Proof: $E[\varepsilon|X]=0$ → Conditional Form

$$E[\varepsilon_i|X_i=x]=0$$

$$E[y_i-\beta_0-\beta_1X_i|X_i=x]=0$$

$$E[y_i|X_i=x]=\beta_0+\beta_1X$$

Population + Sample Mean

$$E[Y]=\int y\, f_y(y)\,dy$$

$$\bar{Y}=\frac{1}{n}\sum_{i}^{n}Y_i$$

Derivation: OLS Coefficients

$$\min_{b_0,b_1}\hat{E}[(Y_i-(b_0+b_1X_i))^2]$$

$$=\min_{b_0,b_1}\frac{1}{n}\sum_{i=1}^{n}[(Y_i^2-2Y_i(b_0+b_1X_i)+(b_0+b_1X_i)^2]$$

$$\frac{\partial}{\partial b_0}\Rightarrow \frac{1}{n}\sum_{i=1}^{n}[-2Y_i+2(b_0+b_1X_i)]=0$$

↳ $\hat{b}_0=\frac{1}{n}\sum^{n}Y_i-\frac{\hat{b}_1}{n}\sum^{n}X_i$

$$\frac{\partial}{\partial b_1}\Rightarrow\frac{1}{n}\sum_{i=1}^{n}[-2Y_iX_i+2(\hat{b}_0+\hat{b}_1X_i)X_i]=0$$

↓ plug in $\hat{b}_0$

$$\boxed{\hat{b}_1=\widehat{cov}(X,Y)/\widehat{Var}(X)}$$

$$\boxed{\hat{b}_0=\bar{Y}-\hat{b}_1\bar{X}}$$

Population & Sample Variance

$$Var(Y)=E[(Y-E[Y])^2]$$

$$\widehat{Var}(Y)=\frac{1}{n-1}\sum_{i=1}^{n}(Y_i-\bar{Y})^2$$

Population & Sample Covariance

$$Cov(X,Y)=E[(Y-E[Y])(X-E[X])]$$

$$\widehat{cov}(X,Y)=\frac{1}{n-1}\sum_{i=1}^{n}[(Y_i-\bar{Y})(X_i-\bar{X})]$$

Derivation: Omitted Variable Bias

$$Y_i=\beta_0+\beta_1X_i+\underbrace{\gamma S_i+\eta_i}_{\varepsilon_i}$$

↳ $\boxed{S_i}$ initially left out of regression

Equations & Proofs

$$\hat{b}_1=Cov(Y,X)/Var(X)$$

$$=\frac{cov(\beta_0,X)^{\to 0}+cov(\beta_1X_i,X_i)+cov(\varepsilon_i,X_i)}{Var(X)}$$

$$=\beta_1+\frac{cov(\gamma S_i,X_i)+cov(\eta_i,X_i)^{\to 0}}{Var(X_i)}$$

$$\boxed{\hat{b}_1=\beta_1+\gamma\frac{cov(S_i,X_i)}{Var(X_i)}}$$

SSR, SST, SSE

SSR	SST	SSE
(sum of squared residuals)	(sum of squares total)	(sum of squared Errors)
$SSR=\sum_{i}^{n}\hat{u}_i^2$	$SST=\sum(Y_i-\bar{Y})^2$	$SSE=\sum_{i}^{n}(Y_i-\hat{Y}_i)^2$

Derivation: $\hat{b}$ matrix

$$\min((Y-Xb)'(Y-Xb))$$

$$\min[(Y'Y-Y'Xb)-(b'X'Y-b'X'Xb)]$$

$$=\min(-2Y'Xb+b'X'Xb)$$

↓ $\boxed{\frac{\partial}{\partial b}}$

$$X'Xb=X'Y$$

$$\boxed{\hat{b}=(X'X)^{-1}X'Y}$$

Derivation: No Perfect collinearity

$$E[\hat{b}|X]=E[(X'X)^{-1}X'Y|X]$$

$$=E[(X'X)^{-1}X'(X\beta+\varepsilon)|X]$$

$$=\beta+E[(X'X)^{-1}X'E[\varepsilon|X]]$$

$$=\boxed{\beta}$$

Homoskedasticity of Error Term

$$V(\hat{b}|X)=(X'X)^{-1}X'V(Y|X)((X'X)X')'$$

$$=(X'X)'X'V(X'\beta+\varepsilon|X)((X'X)^{-1}X')'$$

$$=(X'X)^{-1}X'\sigma^2I((X'X)^{-1}X')'$$

$$=\sigma^2[(X'X)^{-1}X'X(X'X)^{-1}]$$

$$\boxed{=\sigma^2(X'X)^{-1}}$$

Asymptotic Normal Distribution of β

$$\sqrt{n}(\hat{b}-\beta)$$

$$=\sqrt{n}((X'X)^{-1}X'Y-\beta)$$

$$=\sqrt{n}((X'X)^{-1}X'(X\beta+\varepsilon)-\beta)$$

$$=\sqrt{n}(X'X)^{-1}X'\varepsilon$$

$$=\underbrace{\frac{1}{n}\sum^{n}X_iX_i'}_{Q_x\text{ by LLN}}*\underbrace{\frac{1}{\sqrt{n}}\sum^{n}X_i\varepsilon_i}_{N(0,A)\text{ by CLT}}$$

Simultaneity Problem

$$cov(P_t,u_{1t})=Cov(Y+\delta a_t+Z_t^s\lambda^s+U_t|u_{1t})$$

$$=\delta Cov(a_t,u_{1t})+cov(u_{1t},u_{2t})$$

$$=\delta Cov(\alpha+\beta p_t+Z_t^d\lambda^d+u_{1t},u_{1t})+\text{'' ''}$$

$$=\delta(\beta cov(P_t,u_{1t})+Var(u_{1t})+\text{'' ''}$$

$$\boxed{cov(P_t,u_{1t})=\frac{\delta Var(u_{1t})+Cov(u_{1t},u_{2t})}{1-\delta\beta}}$$

Derivation: IV estimator

$$E[Z_i\varepsilon_i]=0$$

$$E[Z_i(Y_i-X_i'\beta)]=0$$

$$E[Z_iY_i-Z_iX_i'\beta]=0$$

$$E[Z_iY_i]-E[Z_iX_i'\beta]=0$$

$$\boxed{\beta=(Z'X)^{-1}Z'Y}$$

Equivalence of simple IV to Just Identif

$$\hat{b}=(\hat{X}'X)\hat{X}Y$$

$$=(X'P_zX)^{-1}XP_zY$$

$$=(X'Z(Z'Z)^{-1}Z'X)^{-1}X'Z(Z'Z)^{-1}Z'Y$$

$$=(Z'X)^{-1}(Z'Z)(X'Z)^{-1}X'Z(Z'Z)^{-1}Z'Y$$

$$\boxed{=(Z'X)^{-1}Z'Y}$$

Proving Random Effects

$$\boxed{A3: \operatorname{Cov}(\tilde{\varepsilon}_{it}, \tilde{\varepsilon}_{js}|X) = 0}$$

$$\operatorname{cov}(\tilde{\varepsilon}_{it}, \tilde{\varepsilon}_{is}|X) = \operatorname{cov}(\varepsilon_{it} - \lambda\bar{\varepsilon}_i, \varepsilon_{is} - \lambda\bar{\varepsilon}_i|X)$$

$$= \operatorname{cov}(\varepsilon_{it}, \varepsilon_{is} - \lambda\bar{\varepsilon}_i|X) - \lambda\operatorname{Cov}(\bar{\varepsilon}_i, \varepsilon_{is} - \lambda\bar{\varepsilon}_i)$$

$$= \boxed{\sigma_\alpha^2 - 2\lambda\operatorname{cov}(\varepsilon_{it}, \bar{\varepsilon}_i) + \lambda^2\operatorname{Var}(\bar{\varepsilon}_i|X)}$$

$$\operatorname{Var}(\bar{\varepsilon}_i|X) = \operatorname{Var}\left(\frac{1}{T}\sum^T \varepsilon_{it}|X\right)$$

$$= \frac{1}{T^2}\operatorname{Var}\left(\sum_{t=1}^T(\alpha_i + \nu_{it})|X\right)$$

$$= \frac{1}{T^2}\operatorname{Var}\left(T\alpha_i + \sum_{t=1}^T \nu_{it}|X\right)$$

$$= \frac{1}{T^2}\left(T^2\sigma_a^2 + T\sigma_v^2\right)$$

$$= \sigma_a^2 + \frac{1}{T}\sigma_v^2$$

$$\boxed{= \frac{1}{T}(\sigma_v^2 + T\sigma_a^2)}$$

Now,

$$\operatorname{Cov}(\varepsilon_{it}, \bar{\varepsilon}_i|X) = \operatorname{cov}\left(\varepsilon_{it}, \frac{1}{T}\sum\varepsilon_s\right)$$

$$= \boxed{\frac{1}{T}(\sigma_v^2 + T\sigma_a^2)}$$

Final

$$\operatorname{cov}(\tilde{\varepsilon}_{it}, \tilde{\varepsilon}_{is}|X)$$

$$\parallel$$

$$\sigma_a^2 - 2\lambda\operatorname{Cov}(\varepsilon_{it}, \bar{\varepsilon}_i) + \lambda^2\operatorname{Var}(\bar{\varepsilon}_i|X)$$

$$= \boxed{0}$$

Proof ①

$$E[\varepsilon_i \mid x_i = x] = 0$$
$$E[y_i - \beta_0 - \beta_1 x_i \mid x_i = x] = 0$$
$$E[y_i \mid x_i = x] - \beta_0 - \beta_1 x = 0$$
$$E[y_i \mid x_i = x] = \beta_0 + \beta_1 x$$

Population & Sample Mean ②

$$E[Y] = \int y f_y(y)\, dy$$

$$\bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i$$

Population & Sample Variance

$$Var(Y) = E[(Y - E[Y])^2]$$

$$\widehat{Var}(Y) = \frac{1}{n-1}\sum_{i=1}^{n}(Y_i - \bar{Y})^2$$

Population & Sample Covariance

$$Cov(X,Y) = E[(X - E[X])(Y - E[Y])]$$

$$\widehat{Cov}(X,Y) = \frac{1}{n-1}\sum(y_i - \bar{Y})(x_i - \bar{X})$$

OLS COEFFICIENT Derivation ③

$$\min_{b_0,b_1} \hat{E}[(Y_i - (b_0 + b_1 X_i))^2]$$

$$= \min_{b_0,b_1} \frac{1}{n}\sum_{i=1}^{n}[Y_i^2 - 2Y_i(b_0 + b_1 X_i) + (b_0 + b_1 X_i)^2]$$

$$\frac{\partial}{\partial b_0}\left\{\frac{1}{n}\sum_{i=1}^{n}[-2Y_i + 2(b_0 + b_1 x_i)] = 0\right.$$

$$\boxed{\hat{b}_0 = \frac{1}{n}\sum Y_i - \frac{\hat{b}_1}{n}\sum x_i}$$

$$\frac{\partial}{\partial b_1}\left\{\frac{1}{n}\sum_{i=1}^{n}[-2Y_i x_i + 2(\hat{b}_0 + \hat{b}_1 x_i)x_i] = 0\right.$$

$$\boxed{\hat{b}_1 = \frac{\sum_{i=1}^{n} Y_i x_i - \frac{1}{n}\sum Y_i \sum x_i}{\sum x_i^2 - \frac{1}{n}(\sum x_i)^2}}$$

$$\Downarrow$$

$$\boxed{\hat{b}_1 = \widehat{Cov}(X,Y) / \widehat{Var}(X)}$$

Asymptotic Normal Distribution of β ⑨

$$\sqrt{n}(\hat{b} - \beta)$$
$$= \sqrt{n}((X'X)^{-1}X'Y - \beta)$$
$$= \sqrt{n}((X'X)^{-1}X'(X\beta + \varepsilon) - \beta)$$
$$= \sqrt{n}(X'X)^{-1}X'\varepsilon$$
$$= \underbrace{\frac{1}{n}\sum_{i=1}^{n} x_i x_i'}_{Q_x \text{ by LLN}} * \underbrace{\frac{1}{\sqrt{n}}\sum_{i=1}^{n} x_i \varepsilon_i}_{N(0,A) \text{ by CLT}}$$

Proof that $b_1 = \beta_1$ & $b_0 = \beta_0$ ④

$$b_1 = \frac{Cov(X,Y)}{Var(X)}$$
$$= \frac{cov(x, \beta_0 + \beta_1 x + \varepsilon)}{var(x)}$$
$$= \frac{0 + \beta_1 var(x) + 0}{var(x)} = \boxed{\beta_1}$$

$$b_0 = E[Y_i] - \beta_1 E[x_i]$$
$$= E[\beta_0 + \beta_1 x_i + \varepsilon_i] - \beta_1 E[x_i]$$
$$= \beta_0 + \beta_1 E[x_i] - \beta_1 E[x_i]$$
$$\boxed{= \beta_0}$$

Derivation of Omitted Variable Bias ⑤

$$Y_i = \beta_0 + \beta_1 x_i + \underbrace{\gamma S_i + n_i}_{\varepsilon_i}$$

$\boxed{S_i}$ initially left out of regression

$$\hat{b}_1 = Cov(Y,x) / Var(x)$$
$$= \frac{Cov(\beta_0, x)^{0} + Cov(\beta_1 x_i, x_i) + cov(\varepsilon_i, x)}{var(x)}$$
$$= \beta_1 \frac{var(x)}{var(x)} + \frac{cov(\gamma S_i, x)}{var(x)} + \frac{cov(n_i, x)^{0}}{var(x)}$$
$$\boxed{\hat{b}_1 = \beta_1 + \gamma \frac{cov(S_i, x_i)}{Var(x_i)}}$$

Derivation of $\hat{b}$ Matrix (Multiple) ⑥

$$\min((Y - Xb)'(Y - Xb))$$
$$\min((Y'Y - Y'Xb) - (b'X'Y - b'X'Xb))$$
$$\min(-2Y'Xb + b'X'Xb)$$
$$\downarrow \frac{\partial}{\partial b}$$
$$X'Xb = X'Y$$
$$\boxed{\hat{b} = (X'X)^{-1}X'Y}$$

No Perfect Collinearity ⑦

$$E[\hat{b} \mid X] = E[(X'X)^{-1}X'Y \mid X]$$
$$= E[(X'X)^{-1}X'(X\beta + \varepsilon) \mid X]$$
$$= E[(X'X)^{-1}X'X\beta + \text{"} \quad \text{"} \varepsilon \mid X]$$
$$= \beta + E[(X'X)^{-1}X'E[\varepsilon \mid X]] \to 0$$
$$= \boxed{\beta}$$

Homoskedasticity of error term ⑧

$$V(\hat{b} \mid X) = (X'X)^{-1}X'V(Y \mid X)((X'X)^{-1}X')'$$
$$= (X'X)^{-1}X'V(X'\beta + \varepsilon \mid X)((X'X)^{-1}X')'$$
$$= (X'X)^{-1}X'\sigma^2 I((X'X)^{-1}X')'$$
$$= \sigma^2(X'X)^{-1}X'X(X'X)^{-1}$$
$$\boxed{= \sigma^2(X'X)^{-1}}$$

SSR, SST, SSE

SSR (sum of square Residuals)	SST (sum of squares total)	SSE (sum of square errors)
$SSR = \sum \hat{n}_i^2$	$SST = \sum(Y_i - \bar{Y})^2$	$SSE = \sum(Y_i - \hat{Y}_i)^2$

P-value for t-test

$$P = F_{t(n-(k+1))}(-|t_{\hat{b}_j}|) +$$
$$1 - F_{t(n-(k+1))}(|t_{\hat{b}_j}|)$$