left(\hat{\sigma}^2\Gamma\left(X'X\right)^{-1}\Gamma'\right)^{-1}\left(\Gamma\hat{\beta}-\Gamma\beta\right) \\ &= \frac{1}{p}\frac{1}{\hat{\sigma}^2}\frac{\sigma^2}{\sigma^2}\left(\Gamma\hat{\beta}-\Gamma\beta\right)'\left(\Gamma\left(X'X\right)^{-1}\Gamma'\right)^{-1}\left(\Gamma\hat{\beta}-\Gamma\beta\right) \\ &= \frac{\left(\Gamma\hat{\beta}-\Gamma\beta\right)'\left(\sigma^2\Gamma\left(X'X\right)\Gamma'\right)^{-1}\left(\Gamma\hat{\beta}-\Gamma\beta\right)/p}{\frac{\hat{\sigma}^2}{\sigma^2}} \\ &= \frac{\left(\Gamma\hat{\beta}-\Gamma\beta\right)'\left(\sigma^2\Gamma\left(X'X\right)^{-1}\Gamma'\right)^{-1}\left(\Gamma\hat{\beta}-\Gamma\beta\right)/p}{\left(\frac{\hat{\varepsilon}'\hat{\varepsilon}}{\sigma^2}\right)/\left(n-k\right)} \end{aligned}$$

Therefore, $F\sim F\left(p,n-k\right)$. ■

There are two alternative characterizations of the F statistic, one involving sums of squared residuals, the other involving R^2 values. In order to derive the first one, it is necessary to characterize the estimator of the restricted model.

Proposition 3 *Consider the following two models (resp. the unrestricted and the restricted models):*

$$\begin{aligned} Y &= X\beta + \varepsilon & (1) \\ Y &= X\beta + \varepsilon,\, \Gamma\beta = \gamma_0 & (2) \end{aligned}$$

Let $\hat{\beta}_{UR}$ be the OLS estimator for (1) and $\hat{\beta}_R$ the OLS estimator for (2). Then $\hat{\beta}_R = \hat{\beta}_{UR} - (X'X)^{-1}\Gamma'\left(\Gamma(X'X)^{-1}\Gamma'\right)^{-1}\left(\Gamma\hat{\beta}_{UR} - \gamma_0\right)$.

Proof. By definition,

$$\begin{aligned} \hat{\beta}_R &= \arg\min_{b:\Gamma b = \gamma_0} (Y - Xb)'(Y - Xb) \\ &= \arg\min_{b:\Gamma b = \gamma_0} \frac{1}{2}(Y - Xb)'(Y - Xb) \end{aligned}$$

The Lagragian for this problem is:

$$\begin{aligned} L(b,\lambda) &= \frac{1}{2}(Y - Xb)'(Y - Xb) + \lambda(\Gamma b - \gamma_0) \\ &= \frac{1}{2}(Y'Y - b'X'Y - Y'Xb + b'X'Xb) + \lambda(\Gamma b - \gamma_0) \\ &= \frac{1}{2}(Y'Y - 2b'X'Y + b'X'Xb) + \lambda(\Gamma b - \gamma_0) \end{aligned}$$

Taking first order conditions, we have:

$$\begin{aligned} (b) &: \frac{1}{2}\left(-2X'Y + 2X'X\hat{\beta}_R\right) + \Gamma'\lambda' = 0 & (3) \\ (\lambda) &: \Gamma\hat{\beta}_R - \gamma_0 = 0 & (4) \end{aligned}$$

Rearranging (3) and solving for $\hat{\beta}_R$, we have:

$$-X'Y + X'X\hat{\beta}_R + \Gamma'\lambda' = 0$$

$$X'X\hat{\beta}_R = X'Y - \Gamma'\lambda'$$

Or

$$\hat{\beta}_R = (X'X)^{-1}X'Y - (X'X)^{-1}\Gamma'\lambda' = \hat{\beta}_{UR} - (X'X)^{-1}\Gamma'\lambda' \quad (5)$$

Plugging this into (4) to solve for λ'

$$0 = \Gamma\hat{\beta}_R - \gamma_0 = \Gamma\hat{\beta}_{UR} - \Gamma(X'X)^{-1}\Gamma'\lambda' - \gamma_0$$

$$\begin{aligned} \Gamma(X'X)^{-1}\Gamma'\lambda' &= \Gamma\hat{\beta}_{UR} - \gamma_0 \\ \lambda' &= \left(\Gamma(X'X)^{-1}\Gamma'\right)^{-1}\left(\Gamma\hat{\beta}_{UR} - \gamma_0\right) \end{aligned}$$

Substituting this back into (5):

$$\hat{\beta}_R = \hat{\beta}_{UR} - (X'X)^{-1}\Gamma'\left(\Gamma(X'X)^{-1}\Gamma'\right)^{-1}\left(\Gamma\hat{\beta}_{UR} - \gamma_0\right)$$

Which is the desired result. ■

Proposition 4 $F = \frac{\left(\hat{\varepsilon}_R'\hat{\varepsilon}_R - \hat{\varepsilon}_{UR}'\hat{\varepsilon}_{UR}\right)/p}{\hat{\varepsilon}_{UR}'\hat{\varepsilon}_{UR}/(n-k)}$

Proof. Let $\hat{\varepsilon}_{UR} \equiv Y - X\hat{\beta}_{UR}$ and $\hat{\varepsilon}_R = Y - X\hat{\beta}_R$. Then,

$$\hat{\varepsilon}_R - \hat{\varepsilon}_{UR} = Y - X\hat{\beta}_R - \left(Y - X\hat{\beta}_{UR}\right) = X\left(\hat{\beta}_{UR} - \hat{\beta}_R\right)$$

It will prove to be fruitful to first derive an expression for $\hat{\varepsilon}_R'\hat{\varepsilon}_R - \hat{\varepsilon}_{UR}'\hat{\varepsilon}_{UR}$.

$$\begin{aligned}(\hat{\varepsilon}_R - \hat{\varepsilon}_{UR})'(\hat{\varepsilon}_R - \hat{\varepsilon}_{UR}) &= \hat{\varepsilon}_R'\hat{\varepsilon}_R - \hat{\varepsilon}_{UR}'\hat{\varepsilon}_R - \hat{\varepsilon}_R'\hat{\varepsilon}_{UR} + \hat{\varepsilon}_{UR}'\hat{\varepsilon}_{UR} \\ &= \hat{\varepsilon}_R'\hat{\varepsilon}_R - 2\hat{\varepsilon}_{UR}'\hat{\varepsilon}_R + \hat{\varepsilon}_{UR}'\hat{\varepsilon}_{UR}\end{aligned}$$

Since $\hat{\varepsilon}_{UR}'\hat{\varepsilon}_R$ is a scalar, $\hat{\varepsilon}_{UR}'\hat{\varepsilon}_R = \left(\hat{\varepsilon}_{UR}'\hat{\varepsilon}_R\right)' = \hat{\varepsilon}_R\hat{\varepsilon}_{UR}$. Rearranging, we have:

$$\begin{aligned}\hat{\varepsilon}_R'\hat{\varepsilon}_R - \hat{\varepsilon}_{UR}'\hat{\varepsilon}_{UR} &= (\hat{\varepsilon}_R - \hat{\varepsilon}_{UR})'(\hat{\varepsilon}_R - \hat{\varepsilon}_{UR}) + 2\hat{\varepsilon}_{UR}'\hat{\varepsilon}_R - 2\hat{\varepsilon}_{UR}'\hat{\varepsilon}_{UR} \\ &= (\hat{\varepsilon}_R - \hat{\varepsilon}_{UR})'(\hat{\varepsilon}_R - \hat{\varepsilon}_{UR}) + 2\varepsilon_{UR}'(\hat{\varepsilon}_R - \hat{\varepsilon}_{UR}) \\ &= (\hat{\varepsilon}_R - \hat{\varepsilon}_{UR})'(\hat{\varepsilon}_R - \hat{\varepsilon}_{UR}) + 2\underbrace{\varepsilon_{UR}'X}_{=0}\left(\hat{\beta}_{UR} - \hat{\beta}_R\right) \\ &= \left(\hat{\beta}_{UR} - \hat{\beta}_R\right)'X'X\left(\hat{\beta}_{UR} - \hat{\beta}_R\right) \qquad (6)\end{aligned}$$

If we rearrange the result of proposition (3), we have

$$\hat{\beta}_{UR} - \hat{\beta}_R = (X'X)^{-1}\Gamma'\left(\Gamma(X'X)^{-1}\Gamma'\right)^{-1}\left(\Gamma\hat{\beta}_{UR} - \gamma_0\right)$$

Substituting this into (6) gives us:

$$\begin{aligned}\hat{\varepsilon}_R'\hat{\varepsilon}_R - \hat{\varepsilon}_{UR}'\hat{\varepsilon}_{UR} &= \left(\Gamma\hat{\beta}_{UR} - \gamma_0\right)'\left(\Gamma(X'X)^{-1}\Gamma'\right)^{-1}\Gamma(X'X)^{-1} \\ &\quad \cdot (X'X)(X'X)^{-1}\Gamma'\left(\Gamma(X'X)^{-1}\Gamma'\right)^{-1}\left(\Gamma\hat{\beta}_{UR} - \gamma_0\right) \\ &= \left(\Gamma\hat{\beta}_{UR} - \gamma_0\right)'\left(\Gamma(X'X)^{-1}\Gamma'\right)^{-1}\Gamma(X'X)^{-1}\Gamma'\left(\Gamma(X'X)^{-1}\Gamma'\right)^{-1}\left(\Gamma\hat{\beta}_{UR} - \gamma_0\right) \\ &= \left(\Gamma\hat{\beta}_{UR} - \gamma_0\right)'\left(\Gamma(X'X)^{-1}\Gamma'\right)^{-1}\left(\Gamma\hat{\beta}_{UR} - \gamma_0\right)\end{aligned}$$

Recall that

$$F = \frac{1}{p}\left(\Gamma\hat{\beta}_{UR} - \gamma_0\right)'\left(\hat{\sigma}^2\Gamma(X'X)^{-1}\Gamma'\right)^{-1}\left(\Gamma\hat{\beta}_{UR} - \gamma_0\right)$$

And therefore,

$$\hat{\varepsilon}_R'\hat{\varepsilon}_R - \hat{\varepsilon}_{UR}'\hat{\varepsilon}_{UR} = p\hat{\sigma}^2F$$

Rearranging, and substituting in $\hat{\sigma}^2 = \frac{\hat{\varepsilon}_{UR}'\hat{\varepsilon}_{UR}}{n-k}$, we have:

$$F = \frac{\left(\hat{\varepsilon}_R'\hat{\varepsilon}_R - \hat{\varepsilon}_{UR}'\hat{\varepsilon}_{UR}\right)/p}{\hat{\varepsilon}_{UR}'\hat{\varepsilon}_{UR}/(n-k)}$$

Which is the desired result. ∎

Proposition 5 $F = \frac{\left(R_{UR}^2 - R_R^2\right)/p}{\left(1 - R_{UR}^2\right)/(n-k)}$

Proof. Recall the formulae for R_{UR}^2 and R_R^2:

$$\begin{aligned}R_{UR}^2 &= 1 - \frac{\hat{\varepsilon}_{UR}'\hat{\varepsilon}_{UR}}{Y'M^0Y} \\ R_R^2 &= 1 - \frac{\hat{\varepsilon}_R'\hat{\varepsilon}_R}{Y'M^0Y}\end{aligned}$$

This gives us:

$$
\begin{aligned}
\frac{\left(R_{UR}^2 - R_R^2\right)/p}{\left(1 - R_{UR}^2\right)/(n-k)} &= \frac{\left(1 - \frac{\hat{\varepsilon}'_{UR}\hat{\varepsilon}_{UR}}{Y'M^0Y} - \left(1 - \frac{\hat{\varepsilon}'_{R}\hat{\varepsilon}_{R}}{Y'M^0Y}\right)\right)/p}{\left(1 - \left(1 - \frac{\hat{\varepsilon}'_{UR}\hat{\varepsilon}_{UR}}{Y'M^0Y}\right)\right)/(n-k)} \\
&= \frac{\left(\frac{\hat{\varepsilon}'_{R}\hat{\varepsilon}_{R} - \hat{\varepsilon}'_{UR}\hat{\varepsilon}_{UR}}{Y'M^0Y}\right)/p}{\left(\frac{\hat{\varepsilon}'_{UR}\hat{\varepsilon}_{UR}}{Y'M^0Y}\right)/(n-k)} \\
&= \frac{\left(\hat{\varepsilon}'_{R}\hat{\varepsilon}_{R} - \hat{\varepsilon}'_{UR}\hat{\varepsilon}_{UR}\right)/p}{\hat{\varepsilon}'_{UR}\hat{\varepsilon}_{UR}/(n-k)}
\end{aligned}
$$

Which, by proposition (4) establishes the result. ■