Instrumental Variables

IV Frontiers

MIXTAPE SESSIONS

Roadmap

- Judge/Examiner IV
  - Approach
  - Cautions
- Shift-Share IV
  - Approach
  - Cautions
- Other Frontiers
  - Diff-in-Diff IV
  - Recentered IV

Approach

A judge (or examiner) IV design leverages the ideosyncratic assignment of individuals to a set of decision-makers

- Kling (2006): sentencing judges
- Doyle (2007): foster care investigators
- Mayestas et al. (2013): SSDI benefit examiners
- Doyle et al. (2015): ambulance companies

Approach

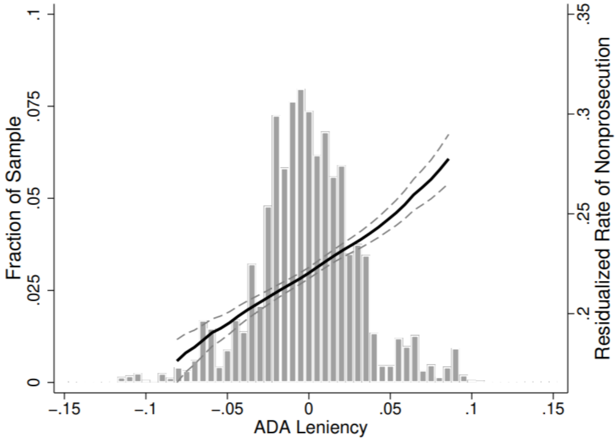
A judge (or examiner) IV design leverages the ideosyncratic assignment of individuals to a set of decision-makers

- Kling (2006): sentencing judges
- Doyle (2007): foster care investigators
- Mayestas et al. (2013): SSDI benefit examiners
- Doyle et al. (2015): ambulance companies

The typical approach is to IV a treatment D_i with a measure of the "leniency" $E[D_i \mid Z_i]$ of one's assigned judge $Z_i = \in \{1, \ldots, J\}$

- E.g. a leave-one-out average, $\hat{L}_i = \frac{1}{|i' \neq i, Z_{i'} = Z_i|} \sum_{i' \neq i, Z_{i'} = Z_i} D_i$

Agan et al. (2021) “Misdemeanor Prosecution”



Note: This figure shows the distribution of our leave-out mean measure of ADA “leniency,” residualized by court-by-month and court-by-day-of-week. More lenient ADAs have higher rates of not prosecuting nonviolent misdemeanor cases. The solid line is a local linear regression of nonprosecution on ADA leniency, along with the 95% confidence interval, estimated from the 1st to 99th percentiles of ADA leniency—a local linear version of our first stage. A case assigned to a more lenient ADA (computed using all cases except the current case and other cases with the same defendant) has a higher likelihood of being not prosecuted.

Agan et al. (2021) “Misdemeanor Prosecution”

	(1) Nonprosecution	(2) ADA Leniency
Number Counts	-0.019***	-0.000
	(0.003)	(0.000)
Number Misdemeanor Counts	0.018***	0.000
	(0.004)	(0.001)
Number of Serious Misdemeanor Counts	-0.102***	-0.000
	(0.006)	(0.000)
Misd Conviction within Past Year	-0.068***	-0.001
	(0.005)	(0.000)
Felony Conviction within Past Year	-0.053***	-0.001
	(0.006)	(0.001)
Citizen	0.042***	-0.000
	(0.004)	(0.000)
Disorderly/Theft	-0.014*	-0.001
	(0.008)	(0.001)
Motor Vehicle	0.105***	-0.000
	(0.009)	(0.000)
Drug	-0.094***	-0.001
	(0.009)	(0.001)
Constant	0.224***	0.001
	(0.009)	(0.002)
Observations	67553	67553
Joint F-Test p-value	0	0.234

Note: This table reports regressions testing the random assignment of cases to arraigning ADAs. ADA leniency is estimated using data from other nonviolent misdemeanor cases assigned to an arraigning ADA following the procedure described in the text. Column (1) reports estimates from an OLS regression of nonprosecution on the variables listed and court-by-time fixed effects. Column (2) reports estimates from an OLS regression of ADA leniency on the variables listed and court-by-time fixed effects. Robust standard errors two-way clustered at the individual and ADA level are reported in parentheses. The p-value reported at the bottom of Columns (1) and (2) is for an F-test of the joint significance of the variables listed with standard errors two-way clustered at the individual and ADA level. $^{***}p < 0.01$,$^{**}p < 0.05$,$^{*}p < 0.10$.

Agan et al. (2021) "Misdemeanor Prosecution"

	OLS		IV	
	(1)	(2)	(3)	(4)
Panel A: Criminal Complaint Within 2 Years				
Not Prosecuted	-0.14***	-0.10***	-0.34***	-0.33***
	(0.01)	(0.01)	(0.10)	(0.11)
			[-0.55, -0.13]	[-0.54, -0.10]
Mean Dep Var Prosecuted	0.37			
Mean Dep Var Prosecuted Compliers	0.57			
Panel B: Misdemeanor Complaint Within 2 Years				
Not Prosecuted	-0.08***	-0.06***	-0.24***	-0.24***
	(0.00)	(0.00)	(0.09)	(0.09)
			[-0.42, -0.06]	[-0.43, -0.05]
Mean Dep Var Prosecuted	0.24			
Mean Dep Var Prosecuted Compliers	0.40			
Panel C: Felony Complaint Within 2 Years				
Not Prosecuted	-0.06***	-0.04***	-0.10*	-0.08
	(0.00)	(0.00)	(0.06)	(0.07)
			[-0.22, 0.03]	[-0.21, 0.06]
Mean Dep Var Prosecuted	0.13			
Mean Dep Var Prosecuted Compliers	0.17			
Observations	67553	67553	67553	67553
Court x Time FE	Yes	Yes	Yes	Yes
Case/Def Covariates	No	Yes	No	Yes

Note: This table reports OLS and two-stage least squares estimates of the impact of nonprosecution on the probability of a subsequent criminal complaint within two years. The regressions are estimated on the sample as described in the notes to Table 1. The dependent variables are identified in the panel headings. Each panel reports the mean of the dependent variable for all prosecuted defendants, and for prosecuted defendants within the set of compliers. See Appendix C.3 for details on the calculation of mean outcomes among prosecuted compliers. Two-stage least squares models instrument for nonprosecution using an ADA leniency measure that is estimated using data from other cases assigned to an arraigning ADA following the procedure described in the text. All specifications control for court-by-month and court-by-day-of-week fixed effects. Robust standard errors two-way clustered at the individual and ADA level are reported in parentheses in Columns (1)-(4). For the IV estimates, confidence intervals based on inversion of the Anderson-Rubin test are shown in brackets. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Caution: Monotonicity

"Strict" first-stage monotonicity requires judges to have a common ordering of individuals for treatment

- E.g. no differences in "skill" at identifying appropriate cases

Caution: Monotonicity

"Strict" first-stage monotonicity requires judges to have a common ordering of individuals for treatment

- E.g. no differences in "skill" at identifying appropriate cases

Imbens and Angrist (& Ridder) saw this coming in 1994!

EXAMPLE 2 (Administrative Screening):[5] Suppose applicants for a social program are screened by two officials. The two officials are likely to have different admission rates, even if the stated admission criteria are identical. Since the identity of the official is probably immaterial to the response, it seems plausible that Condition 1 is satisfied. The instrument is binary so Condition 3 is trivially satisfied. However, Condition 2 requires that if official A accepts applicants with probability $P(0)$, and official B accepts people with probability $P(1) > P(0)$, official B must accept *any* applicant who would have been accepted by official A. This is unlikely to hold if admission is based on a number of criteria. Therefore, in this example we *cannot* use Theorem 1 to identify a local average treatment effect nonparametrically despite the presence of an instrument satisfying Condition 1.

[5] This example was suggested to us by Geert Ridder.

Monotonicity Solutions

Frandsen et al. (2019) formalize a weaker “average monotonicity” condition: intuitively, that skill differences are uncorrelated with TEs

- Similar to de Chaisemartin (2017) “tolerating defiance”
- Also propose non-parametric tests of monotonicity + exclusion (similar to Kitagawa (2015), but with multiple IVs + controls)

Monotonicity Solutions

Frandsen et al. (2019) formalize a weaker “average monotonicity” condition: intuitively, that skill differences are uncorrelated with TEs

- Similar to de Chaisemartin (2017) “tolerating defiance”
- Also propose non-parametric tests of monotonicity + exclusion (similar to Kitagawa (2015), but with multiple IVs + controls)

Other tests include checking whether leniency has the same first stage in different subgroups (Norris, 2021)

- Another solution is to parameterize variation in judge skill and estimate it jointly with TEs (Chan et al. 2021; Arnold et al. 2021)

Caution: Exclusion

“Strict” exclusion requires judges to only affect the outcome through one treatment channel

- E.g. a judge more likely to sentence a defendant to jail does not differentially change sentence conditions

Caution: Exclusion

“Strict” exclusion requires judges to only affect the outcome through one treatment channel

- E.g. a judge more likely to sentence a defendant to jail does not differentially change sentence conditions

Like monotonicity, this can be weakened to an “on average” condition

- Kolesár et al. (2015): exclusion restriction violations are uncorrelated with leniency variation (see also Angrist et al. 2021)
- Need many judges for a “judge-level law of large numbers” to kick in

Adding Treatment Channels

Of course if multiple potential treatment channels are observed they can be included + instrumented by judges

- See Autor/Maestas/Mullen/Strand (2017), which adds a decision-time treatment to Maestas et al. (20913)
- Two instruments: examiner leniency and (leave-out) average examiner decision time

Adding Treatment Channels

Of course if multiple potential treatment channels are observed they can be included + instrumented by judges

- See Autor/Maestas/Mullen/Strand (2017), which adds a decision-time treatment to Maestas et al. (20913)
- Two instruments: examiner leniency and (leave-out) average examiner decision time

Caution though: IV with multiple treatments can be difficult to interpret in a LATE framework (maybe OK as a robustness check)

Caution: Leniency Estimation

stuff on leniency estimation

Roadmap

Judge/Examiner IV

- Approach
- Cautions

Shift-Share IV

- Approach
- Cautions

Other Frontiers

- Diff-in-Diff IV
- Recentered IV

Approach

How shift-share IV works

Cautions

Shift-share IV cautions

Roadmap

Judge/Examiner IV

- Approach
- Cautions

Shift-Share IV

- Approach
- Cautions

Other Frontiers

- Diff-in-Diff IV
- Recentered IV

Diff-in-Diff IV

Stuff about DDIV

Recentered IV

Stuff about Recentered IV