Econometricks: Short guides to econometrics

Trick 05: Simplifying Linear Regressions using Frisch-Waugh-Lowell

Davud Rostam-Afschar (Uni Mannheim)

Content

1. Frisch-Waugh-Lovell theorem in equation algebra

2. Projection and residual maker matrices

3. Frisch-Waugh-Lovell theorem in matrix algebra

Regress y_i on two explanatory variables, where x_i^2 is the variable of interest and x_i^1 (or further variables) are not of interest.

$$y_i = \beta_0 + \beta_2 x_i^2 + \beta_1 x_i^1 + \varepsilon_i.$$

Surprising and useful result:

- We can obtain **exactly the same** coefficients and residuals from a regression of two demeaned variables

$$\tilde{y}_i = \beta_0 + \beta_2 \tilde{x}_i^2 + \varepsilon_i.$$

- We can obtain **exactly the same** coefficient and residuals from a regression of two residualized variables

$$\varepsilon_i^y = \beta_2 \varepsilon_i^2 + \varepsilon_i.$$

Why is the decomposition useful?

Allows breaking a multivariate model with K independent variables into K bivariate models.

- Relationship between two variables from a multivariate model can be shown in a two-dimensional scatter plot
- Absorbs fixed effects to reduce computation time (see reghdfe for Stata)
- Allows to separate variability between the regressors (multicollinearity) and between the residualized variable $\tilde{x}_i^2$ and the dependent variable y_i.
- Understand biases in multivariate models tractably.

How to decompose y_i and x_i^2?

Partial out x_i^1 from y_i and from x_i^2.

- Regress x_i^2 on all x_i^1 and get residuals ε_i^2:

$$x_i^2 = \gamma_0 + \gamma_1 x_i^1 + \varepsilon_i^2,$$

  this implies $Cov(x_i^1, \varepsilon_i^2) = 0$,
- Regress y_i on all x_i^1 and get residuals ε_i^y:

$$y_i = \delta_0 + \delta_1 x_i^1 + \varepsilon_i^y.$$

  This implies $Cov(x_i^1, \varepsilon_i^y) = 0$.

From the residuals and the constants γ_0 and δ_0 generate

- $\tilde{x}_i^2 = \gamma_0 + \varepsilon_i^2$,
- $\tilde{y}_i = \delta_0 + \varepsilon_i^y$.

Finally,

$$\tilde{y}_i = \tilde{\beta}_0 + \tilde{\beta}_1 \tilde{x}_i^2 + \tilde{\varepsilon}_i = \beta_0 + \beta_2 \tilde{x}_i^2 + \varepsilon_i.$$

Decomposition theorem

Theorem

Decomposition theorem. *For multivariate regressions and detrended regressions, e.g.,*

$$y_i = \beta_0 + \beta_2 x_i^2 + \beta_1 x_i^1 + \varepsilon_i,$$

$$\tilde{y}_i = \tilde{\beta}_0 + \tilde{\beta}_1 \tilde{x}_i^2 + \tilde{\varepsilon}_i,$$

the same regression coefficients will be obtained with any non-empty subset of the explanatory variables, such that

$$\tilde{\beta}_1 = \beta_2 \ \text{ and also } \ \tilde{\varepsilon}_i = \varepsilon_i.$$

Examining either set of residuals will convey precisely the same information about the properties of the unobservable stochastic disturbances.

Detrended variables

Show that

$$
\begin{aligned}
y_i &= \beta_0 + \beta_2 x_i^2 + \beta_1 x_i^1 + \varepsilon_i \qquad (1)\\
&= \tilde{y}_i = \tilde{\beta}_0 + \tilde{\beta}_1 \tilde{x}_i^2 + \tilde{\varepsilon}_i. \qquad (2)
\end{aligned}
$$

Plug in the variables $y_i = \delta_0 + \delta_1 x_i^1 + \varepsilon_i^y$ and $x_i^2 = \gamma_0 + \gamma_1 x_i^1 + \varepsilon_i^2$ in the equation (1)

$$
\begin{aligned}
y_i &= \delta_0 + \delta_1 x_i^1 + \varepsilon_i^y = \beta_0 + \beta_2(\gamma_0 + \gamma_1 x_i^1 + \varepsilon_i^2) + \beta_1 x_i^1 + \varepsilon_i\\
\tilde{y}_i &= \delta_0 + \varepsilon_i^y = \beta_0 + \beta_2(\gamma_0 + \varepsilon_i^2) + (\beta_2\gamma_1 - \delta_1 + \beta_1) x_i^1 + \varepsilon_i.
\end{aligned}
$$

Because we partialled out x_i^1 using OLS, x_i^1 is mechanically uncorrelated to ε_i^2 and to ε_i^y. Therefore, the regression coefficient $(\beta_2\gamma_1 - \delta_1 + \beta_1)$ of the partialled out variable x_i^1 is zero. The equation simplifies with $\tilde{x}_i^2 = \gamma_0 + \varepsilon_i^2$ to

$$
\tilde{y}_i = \delta_0 + \varepsilon_i^y = \beta_0 + \beta_2(\gamma_0 + \varepsilon_i^2) + \varepsilon_i.
$$

Detrended variables

Regression anatomy: Only detrending x_i^2 and not y_i. The regression constant, residuals, and the standard errors change but β_2 remains

$$\begin{aligned} y_i = \delta_0 + \delta_1 x_i^1 + \varepsilon_i^y &= (\beta_0 + \delta_1 \bar{x}^1) + \beta_2(\gamma_0 + \varepsilon_i^2) + (\varepsilon_i + \delta_1 x_i^1) \\ y_i &= \kappa + \beta_2 \tilde{x}^2 + \epsilon_i. \end{aligned} \tag{3}$$

Residualized variables

$$\begin{aligned}\tilde{y}_i = \delta_0 + \varepsilon_i^y &= \beta_0 + \beta_2(\gamma_0 + \varepsilon_i^2) + \varepsilon_i \\ \varepsilon_i^y &= \beta_0 - \delta_0 + \beta_2\gamma_0 + \beta_2\varepsilon_i^2 + \varepsilon_i.\end{aligned}$$

The same result of the FWL Theorem holds as well for a regression of the residualized variables because $\beta_0 = \delta_0 - \beta_2\gamma_0$:

$$\varepsilon_i^y = \beta_2\varepsilon_i^2 + \varepsilon_i.$$

Partition of $\boldsymbol{y}$

Least squares partitions the vector $\boldsymbol{y}$ into two orthogonal parts

$$\boldsymbol{y} = \hat{\boldsymbol{y}} + \boldsymbol{e} = \boldsymbol{X}\boldsymbol{b} + \boldsymbol{e} = \boldsymbol{P}\boldsymbol{y} + \boldsymbol{M}\boldsymbol{y}.$$

- $n \times 1$ vector of data $\boldsymbol{y}$
- $n \times n$ projection matrix $\boldsymbol{P}$
- $n \times n$ residual maker matrix $\boldsymbol{M}$
- $n \times 1$ vector of residuals $\boldsymbol{e}$

Projection matrix

$$\begin{aligned} \boldsymbol{Py} &= \boldsymbol{Xb} = \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y} \\ &\rightarrow \boldsymbol{P} = \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'. \end{aligned}$$

Definition

Properties.

- *symmetric such that* $\boldsymbol{P} = \boldsymbol{P}'$*, thus orthogonal*
- *idempotent such that* $\boldsymbol{P} = \boldsymbol{P}^2$*, thus indeed a projection*
- *annihilator matrix* $\boldsymbol{PX} = \boldsymbol{X}$

Example for projection matrix

Example

Show $\boldsymbol{PX} = \boldsymbol{X}(\boldsymbol{X'X})^{-1}\boldsymbol{X'X} = \boldsymbol{X}$.

$$\mathbf{X} = \begin{bmatrix}1 & 0\\1 & 1\\1 & 0\end{bmatrix}; \mathbf{X'X} = \begin{bmatrix}1 & 1 & 1\\0 & 1 & 0\end{bmatrix}\begin{bmatrix}1 & 0\\1 & 1\\1 & 0\end{bmatrix} = \begin{bmatrix}3 & 1\\1 & 1\end{bmatrix}; \mathbf{X'X}^{-1} = \begin{bmatrix}1/2 & -1/2\\-1/2 & 1.5\end{bmatrix};$$

$$\boldsymbol{X}(\boldsymbol{X'X})^{-1}\boldsymbol{X'} = \begin{bmatrix}1 & 0\\1 & 1\\1 & 0\end{bmatrix}\begin{bmatrix}1/2 & -1/2\\-1/2 & 3/2\end{bmatrix}\begin{bmatrix}1 & 1 & 1\\0 & 1 & 0\end{bmatrix} = \begin{bmatrix}1/2 & 0 & 1/2\\0 & 1 & 0\\1/2 & 0 & 1/2\end{bmatrix}$$

$$\boldsymbol{PX} = \begin{bmatrix}1/2 & 0 & 1/2\\0 & 1 & 0\\1/2 & 0 & 1/2\end{bmatrix}\begin{bmatrix}1 & 0\\1 & 1\\1 & 0\end{bmatrix} = \begin{bmatrix}1 & 0\\1 & 1\\1 & 0\end{bmatrix}. \tag{4}$$

Project $\boldsymbol{y}$ on the column space of $\boldsymbol{X}$, i.e. regress $\boldsymbol{y}$ on $\boldsymbol{x}$ and predict $E[\boldsymbol{y}] = \hat{\boldsymbol{y}}$.

$$\boldsymbol{y} = \begin{bmatrix}1\\2\\3\end{bmatrix}; \boldsymbol{Py} = \begin{bmatrix}1/2 & 0 & 1/2\\0 & 1 & 0\\1/2 & 0 & 1/2\end{bmatrix}\begin{bmatrix}1\\2\\3\end{bmatrix} = \hat{\boldsymbol{y}} = \begin{bmatrix}2\\2\\2\end{bmatrix}. \tag{5}$$

Residual maker matrix

$$\begin{aligned}
\boldsymbol{My} &= \boldsymbol{e} = \boldsymbol{y} - \boldsymbol{Xb} = \boldsymbol{y} - \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y} \\
\boldsymbol{My} &= (\boldsymbol{I} - \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}')\boldsymbol{y}
\end{aligned}$$

$$\rightarrow \boldsymbol{M} = \boldsymbol{I} - \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}' = (\boldsymbol{I} - \boldsymbol{P}).$$

Definition

Properties.

- *symmetric such that* $\boldsymbol{M} = \boldsymbol{M}'$
- *idempotent such that* $\boldsymbol{M} = \boldsymbol{M}^2$
- *annihilator matrix* $\boldsymbol{MX} = 0$
- *orthogonal to* $\boldsymbol{P}$*:* $\boldsymbol{PM} = \boldsymbol{MP} = 0$.

Example for residual maker matrix

Example

Show $\boldsymbol{MX} = (\boldsymbol{I} - \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}')\boldsymbol{X} = (\boldsymbol{I} - \boldsymbol{P})\boldsymbol{X} = \boldsymbol{X} - \boldsymbol{X} = 0$.

$$\mathbf{I} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}; \mathbf{X} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 0 \end{bmatrix};$$

$$\boldsymbol{M} = (\boldsymbol{I} - \boldsymbol{P}) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} 1/2 & 0 & 1/2 \\ 0 & 1 & 0 \\ 1/2 & 0 & 1/2 \end{bmatrix} = \begin{bmatrix} 1/2 & 0 & -1/2 \\ 0 & 0 & 0 \\ -1/2 & 0 & 1/2 \end{bmatrix}$$

$$\boldsymbol{MX} = \begin{bmatrix} 1/2 & 0 & -1/2 \\ 0 & 0 & 0 \\ -1/2 & 0 & 1/2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}. \tag{6}$$

Obtain residuals from a projection of $\boldsymbol{y}$ on the column space of $\boldsymbol{X}$, i.e. regress $\boldsymbol{y}$ on $\boldsymbol{x}$ and predict $\boldsymbol{y} - E[\boldsymbol{y}] = \boldsymbol{y} - \hat{\boldsymbol{y}}$.

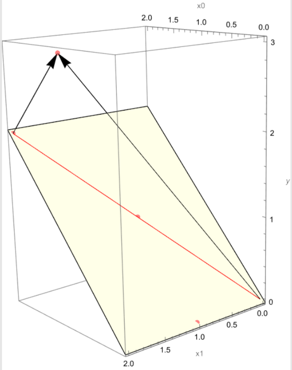
$$\boldsymbol{y} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}; \boldsymbol{My} = \begin{bmatrix} 1/2 & 0 & -1/2 \\ 0 & 0 & 0 \\ -1/2 & 0 & 1/2 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \boldsymbol{y} - \hat{\boldsymbol{y}} = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}. \tag{7}$$

Example for residual maker matrix

Example

Column space of $\mathbf{X}$ is $\mathbf{x}_0$ and $\mathbf{x}_1$.

$$\begin{bmatrix} x_0^1 = 1 & x_1^1 = 0 & y^1 = 1 \\ x_0^2 = 1 & x_1^2 = 1 & y^2 = 2 \\ x_0^3 = 1 & x_1^3 = 0 & y^3 = 3 \end{bmatrix}; \hat{\boldsymbol{y}} = \begin{bmatrix} 2 \\ 2 \\ 2 \end{bmatrix}; \boldsymbol{y} - \hat{\boldsymbol{y}} = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}. \tag{8}$$



The closest point from the vector $\boldsymbol{y}' = [1, 2, 3]$ onto the column space of $\boldsymbol{X}$, is $\hat{\boldsymbol{y}} = \boldsymbol{Xb}$, here $\hat{\boldsymbol{y}}' = [2, 2, 2]$. At this point, we can draw a line orthogonal to the column space of $\boldsymbol{X}$.

Decomposing the normal equations

The normal equations[1] in matrix form are $\boldsymbol{X}'\boldsymbol{X}\boldsymbol{b} = \boldsymbol{X}'\boldsymbol{y}$. If $\boldsymbol{X}$ is partitioned into an interesting segment $\boldsymbol{X}_2$ and an uninteresting $\boldsymbol{X}_1$, normal equations are

$$\begin{bmatrix} \boldsymbol{X}_1'\boldsymbol{X}_1 & \boldsymbol{X}_1'\boldsymbol{X}_2 \\ \boldsymbol{X}_2'\boldsymbol{X}_1 & \boldsymbol{X}_2'\boldsymbol{X}_2 \end{bmatrix} \begin{bmatrix} \boldsymbol{b}_1 \\ \boldsymbol{b}_2 \end{bmatrix} = \begin{bmatrix} \boldsymbol{X}_1'\boldsymbol{y} \\ \boldsymbol{X}_2'\boldsymbol{y} \end{bmatrix}.$$

The multiplication of the two equations can be done separately

$$\begin{bmatrix} \boldsymbol{X}_1'\boldsymbol{X}_1 & \boldsymbol{X}_1'\boldsymbol{X}_2 \end{bmatrix} \begin{bmatrix} \boldsymbol{b}_1 \\ \boldsymbol{b}_2 \end{bmatrix} = \begin{bmatrix} \boldsymbol{X}_1'\boldsymbol{y} \end{bmatrix} \tag{9}$$

$$\begin{bmatrix} \boldsymbol{X}_2'\boldsymbol{X}_1 & \boldsymbol{X}_2'\boldsymbol{X}_2 \end{bmatrix} \begin{bmatrix} \boldsymbol{b}_1 \\ \boldsymbol{b}_2 \end{bmatrix} = \begin{bmatrix} \boldsymbol{X}_2'\boldsymbol{y} \end{bmatrix}. \tag{10}$$

How can we find an expression for $\boldsymbol{b}_2$ that does not involve $\boldsymbol{b}_1$?

[1] It is called a normal equation because $\boldsymbol{y} - \boldsymbol{X}\boldsymbol{b}$ is normal to the range of $\boldsymbol{X}$.

Solving for $\boldsymbol{b}_2$

Idea: Solve equation (9) for $\boldsymbol{b}_1$ in terms of $\boldsymbol{b}_2$, then substituting that solution into the equation (10).

$$\begin{bmatrix} \boldsymbol{X}_1'\boldsymbol{X}_1 & \boldsymbol{X}_1'\boldsymbol{X}_2 \end{bmatrix} \begin{bmatrix} \boldsymbol{b}_1 \\ \boldsymbol{b}_2 \end{bmatrix} = \begin{bmatrix} \boldsymbol{X}_1'\boldsymbol{y} \end{bmatrix}$$

$$\boldsymbol{X}_1'\boldsymbol{X}_1\boldsymbol{b}_1 + \boldsymbol{X}_1'\boldsymbol{X}_2\boldsymbol{b}_2 = \boldsymbol{X}_1'\boldsymbol{y}$$

$$\boldsymbol{X}_1'\boldsymbol{X}_1\boldsymbol{b}_1 = \boldsymbol{X}_1'\boldsymbol{y} - \boldsymbol{X}_1'\boldsymbol{X}_2\boldsymbol{b}_2$$

$$\begin{aligned} \boldsymbol{b}_1 &= (\boldsymbol{X}_1'\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1'\boldsymbol{y} - (\boldsymbol{X}_1'\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1'\boldsymbol{X}_2\boldsymbol{b}_2 \\ &= (\boldsymbol{X}_1'\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1'(\boldsymbol{y} - \boldsymbol{X}_2\boldsymbol{b}_2) \end{aligned}$$

Multiplying out equation (10) gives

$$\begin{bmatrix} \boldsymbol{X}_2'\boldsymbol{X}_1 & \boldsymbol{X}_2'\boldsymbol{X}_2 \end{bmatrix} \begin{bmatrix} \boldsymbol{b}_1 \\ \boldsymbol{b}_2 \end{bmatrix} = \begin{bmatrix} \boldsymbol{X}_2'\boldsymbol{y} \end{bmatrix}$$

$$\boldsymbol{X}_2'\boldsymbol{X}_1\boldsymbol{b}_1 + \boldsymbol{X}_2'\boldsymbol{X}_2\boldsymbol{b}_2 = \boldsymbol{X}_2'\boldsymbol{y}$$

Plugging in the solution for $\boldsymbol{b}_1$ gives

$$\boldsymbol{X}_2'\boldsymbol{X}_1\Big((\boldsymbol{X}_1'\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1'(\boldsymbol{y} - \boldsymbol{X}_2\boldsymbol{b}_2)\Big) + \boldsymbol{X}_2'\boldsymbol{X}_2\boldsymbol{b}_2 = \boldsymbol{X}_2'\boldsymbol{y}. \tag{11}$$

Solving for $\boldsymbol{b}_2$

$$\boldsymbol{X}_2'\boldsymbol{X}_1(\boldsymbol{X}_1'\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1'(\boldsymbol{y}-\boldsymbol{X}_2\boldsymbol{b}_2)+\boldsymbol{X}_2'\boldsymbol{X}_2\boldsymbol{b}_2=\boldsymbol{X}_2'\boldsymbol{y}.$$

The middle part of the first term is $\boldsymbol{X}_1(\boldsymbol{X}_1'\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1'$. This is the projection matrix $\boldsymbol{P}_{X_1}$ from a regression of $\boldsymbol{y}$ on $\boldsymbol{X}_1$.

$$\boldsymbol{X}_2'\boldsymbol{P}_{X_1}\boldsymbol{y}-\boldsymbol{X}_2'\boldsymbol{P}_{X_1}\boldsymbol{X}_2\boldsymbol{b}_2+\boldsymbol{X}_2'\boldsymbol{X}_2\boldsymbol{b}_2=\boldsymbol{X}_2'\boldsymbol{y}.$$

We can multiply by an identity matrix $\boldsymbol{I}$ without changing anything

$$\begin{aligned}
&\boldsymbol{X}_2'\boldsymbol{P}_{X_1}\boldsymbol{y}-\boldsymbol{X}_2'\boldsymbol{P}_{X_1}\boldsymbol{X}_2\boldsymbol{b}_2+\boldsymbol{X}_2'\boldsymbol{I}\boldsymbol{X}_2\boldsymbol{b}_2=\boldsymbol{X}_2'\boldsymbol{I}\boldsymbol{y}.\\
&\boldsymbol{X}_2'\boldsymbol{I}\boldsymbol{y}-\boldsymbol{X}_2'\boldsymbol{P}_{X_1}\boldsymbol{y}=\boldsymbol{X}_2'\boldsymbol{I}\boldsymbol{X}_2\boldsymbol{b}_2-\boldsymbol{X}_2'\boldsymbol{P}_{X_1}\boldsymbol{X}_2\boldsymbol{b}_2.\\
&\boldsymbol{X}_2'(\boldsymbol{I}-\boldsymbol{P}_{X_1})\boldsymbol{y}=\boldsymbol{X}_2'(\boldsymbol{I}-\boldsymbol{P}_{X_1})\boldsymbol{X}_2\boldsymbol{b}_2.
\end{aligned}\tag{12}$$

Now $(\boldsymbol{I}-\boldsymbol{P}_{X_1})$ is the residual maker matrix $\boldsymbol{M}_{X_1}$

$$\boldsymbol{X}_2'\boldsymbol{M}_{X_1}\boldsymbol{y}=\boldsymbol{X}_2'\boldsymbol{M}_{X_1}\boldsymbol{X}_2\boldsymbol{b}_2.$$

Solving for $\boldsymbol{b}_2$ gives

$$\boldsymbol{b}_2 \;=\; (\boldsymbol{X}_2'\boldsymbol{M}_{X_1}\boldsymbol{X}_2)^{-1}\boldsymbol{X}_2'\boldsymbol{M}_{X_1}\boldsymbol{y}.$$

Solving for $\boldsymbol{b}_2$

$$\boldsymbol{b}_2 = (\boldsymbol{X}_2'\boldsymbol{M}_{X_1}\boldsymbol{X}_2)^{-1}\boldsymbol{X}_2'\boldsymbol{M}_{X_1}\boldsymbol{y}.$$

The residualizer matrix is symmetric and idempotent, such that $\boldsymbol{M}_{X_1} = \boldsymbol{M}_{X_1}\boldsymbol{M}_{X_1} = \boldsymbol{M}_{X_1}'\boldsymbol{M}_{X_1}$.

$$\begin{aligned} \boldsymbol{b}_2 &= (\boldsymbol{X}_2'\boldsymbol{M}_{X_1}'\boldsymbol{M}_{X_1}\boldsymbol{X}_2)^{-1}\boldsymbol{X}_2'\boldsymbol{M}_{X_1}'\boldsymbol{M}_{X_1}\boldsymbol{y} \\ &= \left((\boldsymbol{M}_{X_1}\boldsymbol{X}_2)'(\boldsymbol{M}_{X_1}\boldsymbol{X}_2)\right)^{-1}(\boldsymbol{M}_{X_1}\boldsymbol{X}_2)'(\boldsymbol{M}_{X_1}\boldsymbol{y}) \\ &= (\tilde{\boldsymbol{X}}_2'\tilde{\boldsymbol{X}}_2)^{-1}\tilde{\boldsymbol{X}}_2'\tilde{\boldsymbol{y}}. \qquad (13) \\ & \qquad (14) \end{aligned}$$

This is the OLS solution for $\boldsymbol{b}_2$, with $\tilde{\boldsymbol{X}}_2$ instead of $\boldsymbol{X}$ and $\tilde{\boldsymbol{y}}$ instead of $\boldsymbol{y}$.

- $\tilde{\boldsymbol{X}}_2$ are residuals from a regression of $\boldsymbol{X}_2$ on $\boldsymbol{X}_1$
- $\tilde{\boldsymbol{y}}$ are residuals from a regression of $\boldsymbol{y}$ on $\boldsymbol{X}_1$

The solution of the regression coefficients $\boldsymbol{b}_2$ in a regression that includes other regressors $\boldsymbol{X}_1$ is the same as first regressing all of $\boldsymbol{X}_2$ and $\boldsymbol{y}$ on $\boldsymbol{X}_1$, then regressing the residuals from the $\boldsymbol{y}$ regression on the residuals from the $\boldsymbol{X}_2$ regression.

References I

ANGRIST, J. D., AND J.-S. PISCHKE (2009): *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press, Chapter 3.

FILOSO, V. (2013): "Regression Anatomy, Revealed," *The Stata Journal*, 13(1), 92–106.

FRISCH, R., AND F. V. WAUGH (1933): "Partial Time Regressions as Compared with Individual Trends," *Econometrica*, 1(4), 387–401.

GREENE, W. H. (2011): *Econometric Analysis*. Prentice Hall, 5 edn.

LOVELL, M. C. (2008): "A Simple Proof of the FWL Theorem," *The Journal of Economic Education*, 39(1), 88–91.