PS207 Quantitative Causal Inference, Fall 2016

Instrumental Variables

Matto Mildenberger

UC Santa Barbara

Special thanks to Chad Hazlett (UCLA) for select slides, used with permission

The Intuition for (Modern) IV

Basic intuition for all causal IV work can be thought about in terms of a randomized encouragement design.

Suppose you want to test an important hypothesis: in humans, does pizza consumption improves happiness?

- Can we just look at a sample and use $\mathbb{E}[happy|pizza] - \mathbb{E}[happy|no\ pizza]$?
- Turns out your IRB won't let you randomize, because denying pizza for the controls would be a human rights violation
- Instead, suppose we randomly assign half the people in our sample to receive gift cards for pizza.
- Assume we get to observe which ones actually get a pizza thereafter, and then measure happiness for everybody in our sample.

The Intuition for (Modern) IV

Good news: we have an experiment w.r.t. our gift card (aka encouragement, instrument)

- ATE_{card} is identified, $ATE_{card} = \mathbb{E}[happy|card] - \mathbb{E}[happy|no\ card]$
- This is often called the *reduced form* estimate or *intent to treat* effect

But how does this compare to actual causal effect of *pizza* on *happiness*?

Answer: We can back-out effect of the pizza, but only for certain sub-population

Intuition:

- suppose *card* impacts *happiness* only through *pizza* consumption
- our entire ITT estimate comes from just the sub-pop for whom *card* changed their *pizza* consumption.
- ATE_{card} is thus an average over the effect. of *card* on *happy* among that group, with zero effect from those not driven to change their pizza consumption
- to undo this "watering down", we'll see we want to divide the ATE_{card} by Pr(somebody ate more pizza because of the card)

The Intuition for (Modern) IV

To see this, let's consider four groups:

1. *Never-takers.* People with gluten allergies. Won't take the pizza whether they get card or not. What is $\mathbb{E}[Y|card] - \mathbb{E}[Y|no\ card]$ for this group?
2. *Compliers.* Those who will get the pizza if we give them the card, but not if we don't. Since they follow the encouragement, for them the encouragement is the treatment, so $\mathbb{E}[Y|card] - \mathbb{E}[Y|no\ card]$ *is* the ATE *for this group*.
3. *Always-takers* are those who will eat pizza regardless of getting the card. What is $\mathbb{E}[Y|card] - \mathbb{E}[Y|no\ card]$ for this group?
4. *Defiers* are those who feel like eating pizza if they don't get a card but when you give them the card, feel like eating salad. We are just going to assume these people don't exist.

The Intuition for (Modern) IV

The ATE_{card} is averaged over all units,

$$\begin{aligned} ATE_{card} &= (ATE_{card}|\text{never-taker}) \cdot Pr(\text{never-taker}) \\ &+ (ATE_{card}|\text{complier}) \cdot Pr(complier) \\ &+ (ATE_{card}|\text{always-taker}) \cdot Pr(\text{always-taker}) \\ &+ (ATE_{card}|\text{defier}) \cdot Pr(\text{defier}) \end{aligned}$$

$$\frac{ATE_{card}}{Pr(\text{complier})} = (ATE_{card}|\text{complier})$$

Note: $Pr(complier)$ is "effect of instrument on treatment uptake"...

Because instrument is random, it's effect on *pizza* is identified:

$$\text{“}Complier\ Ratio\text{''} = \mathbb{E}[D|Z=1] - \mathbb{E}[D|Z=0]$$

where Z is encouragement (card), D is treatment (pizza)

The Intuition for (Modern) IV

Altogether, we can estimate $\mathbb{E}[\tau|\textit{complier}]$, aka the LATE (local average treatment effect):

$$\begin{aligned} \textit{LATE} &= \frac{\text{Effect of Z on Y}}{\text{Effect of Z on D}} \\ &= \frac{\mathbb{E}[Y|Z=1] - \mathbb{E}[Y|Z=0]}{\mathbb{E}[D|Z=1] - \mathbb{E}[D|Z=0]} \\ &= \frac{cov(Z,Y)}{cov(Z,D)} \end{aligned}$$

This is the IV, 2SLS, or Wald Estimator

The Intuition for (Modern) IV

Lot's of things can go wrong:

- We had to assume no effect of Z on Y except through D.
  Exclusion restriction. (What would go wrong otherwise?)
- We had to assume Z effects D.
  Relevance. (What would go wrong otherwise?)
  - Moreover, Z cannot affect D too weakly. (Why?)
- We had to assume Z was random in to begin with
  Ignorability/ Random first stage. (What would go wrong otherwise?)
- We had to assume no defiers
  Monotonicity. (What would go wrong otherwise?)
- If everything goes right, is the effect you estimate representative?

Good habit to think through any IV problem in this "randomized encouragements" setup

Two Views on Instrumental Variables

1. Traditional Econometric Framework
   - Constant treatment effects
   - Linearity in the case of a multivalued treatment

2. Potential Outcome Model of IV
   - Acknowledges heterogeneous treatment effects
   - Most natural for binary treatments
   - Focuses on Local Average Treatment Effect (LATE)

Identification with Traditional Instrumental Variables

Definition

Two equations:

- $Y = \gamma + \alpha D + \varepsilon$ (Second Stage)
- $D = \tau + \rho Z + \eta$ (First Stage)

Identification Assumption

1. *Exogeneity and Exclusion:* $Cov(Z, \eta) = 0$ *and* $Cov(Z, \varepsilon) = 0$
2. *First Stage:* $\rho \neq 0$
3. $\alpha = Y_{1,i} - Y_{0,i}$ *constant for all units* i.
   (or a constant linear effect for multivalued treatments)

Potential Outcome Model for Instrumental Variables

Definition (Instrument)

Z_i: Binary instrument for unit i.

$$Z_i = \begin{cases} 1 & \text{if unit } i \text{ “encouraged” to receive treatment} \\ 0 & \text{if unit } i \text{ “encouraged” to receive control} \end{cases}$$

Definition (Potential Treatments)

D_{zi} indicates potential treatment status given $Z_i = z$

- e.g. if $D_{1i} = 1$ encouraged to take treatment and takes treatment

Assumption

Observed treatments are realized as

$$D_i = Z_i \cdot D_{1i} + (1 - Z_i) \cdot D_{0i} \text{ so } D_i = \begin{cases} D_{1i} & \text{if } Z_i = 1 \\ D_{0i} & \text{if } Z_i = 0 \end{cases}$$

Potential Outcome Model for Instrumental Variables

Following Angrist, Imbens, and Rubin (1996), we can define:

Definition

- Compliers: $D_{1i} > D_{0i}$ ($D_{0i} = 0$ and $D_{1i} = 1$).
- Always-takers: $D_{1i} = D_{0i} = 1$.
- Never-takers: $D_{1i} = D_{0i} = 0$.
- Defiers: $D_{1i} < D_{0i}$ ($D_{0i} = 1$ and $D_{1i} = 0$).

Problem

Only one of the potential treatment indicators (D_{0i}, D_{1i}) *is observed, so we cannot identify which group any particular individual belongs to*

Who are the Compliers?

Study	Outcome	Treatment	Instrument
Angrist and Evans (1998)	Earnings	More than 2 Children	Multiple Second Birth (Twins)
Angrist and Evans (1998)	Earnings	More than 2 Children	First Two Children are Same Sex
Levitt (1997)	Crime Rates	Number of Policemen	Mayoral Elections
Angrist and Krueger (1991)	Earnings	Years of Schooling	Quarter of Birth
Angrist (1990)	Earnings	Veteran Status	Vietnam Draft Lottery
Miguel, Satyanath and Sergenti (2004)	Civil War Onset	GDP per capita	Lagged Rainfall
Acemoglu, Johnson and Robinson (2001)	Economic performance	Current Institutions	Settler Mortality in Colonial Times
Cleary and Barro (2006)	Religiosity	GDP per capita	Distance from Equator

Potential Outcome Model for Instrumental Variables

Definition (Potential Outcomes)

Given the binary instrument $Z_i \in (1, 0)$ and the binary treatment $D_i \in (1, 0)$ every unit now has four potential outcomes $Y_i(D, Z)$:

- $Y_i(D_i = 1, Z_i = 1)$; $Y_i(D_i = 1, Z_i = 0)$; $Y_i(D_i = 0, Z_i = 1)$; $Y(D_i = 0, Z_i = 0)$

Assumption (1. Ignorability of the Instrument)

$$(Y_{0i}, Y_{1i}, D_{0i}, D_{1i}) \perp\!\!\!\perp Z_i$$

1. *Independence:* $(Y_i(d, z), D_{1i}, D_{0i}) \perp\!\!\!\perp Z_i$ *for all* d, Z*. Implies identifiability of effects of* Z_i *on* Y_i *and* Z_i *on* D_i*.*
2. *Exclusion implied by* $(Y_{0i}, Y_{1i}) \perp\!\!\!\perp Z_i$*.*

   Note, $Y_i(d, 0) = Y(d, 1)$ *for all* $d \rightarrow$ *we can index potential outcomes solely by treatment status:* (Y_{1i}, Y_{0i})

Potential Outcome Model for Instrumental Variables

Estimand (LATE)

$$\alpha_{LATE} = \mathbb{E}[Y_1 - Y_0 | D_1 > D_0]$$

is the Local Average Treatment Effect among compliers.

NB: this estimand varies with the particular instrument Z.

Some special cases:

- When the treatment intake, D, is itself randomized,then $Z = D$ and every individual is a complier
- If one-sided noncompliance s.t. $D_{0i} = 0$:

$$\mathbb{E}[Y_1 | D_1 > D_0] = \mathbb{E}[Y_1 | Z = 1, D_1 = 1] = \mathbb{E}[Y_1 | D = 1]$$

And

$$\mathbb{E}[Y_0 | D_1 > D_0] = \mathbb{E}[Y_0 | D = 1]$$

Thus,

$$\alpha_{LATE} = \mathbb{E}[Y_1 - Y_0 | D_1 > D_0] = \mathbb{E}[Y_1 - Y_0 | D = 1] = \alpha_{ATET}$$

Identification with Instrumental Variables

Identification Assumption

1. *Ignorability of the Instrument:* $(Y_{0i}, Y_{1i}, D_{0i}, D_{1i}) \perp\!\!\!\perp Z_i$
2. *First Stage:* $P(D_{1i} = 1) \neq P(D_{0i} = 1)$
3. *Monotonicity:* $D_{1i} \geq D_{0i}$ *for all* i *(no defiers)*

Identfication Result

$$\begin{aligned}\mathbb{E}[Y_1 - Y_0 | D_1 > D_0] &= \frac{\mathbb{E}[Y|Z=1] - \mathbb{E}[Y|Z=0]}{\mathbb{E}[D|Z=1] - \mathbb{E}[D|Z=0]} \\ &= \frac{\textit{Intent to Treat Effect of } Z \textit{ on } Y}{\textit{First Stage Effect of } Z \textit{ on } D} \\ &= \frac{\textit{Intent to Treat Effect}}{\textit{Proportion of Compliers}}\end{aligned}$$

Identification with Instrumental Variables

Identification Assumption

1. *Ignorability of the Instrument:* $(Y_0, Y_1, D_0, D_1) \perp\!\!\!\perp Z$
2. *First Stage:* $0 < P(Z = 1) < 1$ *and* $P(D_1 = 1) \neq P(D_0 = 1)$
3. *Monotonicity:* $D_1 \geq D_0$

Proof.

$$\frac{E[Y|Z=1] - E[Y|Z=0]}{E[D|Z=1] - E[D|Z=0]} = \frac{E[Y_0 + (Y_1 - Y_0)D_1|Z=1] - E[Y_0 + (Y_1 - Y_0)D_0|Z=0]}{E[D_1|Z=1] - E[D_0|Z=0]}$$

$$= \frac{E[Y_0 + (Y_1 - Y_0)D_1] - E[Y_0 + (Y_1 - Y_0)D_0]}{E[D_1] - E[D_0]} = \frac{E[(Y_1 - Y_0)(D_1 - D_0)]}{E[D_1 - D_0]}$$

$$= \frac{E[Y_1 - Y_0|D_1 > D_0]P(D_1 > D_0) - E[Y_1 - Y_0|D_1 < D_0]P(D_1 < D_0)}{E[D_1 - D_0]} \text{ as } (D_1 - D_0)=(1, 0, -1)$$

$$= \frac{E[Y_1 - Y_0|D_1 > D_0]P[D_1 > D_0]}{P(D_1 > D_0)} = E[Y_1 - Y_0|D_1 > D_0]$$

□

Comments on Identification Assumptions

- Ignorability of the Instrument: $(Y_{0i}, Y_{1i}, D_{0i}, D_{1i}) \perp\!\!\!\perp Z_i$
  - Implies that Z_i is randomly assigned so that the intent to treat effect and first stage effect are causally identified
  - If Z_i effected outcome through channel other than treatment, it would change Y_{0i} and Y_{1i}. So ignorabilty implies exclusion too $(Y(d,0) = Y(d,1) = Y_d)$
  - Exclusion needed to attribute effect of Z on Y to the effect of D alone. Never fully testable!
  - Random assignment is a not a sufficient condition for exclusion. (Why not?)
- First Stage/ Relevance
  - Implies that the instrument Z induces variation in D
  - Fully testable by regressing D on Z
- Monotonicity: $D_1 \geq D_0$
  - Rules out defiers
  - Not testable but often verifiable with qualitative information

Instrumental Variable: Estimators

Estimand (LATE)

$$\mathbb{E}[Y_1 - Y_0 | D_1 > D_0] = \frac{\mathbb{E}[Y|Z=1] - \mathbb{E}[Y|Z=0]}{\mathbb{E}[D|Z=1] - \mathbb{E}[D|Z=0]}$$

Estimator (Wald Estimator)

The sample analog estimator is:

$$\left(\frac{\sum_{i=1}^N Y_i Z_i}{\sum_{i=1}^N Z_i} - \frac{\sum_{i=1}^N Y_i (1-Z_i)}{\sum_{i=1}^N (1-Z_i)} \right) \Bigg/ \left(\frac{\sum_{i=1}^N D_i Z_i}{\sum_{i=1}^N Z_i} - \frac{\sum_{i=1}^N D_i (1-Z_i)}{\sum_{i=1}^N (1-Z_i)} \right)$$

Instrumental Variable: Estimators

Estimator (Two-Stage Least Squares)

Can also implement Wald Estimator using two-stage least squares regression:

1. $D = \pi_0 + \pi Z + \eta$ *(First Stage)*
2. $Y = \beta_0 + \beta D + \varepsilon$ *(Second Stage)*

Implies

$$\begin{aligned} Y &= \beta_0 + \beta(\pi_0 + \pi Z + \eta) + \varepsilon \\ &= (\beta_0 + \beta\pi_0) + \beta\pi Z + (\beta\eta + \varepsilon) \end{aligned}$$

where $cov(\varepsilon, Z) = 0$ *due to exclusion and* $cov(\eta, Z) = 0$ *by exogeneity in first stage*

- β *gives the LATE*
- *If you just regressed* Y *on* Z*, you'd get* $\beta\pi$*. Divide this by* π *and you get* β
- *Don't actually estimate this way, you will get wrong SEs*

Instrumental Variable: Estimators

Estimator (IV Estimator)

Standard implementation is the IV estimator:

$$\hat{\beta}_{IV} = (\boldsymbol{Z}^\top \boldsymbol{D})^{-1} \boldsymbol{Z}^\top \boldsymbol{Y}$$

1. *Typically estimate using ivreg()*
2. *If ID assumptions only hold after conditioning on X, you can add covariates:*
   - *they are added to both stages*
   - $\boldsymbol{X} = [1; X\ \boldsymbol{D}]$ *and* $\boldsymbol{Z} = [1\ \boldsymbol{X}\ \boldsymbol{Z}]$ *and estimated by* $(\boldsymbol{Z}^\top \boldsymbol{X})^{-1} \boldsymbol{Z}^\top \boldsymbol{Y}$
   - *In general,* β_{IV} *estimated in this way does not necessarily have a clear causal interpretation (see Abadie (2003)). More on this later.*

Example: The Vietnam Draft Lottery (Angrist (1990))

- Effect of military service on civilian earnings
- Simple comparison between Vietnam veterans and non-veterans are likely to be a biased measure
- Angrist (1990) used draft-eligibility, determined by the Vietnam era draft lottery, as an instrument for military service in Vietnam
- Draft eligibility is random and affected the probability of enrollment
- Estimate suggest a 15% effect of veteran status on earnings in the period 1981-1984 for white veterans born in 1950-51; although the estimators are quite imprecise

Estimating the Size of the Complier Group

- Since we never observe both potential treatment assignments for the same unit, we cannot identify individual units as compliers
- However, we can easily identify the proportion of compliers in the population using the first stage effect:

$$\begin{aligned} P(D_1 > D_0) &= E[D_1 - D_0] = E[D_1] - E[D_0] \\ &= E[D|Z = 1] - E[D|Z = 0] \end{aligned}$$

Instrumental Variables with Covariates

- Sometimes IV identification assumptions may only hold once we condition on a set of pre-treatment characteristics X.
- In traditional econometric view of IV this poses no additional problems: treatment effects are assumed to be constant, so just include X in the 2SLS model
- But with heterogenous treatment effects, IV/2SLS does not formally identify LATE conditional on X.
- Also hard to interpret, because complier group changes depending on strata of X.
  - One approach: κ-weighting are weights s.t. any model you run gets you a result as if only estimated for compliers (Abadie, 2003)
  - When you use these weights to estimate treatment effect, you get "Local Average Response Function" (LARF)
  - But many investigators still ignore this by taking an econometric view, assuming constant treatment effects

Example: Natural Disasters and Aid

Does getting US disaster aid makes countries cooperate with the US on UN generally assembly votes

- Aid not randomly assigned in general
- Disasters aren't entirely random either, especially seasonal ones
- Non-seasonal disasters – earthquakes and volcanoes – seem totally uncorrelated with "politics"
- ...except maybe through location specific differences in probability (more soon)

So

- Encouragement: non-seasonal natural disasters that kill at least 25 people
- Treatment: getting aid from the US
- Outcome: agreement with the US at UNGA (at various times after disaster/aid)

Let's confirm:

- What is the first-stage?
- What is the reduced-form/intent-to-treat?
- Who are the compliers?

Natural Disasters and Aid with One Year Lag

```
> pd=pdata.frame(d[d$year>1963,],c("country","year"))
> rf.out=lm(daffinity~disaster_25_L1, data=pd)
> summary(rf.out)
Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)    -0.010579   0.001353  -7.821 6.06e-15 ***
disaster_25_L1  0.008352   0.007704   1.084    0.278
---
Multiple R-squared:  0.0001759,Adjusted R-squared:  2.621e-05
F-statistic: 1.175 on 1 and 6680 DF,  p-value: 0.2784

> first.out = lm(allaid_L1~disaster_25_L1, data=pd)
> summary(first.out) #note F stat
Coefficients:
               Estimate Std. Error t value Pr(>|t|)
(Intercept)    0.127977   0.004073   31.42   <2e-16 ***
disaster_25_L1 0.308643   0.023651   13.05   <2e-16 ***
---
Multiple R-squared:  0.02317,Adjusted R-squared:  0.02303
F-statistic: 170.3 on 1 and 7181 DF,  p-value: < 2.2
```

Natural Disaster and Aid with One Year Lag

```
> iv.out=ivreg(daffinity~allaid_L1 | disaster_25_L1, data=pd)
> summary(iv.out)

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.014082   0.003744  -3.761 0.000171 ***
allaid_L1    0.027133   0.025223   1.076 0.282085
---
Multiple R-Squared: -0.01529,Adjusted R-squared: -0.01544
Wald test: 1.157 on 1 and 6680 DF,  p-value: 0.2821
```

Compare to simple Wald Estimator:

```
> rf.out$coef[2]/first.out$coef[2]
disaster_25_L1
    0.02705907
```

Two Extensions

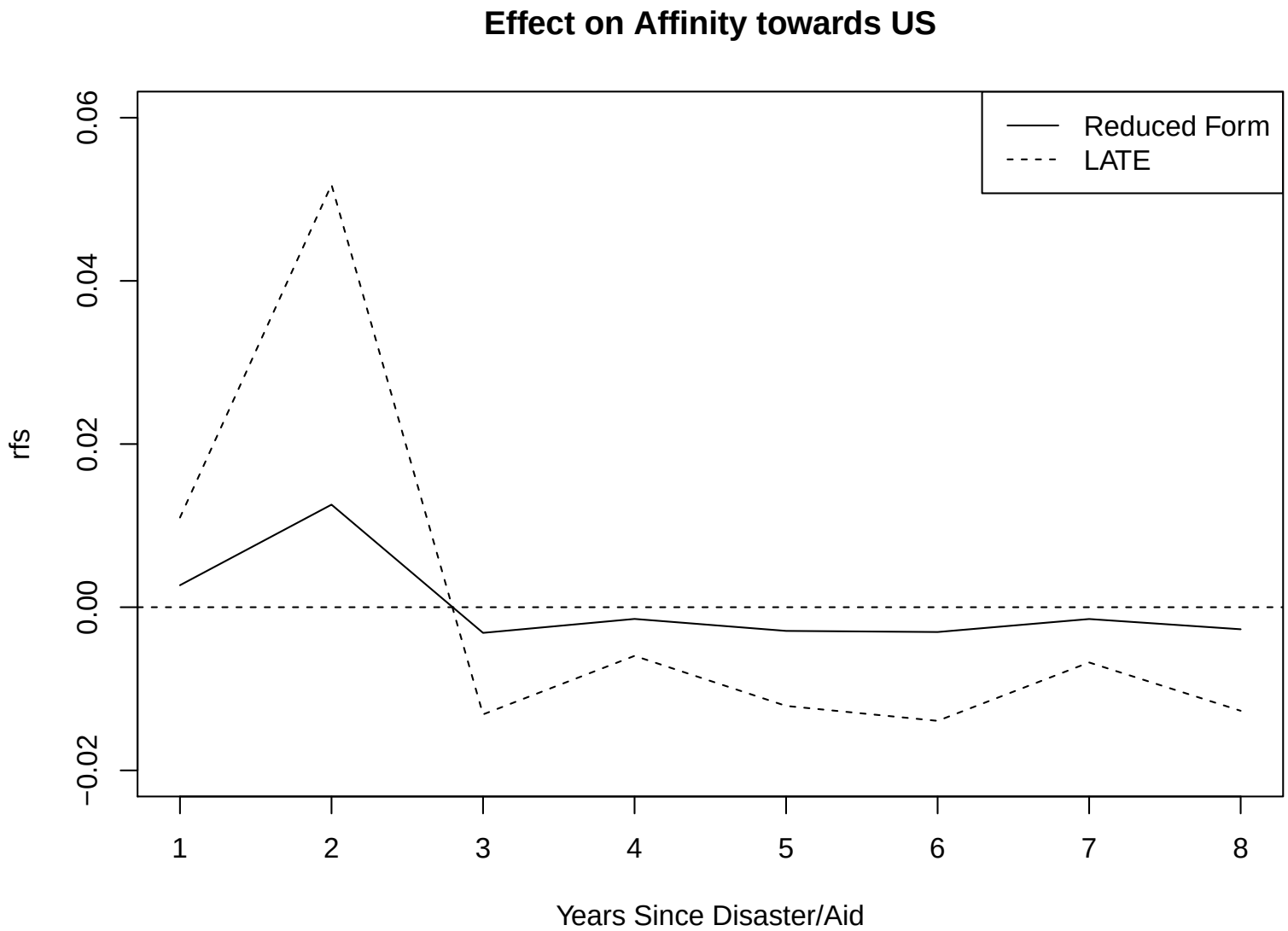
First, we probably want country FEs to strengthen the identification

```
> iv.fe.out=ivreg(daffinity~allaid_L1+country | disaster_25_L1+country, da
> coef(iv.fe.out)[2]
 allaid_L1
0.05312396
```

Second, we don't know what lag to look at so we can look at all of them. Could just run multiple models with different lags, e.g.

```
> rf0.out=lm(daffinity~disaster_25+country, data=pd)
> rf1.out=lm(daffinity~disaster_25_L1+country, data=pd)
> rf2.out=lm(daffinity~disaster_25_L2+country, data=pd)
...
> fs.out=lm(allaid~disaster_25+country, data=pd)
> fs1.out=lm(allaid_L1~disaster_25_L1+country, data=pd)
> fs2.out=lm(allaid_L2~disaster_25_L2+country, data=pd)
```

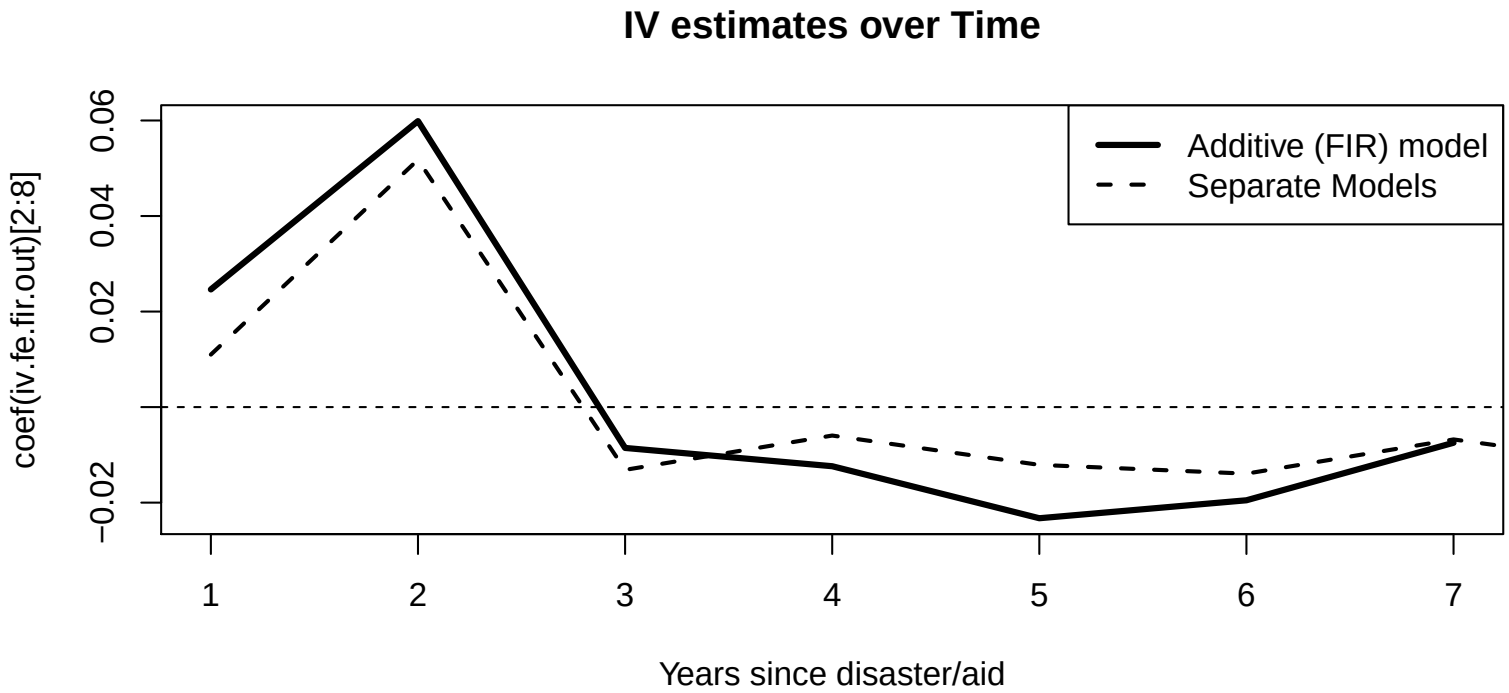
Effects Over Time



Finite-Impulse-Response of Distributed Lag Model

Alternatively, if disasters occur frequently enough that the effects might overlap at some time-points, we can assume additivity and do a “distributed lag” or “finite impulse response” model. e.g.

```
iv.fe.fir.out=ivreg(daffinity~allaid+allaid_L1+allaid_L2+allaid_L3+ ...+co
     |  disaster_25+disaster_25_L1+disaster_25_L2+disaster_25_L3+...+count
```



Summary of Modern IV

- Instruments can give you causal identification where you don't have SOO or other options.
- The key concern is usually the exclusion restriction:
  - together with first-stage ignorability, captured by

$$(Y_{0i}, Y_{1i}, D_{0i}, D_{1i}) \perp\!\!\!\perp Z_i$$

  - a randomized Z_i is a good start, but not sufficient
  - plausibility tests are possible, but cannot be confirmed by observable data
- Generally useful to think about in terms of an "encouragement" to take the treatment and rescaling the ITT by the compliance ratio
- If you do not assume constant treatment effects, then you must recognize the estimate as a LATE among the compliers
- Covariates complicate things: may want to either consider LARF/κ-weighting or rely on constant effect assumption