

Linear Models Lecture 15: Endogeneity

Robert Gulotty

University of Chicago

May 14, 2025

General form of Endogeneity

- Our *structural* model is $Y = \mathbf{x}'\beta + e$.
- If $\mathbb{E}[\mathbf{x}e] \neq 0$, we say \mathbf{x} is endogenous for β .
- We can still define a projection equation, $Y = \mathbf{x}'\beta^* + e^*$, for which $\mathbb{E}[\mathbf{x}e^*] = 0$.

$$\begin{aligned}\beta^* &= (\mathbb{E}[\mathbf{x}\mathbf{x}'])^{-1}\mathbb{E}[\mathbf{x}Y] \\ &= (\mathbb{E}[\mathbf{x}\mathbf{x}'])^{-1}\mathbb{E}[\mathbf{x}(\mathbf{x}'\beta + e)] \\ &= (\mathbb{E}[\mathbf{x}\mathbf{x}'])^{-1}(\mathbb{E}[\mathbf{x}\mathbf{x}'])\beta + (\mathbb{E}[\mathbf{x}\mathbf{x}'])^{-1}\mathbb{E}[\mathbf{x}\mathbf{x}']e \\ &= \beta + (\mathbb{E}[\mathbf{x}\mathbf{x}'])^{-1}\mathbb{E}[\mathbf{x}e]\end{aligned}$$

- So, $\hat{\beta} \xrightarrow{p} \beta^* \neq \beta$ and the least squares estimate is inconsistent.
- This is called endogeneity, or endogeneity bias, or bias due to endogeneity.

General form of Endogeneity

- Our *structural* model is $Y = \mathbf{x}'\beta + e$.
- If $\mathbb{E}[\mathbf{x}e] \neq 0$, we say \mathbf{x} is endogenous for β .
- We can still define a projection equation, $Y = \mathbf{x}'\beta^* + e^*$, for which $\mathbb{E}[\mathbf{x}e^*] = 0$.

$$\begin{aligned}\beta^* &= (\mathbb{E}[\mathbf{x}\mathbf{x}'])^{-1}\mathbb{E}[\mathbf{x}Y] \\ &= (\mathbb{E}[\mathbf{x}\mathbf{x}'])^{-1}\mathbb{E}[\mathbf{x}(\mathbf{x}'\beta + e)] \\ &= (\mathbb{E}[\mathbf{x}\mathbf{x}'])^{-1}(\mathbb{E}[\mathbf{x}\mathbf{x}'])\beta + (\mathbb{E}[\mathbf{x}\mathbf{x}'])^{-1}\mathbb{E}[\mathbf{x}\mathbf{x}']e \\ &= \beta + (\mathbb{E}[\mathbf{x}\mathbf{x}'])^{-1}\mathbb{E}[\mathbf{x}e]\end{aligned}$$

- So, $\hat{\beta} \xrightarrow{p} \beta^* \neq \beta$ and the least squares estimate is inconsistent.
- This is called endogeneity, or endogeneity bias, or bias due to endogeneity.

General form of Endogeneity

- Our *structural* model is $Y = \mathbf{x}'\beta + e$.
- If $\mathbb{E}[\mathbf{x}e] \neq 0$, we say \mathbf{x} is endogenous for β .
- We can still define a projection equation, $Y = \mathbf{x}'\beta^* + e^*$, for which $\mathbb{E}[\mathbf{x}e^*] = 0$.

$$\begin{aligned}\beta^* &= (\mathbb{E}[\mathbf{x}\mathbf{x}'])^{-1}\mathbb{E}[\mathbf{x}Y] \\ &= (\mathbb{E}[\mathbf{x}\mathbf{x}'])^{-1}\mathbb{E}[\mathbf{x}(\mathbf{x}'\beta + e)] \\ &= (\mathbb{E}[\mathbf{x}\mathbf{x}'])^{-1}(\mathbb{E}[\mathbf{x}\mathbf{x}'])\beta + (\mathbb{E}[\mathbf{x}\mathbf{x}'])^{-1}\mathbb{E}[\mathbf{x}\mathbf{x}']e \\ &= \beta + (\mathbb{E}[\mathbf{x}\mathbf{x}'])^{-1}\mathbb{E}[\mathbf{x}e]\end{aligned}$$

- So, $\hat{\beta} \xrightarrow{p} \beta^* \neq \beta$ and the least squares estimate is inconsistent.
- This is called endogeneity, or endogeneity bias, or bias due to endogeneity.

General form of Endogeneity

- Our *structural* model is $Y = \mathbf{x}'\beta + e$.
- If $\mathbb{E}[\mathbf{x}e] \neq 0$, we say \mathbf{x} is endogenous for β .
- We can still define a projection equation, $Y = \mathbf{x}'\beta^* + e^*$, for which $\mathbb{E}[\mathbf{x}e^*] = 0$.

$$\begin{aligned}\beta^* &= (\mathbb{E}[\mathbf{x}\mathbf{x}'])^{-1}\mathbb{E}[\mathbf{x}Y] \\ &= (\mathbb{E}[\mathbf{x}\mathbf{x}'])^{-1}\mathbb{E}[\mathbf{x}(\mathbf{x}'\beta + e)] \\ &= (\mathbb{E}[\mathbf{x}\mathbf{x}'])^{-1}(\mathbb{E}[\mathbf{x}\mathbf{x}'])\beta + (\mathbb{E}[\mathbf{x}\mathbf{x}'])^{-1}\mathbb{E}[\mathbf{x}\mathbf{x}']e \\ &= \beta + (\mathbb{E}[\mathbf{x}\mathbf{x}'])^{-1}\mathbb{E}[\mathbf{x}e]\end{aligned}$$

- So, $\hat{\beta} \xrightarrow{p} \beta^* \neq \beta$ and the least squares estimate is inconsistent.
- This is called endogeneity, or endogeneity bias, or bias due to endogeneity.

General form of Endogeneity

- Our *structural* model is $Y = \mathbf{x}'\beta + e$.
- If $\mathbb{E}[\mathbf{x}e] \neq 0$, we say \mathbf{x} is endogenous for β .
- We can still define a projection equation, $Y = \mathbf{x}'\beta^* + e^*$, for which $\mathbb{E}[\mathbf{x}e^*] = 0$.

$$\begin{aligned}\beta^* &= (\mathbb{E}[\mathbf{x}\mathbf{x}'])^{-1}\mathbb{E}[\mathbf{x}Y] \\ &= (\mathbb{E}[\mathbf{x}\mathbf{x}'])^{-1}\mathbb{E}[\mathbf{x}(\mathbf{x}'\beta + e)] \\ &= (\mathbb{E}[\mathbf{x}\mathbf{x}'])^{-1}(\mathbb{E}[\mathbf{x}\mathbf{x}'])\beta + (\mathbb{E}[\mathbf{x}\mathbf{x}'])^{-1}\mathbb{E}[\mathbf{x}\mathbf{x}']e \\ &= \beta + (\mathbb{E}[\mathbf{x}\mathbf{x}'])^{-1}\mathbb{E}[\mathbf{x}e]\end{aligned}$$

- So, $\hat{\beta} \xrightarrow{p} \beta^* \neq \beta$ and the least squares estimate is inconsistent.
- This is called endogeneity, or endogeneity bias, or bias due to endogeneity.

General form of Endogeneity

- Our *structural* model is $Y = \mathbf{x}'\beta + e$.
- If $\mathbb{E}[\mathbf{x}e] \neq 0$, we say \mathbf{x} is endogenous for β .
- We can still define a projection equation, $Y = \mathbf{x}'\beta^* + e^*$, for which $\mathbb{E}[\mathbf{x}e^*] = 0$.

$$\begin{aligned}\beta^* &= (\mathbb{E}[\mathbf{x}\mathbf{x}'])^{-1}\mathbb{E}[\mathbf{x}Y] \\ &= (\mathbb{E}[\mathbf{x}\mathbf{x}'])^{-1}\mathbb{E}[\mathbf{x}(\mathbf{x}'\beta + e)] \\ &= (\mathbb{E}[\mathbf{x}\mathbf{x}'])^{-1}(\mathbb{E}[\mathbf{x}\mathbf{x}'])\beta + (\mathbb{E}[\mathbf{x}\mathbf{x}'])^{-1}\mathbb{E}[\mathbf{x}\mathbf{x}']e \\ &= \beta + (\mathbb{E}[\mathbf{x}\mathbf{x}'])^{-1}\mathbb{E}[\mathbf{x}e]\end{aligned}$$

- So, $\hat{\beta} \xrightarrow{p} \beta^* \neq \beta$ and the least squares estimate is inconsistent.
- This is called endogeneity, or endogeneity bias, or bias due to endogeneity.

General form of Endogeneity

- Our *structural* model is $Y = \mathbf{x}'\beta + e$.
- If $\mathbb{E}[\mathbf{x}e] \neq 0$, we say \mathbf{x} is endogenous for β .
- We can still define a projection equation, $Y = \mathbf{x}'\beta^* + e^*$, for which $\mathbb{E}[\mathbf{x}e^*] = 0$.

$$\begin{aligned}\beta^* &= (\mathbb{E}[\mathbf{x}\mathbf{x}'])^{-1}\mathbb{E}[\mathbf{x}Y] \\ &= (\mathbb{E}[\mathbf{x}\mathbf{x}'])^{-1}\mathbb{E}[\mathbf{x}(\mathbf{x}'\beta + e)] \\ &= (\mathbb{E}[\mathbf{x}\mathbf{x}'])^{-1}(\mathbb{E}[\mathbf{x}\mathbf{x}'])\beta + (\mathbb{E}[\mathbf{x}\mathbf{x}'])^{-1}\mathbb{E}[\mathbf{x}\mathbf{x}']e \\ &= \beta + (\mathbb{E}[\mathbf{x}\mathbf{x}'])^{-1}\mathbb{E}[\mathbf{x}e]\end{aligned}$$

- So, $\hat{\beta} \xrightarrow{p} \beta^* \neq \beta$ and the least squares estimate is inconsistent.
- This is called endogeneity, or endogeneity bias, or bias due to endogeneity.

Motivation 1: Measurement Error

- We think that $\mathbb{E}[Y|\mathbf{z}] = \mathbf{z}'\beta$, but we don't observe \mathbf{z} .
- Our measured variables $\mathbf{x} = \mathbf{z} + \mathbf{u}$, where \mathbf{u} is $k \times 1$ and independent of e and \mathbf{z} .
- $\text{plim} \frac{\mathbf{z}'\mathbf{u}}{n} = 0$: the measurement error is uncorrelated with the truth.
- $\text{plim} \frac{e'\mathbf{u}}{n} = 0$: the measurement error is uncorrelated with the disturbance.

Motivation 1: Measurement Error

- We think that $\mathbb{E}[Y|\mathbf{z}] = \mathbf{z}'\beta$, but we don't observe \mathbf{z} .
- Our measured variables $\mathbf{x} = \mathbf{z} + \mathbf{u}$, where \mathbf{u} is $k \times 1$ and independent of e and \mathbf{z} .
- $\text{plim} \frac{\mathbf{z}'\mathbf{u}}{n} = 0$: the measurement error is uncorrelated with the truth.
- $\text{plim} \frac{e'\mathbf{u}}{n} = 0$: the measurement error is uncorrelated with the disturbance.

Motivation 1: Measurement Error

- We think that $\mathbb{E}[Y|\mathbf{z}] = \mathbf{z}'\beta$, but we don't observe \mathbf{z} .
- Our measured variables $\mathbf{x} = \mathbf{z} + \mathbf{u}$, where \mathbf{u} is $k \times 1$ and independent of e and \mathbf{z} .
- $\text{plim} \frac{\mathbf{z}'\mathbf{u}}{n} = 0$: the measurement error is uncorrelated with the truth.
- $\text{plim} \frac{e'\mathbf{u}}{n} = 0$: the measurement error is uncorrelated with the disturbance.

Motivation 1: Measurement Error

- We think that $\mathbb{E}[Y|\mathbf{z}] = \mathbf{z}'\beta$, but we don't observe \mathbf{z} .
- Our measured variables $\mathbf{x} = \mathbf{z} + \mathbf{u}$, where \mathbf{u} is $k \times 1$ and independent of e and \mathbf{z} .
- $\text{plim} \frac{\mathbf{z}'\mathbf{u}}{n} = 0$: the measurement error is uncorrelated with the truth.
- $\text{plim} \frac{e'\mathbf{u}}{n} = 0$: the measurement error is uncorrelated with the disturbance.

Motivation 1: Measurement Error

- We can rewrite in terms of observed and unobserved variables.

$$\begin{aligned}Y &= \mathbf{z}'\beta + e \\&= (\mathbf{x} - \mathbf{u})'\beta + e \\&= \mathbf{x}'\beta + e - \mathbf{u}'\beta \\&\equiv \mathbf{x}'\beta + \nu\end{aligned}$$

- However,

$$\mathbb{E}[\mathbf{x}\nu] = \mathbb{E}[(\mathbf{z} + \mathbf{u})(e - \mathbf{u}'\beta)] = \mathbb{E}[\mathbf{z}e] + \mathbb{E}[\mathbf{u}e] - \mathbb{E}[\mathbf{z}\mathbf{u}'\beta] - \mathbb{E}[\mathbf{u}\mathbf{u}'\beta] = -\mathbb{E}[\mathbf{u}\mathbf{u}'\beta] \neq 0$$

Motivation 1: Measurement Error

- We can rewrite in terms of observed and unobserved variables.

$$\begin{aligned} Y &= \mathbf{z}'\beta + e \\ &= (\mathbf{x} - \mathbf{u})'\beta + e \\ &= \mathbf{x}'\beta + e - \mathbf{u}'\beta \\ &\equiv \mathbf{x}'\beta + \nu \end{aligned}$$

- However,

$$\mathbb{E}[\mathbf{x}\nu] = \mathbb{E}[(\mathbf{z} + \mathbf{u})(e - \mathbf{u}'\beta)] = \mathbb{E}[\mathbf{z}e] + \mathbb{E}[\mathbf{u}e] - \mathbb{E}[\mathbf{z}\mathbf{u}'\beta] - \mathbb{E}[\mathbf{u}\mathbf{u}'\beta] = -\mathbb{E}[\mathbf{u}\mathbf{u}'\beta] \neq 0$$

Motivation 1: Measurement Error

- We can rewrite in terms of observed and unobserved variables.

$$\begin{aligned} Y &= \mathbf{z}'\beta + e \\ &= (\mathbf{x} - \mathbf{u})'\beta + e \\ &= \mathbf{x}'\beta + e - \mathbf{u}'\beta \\ &\equiv \mathbf{x}'\beta + \nu \end{aligned}$$

- However,

$$\mathbb{E}[\mathbf{x}\nu] = \mathbb{E}[(\mathbf{z} + \mathbf{u})(e - \mathbf{u}'\beta)] = \mathbb{E}[\mathbf{z}e] + \mathbb{E}[\mathbf{u}e] - \mathbb{E}[\mathbf{z}\mathbf{u}'\beta] - \mathbb{E}[\mathbf{u}\mathbf{u}'\beta] = -\mathbb{E}[\mathbf{u}\mathbf{u}'\beta] \neq 0$$

Motivation 1: Measurement Error

- We can rewrite in terms of observed and unobserved variables.

$$\begin{aligned} Y &= \mathbf{z}'\beta + e \\ &= (\mathbf{x} - \mathbf{u})'\beta + e \\ &= \mathbf{x}'\beta + e - \mathbf{u}'\beta \\ &\equiv \mathbf{x}'\beta + \nu \end{aligned}$$

- However,

$$\mathbb{E}[\mathbf{x}\nu] = \mathbb{E}[(\mathbf{z} + \mathbf{u})(e - \mathbf{u}'\beta)] = \mathbb{E}[\mathbf{z}e] + \mathbb{E}[\mathbf{u}e] - \mathbb{E}[\mathbf{z}\mathbf{u}'\beta] - \mathbb{E}[\mathbf{u}\mathbf{u}'\beta] = -\mathbb{E}[\mathbf{u}\mathbf{u}'\beta] \neq 0$$

Motivation 1: Measurement Error

- We can rewrite in terms of observed and unobserved variables.

$$\begin{aligned} Y &= \mathbf{z}'\beta + e \\ &= (\mathbf{x} - \mathbf{u})'\beta + e \\ &= \mathbf{x}'\beta + e - \mathbf{u}'\beta \\ &\equiv \mathbf{x}'\beta + \nu \end{aligned}$$

- However,

$$\mathbb{E}[\mathbf{x}\nu] = \mathbb{E}[(\mathbf{z} + \mathbf{u})(e - \mathbf{u}'\beta)] = \mathbb{E}[\mathbf{z}e] + \mathbb{E}[\mathbf{u}e] - \mathbb{E}[\mathbf{z}\mathbf{u}'\beta] - \mathbb{E}[\mathbf{u}\mathbf{u}'\beta] = -\mathbb{E}[\mathbf{u}\mathbf{u}'\beta] \neq 0$$

Motivation 1: Measurement Error

- We can rewrite in terms of observed and unobserved variables.

$$\begin{aligned} Y &= \mathbf{z}'\beta + e \\ &= (\mathbf{x} - \mathbf{u})'\beta + e \\ &= \mathbf{x}'\beta + e - \mathbf{u}'\beta \\ &\equiv \mathbf{x}'\beta + \nu \end{aligned}$$

- However,

$$\mathbb{E}[\mathbf{x}\nu] = \mathbb{E}[(\mathbf{z} + \mathbf{u})(e - \mathbf{u}'\beta)] = \mathbb{E}[\mathbf{z}e] + \mathbb{E}[\mathbf{u}e] - \mathbb{E}[\mathbf{z}\mathbf{u}'\beta] - \mathbb{E}[\mathbf{u}\mathbf{u}'\beta] = -\mathbb{E}[\mathbf{u}\mathbf{u}']\beta \neq 0$$

OLS estimates

$$\begin{aligned}\hat{\beta} &= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} \\ &= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'(\mathbf{X}\beta + \mathbf{e} - \mathbf{u}\beta) \\ &= \beta + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{e} - (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{u}\beta\end{aligned}$$

OLS estimates

$$\begin{aligned}\hat{\beta} &= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'y \\ &= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'(\mathbf{X}\beta + e - \mathbf{u}\beta) \\ &= \beta + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'e - (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{u}\beta\end{aligned}$$

OLS estimates

$$\begin{aligned}\hat{\beta} &= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'y \\ &= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'(\mathbf{X}\beta + e - \mathbf{u}\beta) \\ &= \beta + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'e - (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{u}\beta\end{aligned}$$

Asymptotic Bias

$$\begin{aligned}
 \text{plim} \hat{\beta} &= \text{plim}[\beta + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{e} - (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{u}\beta] \\
 &= \beta + \text{plim}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{e} - \text{plim}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{u}\beta \\
 &= \beta + (\text{plim}\frac{\mathbf{X}'\mathbf{X}}{n})^{-1}\text{plim}\frac{\mathbf{X}'\mathbf{e}}{n} - (\text{plim}\frac{\mathbf{X}'\mathbf{X}}{n})^{-1}\text{plim}\frac{\mathbf{X}'\mathbf{u}}{n}\beta \\
 &= \beta + \Sigma_X^{-1}\text{plim}\frac{(\mathbf{Z} + \mathbf{u})'\mathbf{e}}{n} - \Sigma_X^{-1}\text{plim}\frac{(\mathbf{Z} + \mathbf{u})'\mathbf{u}}{n}\beta \\
 &= \beta + \Sigma_X^{-1}0 - \Sigma_X^{-1}\Sigma_u\beta \\
 &= (\underbrace{I - \Sigma_X^{-1}\Sigma_u}_{\text{signal noise}})\beta
 \end{aligned}$$

Asymptotic Bias

$$\begin{aligned}
 \text{plim} \hat{\beta} &= \text{plim}[\beta + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{e} - (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{u}\beta] \\
 &= \beta + \text{plim}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{e} - \text{plim}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{u}\beta \\
 &= \beta + (\text{plim}\frac{\mathbf{X}'\mathbf{X}}{n})^{-1}\text{plim}\frac{\mathbf{X}'\mathbf{e}}{n} - (\text{plim}\frac{\mathbf{X}'\mathbf{X}}{n})^{-1}\text{plim}\frac{\mathbf{X}'\mathbf{u}}{n}\beta \\
 &= \beta + \Sigma_X^{-1}\text{plim}\frac{(\mathbf{Z} + \mathbf{u})'\mathbf{e}}{n} - \Sigma_X^{-1}\text{plim}\frac{(\mathbf{Z} + \mathbf{u})'\mathbf{u}}{n}\beta \\
 &= \beta + \Sigma_X^{-1}0 - \Sigma_X^{-1}\Sigma_u\beta \\
 &= (\underbrace{I - \Sigma_X^{-1}}_{\text{signal}} \underbrace{\Sigma_u}_{\text{noise}})\beta
 \end{aligned}$$

Asymptotic Bias

$$\begin{aligned}
 plim \hat{\beta} &= plim[\beta + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'e - (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{u}\beta] \\
 &= \beta + plim(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'e - plim(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{u}\beta \\
 &= \beta + (plim\frac{\mathbf{X}'\mathbf{X}}{n})^{-1}plim\frac{\mathbf{X}'e}{n} - (plim\frac{\mathbf{X}'\mathbf{X}}{n})^{-1}plim\frac{\mathbf{X}'\mathbf{u}}{n}\beta \\
 &= \beta + \Sigma_X^{-1}plim\frac{(\mathbf{Z} + \mathbf{u})'e}{n} - \Sigma_X^{-1}plim\frac{(\mathbf{Z} + \mathbf{u})'\mathbf{u}}{n}\beta \\
 &= \beta + \Sigma_X^{-1}0 - \Sigma_X^{-1}\Sigma_u\beta \\
 &= (I - \underbrace{\Sigma_X^{-1}}_{\text{signal}}\underbrace{\Sigma_u}_{\text{noise}})\beta
 \end{aligned}$$

Asymptotic Bias

$$\begin{aligned}
 plim \hat{\beta} &= plim[\beta + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{e} - (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{u}\beta] \\
 &= \beta + plim(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{e} - plim(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{u}\beta \\
 &= \beta + (plim \frac{\mathbf{X}'\mathbf{X}}{n})^{-1} plim \frac{\mathbf{X}'\mathbf{e}}{n} - (plim \frac{\mathbf{X}'\mathbf{X}}{n})^{-1} plim \frac{\mathbf{X}'\mathbf{u}}{n} \beta \\
 &= \beta + \Sigma_X^{-1} plim \frac{(\mathbf{Z} + \mathbf{u})'\mathbf{e}}{n} - \Sigma_X^{-1} plim \frac{(\mathbf{Z} + \mathbf{u})'\mathbf{u}}{n} \beta \\
 &= \beta + \Sigma_X^{-1} 0 - \Sigma_X^{-1} \Sigma_u \beta \\
 &= (I - \underbrace{\Sigma_X^{-1}}_{\text{signal}} \underbrace{\Sigma_u}_{\text{noise}}) \beta
 \end{aligned}$$

Asymptotic Bias

$$\begin{aligned}
 plim \hat{\beta} &= plim[\beta + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'e - (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{u}\beta] \\
 &= \beta + plim(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'e - plim(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{u}\beta \\
 &= \beta + (plim \frac{\mathbf{X}'\mathbf{X}}{n})^{-1} plim \frac{\mathbf{X}'e}{n} - (plim \frac{\mathbf{X}'\mathbf{X}}{n})^{-1} plim \frac{\mathbf{X}'\mathbf{u}}{n} \beta \\
 &= \beta + \Sigma_X^{-1} plim \frac{(\mathbf{Z} + \mathbf{u})'e}{n} - \Sigma_X^{-1} plim \frac{(\mathbf{Z} + \mathbf{u})'\mathbf{u}}{n} \beta \\
 &= \beta + \Sigma_X^{-1} 0 - \Sigma_X^{-1} \Sigma_u \beta \\
 &= (I - \underbrace{\Sigma_X^{-1}}_{\text{signal}} \underbrace{\Sigma_u}_{\text{noise}}) \beta
 \end{aligned}$$

Asymptotic Bias

$$\begin{aligned}
 plim\hat{\beta} &= plim[\beta + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'e - (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{u}\beta] \\
 &= \beta + plim(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'e - plim(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{u}\beta \\
 &= \beta + (plim\frac{\mathbf{X}'\mathbf{X}}{n})^{-1}plim\frac{\mathbf{X}'e}{n} - (plim\frac{\mathbf{X}'\mathbf{X}}{n})^{-1}plim\frac{\mathbf{X}'\mathbf{u}}{n}\beta \\
 &= \beta + \Sigma_X^{-1}plim\frac{(\mathbf{Z} + \mathbf{u})'e}{n} - \Sigma_X^{-1}plim\frac{(\mathbf{Z} + \mathbf{u})'\mathbf{u}}{n}\beta \\
 &= \beta + \Sigma_X^{-1}0 - \Sigma_X^{-1}\Sigma_u\beta \\
 &= (\underbrace{I - \Sigma_X^{-1}}_{\text{signal}}\underbrace{\Sigma_u}_{\text{noise}})\beta
 \end{aligned}$$

Examples of Measurement Error

- Note, even if only one variable has measurement error, it affects all of the slope coefficients.
- Survey self reports: income, hours worked, assets, occupation, schooling, age, employment status (see Bound, Brown, and Mathiowetz 2001).
- Government reported data on GDP, trade flows, spending.
- Blood pressure measurements with a sphygmomanometer.

Examples of Measurement Error

- Note, even if only one variable has measurement error, it affects all of the slope coefficients.
- Survey self reports: income, hours worked, assets, occupation, schooling, age, employment status (see Bound, Brown, and Mathiowetz 2001).
- Government reported data on GDP, trade flows, spending.
- Blood pressure measurements with a sphygmomanometer.

Examples of Measurement Error

- Note, even if only one variable has measurement error, it affects all of the slope coefficients.
- Survey self reports: income, hours worked, assets, occupation, schooling, age, employment status (see Bound, Brown, and Mathiowetz 2001).
- Government reported data on GDP, trade flows, spending.
- Blood pressure measurements with a sphygmomanometer.

Examples of Measurement Error

- Note, even if only one variable has measurement error, it affects all of the slope coefficients.
- Survey self reports: income, hours worked, assets, occupation, schooling, age, employment status (see Bound, Brown, and Mathiowetz 2001).
- Government reported data on GDP, trade flows, spending.
- Blood pressure measurements with a sphygmomanometer.

Motivation 2: Selection and Endogeneity (Roy Model)

- Let \mathbf{x} affect potential outcomes e.g., education, experience, age
- Let \mathbf{z} affect the selection decision e.g., $\mathbf{z} \supseteq \mathbf{x}$, e.g., distance to college.
- Potential outcomes: $Y_1 = \mathbf{x}'\beta_1 + e_1$, $Y_0 = \mathbf{x}'\beta_0 + e_0$
- Individuals select into treatment if:

$$Y_1 - Y_0 = (\mathbf{x}'\beta_1 - \mathbf{x}'\beta_0) + (e_1 - e_0) > 0$$

- Define the unobserved selection term: $\eta = e_1 - e_0$
- We cannot observe both Y_1 and Y_0 . We model selection as a monotonic function of a latent index:

$$D = 1\{\mathbf{z}'\gamma + \eta > 0\}$$

- If \mathbf{z} includes variables excluded from \mathbf{x} , we have an **exclusion restriction**.
- If η and e_1 are correlated, then:

$$\mathbb{E}[e_1 \mid D = 1] = \mathbb{E}[e_1 \mid \mathbf{z}'\gamma + \eta > 0] \neq 0 \Rightarrow \text{OLS on treated sample is biased}$$

Motivation 2: Selection and Endogeneity (Roy Model)

- Let \mathbf{x} affect potential outcomes e.g., education, experience, age
- Let \mathbf{z} affect the selection decision e.g., $\mathbf{z} \supseteq \mathbf{x}$, e.g., distance to college.

■ Potential outcomes: $Y_1 = \mathbf{x}'\beta_1 + e_1$, $Y_0 = \mathbf{x}'\beta_0 + e_0$

■ Individuals select into treatment if:

$$Y_1 - Y_0 = (\mathbf{x}'\beta_1 - \mathbf{x}'\beta_0) + (e_1 - e_0) > 0$$

■ Define the unobserved selection term: $\eta = e_1 - e_0$

■ We cannot observe both Y_1 and Y_0 . We model selection as a monotonic function of a latent index:

$$D = 1\{\mathbf{z}'\gamma + \eta > 0\}$$

■ If \mathbf{z} includes variables excluded from \mathbf{x} , we have an **exclusion restriction**.

■ If η and e_1 are correlated, then:

$$\mathbb{E}[e_1 \mid D = 1] = \mathbb{E}[e_1 \mid \mathbf{z}'\gamma + \eta > 0] \neq 0 \Rightarrow \text{OLS on treated sample is biased}$$

Motivation 2: Selection and Endogeneity (Roy Model)

- Let \mathbf{x} affect potential outcomes e.g., education, experience, age
- Let \mathbf{z} affect the selection decision e.g., $\mathbf{z} \supseteq \mathbf{x}$, e.g., distance to college.
- Potential outcomes: $Y_1 = \mathbf{x}'\beta_1 + e_1$, $Y_0 = \mathbf{x}'\beta_0 + e_0$
- Individuals select into treatment if:

$$Y_1 - Y_0 = (\mathbf{x}'\beta_1 - \mathbf{x}'\beta_0) + (e_1 - e_0) > 0$$

- Define the unobserved selection term: $\eta = e_1 - e_0$
- We cannot observe both Y_1 and Y_0 . We model selection as a monotonic function of a latent index:

$$D = 1\{\mathbf{z}'\gamma + \eta > 0\}$$

- If \mathbf{z} includes variables excluded from \mathbf{x} , we have an **exclusion restriction**.
- If η and e_1 are correlated, then:

$$\mathbb{E}[e_1 \mid D = 1] = \mathbb{E}[e_1 \mid \mathbf{z}'\gamma + \eta > 0] \neq 0 \Rightarrow \text{OLS on treated sample is biased}$$

Motivation 2: Selection and Endogeneity (Roy Model)

- Let \mathbf{x} affect potential outcomes e.g., education, experience, age
- Let \mathbf{z} affect the selection decision e.g., $\mathbf{z} \supseteq \mathbf{x}$, e.g., distance to college.
- Potential outcomes: $Y_1 = \mathbf{x}'\beta_1 + e_1$, $Y_0 = \mathbf{x}'\beta_0 + e_0$
- Individuals select into treatment if:

$$Y_1 - Y_0 = (\mathbf{x}'\beta_1 - \mathbf{x}'\beta_0) + (e_1 - e_0) > 0$$

- Define the unobserved selection term: $\eta = e_1 - e_0$
- We cannot observe both Y_1 and Y_0 . We model selection as a monotonic function of a latent index:

$$D = 1\{\mathbf{z}'\gamma + \eta > 0\}$$

- If \mathbf{z} includes variables excluded from \mathbf{x} , we have an **exclusion restriction**.
- If η and e_1 are correlated, then:

$$\mathbb{E}[e_1 \mid D = 1] = \mathbb{E}[e_1 \mid \mathbf{z}'\gamma + \eta > 0] \neq 0 \Rightarrow \text{OLS on treated sample is biased}$$

Motivation 2: Selection and Endogeneity (Roy Model)

- Let \mathbf{x} affect potential outcomes e.g., education, experience, age
- Let \mathbf{z} affect the selection decision e.g., $\mathbf{z} \supseteq \mathbf{x}$, e.g., distance to college.
- Potential outcomes: $Y_1 = \mathbf{x}'\beta_1 + e_1$, $Y_0 = \mathbf{x}'\beta_0 + e_0$
- Individuals select into treatment if:

$$Y_1 - Y_0 = (\mathbf{x}'\beta_1 - \mathbf{x}'\beta_0) + (e_1 - e_0) > 0$$

- Define the unobserved selection term: $\eta = e_1 - e_0$
- We cannot observe both Y_1 and Y_0 . We model selection as a monotonic function of a latent index:

$$D = 1\{\mathbf{z}'\gamma + \eta > 0\}$$

- If \mathbf{z} includes variables excluded from \mathbf{x} , we have an **exclusion restriction**.
- If η and e_1 are correlated, then:

$$\mathbb{E}[e_1 \mid D = 1] = \mathbb{E}[e_1 \mid \mathbf{z}'\gamma + \eta > 0] \neq 0 \Rightarrow \text{OLS on treated sample is biased.}$$

Examples of Endogenous Choice

- Endogeneity arises when both treatment (X) and outcome (Y) are shaped by actors making strategic or optimizing decisions.
- These choices are influenced by unobserved factors that also affect outcomes, leading to omitted variable bias.
- **Education and earnings:** Individuals with higher unobserved ability find education easier to obtain *and* earn higher wages.
- **Crisis bargaining:** Countries with strong but unobserved military capabilities are more likely to deter conflict *and* win concessions.
- **Electoral competition:** High-quality incumbents are more likely to both seek reelection *and* win when they do.

Examples of Endogenous Choice

- Endogeneity arises when both treatment (X) and outcome (Y) are shaped by actors making strategic or optimizing decisions.
- These choices are influenced by unobserved factors that also affect outcomes, leading to omitted variable bias.
- **Education and earnings:** Individuals with higher unobserved ability find education easier to obtain *and* earn higher wages.
- **Crisis bargaining:** Countries with strong but unobserved military capabilities are more likely to deter conflict *and* win concessions.
- **Electoral competition:** High-quality incumbents are more likely to both seek reelection *and* win when they do.

Examples of Endogenous Choice

- Endogeneity arises when both treatment (X) and outcome (Y) are shaped by actors making strategic or optimizing decisions.
- These choices are influenced by unobserved factors that also affect outcomes, leading to omitted variable bias.
- **Education and earnings:** Individuals with higher unobserved ability find education easier to obtain *and* earn higher wages.
- **Crisis bargaining:** Countries with strong but unobserved military capabilities are more likely to deter conflict *and* win concessions.
- **Electoral competition:** High-quality incumbents are more likely to both seek reelection *and* win when they do.

Examples of Endogenous Choice

- Endogeneity arises when both treatment (X) and outcome (Y) are shaped by actors making strategic or optimizing decisions.
- These choices are influenced by unobserved factors that also affect outcomes, leading to omitted variable bias.
- **Education and earnings:** Individuals with higher unobserved ability find education easier to obtain *and* earn higher wages.
- **Crisis bargaining:** Countries with strong but unobserved military capabilities are more likely to deter conflict *and* win concessions.
- **Electoral competition:** High-quality incumbents are more likely to both seek reelection *and* win when they do.

Examples of Endogenous Choice

- Endogeneity arises when both treatment (X) and outcome (Y) are shaped by actors making strategic or optimizing decisions.
- These choices are influenced by unobserved factors that also affect outcomes, leading to omitted variable bias.
- **Education and earnings:** Individuals with higher unobserved ability find education easier to obtain *and* earn higher wages.
- **Crisis bargaining:** Countries with strong but unobserved military capabilities are more likely to deter conflict *and* win concessions.
- **Electoral competition:** High-quality incumbents are more likely to both seek reelection *and* win when they do.

Notation for Endogenous Regressors

- Y is a linear function of exogenous variables \mathbf{x}_1 and endogenous variables \mathbf{x}_2
- Call $\mathbf{y}_2 = \mathbf{x}_2$:

$$Y_1 = \mathbf{x}_1' \beta_1 + \mathbf{y}_2' \beta_2 + e$$

$$\mathbb{E}[\mathbf{y}_2 e] \neq 0$$

Notation for Endogenous Regressors

- Y is a linear function of exogenous variables \mathbf{x}_1 and endogenous variables \mathbf{x}_2
- Call $\mathbf{y}_2 = \mathbf{x}_2$:

$$Y_1 = \mathbf{x}_1' \beta_1 + \mathbf{y}_2' \beta_2 + e$$

$$\mathbb{E}[\mathbf{y}_2 e] \neq 0$$

Notation for Endogenous Regressors

- Y is a linear function of exogenous variables \mathbf{x}_1 and endogenous variables \mathbf{x}_2
- Call $\mathbf{y}_2 = \mathbf{x}_2$:

$$Y_1 = \mathbf{x}_1' \beta_1 + \mathbf{y}_2' \beta_2 + e$$

$$\mathbb{E}[\mathbf{y}_2 e] \neq 0$$

Notation for Endogenous Regressors

- Y is a linear function of exogenous variables \mathbf{x}_1 and endogenous variables \mathbf{x}_2
- Call $\mathbf{y}_2 = \mathbf{x}_2$:

$$Y_1 = \mathbf{x}_1' \beta_1 + \mathbf{y}_2' \beta_2 + e$$

$$\mathbb{E}[\mathbf{y}_2 e] \neq 0$$

Notation for Instruments

- There are l instruments, \mathbf{z} , including the k_1 exogenous variables \mathbf{x}_1 and l_2 excluded exogenous variables \mathbf{z}_2 .

$$E[\mathbf{z}e] = 0 \quad (\text{exogeneity})$$

$$E[\mathbf{z}\mathbf{z}] > 0 \quad (\text{non-redundant})$$

$$\text{rank}(\mathbf{z}\mathbf{x}') = k \quad (\text{relevant})$$

- We can write:

$$Y_1 = \mathbf{z}_1'\beta_1 + \mathbf{y}_2'\beta_2 + e$$

- Measurement Error: The excluded endogenous variables offer an alternative measure of the true variable \mathbf{z} .
- Selection: The excluded exogenous variables affect selection but not potential outcomes.

Notation for Instruments

- There are l instruments, \mathbf{z} , including the k_1 exogenous variables \mathbf{x}_1 and l_2 excluded exogenous variables \mathbf{z}_2 .

$$\mathbf{E}[\mathbf{z}e] = 0 \quad (\text{exogeneity})$$

$$\mathbf{E}[\mathbf{z}\mathbf{z}] > 0 \quad (\text{non-redundant})$$

$$\text{rank}(\mathbf{z}\mathbf{x}') = k \quad (\text{relevant})$$

- We can write:

$$Y_1 = \mathbf{z}_1'\beta_1 + \mathbf{y}_2'\beta_2 + e$$

- Measurement Error: The excluded endogenous variables offer an alternative measure of the true variable \mathbf{z} .
- Selection: The excluded exogenous variables affect selection but not potential outcomes.

Notation for Instruments

- There are l instruments, \mathbf{z} , including the k_1 exogenous variables \mathbf{x}_1 and l_2 excluded exogenous variables \mathbf{z}_2 .

$$\mathbf{E}[\mathbf{z}e] = 0 \quad (\text{exogeneity})$$

$$\mathbf{E}[\mathbf{z}\mathbf{z}] > 0 \quad (\text{non-redundant})$$

$$\text{rank}(\mathbf{z}\mathbf{x}') = k \quad (\text{relevant})$$

- We can write:

$$Y_1 = \mathbf{z}_1'\beta_1 + \mathbf{y}_2'\beta_2 + e$$

- Measurement Error: The excluded endogenous variables offer an alternative measure of the true variable \mathbf{z} .
- Selection: The excluded exogenous variables affect selection but not potential outcomes.

Notation for Instruments

- There are l instruments, \mathbf{z} , including the k_1 exogenous variables \mathbf{x}_1 and l_2 excluded exogenous variables \mathbf{z}_2 .

$$E[\mathbf{z}e] = 0 \quad (\text{exogeneity})$$

$$E[\mathbf{z}\mathbf{z}] > 0 \quad (\text{non-redundant})$$

$$\text{rank}(\mathbf{z}\mathbf{x}') = k \quad (\text{relevant})$$

- We can write:

$$Y_1 = \mathbf{z}_1'\beta_1 + \mathbf{y}_2'\beta_2 + e$$

- Measurement Error: The excluded endogenous variables offer an alternative measure of the true variable \mathbf{z} .
- Selection: The excluded exogenous variables affect selection but not potential outcomes.

Notation for Instruments

- There are l instruments, \mathbf{z} , including the k_1 exogenous variables \mathbf{x}_1 and l_2 excluded exogenous variables \mathbf{z}_2 .

$$E[\mathbf{z}e] = 0 \quad (\text{exogeneity})$$

$$E[\mathbf{z}\mathbf{z}] > 0 \quad (\text{non-redundant})$$

$$\text{rank}(\mathbf{z}\mathbf{x}') = k \quad (\text{relevant})$$

- We can write:

$$Y_1 = \mathbf{z}_1'\beta_1 + \mathbf{y}_2'\beta_2 + e$$

- Measurement Error: The excluded endogenous variables offer an alternative measure of the true variable \mathbf{z} .
- Selection: The excluded exogenous variables affect selection but not potential outcomes.

Reduced Form

- The reduced form transforms $Y_1 = \mathbf{x}'_1\beta_1 + \mathbf{y}'_2\beta_2 + e$ to put the endogenous regressors on the left hand side.
- We construct the fitted regressors and implied transformation using reduced-form coefficients:
 - λ — the reduced-form coefficients from regressing Y_1 on instruments \mathbf{z} ,
 - Γ — the reduced-form coefficients from regressing the endogenous regressors \mathbf{y}_2 on \mathbf{z} ,
 - $\bar{\Gamma}$ — the full projection matrix from regressing all regressors $\mathbf{x} = [\mathbf{x}_1, \mathbf{y}_2]$ on \mathbf{z} .
- Note \mathbf{y}_2 is a vector of k_2 endogenous variables, so:

$$\mathbf{y}_2 = \Gamma'\mathbf{z} + \mathbf{u}_2 = \Gamma'_{12}\mathbf{z}_1 + \Gamma'_{22}\mathbf{z}_2 + \mathbf{u}_2$$

where Γ is a $l \times k_2$ matrix defined by linear projection:

$$\Gamma = \mathbb{E}[\mathbf{z}\mathbf{z}']^{-1}\mathbb{E}[\mathbf{z}\mathbf{y}'_2]$$

Reduced Form

- The reduced form transforms $Y_1 = \mathbf{x}'_1\beta_1 + \mathbf{y}'_2\beta_2 + e$ to put the endogenous regressors on the left hand side.
- We construct the fitted regressors and implied transformation using reduced-form coefficients:
 - λ — the reduced-form coefficients from regressing Y_1 on instruments \mathbf{z} ,
 - Γ — the reduced-form coefficients from regressing the endogenous regressors \mathbf{y}_2 on \mathbf{z} ,
 - $\bar{\Gamma}$ — the full projection matrix from regressing all regressors $\mathbf{x} = [\mathbf{x}_1, \mathbf{y}_2]$ on \mathbf{z} .
- Note \mathbf{y}_2 is a vector of k_2 endogenous variables, so:

$$\mathbf{y}_2 = \Gamma' \mathbf{z} + \mathbf{u}_2 = \Gamma'_{12} \mathbf{z}_1 + \Gamma'_{22} \mathbf{z}_2 + \mathbf{u}_2$$

where Γ is a $l \times k_2$ matrix defined by linear projection:

$$\Gamma = \mathbb{E}[\mathbf{z}\mathbf{z}']^{-1} \mathbb{E}[\mathbf{z}\mathbf{y}'_2]$$

Reduced Form

- The reduced form transforms $Y_1 = \mathbf{x}'_1\beta_1 + \mathbf{y}'_2\beta_2 + e$ to put the endogenous regressors on the left hand side.
- We construct the fitted regressors and implied transformation using reduced-form coefficients:
 - λ — the reduced-form coefficients from regressing Y_1 on instruments \mathbf{z} ,
 - Γ — the reduced-form coefficients from regressing the endogenous regressors \mathbf{y}_2 on \mathbf{z} ,
 - $\bar{\Gamma}$ — the full projection matrix from regressing all regressors $\mathbf{x} = [\mathbf{x}_1, \mathbf{y}_2]$ on \mathbf{z} .
- Note \mathbf{y}_2 is a vector of k_2 endogenous variables, so:

$$\mathbf{y}_2 = \Gamma' \mathbf{z} + \mathbf{u}_2 = \Gamma'_{12} \mathbf{z}_1 + \Gamma'_{22} \mathbf{z}_2 + \mathbf{u}_2$$

where Γ is a $l \times k_2$ matrix defined by linear projection:

$$\Gamma = \mathbb{E}[\mathbf{z}\mathbf{z}']^{-1} \mathbb{E}[\mathbf{z}\mathbf{y}'_2]$$

Reduced Form

- The reduced form transforms $Y_1 = \mathbf{x}'_1\beta_1 + \mathbf{y}'_2\beta_2 + e$ to put the endogenous regressors on the left hand side.
- We construct the fitted regressors and implied transformation using reduced-form coefficients:
 - λ — the reduced-form coefficients from regressing Y_1 on instruments \mathbf{z} ,
 - Γ — the reduced-form coefficients from regressing the endogenous regressors \mathbf{y}_2 on \mathbf{z} ,
 - $\bar{\Gamma}$ — the full projection matrix from regressing all regressors $\mathbf{x} = [\mathbf{x}_1, \mathbf{y}_2]$ on \mathbf{z} .
- Note \mathbf{y}_2 is a vector of k_2 endogenous variables, so:

$$\mathbf{y}_2 = \Gamma' \mathbf{z} + \mathbf{u}_2 = \Gamma'_{12} \mathbf{z}_1 + \Gamma'_{22} \mathbf{z}_2 + \mathbf{u}_2$$

where Γ is a $l \times k_2$ matrix defined by linear projection:

$$\Gamma = \mathbb{E}[\mathbf{z}\mathbf{z}']^{-1}\mathbb{E}[\mathbf{z}\mathbf{y}'_2]$$

Reduced Form

- The reduced form transforms $Y_1 = \mathbf{x}'_1\beta_1 + \mathbf{y}'_2\beta_2 + e$ to put the endogenous regressors on the left hand side.
- We construct the fitted regressors and implied transformation using reduced-form coefficients:
 - λ — the reduced-form coefficients from regressing Y_1 on instruments \mathbf{z} ,
 - Γ — the reduced-form coefficients from regressing the endogenous regressors \mathbf{y}_2 on \mathbf{z} ,
 - $\bar{\Gamma}$ — the full projection matrix from regressing all regressors $\mathbf{x} = [\mathbf{x}_1, \mathbf{y}_2]$ on \mathbf{z} .
- Note \mathbf{y}_2 is a vector of k_2 endogenous variables, so:

$$\mathbf{y}_2 = \Gamma' \mathbf{z} + \mathbf{u}_2 = \Gamma'_{12} \mathbf{z}_1 + \Gamma'_{22} \mathbf{z}_2 + \mathbf{u}_2$$

where Γ is a $l \times k_2$ matrix defined by linear projection:

$$\Gamma = \mathbb{E}[\mathbf{z}\mathbf{z}']^{-1} \mathbb{E}[\mathbf{z}\mathbf{y}'_2]$$

Reduced Form

- The reduced form transforms $Y_1 = \mathbf{x}'_1\beta_1 + \mathbf{y}'_2\beta_2 + e$ to put the endogenous regressors on the left hand side.
- We construct the fitted regressors and implied transformation using reduced-form coefficients:
 - λ — the reduced-form coefficients from regressing Y_1 on instruments \mathbf{z} ,
 - Γ — the reduced-form coefficients from regressing the endogenous regressors \mathbf{y}_2 on \mathbf{z} ,
 - $\bar{\Gamma}$ — the full projection matrix from regressing all regressors $\mathbf{x} = [\mathbf{x}_1, \mathbf{y}_2]$ on \mathbf{z} .
- Note \mathbf{y}_2 is a vector of k_2 endogenous variables, so:

$$\mathbf{y}_2 = \Gamma' \mathbf{z} + \mathbf{u}_2 = \Gamma'_{12} \mathbf{z}_1 + \Gamma'_{22} \mathbf{z}_2 + \mathbf{u}_2$$

where Γ is a $l \times k_2$ matrix defined by linear projection:

$$\Gamma = \mathbb{E}[\mathbf{z}\mathbf{z}']^{-1} \mathbb{E}[\mathbf{z}\mathbf{y}'_2]$$

Reduced Form

- Plugging into the equation for Y_1

$$\begin{aligned}Y_1 &= \mathbf{z}'_1 \beta_1 + (\Gamma'_{12} \mathbf{z}_1 + \Gamma'_{22} \mathbf{z}_2 + \mathbf{u}_2) \beta_2 + e \\&= \mathbf{z}'_1 (\beta_1 + \Gamma'_{12} \beta_2) + \mathbf{z}'_2 (\Gamma'_{22} \beta_2) + (\mathbf{u}'_2 \beta_2 + e) \\&= \mathbf{z}'_1 \lambda_1 + \mathbf{z}'_2 \lambda_2 + u_1 \\&= \mathbf{z}' \lambda + u_1\end{aligned}$$

- We can write $\lambda = \begin{bmatrix} I_{k_1} & \Gamma_{12} \\ 0 & \Gamma_{22} \end{bmatrix} \beta = \begin{bmatrix} I_{k_1} & \Gamma \\ 0 & \end{bmatrix} \beta \equiv \bar{\Gamma} \beta$

Reduced Form

- Plugging into the equation for Y_1

$$\begin{aligned}Y_1 &= \mathbf{z}_1' \beta_1 + (\Gamma_{12}' \mathbf{z}_1 + \Gamma_{22}' \mathbf{z}_2 + \mathbf{u}_2) \beta_2 + e \\&= \mathbf{z}_1' (\beta_1 + \Gamma_{12}' \beta_2) + \mathbf{z}_2' (\Gamma_{22}' \beta_2) + (\mathbf{u}_2' \beta_2 + e) \\&= \mathbf{z}_1' \lambda_1 + \mathbf{z}_2' \lambda_2 + u_1 \\&= \mathbf{z}' \lambda + u_1\end{aligned}$$

- We can write $\lambda = \begin{bmatrix} I_{k_1} & \Gamma_{12}' \\ 0 & \Gamma_{22}' \end{bmatrix} \beta = \begin{bmatrix} I_{k_1} & \Gamma \\ 0 & \end{bmatrix} \beta \equiv \bar{\Gamma} \beta$

Reduced Form

- Plugging into the equation for Y_1

$$\begin{aligned}Y_1 &= \mathbf{z}'_1\beta_1 + (\Gamma'_{12}\mathbf{z}_1 + \Gamma'_{22}\mathbf{z}_2 + \mathbf{u}_2)\beta_2 + e \\&= \mathbf{z}'_1(\beta_1 + \Gamma'_{12}\beta_2) + \mathbf{z}'_2(\Gamma'_{22}\beta_2) + (\mathbf{u}'_2\beta_2 + e) \\&= \mathbf{z}'_1\lambda_1 + \mathbf{z}'_2\lambda_2 + u_1 \\&= \mathbf{z}'\lambda + u_1\end{aligned}$$

- We can write $\lambda = \begin{bmatrix} I_{k_1} & \Gamma'_{12} \\ 0 & \Gamma'_{22} \end{bmatrix} \beta = \begin{bmatrix} I_{k_1} & \Gamma \\ 0 & \Gamma \end{bmatrix} \beta \equiv \bar{\Gamma}\beta$

Reduced Form

- Plugging into the equation for Y_1

$$\begin{aligned}Y_1 &= \mathbf{z}'_1\beta_1 + (\Gamma'_{12}\mathbf{z}_1 + \Gamma'_{22}\mathbf{z}_2 + \mathbf{u}_2)\beta_2 + e \\&= \mathbf{z}'_1(\beta_1 + \Gamma'_{12}\beta_2) + \mathbf{z}'_2(\Gamma'_{22}\beta_2) + (\mathbf{u}'_2\beta_2 + e) \\&= \mathbf{z}'_1\lambda_1 + \mathbf{z}'_2\lambda_2 + u_1 \\&= \mathbf{z}'\lambda + u_1\end{aligned}$$

- We can write $\lambda = \begin{bmatrix} I_{k_1} & \Gamma'_{12} \\ 0 & \Gamma'_{22} \end{bmatrix} \beta = \begin{bmatrix} I_{k_1} & \Gamma \\ 0 & \Gamma \end{bmatrix} \beta \equiv \bar{\Gamma}\beta$

Reduced Form

- Plugging into the equation for Y_1

$$\begin{aligned}Y_1 &= \mathbf{z}'_1\beta_1 + (\Gamma'_{12}\mathbf{z}_1 + \Gamma'_{22}\mathbf{z}_2 + \mathbf{u}_2)\beta_2 + e \\&= \mathbf{z}'_1(\beta_1 + \Gamma'_{12}\beta_2) + \mathbf{z}'_2(\Gamma'_{22}\beta_2) + (\mathbf{u}'_2\beta_2 + e) \\&= \mathbf{z}'_1\lambda_1 + \mathbf{z}'_2\lambda_2 + u_1 \\&= \mathbf{z}'\lambda + u_1\end{aligned}$$

- We can write $\lambda = \begin{bmatrix} I_{k_1} & \Gamma_{12} \\ 0 & \Gamma_{22} \end{bmatrix} \beta = \begin{bmatrix} I_{k_1} & \Gamma \\ 0 & \end{bmatrix} \beta \equiv \bar{\Gamma}\beta$

Reduced Form

- β_1 and β_2 are the structural parameters.
- Γ and λ are the reduced form parameters.

$$Y_1 = \lambda'Z + u_1$$

$$Y_2 = \Gamma'Z + u_2$$

Reduced Form

- β_1 and β_2 are the structural parameters.
- Γ and λ are the reduced form parameters.

$$Y_1 = \lambda'Z + u_1$$

$$Y_2 = \Gamma'Z + u_2$$

Reduced Form

- β_1 and β_2 are the structural parameters.
- Γ and λ are the reduced form parameters.

$$Y_1 = \lambda'Z + u_1$$

$$Y_2 = \Gamma'Z + u_2$$

Identified and Over-identified

- A parameter is **identified** if it is a unique function of the probability distribution of the observables, for example, population moments.
- Recall, the linear projection model is identified if $E[\mathbf{x}\mathbf{x}'] > 0$.
- We say the Instrumental Variables model is **just-identified** if $l = k$ and **over-identified** if $l > k$

Identified and Over-identified

- A parameter is **identified** if it is a unique function of the probability distribution of the observables, for example, population moments.
- Recall, the linear projection model is identified if $E[\mathbf{x}\mathbf{x}'] > 0$.
- We say the Instrumental Variables model is **just-identified** if $l = k$ and **over-identified** if $l > k$

Identified and Over-identified

- A parameter is **identified** if it is a unique function of the probability distribution of the observables, for example, population moments.
- Recall, the linear projection model is identified if $E[\mathbf{x}\mathbf{x}'] > 0$.
- We say the Instrumental Variables model is **just-identified** if $l = k$ and **over-identified** if $l > k$

Relevance

- Recall $\mathbf{x} = [\mathbf{x}_1, \mathbf{y}_2]$,

$$\bar{\Gamma} = \begin{bmatrix} I_{k_1} & \Gamma \\ \mathbf{0} & \end{bmatrix} = \begin{bmatrix} I_{k_1} & \mathbb{E}[\mathbf{z}\mathbf{z}']^{-1}\mathbb{E}[\mathbf{z}\mathbf{y}_2'] \\ \mathbf{0} & \end{bmatrix}$$

$$= \mathbb{E}[\mathbf{z}\mathbf{z}']^{-1} \left(\begin{bmatrix} \mathbb{E}[\mathbf{z}\mathbf{z}'] I_{k_1} \\ \mathbf{0} \end{bmatrix} \quad \mathbb{E}[\mathbf{z}\mathbf{y}_2'] \right)$$

$$= \mathbb{E}[\mathbf{z}\mathbf{z}']^{-1}\mathbb{E}[\mathbf{z}\mathbf{x}']$$

$$\lambda = \mathbb{E}[\mathbf{z}\mathbf{z}']^{-1}\mathbb{E}[\mathbf{z}\mathbf{x}']\beta$$

$$\mathbb{E}[\mathbf{z}\mathbf{z}']^{-1}\mathbb{E}[\mathbf{z}\mathbf{y}_1] = \mathbb{E}[\mathbf{z}\mathbf{z}']^{-1}\mathbb{E}[\mathbf{z}\mathbf{x}']\beta$$

$$\mathbb{E}[\mathbf{z}\mathbf{y}_1] = \mathbb{E}[\mathbf{z}\mathbf{x}']\beta$$

- To have a solution, $\text{rank}(\mathbb{E}[\mathbf{z}\mathbf{x}']) = k$, what is called the **relevance condition**.
- If $\bar{\Gamma}$ is rank k , then $\beta = (\bar{\Gamma}'\bar{\Gamma})^{-1}\bar{\Gamma}'\lambda$

Relevance

- Recall $\mathbf{x} = [\mathbf{x}_1, \mathbf{y}_2]$,

$$\bar{\Gamma} = \begin{bmatrix} I_{k_1} & \Gamma \\ \mathbf{0} & \end{bmatrix} = \begin{bmatrix} I_{k_1} & \mathbb{E}[\mathbf{z}\mathbf{z}']^{-1}\mathbb{E}[\mathbf{z}\mathbf{y}_2'] \\ \mathbf{0} & \end{bmatrix}$$

$$= \mathbb{E}[\mathbf{z}\mathbf{z}']^{-1} \left(\begin{bmatrix} \mathbb{E}[\mathbf{z}\mathbf{z}'] I_{k_1} \\ \mathbf{0} \end{bmatrix} \quad \mathbb{E}[\mathbf{z}\mathbf{y}_2'] \right)$$

$$= \mathbb{E}[\mathbf{z}\mathbf{z}']^{-1}\mathbb{E}[\mathbf{z}\mathbf{x}']$$

$$\lambda = \mathbb{E}[\mathbf{z}\mathbf{z}']^{-1}\mathbb{E}[\mathbf{z}\mathbf{x}']\beta$$

$$\mathbb{E}[\mathbf{z}\mathbf{z}']^{-1}\mathbb{E}[\mathbf{z}\mathbf{y}_1] = \mathbb{E}[\mathbf{z}\mathbf{z}']^{-1}\mathbb{E}[\mathbf{z}\mathbf{x}']\beta$$

$$\mathbb{E}[\mathbf{z}\mathbf{y}_1] = \mathbb{E}[\mathbf{z}\mathbf{x}']\beta$$

- To have a solution, $\text{rank}(\mathbb{E}[\mathbf{z}\mathbf{x}']) = k$, what is called the **relevance condition**.
- If $\bar{\Gamma}$ is rank k , then $\beta = (\bar{\Gamma}'\bar{\Gamma})^{-1}\bar{\Gamma}'\lambda$

Relevance

- Recall $\mathbf{x} = [\mathbf{x}_1, \mathbf{y}_2]$,

$$\bar{\Gamma} = \begin{bmatrix} \mathbf{I}_{k_1} & \Gamma \\ \mathbf{0} & \end{bmatrix} = \begin{bmatrix} \mathbf{I}_{k_1} & \mathbb{E}[\mathbf{z}\mathbf{z}']^{-1}\mathbb{E}[\mathbf{z}\mathbf{y}_2'] \\ \mathbf{0} & \end{bmatrix}$$

$$= \mathbb{E}[\mathbf{z}\mathbf{z}']^{-1} \left(\begin{bmatrix} \mathbb{E}[\mathbf{z}\mathbf{z}'] \mathbf{I}_{k_1} \\ \mathbf{0} \end{bmatrix} \quad \mathbb{E}[\mathbf{z}\mathbf{y}_2'] \right)$$

$$= \mathbb{E}[\mathbf{z}\mathbf{z}']^{-1} \mathbb{E}[\mathbf{z}\mathbf{x}']$$

$$\lambda = \mathbb{E}[\mathbf{z}\mathbf{z}']^{-1} \mathbb{E}[\mathbf{z}\mathbf{x}'] \beta$$

$$\mathbb{E}[\mathbf{z}\mathbf{z}']^{-1} \mathbb{E}[\mathbf{z}\mathbf{y}_1] = \mathbb{E}[\mathbf{z}\mathbf{z}']^{-1} \mathbb{E}[\mathbf{z}\mathbf{x}'] \beta$$

$$\mathbb{E}[\mathbf{z}\mathbf{y}_1] = \mathbb{E}[\mathbf{z}\mathbf{x}'] \beta$$

- To have a solution, $\text{rank}(\mathbb{E}[\mathbf{z}\mathbf{x}']) = k$, what is called the **relevance condition**.
- If $\bar{\Gamma}$ is rank k , then $\beta = (\bar{\Gamma}'\bar{\Gamma})^{-1}\bar{\Gamma}'\lambda$

Relevance

- Recall $\mathbf{x} = [\mathbf{x}_1, \mathbf{y}_2]$,

$$\bar{\Gamma} = \begin{bmatrix} \mathbf{I}_{k_1} & \Gamma \\ \mathbf{0} & \end{bmatrix} = \begin{bmatrix} \mathbf{I}_{k_1} & \mathbb{E}[\mathbf{z}\mathbf{z}']^{-1}\mathbb{E}[\mathbf{z}\mathbf{y}_2'] \\ \mathbf{0} & \end{bmatrix}$$

$$= \mathbb{E}[\mathbf{z}\mathbf{z}']^{-1} \left(\begin{bmatrix} \mathbb{E}[\mathbf{z}\mathbf{z}'] \mathbf{I}_{k_1} \\ \mathbf{0} \end{bmatrix} \quad \mathbb{E}[\mathbf{z}\mathbf{y}_2'] \right)$$

$$= \mathbb{E}[\mathbf{z}\mathbf{z}']^{-1} \mathbb{E}[\mathbf{z}\mathbf{x}']$$

$$\lambda = \mathbb{E}[\mathbf{z}\mathbf{z}']^{-1} \mathbb{E}[\mathbf{z}\mathbf{x}'] \beta$$

$$\mathbb{E}[\mathbf{z}\mathbf{z}']^{-1} \mathbb{E}[\mathbf{z}\mathbf{y}_1] = \mathbb{E}[\mathbf{z}\mathbf{z}']^{-1} \mathbb{E}[\mathbf{z}\mathbf{x}'] \beta$$

$$\mathbb{E}[\mathbf{z}\mathbf{y}_1] = \mathbb{E}[\mathbf{z}\mathbf{x}'] \beta$$

- To have a solution, $\text{rank}(\mathbb{E}[\mathbf{z}\mathbf{x}']) = k$, what is called the **relevance condition**.
- If $\bar{\Gamma}$ is rank k , then $\beta = (\bar{\Gamma}'\bar{\Gamma})^{-1}\bar{\Gamma}'\lambda$

Relevance

- Recall $\mathbf{x} = [\mathbf{x}_1, \mathbf{y}_2]$,

$$\bar{\Gamma} = \begin{bmatrix} I_{k_1} & \Gamma \\ \mathbf{0} & \end{bmatrix} = \begin{bmatrix} I_{k_1} & \mathbb{E}[\mathbf{z}\mathbf{z}']^{-1}\mathbb{E}[\mathbf{z}\mathbf{y}_2'] \\ \mathbf{0} & \end{bmatrix}$$

$$= \mathbb{E}[\mathbf{z}\mathbf{z}']^{-1} \left(\begin{bmatrix} \mathbb{E}[\mathbf{z}\mathbf{z}'] I_{k_1} \\ \mathbf{0} \end{bmatrix} \quad \mathbb{E}[\mathbf{z}\mathbf{y}_2'] \right)$$

$$= \mathbb{E}[\mathbf{z}\mathbf{z}']^{-1}\mathbb{E}[\mathbf{z}\mathbf{x}']$$

$$\lambda = \mathbb{E}[\mathbf{z}\mathbf{z}']^{-1}\mathbb{E}[\mathbf{z}\mathbf{x}']\beta$$

$$\mathbb{E}[\mathbf{z}\mathbf{z}']^{-1}\mathbb{E}[\mathbf{z}\mathbf{y}_1] = \mathbb{E}[\mathbf{z}\mathbf{z}']^{-1}\mathbb{E}[\mathbf{z}\mathbf{x}']\beta$$

$$\mathbb{E}[\mathbf{z}\mathbf{y}_1] = \mathbb{E}[\mathbf{z}\mathbf{x}']\beta$$

- To have a solution, $\text{rank}(\mathbb{E}[\mathbf{z}\mathbf{x}']) = k$, what is called the **relevance condition**.
- If $\bar{\Gamma}$ is rank k , then $\beta = (\bar{\Gamma}'\bar{\Gamma})^{-1}\bar{\Gamma}'\lambda$

Relevance

- Recall $\mathbf{x} = [\mathbf{x}_1, \mathbf{y}_2]$,

$$\bar{\Gamma} = \begin{bmatrix} I_{k_1} & \Gamma \\ \mathbf{0} & \end{bmatrix} = \begin{bmatrix} I_{k_1} & \mathbb{E}[\mathbf{z}\mathbf{z}']^{-1}\mathbb{E}[\mathbf{z}\mathbf{y}_2'] \\ \mathbf{0} & \end{bmatrix}$$

$$= \mathbb{E}[\mathbf{z}\mathbf{z}']^{-1} \left(\begin{bmatrix} \mathbb{E}[\mathbf{z}\mathbf{z}'] I_{k_1} \\ \mathbf{0} \end{bmatrix} \quad \mathbb{E}[\mathbf{z}\mathbf{y}_2'] \right)$$

$$= \mathbb{E}[\mathbf{z}\mathbf{z}']^{-1}\mathbb{E}[\mathbf{z}\mathbf{x}']$$

$$\lambda = \mathbb{E}[\mathbf{z}\mathbf{z}']^{-1}\mathbb{E}[\mathbf{z}\mathbf{x}']\beta$$

$$\mathbb{E}[\mathbf{z}\mathbf{z}']^{-1}\mathbb{E}[\mathbf{z}\mathbf{y}_1] = \mathbb{E}[\mathbf{z}\mathbf{z}']^{-1}\mathbb{E}[\mathbf{z}\mathbf{x}']\beta$$

$$\mathbb{E}[\mathbf{z}\mathbf{y}_1] = \mathbb{E}[\mathbf{z}\mathbf{x}']\beta$$

- To have a solution, $\text{rank}(\mathbb{E}[\mathbf{z}\mathbf{x}']) = k$, what is called the **relevance condition**.
- If $\bar{\Gamma}$ is rank k , then $\beta = (\bar{\Gamma}'\bar{\Gamma})^{-1}\bar{\Gamma}'\lambda$

Estimation: Instrumental Variables Estimator

- If $l = k$, we can write a system of $l = k$ equations and k unknowns:

$$\mathbb{E}[\mathbf{z}e] = 0$$

$$\mathbb{E}[\mathbf{z}(Y_1 - \mathbf{x}'\beta)] = 0$$

$$\mathbb{E}[\mathbf{z}Y_1] - \mathbb{E}[\mathbf{z}\mathbf{x}']\beta = 0$$

$$\beta = (\mathbb{E}[\mathbf{z}\mathbf{x}'])^{-1}\mathbb{E}[\mathbf{z}Y_1']$$

$$\hat{\beta}_{iv} = (\mathbf{Z}'\mathbf{X})^{-1}\mathbf{Z}'\mathbf{y}_1$$

Estimation: Instrumental Variables Estimator

- If $l = k$, we can write a system of $l = k$ equations and k unknowns:

$$\mathbb{E}[\mathbf{z}e] = 0$$

$$\mathbb{E}[\mathbf{z}(Y_1 - \mathbf{x}'\beta)] = 0$$

$$\mathbb{E}[\mathbf{z}Y_1] - \mathbb{E}[\mathbf{z}\mathbf{x}']\beta = 0$$

$$\beta = (\mathbb{E}[\mathbf{z}\mathbf{x}'])^{-1}\mathbb{E}[\mathbf{z}Y_1']$$

$$\hat{\beta}_{iv} = (\mathbf{Z}'\mathbf{X})^{-1}\mathbf{Z}'\mathbf{y}_1$$

Estimation: Instrumental Variables Estimator

- If $l = k$, we can write a system of $l = k$ equations and k unknowns:

$$\mathbb{E}[\mathbf{z}e] = 0$$

$$\mathbb{E}[\mathbf{z}(Y_1 - \mathbf{x}'\beta)] = 0$$

$$\mathbb{E}[\mathbf{z}Y_1] - \mathbb{E}[\mathbf{z}\mathbf{x}']\beta = 0$$

$$\beta = (\mathbb{E}[\mathbf{z}\mathbf{x}'])^{-1}\mathbb{E}[\mathbf{z}Y_1]$$

$$\hat{\beta}_{iv} = (\mathbf{Z}'\mathbf{X})^{-1}\mathbf{Z}'\mathbf{y}_1$$

Estimation: Instrumental Variables Estimator

- If $l = k$, we can write a system of $l = k$ equations and k unknowns:

$$\mathbb{E}[\mathbf{z}e] = 0$$

$$\mathbb{E}[\mathbf{z}(Y_1 - \mathbf{x}'\beta)] = 0$$

$$\mathbb{E}[\mathbf{z}Y_1] - \mathbb{E}[\mathbf{z}\mathbf{x}']\beta = 0$$

$$\beta = (\mathbb{E}[\mathbf{z}\mathbf{x}'])^{-1}\mathbb{E}[\mathbf{z}Y_1']$$

$$\hat{\beta}_{iv} = (\mathbf{Z}'\mathbf{X})^{-1}\mathbf{Z}'\mathbf{y}_1$$

Estimation: Instrumental Variables Estimator

- If $l = k$, we can write a system of $l = k$ equations and k unknowns:

$$\mathbb{E}[\mathbf{z}e] = 0$$

$$\mathbb{E}[\mathbf{z}(Y_1 - \mathbf{x}'\beta)] = 0$$

$$\mathbb{E}[\mathbf{z}Y_1] - \mathbb{E}[\mathbf{z}\mathbf{x}']\beta = 0$$

$$\beta = (\mathbb{E}[\mathbf{z}\mathbf{x}'])^{-1}\mathbb{E}[\mathbf{z}Y_1']$$

$$\hat{\beta}_{iv} = (\mathbf{Z}'\mathbf{X})^{-1}\mathbf{Z}'\mathbf{y}_1$$

Estimation: Indirect Least Squares

- We can rewrite in reduced form:

$$\beta = \bar{\Gamma}^{-1}\lambda$$

$$\hat{\beta}_{ils} = \hat{\bar{\Gamma}}^{-1}\hat{\lambda}$$

$$= ((Z'Z)^{-1}(Z'X))^{-1}(Z'Z)^{-1}(Z'y_1)$$

$$= (Z'X)^{-1}(Z'Z)(Z'Z)^{-1}(Z'y_1)$$

$$= (Z'X)^{-1}(Z'y_1)$$

$$= \hat{\beta}_{iv}$$

Estimation: Indirect Least Squares

- We can rewrite in reduced form:

$$\begin{aligned}\beta &= \bar{\Gamma}^{-1}\lambda \\ \hat{\beta}_{ils} &= \hat{\bar{\Gamma}}^{-1}\hat{\lambda} \\ &= ((\mathbf{Z}'\mathbf{Z})^{-1}(\mathbf{Z}'\mathbf{X}))^{-1}(\mathbf{Z}'\mathbf{Z})^{-1}(\mathbf{Z}'\mathbf{y}_1) \\ &= (\mathbf{Z}'\mathbf{X})^{-1}(\mathbf{Z}'\mathbf{Z})(\mathbf{Z}'\mathbf{Z})^{-1}(\mathbf{Z}'\mathbf{y}_1) \\ &= (\mathbf{Z}'\mathbf{X})^{-1}(\mathbf{Z}'\mathbf{y}_1) \\ &= \hat{\beta}_{iv}\end{aligned}$$

Estimation: Indirect Least Squares

- We can rewrite in reduced form:

$$\begin{aligned}\beta &= \bar{\Gamma}^{-1}\lambda \\ \hat{\beta}_{ils} &= \hat{\bar{\Gamma}}^{-1}\hat{\lambda} \\ &= ((\mathbf{Z}'\mathbf{Z})^{-1}(\mathbf{Z}'\mathbf{X}))^{-1}(\mathbf{Z}'\mathbf{Z})^{-1}(\mathbf{Z}'\mathbf{y}_1) \\ &= (\mathbf{Z}'\mathbf{X})^{-1}(\mathbf{Z}'\mathbf{Z})(\mathbf{Z}'\mathbf{Z})^{-1}(\mathbf{Z}'\mathbf{y}_1) \\ &= (\mathbf{Z}'\mathbf{X})^{-1}(\mathbf{Z}'\mathbf{y}_1) \\ &= \hat{\beta}_{iv}\end{aligned}$$

Estimation: Indirect Least Squares

- We can rewrite in reduced form:

$$\begin{aligned}\beta &= \bar{\Gamma}^{-1}\lambda \\ \hat{\beta}_{ils} &= \hat{\bar{\Gamma}}^{-1}\hat{\lambda} \\ &= ((\mathbf{Z}'\mathbf{Z})^{-1}(\mathbf{Z}'\mathbf{X}))^{-1}(\mathbf{Z}'\mathbf{Z})^{-1}(\mathbf{Z}'\mathbf{y}_1) \\ &= (\mathbf{Z}'\mathbf{X})^{-1}(\mathbf{Z}'\mathbf{Z})(\mathbf{Z}'\mathbf{Z})^{-1}(\mathbf{Z}'\mathbf{y}_1) \\ &= (\mathbf{Z}'\mathbf{X})^{-1}(\mathbf{Z}'\mathbf{y}_1) \\ &= \hat{\beta}_{iv}\end{aligned}$$

Estimation: Indirect Least Squares

- We can rewrite in reduced form:

$$\begin{aligned}\beta &= \bar{\Gamma}^{-1}\lambda \\ \hat{\beta}_{ils} &= \hat{\bar{\Gamma}}^{-1}\hat{\lambda} \\ &= ((\mathbf{Z}'\mathbf{Z})^{-1}(\mathbf{Z}'\mathbf{X}))^{-1}(\mathbf{Z}'\mathbf{Z})^{-1}(\mathbf{Z}'\mathbf{y}_1) \\ &= (\mathbf{Z}'\mathbf{X})^{-1}(\mathbf{Z}'\mathbf{Z})(\mathbf{Z}'\mathbf{Z})^{-1}(\mathbf{Z}'\mathbf{y}_1) \\ &= (\mathbf{Z}'\mathbf{X})^{-1}(\mathbf{Z}'\mathbf{y}_1) \\ &= \hat{\beta}_{iv}\end{aligned}$$

A Simple IV Estimator: The Wald Estimator

- Consider the scalar model: $Y = X\beta + e$, let $Z \in \{0, 1\}$ be a binary instrument.
- The IV estimator is:

$$\hat{\beta}^{IV} = \frac{\text{Cov}(Z, Y)}{\text{Cov}(Z, X)}$$

$$\begin{aligned}\text{Cov}(Z, Y) &= \mathbb{E}[ZY] - \text{Pr}(Z = 1) * \mathbb{E}[Y] \\ &= p\mathbb{E}[Y|Z = 1] - p * [\mathbb{E}[Y|Z = 1] * p + [\mathbb{E}[Y|Z = 0] * (1 - p)]] \\ &= p\mathbb{E}[Y|Z = 1] - pp\mathbb{E}[Y|Z = 1] - p[\mathbb{E}[Y|Z = 0](1 - p)] \\ &= p(1 - p) * ([\mathbb{E}[Y|Z = 1]] - [\mathbb{E}[Y|Z = 0]]) \\ \text{Cov}(Z, X) &= p(1 - p) * ([\mathbb{E}[X|Z = 1]] - [\mathbb{E}[X|Z = 0]])\end{aligned}$$

- The ratio gives the **Wald estimator**:

$$\hat{\beta}^{IV} = \frac{\mathbb{E}[Y | Z = 1] - \mathbb{E}[Y | Z = 0]}{\mathbb{E}[X | Z = 1] - \mathbb{E}[X | Z = 0]}$$

A Simple IV Estimator: The Wald Estimator

- Consider the scalar model: $Y = X\beta + e$, let $Z \in \{0, 1\}$ be a binary instrument.
- The IV estimator is:

$$\hat{\beta}^{IV} = \frac{\text{Cov}(Z, Y)}{\text{Cov}(Z, X)}$$

$$\begin{aligned}\text{Cov}(Z, Y) &= \mathbb{E}[ZY] - \text{Pr}(Z = 1) * \mathbb{E}[Y] \\ &= p\mathbb{E}[Y|Z = 1] - p * [\mathbb{E}[Y|Z = 1] * p + [\mathbb{E}[Y|Z = 0] * (1 - p)]] \\ &= p\mathbb{E}[Y|Z = 1] - pp\mathbb{E}[Y|Z = 1] - p[\mathbb{E}[Y|Z = 0](1 - p)] \\ &= p(1 - p) * ([\mathbb{E}[Y|Z = 1]] - [\mathbb{E}[Y|Z = 0]]) \\ \text{Cov}(Z, X) &= p(1 - p) * ([\mathbb{E}[X|Z = 1]] - [\mathbb{E}[X|Z = 0]])\end{aligned}$$

- The ratio gives the **Wald estimator**:

$$\hat{\beta}^{IV} = \frac{\mathbb{E}[Y | Z = 1] - \mathbb{E}[Y | Z = 0]}{\mathbb{E}[X | Z = 1] - \mathbb{E}[X | Z = 0]}$$

A Simple IV Estimator: The Wald Estimator

- Consider the scalar model: $Y = X\beta + e$, let $Z \in \{0, 1\}$ be a binary instrument.
- The IV estimator is:

$$\hat{\beta}^{IV} = \frac{\text{Cov}(Z, Y)}{\text{Cov}(Z, X)}$$

$$\begin{aligned}\text{Cov}(Z, Y) &= \mathbb{E}[ZY] - \text{Pr}(Z = 1) * \mathbb{E}[Y] \\ &= p\mathbb{E}[Y|Z = 1] - p * [\mathbb{E}[Y|Z = 1] * p + [\mathbb{E}[Y|Z = 0] * (1 - p)]] \\ &= p\mathbb{E}[Y|Z = 1] - pp\mathbb{E}[Y|Z = 1] - p[\mathbb{E}[Y|Z = 0](1 - p)] \\ &= p(1 - p) * ([\mathbb{E}[Y|Z = 1]] - [\mathbb{E}[Y|Z = 0]]) \\ \text{Cov}(Z, X) &= p(1 - p) * ([\mathbb{E}[X|Z = 1]] - [\mathbb{E}[X|Z = 0]])\end{aligned}$$

- The ratio gives the **Wald estimator**:

$$\hat{\beta}^{IV} = \frac{\mathbb{E}[Y | Z = 1] - \mathbb{E}[Y | Z = 0]}{\mathbb{E}[X | Z = 1] - \mathbb{E}[X | Z = 0]}$$

A Simple IV Estimator: The Wald Estimator

- Consider the scalar model: $Y = X\beta + e$, let $Z \in \{0, 1\}$ be a binary instrument.
- The IV estimator is:

$$\hat{\beta}^{IV} = \frac{\text{Cov}(Z, Y)}{\text{Cov}(Z, X)}$$

$$\begin{aligned}\text{Cov}(Z, Y) &= \mathbb{E}[ZY] - \text{Pr}(Z = 1) * \mathbb{E}[Y] \\ &= p\mathbb{E}[Y|Z = 1] - p * [\mathbb{E}[Y|Z = 1] * p + [\mathbb{E}[Y|Z = 0] * (1 - p)]] \\ &= p\mathbb{E}[Y|Z = 1] - pp\mathbb{E}[Y|Z = 1] - p[\mathbb{E}[Y|Z = 0](1 - p)] \\ &= p(1 - p) * ([\mathbb{E}[Y|Z = 1]] - [\mathbb{E}[Y|Z = 0]]) \\ \text{Cov}(Z, X) &= p(1 - p) * ([\mathbb{E}[X|Z = 1]] - [\mathbb{E}[X|Z = 0]])\end{aligned}$$

- The ratio gives the **Wald estimator**:

$$\hat{\beta}^{IV} = \frac{\mathbb{E}[Y | Z = 1] - \mathbb{E}[Y | Z = 0]}{\mathbb{E}[X | Z = 1] - \mathbb{E}[X | Z = 0]}$$

A Simple IV Estimator: The Wald Estimator

- Consider the scalar model: $Y = X\beta + e$, let $Z \in \{0, 1\}$ be a binary instrument.
- The IV estimator is:

$$\hat{\beta}^{IV} = \frac{\text{Cov}(Z, Y)}{\text{Cov}(Z, X)}$$

$$\begin{aligned}\text{Cov}(Z, Y) &= \mathbb{E}[ZY] - \text{Pr}(Z = 1) * \mathbb{E}[Y] \\ &= p\mathbb{E}[Y|Z = 1] - p * [\mathbb{E}[Y|Z = 1] * p + [\mathbb{E}[Y|Z = 0] * (1 - p)]] \\ &= p\mathbb{E}[Y|Z = 1] - pp\mathbb{E}[Y|Z = 1] - p[\mathbb{E}[Y|Z = 0](1 - p)] \\ &= p(1 - p) * ([\mathbb{E}[Y|Z = 1]) - [\mathbb{E}[Y|Z = 0]] \\ \text{Cov}(Z, X) &= p(1 - p) * ([\mathbb{E}[X|Z = 1]) - [\mathbb{E}[X|Z = 0]]\end{aligned}$$

- The ratio gives the **Wald estimator**:

$$\hat{\beta}^{IV} = \frac{\mathbb{E}[Y | Z = 1] - \mathbb{E}[Y | Z = 0]}{\mathbb{E}[X | Z = 1] - \mathbb{E}[X | Z = 0]}$$

Finite-Sample Wald Estimator

- Suppose we observe n observations $\{(Y_i, X_i, Z_i)\}_{i=1}^n$ with $Z_i \in \{0, 1\}$

- Let:

$$\bar{Y}_1 = \frac{1}{n_1} \sum_{i: Z_i=1} Y_i$$

$$\bar{Y}_0 = \frac{1}{n_0} \sum_{i: Z_i=0} Y_i$$

$$\bar{X}_1 = \frac{1}{n_1} \sum_{i: Z_i=1} X_i$$

$$\bar{X}_0 = \frac{1}{n_0} \sum_{i: Z_i=0} X_i$$

where $n_1 = \sum_i Z_i$ and $n_0 = \sum_i (1 - Z_i)$

- Then the IV estimator is:

$$\hat{\beta}^{IV} = \frac{\bar{Y}_1 - \bar{Y}_0}{\bar{X}_1 - \bar{X}_0}$$

- This is the difference in average outcomes divided by the difference in average treatment intensity.

Finite-Sample Wald Estimator

- Suppose we observe n observations $\{(Y_i, X_i, Z_i)\}_{i=1}^n$ with $Z_i \in \{0, 1\}$
- Let:

$$\bar{Y}_1 = \frac{1}{n_1} \sum_{i:Z_i=1} Y_i$$

$$\bar{X}_1 = \frac{1}{n_1} \sum_{i:Z_i=1} X_i$$

$$\bar{Y}_0 = \frac{1}{n_0} \sum_{i:Z_i=0} Y_i$$

$$\bar{X}_0 = \frac{1}{n_0} \sum_{i:Z_i=0} X_i$$

where $n_1 = \sum_i Z_i$ and $n_0 = \sum_i (1 - Z_i)$

- Then the IV estimator is:

$$\hat{\beta}^{IV} = \frac{\bar{Y}_1 - \bar{Y}_0}{\bar{X}_1 - \bar{X}_0}$$

- This is the difference in average outcomes divided by the difference in average treatment intensity.

Standard Errors of the IV estimator (under homoskedasticity)

$$\begin{aligned}Avar(\hat{\beta}_{IV}) &= plim(\hat{\beta}_{IV} - \beta)(\hat{\beta}_{IV} - \beta)' \\&= plim(\mathbf{z}'\mathbf{x})^{-1}\mathbf{z}'ee'\mathbf{z}(\mathbf{x}'\mathbf{z})^{-1} \\&= \sigma_e^2 plim\left(\frac{\mathbf{z}'\mathbf{x}}{n}\right)^{-1} plim\left(\frac{\mathbf{z}'\mathbf{z}}{n}\right) plim\left(\frac{\mathbf{x}'\mathbf{z}}{n}\right)^{-1} \\&= \sigma_e^2 (plim\frac{\mathbf{z}'\mathbf{x}}{n})^{-1} (plim\frac{\mathbf{z}'\mathbf{z}}{n}) (plim\frac{\mathbf{x}'\mathbf{z}}{n})^{-1} \\&= \sigma_e^2 \Sigma_{zx}^{-1} \Sigma_z \Sigma_{xz}^{-1} \\ \widehat{Avar}(\hat{\beta}_{IV}) &= \sigma^2 (\mathbf{Z}'\mathbf{X})^{-1} \mathbf{Z}'\mathbf{Z}(\mathbf{X}'\mathbf{Z})^{-1} \\ \hat{\sigma}^2 &= n^{-1}(\mathbf{y} - \mathbf{X}\hat{\beta}_{iv})'(\mathbf{y} - \mathbf{X}\hat{\beta}_{iv})\end{aligned}$$

Standard Errors of the IV estimator (under homoskedasticity)

$$\begin{aligned}Avar(\hat{\beta}_{IV}) &= plim(\hat{\beta}_{IV} - \beta)(\hat{\beta}_{IV} - \beta)' \\&= plim(\mathbf{Z}'\mathbf{X})^{-1}\mathbf{Z}'ee'\mathbf{Z}(\mathbf{X}'\mathbf{Z})^{-1} \\&= \sigma_e^2 plim\left(\frac{\mathbf{Z}'\mathbf{X}}{n}\right)^{-1} plim\left(\frac{\mathbf{Z}'\mathbf{Z}}{n}\right) plim\left(\frac{\mathbf{X}'\mathbf{Z}}{n}\right)^{-1} \\&= \sigma_e^2 (plim\frac{\mathbf{Z}'\mathbf{X}}{n})^{-1} (plim\frac{\mathbf{Z}'\mathbf{Z}}{n}) (plim\frac{\mathbf{X}'\mathbf{Z}}{n})^{-1} \\&= \sigma_e^2 \Sigma_{ZX}^{-1} \Sigma_Z \Sigma_{XZ}^{-1} \\ \widehat{Avar}(\hat{\beta}_{IV}) &= \sigma^2 (\mathbf{Z}'\mathbf{X})^{-1} \mathbf{Z}'\mathbf{Z}(\mathbf{X}'\mathbf{Z})^{-1} \\ \hat{\sigma}^2 &= n^{-1}(\mathbf{y} - \mathbf{X}\hat{\beta}_{iv})'(\mathbf{y} - \mathbf{X}\hat{\beta}_{iv})\end{aligned}$$

Standard Errors of the IV estimator (under homoskedasticity)

$$\begin{aligned}Avar(\hat{\beta}_{IV}) &= plim(\hat{\beta}_{IV} - \beta)(\hat{\beta}_{IV} - \beta)' \\&= plim(\mathbf{z}'\mathbf{x})^{-1}\mathbf{z}'ee'\mathbf{z}(\mathbf{x}'\mathbf{z})^{-1} \\&= \sigma_e^2 plim\left(\frac{\mathbf{z}'\mathbf{x}}{n}\right)^{-1} plim\left(\frac{\mathbf{z}'\mathbf{z}}{n}\right) plim\left(\frac{\mathbf{x}'\mathbf{z}}{n}\right)^{-1} \\&= \sigma_e^2 (plim\frac{\mathbf{z}'\mathbf{x}}{n})^{-1} (plim\frac{\mathbf{z}'\mathbf{z}}{n}) (plim\frac{\mathbf{x}'\mathbf{z}}{n})^{-1} \\&= \sigma_e^2 \Sigma_{zx}^{-1} \Sigma_z \Sigma_{xz}^{-1} \\ \widehat{Avar}(\hat{\beta}_{IV}) &= \sigma^2 (\mathbf{Z}'\mathbf{X})^{-1} \mathbf{Z}'\mathbf{Z}(\mathbf{X}'\mathbf{Z})^{-1} \\ \hat{\sigma}^2 &= n^{-1}(\mathbf{y} - \mathbf{X}\hat{\beta}_{iv})'(\mathbf{y} - \mathbf{X}\hat{\beta}_{iv})\end{aligned}$$

Standard Errors of the IV estimator (under homoskedasticity)

$$\begin{aligned}Avar(\hat{\beta}_{IV}) &= plim(\hat{\beta}_{IV} - \beta)(\hat{\beta}_{IV} - \beta)' \\&= plim(\mathbf{z}'\mathbf{x})^{-1}\mathbf{z}'ee'\mathbf{z}(\mathbf{x}'\mathbf{z})^{-1} \\&= \sigma_e^2 plim\left(\frac{\mathbf{z}'\mathbf{x}}{n}\right)^{-1} plim\left(\frac{\mathbf{z}'\mathbf{z}}{n}\right) plim\left(\frac{\mathbf{x}'\mathbf{z}}{n}\right)^{-1} \\&= \sigma_e^2 (plim\frac{\mathbf{z}'\mathbf{x}}{n})^{-1} (plim\frac{\mathbf{z}'\mathbf{z}}{n}) (plim\frac{\mathbf{x}'\mathbf{z}}{n})^{-1} \\&= \sigma_e^2 \Sigma_{zx}^{-1} \Sigma_z \Sigma_{xz}^{-1}\end{aligned}$$

$$\begin{aligned}\widehat{Avar}(\hat{\beta}_{IV}) &= \sigma^2 (\mathbf{Z}'\mathbf{X})^{-1} \mathbf{Z}'\mathbf{Z} (\mathbf{X}'\mathbf{Z})^{-1} \\ \hat{\sigma}^2 &= n^{-1} (\mathbf{y} - \mathbf{X}\hat{\beta}_{iv})' (\mathbf{y} - \mathbf{X}\hat{\beta}_{iv})\end{aligned}$$

Standard Errors of the IV estimator (under homoskedasticity)

$$\begin{aligned}Avar(\hat{\beta}_{IV}) &= plim(\hat{\beta}_{IV} - \beta)(\hat{\beta}_{IV} - \beta)' \\&= plim(\mathbf{z}'\mathbf{x})^{-1}\mathbf{z}'ee'\mathbf{z}(\mathbf{x}'\mathbf{z})^{-1} \\&= \sigma_e^2 plim\left(\frac{\mathbf{z}'\mathbf{x}}{n}\right)^{-1} plim\left(\frac{\mathbf{z}'\mathbf{z}}{n}\right) plim\left(\frac{\mathbf{x}'\mathbf{z}}{n}\right)^{-1} \\&= \sigma_e^2 (plim\frac{\mathbf{z}'\mathbf{x}}{n})^{-1} (plim\frac{\mathbf{z}'\mathbf{z}}{n}) (plim\frac{\mathbf{x}'\mathbf{z}}{n})^{-1} \\&= \sigma_e^2 \Sigma_{zx}^{-1} \Sigma_z \Sigma_{xz}^{-1}\end{aligned}$$

$$\begin{aligned}\widehat{Avar}(\hat{\beta}_{IV}) &= \sigma^2 (\mathbf{Z}'\mathbf{X})^{-1} \mathbf{Z}'\mathbf{Z} (\mathbf{X}'\mathbf{Z})^{-1} \\ \hat{\sigma}^2 &= n^{-1} (\mathbf{y} - \mathbf{X}\hat{\beta}_{iv})' (\mathbf{y} - \mathbf{X}\hat{\beta}_{iv})\end{aligned}$$

Standard Errors of the IV estimator (under homoskedasticity)

$$\begin{aligned}Avar(\hat{\beta}_{IV}) &= plim(\hat{\beta}_{IV} - \beta)(\hat{\beta}_{IV} - \beta)' \\&= plim(\mathbf{z}'\mathbf{x})^{-1}\mathbf{z}'ee'\mathbf{z}(\mathbf{x}'\mathbf{z})^{-1} \\&= \sigma_e^2 plim\left(\frac{\mathbf{z}'\mathbf{x}}{n}\right)^{-1} plim\left(\frac{\mathbf{z}'\mathbf{z}}{n}\right) plim\left(\frac{\mathbf{x}'\mathbf{z}}{n}\right)^{-1} \\&= \sigma_e^2 \left(plim\frac{\mathbf{z}'\mathbf{x}}{n}\right)^{-1} \left(plim\frac{\mathbf{z}'\mathbf{z}}{n}\right) \left(plim\frac{\mathbf{x}'\mathbf{z}}{n}\right)^{-1} \\&= \sigma_e^2 \Sigma_{zx}^{-1} \Sigma_z \Sigma_{xz}^{-1} \\ \widehat{Avar}(\hat{\beta}_{IV}) &= \sigma^2 (\mathbf{Z}'\mathbf{X})^{-1} \mathbf{Z}'\mathbf{Z}(\mathbf{X}'\mathbf{Z})^{-1} \\ \hat{\sigma}^2 &= n^{-1}(\mathbf{y} - \mathbf{X}\hat{\beta}_{iv})'(\mathbf{y} - \mathbf{X}\hat{\beta}_{iv})\end{aligned}$$

Standard Errors of the IV estimator (under homoskedasticity)

$$\begin{aligned}Avar(\hat{\beta}_{IV}) &= plim(\hat{\beta}_{IV} - \beta)(\hat{\beta}_{IV} - \beta)' \\&= plim(\mathbf{z}'\mathbf{x})^{-1}\mathbf{z}'ee'\mathbf{z}(\mathbf{x}'\mathbf{z})^{-1} \\&= \sigma_e^2 plim\left(\frac{\mathbf{z}'\mathbf{x}}{n}\right)^{-1} plim\left(\frac{\mathbf{z}'\mathbf{z}}{n}\right) plim\left(\frac{\mathbf{x}'\mathbf{z}}{n}\right)^{-1} \\&= \sigma_e^2 \left(plim\frac{\mathbf{z}'\mathbf{x}}{n}\right)^{-1} \left(plim\frac{\mathbf{z}'\mathbf{z}}{n}\right) \left(plim\frac{\mathbf{x}'\mathbf{z}}{n}\right)^{-1} \\&= \sigma_e^2 \Sigma_{zx}^{-1} \Sigma_z \Sigma_{xz}^{-1} \\ \widehat{Avar}(\hat{\beta}_{IV}) &= \sigma^2 (\mathbf{Z}'\mathbf{X})^{-1} \mathbf{Z}'\mathbf{Z}(\mathbf{X}'\mathbf{Z})^{-1} \\ \hat{\sigma}^2 &= n^{-1}(\mathbf{y} - \mathbf{X}\hat{\beta}_{iv})'(\mathbf{y} - \mathbf{X}\hat{\beta}_{iv})\end{aligned}$$

Two-Stage Least Squares

- The logic of IV extends to cases where $l \geq k$.

$$Y_1 = \mathbf{z}'\bar{\Gamma}\beta + u_1$$

$$\mathbb{E}[\mathbf{z}u_1] = 0$$

$$\begin{aligned}\hat{\beta}_{2sls} &= (\hat{\Gamma}'\mathbf{z}'\mathbf{z}\hat{\Gamma})^{-1}(\hat{\Gamma}'\mathbf{z}\mathbf{y}_1) \\ &= (\mathbf{x}'\mathbf{z}(\mathbf{z}'\mathbf{z})^{-1}\mathbf{z}'\mathbf{z}(\mathbf{z}'\mathbf{z})^{-1}\mathbf{z}'\mathbf{x})^{-1}\mathbf{x}'\mathbf{z}(\mathbf{z}'\mathbf{z})^{-1}\mathbf{z}'\mathbf{y}_1) \\ &= (\mathbf{x}'\mathbf{z}(\mathbf{z}'\mathbf{z})^{-1}\mathbf{z}'\mathbf{x})^{-1}\mathbf{x}'\mathbf{z}(\mathbf{z}'\mathbf{z})^{-1}\mathbf{z}'\mathbf{y}_1) \\ &= (\mathbf{x}'\mathbf{P}_z'\mathbf{x})^{-1}(\mathbf{x}'\mathbf{P}_z\mathbf{y}_1) \\ &= (\mathbf{x}'\mathbf{P}_z\mathbf{P}_z'\mathbf{x})^{-1}\mathbf{x}'\mathbf{P}_z\mathbf{y}_1) \\ &= (\hat{\mathbf{x}}'\hat{\mathbf{x}})^{-1}(\hat{\mathbf{x}}'\mathbf{y}_1)\end{aligned}$$

- The regression of Y_1 on the fitted values $\hat{\mathbf{x}}$.

Two-Stage Least Squares

- The logic of IV extends to cases where $l \geq k$.

$$\begin{aligned}Y_1 &= \mathbf{z}'\bar{\Gamma}\beta + u_1 \\ \mathbb{E}[\mathbf{z}u_1] &= 0 \\ \hat{\beta}_{2sls} &= (\hat{\Gamma}'\mathbf{z}'\mathbf{z}\hat{\Gamma})^{-1}(\hat{\Gamma}'\mathbf{z}\mathbf{y}_1) \\ &= (\mathbf{x}'\mathbf{z}(\mathbf{z}'\mathbf{z})^{-1}\mathbf{z}'\mathbf{z}(\mathbf{z}'\mathbf{z})^{-1}\mathbf{z}'\mathbf{x})^{-1}\mathbf{x}'\mathbf{z}(\mathbf{z}'\mathbf{z})^{-1}\mathbf{z}'\mathbf{y}_1 \\ &= (\mathbf{x}'\mathbf{z}(\mathbf{z}'\mathbf{z})^{-1}\mathbf{z}'\mathbf{x})^{-1}\mathbf{x}'\mathbf{z}(\mathbf{z}'\mathbf{z})^{-1}\mathbf{z}'\mathbf{y}_1 \\ &= (\mathbf{x}'\mathbf{P}_z'\mathbf{x})^{-1}(\mathbf{x}'\mathbf{P}_z\mathbf{y}_1) \\ &= (\mathbf{x}'\mathbf{P}_z\mathbf{P}_z'\mathbf{x})^{-1}\mathbf{x}'\mathbf{P}_z\mathbf{y}_1 \\ &= (\hat{\mathbf{x}}'\hat{\mathbf{x}})^{-1}(\hat{\mathbf{x}}'\mathbf{y}_1)\end{aligned}$$

- The regression of Y_1 on the fitted values $\hat{\mathbf{x}}$.

Two-Stage Least Squares

- The logic of IV extends to cases where $l \geq k$.

$$\begin{aligned}Y_1 &= \mathbf{z}'\bar{\Gamma}\beta + u_1 \\ \mathbb{E}[\mathbf{z}u_1] &= 0 \\ \hat{\beta}_{2sls} &= (\hat{\Gamma}'\mathbf{z}'\mathbf{z}\hat{\Gamma})^{-1}(\hat{\Gamma}'\mathbf{z}\mathbf{y}_1) \\ &= (\mathbf{x}'\mathbf{z}(\mathbf{z}'\mathbf{z})^{-1}\mathbf{z}'\mathbf{z}(\mathbf{z}'\mathbf{z})^{-1}\mathbf{z}'\mathbf{x})^{-1}\mathbf{x}'\mathbf{z}(\mathbf{z}'\mathbf{z})^{-1}\mathbf{z}'\mathbf{y}_1 \\ &= (\mathbf{x}'\mathbf{z}(\mathbf{z}'\mathbf{z})^{-1}\mathbf{z}'\mathbf{x})^{-1}\mathbf{x}'\mathbf{z}(\mathbf{z}'\mathbf{z})^{-1}\mathbf{z}'\mathbf{y}_1 \\ &= (\mathbf{x}'\mathbf{P}_z'\mathbf{x})^{-1}(\mathbf{x}'\mathbf{P}_z\mathbf{y}_1) \\ &= (\mathbf{x}'\mathbf{P}_z\mathbf{P}_z'\mathbf{x})^{-1}\mathbf{x}'\mathbf{P}_z\mathbf{y}_1 \\ &= (\hat{\mathbf{x}}'\hat{\mathbf{x}})^{-1}(\hat{\mathbf{x}}'\mathbf{y}_1)\end{aligned}$$

- The regression of Y_1 on the fitted values $\hat{\mathbf{x}}$.

Two-Stage Least Squares

- The logic of IV extends to cases where $l \geq k$.

$$\begin{aligned}Y_1 &= \mathbf{z}'\bar{\Gamma}\beta + u_1 \\ \mathbb{E}[\mathbf{z}u_1] &= 0 \\ \hat{\beta}_{2sls} &= (\hat{\Gamma}'\mathbf{z}'\mathbf{z}\hat{\Gamma})^{-1}(\hat{\Gamma}'\mathbf{z}\mathbf{y}_1) \\ &= (\mathbf{x}'\mathbf{z}(\mathbf{z}'\mathbf{z})^{-1}\mathbf{z}'\mathbf{z}(\mathbf{z}'\mathbf{z})^{-1}\mathbf{z}'\mathbf{x})^{-1}\mathbf{x}'\mathbf{z}(\mathbf{z}'\mathbf{z})^{-1}\mathbf{z}'\mathbf{y}_1 \\ &= (\mathbf{x}'\mathbf{z}(\mathbf{z}'\mathbf{z})^{-1}\mathbf{z}'\mathbf{x})^{-1}\mathbf{x}'\mathbf{z}(\mathbf{z}'\mathbf{z})^{-1}\mathbf{z}'\mathbf{y}_1 \\ &= (\mathbf{x}'\mathbf{P}'_z\mathbf{x})^{-1}(\mathbf{x}'\mathbf{P}_z\mathbf{y}_1) \\ &= (\mathbf{x}'\mathbf{P}_z\mathbf{P}'_z\mathbf{x})^{-1}\mathbf{x}'\mathbf{P}_z\mathbf{y}_1 \\ &= (\hat{\mathbf{x}}'\hat{\mathbf{x}})^{-1}(\hat{\mathbf{x}}'\mathbf{y}_1)\end{aligned}$$

- The regression of Y_1 on the fitted values $\hat{\mathbf{x}}$.

Two-Stage Least Squares

- The logic of IV extends to cases where $l \geq k$.

$$\begin{aligned}Y_1 &= \mathbf{z}'\bar{\Gamma}\beta + u_1 \\ \mathbb{E}[\mathbf{z}u_1] &= 0 \\ \hat{\beta}_{2sls} &= (\hat{\Gamma}'\mathbf{z}'\mathbf{z}\hat{\Gamma})^{-1}(\hat{\Gamma}'\mathbf{z}\mathbf{y}_1) \\ &= (\mathbf{x}'\mathbf{z}(\mathbf{z}'\mathbf{z})^{-1}\mathbf{z}'\mathbf{z}(\mathbf{z}'\mathbf{z})^{-1}\mathbf{z}'\mathbf{x})^{-1}\mathbf{x}'\mathbf{z}(\mathbf{z}'\mathbf{z})^{-1}\mathbf{z}'\mathbf{y}_1 \\ &= (\mathbf{x}'\mathbf{z}(\mathbf{z}'\mathbf{z})^{-1}\mathbf{z}'\mathbf{x})^{-1}\mathbf{x}'\mathbf{z}(\mathbf{z}'\mathbf{z})^{-1}\mathbf{z}'\mathbf{y}_1 \\ &= (\mathbf{x}'\mathbf{P}_z'\mathbf{x})^{-1}(\mathbf{x}'\mathbf{P}_z\mathbf{y}_1) \\ &= (\mathbf{x}'\mathbf{P}_z\mathbf{P}_z'\mathbf{x})^{-1}\mathbf{x}'\mathbf{P}_z\mathbf{y}_1 \\ &= (\hat{\mathbf{x}}'\hat{\mathbf{x}})^{-1}(\hat{\mathbf{x}}'\mathbf{y}_1)\end{aligned}$$

- The regression of Y_1 on the fitted values $\hat{\mathbf{x}}$.

Two-Stage Least Squares

- The logic of IV extends to cases where $l \geq k$.

$$\begin{aligned}Y_1 &= \mathbf{z}'\bar{\Gamma}\beta + u_1 \\ \mathbb{E}[\mathbf{z}u_1] &= 0 \\ \hat{\beta}_{2sls} &= (\hat{\Gamma}'\mathbf{z}'\mathbf{z}\hat{\Gamma})^{-1}(\hat{\Gamma}'\mathbf{z}\mathbf{y}_1) \\ &= (\mathbf{x}'\mathbf{z}(\mathbf{z}'\mathbf{z})^{-1}\mathbf{z}'\mathbf{z}(\mathbf{z}'\mathbf{z})^{-1}\mathbf{z}'\mathbf{x})^{-1}\mathbf{x}'\mathbf{z}(\mathbf{z}'\mathbf{z})^{-1}\mathbf{z}'\mathbf{y}_1 \\ &= (\mathbf{x}'\mathbf{z}(\mathbf{z}'\mathbf{z})^{-1}\mathbf{z}'\mathbf{x})^{-1}\mathbf{x}'\mathbf{z}(\mathbf{z}'\mathbf{z})^{-1}\mathbf{z}'\mathbf{y}_1 \\ &= (\mathbf{x}'\mathbf{P}'_z\mathbf{x})^{-1}(\mathbf{x}'\mathbf{P}_z\mathbf{y}_1) \\ &= (\mathbf{x}'\mathbf{P}_z\mathbf{P}'_z\mathbf{x})^{-1}\mathbf{x}'\mathbf{P}_z\mathbf{y}_1 \\ &= (\hat{\mathbf{x}}'\hat{\mathbf{x}})^{-1}(\hat{\mathbf{x}}'\mathbf{y}_1)\end{aligned}$$

- The regression of Y_1 on the fitted values $\hat{\mathbf{x}}$.

Standard Errors of the 2SLS estimator

$$\begin{aligned}y &= \mathbf{X}\beta + e \\&= \hat{\mathbf{X}} + \hat{u})\beta + e \\&= \hat{\mathbf{X}}\beta + (e + \hat{u}\beta) \\&= \hat{\mathbf{X}}\beta + \varepsilon\end{aligned}$$

$$\begin{aligned}\widehat{Avar}(\hat{\beta}_{2SLS}) &= \sigma_{\varepsilon}^2(\hat{\mathbf{X}}'\hat{\mathbf{X}})^{-1} \\&= \sigma_{\varepsilon}^2(\mathbf{Z}'\mathbf{X})^{-1}\mathbf{Z}'\mathbf{Z}(\mathbf{X}'\mathbf{Z})^{-1} \\&= \underbrace{\sigma_{\varepsilon}^2(\mathbf{Z}'\mathbf{X})^{-1}\mathbf{Z}'\mathbf{Z}(\mathbf{X}'\mathbf{Z})^{-1}}_{Avar(\hat{\beta}_{IV})} + var(\hat{u}\beta)(\mathbf{Z}'\mathbf{X})^{-1}\mathbf{Z}'\mathbf{Z}(\mathbf{X}'\mathbf{Z})^{-1}\end{aligned}$$

Standard Errors of the 2SLS estimator

$$\begin{aligned}y &= \mathbf{X}\beta + e \\&= \hat{\mathbf{X}} + \hat{u})\beta + e \\&= \hat{\mathbf{X}}\beta + (e + \hat{u}\beta) \\&= \hat{\mathbf{X}}\beta + \varepsilon\end{aligned}$$

$$\begin{aligned}\widehat{Avar}(\hat{\beta}_{2SLS}) &= \sigma_{\varepsilon}^2(\hat{\mathbf{X}}'\hat{\mathbf{X}})^{-1} \\&= \sigma_{\varepsilon}^2(\mathbf{Z}'\mathbf{X})^{-1}\mathbf{Z}'\mathbf{Z}(\mathbf{X}'\mathbf{Z})^{-1} \\&= \underbrace{\sigma_{\varepsilon}^2(\mathbf{Z}'\mathbf{X})^{-1}\mathbf{Z}'\mathbf{Z}(\mathbf{X}'\mathbf{Z})^{-1}}_{Avar(\hat{\beta}_{IV})} + var(\hat{u}\beta)(\mathbf{Z}'\mathbf{X})^{-1}\mathbf{Z}'\mathbf{Z}(\mathbf{X}'\mathbf{Z})^{-1}\end{aligned}$$

Standard Errors of the 2SLS estimator

$$\begin{aligned}y &= \mathbf{X}\beta + e \\&= \hat{\mathbf{X}} + \hat{u})\beta + e \\&= \hat{\mathbf{X}}\beta + (e + \hat{u}\beta) \\&= \hat{\mathbf{X}}\beta + \varepsilon\end{aligned}$$

$$\begin{aligned}\widehat{Avar}(\hat{\beta}_{2SLS}) &= \sigma_{\varepsilon}^2(\hat{\mathbf{X}}'\hat{\mathbf{X}})^{-1} \\&= \sigma_{\varepsilon}^2(\mathbf{Z}'\mathbf{X})^{-1}\mathbf{Z}'\mathbf{Z}(\mathbf{X}'\mathbf{Z})^{-1} \\&= \underbrace{\sigma_{\varepsilon}^2(\mathbf{Z}'\mathbf{X})^{-1}\mathbf{Z}'\mathbf{Z}(\mathbf{X}'\mathbf{Z})^{-1}}_{Avar(\hat{\beta}_{IV})} + var(\hat{u}\beta)(\mathbf{Z}'\mathbf{X})^{-1}\mathbf{Z}'\mathbf{Z}(\mathbf{X}'\mathbf{Z})^{-1}\end{aligned}$$

Standard Errors of the 2SLS estimator

$$\begin{aligned}y &= \mathbf{X}\beta + e \\&= \hat{\mathbf{X}} + \hat{u})\beta + e \\&= \hat{\mathbf{X}}\beta + (e + \hat{u}\beta) \\&= \hat{\mathbf{X}}\beta + \varepsilon\end{aligned}$$

$$\begin{aligned}\widehat{Avar}(\hat{\beta}_{2SLS}) &= \sigma_{\varepsilon}^2(\hat{\mathbf{X}}'\hat{\mathbf{X}})^{-1} \\&= \sigma_{\varepsilon}^2(\mathbf{Z}'\mathbf{X})^{-1}\mathbf{Z}'\mathbf{Z}(\mathbf{X}'\mathbf{Z})^{-1} \\&= \underbrace{\sigma_{\varepsilon}^2(\mathbf{Z}'\mathbf{X})^{-1}\mathbf{Z}'\mathbf{Z}(\mathbf{X}'\mathbf{Z})^{-1}}_{Avar(\hat{\beta}_{IV})} + var(\hat{u}\beta)(\mathbf{Z}'\mathbf{X})^{-1}\mathbf{Z}'\mathbf{Z}(\mathbf{X}'\mathbf{Z})^{-1}\end{aligned}$$

Standard Errors of the 2SLS estimator

$$\begin{aligned}y &= \mathbf{X}\beta + e \\&= \hat{\mathbf{X}} + \hat{u})\beta + e \\&= \hat{\mathbf{X}}\beta + (e + \hat{u}\beta) \\&= \hat{\mathbf{X}}\beta + \varepsilon\end{aligned}$$

$$\begin{aligned}\widehat{Avar}(\hat{\beta}_{2SLS}) &= \sigma_{\varepsilon}^2(\hat{\mathbf{X}}'\hat{\mathbf{X}})^{-1} \\&= \sigma_{\varepsilon}^2(\mathbf{Z}'\mathbf{X})^{-1}\mathbf{Z}'\mathbf{Z}(\mathbf{X}'\mathbf{Z})^{-1} \\&= \underbrace{\sigma_{\varepsilon}^2(\mathbf{Z}'\mathbf{X})^{-1}\mathbf{Z}'\mathbf{Z}(\mathbf{X}'\mathbf{Z})^{-1}}_{Avar(\hat{\beta}_{IV})} + var(\hat{u}\beta)(\mathbf{Z}'\mathbf{X})^{-1}\mathbf{Z}'\mathbf{Z}(\mathbf{X}'\mathbf{Z})^{-1}\end{aligned}$$

Standard Errors of the 2SLS estimator

$$\begin{aligned}y &= \mathbf{X}\beta + e \\&= \hat{\mathbf{X}} + \hat{u})\beta + e \\&= \hat{\mathbf{X}}\beta + (e + \hat{u}\beta) \\&= \hat{\mathbf{X}}\beta + \varepsilon\end{aligned}$$

$$\begin{aligned}\widehat{Avar}(\hat{\beta}_{2SLS}) &= \sigma_{\varepsilon}^2(\hat{\mathbf{X}}'\hat{\mathbf{X}})^{-1} \\&= \sigma_{\varepsilon}^2(\mathbf{Z}'\mathbf{X})^{-1}\mathbf{Z}'\mathbf{Z}(\mathbf{X}'\mathbf{Z})^{-1} \\&= \underbrace{\sigma_{\varepsilon}^2(\mathbf{Z}'\mathbf{X})^{-1}\mathbf{Z}'\mathbf{Z}(\mathbf{X}'\mathbf{Z})^{-1}}_{Avar(\hat{\beta}_{IV})} + var(\hat{u}\beta)(\mathbf{Z}'\mathbf{X})^{-1}\mathbf{Z}'\mathbf{Z}(\mathbf{X}'\mathbf{Z})^{-1}\end{aligned}$$

2SLS in Just-Identified Case: No Finite Moments

- In the just-identified IV model with one instrument Z and one endogenous regressor X ,

$$\hat{\beta}^{IV} = \frac{\sum_i Z_i Y_i}{\sum_i Z_i X_i}$$

- Even under standard assumptions (e.g., i.i.d. errors, homoskedasticity, valid instruments), the denominator $\sum_i Z_i X_i$ is random.
- The distribution of $\hat{\beta}^{IV}$ has heavy (Cauchy-like) tails.
- This implies:

$$\mathbb{E}[(\hat{\beta}^{IV})^k] = \infty \quad \text{for all integers } k \geq 1$$

- Applies even with strong instruments.

Bootstrap Fails Under Weak Identification

- Standard bootstrap relies on the estimator having a smooth, well-behaved limiting distribution.
- In just-identified IV, the heavy-tailed distribution of $\hat{\beta}^{IV}$ violates these conditions.
- Bootstrap t-statistics and confidence intervals can be severely misleading.
- Instead, use the empirical quantiles of the bootstrapped replications.