150C Causal Inference
Instrumental Variables: Modern Perspective with Heterogeneous Treatment Effects

Jonathan Mummolo

May 22, 2017
Two Views on Instrumental Variables

1. Traditional Econometric Framework
   - Constant treatment effects
   - Linearity in the case of a multivalued treatment

2. Potential Outcome Model of IV
   - Heterogeneous treatment effects
   - Focus in Local Average Treatment Effect (LATE)
Identification with Traditional Instrumental Variables

Definition

Two equations:
- \( Y = \gamma + \alpha D + \varepsilon \) (Second Stage)
- \( D = \tau + \rho Z + \eta \) (First Stage)

Identification Assumption

1. **Exogeneity and Exclusion**: \( \text{Cov}(Z, \eta) = 0 \) and \( \text{Cov}(Z, \varepsilon) = 0 \)
2. **First Stage**: \( \rho \neq 0 \)
3. **\( \alpha \)** = \( Y_{1,i} - Y_{0,i} \) constant for all units \( i \).
   *Or in the case of a multivalued treatment with \( s \) levels we assume \( \alpha = Y_{s,i} - Y_{s-1,i} \).*
True model: $Y = D\alpha + X\beta + \varepsilon$

Given the IV assumptions, we could regress: $Y = Z\rho + \omega$ and obtain an unbiased effect $\hat{\rho}$, the effect of $Z$ on $Y$

But we can also obtain an unbiased estimate of $\beta$, the effect of $D$ on $Y$ by using only the exogenous variation in $D$ that is induced by $Z$

Assume $\text{Cov}[\nu = \varepsilon + X\beta, Z] = 0$. 
Outline

Instrumental Variables with Potential Outcomes (No Covariates)
- Identification
- Estimation
- Examples
- Size of Complier Group
Potential Outcome Model for Instrumental Variables

**Definition (Instrument)**

\[ Z_i: \text{Binary instrument for unit } i. \]

\[ Z_i = \begin{cases} 
1 & \text{if unit } i \text{ “encouraged” to receive treatment} \\
0 & \text{if unit } i \text{ “encouraged” to receive control} 
\end{cases} \]

**Definition (Potential Treatments)**

\[ D_z \text{ indicates potential treatment status given } Z = z \]

- \( D_1 = 1 \) encouraged to take treatment and takes treatment

**Assumption**

*Observed treatments are realized as*

\[ D = Z \cdot D_1 + (1 - Z) \cdot D_0 \] so \( D_i = \begin{cases} 
D_{1i} & \text{if } Z_i = 1 \\
D_{0i} & \text{if } Z_i = 0 
\end{cases} \]
Following Angrist, Imbens, and Rubin (1996), we can define:

**Definition**
- **Compliers**: $D_1 > D_0$ ($D_0 = 0$ and $D_1 = 1$).
- **Always-takers**: $D_1 = D_0 = 1$.
- **Never-takers**: $D_1 = D_0 = 0$.
- **Defiers**: $D_1 < D_0$ ($D_0 = 1$ and $D_1 = 0$).

**Problem**

*Only one of the potential treatment indicators $(D_0, D_1)$ is observed, so we cannot identify which group any particular individual belongs to.*
### Who are the Compliers?

<table>
<thead>
<tr>
<th>Study</th>
<th>Outcome</th>
<th>Treatment</th>
<th>Instrument</th>
</tr>
</thead>
<tbody>
<tr>
<td>Angrist and Evans (1998)</td>
<td>Earnings</td>
<td>More than 2 Children</td>
<td>Multiple Second Birth (Twins)</td>
</tr>
<tr>
<td>Angrist and Evans (1998)</td>
<td>Earnings</td>
<td>More than 2 Children</td>
<td>First Two Children are Same Sex</td>
</tr>
<tr>
<td>Levitt (1997)</td>
<td>Crime Rates</td>
<td>Number of Policemen</td>
<td>Mayoral Elections</td>
</tr>
<tr>
<td>Angrist and Krueger (1991)</td>
<td>Earnings</td>
<td>Years of Schooling</td>
<td>Quarter of Birth</td>
</tr>
<tr>
<td>Angrist (1990)</td>
<td>Earnings</td>
<td>Veteran Status</td>
<td>Vietnam Draft Lottery</td>
</tr>
<tr>
<td>Miguel, Satyanath and Sergenti (2004)</td>
<td>Civil War Onset</td>
<td>GDP per capita</td>
<td>Lagged Rainfall</td>
</tr>
<tr>
<td>Acemoglu, Johnson and Robinson (2001)</td>
<td>Economic performance</td>
<td>Current Institutions</td>
<td>Settler Mortality in Colonial Times</td>
</tr>
<tr>
<td>Cleary and Barro (2006)</td>
<td>Religiosity</td>
<td>GDP per capita</td>
<td>Distance from Equator</td>
</tr>
</tbody>
</table>
**Definition (Potential Outcomes)**

Given the binary instrument $Z_i \in (1, 0)$ and the binary treatment $D_i \in (1, 0)$ every unit now has four potential outcomes $Y_i(D, Z)$:

- $Y(D = 1, Z = 1)$; $Y(D = 1, Z = 0)$; $Y(D = 0, Z = 1)$; $Y(D = 0, Z = 0)$

  e.g. the causal effect of the treatment given the unit’s realized encouragement status is given by $Y(D = 1, Z_i) - Y(D = 0, Z_i)$.

**Assumption (Ignorability)**

**Ignorability of the Instrument:** $(Y_0, Y_1, D_0, D_1) \perp\!\!\!\!\!\!\perp Z$

- **Independence:** $(Y(D, Z), D_1, D_0) \perp\!\!\!\!\!\!\perp Z$ which implies that causal effects of $Z$ on $Y$ and $Z$ on $D$ are identified.

- **Exclusion:** $Y(D, 0) = Y(D, 1)$ for $D = 0, 1$ so we can simply define potential outcomes indexed solely by treatment status: $(Y_1, Y_0)$
**Potential Outcome Model for Instrumental Variables**

**Estimand (LATE)**

\[ \alpha_{LATE} = E[Y_1 - Y_0 | D_1 > D_0] \] is defined as the Local Average Treatment Effect for Compliers

- This estimand varies with the particular instrument Z

**Proposition (Special Cases)**

- When the treatment intake, D, is itself randomized,
Estimand (LATE)

\[ \alpha_{LATE} = E[Y_1 - Y_0 | D_1 > D_0] \]

is defined as the Local Average Treatment Effect for Compliers

- This estimand varies with the particular instrument Z

Proposition (Special Cases)

- When the treatment intake, D, is itself randomized, then Z = D and every individual is a complier

- Given one-sided noncompliance, \( D_0 = 0 \):
Potential Outcome Model for Instrumental Variables

**Estimand (LATE)**

\[ \alpha_{LATE} = E[Y_1 - Y_0|D_1 > D_0] \]

is defined as the Local Average Treatment Effect for Compliers

- This estimand varies with the particular instrument Z

**Proposition (Special Cases)**

- When the treatment intake, D, is itself randomized, then \( Z = D \) and every individual is a complier

- Given one-sided noncompliance, \( D_0 = 0 \):

\[
E[Y_1|D_1 > D_0] = E[Y_1|D_1 = 1] = E[Y_1|Z = 1, D_1 = 1] = E[Y_1|D = 1], \text{ and}
\]

\[
E[Y_0|D_1 > D_0] = E[Y_0|D = 1]
\]

so \( \alpha_{LATE} = E[Y_1 - Y_0|D_1 > D_0] = E[Y_1 - Y_0|D = 1] = \alpha_{ATET} \)
Outline

1. Instrumental Variables with Potential Outcomes (No Covariates)
   - Identification
   - Estimation
   - Examples
   - Size of Complier Group
Identification with Instrumental Variables

Identification Assumption

1. **Ignorability of the Instrument**: \((Y_0, Y_1, D_0, D_1) \perp\!\!\!\!\!\!\perp Z\)
2. **First Stage**: \(0 < P(Z = 1) < 1\) and \(P(D_1 = 1) \neq P(D_0 = 1)\)
3. **Monotonicity**: \(D_1 \geq D_0\)

Identification Result

\[
E[Y_1 - Y_0 | D_1 > D_0] = \frac{E[Y|Z = 1] - E[Y|Z = 0]}{E[D|Z = 1] - E[D|Z = 0]}
\]

\[
= \frac{\text{Intent to Treat Effect of } Z \text{ on } Y}{\text{First Stage Effect of } Z \text{ on } D}
\]

\[
= \frac{\text{Intent to Treat Effect}}{\text{Proportion of Compliers}}
\]
Identification with Instrumental Variables

Identification Assumption

1. **Ignorability of the Instrument**: \((Y_0, Y_1, D_0, D_1) \perp Z\)
2. **First Stage**: \(0 < P(Z = 1) < 1\) and \(P(D_1 = 1) \neq P(D_0 = 1)\)
3. **Monotonicity**: \(D_1 \geq D_0\)

Proof.

\[
\frac{E[Y|Z = 1] - E[Y|Z = 0]}{E[D|Z = 1] - E[D|Z = 0]} = \frac{E[Y_0 + (Y_1 - Y_0)D_1|Z = 1] - E[Y_0 + (Y_1 - Y_0)D_0|Z = 0]}{E[D_1|Z = 1] - E[D_0|Z = 0]}
\]

\[
= \frac{E[Y_0 + (Y_1 - Y_0)D_1] - E[Y_0 + (Y_1 - Y_0)D_0]}{E[D_1] - E[D_0]}
\]

\[
= \frac{E[Y_1 - Y_0|D_1 > D_0]P(D_1 > D_0) - E[Y_1 - Y_0|D_1 < D_0]P(D_1 < D_0)}{E[D_1] - E[D_0]}
\]

As \((D_1 - D_0) = (1, 0, -1)\)
Identification Assumptions

- **Ignorability of the Instrument:** \((Y_0, Y_1, D_0, D_1) \perp \perp Z\)
  - Implies that \(Z\) is randomly assigned so that the intent to treat effect and first stage effect are causally identified
  - \(Y(d, z)\) implies exclusion restriction so that \(Y(d, 0) = Y(d, 1)\) for \(d = (1, 0)\). Rules out independent effect of \(Z\) on \(Y\)
  - Allows to attribute correlation between \(Z\) and \(Y\) to the effect of \(D\) alone; assumption is not testable
    - Random assignment is not a sufficient condition for exclusion.

- **First Stage:** \(0 < P(Z = 1) < 1\) and \(P(D_1 = 1) \neq P(D_0 = 1)\)
  - Implies that the instrument \(Z\) induces variation in \(D\)
  - Testable by regressing \(D\) on \(Z\)

- **Monotonicity:** \(D_1 \geq D_0\)
  - Rules out defiers
  - Often easy to assess from institutional knowledge
Outline

1. Instrumental Variables with Potential Outcomes (No Covariates)
   - Identification
   - Estimation
   - Examples
   - Size of Complier Group
Estimand (LATE)

\[ E[Y_1 - Y_0|D_1 > D_0] = \frac{E[Y|Z = 1] - E[Y|Z = 0]}{E[D|Z = 1] - E[D|Z = 0]} \left( \frac{\text{cov}(Y, Z)}{\text{cov}(D, Z)} \right) \]

Estimator (Wald Estimator)

The sample analog estimator is:

\[ \left( \frac{\sum_{i=1}^{N} Y_i Z_i}{\sum_{i=1}^{N} Z_i} - \frac{\sum_{i=1}^{N} Y_i (1 - Z_i)}{\sum_{i=1}^{N} (1 - Z_i)} \right) \div \left( \frac{\sum_{i=1}^{N} D_i Z_i}{\sum_{i=1}^{N} Z_i} - \frac{\sum_{i=1}^{N} D_i (1 - Z_i)}{\sum_{i=1}^{N} (1 - Z_i)} \right) \]
Instrumental Variables: Estimators

**Estimand (LATE)**

\[
E[Y_1 - Y_0 | D_1 > D_0] = \frac{E[Y | Z = 1] - E[Y | Z = 0]}{E[D | Z = 1] - E[D | Z = 0]} \left( = \frac{\text{cov}(Y, Z)}{\text{cov}(D, Z)} \right)
\]

**Estimator (Wald Estimator as IV Regression)**

*Can also implement Wald Estimator using an IV regression:*

\[
Y = \mu + \alpha D + \varepsilon
\]

where \( E[\varepsilon | Z] = 0 \), so \( \alpha = \text{cov}(Y, Z)/\text{cov}(D, Z) \)

*To estimate \( \alpha \) we run the simple IV regression of \( Y \) on a constant and \( D \) and instrument \( D \) with \( Z \).*
### Estimand (LATE)

\[
E[Y_1 - Y_0 | D_1 > D_0] = \frac{E[Y|Z = 1] - E[Y|Z = 0]}{E[D|Z = 1] - E[D|Z = 0]} \left( = \frac{\text{cov}(Y,Z)}{\text{cov}(D,Z)} \right)
\]

### Estimator (Two Stage Least Squares)

*If identification assumptions only hold after conditioning on X, covariates are often introduced using 2SLS regression:*

\[
Y = \mu + \alpha D + X'\beta + \epsilon,
\]

where \(E[\epsilon|X,Z] = 0\). Now \(\alpha\) and \(\beta\) are computed regressing \(Y\) on \(D\) and \(X\), and using \(Z\) and \(X\) as instruments.

*In general, \(\alpha\) estimated in this way does not necessarily have a clear causal interpretation (see Abadie (2003))*
Outline

1. Instrumental Variables with Potential Outcomes (No Covariates)
   - Identification
   - Estimation
   - Examples
   - Size of Complier Group
Example: The Vietnam Draft Lottery (Angrist (1990))

- Effect of military service on civilian earnings
- Simple comparison between Vietnam veterans and non-veterans are likely to be a biased measure
- Angrist (1990) used draft-eligibility, determined by the Vietnam era draft lottery, as an instrument for military service in Vietnam
- Draft eligibility is random and affected the probability of enrollment
- Estimate suggest a 15% effect of veteran status on earnings in the period 1981-1984 for white veterans born in 1950-51; although the estimators are quite imprecise
### Wald Estimates for Vietnam Draft Lottery (Angrist (1990))

<table>
<thead>
<tr>
<th>Cohort</th>
<th>Year</th>
<th>Draft-Eligibility Effects in Current $</th>
<th></th>
<th>Service Effect in 1978 $</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Draft-Eligibility</td>
<td>Adjusted FICA Earnings</td>
<td>Total W-2 Earnings</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Earnings (1)</td>
<td>Earnings (2)</td>
</tr>
<tr>
<td>1950</td>
<td>1981</td>
<td>-435.8</td>
<td>-487.8</td>
<td>-589.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(210.5)</td>
<td>(237.6)</td>
<td>(299.4)</td>
</tr>
<tr>
<td></td>
<td>1982</td>
<td>-320.2</td>
<td>-396.1</td>
<td>-305.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(235.8)</td>
<td>(281.7)</td>
<td>(345.4)</td>
</tr>
<tr>
<td></td>
<td>1983</td>
<td>-349.5</td>
<td>-450.1</td>
<td>-512.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(261.6)</td>
<td>(302.0)</td>
<td>(441.2)</td>
</tr>
<tr>
<td></td>
<td>1984</td>
<td>-484.3</td>
<td>-638.7</td>
<td>-1,143.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(286.8)</td>
<td>(336.5)</td>
<td>(492.2)</td>
</tr>
<tr>
<td>1951</td>
<td>1981</td>
<td>-358.3</td>
<td>-428.7</td>
<td>-71.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(203.6)</td>
<td>(224.5)</td>
<td>(423.4)</td>
</tr>
<tr>
<td></td>
<td>1982</td>
<td>-117.3</td>
<td>-278.5</td>
<td>-72.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(229.1)</td>
<td>(264.1)</td>
<td>(372.1)</td>
</tr>
<tr>
<td></td>
<td>1983</td>
<td>-314.0</td>
<td>-452.2</td>
<td>-896.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(253.2)</td>
<td>(289.2)</td>
<td>(426.3)</td>
</tr>
<tr>
<td></td>
<td>1984</td>
<td>-398.4</td>
<td>-573.3</td>
<td>-809.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(279.2)</td>
<td>(331.1)</td>
<td>(380.9)</td>
</tr>
<tr>
<td>1952</td>
<td>1981</td>
<td>-342.8</td>
<td>-392.6</td>
<td>-440.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(206.8)</td>
<td>(228.6)</td>
<td>(265.0)</td>
</tr>
<tr>
<td></td>
<td>1982</td>
<td>-235.1</td>
<td>-255.2</td>
<td>-514.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(232.3)</td>
<td>(264.5)</td>
<td>(296.5)</td>
</tr>
<tr>
<td></td>
<td>1983</td>
<td>-437.7</td>
<td>-500.0</td>
<td>-915.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(257.5)</td>
<td>(294.7)</td>
<td>(395.2)</td>
</tr>
<tr>
<td></td>
<td>1984</td>
<td>-436.0</td>
<td>-560.0</td>
<td>-767.2</td>
</tr>
</tbody>
</table>
Example: Minneapolis Domestic Violence Experiment

Minneapolis Domestic Violence Experiment was first field experiment to examine effectiveness of methods used by police to reduce domestic violence (Sherman and Berk 1984).

**Sample**: 314 cases of male-on-female spousal assault in two high-density precincts, in which both parties present at scene. 51 patrol officers participated in the study.

**Treatments**: Random assignment of cases to one of three approaches:
- Send the abuser away for eight hours
- Advice and mediation of disputes
- Make an arrest

**Outcome**: 6-month follow-up period, with both victims and offenders, as well as official records consulted to determine whether or not re-offending had occurred.
Table 1: Assigned and Delivered Treatments in Spousal Assault Cases

<table>
<thead>
<tr>
<th>Assigned Treatment</th>
<th>Delivered Treatment</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Arrest</td>
<td>Advise</td>
<td>Separate</td>
<td>Total</td>
<td></td>
</tr>
<tr>
<td>Arrest</td>
<td>98.9 (91)</td>
<td>0.0 (0)</td>
<td>1.1 (1)</td>
<td>29.3 (92)</td>
<td></td>
</tr>
<tr>
<td>Advise</td>
<td>17.6 (19)</td>
<td>77.8 (84)</td>
<td>4.6 (5)</td>
<td>34.4 (108)</td>
<td></td>
</tr>
<tr>
<td>Separate</td>
<td>22.8 (26)</td>
<td>4.4 (5)</td>
<td>72.8 (83)</td>
<td>36.3 (114)</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>43.4 (136)</td>
<td>28.3 (89)</td>
<td>28.3 (89)</td>
<td>100.0 (314)</td>
<td></td>
</tr>
</tbody>
</table>

Notes: The table shows statistics from Sherman and Berk (1984), Table 1.
Note that with a single endogenous variable and a single instrument, the causal effect of $D_i$ in the causal model is the ratio of reduced-form to first-stage effects:

$$
\frac{\beta}{\gamma} = 1.
$$

In a randomized trial with imperfect compliance, the reduced-form effect is also the ITT effect. More generally, 2SLS second-stage estimates can be understood as a re-scaling of the reduced form. It can also be shown that the significance levels for the reduced-form and the second-stage are asymptotically the same under the null hypothesis of no treatment effect. Hence, the workingman’s IV motto: *If you can’t see your causal effect in the reduced form, it ain’t there.*

On final reason for looking at the reduced form is that $Y$ in contrast with the 2SLS estimates themselves $Y$ the reduced form estimates have all the attractive statistical properties of any ordinary least squares regression estimates. In particular, estimates of reduced form regression coefficients are unbiased (i.e., centered on the population parameter in repeated samples) and that the statistical theory that justifies statistical inference for these coefficients (i.e., confidence intervals and hypothesis testing) does not require large samples. 2SLS estimates on the other hand, are not unbiased, although they are consistent. This means that in large samples, the sample estimates can be expected to be close to the target population parameter. Moreover, the statistical theory that justifies confidence intervals and hypothesis testing for 2SLS requires that samples be large enough for a reasonably good asymptotic approximation (in particular, for application of central limit theorems).

How large a sample is large enough for asymptotic statistical theory to work? Unfortunately, there is no general answer to this question. Various theoretical arguments and simulations studies have shown, however, that the asymptotic approximations used for 2SLS inference are usually reasonably accurate in models where the number of instruments is equal to (or not much more than) the number of endogenous variables (as would be the case in studies using randomly assigned intention to treat as an instrument for treatment delivered). Also, that the key to $t^2.1$

Table 2. First stage and reduced forms for Model 1.

<table>
<thead>
<tr>
<th>Endogenous variable is coddled</th>
<th>First stage</th>
<th>Reduced form (ITT)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)*</td>
</tr>
<tr>
<td>Coddled-assigned</td>
<td>0.786 (0.043)</td>
<td>0.773 (0.043)</td>
</tr>
<tr>
<td>Weapon</td>
<td>−0.064 (0.045)</td>
<td>−0.004 (0.042)</td>
</tr>
<tr>
<td>Chem. influence</td>
<td>−0.088 (0.040)</td>
<td></td>
</tr>
<tr>
<td>Dep. var. mean</td>
<td>0.567</td>
<td>0.178</td>
</tr>
<tr>
<td>(Coddled–delivered)</td>
<td></td>
<td>(V Failed)</td>
</tr>
</tbody>
</table>

The table reports OLS estimates of the first-stage and reduced form for Model 1 in the text. *Other covariates include year and quarter dummies, and dummies for non-white and mixed race.*
Instrumental Variables with Potential Outcomes (No Covariates) Examples

Treatment Effect in Minneapolis Experiment

Table 3. OLS and 2SLS estimates for Model 1.

<table>
<thead>
<tr>
<th>Endogenous variable is coddled</th>
<th>OLS</th>
<th>IV/2SLS</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td>(2)*</td>
<td>(3)</td>
</tr>
<tr>
<td>Coddled–delivered</td>
<td>0.087 (0.044)</td>
<td>0.145 (0.060)</td>
</tr>
<tr>
<td>Weapon</td>
<td>0.010 (0.043)</td>
<td>0.005 (0.043)</td>
</tr>
<tr>
<td>Chem. influence</td>
<td>0.057 (0.039)</td>
<td>0.064 (0.039)</td>
</tr>
</tbody>
</table>

The Table reports OLS and 2SLS estimates of the structural equation in Model 1.
*Other covariates include year and quarter dummies, and dummies for non-white and mixed race.
Outline

1. Instrumental Variables with Potential Outcomes (No Covariates)
   - Identification
   - Estimation
   - Examples
   - Size of Complier Group
Since we never observe both potential treatment assignments for the same unit, we cannot identify individual units as compliers. However, we can easily identify the proportion of compliers in the population using the first stage effect:

\[
P(D_1 > D_0) = E[D_1 - D_0] = E[D_1] - E[D_0] = E[D|Z = 1] - E[D|Z = 0]
\]

Using a similar logic we can identify the proportion of compliers among the treated or controls only. For example:

\[
P(D_1 > D_0|D = 1) = \frac{P(Z = 1)(E[D|Z = 1] - E[D|Z = 0])}{P(D = 1)}
\]
### Table 4.4.2
Probabilities of compliance in instrumental variables studies

| Source                         | Endogenous Variable (d) | Instrument (z) | Sample (4)                          | P[d = 1] (5) | First Stage, P[d_1 > d_0] (6) | P[z = 1] (7) | P[d_1 > d_0 | d = 0] (8) | P[d_1 > d_0 | d = 0] (9) |
|-------------------------------|-------------------------|----------------|-------------------------------------|--------------|-----------------------------|--------------|--------------|-------------|--------------|
|                               |                         |                | Non-white men born in 1950          | .163         | .060                        | .534         | .197         | .033        |
| Angrist and Krueger (1991)    | High school graduate    | First two children are same sex | Men born between 1930 and 1939 | .770         | .016                        | .509         | .011         | .034        |
| Acemoglu and Angrist (2000)   | High school graduate    | Third- or fourth-quarter birth | White men aged 40-49               | .617         | .037                        | .300         | .018         | .068        |

**Notes:** The table computes the absolute and relative size of the complier population for a number of instrumental variables. The first stage, reported in column 6, gives the absolute size of the complier group. Columns 8 and 9 show the size of the complier population relative to the treated and untreated populations.
Precision for LATE Estimation

- When $N$ is large the standard error on the instrumental variable estimator of the LATE is approximately

\[ SE_{\text{LATE}} \approx \frac{SE_{\text{ITT}}}{\text{Compliance Ratio}} \]

- In JTPA data we get $330/\cdot.62 = 532$ which is close to the standard error estimate from the instrumental variable regression of 526.

- Two estimates converge if there is perfect compliance.

- Otherwise, all else equal, the standard error on the LATE decreases linearly with the compliance!
  - If compliance ratio drops from 100% to 10%, the LATE standard error increases by a factor of 10

- Always wise to conduct a pilot to test the encouragement.

- Design it to boost compliance, but do not violate exclusion restriction.