Causal Inference: Instrumental Variables and 2SLS Estimation

Ken Stiller

20th June 2024

Overview

Introduction

000

Previous three weeks were about "selection-on-observables": how to estimate treatment effects by controlling for all relevant covariates, and exploiting variation in time and units in a "difference-in-differences" .

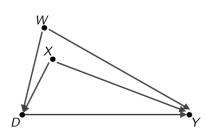
This week we consider situations where:

- ▶ Treatment depends on unobservables, i.e. CIA does not hold
- ▶ But treatment also depends on an as-if random variable Z_i that only affects the outcome through treatment (at least conditional on covariates).

This special variable Z_i is an **instrument**: it changes D_i , and we can use this change to measure the effect of D_i .

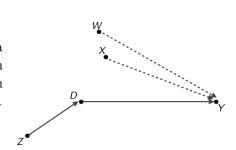
To estimate the effect of D on Y, we must observe and control for X and W.

●00



If D is completely determined by a randomization process Z, so we can measure the effect of D on Y, even if X and W are not observed (e.g. through DIGM).

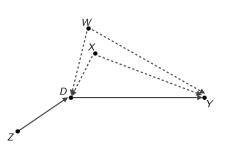
Introduction



Graphical overview: instrumental variables

If D is partly determined by random Z, and Z does not affect Y in any other way, we can measure the effect of D on Y. This is the case even if X and W are not observed (through IV techniques).

In this case, we call Z the instrument.



Treatment assigned vs. treatment received

In an experiment, we can distinguish between treatment assigned Z_i and treatment received D_i .

We previously (implicitly) assumed $D_i = Z_i$. But in practice there may be **non-compliance**:

- ► GOTV canvassing experiment in which some people don't answer the door We discussed this in the lecture on randomized experiments
- ▶ lottery for school places in which some lottery winners do not attend
- ▶ draft lottery for military in which some are drafted but do not serve, some not drafted but serve

Introduction

800

One-sided and two-sided non-compliance

	Control Treatment	
GOTV canvassing experiment	No visit from canvasser	Visit from canvasser (but some people don't answer door)
Lottery for school places	No place offered at school	Place offered at school (but some people don't attend)
Military draft	Not drafted (but some volunteer and serve anyway)	Drafted (but some receive exemption or deferment)

Intention-to-treat (ITT)

Denote by $Y_{i,Z1}$ and $Y_{i,Z0}$ i's potential outcomes if assigned to treatment vs. control.

Intention-to-treat (ITT) effect defined as ITT $\equiv E[Y_{i,Z1} - Y_{i,Z0}]$.

If Z_i is randomized, $E[Y_i|Z_i=1]-E[Y_i|Z_i=0]$ is an unbiased estimator of the ATE.

If Z_i is randomized but there is non-compliance (i.e. $Z_i \neq D_i$ for some i), $E[Y_i|D_i=1]-E[Y_i|D_i=0]$ (DIGM) will generally not be an unbiased estimator of the ATE.

Consider the examples of lottery for private school places.

Instrumental variables methods (IV) let us use ITT (effect of treatment assignment) to estimate an ATE (effect of treatment).

Assigned to control $(Z_i = 0)$

Introduction

		Not treated $(D_i = 0)$	Treated $(D_i = 1)$
Assigned to treatment $(Z_i = 1)$	Not treated $(D_i = 0)$	Never taker (N)	Defier (D)
	Treated $(D_i = 1)$	Complier (C)	Always taker (A)

- ► Can we identify the compliance type of an individual?
- ► Can we measure the proportion of each compliance type $(\pi_A, \pi_C, \pi_D, \pi_N)$? (Only with some further assumptions.)

Introduction

000

Estimating compliance frequencies

Who gets treated when $Z_i = 0$? Always-takers and defiers. Who gets treated when $Z_i = 1$? Always-takers and compliers.

Assumption: treatment assignment Z_i is random \Longrightarrow compliance-type proportions same for $Z_i = 0$ and $Z_i = 1$.

$$E[D_i = 1 | Z_i = 0] = \pi_A + \pi_D$$

$$E[D_i = 1 | Z_i = 1] = \pi_A + \pi_C$$

Can't estimate π_A , π_D , π_C , or π_N .

Estimating compliance frequencies (2)

Further assumption: $\pi_D = 0$ (no defiers).

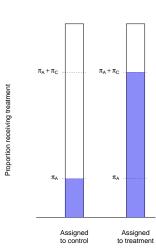
This implies

Introduction

$$E[D_i = 1 | Z_i = 0] = \pi_A$$

and

$$E[D_i = 1|Z_i = 1] - E[D_i = 1|Z_i = 0] = \pi_0$$



Further assumption: $\pi_D = 0$ (no defiers).

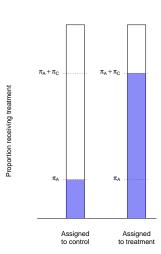
This implies

Introduction

$$E[D_i = 1|Z_i = 0] = \pi_A$$

and

$$E[D_i = 1|Z_i = 1] - E[D_i = 1|Z_i = 0] = \pi_C$$



We can decompose the ITT by **compliance type**.

Let π_G and ITT_G be proportion and ITT for compliance type $G \in \{C, A, N, D\}$.

Then by definition

$$ITT = \pi_C ITT_C + \pi_A ITT_A + \pi_N ITT_N + \pi_D ITT_D$$
 (1)

Let's assume

- ▶ No defiers (monotonicity).
- **Exclusion restriction:** Treatment assigned only affects outcomes by affecting treatment received.

We can decompose the ITT by **compliance type**.

Let π_G and ITT_G be proportion and ITT for compliance type $G \in \{C, A, N, D\}.$

Then by definition

$$ITT = \pi_C ITT_C + \pi_A ITT_A + \pi_N ITT_N + \pi_D ITT_D$$
 (1)

Let's assume

000

- ▶ No defiers (monotonicity).
- **Exclusion restriction:** Treatment assigned only affects outcomes by affecting treatment received.

000

No defiers tells us that $\pi_D = 0$. Exclusion restriction tells us that $ITT_A = ITT_N = 0$. So:

$$ITT = \pi_C ITT_C + \pi_A \mathbf{0} + \pi_N \mathbf{0} + \mathbf{0}ITT_D.$$
 (2)

$$ITT = \pi_C CATE_C, \tag{3}$$

No defiers tells us that $\pi_D = 0$. Exclusion restriction tells us that $ITT_A = ITT_N = 0$. So:

$$ITT = \pi_C ITT_C + \pi_A \mathbf{0} + \pi_N \mathbf{0} + \mathbf{0}ITT_D.$$
 (2)

Exclusion also tells us that, for compliers, the effect of *treatment assignment* on outcomes is the same as the effect of *treatment* on outcomes:

$$ITT = \pi_C CATE_C, (3)$$

where $CATE_C$ is the conditional average treatment effect for compliers.

LATE and the Wald estimator

Assuming $\pi_C > 0$ (non-zero complier proportion), the conditional average treatment effect for compliers or local average treatment effect (LATE) is

$$LATE = CATE_C = \frac{ITT}{\pi_C}$$
 (4)

If in addition Z_i is randomly assigned, we have an unbiased estimator for the above - the **Wald estimator**:

$$CA\hat{T}E_C = \frac{E[Y_i|Z_i=1] - E[Y_i|Z_i=0]}{E[D_i|Z_i=1] - E[D_i|Z_i=0]} = \frac{\text{effect of } Z_i \text{ on } Y_i}{\text{effect of } Z_i \text{ on } D_i} = \frac{ITT_Y}{ITT_D}$$
(5)

LATE and the Wald estimator

Assuming $\pi_C > 0$ (non-zero complier proportion), the conditional average treatment effect for compliers or local average treatment effect (LATE) is

$$LATE = CATE_C = \frac{ITT}{\pi_C}$$
 (4)

If in addition Z_i is randomly assigned, we have an unbiased estimator for the above - the **Wald estimator**:

$$CA\hat{T}E_C = \frac{E[Y_i|Z_i=1] - E[Y_i|Z_i=0]}{E[D_i|Z_i=1] - E[D_i|Z_i=0]} = \frac{\text{effect of } Z_i \text{ on } Y_i}{\text{effect of } Z_i \text{ on } D_i} = \frac{\text{ITT}_Y}{\text{ITT}_D}$$
(5)

LATE and the Wald estimator

Four assumptions used (not including SUTVA):

- ► No defiers (monotonicity)
- ▶ Exclusion restriction (Z_i affects Y_i only through D_i)
- ► Non-zero complier proportion
- ightharpoonup Random assignment of Z_i

Applications of IV

IV methods can be seen as a remedy for a **broken experiment**, i.e. failure to obtain 100% compliance.

More positively, IV methods can be used as part of the design of using natural experiments in which some random variation in Z_i creates some variation in D_i and then (given exclusion restriction) measures effect of D_i on some outcome Y_i .

Encouragement design example

Proposition 209: 1996 ballot proposition to end race-based preferences (affirmative action) in California government policies

Research question (Albertson and Lawrence 2009): Could watching a TV program affect citizens' attitudes toward Prop. 209?



Albertson and Lawrence 2009: Design

- ▶ Representative sample of households in Orange County, CA, interviewed by phone in October 1996
- ▶ All respondents told there will be a follow-up interview after the election
- Random subset of respondents told to watch upcoming TV debate on Prop. 209
- ▶ In follow-up, asked if they watched the debate; supported Prop. 209; felt knowledgeable about Prop. 209

In this design:

- ▶ What are Z_i , D_i , Y_i ?
- ▶ What does intention-to-treat (ITT) effect mean?
- ▶ What is the exclusion restriction?
- \blacktriangleright What does the LATE (CATE_C) measure?

Albertson and Lawrence 2009: Data

	$Z_i = 0$	$Z_i = 1$	Difference
Watched TV program	0.052	0.48	0.428
Know about Prop. 209	3.251	3.293	0.041
Support Prop. 209	0.654	0.651	-0.003
			1

- ▶ What is π_C (proportion of compliers)?
- ▶ What is the ITT?

Introduction

800

▶ What is LATE i.e. CATE $_C$?

Examples

Taking stock

We assumed binary treatment assignment and binary treatment.

Given random treatment assignment, we can

- \triangleright estimate the **intention-to-treat** effect (ITT) by comparing average Y_i among units assigned to treatment and units assigned to control
- \blacktriangleright estimate the **proportion of compliers** (π_C) by comparing average D_i among units assigned to treatment and units assigned to control
- \triangleright estimate the **LATE** (CATE_C) by dividing the ITT by the proportion of compliers

Can we generalize this somehow?

- \blacktriangleright non-binary treatment (D_i)
- \blacktriangleright non-binary instrument (Z_i)
- \triangleright covariates (e.g. because non-random Z_i)
- ▶ more than one instrument

Another way to get LATE

0000000000

The LATE approach

We estimated the LATE with

$$\frac{E[Y_i|Z_i=1] - E[Y_i|Z_i=0]}{E[D_i|Z_i=1] - E[D_i|Z_i=0]} = \frac{ITT_Y}{ITT_D}$$

Here is another way:

Introduction

000

- ▶ Regress D_i on Z_i , get fitted values \hat{D}_i
- ightharpoonup Regress Y_i on \hat{D}_i

This is called two-stage least squares.

Examples

Considerations

000

Regressing Y on Z gives you ITT_V .

Given binary Z, ITT_Y is an underestimate of $CATE_C$.

Wald estimator inflates ITT_Y by dividing by compliance rate $(E[D_i|Z_i=1]-E[D_i|Z_i=0]).$

TSLS inflates ITT_Y by replacing Z by \hat{D} , i.e. regressing Y on \hat{D} instead of Y on Z.

Now we can generalize

000

Wald estimator is limited to binary D_i and Z_i :

$$\lambda = \frac{E[Y_i|Z_i=1] - E[Y_i|Z_i=0]}{E[D_i|Z_i=1] - E[D_i|Z_i=0]} = \frac{ITT_Y}{ITT_D}$$

Two-stage least squares is a much more general procedure:

First stage:
$$D_i = \alpha_1 + \phi Z_i + \beta_1 X_{1i} + \gamma_1 X_{2i} + e_{1i}$$

Second stage: $Y_i = \alpha_2 + \lambda \hat{D}_i + \beta_2 X_{1i} + \gamma_2 X_{2i} + e_{2i}$

where Z_i and D_i might not be binary and you can include covariates e.g. X_{1i}, X_{2i} same covariates must be included in the first and second stage

Two-stage least squares: terminology

Terminology:

Reduced form: $Y_i = \alpha_0 + \rho Z_i + \beta_0 X_{1i} + \gamma_0 X_{2i} + e_{0i}$ First stage: $D_i = \alpha_1 + \phi Z_i + \beta_1 X_{1i} + \gamma_1 X_{2i} + e_{1i}$ Second stage: $Y_i = \alpha_2 + \lambda \hat{D}_i + \beta_2 X_{1i} + \gamma_2 X_{2i} + e_{2i}$

NB: λ is the LATE. Must use same covariates in first stage and second stage.

Two-stage least squares: assumptions

Key assumptions (Wald assumptions with covariates and without "complier" terminology):

- ▶ Non-zero first-stage: instrument affects treatment, conditional on covariates ($\phi \neq 0$ in first stage)
- ▶ Independence (exogeneity, ignorability): instrument unrelated to potential outcomes, conditional on covariates (no OVB on ρ in reduced form or ϕ in first stage)
- ▶ Exclusion restriction: instrument only affects outcome through treatment, conditional on covariates
- ▶ Monotonicity: instrument's effect on treatment is weakly positive or weakly negative for all units

Colantone and Stanig (2018): Globalization and the Brexit vote?

- ▶ Question: Did economic globalization lead to support for the Leave option in the Brexit referendum?
- ightharpoonup Treatment: Import Shock is the strength of the Chinese import shock at the regional level i
- ightharpoonup Outcome: Leave Share (for Brexit) in NUTS3 region i

To consider:

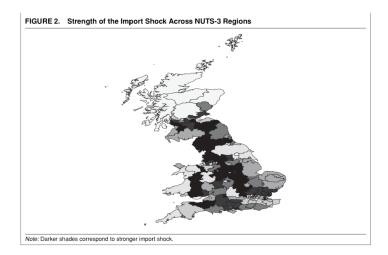
- ▶ What about just regressing outcome on treatment?
- ▶ What covariates might remove the bias in that regression?
- ► How might an IV approach help?

Examples

Colantone and Stanig (2018): Plausible Instrument

- ► The role of China in the import shock: in two decades the import share went from 1% to around 9% since China's WTO membership
- ▶ The literature (Autor, Dorn, and Hanson, 2013) has identified reasons why the surge in Chinese competition constitutes an excellent exogenous source of identification.
- ▶ Chinese import shock: the measure has a very intuitive interpretation: for given changes in nation-level imports per worker, the Chinese shock will be stronger in those regions in which a larger share of workers was initially employed in industries witnessing larger subsequent increases in imports from China.
- ▶ They instrument the import shock using the growth in imports from China to the United States across industries, which is due to the exogenous changes in supply conditions in China, rather than to domestic factors in the United Kingdom that could be

Colantone and Stanig (2018): Plausible Instrument



Introduction

800

Colantone and Stanig (2018): Results

VARIABLES	(1) Leave Share	(2) Leave Share	(3) Leave Share	(4) Leave Share	(5) Leave Share	(6) Leave Share
Import Shock	12.233** [4.763]	12.225*** [4.091]	12.965*** [4.543]	12.085*** [3.890]	11.073*** [3.861]	12.299**
Immigrant Share	[00]	[]	[]	-0.490*** [0.165]	-0.513*** [0.155]	-0.491* [0.154]
Immigrant Arrivals				-0.066 [0.741]	0.496 [0.801]	-0.058 [0.691]
NUTS-1 Fixed Effects	Υ	Υ	Υ	Υ ΄	Y	Y .
NUTS-2 Random Intercepts	N	Υ	N	N	Υ	N
Observations	167	167	167	167	167	167
R-Squared	0.57		0.57	0.65		0.65
Kleibergen-Paap F Statistic			662.7			614
Number of Groups		39			39	
Model	Linear	Hierarchical	IV	Linear	Hierarchical	IV

Martin and Yurukoglu (2017) on impact of Fox News in USA

- ▶ Question: "how much does consuming slanted news, like the Fox News Channel, change individuals' partisan voting preferences?"
- ► Treatment: Minutes spent watching Fox News Channel, based on surveys
- ▶ Outcome: Voting in presidential election, based on aggregate zip code-level results

To consider:

- ▶ What about just regressing outcome on treatment?
- ▶ What covariates might remove the bias in that regression?
- ► How might an IV approach help?

Martin & Yurukoglu (2)

Instrument: channel position of Fox News on the cable lineup

Evaluate:

- ► **Independence** (exogeneity, ignorability): instrument unrelated to potential outcomes, conditional on covariates
- **Exclusion restriction:** instrument only affects outcome through treatment, conditional on covariates
- ► Monotonicity: instrument's effect on treatment is weakly positive or weakly negative for all units



Examples

0000

Introduction

800

TABLE 2—FIRST-STAGE REGRESSIONS: NIELSEN DATA

		FNC minutes per week					
	(1)	(2)	(3)	(4)	(5)	(6)	
FNC position	-0.146	-0.075	-0.174	-0.167	-0.097	-0.111	
	(0.043)	(0.039)	(0.028)	(0.025)	(0.033)	(0.030)	
MSNBC position	0.078	0.073	0.064	0.070	0.019	0.020	
	(0.036)	(0.032)	(0.025)	(0.022)	(0.034)	(0.035)	
Has MSNBC only	1.904 (3.697)	1.137 (3.713)	-3.954 (4.255)	-2.804 (3.416)	-1.220 (6.180)	-1.562 (5.397)	
Has FNC only	31.423	26.526	23.460	22.011	15.141	15.069	
	(2.677)	(2.546)	(2.278)	(1.864)	(2.697)	(2.314)	
Has both	24.859	23.118	18.338	16.168	15.159	14.486	
	(2.919)	(2.687)	(2.361)	(1.991)	(3.216)	(2.842)	
Satellite FNC minutes	, ,	, ,	, ,	0.197 (0.013)	, ,	0.173 (0.015)	
Fixed effects Cable controls Demographics Robust <i>F</i> -stat Number of clusters	Year	State-year	State-year	State-year	County-year	County-year	
	Yes	Yes	Yes	Yes	Yes	Yes	
	None	None	Extended	Extended	Extended	Extended	
	11.39	3.72	39.02	44.7	8.86	13.43	
	5,789	5,789	4,830	4,761	4,839	4,770	
Observations R^2	71,150	71,150	59,541	52,053	59,684	52,165	
	0.030	0.074	0.213	0.377	0.428	0.544	

Notes: Cluster-robust standard errors in parentheses (clustered by cable system). Instrument is the ordinal position of FNC on the local system. The omitted category for the availability dummies is systems where neither FNC nor

Martin & Yurukoglu: second-stage

Introduction

000

TABLE 4—SECOND STAGE REGRESSIONS: ZIP CODE VOTING DATA

	2008 McCain vote percentage				
	(1)	(2)	(3)	(4)	
Predicted FNC minutes	0.152	0.120	0.157	0.098	
	(0.056, 0.277)	(0.005, 0.248)	(-0.126, 0.938)	(-0.121, 0.429)	
Satellite FNC minutes		-0.021 $(-0.047, 0.001)$		-0.015 $(-0.073, 0.022)$	
Fixed effects Cable system controls Demographics Number of clusters	State	State	County	County	
	Yes	Yes	Yes	Yes	
	Extended	Extended	Extended	Extended	
	4,814	3,993	4,729	4,001	
Observations \mathbb{R}^2	17,400	12,417	17,283	12,443	
	0.833	0.841	0.907	0.919	

Notes: The first stage is estimated using viewership data for all Nielsen TV households. See first-stage tables for description of instruments and control variables. Observations in the first stage are weighted by the number of survey individuals in the zip code according to Nielsen. Confidence intervals are generated from 1,000 independent STID-block-bootstraps of the first and second stage datasets. Reported lower and upper bounds give the central 95 percent interval of the relevant bootstrapped statistic.

Considerations

000

Why we should be skeptical of most IV designs

IV designs must convince us of two key untestable assumptions:

- \blacktriangleright The instrument Z_i satisfies **independence**, i.e. the CIA is met with respect to D_i and Y_i , e.g. because Z_i is random
- \blacktriangleright The instrument Z_i satisfies **exclusion**, i.e. it only affects Y_i through D_i

When Z_i is randomly determined in an experiment, it's easier to accept **independence** and think hard and discuss the **exclusion** assumption.

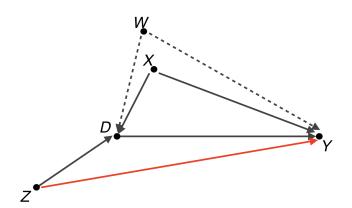
In an observational study, one should be skeptical about both.

- ▶ Is the CIA really satisfied in the reduced form?
- ▶ Is D_i really the only channel through which Z_i affects Y_i ?

Exclusion restriction violation DAG

Introduction

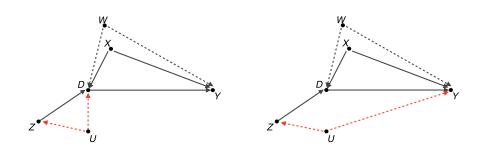
800



Independence violations DAGs

Introduction

800



(where U is an unobserved covariate)

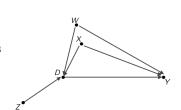
Can we test the exclusion restriction?

What about

- ightharpoonup regress Y on D and Z
- ightharpoonup conclude exclusion is valid if coefficient on Z is 0

Unfortunately this doesn't work, because D is also affected by X and W:

- ightharpoonup if X and W are observed, you don't need IV to estimate effect of D on Y
- ▶ if they are not observed, by controlling for D you induce an association between Z and X and/or W, leading to collider bias



Introduction

000

Suppose Z_i and X_i are not correlated with each other, but both increase the probability of $D_i = 1$.

Conditional on D_i , Z_i and X_i may be negatively correlated:

- ▶ e.g. height (Z_i) and athletic ability (X_i) among basketball players who become pros $(D_i = 1)$
- ▶ e.g. low channel position of Fox (Z_i) and political conservatism (X_i) among people who watch Fox News $(D_i = 1)$

Therefore, regressing some Y on D_i and Z_i (but not X_i), coefficient on Z_i is contaminated by effect of omitted X_i on Y_i .

You could find an effect of Z_i even if the exclusion restriction actually holds (Gerber & Green pg. 199).

Can we test the exclusion restriction? (2)

Two main strategies:

Introduction

000

- ▶ placebo population test of the reduced form: move the (reduced-form) analysis to a different population in which the instrument should not affect the treatment, show zero estimated (reduced form) effect (e.g. Acharya, Blackwell, and Sen 2016 JOP)
- ▶ placebo outcome test of the first stage: replace the treatment with something that should not be affected by instrument but may suffer from same OVB (e.g. lagged treatment), show zero estimated first-stage effect (e.g. Meredith 2013 APSR)

Think about applying to the Colantone and Stanig research?

Examples

Further thoughts on IV

- ► Could you use IVs in your own research? Examples.
- ▶ A good instrument is hard to find another reason to start by looking for randomness.
- ► In observational studies, a variable that satisfies **independence** (CIA) is a rare and wonderful thing. Usually **exclusion** is doubtful, but you can gauge its effect and speculate about channels.
- Now that you know about instrumental variables, you should always define what you refer to as "IV": instead, you could say "treatment", "control variable", "covariate", "regressor"
- ▶ The recent econometrics literature is very skeptical of the internal validity of IVs, see for example: Lal, Apoorva, Mackenzie William Lockhart, Yiqinq Xu, and Ziwen Zu. (2021) "How Much Should We Trust Instrumental Variable Estimates in Political Science? Practical Advice based on Over 60 Replicated Studies."