Frisch-Waugh-Lovell Theorem

Research Module Econometrics and Statistics,
University of Bonn

November 4, 2019

Outline:

1. Context
2. Proof
3. Conclusion

Context

- Interested in $\mathbf{X}_2$'s effect on $\mathbf{y}$ in bi-variate linear model

Under standard assumptions, classical linear model given by

$$\mathbf{y} = \mathbf{X}_1\mathbf{b}_1 + \mathbf{X}_2\mathbf{b}_2 + \hat{\varepsilon}$$

Context

- Interested in $\mathbf{X}_2$'s effect on $\mathbf{y}$ in bi-variate linear model

Under standard assumptions, classical linear model given by

$$\mathbf{y} = \mathbf{X}_1\mathbf{b}_1 + \mathbf{X}_2\mathbf{b}_2 + \hat{\varepsilon}$$

- Frisch and Waugh (1933) offer alternative procedure
- Two-step procedure:
  1. partial out effect of $\mathbf{X}_1$ on both $\mathbf{X}_2$ and $\mathbf{y}$ separately
  2. regress residuals from $\mathbf{y}$ on $\mathbf{X}_1$ onto residuals from $\mathbf{X}_2$ on $\mathbf{X}_1$

  ... inducing the following (partitioned) model

$$\mathbf{y} - \hat{\mathbf{y}} = (\mathbf{X}_2 - \hat{\mathbf{X}}_2)\hat{\boldsymbol{\beta}}_2 + \hat{\mathbf{v}}$$

with $\hat{\mathbf{y}}$ and $\hat{\mathbf{X}}_2$ depicting fitted values obtained from step 1.

Frisch-Waugh-Lovell theorem

- Rewriting the partitioned model to

$$\mathbf{M}_1\mathbf{y} = \mathbf{M}_1\mathbf{X}_2\hat{\boldsymbol{\beta}}_2 + \hat{\mathbf{v}}$$

with $\mathbf{M}_1 \equiv \mathbf{I}_n - \mathbf{X}_1(\mathbf{X}_1'\mathbf{X}_1)^{-1}\mathbf{X}_1' \equiv \mathbf{I}_n - \mathbf{P}_1$

Frisch-Waugh-Lovell theorem

- Rewriting the partitioned model to

$$\mathbf{M}_1\mathbf{y} = \mathbf{M}_1\mathbf{X}_2\hat{\boldsymbol{\beta}}_2 + \hat{\mathbf{v}}$$

with $\mathbf{M}_1 \equiv \mathbf{I}_n - \mathbf{X}_1(\mathbf{X}_1'\mathbf{X}_1)^{-1}\mathbf{X}_1' \equiv \mathbf{I}_n - \mathbf{P}_1$

- Reminding ourselves of the classical linear model

$$\mathbf{y} = \mathbf{X}_1\mathbf{b}_1 + \mathbf{X}_2\mathbf{b}_2 + \hat{\boldsymbol{\varepsilon}}$$

brings us to *Proposition 3.1.5*

Proposition 3.1.5 (Frisch-Waugh-Lovell theorem)

For the two underlying models, the following is true: $\hat{\boldsymbol{\beta}}_2 = \mathbf{b}_2$ and $\hat{\boldsymbol{\varepsilon}} = \hat{\mathbf{v}}$.

Proof

- First, prove that $\hat{\boldsymbol{\beta}}_2 = \mathbf{b}_2$
- OLS estimators are given by

$$\begin{aligned}
\hat{\boldsymbol{\beta}}_2 &= \left((\mathbf{M}_1\mathbf{X}_2)'(\mathbf{M}_1\mathbf{X}_2)\right)^{-1}(\mathbf{M}_1\mathbf{X}_2)'\mathbf{M}_1\mathbf{y} \\
&= \left(\mathbf{X}_2'\mathbf{M}_1\mathbf{X}_2\right)^{-1}\mathbf{X}_2'\mathbf{M}_1\mathbf{y} \\
\mathbf{b} &= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}
\end{aligned}$$

Proof

- First, prove that $\hat{\boldsymbol{\beta}}_2 = \mathbf{b}_2$
- OLS estimators are given by

$$\hat{\boldsymbol{\beta}}_2 = \left((\mathbf{M}_1\mathbf{X}_2)'(\mathbf{M}_1\mathbf{X}_2)\right)^{-1}(\mathbf{M}_1\mathbf{X}_2)'\mathbf{M}_1\mathbf{y}$$
$$= \left(\mathbf{X}_2'\mathbf{M}_1\mathbf{X}_2\right)^{-1}\mathbf{X}_2'\mathbf{M}_1\mathbf{y}$$
$$\mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$$

- Rewrite OLS estimator of classical linear model

$$\mathbf{X}'\mathbf{X}\mathbf{b} = \mathbf{X}'\mathbf{y}$$

$$\begin{pmatrix} \mathbf{X}_1'\mathbf{X}_1 & \mathbf{X}_1'\mathbf{X}_2 \\ \mathbf{X}_2'\mathbf{X}_1 & \mathbf{X}_2'\mathbf{X}_2 \end{pmatrix} \begin{pmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{X}_1'\mathbf{y} \\ \mathbf{X}_2'\mathbf{y} \end{pmatrix}$$

- System of two equations and two unknowns

Proof cont'd

- Want to solve for $\mathbf{b}_2$
- In first equation write $\mathbf{b}_1$ as a function of $\mathbf{b}_2$

$$\mathbf{b}_1 = \left(\mathbf{X}_1'\mathbf{X}_1\right)^{-1}\left(\mathbf{X}_1'\mathbf{y} - \mathbf{X}_1'\mathbf{X}_2\mathbf{b}_2\right)$$

- Plug this into second equation and obtain

$$\mathbf{b}_2 = \left(\mathbf{X}_2'\mathbf{M}_1\mathbf{X}_2\right)^{-1}\mathbf{X}_2'\mathbf{M}_1\mathbf{y}$$

- Remember OLS estimator of partitioned model

$$\hat{\boldsymbol{\beta}}_2 = \left(\mathbf{X}_2'\mathbf{M}_1\mathbf{X}_2\right)^{-1}\mathbf{X}_2'\mathbf{M}_1\mathbf{y}$$

Proof cont'd

- Now, want to show that $\hat{\boldsymbol{\varepsilon}} = \hat{\mathbf{v}}$
- From partitioned model we have

$$\hat{\mathbf{v}} = \mathbf{M}_1\mathbf{y} - \mathbf{M}_1\mathbf{X}_2\hat{\boldsymbol{\beta}}_2$$

$$\begin{aligned}
\hat{\boldsymbol{\varepsilon}} &= \mathbf{y} - \mathbf{X}_1\mathbf{b}_1 - \mathbf{X}_2\mathbf{b}_2 \\
&= \mathbf{y} - \mathbf{X}_1\left(\mathbf{X}_1'\mathbf{X}_1\right)^{-1}\left(\mathbf{X}_1'\mathbf{y} - \mathbf{X}_1'\mathbf{X}_2\mathbf{b}_2\right) - \mathbf{X}_2\mathbf{b}_2 \\
&= \mathbf{y} - \mathbf{P}_1\mathbf{y} - \left(\mathbf{X}_2\mathbf{b}_2 - \mathbf{P}_1\mathbf{X}_2\mathbf{b}_2\right) \\
&= \mathbf{M}_1\mathbf{y} - \mathbf{M}_1\mathbf{X}_2\mathbf{b}_2 \\
&= \hat{\mathbf{v}}
\end{aligned}$$

Conclusion

- Basic intuition

A multivariate model with K regressors can be broken into K different bi-variate models

- Usefulness of the theorem

Sheds light on the *machinery* of multivariate OLS
Examples: understanding the problems of multicollinearity and omitted variable bias

Conclusion

- A classical application: de-trending a time series

Time series with K explanatory variables and a linear time trend, $\mathbf{t} = \mathbf{1}, \mathbf{2}, \ldots$:

$$\mathbf{y}_t = \mathbf{b}_0 + \mathbf{dt} + \mathbf{X}_{1t}\mathbf{b}_1 + \mathbf{X}_{2t}\mathbf{b}_2 + \ldots + \mathbf{X}_{kt}\mathbf{b}_k + \hat{\varepsilon}_t$$

- Alternative to fitting by least squares:
  1. partial out effect of time trend on $\mathbf{X}_{it}$ and $\mathbf{y}_t$ by regressing each on the time variable
  2. use the residuals from these least squares regressions to calculate the de-trended variables
  3. run the de-trended regression