Session V
Instrumental variables

Evaluating public policies

Arthur Heim (PSE & Cnaf)

March 27, 2023*

Outline

Introduction

What's an instrumental variable

- We consider settings with unobservables that affect treatment and outcomes.
- Econometricians have come a long way to define settings and hypotheses that get you close to a randomized experiment. That's why these methods are sometimes refered to as "quasi-experiments" or "natural experiments"
- This lecture is about **instrumental variables** (IV), one of the most popular method used in empirical work to estimate the impact of a policy.
- An instrument is a variable that affect treatment but has no direct effect on outcomes.
- When we have such variables, we can use the variation of the treatment that is caused by the instrument to identify the effect of the treatment on the outcomes.

What's an instrumental variable

Consider a typical "population regression"

$$Y_i = \alpha + \beta D_i + \boldsymbol{X'}\gamma + \mu_i \tag{1}$$

- To estimate this model with OLS on a random sample, a critical assumption is that $cov(\mu_i, D_i|\boldsymbol{X}) = 0$ i.e. **exogeneity**
- This assumption is likely to be violated if:
  - **Unobserved heterogeneity**: we may not observe all confounding variables
  - D_i may be **measured with error**
  - **Simultaneity** or **reverse causality**

What's an instrumental variable

What do IV do ?

In theory, instrumental variables offer a way to

- break the correlation $cov(D_i, \mu_i)$
- and obtain a consistent causal estimate of the treatment on Y_i

Instrumental variable is based on two conditions:

1. **First stage:** the instrument predicts treatment well: $cov(Z, D) \neq 0$
2. **Exogeneity:** Instrument Z is exogenous, unrelated to the structural error $cov(Z, \mu) = 0$
3. **exclusion restriction:** The instrument has an effect on the outcome only through the treatment variable.

What's an instrumental variable

The first-stage relationship is testable

- We can run a regression of D on Z
- it is also possible to include covariates

The exclusion restriction is not testable

- It is an **identification assumption**
- We need to make a convincing argument in favor of it
- This is difficult and the reason for heated debates in seminars

Some say: friends tell their friends not to use IV...

What's an instrumental variable

Intuition behind instrumental variables

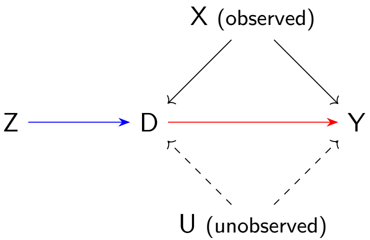
- People/firms make optimal choices that affect their treatment status
- Z is a shock that changes the behavior of at least some people/firms
- Z has to be unrelated to people's characteristics
- i.e. it should be assigned as good as randomly

Instrumental variable and experiments

- The instrument Z is a treatment/incentive that is offered to units/people
- D measures if the unit actually takes up the treatment
- Instrumental variable work to deal with imperfect compliance in experiment
- instrument Z should be as good as randomly assigned

What's an instrumental variable

Figure 1: Instrumental variable DAG



Other interpretation of instrumental variables

- Regressing Y on D and X with OLS uses all the variation in D to explain Y, including the one caused by U. Hence the bias.
- IV uses only the variation in D that is related to Z
- So this means less variation is used, but at least Z is unrelated to U

Outline

1. Introduction

2. The theory of instrumental variables
   - Instrumental variable equations: Two stage least square (2SLS)
   - Conceptual example
   - When the instrument and treatment are binary
   - General model in matrix notation
   - Variance of OLS and 2SLS
   - Instrumental variables' troublesomenesses
   - Weak instruments: inconsistency

3. Illustration: Angrist and Evans (1998) on child penalty

Instrumental variable equations: Two stage least square (2SLS)

- We consider a linear representation of the relationship between outcome Y, treatment D and potential covariates $\boldsymbol{X}$:

$$Y_i = \alpha + \beta D_i + \boldsymbol{X_i'}\gamma + \mu_i \tag{2}$$

- This is the structural equation and writing it like that is already an assumption
- We call the **"First stage"** the relationship between treatment D and the instrument Z (and potential covariates) in a linear equation:

$$D_i = \delta_0 + \delta_1 Z_i + \boldsymbol{X_i'}\rho + \epsilon_i \tag{3}$$

- We call the **"second stage"** the regression of the outcome on $\boldsymbol{X}$ and the prediction of the first stage $\hat{D}_i$:

$$Y_i = \tilde{\alpha} + \tilde{\beta}\hat{D}_i + \boldsymbol{X_i'}\kappa + \varepsilon_i \tag{4}$$

- We call the **"reduced form"** the regression of the outcome on the instrument Z and potential covariates $\boldsymbol{X}$

$$Y_i = \lambda_0 + \lambda_1 Z_i + \boldsymbol{X_i'}\tau + \upsilon_i \tag{5}$$

The theory of instrumental variables

Instrumental variable equations: Two stage least square (2SLS)

- It can be shown that the following estimator is a consistent estimator of β under independence and exclusion $cov(Z, \mu_i) = 0$

$$\hat{\beta_{IV}} = \frac{cov(Y, Z)}{cov(D, Z)} = \frac{\hat{\lambda_1}}{\hat{\delta_1}} \tag{6}$$

- This estimator is nothing but the **reduced form coefficient** $\hat{\lambda_1} = \frac{cov(Y,Z)}{var(Z)}$ **divided by the first stage coefficient** $\hat{\delta_1} = \frac{cov(D,Z)}{var(Z)}$
- Two stage least square (2SLS) estimate a first and second stage that estimate the structural relationship and retrieve a consistent estimate of β

The theory of instrumental variables

Conceptual example

Let's consider the relationship between income Y and education S where we assume there's ability bias such that we can write this linear structural relationship: $Y_i = \alpha + \delta S_i + \gamma A_i + \mu_i$

- When A_i cannot be observed, when we estimate $\hat{Y}_i = \hat{\alpha} + \hat{\delta} S_i + \mu_i$ using OLS we get biased estimates of the coefficient of interest δ
- (In lecture II we showed that $\hat{\delta} = \delta + \gamma \frac{Cov(A_i, S_i)}{\sigma_S^2}$)
- Suppose that we have a valid instrument Z that predict schooling S well
- Z is a valid instrument if $Z_i \perp A_i, \mu_i$

The theory of instrumental variables

Conceptual example

- We write the **first stage** equation : $S_i = \eta + \pi Z_i + \epsilon_i$
- Coefficient π corresponds to the share of variation in S caused by Z.
- We can also write the **reduced form**: $Y_i = c + \beta Z_i + \xi_i$
- Often, the reduced form is interesting per se (for instance, instrument can be eligibility for a policy and be interepreted as the "intention to treat").
- Let's replace S_i in the structural equation by the first stage:

$$Y_i = \alpha + \delta(\eta + \pi Z_i + \epsilon_i) + \gamma A_i + \mu_i$$

$$\Leftrightarrow Y_i = \underbrace{\alpha + \delta\eta}_{c} + \underbrace{\delta\pi}_{\beta} Z_i + \underbrace{\delta\epsilon_i + \gamma A_i + \mu_i}_{\xi}$$

The theory of instrumental variables

Conceptual example

$$Y_i = \underbrace{\alpha + \delta\eta}_{c} + \underbrace{\delta\pi}_{\beta} Z_i + \underbrace{\delta\epsilon_i + \gamma A_i + \mu_i}_{\xi}$$

- The structural parameter δ is identified by dividing the reduced form coefficient by the first stage coefficient:

$$\delta = \frac{\beta}{\pi} = \frac{Cov(Y_i, Z_i)/\sigma^2_{Z_i}}{Cov(S_i, Z_i)/\sigma^2_{Z_i}} = \frac{Cov(Y_i, Z_i)}{Cov(S_i, Z_i)}$$

The theory of instrumental variables

Conceptual example

- It's a two step procedure
  1. Estimate first stage and predict $\hat{S}_i$
  2. Estimate structural model replacing S by the first stage prediction $\hat{S}_i$ = **Second stage**
- That's what two stage least square do.
- Don't do these steps separately, if you do, the standard errors of your second stage will be wrong for they do not take into account the fact that the regressor is a prediction from the first stage. These are **forbidden regressions**.
- Implement 2SLS directly using the appropriate commands. in R the *estimater* package has the *iv_robust* function that does that well, or *fixest*'s function *feols*.

The theory of instrumental variables

When the instrument and treatment are binary

- When Z_i is a dummy (0,1) and D_i denotes treatment participation
- We can write the standard regression equation for treatment effect: $Y_i = \alpha + \delta D_i + \varepsilon_i$
- Let's rewrite it in expectation:

$$\mathbb{E}[Y_i|Z_i = 1] = \underbrace{\mathbb{E}[\alpha|Z_i = 1]}_{=\alpha} + \mathbb{E}[D_i|Z_i = 1]\delta + \mathbb{E}[\varepsilon|Z_i = 1]$$

$$\mathbb{E}[Y_i|Z_i = 0] = \underbrace{\mathbb{E}[\alpha|Z_i = 0]}_{=\alpha} + \mathbb{E}[D_i|Z_i = 0]\delta + \mathbb{E}[\varepsilon|Z_i = 0]$$

If Z is exogenous then $\mathbb{E}[\varepsilon|Z_i = 1] = \mathbb{E}[\varepsilon|Z_i = 0] = 0$

$$\Rightarrow \delta = \frac{\mathbb{E}[Y_i|Z_i = 1] - \mathbb{E}[Y_i|Z_i = 0]}{\mathbb{E}[D_i|Z_i = 1] - \mathbb{E}[D_i|Z_i = 0]} = \frac{\text{reduced form}}{\text{1st stage}} = \frac{ITT}{\text{take-up}}$$

- This estimator is called the Wald estimator.

The theory of instrumental variables

General model in matrix notation

- Reminder: The OLS estimator in matrix form is:

$$\hat{\beta}_{OLS} = (\boldsymbol{X'X})^{-1}\boldsymbol{X'Y}$$

- The First stage is a regression of X on Z so it would be in matrix form:

$$\hat{\beta}_{FS} = (\boldsymbol{Z'Z})^{-1}\boldsymbol{Z'X}$$

- $\boldsymbol{P_Z} = \boldsymbol{Z}(\boldsymbol{Z'Z})^{-1}\boldsymbol{Z'}$ is the **projection matrix** of the first stage.
- The 2SLS estimator includes a projection of X on Z:

$$\begin{aligned}\hat{\beta}_{2SLS} =& (\boldsymbol{X'Z}(\boldsymbol{Z'Z})^{-1}\boldsymbol{Z'X})^{-1}\boldsymbol{X'Z}(\boldsymbol{Z'Z})^{-1}\boldsymbol{Z'Y} \\ \hat{\beta}_{2SLS} =& (\boldsymbol{X'P_ZX})^{-}1\boldsymbol{X'P_ZY}\end{aligned} \tag{7}$$

The theory of instrumental variables

Variance of OLS and 2SLS

- Assuming homoskedasticity, the asymptotic variance of the OLS estimator with K regressors in $\boldsymbol{X}$ is:

$$\widehat{Avar}(\hat{\beta}_{OLS}) = \hat{\sigma}^2(\boldsymbol{X'X})^{-1}$$

with $\hat{\sigma}^2 = \frac{\hat{u}'\hat{u}}{N-K}$

- The 2SLS equivalent is:

$$\begin{aligned} \widehat{Avar}(\hat{\beta}_{2SLS}) &= \hat{\sigma}^2(\boldsymbol{\hat{X}'\hat{X}})^{-1} \\ &= \hat{\sigma}^2(\boldsymbol{X'Z}(\boldsymbol{Z'Z})^{-1}\boldsymbol{Z'X})^{-1} \end{aligned} \tag{8}$$

- A weak first stage correlation will **increase the variance of the estimator** and thus standard errors and reduce precision.

The theory of instrumental variables

Instrumental variables' troublesomenesses

1. Instruments may have have small effects: **weak instruments**
   - Larger variance
   - Inconsistency
2. **Exclusion restriction**: cannot be tested
3. Small sample biais
4. Heterogenous treatment effects
5. Modeling issues with 2SLS.
6. Interpretations with multiple instruments.

The theory of instrumental variables

Weak instruments: inconsistency

- Consider the simultaneous equations model

$$y_i = \alpha + \beta x_i + \varepsilon_i$$
$$x_i = \mu + \pi z_i + v_i.$$

- The OLS and IV estimators are given by

$$\widehat{\beta}_{OLS} = \frac{\operatorname{cov}(y_i, x_i)}{\operatorname{var}(x_i)}$$
$$\widehat{\beta}_{2SLS} = \frac{\operatorname{cov}(y_i, \widehat{x}_i)}{\operatorname{var}(\widehat{x}_i)}$$

- and the plims of the estimators are

$$\operatorname{plim} \widehat{\beta}_{OLS} = \beta + \frac{\sigma_{x\varepsilon}}{\sigma_x^2}$$
$$\operatorname{plim} \widehat{\beta}_{2SLS} = \beta + \frac{\sigma_{\widehat{x}\varepsilon}}{\sigma_{\widehat{x}}^2}.$$

The theory of instrumental variables

Weak instruments: inconsistency

- This yields

$$\frac{\text{plim}\,\widehat{\beta}_{2SLS} - \beta}{\text{plim}\,\widehat{\beta}_{OLS} - \beta} = \frac{\sigma_{\widehat{x}\varepsilon}/\sigma_{x\varepsilon}}{\sigma^2_{\widehat{x}}/\sigma^2_{x}} = \frac{\sigma_{\widehat{x}\varepsilon}/\sigma_{x\varepsilon}}{R^2_{xz}}$$

- The inconsistency of the 2SLS estimator relative to the OLS estimator is related to the **relative endogeneity** of z and x.
- Notice that R^2_{xz}, is the R^2 from the first stage regression.
  - The instrument z may be almost as good as randomly assigned but not quite.
  - Hence, $\sigma_{\widehat{x}\varepsilon}$ may be small but not quite zero.
- However, even if $\sigma_{\widehat{x}\varepsilon}$ is small compared to $\sigma_{x\varepsilon}$, the **relative inconsistency of the 2SLS estimator** may still be important as long as R^2_{xz} is also small, i.e. as long as the **correlation of z and x is low**.

The theory of instrumental variables

Weak instruments: what should we do ?

- We would test the joint significance of your instruments' coefficients via an F-test.
- In case you only have one instrument, this F-statistic is equivalent to the square of the t-statistic of your instrument's coefficient in the first stage.
- Old Rule of thumb: F-stat > 10 (Stock, Wright, and Yogo 2002) prove to be too low (Keane and Neal 2022)
- If you consider using more than one instrument (Avoid if you can - see sections after), show first the **results of the just-identified model**, using your best instrument.
- Monte Carlo simulations show that just-identified 2SLS is approximately unbiased
- But just-identified estimates are also unstable and imprecise
- **Show the F-Statistic of the first stage**
- Keane and Neal (2021) wrote a very clear paper on what to do with weak instruments

  > *A first-stage F well above 10 is necessary to give high confidence that 2SLS will outperform OLS. Otherwise, OLS combined with controls for sources of endogeneity may be a superior research strategy to IV.*

Outline

Illustration: Angrist and Evans (1998) on child penalty

Context and research question

- In the US, the number of children per women decreased while at the same time women participation in the labor market increased.
- Women who have children are less likely to work, and when they do, their wage is lower than women who don't have children, and even lower than men, father or not.

⋆ Is it because having a child cause women to work less ? or is it some women with specific latent caracteristics that choose to have children and the differences reflect this selection process (or both, and so how much effect is selection, how much is causal effet ?)

- Ideas on how to solve this problem ?

Illustration: Angrist and Evans (1998) on child penalty

Context and research question

- In the US, the number of children per women decreased while at the same time women participation in the labor market increased.
- Women who have children are less likely to work, and when they do, their wage is lower than women who don't have children, and even lower than men, father or not.
- ⋆ Is it because having a child cause women to work less ? or is it some women with specific latent caracteristics that choose to have children and the differences reflect this selection process (or both, and so how much effect is selection, how much is causal effet ?)
- Ideas on how to solve this problem ?

Identification strategy: gender preferences

- Angrist and Evans (1998) notice that parents with 2 children are more likely to have a third child when their two first children have the same gender.

Context and research question

How they construct instruments

- **Intuition**: Parents have a preference for gender diversity among their children and are more likely to have a third child when their two first children have the same gender.
- **it can be seen in the data** and could be used as an instrument.

Is it a good instrument ?

Context and research question

How they construct instruments

- **Intuition**: Parents have a preference for gender diversity among their children and are more likely to have a third child when their two first children have the same gender.
- **it can be seen in the data** and could be used as an instrument.

Is it a good instrument ?

- it predicts the probability of having a third child $\Rightarrow$ First stage $\surd$
- children's gender is not correlated with labour market outcomes $\Rightarrow$ Exogeneity $\surd$
- Children's gender has no effect on outcomes but through the effect on the probability of having a third child $\Rightarrow$ Exclusion $\surd$

Context and research question

Data and estimation

- Standard linear equation : $Y_i = \alpha + \beta C_i + \boldsymbol{X_i'}\boldsymbol{\gamma} + \varepsilon_i$
- Y can be income, employment... C is a dummy for having 3 children or more ;
- instrument Z equals 1 when the two first children have the same gender, 0 otherwise.
- They use census data from 1980 and 1990 in the US and restrict the sample to women aged 21-35, maried or not with at least two children
- In these data, 27 % of women aged 21 to 27 have at least two children, ad 50 % of women aged 27 to 35.
- They remove older women because it's likely that their children are old enough to leave and wouldn't be accounted in the data.

Illustration: Angrist and Evans (1998) on child penalty

Replication

- Data from the 1980 census used by Angrist and Evans (1998) are available in the package

```
library(ivmte)
AngristEvans <- ivmte::AE
```

- Let's:
  1. Run a linear regression of $worked$ on $morekids$ and controls if we want
  2. See whether $samesex$ has an effect on $morekids$
  3. Estimate the reduced form
  4. Run the forbidden regression
  5. Run 2SLS regression and estimate the effect of having a third child on women participation to the labour market

Illustration: Angrist and Evans (1998) on child penalty

Replication

```r
# Run a linear regression of $worked$ on $morekids$ and controls if we
# want
naive <- lm_robust(worked ~ morekids, data = AngristEvans)
naivecov <- lm_robust(worked ~ morekids + black + hisp, data = AngristEvans)
# See whether $samesex$ has an effect on $morekids$
Firststage <- lm_robust(morekids ~ samesex, data = AngristEvans)
# Estimate the reduced form
reduced <- lm_robust(worked ~ samesex, data = AngristEvans)
# Run the forbidden regression
AngristEvans$predicted_child <- Firststage$fitted.values
forbidden <- lm_robust(worked ~ predicted_child, data = AngristEvans)
# Run 2SLS regression and estimate the effect of having a third child on
# women participation to the labour market
IV <- iv_robust(worked ~ morekids | samesex, data = AngristEvans)
IVcov <- iv_robust(worked ~ morekids + black + hisp | samesex + black + hisp,
    data = AngristEvans)
```

Illustration: Angrist and Evans (1998) on child penalty

With fixest:

```
# Run a linear regression of $worked$ on $morekids$ and controls if we
# want
library(fixest)
naive <- feols(worked ~ morekids, data = AngristEvans, vcov = "hetero")
naivecov <- feols(worked ~ morekids + black + hisp, data = AngristEvans, vcov =
# See whether $samesex$ has an effect on $morekids$
Firststage <- feols(morekids ~ samesex, data = AngristEvans, vcov = "hetero")
# Estimate the reduced form
reduced <- feols(worked ~ samesex, data = AngristEvans, vcov = "hetero")
# Run the forbidden regression
AngristEvans$predicted_child <- Firststage$fitted.values
forbidden <- feols(worked ~ predicted_child, data = AngristEvans, vcov = "heter
# Run 2SLS regression and estimate the effect of having a third child on
# women participation to the labour market
IV <- feols(worked ~ 1 | morekids ~ samesex, data = AngristEvans, vcov = "heter
IVcov <- feols(worked ~ black + hisp | morekids ~ samesex, data = AngristEvans,
    vcov = "hetero")
```

Illustration: Angrist and Evans (1998) on child penalty

Replication

	OLS	Covariates	First stage	Reduced form	Forbidden regression	2SLS	2SLS cov
(Intercept)	0.582***	0.572***	0.302***	0.538***	0.563***	0.563***	0.552***
	(0.001)	(0.001)	(0.001)	(0.002)	(0.012)	(0.012)	(0.012)
morekids	−0.142***	−0.144***					
	(0.002)	(0.002)					
black		0.172***					0.169***
		(0.004)					(0.004)
hisp		−0.014**					−0.023***
		(0.007)					(0.009)
samesex			0.059***	−0.005**			
			(0.002)	(0.002)			
predicted_child					−0.085**		
					(0.037)		
fit_morekids						−0.085**	−0.083**
						(0.037)	(0.037)
Num.Obs.	209 133	209 133	209 133	209 133	209 133	209 133	209 133
R2	0.018	0.026	0.004	0.000	0.000	0.015	0.022
R2 Adj.	0.018	0.026	0.004	0.000	0.000	0.015	0.022
RMSE	0.49	0.49	0.47	0.50	0.50	0.49	0.49
Std.Errors	Heteroskedasticity-robust	Heteroskedasticity-robust	Heteroskedasticity-robust	Heteroskedasticity-robust	Heteroskedasticity-robust	Heteroskedasticity-robust	Heteroskedasticity-robust

Illustration: Angrist and Evans (1998) on child penalty

Results: interpretations

- Women with 3 children are -14.23 percentage points less likely to work than women with two children
- Women with same sex children are 5.89 percentage points more likely to have a third child than those with children of both gender.
- From the reduced form we see that having same sex children reduces the probability to work by -0.5 percentage point.
- If an increase in the number of children by 6 percentage points reduces the probability to work by -0.5 percentage points...
- Then an **increase by one child** reduces the probability to work by $\frac{-0.5}{6} = -8.5$
- For this reason, we often say that we "scale up" the reduced form by the first stage

Results: interpretations

So we have that $\hat{\beta_{IV}} < \hat{\beta_{OLS}}$. Does this make sense?

- **Explanation 1**: OLS estimator is upward biased (i.e. closer to zero)
  - there could be an omitted variable (for example family wealth)
  - both the correlation with kids and the direct effect on hours need to have the same sign
  - e.g. $cov(wealth; kids) > 0$ and $cov(wealth; hours jkids) > 0$ or both negative
- **Explanation 2:** IV effect measures the effect for a specific population
  - only 1 in 14 families "respond" to the instrument
  - families who respond may not be the average family...

Outline

Local average treatment effect

What do you think ?

Suppose we have two valid instruments for the same public policy

Local average treatment effect

What do you think ?

Suppose we have two valid instruments for the same public policy

- Would the IV estimator be the same with both instruments ?

Local average treatment effect

What do you think ?

Suppose we have two valid instruments for the same public policy

- Would the IV estimator be the same with both instruments ?
- Under what condition(s) will the estimates be equal?

Local average treatment effect

What do you think ?

Suppose we have two valid instruments for the same public policy

- Would the IV estimator be the same with both instruments ?
- Under what condition(s) will the estimates be equal?

Answer:

- If the treatment effect is **constant** for everyone
- If those who react to one instrument are the same as those who react to the other instrument
- If treatment effect is heterogenous, we need additional assumption to retrieve causal parameters.

Local average treatment effect

Implicit homogenous treatment effect assumption so far

- You probably noticed that contrary to other sessions we didn't use potential outcomes notation so far.
- We derived IV estimator from regression notations and actually implicitely assumed that the treatment effect we estimate is constant.

Allowing heterogneous treatment effect

- With heterogenous treatment effect, the IV assumptions are not sufficient to retrieve causal effects. Fomral proof
- We need to change notation and add an additional assumption to use instrumental variables
- Consider a case with a binary instrument $Z_i \in \{0, 1\}$ the the treatment statuses
  - D_{1i} = i's treatment status when $Z_i = 1$
  - D_{0i} = i's treatment status when $Z_i = 0$
- The **observed treatment status** is

$$D_i = D_{0i} + (D_{1i} - D_{0i}) Z_i = \pi_0 + \pi_{1i} Z_i + \zeta_i$$

Allowing heterogneous treatment effect

First stage

- The previous notation gives us the first stage

$$\begin{aligned} D_i &= D_{0i} + (D_{1i} - D_{0i})Z_i \\ &= \pi_0 + \pi_{1i} Z_i + \zeta_i \end{aligned}$$

- $\mathbb{E}[D_{0i}] = \pi_0$
- $(D_{1i} - D_{0i}) = \pi_{1i}$ is the individual effect of the instrument on treatment participation
- A first assumption for IV to work is to have a first stage i.e. $\mathbb{E}[D_{1i} - D_{0i}] \neq 0$

Monotonicity

- "If the instrument has no effect on some individuals, all those affected are in the same direction"
- $\pi_{1i} \geq 0 \ \mid \ \pi_{1i} \leq 0 \ \forall \ i$

Allowing heterogneous treatment effect

An unkwnow partition of the population

We can divide the population into four groups depending on their reaction to the instrument:

Table 1: Four sub-populations

Types	D_{0i}	D_{1i}
Never takers	0	0
Defiers	1	0
Compliers	0	1
Always takers	1	1

- Monotonicity implies that there is no *defiers* in the population
- From any dataset, it is impossible to see who belongs to what group

Allowing heterogneous treatment effect

Independence

- The instrument is independent of potential outcomes and potential treatment status: $\leftrightarrow$ Z_i assigned *as good as random*.

$$\Big(Y_i(D_1,1), Y_i(D_0,0), D_{1i}, D_{0i}\Big) \perp Z_i$$

- Independence is enough for a causal interpretation of the reduced form

$$\begin{aligned}\mathbb{E}[Y_i|Z_i=1]-\mathbb{E}[Y_i|Z_i=0] &= \mathbb{E}[Y_i(D_{1i},1)|Z_i=1]-\mathbb{E}[Y_i(D_{0i},0)|Z_i=1] \\ &= \mathbb{E}[Y_i(D_{1i},1)]-\mathbb{E}[Y_i(D_{0i},0)]\end{aligned}$$

- Independence also means the *first stage* captures the causal effect of Z on treatment D.

$$\mathbb{E}[D_i|Z_i=1]-\mathbb{E}[D_i|Z_i=1]=\mathbb{E}[D_{1i}-D_{0i}]$$

Allowing heterogneous treatment effect

Exclusion

- Exclusion means there is only **one causal path** between the instrument and the outcome, and that is through the effect on treatment status.
- We **assume** the variation in the ITT solely comes from the effect of compliers
- Fromally, it meanse that potential outcome $Y_i(d, z)$ is only a function of d

$$Y_i(d, 1) = Y_i(d, 0) \equiv Y_{di} \text{ pour } d = 0, 1.$$

- So we can rewrite observed outcomes as a function of these potential outcomes

$$\begin{aligned} Y_i &= Y_(0, Z_i) + \big(Y_i(1, Z_i) - Y_i(0, Z_i)\big) D_i \\ &= Y_{0i} + (Y_{1i} - Y_{0i}) D_i \\ &= \alpha + \rho_i D_i + \varepsilon_i \end{aligned}$$

- With $E(Y_{0i}) = \alpha$
- $(Y_{1i} - Y_{0i}) = \rho_i$ is the individual treatment effect

The LATE theorem (Imbens and Angrist 1994; Angrist, Imbens, and Rubin 1996)

What we are looking for

If the following conditions are satisfied

- **Independence** $\left(Y_i(D_1,1), Y_i(D_0,0), D_{1i}, D_{0i}\right) \perp Z_i$
- **exclusion** $Y_i(d,1) = Y_i(d,0) \equiv Y_{di}$ pour $d = 0, 1.$
- **first-stage** $\mathbb{E}[D_{1i} - D_{0i}] \neq 0$
- **monotonicity** $D_{1i} - D_{0i} \geq 0 \ \mid \ D_{1i} - D_{0i} \leq 0 \ \forall \ i$

Then the WALD ratio estimates the **local average treatment effect**:

$$\begin{aligned}
\frac{\mathbb{E}[Y_i|Z_i=1] - \mathbb{E}[Y_i|Z_i=0]}{\mathbb{E}[D_i|Z_i=1] - \mathbb{E}[D_i|Z_i=0]} &= \mathbb{E}[Y_{1i} - Y_{0i}|D_{1i} > D_{0i}] \\
&= \mathbb{E}[\rho_i|\pi_{1i} > 0]
\end{aligned}$$

The LATE theorem (Imbens and Angrist 1994; Angrist, Imbens, and Rubin 1996)

What is this local average treatment effect?

- Under these 4 hypotheses, the Wald ratio or 2SLS coefficient are consistent estimates of the LATE, the **average causal effect on the compliers**, and them only.
- Intuitively this makes sense because compliers are the only group on which the data can be informative:
  - Compliers are the only group with units observed in both treatments (given that defiers have been ruled out).
  - Always takers and never-takers are observed only in one treatment.
- The LATE is a controversial parameter,
  - It is defined for an unobservable sub-population
  - It is instrument dependent
- Therefore, it is no longer clear which interesting policy question it can answer.

Local average treatment effect

Better LATE than nothing ?

- Under A1-A4, IV ensures internal validity of the LATE
- But LATE has no (or little) external validity. Without further assumptions
  - we cannot generalize to the population
  - we cannot generalize to different contexts
- Despite these shortcomings, LATE is often the best we can do
- Similar estimates from different contexts increase external validity
- There are many relevant positive and normative questions for which
- the LATE seems to be an interesting parameter in addition to being
- the only one we can identify without making unreasonable assumptions (Imbens 2010)

Local average treatment effect

Instrumental variables in experiments

- In experiments, it is often the case that compliance is imperfect and comparison between treated and untreated units are biased by endogenous participation
- With random assignment, it's always possible to estimate the intention to treat (ITT) which is the reduced form
- The Wald estimand will give us the average treatment effect on compliers
- Special case: when non-compliance is one sided (non compliers only in the assigned-to-treatment group), then the instrumental variable estimator retrieves the Average treatment effect on the treated (Frölich and Melly 2013).

IV with covariates

The IV assumptions are different

- First stage: $cov(Z, D|X) \neq 0$ has to be sufficiently strong
- Exclusion restriction: $cov(Z, u|X) = 0$ has to hold conditional on X
- Monotonicity also has to hold conditional on X

Functional form is important !

- Saturated 2SLS in both 1st and second stage is **necessary** for 2SLS to retrieve the LATE with covariates (Blandhol et al. 2022)
- However, researchers almost never do that. The LATE interpretation of 2LS is rarely well formulated.
- Other methods may perform better.

IV with multiple instruments

It is possible to have multiple instrumental variables

- 2SLS combines instruments to get a single (more precise) estimate

Assumptions needed

- Each instrument must be as good as randomly assigned
- Each instrument needs to satisfy the exclusion restriction
- The joint first stage has to be strong enough

A model is **just identified** if there are as many instruments as regressors, and is overidentified when there is more instruments than regressors.

Local average treatment effect

Continuous instruments

- It is also possible to obtain an IV estimate with a continuous instrument and/or treatment
- The assumptions (first stage, exclusion restriction, monotonicity) remain the same
- The LATE is more difficult to interpret
- units differ in their compliance intensity
- i.e. some react to the instrument more than others
- LATE is the weighted average of unit causal effects over the support of D
- weights are determined by the share of compliers in each bin of D
- Often useful to use the binary case as reference (high/low intensity of treatment and instrument)

Outline

Conclusions on instrumental variables

Where do we find good instruments ?

- Theory combined with clever datas collection. Examples
  - Distance from job training centers
  - College openings
- Variation in policies. This requires a deep institutional knowledge. Examples
  - assignment to judges with different severity
  - differences in budgets across job training centers
  - school lotteries and other algorithm-based assignments
  - ...
- Nature. Examples
  - different levels of pollution in different places
  - sex of the first two children
  - ...

Cookbook for IV

Explain your identification clearly

- start with the ideal experiment; why is your setting different? Why is your regressor endogenous?
- Explain theoretically why there should be a first stage and what coefficient we should expect
- Explain why the instrument is as good as randomly assigned
- Explain theoretically why the exclusion restriction holds in your setting

Show and discuss the first stage

- Best to start with a raw correlation
- Do the sign and magnitude make sense?
- Assess the strength of the instrument using state-of-the-art

Cookbook for IV

Bring supportive evidence for instrument validity

- Show that the instrument does not predict pre-treatment characteristics
- Can you provide evidence in support of the exclusion restriction?
- Compare complier characteristics with never-takers'.

Bring supportive evidence for instrument validity

- Show the OLS and 2SLS results, both with varying sets of controls
- Comment on the differences between both (bias, LATE, etc)
- Show the reduced form
- "If the reduced form isn't there, the effect isn't there" (Angrist and Pischke 2008)

Bibliography I

- Angrist, Joshua, and William Evans. 1998. "Children and Their Parents' Labor Supply: Evidence from Exogenous Variation in Family Size." *The American Economic Review , Jun., 1998, Vol. 88, No. 3* (June): w5778.
- Angrist, Joshua D, Guido W Imbens, and Donald B Rubin. 1996. "Identification of Causal Effects Using Instrumental Variables." *Journal of the American Statistical Association* 91 (434): 444–455.
- Angrist, Joshua D., and Jörn-Steffen Pischke. 2008. *Mostly Harmless Econometrics: An Empiricist's Companion.* Princeton University Press.
- Blandhol, Christine, Magne Mogstad, John Bonney, and Alexander Torgovitsky. 2022. "When Is TSLS Actually LATE?" Preprint, February 9, 2022.
- Frölich, Markus, and Blaise Melly. 2013. "Identification of Treatment Effects on the Treated with One-Sided Non-Compliance." *Econometric Reviews* 32, no. 3 (March 1, 2013): 384–414.
- Imbens, Guido W. 2010. "Better LATE Than Nothing: Some Comments on Deaton (2009) and Heckman and Urzua (2009)." *Journal of Economic Literature* 48, no. 2 (June): 399–423.
- Imbens, Guido W, and Joshua D Angrist. 1994. "Identification and Estimation of Local Average Treatment Effects," 12.
- Keane, Michael, and Timothy Neal. 2021. "A Practical Guide to Weak Instruments," 37.
- ———. 2022. "A Practical Guide to Weak Instruments." *Discussion Papers,* Discussion Papers, nos. 2021-05d (June).

Bibliography II

- Stock, James H, Jonathan H Wright, and Motohiro Yogo. 2002. “A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments.” *Journal of Business & Economic Statistics* 20, no. 4 (October): 518–529.

Outline

Demonstrating the LATE theorem

Demonstrating the LATE theorem

Main instrumental variable hypotheses

1. **Independance** $\left(Y_i(D_1,1), Y_i(D_0,0), D_{1i}, D_{0i}\right) \perp Z_i$
2. **exclusion** $Y_i(d,1) = Y_i(d,0) \equiv Y_{di}$ pour $d = 0, 1.$
3. **first-stage** $\mathbb{E}[D_{1i} - D_{0i}] \neq 0$

We use exclusion to write:

$$\mathbb{E}[Y_i|Z_i = 1] = \mathbb{E}[Y_{0i} + (Y_{1i} - Y_{0i})D_i|Z_i]$$

And with independence:

$$\Rightarrow \mathbb{E}[Y_i|Z_i = 1] = \mathbb{E}[Y_{0i} + (Y_{1i} - Y_{0i})D_{1i}]$$

Similarly we write:

$$\mathbb{E}[Y_i|Z_i = 0] = \mathbb{E}[Y_{0i} + (Y_{1i} - Y_{0i})D_i]$$

Demonstrating the LATE theorem

Reduced form

$$
\begin{aligned}
\mathbb{E}[Y_i|Z_i=1]-\mathbb{E}[Y_i|Z_i=0] = \quad & \color{red}{\mathbb{E}[Y_{0i}]} + \mathbb{E}[(Y_{1i}-Y_{0i})D_{1i}] \\
- \quad & \color{red}{\mathbb{E}[Y_{0i}]} + \mathbb{E}[(Y_{1i}-Y_{0i})D_{0i}] \\
= \quad & \mathbb{E}[(Y_{1i}-Y_{0i})(D_{1i}-D_{0i})]
\end{aligned}
$$

We use the law of iterated expectation

$$
\begin{aligned}
\mathbb{E}[(Y_{1i}-Y_{0i})(D_{1i}-D_{0i})] = {} & \mathbb{E}[(Y_{1i}-Y_{0i})|D_{1i}>D_{0i}]Pr(D_{1i}>D_{0i}) \\
& - \mathbb{E}[(Y_{1i}-Y_{0i})|D_{1i}<D_{0i}]\underbrace{Pr(D_{1i}<D_{0i})}_{\text{Defiers}}
\end{aligned}
$$

Without monotonicity, we may have situations where the reduced form is null or negative whereas the true effect is positive for everyone.

Demonstrating the LATE theorem

Assuming monotonicity

$$\begin{aligned}\mathbb{E}[Y_i|Z_i=1]-\mathbb{E}[Y_i|Z_i=0] &= \mathbb{E}[(Y_{1i}-Y_{0i})|D_{1i}>D_{0i}]Pr(D_{1i}>D_{0i}) \\ &\quad -\mathbb{E}[(Y_{1i}-Y_{0i})|D_{1i}<D_{0i}]\underbrace{Pr(D_{1i}<D_{0i})}_{=0} \\ &= \underbrace{\mathbb{E}[(Y_{1i}-Y_{0i})|D_{1i}>D_{0i}]}_{\text{effect on compliers}}Pr(D_{1i}>D_{0i})\end{aligned}$$

Similarly we find

$$\begin{aligned}\mathbb{E}[D_i|Z_i=1]-\mathbb{E}[D_i|Z_i=0] &= \mathbb{E}[D_{1i}-D_{0i}] \\ &= Pr(D_{1i}>D_{0i})\end{aligned}$$

Therefore, the Wald estimand identifies the average treatment effect on compliers:

$$\beta_{Wald}=\frac{\mathbb{E}[Y_i|Z_i=1]-\mathbb{E}[Y_i|Z_i=0]}{\mathbb{E}[D_i|Z_i=1]-\mathbb{E}[D_i|Z_i=0]}=\mathbb{E}[(Y_{1i}-Y_{0i})|D_{1i}>D_{0i}]$$