Econometrics II

Lecture 2 - Instrumental Variables (Part I)

Lasha Chochua

2026

ISET

Introduction

- **Endogeneity** and **instrumental variable** concepts are central to econometrics
  - Represent a major shift from traditional statistics
  - Arise from models of **simultaneous equations**
  - Classic example: supply/demand model of price determination
- **Historical context**:
  - Identification problem in simultaneous equations:
    - First studied by *Philip Wright* (1915) and *Working* (1927)
  - Method of instrumental variables introduced in a 1928 book by *Philip Wright*
    - Sometimes credited to his son *Sewell Wright*
  - Term **"instrumental variables"** coined by *Reiersøl* (1945)

Overview: Endogeneity in Linear Models

- We say there is **endogeneity** in the linear model:

$$Y = X'\beta + e \tag{1}$$

where β is the **parameter of interest**, if:

$$\mathbb{E}[Xe] \neq 0 \tag{2}$$

- Equation (1) is a **structural equation**, and β a **structural parameter**
- When (2) holds, X is said to be **endogenous** for β

Note: A structural equation represents a theoretical relationship between variables based on economic theory and causal mechanisms.

Overview: Projection vs Structural Parameters

- Endogeneity cannot arise if the coefficient is from a **linear projection**:

$$\beta^* = \mathbb{E}[XX']^{-1}\mathbb{E}[XY]$$

  - Projection equation:

$$Y = X'\beta^* + e^*, \quad \mathbb{E}[Xe^*] = 0$$

- But the problem is that under endogeneity (2), $\beta^* \neq \beta$:

$$\begin{aligned}\beta^* &= \left(\mathbb{E}[XX']\right)^{-1}\mathbb{E}[XY] \\ &= \left(\mathbb{E}[XX']\right)^{-1}\mathbb{E}[X(X'\beta + e)] \\ &= \beta + \left(\mathbb{E}[XX']\right)^{-1}\mathbb{E}[Xe] \neq \beta\end{aligned}$$

Overview: Consequences of Endogeneity

- **Structural** $\neq$ **projection** $\rightarrow$ different estimation goals
- Endogeneity implies **OLS is inconsistent** for β
- But OLS is consistent for projection coefficient β^*:

$$\hat{\beta} \xrightarrow{p} \left(\mathbb{E}[XX']\right)^{-1} \mathbb{E}[XY] = \beta^* \neq \beta$$

- This is often called **endogeneity bias** or **estimation bias**
  - Actually, the issue is **inconsistency**, not classical bias
- Since β is the parameter of interest, we need **alternative estimation methods**

Example 1: Measurement Error in the Regressor

- Suppose (Y, Z) are joint random variables and:
  - $\mathbb{E}[Y \mid Z] = Z'\beta$
  - β is the **structural parameter**
- We **do not** observe Z, but instead observe:

$$X = Z + u$$

  - u is a $k \times 1$ measurement error
  - u is independent of both Z and the structural error e
  - $\mathbb{E}[u] = 0 \rightarrow$ **classical measurement error**. X is noisy but unbaised measure of Z.
- Substituting into the structural equation:

$$Y = Z'\beta + e = (X - u)'\beta + e = X'\beta + \nu$$

where $\nu = e - u'\beta$

Measurement Error in the Regressor

- So we observe the equation:
$$Y = X'\beta + \nu$$
  - But: ν is **not** a projection error
  - Indeed,
$$\mathbb{E}[X\nu] = \mathbb{E}[(Z + u)(e - u'\beta)] = -\mathbb{E}[uu']\beta \neq 0$$
if β and $\mathbb{E}[uu']$ do not equal zero.

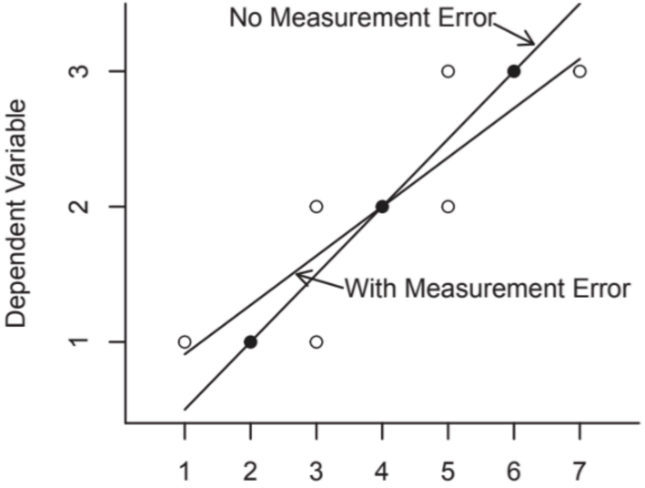
Measurement Error $\rightarrow$ Attenuation Bias

- If $\mathbb{E}[X\nu] \neq 0$, then **OLS is inconsistent** for β
- Let $k = 1$ (scalar case)
  Then the projection coefficient is:

$$\beta^* = \beta + \frac{\mathbb{E}[X\nu]}{\mathbb{E}[X^2]} = \beta \left(1 - \frac{\mathbb{E}[u^2]}{\mathbb{E}[X^2]}\right)$$

- So $\beta^* < \beta$ in magnitude $\rightarrow$ **attenuation bias**

Measurement Error Illustration



Example 2: Supply and Demand

- Q and P (quantity and price) are jointly determined by:
  - **Demand equation**:

$$Q = -\beta_1 P + e_1$$

  - **Supply equation**:

$$Q = \beta_2 P + e_2$$

- Let $e = (e_1, e_2)$, with:
  - $\mathbb{E}[e] = 0$, $\mathbb{E}[ee'] = I_2$

Supply and Demand

- Solve for (Q, P) using matrix form:

$$\begin{bmatrix} 1 & \beta_1 \\ 1 & -\beta_2 \end{bmatrix} \begin{pmatrix} Q \\ P \end{pmatrix} = \begin{pmatrix} e_1 \\ e_2 \end{pmatrix}$$

- Invert the system:

$$\begin{aligned} \begin{pmatrix} Q \\ P \end{pmatrix} &= \begin{bmatrix} \beta_2 & \beta_1 \\ 1 & -1 \end{bmatrix} \begin{pmatrix} e_1 \\ e_2 \end{pmatrix} \cdot \frac{1}{\beta_1 + \beta_2} \\ &= \begin{pmatrix} \frac{\beta_2 e_1 + \beta_1 e_2}{\beta_1 + \beta_2} \\ \frac{e_1 - e_2}{\beta_1 + \beta_2} \end{pmatrix} \end{aligned}$$

Simultaneous Equations Bias

- Projecting Q on P gives:
$$Q = \beta^* P + e^*, \quad \mathbb{E}[Pe^*] = 0$$
- Projection coefficient:
$$\beta^* = \frac{\mathbb{E}[PQ]}{\mathbb{E}[P^2]} = \frac{\beta_2 - \beta_1}{2}$$
- So $\beta^* \neq \beta_1$, $\beta^* \neq \beta_2$
- This is called **simultaneous equations bias**
  - Arises when Y and X are jointly determined (as in equilibrium models)
- **Key idea**: OLS cannot recover either structural slope
- This is a classic case of **endogeneity due to simultaneity**

Simultaneous Equations Bias (Cont.)

- From the reduced form:

$$Q = \frac{\beta_2 e_1 + \beta_1 e_2}{\beta_1 + \beta_2}, \qquad P = \frac{e_1 - e_2}{\beta_1 + \beta_2}$$

- Using $\mathbb{E}[e_1^2] = \mathbb{E}[e_2^2] = 1$ and $\mathbb{E}[e_1 e_2] = 0$:

$$\mathbb{E}[P^2] = \frac{\mathbb{E}[(e_1 - e_2)^2]}{(\beta_1 + \beta_2)^2} = \frac{2}{(\beta_1 + \beta_2)^2}$$

$$\mathbb{E}[PQ] = \frac{\mathbb{E}[(e_1 - e_2)(\beta_2 e_1 + \beta_1 e_2)]}{(\beta_1 + \beta_2)^2} = \frac{\beta_2 - \beta_1}{(\beta_1 + \beta_2)^2}$$

- Therefore:

$$\beta^* = \frac{\mathbb{E}[PQ]}{\mathbb{E}[P^2]} = \frac{\beta_2 - \beta_1}{(\beta_1 + \beta_2)^2} \cdot \frac{(\beta_1 + \beta_2)^2}{2} = \frac{\beta_2 - \beta_1}{2}$$

Example 3: Choice Variables as Regressors

- Consider the wage equation:

$$\log(wage) = \beta \cdot education + e$$

- β: average **causal** effect of education on wages
- **Endogeneity source**:
  - Individuals with high unobserved ability **self-select** into more education
  - So e (unobserved ability) is positively correlated with education
  - $\Rightarrow$ **education is endogenous**
- Consequence:
  - Projection coefficient β^* is **upward biased** relative to structural β
  - OLS will **overestimate** the causal effect

Choice Variables: Endogeneity Insight

- This kind of endogeneity arises when both Y and X are **choices by the same agent**
  - Even if made at different times
- General rule:
  - When both dependent and regressor variables are **choice variables**, they should be treated as **endogenous**

Endogenous Regressors

- **Endogeneity**: a regressor X is correlated with the error term e
- **Exogeneity**: X is **exogenous** for β if $\mathbb{E}[Xe] = 0$
- In economics:
  - X is **endogenous** if it is **jointly** determined with Y
  - X is **exogenous** if it is determined separately from Y
- Partition regressors: $X = (X_1, X_2)$
  - X_1: **exogenous regressors** (k_1-dimensional)
  - X_2: **endogenous regressors** (k_2-dimensional)
- Sometimes we call X_2 the **endogenous right-hand-side variable**

Structural Equation with Endogenous Regressors

- Partition parameters: $\beta = (\beta_1, \beta_2)$
- Then the atructural model becomes:

$$Y = X_1'\beta_1 + X_2'\beta_2 + e \tag{3}$$

- Alternative notation:
  - Let $Y_2 = X_2$ and rename Y as Y_1
  - Then the structural equation becomes:

$$Y_1 = X_1'\beta_1 + Y_2'\beta_2 + e \tag{4}$$

Structural Equation with Endogenous Regressors (Cont.)

- Use $\vec{Y} = (Y_1, Y_2')'$ to represent the set of endogeneous variables.
- The assumptions regarding the regressors and regression error are:

$$\mathbb{E}[X_1 e] = 0, \quad \mathbb{E}[Y_2 e] \neq 0$$

- Y_2 captures endogenous variables such as:
    - Simultaneous regressors
    - Choice variables
    - Mis-measured regressors
- Often k_2 (number of endogenous regressors) is small (1 or 2)

Structural Equation with Endogenous Regressors (Cont.)

- Use $\vec{Y} = (Y_1, Y_2')'$ to represent the set of endogeneous variables.
- The assumptions regarding the regressors and regression error are:

$$\mathbb{E}[X_1 e] = 0, \quad \mathbb{E}[Y_2 e] \neq 0$$

- Y_2 captures endogenous variables such as:
  - Simultaneous regressors
  - Choice variables
  - Mis-measured regressors
- Often k_2 (number of endogenous regressors) is small (1 or 2)
- **Note:** e is a scalar.

Instruments

- To consistently estimate β, we need additional information
- One common source: **instruments** (instrumental variables)

Definition 1 (Instrumental Variable Conditions)
The $\ell \times 1$ random vector Z is an *instrumental variable* for the structural equation

$$Y = X_1'\beta_1 + X_2'\beta_2 + e$$

if:

$$\mathbb{E}[Ze] = 0 \qquad (5)$$

$$\mathbb{E}[ZZ'] > 0 \qquad (6)$$

$$\text{rank}(\mathbb{E}[ZX']) = k \qquad (7)$$

These are the *exogeneity*, *non-degeneracy*, and *relevance* conditions.

Instrumental Variable Conditions: Intuition

- **Exogeneity** $\mathbb{E}[Ze] = 0$: the instrument is **uncorrelated with the error**
  - Z must not be a direct cause of Y (only through X)
  - Violation: using ability as an instrument for education in a wage equation
  - Ability affects wages *directly* — not a valid instrument
- **Non-degeneracy** $\mathbb{E}[ZZ'] > 0$: the instrument has **variation**
  - Z cannot be constant or collinear
  - Technically ensures the moment matrix is invertible
  - Rarely an issue in practice – just rules out degenerate cases
- **Relevance** rank$(\mathbb{E}[ZX']) = k$: Z provides **independent identifying variation** for each regressor
  - In the simple case $(k = \ell = 1)$: reduces to $\mathbb{E}[ZX] \neq 0$ – nonzero correlation
  - In general: instruments must collectively span enough directions to identify all k parameters
  - Violation: **weak instruments** – Z barely correlated with X_2, causing severe finite-sample bias

Instrumental Variable Conditions: Intuition

- **Exogeneity** $\mathbb{E}[Ze] = 0$: the instrument is **uncorrelated with the error**
  - Z must not be a direct cause of Y (only through X)
  - Violation: using ability as an instrument for education in a wage equation
  - Ability affects wages *directly* — not a valid instrument
- **Non-degeneracy** $\mathbb{E}[ZZ'] > 0$: the instrument has **variation**
  - Z cannot be constant or collinear
  - Technically ensures the moment matrix is invertible
  - Rarely an issue in practice – just rules out degenerate cases
- **Relevance** rank$(\mathbb{E}[ZX']) = k$: Z provides **independent identifying variation** for each regressor
  - In the simple case $(k = \ell = 1)$: reduces to $\mathbb{E}[ZX] \neq 0$ – nonzero correlation
  - In general: instruments must collectively span enough directions to identify all k parameters
  - Violation: **weak instruments** – Z barely correlated with X_2, causing severe finite-sample bias

Key insight: Exogeneity and relevance pull in opposite directions – a stronger correlation with X risks correlation with e. Finding valid instruments is hard.

Instruments and Regressors

- The **exogenous regressors** X_1 satisfy $\mathbb{E}[X_1 e] = 0$ (condition 5)
  - So X_1 can be included as instruments
- Thus $X_1 \subset Z$

$$Z = \begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix} = \begin{pmatrix} X_1 \\ Z_2 \end{pmatrix} \quad \begin{matrix} k_1 \times 1 \\ \ell_2 \times 1 \end{matrix} \tag{8}$$

- **Note:** $\ell_2 = \ell - k_1$

Instrument Sets and Identification

- Let $X_1 = Z_1$ (included exogenous variables)
  Let $Z_2 =$ excluded exogenous variables (uncorrelated with e, but not in the structural equation)
- Then the structural equation becomes:

$$Y_1 = Z_1'\beta_1 + Y_2'\beta_2 + e \tag{9}$$

- This clarifies roles:
  - Z_1: **exogenous regressors**
  - Y_2: **endogenous regressors**
  - Z_2: **instrumental variables**
- Identification:
  - **Just-identified** if $\ell = k \quad (= k_1 + k_2)$
  - **Over-identified** if $\ell > k$

Instrumental Variables

- To be a valid instrument Z_2, a variable must:
  - Satisfy $\mathbb{E}[Z_2 e] = 0$ (exogeneity)
  - Satisfy $\mathbb{E}[Z_2 X_2'] \neq 0$ (relevance – correlated with the endogenous regressor)

Examples of Valid Instruments

- Let's revisit the **three core examples** and examine what qualifies as a valid instrument Z_2 in each.

1. Measurement Error in the Regressor

- Suppose X is a **mis-measured version of** Z
- A valid instrument Z_2 is an **alternative measurement** of Z
- For Z_2 to be valid:
  - Its **measurement error must be independent** of that in X
  - Z_2 must be correlated with Z, but uncorrelated with e

Examples of Valid Instruments (cont.)

2. Supply and Demand

- In a **demand equation**:
  - Valid Z_2: a variable that affects **supply but not demand**
  - It alters the equilibrium (P, Q), but does **not directly affect** P **except through** Q
  - Example: a variable related to **production costs**
- In a **supply equation**:
  - Valid Z_2: a variable that affects **demand but not supply**
  - It influences the equilibrium (P, Q), but affects **price only via quantity**

Examples of Valid Instruments (cont.)

3. Choice Variable as Regressor

- Example: X is a **choice variable** (e.g. education)
- Valid Z_2: affects the **choice** of the regressor (e.g. education)
  - But does **not directly affect the outcome** (e.g. wages)
  - Only affects the outcome **indirectly through** X
- Key condition: Z_2 is **excluded** from the structural equation, but still predicts X
- This ensures both **relevance** and **exogeneity**

Example: College Proximity (Card 1995)

- **Idea**: Living near a college lowers cost $\Rightarrow$ more likely to attend college
  - But **does not directly affect** market skills or wage
- So, **college proximity** can be a valid **instrument** for **education** in a wage regression
  - Satisfies exogeneity: affects $education$ but not $wage$ directly
  - Satisfies relevance: increases likelihood of attending college
- Used throughout the lecture to illustrate IV concepts

Card's Wage Regression: NLSYM Data

- Data source: **National Longitudinal Survey of Young Men (1976)**
- Sample: 3,010 observations (after dropping rows with missing $wage$)
- **Dependent variable**:
  - $\log$(weekly earnings)
- **Regressors**:
  - $education$: years of schooling
  - $experience$: $age - (education + 6)$
  - $experience^2/100$
  - $Black$: race indicator
  - $south$: indicator for living in the South
  - $urban$: metro area indicator
- **You will replicate this paper.**

College Proximity as an Instrument

- **Education is endogenous**: it's a **choice variable**
  - Affected by unobserved ability $\rightarrow$ correlated with e
  - OLS is biased for causal effect of education on wages
- Labor economics predicts:
  - Ability, education, and wages are **positively correlated**
  - $\Rightarrow$ OLS likely **upward biased**
  - But sign of bias is uncertain due to other regressors/sources

College Proximity Measures (Card 1995)

- Card proposes:
  Instrument = grew up near a 4-year college
- Three dummy variables used as instruments:
  - *college*: grew up in same county as **any** 4-year college
  - *public*: grew up in same county as a **public** 4-year college
  - *private*: grew up in same county as a **private** 4-year college
- These affect **likelihood of attending college** but not wages directly $\Rightarrow$ satisfy IV relevance and exclusion restrictions

Card (1995) - Results

	OLS	IV(a)	IV(b)	2SLS(a)	2SLS(b)	LIML
education	0.074	0.132	0.133	0.161	0.160	0.164
	(0.004)	(0.049)	(0.051)	(0.040)	(0.041)	(0.042)
experience	0.084	0.107	0.056	0.119	0.047	0.120
	(0.007)	(0.021)	(0.026)	(0.018)	(0.025)	(0.019)
experience^2/100	−0.224	−0.228	−0.080	−0.231	−0.032	−0.231
	(0.032)	(0.035)	(0.133)	(0.037)	(0.127)	(0.037)
Black	−0.190	−0.131	−0.103	−0.102	−0.064	−0.099
	(0.017)	(0.051)	(0.075)	(0.044)	(0.061)	(0.045)
south	−0.125	−0.105	−0.098	−0.095	−0.086	−0.094
	(0.015)	(0.023)	(0.0284)	(0.022)	(0.026)	(0.022)
urban	0.161	0.131	0.108	0.116	0.083	0.115
	(0.015)	(0.030)	(0.049)	(0.026)	(0.041)	(0.027)
Sargan				0.82	0.52	0.82
p-value				0.37	0.47	0.37

1. **IV(a)**:
   - Instrument: *college*
   - For: $education$
2. **IV(b)**:
   - Instruments: *college*, age, $age^2/100$
   - For: $education$, $experience$, $experience^2/100$
3. **2SLS(a)**:
   - Instruments: *public*, *private*
   - For: $education$
4. **2SLS(b)**:
   - Instruments: *public*, *private*, age, age^2
   - For: $education$, $experience$, $experience^2/100$
5. **LIML**:
   - Instruments: *public*, *private*
   - For: $education$

Reduced Form Setup

- The reduced form expresses endogenous regressors Y_2 as functions of instruments Z
- Recall:
  - Y_2: $k_2 \times 1$ vector of endogenous regressors
  - Z: $\ell \times 1$ vector of instruments
- The reduced form is:

$$Y_2 = \Gamma' Z + u_2 \tag{10}$$

- Γ: $\ell \times k_2$ matrix of projection coefficients
- u_2: $k_2 \times 1$ projection error, satisfies $\mathbb{E}[Zu_2'] = 0$
- Γ is given by:

$$\Gamma = \mathbb{E}[ZZ']^{-1}\mathbb{E}[ZY_2'] \tag{11}$$

Reduced Form with Partitioned Instruments

- Partition instruments and projection matrix:
  - $Z = \begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix}$ with
    - Z_1: $k_1 \times 1$ (included exogenous variables)
    - Z_2: $\ell_2 \times 1$ (excluded instruments)
  - $\Gamma' = \begin{pmatrix} \Gamma'_{12} & \Gamma'_{22} \end{pmatrix}$
- Then (10) becomes:

$$Y_2 = \Gamma'_{12} Z_1 + \Gamma'_{22} Z_2 + u_2$$

Reduced Form with Partitioned Instruments

- Partition instruments and projection matrix:
  - $Z = \begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix}$ with
    - Z_1: $k_1 \times 1$ (included exogenous variables)
    - Z_2: $\ell_2 \times 1$ (excluded instruments)
  - $\Gamma' = \begin{pmatrix} \Gamma'_{12} & \Gamma'_{22} \end{pmatrix}$
- Then (10) becomes:

$$Y_2 = \Gamma'_{12} Z_1 + \Gamma'_{22} Z_2 + u_2$$

- $\Gamma' = \begin{pmatrix} \Gamma'_{12} & \Gamma'_{22} \end{pmatrix}$
  - Γ_{12}: $k_1 \times k_2$ (coefficients on Z_1)
  - Γ_{22}: $\ell_2 \times k_2$ (coefficients on Z_2)

Reduced Form with Partitioned Instruments (Cont.)

$$Y_2 = \Gamma_{12}' Z_1 + \Gamma_{22}' Z_2 + u_2$$

$$Y_2 \in \mathbb{R}^{k_2}, \quad Z_1 \in \mathbb{R}^{k_1}, \quad Z_2 \in \mathbb{R}^{\ell_2},$$
$$\Gamma_{12}' \in \mathbb{R}^{k_2 \times k_1}, \quad \Gamma_{22}' \in \mathbb{R}^{k_2 \times \ell_2}, \quad u_2 \in \mathbb{R}^{k_2}$$

Reduced Form for Y_1

- We substitute the reduced form of Y_2 into the structural equation (9):

$$Y_1 = Z_1'\beta_1 + (\Gamma_{12}'Z_1 + \Gamma_{22}'Z_2 + u_2)'\beta_2 + e \tag{12}$$

$$= Z_1'\lambda_1 + Z_2'\lambda_2 + u_1 \quad \text{Dimension?} \tag{13}$$

- where:

$$\lambda_1 = \beta_1 + \Gamma_{12}\beta_2 \quad (k_1 \times 1) \tag{14}$$

$$\lambda_2 = \Gamma_{22}\beta_2 \quad (\ell_2 \times 1) \tag{15}$$

$$u_1 = u_2'\beta_2 + e$$

Matrix Notation for Reduced Form Coefficients

- Define the combined reduced-form coefficient:

$$\lambda = \bar{\Gamma}\beta \tag{16}$$

- Where:

$$\bar{\Gamma} = \begin{bmatrix} I_{k_1} & \Gamma_{12} \\ 0 & \Gamma_{22} \end{bmatrix}$$

- This notation helps unify both reduced form equations under a common projection structure.

Full Reduced Form System (Stacked)

- We now write the reduced form for both Y_1 and Y_2:

$$Y_1 = \lambda' Z + u_1$$
$$Y_2 = \Gamma' Z + u_2$$

- Or in stacked matrix form:

$$\vec{Y} = \begin{bmatrix} \lambda_1' & \lambda_2' \\ \Gamma_{12}' & \Gamma_{22}' \end{bmatrix} Z + u, \quad \text{where } \mathbf{u} = (u_1, u_2')' \tag{17}$$

Estimating Reduced Form Coefficients

- The relationships in Equations (16) show how reduced form parameters (Γ, λ) relate to the structural parameters (β_1, β_2)
- The reduced form equations are **linear projections**, so their coefficients can be estimated by **least squares**
- The least squares estimators for: $Y_2 = \Gamma' Z + u_2$ and $Y_1 = \lambda' Z + u_1$ are:

$$\hat{\Gamma} = \left(\sum_{i=1}^{n} Z_i Z_i' \right)^{-1} \left(\sum_{i=1}^{n} Z_i Y_{2i}' \right) \tag{18}$$

$$\hat{\lambda} = \left(\sum_{i=1}^{n} Z_i Z_i' \right)^{-1} \left(\sum_{i=1}^{n} Z_i Y_{1i} \right) \tag{19}$$

- These use the usual OLS formula for projection.

Identification via Reduced Form

- A parameter is **identified** if it is a unique function of the distribution of observables
- One way to prove identification is to express parameters in terms of **population moments**
- For example, the reduced form coefficients are identified:

$$\Gamma = \mathbb{E}[ZZ']^{-1}\mathbb{E}[ZY_2'] \tag{20}$$

$$\lambda = \mathbb{E}[ZZ']^{-1}\mathbb{E}[ZY_1] \tag{21}$$

- These are uniquely determined under Definition 1 (if $\mathbb{E}[ZZ']$ is invertible)
- We're interested in **identifying the structural parameter** β

Identifying the Structural Parameter β

- β relates to reduced form parameters via:

$$\lambda = \bar{\Gamma}\beta \tag{16}$$

- This is a system of ℓ equations in k unknowns
  $\rightarrow$ Uniquely solvable **iff**:

$$\text{rank}(\bar{\Gamma}) = k \tag{22}$$

- Since $\bar{\Gamma} = \mathbb{E}[ZZ']^{-1}\mathbb{E}[ZX']$, we substitute into (16):

$$\mathbb{E}[ZZ']^{-1}\mathbb{E}[ZY_1] = \mathbb{E}[ZZ']^{-1}\mathbb{E}[ZX']\beta$$

Identifying the Structural Parameter β (Cont.)

- Multiply both sides by $\mathbb{E}[ZZ']$:

$$\mathbb{E}[ZY_1] = \mathbb{E}[ZX']\beta$$

- This is a system of ℓ equations in k unknowns

  $\rightarrow$ Has unique solution **iff**:

$$\text{rank}(\mathbb{E}[ZX']) = k \tag{23}$$

- Equations (22) and (23) are **equivalent**
  $\rightarrow$ This is the **relevance condition** for identification

Solving for β under Identification

- In the just-identified case ($\ell = k$), Equation (22) implies $\bar{\Gamma}$ is invertible, so:

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

- When $\ell > k$, solve the system $\lambda = \bar{\Gamma}\beta$ by least squares:

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

- This system has ℓ equations in k unknowns and no error, so it's solvable if:

$$\text{rank}(\bar{\Gamma}) = k$$

- Since $\bar{\Gamma}$ has the block form:

$$\bar{\Gamma} = \begin{bmatrix} I_{k_1} & \Gamma_{12} \\ 0 & \Gamma_{22} \end{bmatrix}$$

- Then: $\text{rank}(\bar{\Gamma}) = k$ **if and only if** $\text{rank}(\Gamma_{22}) = k_2$

Explicit Solutions for β_1 and β_2

- The model is identified if Γ_{22} has full rank (k_2), since that block determines β_2
  From (15): $\lambda_2 = \Gamma_{22}\beta_2$
- Solve for β_2:

$$\beta_2 = (\Gamma_{22}'\Gamma_{22})^{-1}\Gamma_{22}'\lambda_2$$

- From (14): $\lambda_1 = \beta_1 + \Gamma_{12}\beta_2$
- Plug in β_2 to solve for β_1:

$$\beta_1 = \lambda_1 - \Gamma_{12}(\Gamma_{22}'\Gamma_{22})^{-1}\Gamma_{22}'\lambda_2$$

- In the **just-identified case** ($\ell_2 = k_2$), this simplifies to (as Γ_{22} is square $k_2 \times k_2$ and ivertible matrix):

$$\beta_2 = \Gamma_{22}^{-1}\lambda_2, \quad \beta_1 = \lambda_1 - \Gamma_{12}\Gamma_{22}^{-1}\lambda_2$$

Instrumental Variables Estimator

- Consider the just-identified case: $\ell = k$
- The IV condition implies $\mathbb{E}[Ze] = 0$
- Let $e = Y_1 - X'\beta$, then:

$$\mathbb{E}[Z(Y_1 - X'\beta)] = 0$$

- Expand:

$$\mathbb{E}[ZY_1] - \mathbb{E}[ZX']\beta = 0$$

- Solve for β:

$$\beta = (\mathbb{E}[ZX'])^{-1}\,\mathbb{E}[ZY_1]$$

- Requires $\mathbb{E}[ZX']$ to be invertible (see equations 7 and 23)

Sample IV Estimator

- Replace population moments with sample averages:

$$\hat{\beta}_{\text{IV}} = \left(\frac{1}{n}\sum_{i=1}^{n} Z_i X_i'\right)^{-1} \left(\frac{1}{n}\sum_{i=1}^{n} Z_i Y_{1i}\right)$$

- Which simplifies to:

$$\hat{\beta}_{\text{IV}} = \left(\sum_{i=1}^{n} Z_i X_i'\right)^{-1} \left(\sum_{i=1}^{n} Z_i Y_{1i}\right) \tag{24}$$

- More generally, for any instrument $W \in \mathbb{R}^k$:

$$\hat{\beta}_{\text{IV}} = \left(\sum_{i=1}^{n} W_i X_i'\right)^{-1} \left(\sum_{i=1}^{n} W_i Y_{1i}\right)$$

- This generalizes the IV estimator to **arbitrary instruments** W

Indirect Least Squares (ILS) Estimator

- When $\ell = k$, we can write:

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

- Replacing $\bar{\Gamma}$ and λ with their least squares estimators (from 18–19), we define the: **Indirect Least Squares (ILS) estimator**:

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

- Using matrix algebra:

$$\begin{aligned} \hat{\beta}_{\text{ILS}} &= \left((\mathbf{Z}'\mathbf{Z})^{-1}(\mathbf{Z}'\mathbf{X})\right)^{-1}\left((\mathbf{Z}'\mathbf{Z})^{-1}(\mathbf{Z}'\mathbf{Y}_1)\right) \\ &= (\mathbf{Z}'\mathbf{X})^{-1}(\mathbf{Z}'\mathbf{Z})(\mathbf{Z}'\mathbf{Z})^{-1}(\mathbf{Z}'\mathbf{Y}_1) \\ &= (\mathbf{Z}'\mathbf{X})^{-1}(\mathbf{Z}'\mathbf{Y}_1) \end{aligned}$$

- So $\hat{\beta}_{\text{ILS}} = \hat{\beta}_{\text{IV}}$ – the two estimators are **identical**

IV Residuals and Orthogonality

- Define IV residuals:

$$\hat{e}_i = Y_{1i} - X_i' \hat{\beta}_{\text{IV}}$$

- These satisfy the moment condition:

$$\mathbf{Z}'\hat{\mathbf{e}} = \mathbf{Z}'\mathbf{Y}_1 - \mathbf{Z}'\mathbf{X}(\mathbf{Z}'\mathbf{X})^{-1}(\mathbf{Z}'\mathbf{Y}_1) = 0 \quad (25)$$

- Interpretation:
  - Residuals are orthogonal to all instruments in Z
  - If Z includes a constant, residuals sum to zero
  - Residuals are also uncorrelated with **included and excluded instruments**

Example: IV Regression with College Instrument

- To illustrate IV regression, we estimate the **reduced form equations**:
  - Treat **education** as an **endogenous regressor**
  - Use **college** as the **instrumental variable**
- The reduced form equations are for:
  - $\log(\text{wage})$
  - *education*
- These are reported in the **first and second columns of next Table**

IV Regression - Example

	log(wage)	education	education	experience	$experience^2/100$	education
experience	0.053	−0.410				−0.413
	(0.007)	(0.032)				(0.032)
$experience^2/100$	−0.219	0.073				0.093
	(0.033)	(0.170)				(0.171)
black	−0.264	−1.006	−1.468	1.468	0.282	−1.006
	(0.018)	(0.088)	(0.115)	(0.115)	(0.026)	(0.088)
south	−0.143	−0.291	−0.460	0.460	0.112	−0.267
	(0.017)	(0.078)	(0.103)	(0.103)	(0.022)	(0.079)
urban	0.185	0.404	0.835	−0.835	−0.176	0.400
	(0.017)	(0.085)	(0.112)	(0.112)	(0.025)	(0.085)
college	0.045	0.337	0.347	−0.347	−0.073	
	(0.016)	(0.081)	(0.109)	(0.109)	(0.023)	
public						0.430
						(0.086)
private						0.123
						(0.101)
age			1.061	−0.061	−0.555	
			(0.296)	(0.296)	(0.065)	
$age^2/100$			−1.876	1.876	1.313	
			(0.516)	(0.516)	(0.116)	
F		17.51	8.22	1581	1112	13.87

Interpreting IV Estimates for Education

- Focus: coefficient on **education**, using **college** as the excluded instrument
  - Estimate: 0.337 (small standard error)
  - Interpretation: Growing up near a 4-year college raises education by 0.3 years
- Since the model is just-identified:
  - $\hat{\beta}_{\text{ILS}} = \hat{\lambda}/\hat{\Gamma}$ (ratio form)
  - e.g. $0.045/0.337 \approx 0.13 \rightarrow$ implies a **13% return** to one year of education
- This is larger than the **7% return** from OLS (previous Table, column 1)
- Why might the IV estimate be larger?
  - OLS may be **upward biased** due to endogenous schooling choice
  - But IV should correct this, so result seems surprising

Extending Instruments: Experience and Age

- If **education** is endogenous, then **experience = age − education** is also endogenous
- Card (1995) suggests using: **age** and **age²** as instruments for **experience** and **experience²**
- These are valid instruments:
  - Exogenous (not choice variables)
  - Correlated with the endogenous regressors
- Observation:
  - Coefficients for **education** and **experience** equations are mirror images, except for **age**
  - Due to linear relation: $\text{age} = \text{education} + \text{experience}$
- New result:
  - Return to schooling stable across instruments
  - **Quadratic return to experience (experience²)** flattens

Demeaned Representation of IV Estimator

- Does the **demeaned representation** from OLS carry over to IV?
- Start by writing the linear projection equation as:

$$Y_1 = X'\beta + \alpha + e$$

- Likewise, partition the instrument as $(1, Z)$ where Z does **not** contain a constant
- For the ith observation:

$$Y_{1i} = X_i'\widehat{\beta}_{\text{IV}} + \hat{a}_{\text{IV}} + \hat{e}_i$$

- Orthogonality (25) implies:

$$\sum_{i=1}^{n}(Y_{1i} - X_i'\widehat{\beta}_{\text{IV}} - \hat{a}_{\text{IV}}) = 0$$

$$\sum_{i=1}^{n} Z_i(Y_{1i} - X_i'\widehat{\beta}_{\text{IV}} - \hat{a}_{\text{IV}}) = 0$$

Solving for the Demeaned IV Estimator

- First equation implies:

$$\hat{a}_{\text{IV}} = \bar{Y}_1 - \bar{X}'\hat{\beta}_{\text{IV}}$$

- Substituting into second equation gives:

$$\sum_{i=1}^{n} Z_i \left((Y_{1i} - \bar{Y}_1) - (X_i - \bar{X})'\hat{\beta}_{\text{IV}} \right)$$

- Solving for $\hat{\beta}_{\text{IV}}$ yields:

$$\hat{\beta}_{\text{IV}} = \left(\sum_{i=1}^{n} Z_i (X_i - \bar{X})' \right)^{-1} \left(\sum_{i=1}^{n} Z_i (Y_{1i} - \bar{Y}_1) \right)$$

Solving for the Demeaned IV Estimator (Cont.)

- Which simplifies to:

$$\hat{\beta}_{\mathsf{IV}} = \left(\sum_{i=1}^{n} (Z_i - \bar{Z})(X_i - \bar{X})' \right)^{-1} \left(\sum_{i=1}^{n} (Z_i - \bar{Z})(Y_{1i} - \bar{Y}_1) \right) \tag{26}$$

- The IV estimator depends **only on demeaned data**

Notation Reference: Scalars

Symbol	Type	Definition
Y_1	scalar	Outcome variable (e.g. log wage)
e	scalar	Structural error term
u_1	scalar	Reduced form error for Y_1; equals $u_2'\beta_2 + e$
k_1	scalar	Number of exogenous regressors
k_2	scalar	Number of endogenous regressors
k	scalar	Total regressors; $k = k_1 + k_2$
ℓ	scalar	Total number of instruments
ℓ_2	scalar	Number of excluded instruments; $\ell_2 = \ell - k_1$
n	scalar	Sample size

Notation Reference: Vectors

Symbol	Dimension	Definition
X	$k \times 1$	Full regressor vector; $X = (X_1', X_2')'$
X_1	$k_1 \times 1$	Exogenous regressors
$X_2 = Y_2$	$k_2 \times 1$	Endogenous regressors
Z	$\ell \times 1$	Full instrument vector; $Z = (Z_1', Z_2')'$
Z_1	$k_1 \times 1$	Included exogenous instruments
Z_2	$\ell_2 \times 1$	Excluded instruments (the "true" IVs)
β	$k \times 1$	Structural parameter vector; $\beta = (\beta_1', \beta_2')'$
λ	$\ell \times 1$	Reduced form coefficients for Y_1 on Z
λ_1	$k_1 \times 1$	Reduced form coefficients on Z_1; $\lambda_1 = \beta_1 + \Gamma_{12}\beta_2$
λ_2	$\ell_2 \times 1$	Reduced form coefficients on Z_2; $\lambda_2 = \Gamma_{22}\beta_2$
u_2	$k_2 \times 1$	Reduced form error for Y_2
$\vec{Y}$	$(1 + k_2) \times 1$	Stacked endogenous variables; $\vec{Y} = (Y_1, Y_2')'$

Notation Reference: Matrices

Symbol	Dimension	Definition
Γ	$\ell \times k_2$	Reduced form coefficients for Y_2 on Z
Γ_{12}	$k_1 \times k_2$	Coefficients on Z_1 in reduced form for Y_2
Γ_{22}	$\ell_2 \times k_2$	Coefficients on Z_2 in reduced form for Y_2; rank$(\Gamma_{22}) = k_2$ for identification
$\bar{\Gamma}$	$\ell \times k$	Stacked matrix; $\bar{\Gamma} = \begin{bmatrix} I_{k_1} & \Gamma_{12} \\ 0 & \Gamma_{22} \end{bmatrix}$
$\mathbf{P}_Z$	$n \times n$	Projection matrix; $\mathbf{P}_Z = \mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'$
$\mathbf{Z}'\mathbf{Z}$	$\ell \times \ell$	Sample moment matrix of instruments
$\mathbf{Z}'\mathbf{X}$	$\ell \times k$	Sample cross-moment of instruments and regressors
$\mathbf{Z}'\mathbf{Y}_1$	$\ell \times 1$	Sample cross-moment of instruments and outcome

Motivation: From IV to 2SLS

- The IV estimator from the previous section assumed **exact identification**: $\ell = k$
- Now we allow the **general case** $\ell \geq k$ (overidentification)
- Recall the reduced-form equation (13) and (16):

$$Y_1 = Z'\bar{\Gamma}\beta + u_1, \qquad \mathbb{E}[Zu_1] = 0$$

- Define $W = \bar{\Gamma}'Z$. Then:

$$Y_1 = W'\beta + u_1, \qquad \mathbb{E}[Wu_1] = 0$$

- Intuition: Z is a set of **candidate instruments**; $W = \bar{\Gamma}'Z$ is a $1 \times k$-dimensional set of **linear combinations** of Z

From Infeasible to Feasible Estimation

- **If $\bar{\Gamma}$ were known**, we would estimate β by OLS of Y_1 on $W = \bar{\Gamma}' Z$:

$$\hat{\beta} = (\mathbf{W}'\mathbf{W})^{-1}(\mathbf{W}'\mathbf{Y}) = \left(\bar{\Gamma}'\mathbf{Z}'\mathbf{Z}\bar{\Gamma}\right)^{-1}\left(\bar{\Gamma}'\mathbf{Z}'\mathbf{Y}_1\right)$$

- But $\bar{\Gamma}$ is **unknown** – we must estimate it from the reduced form
- The reduced-form estimator is:

$$\hat{\bar{\Gamma}} = (\mathbf{Z}'\mathbf{Z})^{-1}(\mathbf{Z}'\mathbf{X})$$

- Replacing $\bar{\Gamma}$ with $\hat{\bar{\Gamma}}$ gives the **2SLS estimator**

Deriving the 2SLS Estimator

Substituting $\hat{\hat{\Gamma}} = (\mathbf{Z}'\mathbf{Z})^{-1}(\mathbf{Z}'\mathbf{X})$:

$$\hat{\beta}_{\text{2sls}} = \left(\hat{\hat{\Gamma}}'\mathbf{Z}'\mathbf{Z}\hat{\hat{\Gamma}}\right)^{-1}\left(\hat{\hat{\Gamma}}'\mathbf{Z}'\mathbf{Y}_1\right)$$

$$= \left(\mathbf{X}'\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{X}\right)^{-1}\mathbf{X}'\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{Y}_1$$

$$\boxed{\hat{\beta}_{\text{2sls}} = \left(\mathbf{X}'\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{X}\right)^{-1}\mathbf{X}'\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{Y}_1} \tag{27}$$

- This is the **Two-Stage Least Squares (2SLS) estimator**
- Originally proposed by Theil (1953) and Basmann (1957)
- Standard estimator for **linear equations with instruments**

2SLS Reduces to IV When $k = \ell$

- When the model is **just-identified** ($k = \ell$), $\mathbf{X}'\mathbf{Z}$ and $\mathbf{Z}'\mathbf{X}$ are square, so:

$$\left(\mathbf{X}'\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{X}\right)^{-1} = (\mathbf{Z}'\mathbf{X})^{-1}(\mathbf{Z}'\mathbf{Z})(\mathbf{X}'\mathbf{Z})^{-1}$$

- Substituting back into (27):

$$\begin{aligned}\widehat{\beta}_{\text{2sls}} &= (\mathbf{Z}'\mathbf{X})^{-1}(\mathbf{Z}'\mathbf{Z})(\mathbf{X}'\mathbf{Z})^{-1}\mathbf{X}'\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{Y}_1 \\ &= (\mathbf{Z}'\mathbf{X})^{-1}(\mathbf{Z}'\mathbf{Z})(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{Y}_1 \\ &= (\mathbf{Z}'\mathbf{X})^{-1}\mathbf{Z}'\mathbf{Y}_1 = \widehat{\beta}_{\text{iv}}\end{aligned}$$

- So **2SLS** $\equiv$ **IV** in the just-identified case – 2SLS is a generalization of IV

Alternative Representation: Projection Matrix (1)

- Define the **projection matrix** onto the column space of Z:

$$\mathbf{P}_Z = \mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'$$

- We can write the 2SLS estimator more compactly as:

$$\hat{\beta}_{\text{2sls}} = (\mathbf{X}'\mathbf{P}_Z\mathbf{X})^{-1}\mathbf{X}'\mathbf{P}_Z\mathbf{Y}_1$$

- Useful for **theoretical derivations** but **not for computation** – $\mathbf{P}_Z$ is $n \times n$, too large when n is large
- Define **fitted values** of $\mathbf{X}$ from the reduced form $\hat{\mathbf{X}} = \mathbf{P}_Z\mathbf{X} = \mathbf{Z}\hat{\hat{\Gamma}}$
- Then since $\mathbf{P}_Z$ is **idempotent** ($\mathbf{P}_Z\mathbf{P}_Z = \mathbf{P}_Z$):

$$\hat{\beta}_{\text{2sls}} = (\mathbf{X}'\mathbf{P}_Z\mathbf{P}_Z\mathbf{X})^{-1}\mathbf{X}'\mathbf{P}_Z\mathbf{Y}_1 = (\hat{\mathbf{X}}'\hat{\mathbf{X}})^{-1}\hat{\mathbf{X}}'\mathbf{Y}_1$$

- This is **OLS of Y_1 on $\hat{\mathbf{X}}$** – the source of the "two-stage" name!

The Two Stages Explicitly (2)

- Regress $\mathbf{X}$ on $\mathbf{Z}$ to obtain $\hat{\hat{\Gamma}} = (\mathbf{Z}'\mathbf{Z})^{-1}(\mathbf{Z}'\mathbf{X})$ and $\hat{\mathbf{X}} = \mathbf{Z}\hat{\hat{\Gamma}} = \mathbf{P}_Z\mathbf{X}$
- This is an **IV estimator** using $\hat{X}$ as an instrument for X:

$$\hat{\beta}_{\text{2sls}} = (\hat{X}'X)^{-1}\hat{X}'Y_1$$

- The source of the "two-stage" name – it can be computed as follows:
  - **Stage 1:** Regress X on Z: obtain $\hat{\hat{\Gamma}} = (\mathbf{Z}'\mathbf{Z})^{-1}(\mathbf{Z}'\mathbf{X})$ and $\hat{\mathbf{X}} = \mathbf{Z}\hat{\hat{\Gamma}} = \mathbf{P}_Z X$
  - **Stage 2:** Regress Y_1 on $\hat{X}$: obtain $\hat{\beta}_{\text{2sls}} = (\hat{\mathbf{X}}'\hat{\mathbf{X}})^{-1}\hat{\mathbf{X}}'\mathbf{Y}_1$

What Gets Replaced in the Second Stage? (3)

- Recall $\mathbf{X} = [\mathbf{Z}_1, Y_2]$ and $\mathbf{Z} = [\mathbf{Z}_1, \mathbf{Z}_2]$
- Notice $\hat{X}_1 = \mathbf{P}_Z \mathbf{Z}_1 = \mathbf{Z}_1$ since $\mathbf{Z}_1$ lies in the span of $\mathbf{Z}$
- Therefore: $\hat{X} = [\hat{X}_1, \hat{Y}_2] = [\mathbf{Z}_1, \hat{Y}_2]$
- In Stage 2 we regress Y_1 on $\mathbf{Z}_1$ and $\hat{Y}_2$ – only **endogenous variables** Y_2 are replaced by their fitted values:

$$\hat{Y}_2 = \hat{\bar{\bar{\Gamma}}}'_{12} \mathbf{Z}_1 + \hat{\bar{\bar{\Gamma}}}'_{22} \mathbf{Z}_2$$

- The included exogenous variables $\mathbf{Z}_1$ pass through **unchanged**

2SLS Residuals

- Define **2SLS residuals**: $\hat{e}_i = Y_{1i} - X_i'\hat{\beta}_{\text{2sls}}$
- When **overidentified** ($\ell > k$): there are ℓ instruments but only k moment conditions are satisfied – so $\mathbf{Z}'\hat{e} \neq 0$ in general
- However, the fitted values $\hat{X} = \mathbf{P}_Z X$ **are** orthogonal to residuals:

$$\hat{\mathbf{X}}'\hat{\mathbf{e}} = \hat{\hat{\Gamma}}'\mathbf{Z}'\hat{\mathbf{e}}$$

$$= \mathbf{X}'\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\hat{e}$$

$$= \mathbf{X}'\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{Y}_1 - \mathbf{X}'\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{X}\hat{\beta}_{\text{2sls}} = 0$$

Card's Proximity Example (Public & Private Instruments)

- Original instrument: `college` = grew up near any 4-year college
- New approach: Use **two instruments**
  - `public`: Grew up near a **public** 4-year college
  - `private`: Grew up near a **private** 4-year college
- One endogenous regressor: `education`
- Instruments: `public`, `private`
- Reduced form:
  - Coefficient on `public` is **larger**
  - Coefficient on `private` is **much smaller**
  - Education proximity impact is mainly via **public colleges**

Card's Proximity Example (Public & Private Instruments)

- **2SLS Estimate**:
  - From First Table (col. 4):
    - $\hat{\beta}_{\text{education}} = 0.161$
    - 16% return to a year of education
    - Roughly **double** the OLS estimate (7%)

Card's Proximity Example (Public & Private Instruments)

	OLS	IV(a)	IV(b)	2SLS(a)	2SLS(b)	LIML
education	0.074	0.132	0.133	0.161	0.160	0.164
	(0.004)	(0.049)	(0.051)	(0.040)	(0.041)	(0.042)
experience	0.084	0.107	0.056	0.119	0.047	0.120
	(0.007)	(0.021)	(0.026)	(0.018)	(0.025)	(0.019)
$\text{experience}^2/100$	−0.224	−0.228	−0.080	−0.231	−0.032	−0.231
	(0.032)	(0.035)	(0.133)	(0.037)	(0.127)	(0.037)
Black	−0.190	−0.131	−0.103	−0.102	−0.064	−0.099
	(0.017)	(0.051)	(0.075)	(0.044)	(0.061)	(0.045)
south	−0.125	−0.105	−0.098	−0.095	−0.086	−0.094
	(0.015)	(0.023)	(0.0284)	(0.022)	(0.026)	(0.022)
urban	0.161	0.131	0.108	0.116	0.083	0.115
	(0.015)	(0.030)	(0.049)	(0.026)	(0.041)	(0.027)
Sargan				0.82	0.52	0.82
p-value				0.37	0.47	0.37

Adding More Endogenous Variables

- Now treat:
  - `education`, `experience`, `experience^2` as endogenous
  - Use **four instruments**: `public`, `private`, `age`, `age^2`
- From Table (col. 5):
  - Education return stays ~16%
  - Return to experience flattens (quadratic term shrinks)
- Can we use all three college instruments: `college`, `public`, `private`?

  No:

  - `college = public + private`
  - Linear dependence → **violates full rank** (fails 6)

Reading

- **Hansen (2022), Econometrics.**
  - Sections: 12.1 - 12.10 and 12.12