Mathematical Proof

The Frisch-Waugh-Lovell Theorem

1 Mathematical Foundation

1.1 Problem Setup

Consider the linear regression model:

$$\boxed{y = X_1\beta_1 + X_2\beta_2 + u} \tag{1}$$

where the components are defined as follows:

- $y \in \mathbb{R}^{n\times 1}$ is the vector of outcomes
- $X_1 \in \mathbb{R}^{n\times k_1}$ is the matrix of regressors of interest with $\text{rank}(X_1) = k_1$
- $X_2 \in \mathbb{R}^{n\times k_2}$ is the matrix of control variables with $\text{rank}(X_2) = k_2$
- $\beta_1 \in \mathbb{R}^{k_1\times 1}$ is the parameter vector of primary interest
- $\beta_2 \in \mathbb{R}^{k_2\times 1}$ is the parameter vector for control variables
- $u \in \mathbb{R}^{n\times 1}$ is the error vector with $\mathbb{E}[u|X_1, X_2] = 0$

Definition 1.1 (Projection and Annihilator Matrices). For a full-rank matrix $X_2 \in \mathbb{R}^{n\times k_2}$, we define:

$$P_{X_2} = X_2(X_2'X_2)^{-1}X_2' \quad \text{(Projection matrix)} \tag{2}$$

$$M_{X_2} = I_n - P_{X_2} = I_n - X_2(X_2'X_2)^{-1}X_2' \quad \text{(Annihilator matrix)} \tag{3}$$

These matrices satisfy the following fundamental properties:

(i) P_{X_2} and M_{X_2} are symmetric: $P_{X_2}' = P_{X_2}$, $M_{X_2}' = M_{X_2}$

(ii) P_{X_2} and M_{X_2} are idempotent: $P_{X_2}^2 = P_{X_2}$, $M_{X_2}^2 = M_{X_2}$

(iii) $M_{X_2}X_2 = 0$ and $P_{X_2}X_2 = X_2$

(iv) $P_{X_2} + M_{X_2} = I_n$

2 Main Result

2.1 The Frisch-Waugh-Lovell Theorem

Theorem 2.1 (Frisch-Waugh-Lovell Theorem). The OLS estimate of β_1 in the full regression of y on $[X_1 \quad X_2]$ is identical to the OLS estimate obtained from the following two-step partialling-out procedure:

Step 1: Regress y on X_2 and obtain residuals: $\tilde{y} = M_{X_2} y$

Step 2: Regress X_1 on X_2 and obtain residuals: $\tilde{X}_1 = M_{X_2} X_1$

Step 3: Regress $\tilde{y}$ on $\tilde{X}_1$ to obtain: $\hat{\beta}_1^{\text{FWL}} = (\tilde{X}_1{}' \tilde{X}_1)^{-1} \tilde{X}_1{}' \tilde{y}$

Formal Statement:

$$\boxed{\hat{\beta}_1 = \hat{\beta}_1^{\text{FWL}} = (\tilde{X}_1{}' \tilde{X}_1)^{-1} \tilde{X}_1{}' \tilde{y}} \tag{4}$$

2.2 Proof

Proof. We establish the equivalence by demonstrating that both approaches yield identical coefficient estimates through rigorous matrix algebra.

Part I: Full Regression Setup

The full regression model in partitioned form is:

$$y = [X_1 \quad X_2] \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} + u = X\beta + u \tag{5}$$

where $X = [X_1 \quad X_2] \in \mathbb{R}^{n \times (k_1 + k_2)}$ and $\beta = [\beta_1' \quad \beta_2']' \in \mathbb{R}^{(k_1+k_2) \times 1}$.
The OLS estimator is given by:

$$\hat{\beta} = (X'X)^{-1} X'y = \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix} \tag{6}$$

Part II: Matrix Partitioning

We partition the cross-product matrices:

$$X'X = \begin{bmatrix} X_1'X_1 & X_1'X_2 \\ X_2'X_1 & X_2'X_2 \end{bmatrix} = \begin{bmatrix} A & B \\ B' & D \end{bmatrix} \tag{7}$$

$$X'y = \begin{bmatrix} X_1'y \\ X_2'y \end{bmatrix} \tag{8}$$

where:

$$A = X_1'X_1 \in \mathbb{R}^{k_1 \times k_1}, \quad B = X_1'X_2 \in \mathbb{R}^{k_1 \times k_2}, \quad D = X_2'X_2 \in \mathbb{R}^{k_2 \times k_2} \tag{9}$$

Part III: Partitioned Matrix Inverse

Using the block matrix inversion formula:

$$(X'X)^{-1} = \begin{bmatrix} (A - BD^{-1}B')^{-1} & -(A - BD^{-1}B')^{-1}BD^{-1} \\ -D^{-1}B'(A - BD^{-1}B')^{-1} & D^{-1} + D^{-1}B'(A - BD^{-1}B')^{-1}BD^{-1} \end{bmatrix} \tag{10}$$

Part IV: Key Algebraic Identity

We establish the fundamental relationship:

$$\begin{aligned} A - BD^{-1}B' &= X_1'X_1 - X_1'X_2(X_2'X_2)^{-1}X_2'X_1 && (11)\\ &= X_1'(I_n - X_2(X_2'X_2)^{-1}X_2')X_1 && (12)\\ &= X_1'M_{X_2}X_1 && (13) \end{aligned}$$

Part V: Extracting $\hat{\beta}_1$

From the normal equations $(X'X)\hat{\beta} = X'y$, the first block gives us:

$$\begin{aligned} \hat{\beta}_1 &= (A - BD^{-1}B')^{-1}(X_1'y - BD^{-1}X_2'y) && (14)\\ &= (X_1'M_{X_2}X_1)^{-1}(X_1'y - X_1'X_2(X_2'X_2)^{-1}X_2'y) && (15)\\ &= (X_1'M_{X_2}X_1)^{-1}X_1'(I_n - X_2(X_2'X_2)^{-1}X_2')y && (16)\\ &= (X_1'M_{X_2}X_1)^{-1}X_1'M_{X_2}y && (17) \end{aligned}$$

Part VI: Two-Step Procedure Analysis

The partialling-out procedure yields:

$$\begin{aligned} \tilde{y} &= M_{X_2}y \quad \text{(Step 1)} && (18)\\ \tilde{X}_1 &= M_{X_2}X_1 \quad \text{(Step 2)} && (19)\\ \hat{\beta}_1^{\text{FWL}} &= (\tilde{X}_1{}'\tilde{X}_1)^{-1}\tilde{X}_1{}'\tilde{y} \quad \text{(Step 3)} && (20) \end{aligned}$$

Part VII: Establishing Equivalence

Substituting the definitions from Steps 1 and 2:

$$\begin{aligned} \hat{\beta}_1^{\text{FWL}} &= ((M_{X_2}X_1)'(M_{X_2}X_1))^{-1}(M_{X_2}X_1)'(M_{X_2}y) && (21)\\ &= (X_1'M_{X_2}'M_{X_2}X_1)^{-1}X_1'M_{X_2}'M_{X_2}y && (22) \end{aligned}$$

Part VIII: Applying Matrix Properties

Using the symmetry and idempotency of M_{X_2} from Definition 1.1:

$$\begin{aligned} M_{X_2}' &= M_{X_2} \quad \text{(symmetry)} && (23)\\ M_{X_2}M_{X_2} &= M_{X_2} \quad \text{(idempotency)} && (24) \end{aligned}$$

Therefore:

$$\begin{aligned} \hat{\beta}_1^{\text{FWL}} &= (X_1'M_{X_2}M_{X_2}X_1)^{-1}X_1'M_{X_2}M_{X_2}y && (25)\\ &= (X_1'M_{X_2}X_1)^{-1}X_1'M_{X_2}y && (26) \end{aligned}$$

Part IX: Final Equivalence

Comparing equations (17) and (26):

$$\hat{\beta}_1 = (X_1' M_{X_2} X_1)^{-1} X_1' M_{X_2} y = \hat{\beta}_1^{\text{FWL}} \tag{27}$$

This establishes the desired result:

$$\boxed{\hat{\beta}_1 = (\tilde{X}_1{}' \tilde{X}_1)^{-1} \tilde{X}_1{}' \tilde{y}} \tag{28}$$

$\square$