Statistical Modeling and Causal Inference with R

Week 5: Instrumental variable estimation

Manuel Bosancianu

Max Schaub

October 5, 2020

Hertie School of Governance

Today's focus

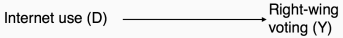
- ✓ Why instrumental variables? An example.
- ✓ Defining instrumental variables
- ✓ Analyze experiments affected by non-compliance using IVs
- ✓ Use of IVs in observational studies

Why instrumental variables? An example.

Why instrumental variables? An example.

Example: Estimate effect of internet use on right-wing support (Schaub & Morisi, 2020).

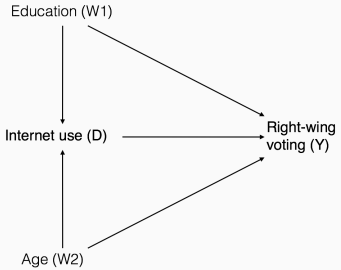
- ✓ Rise of right-wing populism coincides with internet age
- ✓ Internet way for right-wing actors to circumvent gatekeepers in media
- ✓ Targeted online campaigns during election time



Why instrumental variables? An example.

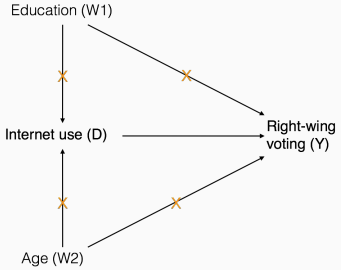
The problem is tht there are many plausible confounders, some of them observable...

- ✓ Education (educated use internet more, vote less often for right)
- ✓ Age (young use internet more, vote more often for right)
- ✓ ...



Why instrumental variables? An example.

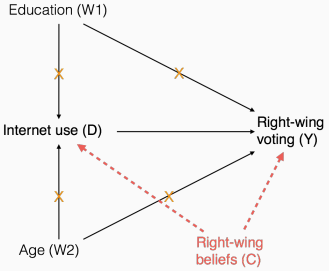
Observable confounders can be controlled for by conditioning on them...



Why instrumental variables? An example.

…but this is impossible for unobservable confounders, e.g. right-wing mindset!

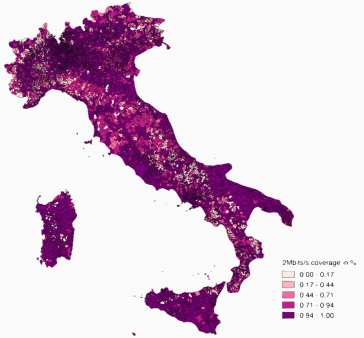
- ✓ Right-wing ideology predicts right-wing voting, and plausibly also lets people seek out right-wing content online, leading to higher internet consumption



Why instrumental variables? An example.

We solve this problem with an instrumental variable: a variable that changes the endogenous predictor internet use – and therefore the outcome – but is unrelated to possible confounders, esp. right-wing beliefs.

Here, local broadband availability is used as such an instrument:



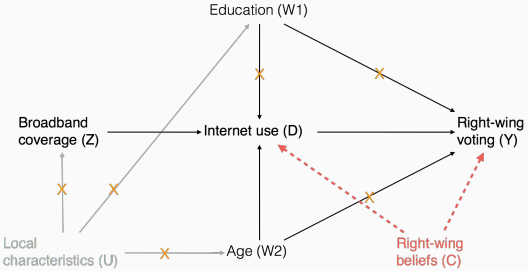
Why instrumental variables? An example.

Justification for relevance of the instrument and independence of confounders:

- ✓ The availability of high-speed internet increases internet consumption, and makes possible the delivery of higher quality, more interesting, and more persuasive content.
- ✓ At the same time, broadband availability is unlikely to be correlated with right-wing voting through any other channel than increasing internet use, given some controls such as economic prowess of the region and population density/degree of urbanization.
- ✓ In particular, it is only implausibly correlated with right-wing ideology.

Why instrumental variables? An example.

In the DAG, this IV appears as a variable that is causally connected to internet use, but shares no open paths with the potential confounders.



Defining instrumental variables

Defining IVs: Basic idea

We are interested in the causal effect of D on Y, but the relationship is subject to confounding, measurement problems, or non-compliance.

Basic idea of instrumental variables (IV):
Split the variation of D into two parts

- ✓ one potentially related to (potentially unobserved) confounders W
- ✓ one truly exogenous, i.e. caused by other factors unrelated to the confounders
- ✓ To find the portion of D unrelated to confounders, one needs a variable Z (the instrumental variable) that is (as if) randomly assigned and related to D

Defining IVs through their requirements

An instrumental variable (Z) is a variable that:

1. Is substantially related to the treatment/independent variable (D) $\rightarrow$ the first stage

   The first stage is shown with a positive correlation between D and Z

2. Only influences the outcome Y through its effect on D, i.e. there are no other causal paths running from Z to Y other than through D $\rightarrow$ the exclusion restriction.

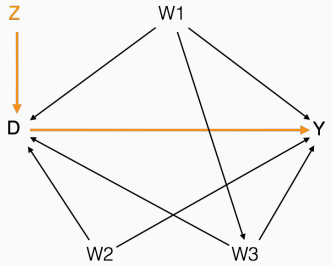
   The exclusion restriction is not fully testable, but it implies that:

   - ✓ Z is not correlated with any confounder influencing the relationship between the treatment/independent variable (D) and the outcome (Y).
   - ✓ Z should not predict treatments or outcomes similar to D or Y (placebo/falsification criterion).

Defining IVs as DAG

An instrumental variable (Z) is a variable that:

1. Is connected through a single causal path to the endogenous predictor D
2. Is not a descendant of any confounder (W1, W2, etc.)



Defining IVs in the POF

The basic setup for IVs in the potential outcomes framework:

The instrument/(randomized) treatment assignment: Z_i
The potential treatment variables: $D_i|Z_i = 1$ and $D_i|Z_i = 0$

1. $D_i|Z_i = 1$: treatment status if assigned to treatment
2. $D_i|Z_i = 0$: treatment status if not assigned to treatment

The potential outcomes: $[Y_i = Y_i|Z_i = z, D_i = d]$
with $z \in \{0, 1\}$ and $d \in \{0, 1\}$

Defining IVs in the POF

Based on the POF notation, we can define several quantities of interest:

The first stage: $E[D_i|Z_i = 1] - E[D_i|Z_i = 0]$

The reduced form: $E[Y_i|Z_i = 1] - E[Y_i|Z_i = 0] = ITT$

The Wald estimator: $$\frac{E[Y_i|Z_I = 1] - E[Y_i|Z_I = 0]}{E[D_i|Z_I = 1] - E[D_i|Z_I = 0]} = LATE$$

First stage: the effect of Z on D/the share of individuals *i* that *comply*, i.e. who change their treatment status if assigned to the treatment.

Reduced form/ intent-to-treat (ITT) effect: effect of being assigned to the treatment, estimates the encouragement or program effect; causally identified because randomly assigned, but unlike ATE in case of partial compliance (a first-stage<1)

Wald estimator: the ITT scaled up by the first-stage, estimates the local average treatment effect (LATE) under certain assumptions

IV assumptions

The POF notation also allows us to state assumptions invoked by IV analyses more formally:

- ✓ Relevance/first stage: $Z \not\perp\!\!\!\perp D$
- ✓ Exclusion restriction: $Z \perp\!\!\!\perp W$ and $Z \perp\!\!\!\perp Y | D, W$
- ✓ Stable Unit Treatment Assumption (SUTVA): no spillovers between treated and untreated units
- ✓ Monotonicity: Effect of treatment only in one direction (either positive or negative)
- ✓ Homogeneity: Constant treatment effect assumption

Defining IVs in the POF

Distinguishing between treatment assignment Z and actual treatment received also allows us to define four principle strata or compliance types :

- ✓ compliers: $[D_i|Z_i = 1] = 1$ and $[D_i|Z_i = 0] = 0$
- ✓ non-compliers
  - always-takers: $[D_i|Z_i = 1] = [D_i|Z_i = 0] = 1$
  - never-takers: $[D_i|Z_i = 1] = [D_i|Z_i = 0] = 0$
  - defiers: $[D_i|Z_i = 1] = 0$ and $[D_i|Z_i = 0] = 1$

Main uses of instrumental variables

Two main uses for instrumental variables:

1. Analyze experiments affected by non-compliance
2. Make causal estimates for relationships affected by unmeasured or unmeasurable confounders in observational studies – by exploiting natural experiments (cp. the example in the beginning)

Analyze experiments affected by non-compliance using IVs

Experiments with non-compliance

Example: micro-finance experiment, where 100 individuals each were assigned to treatment and control.
In a perfect world, this would look as follows:

	Assigned to control $Z_i = 0$	Assigned to treatment $Z_i = 1$
Did not take loan $D_i = 0$	100	–
Took out loan $D_i = 1$	–	100

Observed treatment strata in perfect micro-finance experiment

All individuals would be compliers: those assigned to treatment took out loan, those assigned to control did not, i.e. $D_i = Z_i$.
Given random assignment, we could therefore calculate
$NATE = ATE = ITT = E[Y_i|D_i = 1] - E[Y_i|D_i = 0] = E[Y_i|Z_i = 1] - E[Y_i|Z_i = 0]$

Experiments with non-compliance

In the real world, experiments are often affected by imperfect compliance: some people might fail or refuse to take up the treatment, while others might 'consume' it anyway. The data from a real-world experiment might therefore look as follows:

	Assigned to control $Z_i = 0$	Assigned to treatment $Z_i = 1$
Did not take loan $D_i = 0$	80	10
Took out loan $D_i = 1$	20	90

Observed treatment strata in real-world micro-finance experiment

With such selective non-compliance $NATE \neq ATE$ because potential outcomes of compliers might differ from those of non-compliers, i.e.
$E[Y_{i0}|D = 1] - E[Y_{i0}|D = 0] \equiv E[u_i|D_i] \neq 0$.

Recovering causal effects under non-compliance

What can be done? What estimators can we recover?

1. The Intent to Treat (ITT) effect:
   $ITT = E[Y_i|Z_i = 1] - E[Y_i|Z_i = 0]$

   First, with random treatment assignment, we can always recover the ITT (sometimes also called encouragment or program effect), no matter whether treatment was received or not. Note: This can be useful for policy advice, assuming that non-compliance is part of all policy implementations.

2. The local average treatment effect (LATE) for the compliers:
   $$LATE = \frac{E[Y_i|Z_I = 1] - E[Y_i|Z_I = 0]}{E[D_i|Z_I = 1] - E[D_i|Z_I = 0]}$$
   We may also recover the LATE (sometimes also called Complier Average Causal Effect (CACE), i.e. the causal effect of the treatment for the compliers only.

Recovering causal effects under non-compliance

However, for the LATE to be estimable we have to invoke the monotonicity and exclusion restrictions introduced above. Why is that so?

The problem: instead of the principal strata, which would allow us to identify the compliers, we can only ever see the observed strata, where different compliance types may mix.

So instead of observing the principal strata:

		$Z_i = 0$	
		$D_i = 1$	$D_i = 0$
$Z_i = 1$	$D_i = 0$	Defier	Never-taker
	$D_i = 1$	Always-taker	Complier

Principal strata/compliance types

Recovering causal effects under non-compliance

We only ever see the observed strata, where different compliance types may mix:

	$Z_i = 0$	$Z_i = 1$
$D_i = 0$	Complier/Never-taker	Never-taker/Defier
$D_i = 1$	Always-taker/Defier	Complier/Always-taker

Observable treatment responses among subjects

Recovering causal effects under non-compliance

By invoking the montonicity assumption – the assuption that assignment to treatment *never dissuades* someone from taking the treatment – we can 'assume away' defiers:

	$Z_i = 0$	$Z_i = 1$
$D_i = 0$	Complier/Never-taker	Never-taker/~~Defier~~
$D_i = 1$	Always-taker/~~Defier~~	Complier/Always-taker

Our observed strata are now composed of individuals who either are unreactive to the treatment (always-takers and never-takers) or change their behavior in a consistent way in reaction to the treatment, and to treatment only – the exclusion restriction.

Recovering causal effects under non-compliance

Another way of conceiving the problem is that the ITT consists of four subgroups, corresponding to the principal strata and occuring with a certain probability *Pr*:

$$\begin{aligned}\text{ITT} &= \text{ITT}_c \times \Pr(\text{compliers}) + \text{ITT}_a \times \Pr(\text{always-takers}) \\ &\quad + \text{ITT}_n \times \Pr(\text{never-takers}) + \text{ITT}_d \times \Pr(\text{defiers})\end{aligned}$$

Under monotonicity and the exclusion restriction, this simplifies to:

$$\begin{aligned}\text{ITT} &= \text{ITT}_c \times \Pr(\text{compliers}) + \text{ITT}_a \times \Pr(\text{always-takers}) \\ &\quad + \text{ITT}_n \times \Pr(\text{never-takers}) + 0 \;\; [\textbf{monotonicity}] \\ &= \text{ITT}_c \times \Pr(\text{compliers}) + 0 \times \Pr(\text{always-takers}) \\ &\quad + 0 \times \Pr(\text{never-takers}) \;\; [\textbf{exclusion restriction}] \\ &= \text{ITT}_c \times \Pr(\text{compliers}) \;\; [\textbf{rearrange}] \\ \text{ITT}_c &= \frac{\text{ITT}}{\Pr(\text{compliers})} = \text{LATE}\end{aligned}$$

Recovering causal effects under non-compliance

Sometimes, we know that strictly only units assigned to treatment could get access to the treatment, e.g. in medical trials, where the administration of a new drug was confined to a controlled setting (but be mindful that these settings are rare).

In these cases, the situation simplifies, because we *know* that there are no always-takers. Our uncertainty with regard to the observed strata is therefore reduced to the question who are compliers, and who are never-takers:

	$Z_i = 0$	$Z_i = 1$
$D_i = 0$	Complier/Never-taker	Never-taker/~~Defier~~
$D_i = 1$	~~Always-taker~~/~~Defier~~	Complier/~~Always-taker~~

In this case, our Wald/LATE estimator therefore recovers the average-treatment effect for the treated, the ATT!

Estimating ITT and LATE in practice

Estimating ITT and LATE in practice

How to calculate ITT and LATE?

Given information on a) the frequency of observations, and b) the outcomes in each observed strata, and invoking our assumptions, we can estimate the ITT and LATE 'by hand' as follows:

Observed obs in treatment strata		$Z_i = 0$	$Z_i = 1$
	$D_i = 0$	80	10
	$D_i = 1$	20	90
Observed outcomes $E[Y_i \mid D_i, Z_i]$		$Z_i = 0$	$Z_i = 1$
	$D_i = 0$	500	1000
	$D_i = 1$	450	800

ITT = $E[Y_i|Z_i = 1] - E[Y_i|Z_i = 0] = (1000 \times 0.1 + 800 \times 0.9) - (500 * 0.8 + 450 \times 0.2) = 330$
First stage = $E[D_i|Z_i = 1] - E[D_i|Z_i = 0] = 90/(10 + 90) - 20/(20 + 80) = 0.7$
LATE/Wald estimator = ITT/First stage = $330/0.7 = 471$

Estimating ITT and LATE in practice

Instead of calculating the first-stage and LATE by hand, this can be done within the regression framework with a sequence of two regressions that define the two-stage-least-square estimator (abbreviated as 2SLS or TSLS):

1. Calculate the first stage by regressing the treatment status on the treatment assigment:
   $D = \gamma_0 + \gamma_1 Z_i + v_i$
2. Calculate the predicted values $\hat{D}$
3. Regress the outcome Y on the predicted values $\hat{D}$:
   $Y_i = \beta_0 + \beta^{2sls}\hat{D} + u_i$ β^{2sls} is the LATE estimator!

The ITT/reduced form is simply:
$Y_i = \theta_0 + \theta_1 Z_i + w_i$

The estimate for β^{2sls} can also be obtained by $\frac{\theta_1}{\gamma_1}$.

Estimating ITT and LATE in practice

In practice, the 2SLS estimator is much more commonly used because:

1. Allows for the inclusion of controls ($\rightarrow$ include all controls in both the first and the second stage)
2. Accomodates continuous treatments and instruments
3. Provides standard errors/certainty bounds for your estimate
   (Note: Use a dedicated software package to conduct 2SLS rather than calculating the first and second stage by hand, since otherwise your estimates for the standard errors in the second-stage will ignore the additional uncertainty introduced by the estimation of $\hat{D}$ in the first stage)

The ITT/reduced form can be estimated using OLS.

Use of IVs in observational studies

Use IVs to analyze natural experiments

A second major use of IVs is in observational studies where the causal relationship under consideration is affected by unobserved/unobservable confounders, e.g. the example on internet use and right-wing voting.

In these situations, a natural experiments sometimes provides the instrument that is needed for identification.

Natural experiments: situations where nature (including the natural environment, people, administrative rules) happened to cause plausibly as-if random assignment (the instrumental variable!) of your endogenous predictor of interest, but where in contrast to a true experiment the researcher has no control over design and assignment (Dunning, 2012, 17).

In general, all the lessons regarding assumptions, estimators etc. from above apply to the analysis of natural experiments as well.

In practice, you usually face two issues: a) Finding a good instrument, and b) convincing your audience that your exclusion restriction holds

Examples of IVs

Where to find good IVs/natural experiments? Read, read, and keep your eyes open!

Examples of natural experiments:

- ✓ Military draft lottery (Z) → Military service (D) → Lifetime earnings (Y)
  (Angrist, 1990)
- ✓ Distance to African coastline (Z) → Transatlantic slavery (D) → Trust (Y)
  (Nunn & Wantchekon, 2011)
- ✓ Rainfall (Z) → Economic decline (D) → Civil war (Y)
  (Miguel, Satyanath, & Sergenti, 2004)
- ✓ Disease environment (Z) → Institutions (D) → Economic development (Y)
  (Acemoglu, Johnson, & Robinson, 2001)
- ✓ Convicts' placement in Australia (Z) → Sex ratio (D) → Gender norms (Y)
  (Grosjean & Khattar, 2019)
- ✓ River gradient (Z) → Dam construction (D) → Income distribution (Y)
  (Duflo & Pande, 2007)

A few practical points

1. In observational studies, IVs are often only plausible given controls.
2. Your first-stage may be weak, which can lead to biased estimates; try to find a strong instrument, and always show the first-stage.
3. The fact that the exclusion restriction holds is something you should carefully argue and demonstrate where possible e.g. by presenting tests showing the independence of instrument from observables, and by showing placebo results/falsification tests.
4. As with the experiments with non-compliance, the reduced form is the intent-to-treat effect, and the 2SLS estimate is the LATE for compliers – try to develop a deep understanding what this means in your case and when interpreting your estimates.

Keep these in mind when designing your own research and when assessing others!

Thank you for watching, and see you next Monday!

References i

Acemoglu, D., Johnson, S., & Robinson, J. A. (2001, December). The Colonial Origins of Comparative Development: An Empirical Investigation. *American Economic Review*, *91*(5), 1369–1401.

Angrist, J. D. (1990). Lifetime earnings and the Vietnam era draft lottery: Evidence from social security administrative records. *The American Economic Review*, 313–336.

Duflo, E., & Pande, R. (2007, May). Dams. *The Quarterly Journal of Economics*, *122*(2), 601–646.

Dunning, T. (2012). *Natural experiments in the social sciences: A design-based approach*. Cambridge; New York: Cambridge University Press.

Grosjean, P., & Khattar, R. (2019, March). It's Raining Men! Hallelujah? The Long-Run Consequences of Male-Biased Sex Ratios. *The Review of Economic Studies*, *86*(2), 723–754.

References ii

Miguel, E., Satyanath, S., & Sergenti, E. (2004, August). Economic Shocks and Civil Conflict: An Instrumental Variables Approach. *Journal of Political Economy*, *112*(4), 725–753.

Nunn, N., & Wantchekon, L. (2011, December). The Slave Trade and the Origins of Mistrust in Africa. *The American Economic Review*, *101*(7), 3221–3252.

Schaub, M., & Morisi, D. (2020). Voter Mobilisation in the Echo Chamber: Broadband Internet and the Rise of Populism in Europe. *European Journal of Political Research*, *59*(4).