Instrumental Variables

ECON 31720 Applied Microeconometrics

Francesco Ruggieri

The University of Chicago

October 28, 2020

Framework for Instrumental Variables

- $Y \in \mathbb{R}$ is an **outcome** of interest
- $X \in \mathbb{R}^{d_x}$ is a vector of **observed determinants** of Y that may be **partitioned** into
  - $D \in \mathbb{R}^{d_d}$, a vector of observed determinants of interest
  - $W \in \mathbb{R}^{d_w}$, a vector of control variables that typically include a deterministic constant
- $U \in \mathbb{R}$ encompasses all of the **unobserved determinants** of Y
- A **linear all-causes model** of the observed and unobserved determinants of the outcome:

$$Y = X'\beta + U = D'\alpha + W'\gamma + U$$

- Observed and unobserved determinants of Y are **systematically related**: $\mathbb{E}[XU] \neq 0_{d_x}$

Framework for Instrumental Variables

Because U has a causal interpretation and $\mathbb{E}[XU] \neq 0_{d_x}$:

- The **orthogonality condition** imposed by **linear regression**, $\mathbb{E}[XU] = 0_{d_x}$, does not logically match the systematic relationship between observed and unobserved determinants of Y
- Solution: consider a vector of **instrumental variables**, $Z \in \mathbb{R}^{d_z}$, such that $\mathbb{E}[ZU] = 0_{d_z}$

To identify the vector of causal parameters β, make the following **assumptions**:

1. **Exclusion**: Z is not a direct determinant of Y, i.e., $Y = X'\beta + Z'\eta + U$ is such that $\eta = 0_{d_z}$
2. **Exogeneity**: Z and U are orthogonal, i.e., $\mathbb{E}[ZU] = 0_{d_z}$
3. **Relevance**: $\mathbb{E}[ZX']$ has full rank

Framework for Instrumental Variables

Under these assumptions, the orthogonality between Z and U can be restated as

$$\mathbb{E}[ZU] = 0_{d_z} \iff \mathbb{E}[Z(Y - X'\beta)] = 0_{d_z} \iff \mathbb{E}[ZY] = \mathbb{E}[ZX']\beta$$

1. if $d_z = d_x$, $\mathbb{E}[ZX']$ is an invertible matrix, and the **Instrumental Variables** estimand is

$$\beta_{\text{IV}} \equiv \mathbb{E}[ZX']^{-1}\mathbb{E}[ZY]$$

2. if $d_z > d_x$, pre-multiply $\mathbb{E}[ZX']$ by a $d_x \times d_z$ matrix of **deterministic constants** c, so that

$$\beta_{\text{IV}} \equiv \mathbb{E}[cZX']^{-1}\mathbb{E}[cZY]$$

   If c is chosen to be the transpose of the matrix of first-stage **regression coefficients**,

$$\beta_{\text{TSLS}} \equiv \mathbb{E}[\pi' ZX']^{-1}\mathbb{E}[\pi' ZY]$$

   is the **Two-Stage Least Squares** estimand, where $\pi \equiv \mathbb{E}[ZZ']^{-1}\mathbb{E}[ZX']$

The Two-Sample Two-Stage Least Squares Estimator

- Consider $\{Y_i, X_i, Z_i\}_{i=1}^n$, a **sample** of i.i.d. draws from the joint distribution of (Y, X, Z)

- The **Two-Stage Least Squares estimator** of β is the sample analog of β_{TSLS}:

$$\widehat{B}_{\text{TSLS}} \equiv \left(\frac{1}{n} \sum_{i=1}^{n} \widehat{\Pi}' Z_i X_i' \right)^{-1} \left(\frac{1}{n} \sum_{i=1}^{n} \widehat{\Pi}' Z_i Y_i \right)$$

- The Weak Law of Large Numbers and the Continuous Mapping Theorem imply that $\widehat{B}_{\text{TSLS}} \xrightarrow{p} \beta$

- In addition, by the Central Limit Theorem and the Continuous Mapping Theorem,

$$\sqrt{n} \left(\widehat{B}_{\text{TSLS}} - \beta \right) \xrightarrow{d} \mathcal{N} \left(0, \mathbb{E} \left[\pi' Z X' \right]^{-1} \pi' \text{Var} \left[Z U \right] \pi \mathbb{E} \left[X Z' \pi \right]^{-1} \right)$$

- The Two-Stage Least Squares estimator is **consistent** for β and **asymptotically normal**

The Two-Sample Two-Stage Least Squares Estimator

- Suppose a single sample from the **joint distribution** of (Y, X, Z) were **not available**
  - In other words, no sample contains joint information on Y, X, and Z
- However, **two independent samples** are available: $\{Y_i^A, Z_i^A\}_{i=1}^{n_A}$ and $\{X_i^B, Z_i^B\}_{i=1}^{n_B}$
  - Importantly, both samples include information on the vector of instrumental variables
- A classic example of this setting is Angrist and Krueger (1992)

1. Framework for Instrumental Variables
2. The Two-Sample Two-Stage Least Squares Estimator
   - Angrist and Krueger (1992)
3. Framework for Instrumental Variables with Heterogeneity
4. A LATE Extension: Multiple Unordered Treatments
   - Kirkeboen, Leuven, and Mogstad (2016)
   - Kline and Walters (2016)
5. Summary

Angrist and Krueger (1992)

- Goal: estimate the effect of age at school entry on educational attainment
- Setting: states allow children to **enroll** in primary school if their age is 6 at given date **cutoffs**
- In addition, students are allowed to **leave** school as soon as they turn 16
- Assumption: the share of students dropping out at 16 is fixed and independent of birth date
- Implication: students born **earlier** in the year attain, on average, **less education**
- Angrist and Krueger (1992) instruments entry age with **quarter-of-birth indicators**

Angrist and Krueger (1992)

- A dataset containing age of school entry (X) and years of schooling (Y) is **not available**
- Angrist and Krueger (1992) uses **two distinct samples**:
  1. The **1960** Census to compute age at entry (and quarter of birth)
  2. The **1980** Census to back out years of completed schooling and (quarter of birth)
- The authors propose a **Two-Sample Two-Stage Least Squares** estimator:

$$\widehat{B}_{\text{TSTSLS}} \equiv \left(\frac{1}{n_{60}} \sum_{i=1}^{n_{60}} \widehat{\Pi}'_{60} Z_i^{60} X_i'^{60} \right)^{-1} \left(\frac{1}{n_{80}} \sum_{i=1}^{n_{80}} \widehat{\Pi}'_{60} Z_i^{80} Y_i^{80} \right)$$

where

$$\widehat{\Pi}_{60} \equiv \left(\frac{1}{n_{60}} \sum_{i=1}^{n_{60}} Z_i^{60} Z_i'^{60} \right)^{-1} \left(\frac{1}{n_{60}} \sum_{i=1}^{n_{60}} Z_i^{60} X_i'^{60} \right)$$

Angrist and Krueger (1992)

- More in general, given the two samples A and B defined above, consider

$$\widehat{B}_{\text{TSTSLS}} \equiv \left(\frac{1}{n_B} \sum_{i=1}^{n_B} \widehat{\Pi}'_B Z_i^B X_i'^B \right)^{-1} \left(\frac{1}{n_A} \sum_{i=1}^{n_A} \widehat{\Pi}'_B Z_i^A Y_i^A \right)$$

- If both samples contain **independent and identically distributed** random variables,

$$\widehat{B}_{\text{TSTSLS}} \xrightarrow{p} \beta_{\text{TSLS}} \equiv \mathbb{E}\left[\pi' Z X'\right]^{-1} \mathbb{E}\left[\pi' Z Y\right]$$

applying the Weak Law of Large Numbers and the Continuous Mapping Theorem

Angrist and Krueger (1992)

- **Alternative consistent estimators** can be constructed exploiting these two samples:

$$\begin{aligned}\widehat{B}^{(1)}_{\text{TSTSLS}} &= \left(\frac{1}{n_B}\sum_{i=1}^{n_B}\widehat{\Pi}'_B Z_i^B X_i'^B\right)^{-1}\left(\frac{1}{n_A}\sum_{i=1}^{n_A}\widehat{\Pi}'_B Z_i^A Y_i^A\right)\\ &= \left(\frac{1}{n_B}\sum_{i=1}^{n_B}\widehat{\Pi}'_B Z_i^B\left(\widehat{\Pi}'_B Z_i^B + R_i^B\right)'\right)^{-1}\left(\frac{1}{n_A}\sum_{i=1}^{n_A}\widehat{\Pi}'_B Z_i^A Y_i^A\right)\\ &= \left(\frac{1}{n_B}\sum_{i=1}^{n_B}\widehat{\Pi}'_B Z_i^B Z_i'^B \widehat{\Pi}_B\right)^{-1}\left(\frac{1}{n_A}\sum_{i=1}^{n_A}\widehat{\Pi}'_B Z_i^A Y_i^A\right)\end{aligned}$$

- **Replacing Z^B with Z^A** in the first matrix yields the alternative estimator for β,

$$\widehat{B}^{(2)}_{\text{TSTSLS}} \equiv \left(\frac{1}{n_A}\sum_{i=1}^{n_A}\widehat{\Pi}'_B Z_i^A Z_i'^A \widehat{\Pi}_B\right)^{-1}\left(\frac{1}{n_A}\sum_{i=1}^{n_A}\widehat{\Pi}'_B Z_i^A Y_i^A\right)$$

Angrist and Krueger (1992)

- The matrix of **first-stage regression coefficients** can be estimated using sample A too:

$$\widehat{\Pi}_{AB} = \left(\frac{1}{n_A}\sum_{i=1}^{n_A} Z_i^A Z_i'^A\right)^{-1}\left(\frac{1}{n_B}\sum_{i=1}^{n_B} Z_i^B X_i'^B\right)$$

- Two **additional consistent estimators** for β are therefore

$$\widehat{B}^{(3)}_{\text{TSTSLS}} \equiv \left(\frac{1}{n_B}\sum_{i=1}^{n_B} \widehat{\Pi}'_{AB} Z_i^B Z_i'^B \widehat{\Pi}_{AB}\right)^{-1}\left(\frac{1}{n_A}\sum_{i=1}^{n_A} \widehat{\Pi}'_{AB} Z_i^A Y_i^A\right)$$

$$\widehat{B}^{(4)}_{\text{TSTSLS}} \equiv \left(\frac{1}{n_A}\sum_{i=1}^{n_A} \widehat{\Pi}'_{AB} Z_i^A Z_i'^A \widehat{\Pi}_{AB}\right)^{-1}\left(\frac{1}{n_A}\sum_{i=1}^{n_A} \widehat{\Pi}'_{AB} Z_i^A Y_i^A\right)$$

- Though consistent, these estimators are **numerically distinct** in **finite samples**

Angrist and Krueger (1992)

Moreover, if $d_z = d_x$, two-sample IV and two-sample TSLS are **not always equivalent**:

$$
\begin{aligned}
\widehat{B}^{(2)}_{\text{TSTSLS}} &\equiv \left(\frac{1}{n_A}\sum_{i=1}^{n_A}\widehat{\Pi}'_B Z_i^A Z_i'^A \widehat{\Pi}_B\right)^{-1}\left(\frac{1}{n_A}\sum_{i=1}^{n_A}\widehat{\Pi}'_B Z_i^A Y_i^A\right) \\
&= \widehat{\Pi}_B^{-1}\left(\frac{1}{n_A}\sum_{i=1}^{n_A} Z_i^A Z_i'^A\right)^{-1}\widehat{\Pi}_B'^{-1}\widehat{\Pi}'_B\left(\frac{1}{n_A}\sum_{i=1}^{n_A} Z_i^A Y_i^A\right) \\
&= \left(\frac{1}{n_B}\sum_{i=1}^{n_B} Z_i^B X_i'^B\right)^{-1}\underbrace{\left(\frac{1}{n_B}\sum_{i=1}^{n_B} Z_i^B Z_i'^B\right)\left(\frac{1}{n_A}\sum_{i=1}^{n_A} Z_i^A Z_i'^A\right)^{-1}}_{\neq i_{d_z\times d_z}}\left(\frac{1}{n_A}\sum_{i=1}^{n_A} Z_i^A Y_i^A\right) \\
&\neq \widehat{B}_{\text{TSIV}} \equiv \left(\frac{1}{n_B}\sum_{i=1}^{n_B} Z_i^B X_i'^B\right)^{-1}\left(\frac{1}{n_A}\sum_{i=1}^{n_A} Z_i^A Y_i^A\right)
\end{aligned}
$$

Angrist and Krueger (1992)

- Angrist and Krueger (1992) makes two additional assumptions:
  1. Moments estimated from A are **independent** from moments estimated from B
  2. Let $n_B\left[n_A\right]$ denote n_B as a function of n_A. Then the **ratio** n_A and n_B is **constant**:

$$\lim_{n_A \to \infty} \frac{n_A}{n_B\left[n_A\right]} = k \in \mathbb{R}$$

- Under these assumptions and focusing, for simplicity, on the **two-sample IV estimator**:

$$\begin{aligned} g\left(\beta\right) &\equiv \frac{1}{n_A}\sum_{i=1}^{n_A} Z_i^A Y_i^A - \frac{1}{n_B}\sum_{i=1}^{n_B} Z_i^B X_i'^B \beta \\ &= \frac{1}{n_A}\sum_{i=1}^{n_A} Z_i^A Y_i^A - \mathbb{E}\left[ZX'\right]\beta - \sqrt{\frac{n_A}{n_B}}\left(\frac{1}{\sqrt{n_A n_B}}\sum_{i=1}^{n_B} Z_i^B X_i'^B \beta - \sqrt{\frac{n_B}{n_A}}\mathbb{E}\left[ZX'\right]\beta\right) \end{aligned}$$

Angrist and Krueger (1992)

- Exploiting the two previous assumptions, $\sqrt{n_A}$ can be used as a **normalization**:

$$\sqrt{n_A} g(\beta) = \sqrt{n_A}\left(\frac{1}{n_A}\sum_{i=1}^{n_A} Z_i^A Y_i^A - \mathbb{E}[ZX']\beta\right) - \sqrt{k n_B}\left(\frac{1}{n_B}\sum_{i=1}^{n_B} Z_i^B X_i'^B \beta - \mathbb{E}[ZX']\beta\right)$$
$$\xrightarrow{d} \mathcal{N}(0, \phi_A + k\omega_B) = \mathcal{N}(0, \Phi)$$

- Thus, the two-sample Instrumental Variables estimator is **asymptotically normal**
  - Indeed, $\widehat{B}_{\text{TSIV}} - \beta$ is proportional to $g(\beta)$ and Slutsky's Theorem implies the result

- The authors propose a TSIV estimator that uses Φ as a **GMM weighting matrix**

- Inoue and Solon (2010) shows that estimators such as $\widehat{B}^{(2)}_{\text{TSTSLS}}$ are **more efficient**

Framework for Instrumental Variables with Heterogeneity

- $Y \in \mathbb{R}$ is an **outcome** of interest, $D \in \{0,1\}$ is a binary **treatment**
- D and Y are linked by **potential outcomes** $Y(0)$ and $Y(1)$
- Agents **choose** whether to **sort** into the treated or untreated arm
- This self-selection is thought to be based on **unobserved determinants** of the outcome:

$$Y(0), Y(1) \not\perp\!\!\!\perp D$$

- Suppose this self-selection could be shifted by an **instrumental variable** $Z \in \{0,1\}$
- Z and D are linked by **potential treatments** $D(0)$ and $D(1)$

Framework for Instrumental Variables with Heterogeneity

- Goal: estimate **some feature** of the distribution of the random variable $Y(1) - Y(0)$
  - The **effect** of D on Y is heterogeneous across agents

- Unobservables induce agents to **choose** $D = 0$ or $D = 1$, so Z must satisfy four assumptions:
  1. **Exclusion**: $Y(d, z) = Y(d)\ \forall d, z$
  2. **Exogeneity**: $(Y(0), Y(1), D(0), D(1)) \perp\!\!\!\perp Z$
  3. **Relevance**: $\mathrm{Cov}\left[D, Z\right] \neq 0$. If exogeneity holds, relevance implies $\mathbb{P}\left(D(0) = D(1)\right) < 1$
  4. **Monotonicity**: $\mathbb{P}\left(D(1) \geq D(0)\right) = 1$ or $\mathbb{P}\left(D(0) \geq D(1)\right) = 1$

- Under these assumptions, the IV estimand identifies the **Local Average Treatment Effect**:

$$\beta_{\text{IV}} \equiv \frac{\mathrm{Cov}\left[Z, Y\right]}{\mathrm{Cov}\left[Z, D\right]} = \frac{\mathbb{E}\left[Y|Z = 1\right] - \mathbb{E}\left[Y|Z = 0\right]}{\mathbb{E}\left[D|Z = 1\right] - \mathbb{E}\left[D|Z = 0\right]} = \mathbb{E}\left[Y(1) - Y(0)|D(1) > D(0)\right]$$

1. Framework for Instrumental Variables
2. The Two-Sample Two-Stage Least Squares Estimator
   - Angrist and Krueger (1992)
3. Framework for Instrumental Variables with Heterogeneity
4. A LATE Extension: Multiple Unordered Treatments
   - Kirkeboen, Leuven, and Mogstad (2016)
   - Kline and Walters (2016)
5. Summary

A LATE Extension: Multiple Unordered Treatments

- Extend the LATE framework to the case in which there exist **multiple unordered treatments**
  - Treatment states cannot be logically ranked
  - Examples are **discrete choice problems** of field of study, occupation, location, etc.
- For simplicity, let us focus on the case in which D is trinary, i.e., $D \in \{0, 1, 2\}$
- This setting implies three **treatment state indicators**:

$$D_0 \equiv \mathbb{I}[D = 0] \quad D_1 \equiv \mathbb{I}[D = 1] \quad D_2 \equiv \mathbb{I}[D = 2]$$

- Suppose there exists a trinary **instrument**, $Z \in \{0, 1, 2\}$, that shifts self-selection into D
- This setting again implies three **instrument indicators**:

$$Z_0 \equiv \mathbb{I}[Z = 0] \quad Z_1 \equiv \mathbb{I}[Z = 1] \quad Z_2 \equiv \mathbb{I}[Z = 2]$$

Kirkeboen, Leuven, and Mogstad (2016)

- Kirkeboen, Leuven, and Mogstad (2016) studies the effect of major choice on earnings

- A standard exclusion restriction implies **three** potential outcome random variables

- Each indicator D_j is associated with **two** potential treatment states

- **Observed** and **potential** outcomes are **linked** as follows:

$$
\begin{aligned}
Y &= Y(0) + (Y(1) - Y(0))\, D_1 + (Y(2) - Y(0))\, D_2 \\
D_j &= D_j(0) + (D_j(1) - D_j(0))\, Z_1 + (D_j(2) - D_j(0))\, Z_2 \quad \text{for } j \in \{1, 2\}
\end{aligned}
$$

- The Imbens and Angrist (1994) **monotonicity** assumption in this setting is

$$
D_1(1) \geq D_1(0) \quad \text{and} \quad D_2(2) \geq D_2(0)
$$

Being assigned instrument $Z = j$ does not make it less likely to choose major $D = j$

Kirkeboen, Leuven, and Mogstad (2016)

- A **raw comparison** of earnings by major is contaminated by **selection bias**:

$$\begin{aligned}\mathbb{E}[Y|D=2]-\mathbb{E}[Y|D=0] &= \mathbb{E}[Y(2)|D=2]-\mathbb{E}[Y(0)|D=0]\\ &= \underbrace{\mathbb{E}[Y(2)-Y(0)|D=2]}_{\text{payoff}}+\underbrace{\mathbb{E}[Y(0)|D=2]-\mathbb{E}[Y(0)|D=0]}_{\text{selection bias}}\end{aligned}$$

- Even if one could eliminate selection bias, the "payoff" would still be **hard to interpret**:

$$\begin{aligned}\mathbb{E}[Y(2)-Y(0)|D=2] &= \mathbb{E}\left[Y(2)-Y(0)|D=2,D_{/2}=0\right]\times\mathbb{P}\left(D_{/2}=0|D=2\right)\\ &+\mathbb{E}\left[Y(2)-Y(0)|D=2,D_{/2}=1\right]\times\mathbb{P}\left(D_{/2}=1|D=2\right)\end{aligned}$$

  where $D_{/2}$ denotes one's **next-best alternative**

- Absent selection bias, the OLS estimand is still a **weighted average** of "different" payoffs

Kirkeboen, Leuven, and Mogstad (2016)

- As usual, the issue of **selection bias** can be addressed with **instrumental variables**
- Is IV sufficient to identify parameters with a **clear economic interpretation**?
- Consider the **linear all-causes model**

$$Y = \beta_0 + \beta_1 D_1 + \beta_2 D_2 + U \quad \text{with} \quad \mathbb{E}[D_1 U] \neq 0 \text{ and } \mathbb{E}[D_2 U] \neq 0$$

- Re-express U in terms of **potential outcomes** and **potential treatments**:

$$\begin{aligned} U &\equiv Y(0) - \beta_0 + (Y(1) - Y(0) - \beta_1) D_1 + (Y(2) - Y(0) - \beta_2) D_2 \\ &\equiv Y(0) - \beta_0 \\ &\quad + (Y(1) - Y(0) - \beta_1)(D_1(0) + (D_1(1) - D_1(0)) Z_1 + (D_1(2) - D_1(0)) Z_2) \\ &\quad + (Y(2) - Y(0) - \beta_2)(D_2(0) + (D_2(1) - D_2(0)) Z_1 + (D_2(2) - D_2(0)) Z_2) \end{aligned}$$

Kirkeboen, Leuven, and Mogstad (2016)

- Define the **payoffs** $\Delta^1 \equiv Y(1) - Y(0)$ and $\Delta^2 \equiv Y(2) - Y(0)$

- Using this expression for U, the IV **orthogonality conditions** can be written as

$$\mathbb{E}[Z_1 U] = \mathbb{E}\left[\left(\Delta^1 - \beta_1\right)\left(D_1(1) - D_1(0)\right) + \left(\Delta^2 - \beta_2\right)\left(D_2(1) - D_2(0)\right)\right] = 0$$
$$\mathbb{E}[Z_2 U] = \mathbb{E}\left[\left(\Delta^1 - \beta_1\right)\left(D_1(2) - D_1(0)\right) + \left(\Delta^2 - \beta_2\right)\left(D_2(2) - D_2(0)\right)\right] = 0$$

- Solving this system of equations for β_1 and β_2 yields two **linear combinations** of
  - $\Delta^1 \equiv Y(1) - Y(0)$, the payoff of major 1 relative to major 0
  - $\Delta^2 \equiv Y(2) - Y(0)$, the payoff of major 2 relative to major 0
  - $\Delta^2 - \Delta^1 \equiv Y(2) - Y(1)$, the payoff of major 2 relative to major 1

- Thus, IV identifies **weighted averages** of **payoffs** to choosing different fields

Kirkeboen, Leuven, and Mogstad (2016)

For IV to identify interpretable parameters of interest, additional **assumptions** are needed:

1. **Constant effects**, i.e., payoffs to major choice are homogeneous across agents:

$$\beta_1 = \Delta^1 \equiv Y(1) - Y(0) \qquad \beta_2 = \Delta^2 \equiv Y(2) - Y(0)$$

2. **Restricting preferences** to $D_2(0) = D_2(1)$ and $D_1(0) = D_1(2)$:

$$\beta_1 = \mathbb{E}\left[\Delta^1 | D_1(1) - D_1(0) = 1\right] \qquad \beta_2 = \mathbb{E}\left[\Delta^2 | D_2(2) - D_2(0) = 1\right]$$

3. **Irrelevance and Next-Best Alternative**, i.e., $D_1(1) = D_1(0) = 0 \implies D_2(1) = D_2(0)$ and $D_2(2) = D_2(0) = 0 \implies D_1(2) = D_1(0)$:

$$\beta_1 = \mathbb{E}\left[\Delta^1 | D_1(1) - D_1(0) = 1, D_2(0) = 0\right] \qquad \beta_2 = \mathbb{E}\left[\Delta^2 | D_2(2) - D_2(0) = 1, D_1(0) = 0\right]$$

   Pair this assumption with info on next-best alternatives to identify **field-specific LATEs**.

Kline and Walters (2016)

- Kline and Walters (2016) studies patterns of substitution across public assistance programs
- Setting: the **Head Start Impact Study**, a 2002-2006 national longitudinal study
- Each Head Start applicant participates in one of **three** possible **treatments**, $D \in \{h, c, n\}$
  - h, c, n denote Head Start, other pre-school programs, and home care, respectively
- A binary **instrument** $Z \in \{0, 1\}$ indicates receipt of a **Head Start offer**
- The authors impose a theoretical **restriction on substitution patterns**:

$$D(1) \neq D(0) \implies D(1) = h$$

Receiving a Head Start offer does **not** induce any agent to **switch** between n and c

Kline and Walters (2016)

This restriction implies that Head Start applicants can be **partioned** into **five** groups:

1. ***n*-compliers**, $D(1) = h, D(0) = n$, **switch** from home care to Head Start
2. ***c*-compliers**, $D(1) = h, D(0) = c$, **switch** from other programs to Head Start
3. ***n*-never takers**, $D(1) = D(0) = n$, **never attend** Head Start and choose home care
4. ***c*-never takers**, $D(1) = D(0) = c$, **never attend** Head Start and choose other programs
5. ***h*-always takers**, $D(1) = D(0) = h$, **manage to enroll** in Head Start in any case

Kline and Walters (2016)

- Consider the **linear all-causes model** $Y = \alpha + \beta S + U$ with $\mathbb{E}[SU] \neq 0$
  - $S = 1$ if $D = h$, i.e., an applicant **participates** in the Head Start program

- The **Instrumental Variables** estimand of β is

$$\begin{aligned}\beta_{\text{IV}} = \frac{\text{Cov}[Z, Y]}{\text{Cov}[Z, S]} &= \frac{\mathbb{E}[Y|Z=1] - \mathbb{E}[Y|Z=0]}{\mathbb{E}[S|Z=1] - \mathbb{E}[S|Z=0]} \\ &= \frac{\mathbb{E}[Y|Z=1] - \mathbb{E}[Y|Z=0]}{\mathbb{E}[\mathbb{I}[D=h]|Z=1] - \mathbb{E}[\mathbb{I}[D=h]|Z=0]} \\ &= \mathbb{E}[Y(h) - Y(D(0))|D(1) = h, D(0) \neq h] \\ &\equiv \text{LATE}_h\end{aligned}$$

- LATE_h is the **average effect of Head Start** among **compliers**, where compliers
  - Have **different counterfactual choices**, i.e., include both n-compliers and c-compliers

Summary

- If a single sample containing draws from (Y, X, Z) is not available, the **Two-Sample Two-Stage Least Squares estimator** is still consistent and asymptotically normal

- An important **extension** of the LATE framework is **multiple unordered treatments**:
  - Kirkeboen, Leuven, and Mogstad (2016) shows that Instrumental Variables may eliminate selection bias, but does **not necessarily identify economically interpretable parameters**

  - Kline and Walters (2016) shows how theoretically **restricting agents' behavior** may allow one to identify a **salient** Local Average Treatment Effect when agents face a discrete choice