MY457/MY557: Causal Inference for Observational and Experimental Studies

Week 10: Instrumental Variables 2

Daniel de Kadt

Department of Methodology
LSE

Winter Term 2024

Course Outline

- **Week 1:** The potential outcomes framework
- **Week 2:** Randomized experiments
- **Week 3:** Selection on observables I
- **Week 4:** Selection on observables II
- **Week 5:** Selection on observables III
- Week 6: Reading week
- **Week 7:** Difference-in-differences I
- **Week 8:** Difference-in-differences II
- **Week 9:** Instrumental variables I
- **Week 10:** Instrumental variables II
- **Week 11:** Regression discontinuity

Today

Table of Contents

IV with a Continuous Treatment

So far we have formalised cases with binary D and binary Z.

Often we have continuous versions. Focus on continuous D.

$\rightsquigarrow$ for continuous Z, LATE interpretation roughly applies.

In this setting, we can recast our unit-level potential outcomes for the outcome variable as:

$$Y_{S_{Z_i} i} \equiv f_i(s)$$

where S_i is a continuous treatment $S_i \in \{1, \ldots, \bar{s}\}$.

Then our Wald estimator can be re-written as estimating the Average Causal Response (ACR):

$$\tau_{ACR} = \frac{\mathbb{E}(Y_i \mid Z_i = 1) - \mathbb{E}(Y_i \mid Z_i = 0)}{\mathbb{E}(S_i \mid Z_i = 1) - \mathbb{E}(S_i \mid Z_i = 0)}$$

What does this represent?

IV with a Continuous Treatment

Under the typical IV assumptions:

$$
\begin{aligned}
\tau_{ACR} &= \frac{\mathbb{E}(Y_i \mid Z_i = 1) - \mathbb{E}(Y_i \mid Z_i = 0)}{\mathbb{E}(S_i \mid Z_i = 1) - \mathbb{E}(S_i \mid Z_i = 0)} \\
&= \sum_{s=1}^{\bar{s}} w_s \times \mathbb{E}[Y_{si} - Y_{s-1,i} \mid S_{1i} \geq s > S_{0i}]
\end{aligned}
$$

where

$$
w_s = \frac{\Pr[S_{1i} \geq s > S_{0i}]}{\sum_{j=1}^{\bar{s}} \Pr[S_{1i} \geq j > S_{0i}]}
$$

Read: The ACR is a weighted average of causal responses along the span of S, with up-weighting parts where compliance is high.

Interpretation: Each causal response is itself the LATE for compliers who move from a treatment intensity lower than s to at least s.

Table of Contents

Motivating Example: Examiner Instruments

Does short-term incarceration affect downstream political participation?

Any naïve study of this would be seriously confounded:

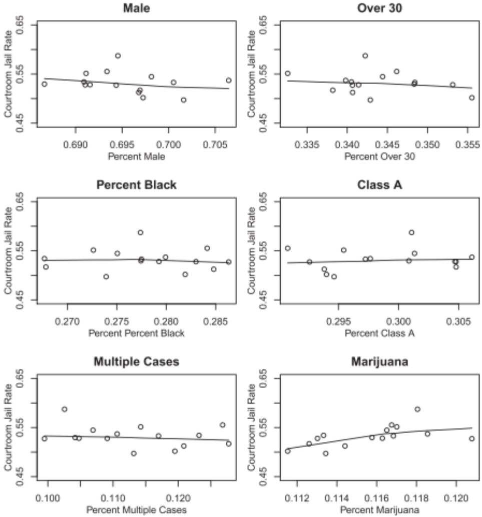
- Generally, those who spend time in jail are dissimilar to those who do not.
- Could study only those prosecuted, using legal guilt as instrument. But juries (determining legal guilt) are likely systematically biased.
- Could study only those found guilty, studying jail-time directly. But judges (sentencing people to jail-time) are likely systematically biased too.

White (2019) studies Harris County (Texas) where:

- First time defendants are randomly assigned to courtrooms
- Courtrooms differ in the propensity for defendants to serve jail-time

Motivating Example: Examiner Instruments

FIGURE 1. Scatterplots of Pre-treatment Case Characteristics Against Courtroom Incarceration Rates.



Setup: Examiner Instruments

Let's generalise this research setting:

- N units (typically individuals), indexed by $i = 1, \ldots, N$
- K examiners who have control over...
- ... our (binary, for now) treatment status $D_i \in \{0, 1\}$
- which (we believe) affects outcome of interest Y_i

Each i is assigned an examiner in a known way s.t. they receive status $Q_i \in \{0, \ldots, K-1\}$, from which we can write $Z_{ki} = \mathbf{1}[Q_i = k]$

Given this we can define potential outcomes:

- D_{qi}: p.o. of treatment for i under examiner status q
- Y_{qi}: p.o. of outcome for i under examiner status q
- Y_{dqi}: p.o. of outcome for i under treatment d and examiner status q

This is now a bit different to anything we've seen before. We have high-dimensional instrument(s) for a single treatment.

Identification Assumptions

What do our IV assumptions look like in the examiner design?

1. SUTVA
2. Relevance: Does every examiner assign D differently?
3. Ignorability: May be given, but often only holds conditionally.
4. Exclusion: What else can examiners control?

   $\rightsquigarrow$ Can be weakened to an on average exclusion restriction, s.t. any direct examiner effects are independent of their first-stage effects.
5. Monotonicity: Examiner behaviour must be ordered.

   $\rightsquigarrow$ If Examiner 1 more lenient on average, must be weakly more lenient for every unit!

   $\rightsquigarrow$ Example violation would be differential weights on unit-characteristics, such as examiner-varying racial or gender biases.

   $\rightsquigarrow$ Can be weakened to average monotonicity: units may violate monotonicity with some examiners, so long as they comply on average.

See Frandsen et al (2023) for testing monotonicity and exclusion.

Latent Estimation Approach

Intuitively, we want to use the fact that different examiners are different – that is, variation in treatment assignment between examiners.

Intuition is to estimate:

$$\hat{L_{ik}} = \frac{\sum_{j \neq i} \mathbf{1}[Q_j = Q_i] D_j}{\sum_{j \neq i} \mathbf{1}[Q_j = Q_i]}$$

Read: L_k is the latent examiner-specific treatment assignment, and $\hat{L_{ik}}$ is our estimate thereof ignoring the treatment outcome for unit i (we can actually drop the k indexing). This is called a leave-one-out procedure.

We then instrument for D_i using $\hat{L_{ik}}$. What might we be missing?

1. Under what constraints is examiner assignment random?
2. How variable are estimates of $\hat{L_{ik}}$?

Examiner Fixed-Effects Approach

Observe that L_k is an examiner-specific time-invariant property. L_k is thus subsumed under examiner fixed-effects.

Can use $Z_1, \ldots, Z_k$ as multiple dummy variable instruments in first stage.

$\rightsquigarrow$ This is over-identified IV – we have more instruments than endogenous regressors (contrast with under- or just-identified IV).

Include whatever other covariates or fixed-effects necessary as controls.

$\rightsquigarrow$ In the latent estimation approach, these need to be dealt with when estimating the latent property, resolving issue 1 on the previous slide.

Estimation using Jacknife Instrumental Variables Estimation (JIVE) (Angrist et al, 1999) or UJIVE (Kolesar, 2013), leave-one-out estimators.

Face a many weak instruments problem: With many examiners we have many instruments, some of which may be weak $\rightsquigarrow$ significant bias.

Data for these two approaches might look something like this:

i	Y_i	D_i	Q_i	$\hat{L_{ik}}$	K_i^1	K_i^2	$\dots$	K_i^K
1	1	1	1	$\hat{l_{11}}$	1	0	$\dots$	0
2	0	1	2	$\hat{l_{22}}$	0	1	$\dots$	0
3	1	0	1	$\hat{l_{31}}$	1	0	$\dots$	0
4	1	1	K	$\hat{l_{4K}}$	0	0	$\dots$	1
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
N	0	1	1	$\hat{l_{N1}}$	1	0	$\dots$	0

Overidentification here is given by the fact that all K_i^1 through K_i^K variables (or dimensions) are jointly determining a unit's D_i status. That is, there is a many-to-one instrument-to-endogenous-regressor mapping.

Back to the Motivating Example: Estimation and Results

White (2019) uses the first approach (latent estimation):

1. Estimate courtroom-year average incarceration rate
   (in appendix shows courtroom-year FE approach)
2. 2SLS model with courtroom-year rate as instrument and year-dummies
3. Subset by race of defendant for heterogeneous effects

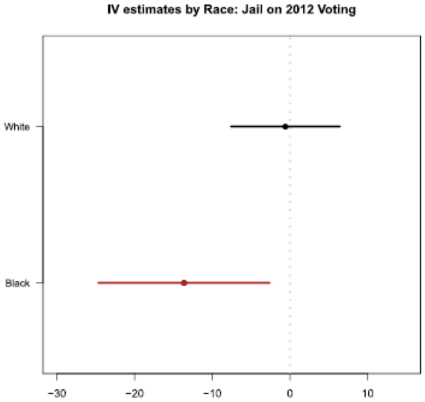
TABLE 3. Jail Sentences on 2012 Voting

	Dependent variable	
	Jail (1)	Voted 2012 (2)
Court jail average (Yr)	1.000*	
	(0.051)	
Jail		−0.045
		(0.034)
Constant	−0.0001	0.142*
	(0.029)	(0.019)
Year dummies	Yes	Yes
Observations	113,367	113,367
Adjusted R^2	0.004	0.017
F statistic	97.948* (df = 5; 113,361)	

Note: $^*p < 0.05$.

Back to the Motivating Example: Heterogeneous Effects

FIGURE 2. Jail's Effect on Voter Turnout (2SLS Estimates), by Race of Defendant.



Note: A coefficient of −0.13 indicates a turnout decrease of 13 percentage points (among compliers).

Table of Contents

A Motivating Example: Shift-Share Instruments

How does in-migration of new minority groups affect social position of pre-existing minority groups?

Fouka, Mazumder, & Tabellini (2022) study the context of the First Great Migration in the USA at the MSA-level:

- 1915 - 1930, 1.5m Black Americans left Southern for Northern states.
- Did this affect white Americans' views of European immigrants?

Much like Ananat (2011), many potential confounders exist. Black Americans may move to...

- ... more liberal MSAs
- ... wealthier MSAs
- ... etc.

Authors use a shift-share instrumental variables (SSIV) approach.

Setup: The Endogenous Model

Let's map the previous example to a simple model where $l = 1, \ldots, L$ are units (labour markets, areas, MSAs etc.):

$$y_l = \alpha + \beta x_l + \varepsilon_l$$

β is the effect of x_l on y_l, but assume it is confounded ($\mathbb{E}[x_l \varepsilon_l] \neq 0$).

$\rightsquigarrow$ We are using the classical IV approach here. Both ignorability and exclusion fall under a single 'valid instrument' assumption. For instrument z_i, satisfied if $\mathbb{E}[z_i \varepsilon_i] = 0$, equivalently $Cov(z_i, \varepsilon_i) = 0$ or $\mathbb{E}[\varepsilon_i \mid z_i] = 0$.

Now imagine we had a second dimension of observation for our units, across $k = 1, \ldots, K$ which are types (industry, population groups, migrants from specific places, etc.).

Each unit l can be observed for all or multiple k, so we have y_{lk} and x_{lk}.

Setup: Shift-Share Instruments

Given both y_{lk} and x_{lk}, we have two sources of potential variation:

1. Shares: A unit- and type-varying (lk-varying) baseline variable, z_{lk}
2. Shift(s): A type-varying (k-varying) change variable, g_k

Interacting the above gives us a shift-share variable: $S_l = \sum_k z_{lk} \times g_k$

Read: The combination of exposure (share) to a shock (shift), capturing unit-varying shock-exposures.

SSIV Intuition: Use this interaction to instrument for x_l.

Shift-share instruments are sometimes called Bartik instruments, after Bartik (1991).

Back to the Motivating Example: Shift-Share Instrument

Fouka, Mazumder, & Tabellini (2022) construct their instrument as:

1. Shares: z_{lk}, the pre-shift share of Black American migrants born in state k living in MSA l
2. Shift(s): g_k, the share of Black Americans born in state k who left
3. Shift-share: $S_l = \sum_k z_{lk} \times g_k$, the predicted post-shift sum of Black American migrants in MSA l

Read: Their instrument captures the predicted level of Black American in-migration in an MSA, based on some randomness in the timing of out-migration from different states.

They then use this to instrument for actual Black American in-migration by decade.

Back to the Motivating Example: First Stage

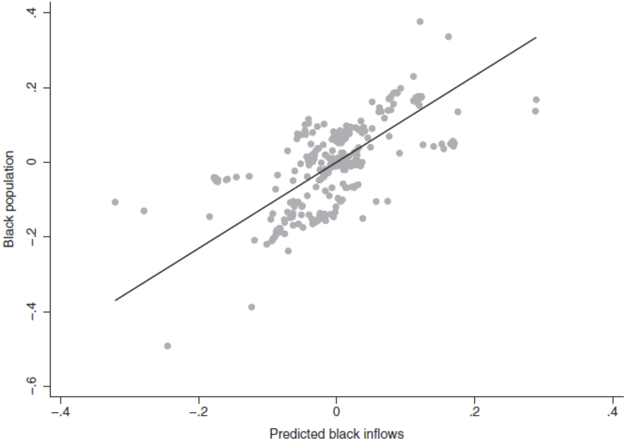


FIGURE 3

First stage

Notes: The figure shows the relationship between actual and predicted Black population for the years 1910–30. Each point represents the residual change in an MSA's actual (y-axis) and predicted (x-axis) number of Black migrants after partialling out total MSA population and MSA and region by decade fixed effects.

Back to the Motivating Example: Results

TABLE 2

Effects of Black inflows on intermarriage and naturalization rates

	(1) OLS	(2) OLS	(3) 2SLS	(4) 2SLS	(5) 2SLS	(6) 2SLS
	Panel A: Married w/ native *(1910 mean: 0.071)*					
Black population	0.007	0.005	0.012**	0.013***	0.012**	0.019
	(0.006)	(0.004)	(0.005)	(0.005)	(0.005)	(0.017)
Observations	9,323,126	9,323,126	9,323,126	9,323,109	9,323,109	88,892
F-stat			23.33	23.71	23.83	32.75
	Panel B: Naturalized *(1910 mean: 0.491)*					
Black population	0.056**	0.030*	0.033**	0.038***	0.033**	0.144***
	(0.025)	(0.016)	(0.015)	(0.013)	(0.014)	(0.031)
Observations	15,267,846	15,267,846	15,267,846	15,267,844	15,267,844	80,866
F-stat			24.23	24.38	24.51	32.74
Individual controls	X	X	X	X	X	
Region × Decade		X	X	X	X	
MSA × Origin region				X	X	
Origin region × Decade					X	
Linked sample						X

Notes: The table presents results for immigrant men living in the 108 non-southern MSAs for which the instrument could be constructed in census years 1910, 1920, and 1930. In Panel A, the sample is restricted to married men. *Married w/ native* is a dummy equal to 1 if the individual is married to a native-born spouse of native-born parentage. Columns 1–5 present results obtained from the repeated cross-sections, while column 6 shows results from the linked panel of men who always remained in the same MSA in the three census years. Columns 1–2 (resp. 3–6) present OLS (resp. 2SLS results). Individual controls include fixed effects for age, years in the U.S. and origin region. All regressions control for MSA and year fixed effects and for total MSA population. Regressions in column 6 include individual fixed effects. In column 6 of Panel A, the sample is restricted to men who were not married in the previous decade (Panel A) or who were not naturalized in the previous decade (Panel B). F-stat refers to the KP F-stat for weak instruments. Robust standard errors, clustered at the MSA level, in parentheses. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Identifying Assumptions: Shares Perspective

There are two broad perspectives on identification in shift-share settings.

First relies on properties of the shares (Goldsmith-Pinkham et al (2020)):

1. Relevance of the share:

   $\rightsquigarrow$ For all k, z_{lk} must predict x_l

2. Strict exogeneity of the share to the error term, $\mathbb{E}[z_{lk}\varepsilon_l] = 0$

   $\rightsquigarrow$ This could be a conditional statement, or indexed over time

   $\rightsquigarrow$ Intuition is that exogeneity is to changes, not the level of y (think diff-in-diff)

With these two assumptions (plus SUTVA), two stage least squares with the shift-share instrument is consistent.

Identifying Assumptions: Shifts Perspective

Second approach concerns shifts ('shocks' in Borusyak et al (2022)):

1. Exogeneity of the shocks to confounding in the shares, $\mathbb{E}[g_k|\bar{\varepsilon}z] = \mu$ for all k

   $\rightsquigarrow$ Every shock g_k has the same mean μ, regardless of observables ($\bar{\varepsilon}$) or shares (z). Essentially, shocks are as-if randomly assigned.
2. Many uncorrelated shocks: $\mathbb{E}[\sum_k z_k^2] \rightarrow 0$ and $Cov(g_k, g_{k'}) = 0$ for all (k, k') with $k \neq k'$.

   $\rightsquigarrow$ Intuition is that with very many exogenous shocks, residual confounding bias from shares averages out.

With these two assumptions plus relevance and SUTVA, two stage least squares is consistent.

Shifts or Shares?

Consider a few settings:

1. Cross-sectional (no time-variation), two types ($K = 2$):
   - $\rightsquigarrow$ Identification tends to rely on exogeneity of shares
   - $\rightsquigarrow$ Design is akin to a two-period first-differenced diff-in-diff
2. Multiple (T) time periods, two types:
   - $\rightsquigarrow$ Identification tends to rely on shares
   - $\rightsquigarrow$ Multi-period diff-in-diff concerns apply.
3. Many ($K > 2$) types, cross-sectional or multiple T:
   - $\rightsquigarrow$ For very many K, rely on exogeneity of shocks

See Goldsmith-Pinkham et al (2020) (focused on shares) and Borusyak et al. (2022) (focused on shifts) for various falsification tests.

In general, be explicit (in reasoning and language) about what variation you think facilitates identification.

Table of Contents

A Motivating Example: Recentered Instruments

Does distributing government funds affect election outcomes?

This is hard to study because... ok by now you're tired of me saying: confounding. But again, it's a problem. Targeted areas might be...

- ... wealthier (or not)
- ... politically aligned
- ... etc.

Gulotty & Strezhnev (2024) study the US Department of Agriculture's (USDA) Market Facilitation Program (MFP) and the 2020 US election.

In 2019, the MFP allocated $14bn to agricultural producers using a formula based on:

1. Measures of agricultural production $\leftarrow$ endogenous
2. Product-specific trade damage estimates $\leftarrow$ plausibly random?

Setup: Recentered IV

Consider a unit-varying z_i (a candidate treatment or instrument) that is generated by formula s.t. multiple inputs drive variation.

As in the SSIV case, some variation in z_i may be (as good as) random while some is systematic.

Borusyak & Hull (2023) propose purging systematic component via recentering:

1. Specify the formula that generates z_i precisely
2. Using knowledge of random component, define counterfactual shocks
3. Calculate $\mu_i =$ average z_i across simulations of counterfactual shocks

   $\rightsquigarrow$ this represents the systematic component

   $\rightsquigarrow$ can think of it as akin to a propensity score
4. Recenter z_i to create $\tilde{z}_i = z_i - \mu_i$
5. Instrument with $\tilde{z}_i$ (or use z_i directly, controlling for μ_i)
6. Use the counterfactual shocks for randomization inference

Back to the Motivating Example: Treatment via Formula

Recall Gulotty & Strezhnev (2024) wish to study the effect of MFP agricultural transfers to counties.

The authors are able to reconstruct the USDA formula for MFP as:

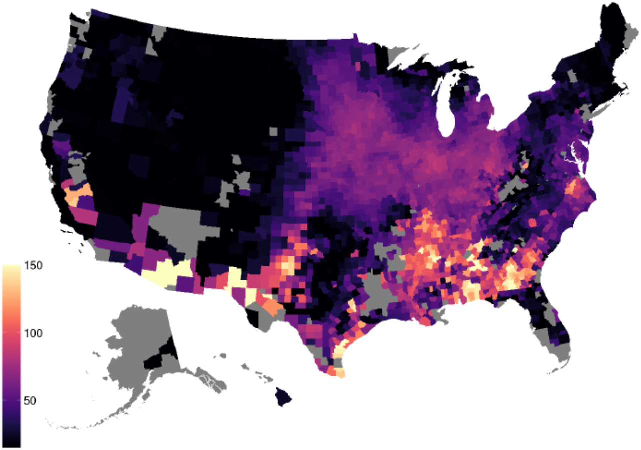
$$(\text{County Rate})_{\$/\text{acre}} = \frac{\sum_{c=1}^{C} (\text{County Acres})^{c}_{\text{acre}} \times (\text{County Yield})^{c}_{\text{unit/acre}} \times (\text{Crop Rate})^{c}_{\$/\text{unit}}}{\sum_{c=1}^{C} (\text{County Acres})^{c}_{\text{acre}}}$$

Where the critical component $(\text{Crop Rate})_{\$/unit}$ is given by:

$$(\text{Crop Rate})_{\$/\text{unit}} = \frac{(\text{Value of Exports})_{\$} \times (\%\text{ Reduction in exports})}{(\text{Annual Production})_{\text{unit}}}$$

The numerator are product-specific trade damage estimates which rely on historical export volatility, plausibly unconnected to domestic politics.

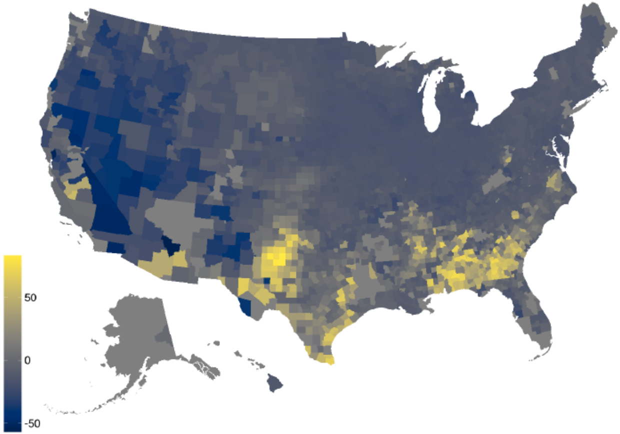
Back to the Motivating Example: Treatment via Formula



Notes: 2,892 counties. In the few cases where separate rates were calculated for parts of the same county, we combined the USDA's total acreage and payment calculations for each component to construct an "average" county rate.

Figure 1: Map of actual USDA 2019 non-specialty crop MFP payment rates (dollars per acre)

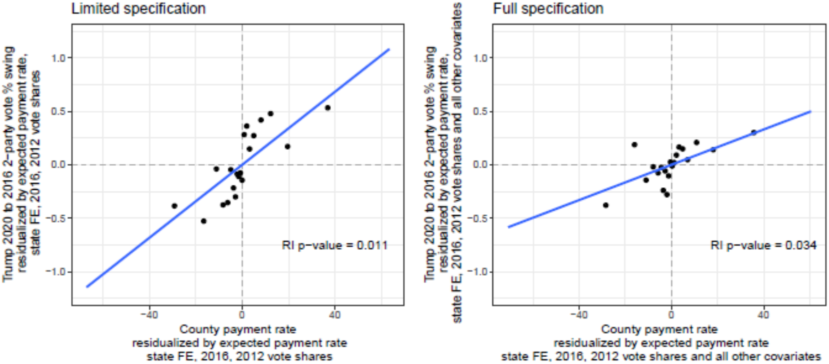
Back to the Motivating Example: Recentered Instrument



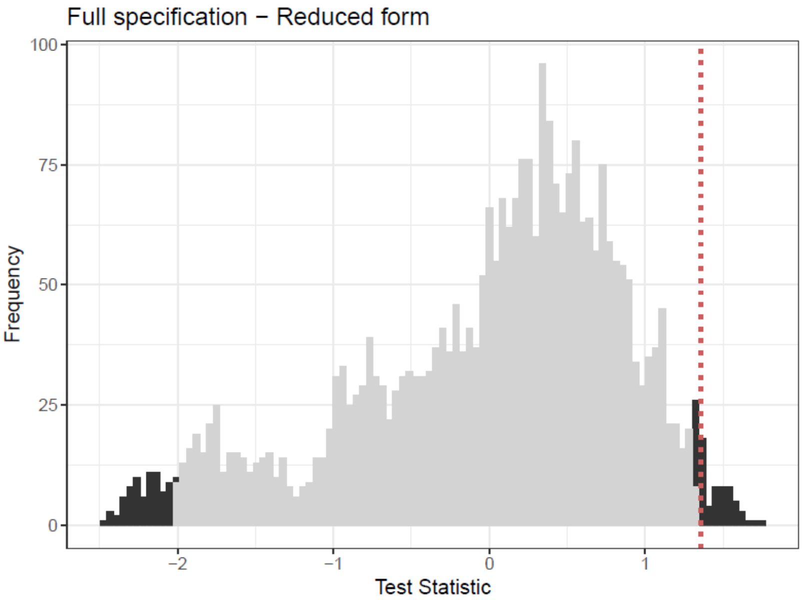
Notes: 2,877 counties with non-zero eligible acres based on FSA crop acreage data. "Expected" county rates calculated via 3000 random permutations of the per-acre commodity excess rate shocks.

Figure 4: Map of recentered county rates (dollars per acre)

Back to the Motivating Example: Results



Back to the Motivating Example: Randomization Inference



Recentered IV Identification Assumptions

For $i = 1, \ldots, N$, assume a simple model:

$$y_i = \beta x_i + \varepsilon_i$$

Consider a candidate instrument for x_i, $z_i = f_i(g, w)$, where:

- g is a vector of shocks
- w represents possibly endogenous variables (e.g. shares)
- $f_i(\cdot)$ maps the shocks and endogenous variables to z_i

  Note: Again, z_i need not be an instrument, could be a treatment

Now, assume:

1. Exogenous shocks s.t. $g \perp\!\!\!\perp \varepsilon \mid w$
2. Distribution $G(g|w)$ is known $\leftarrow$ this is important!

Recentered IV Identification Result

We can write down the expected value (average) of all potential shocks as:

$$\mu_i = \mathbb{E}[f_i(g, w)]$$

Under assumption 1, can show that:

$$\mathbb{E}\left[\frac{1}{N}\sum_i z_i \varepsilon_i\right] = \mathbb{E}\left[\frac{1}{N}\sum_i \mu_i \varepsilon_i\right]$$

And thus for the recentered instrument:

$$\mathbb{E}\left[\frac{1}{N}\sum_i \tilde{z}_i \varepsilon_i\right] = \mathbb{E}\left[\frac{1}{N}\sum_i z_i \varepsilon_i\right] - \mathbb{E}\left[\frac{1}{N}\sum_i \mu_i \varepsilon_i\right] = 0$$

Under assumption 1 β is identified by $\tilde{z}_i$ (assuming relevance of $\tilde{z}_i$), and under assumption 2 it can be unbiasedly estimated with data.

Thanks to Kiril Borusyak, Paul Goldsmith-Pinkham, and Peter Hull, whose publicly available materials I leaned heavily on for this week.