LATE

Jaakko Meriläinen | Econometrics @ SSE, Fall 2023

Set-Up

- Suppose we are interested in studying the effect of family size on labor supply (Y_i).
- This is tricky to study!
- Let us consider something similar to what Angrist and Evans (1998, AER) do.
  - We have a sample of families with two or more children. We observe their # of children by gender and parents' labor supply.
  - We decide to use the gender composition of the first to children as an instrumental variable for having three or more children (treatment variable is $T_i = 1[Three\ \ children]_i$ - this is *treatment*).
  - We hypothesize that families with two children of the same gender are more likely to have a third child, because maybe they prefer having a mix (instrument is $Z_i = 1[Two\ first\ \ of\ \ same\ \ gender]_i$ - this is *treatment assignment*).

Potential Outcomes (and Treatments)

- Potential outcomes of an individual i are

$$Y_i(0,0) \ \mathit{if} \ T_i = 0 \ \mathit{and} \ Z_i = 0$$

$$Y_i(1,0) \ \mathit{if} \ T_i = 1 \ \mathit{and} \ Z_i = 0$$

$$Y_i(0,1) \ \mathit{if} \ T_i = 0 \ \mathit{and} \ Z_i = 1$$

$$Y_i(1,1) \ \mathit{if} \ T_i = 1 \ \mathit{and} \ Z_i = 1$$

- Potential treatments of an individual i are

$$T_0 \ \mathit{if} \ Z_i = 0$$

$$T_1 \ \mathit{if} \ Z_i = 1$$

Compliance Groups

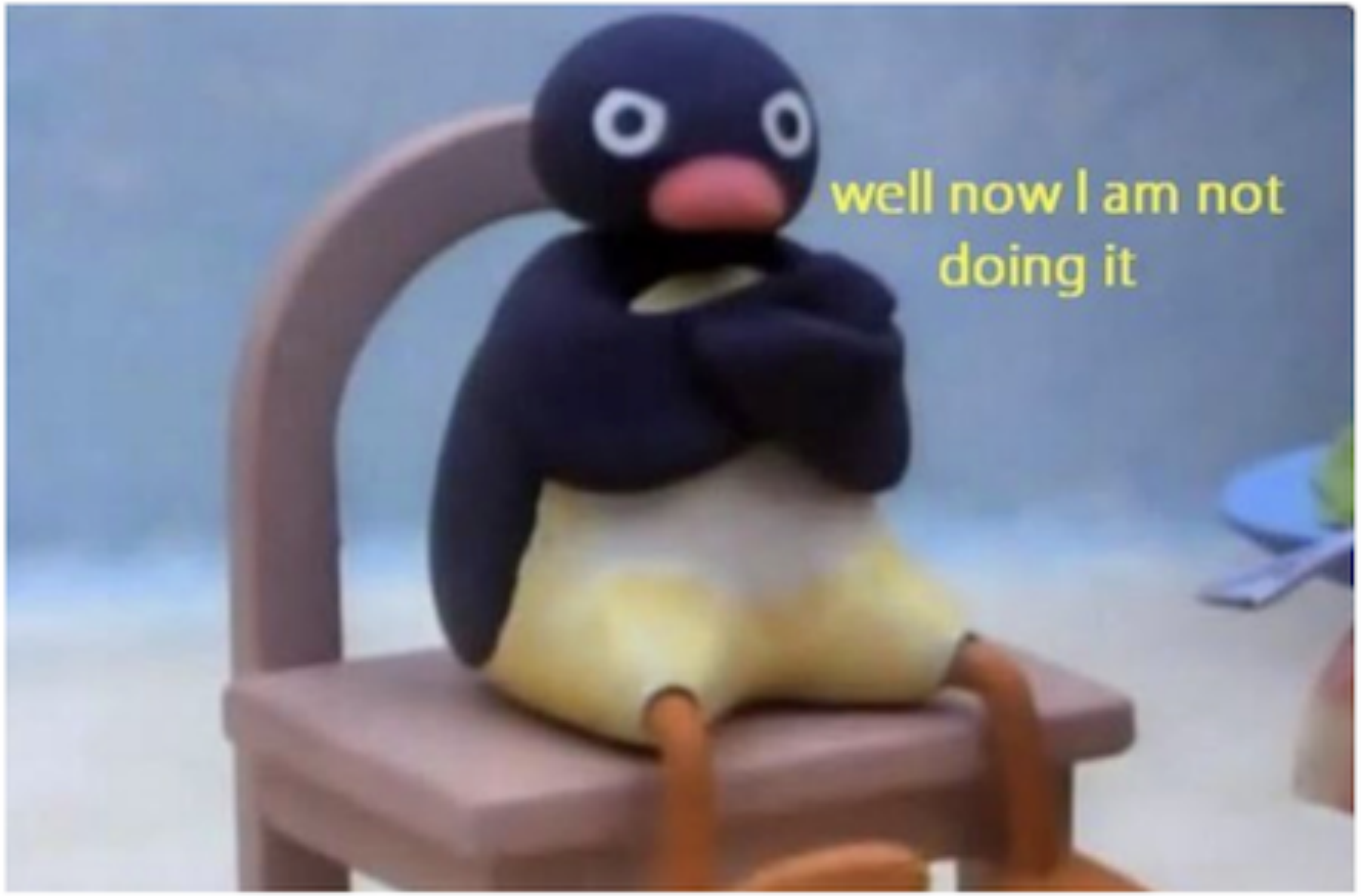
- We can think of four types of compliance groups:
  - Always-takers who always get treated despite treatment assignment (gender composition of the first two kids): $T_{i1} = T_{i0} = 1$
  - Never-takers who never get treated despite treatment assignment: $T_{i1} = T_{i0} = 0$
  - Compliers who get treated according to the treatment assignment: $T_{i1} > T_{i0}$
  - Defiers who get treated opposite to the treatment assignment: $T_{i1} < T_{i0}$
- Remember that we cannot tell which group an observation belongs to (without some additional assumptions)!

Researcher: You, are in the control group. No need to take the treatment

Defier: But I want it!

Researcher: Just kidding, you are in the treatment group. Here it is

Defier:



IV Conditions

- If the four conditions hold, we could causally identify the **local average treatment effect (LATE)** of having three vs. two children on labor supply:

  1. **Independence:** The instrument is as good as randomly assigned.
  2. **Exclusion restriction:** The instrument only affects the outcome through the treatment.
  3. **Relevance**: The instrument does affect the treatment.
  4. **Monotonicity:** There are no defiers.

- A small exercise to you: How would you express these things using the potential outcome notation from the third slide? See slides for Lecture 9!

SUTVA

- In *Causal Inference: The Mixtape*, you can see a fifth assumption: **SUTVA** (stable unit treatment value assumption).
- You might also remember this from the first lecture.
- We typically require that a treatment only affects the treated individuals and does not spill over to the control group.
- This is the SUTVA assumption.
- If there are spillovers, you may still be able to say something about the presence and direction of a causal effect.

Evidence of Independence and Relevance (?)

Sex of first two children in families with two or more children	All women				Married women			
	1980 PUMS (394,835 observations)		1990 PUMS (380,007 observations)		1980 PUMS (254,654 observations)		1990 PUMS (301,588 observations)	
	Fraction of sample	Fraction that had another child	Fraction of sample	Fraction that had another child	Fraction of sample	Fraction that had another child	Fraction of sample	Fraction that had another child
one boy, one girl	0.494	0.372 (0.001)	0.495	0.344 (0.001)	0.494	0.346 (0.001)	0.497	0.331 (0.001)
two girls	0.242	0.441 (0.002)	0.241	0.412 (0.002)	0.239	0.425 (0.002)	0.239	0.408 (0.002)
two boys	0.264	0.423 (0.002)	0.264	0.401 (0.002)	0.266	0.404 (0.002)	0.264	0.396 (0.002)
(1) one boy, one girl	0.494	0.372 (0.001)	0.495	0.344 (0.001)	0.494	0.346 (0.001)	0.497	0.331 (0.001)
(2) both same sex	0.506	0.432 (0.001)	0.505	0.407 (0.001)	0.506	0.414 (0.001)	0.503	0.401 (0.001)
difference (2) − (1)	—	0.060 (0.002)	—	0.063 (0.002)	—	0.068 (0.002)	—	0.070 (0.002)

Notes: The samples are the same as in Table 2. Standard errors are reported in parentheses.

Evidence of Independence and Relevance (?)

The probability of a boy or a girl should be around 0.5 — seems to be the case here!

Sex of first two children in families with two or more children	All women: 1980 PUMS (394,835 observations) Fraction of sample	Fraction that had another child	All women: 1990 PUMS (380,007 observations) Fraction of sample	Fraction that had another child	Married women: 1980 PUMS (254,654 observations) Fraction of sample	Fraction that had another child	Married women: 1990 PUMS (301,588 observations) Fraction of sample	Fraction that had another child
one boy, one girl	0.494	0.372 (0.001)	0.495	0.344 (0.001)	0.494	0.346 (0.001)	0.497	0.331 (0.001)
two girls	0.242	0.441 (0.002)	0.241	0.412 (0.002)	0.239	0.425 (0.002)	0.239	0.408 (0.002)
two boys	0.264	0.423 (0.002)	0.264	0.401 (0.002)	0.266	0.404 (0.002)	0.264	0.396 (0.002)
(1) one boy, one girl	0.494	0.372 (0.001)	0.495	0.344 (0.001)	0.494	0.346 (0.001)	0.497	0.331 (0.001)
(2) both same sex	0.506	0.432 (0.001)	0.505	0.407 (0.001)	0.506	0.414 (0.001)	0.503	0.401 (0.001)
difference (2) − (1)	—	0.060 (0.002)	—	0.063 (0.002)	—	0.068 (0.002)	—	0.070 (0.002)

Notes: The samples are the same as in Table 2. Standard errors are reported in parentheses.

Evidence of Independence and Relevance (?)

There seems to be a first stage, too…

Sex of first two children in families with two or more children	All women				Married women			
	1980 PUMS (394,835 observations)		1990 PUMS (380,007 observations)		1980 PUMS (254,654 observations)		1990 PUMS (301,588 observations)	
	Fraction of sample	Fraction that had another child	Fraction of sample	Fraction that had another child	Fraction of sample	Fraction that had another child	Fraction of sample	Fraction that had another child
one boy, one girl	0.494	0.372 (0.001)	0.495	0.344 (0.001)	0.494	0.346 (0.001)	0.497	0.331 (0.001)
two girls	0.242	0.441 (0.002)	0.241	0.412 (0.002)	0.239	0.425 (0.002)	0.239	0.408 (0.002)
two boys	0.264	0.423 (0.002)	0.264	0.401 (0.002)	0.266	0.404 (0.002)	0.264	0.396 (0.002)
(1) one boy, one girl	0.494	0.372 (0.001)	0.495	0.344 (0.001)	0.494	0.346 (0.001)	0.497	0.331 (0.001)
(2) both same sex	0.506	0.432 (0.001)	0.505	0.407 (0.001)	0.506	0.414 (0.001)	0.503	0.401 (0.001)
difference (2) − (1)	—	0.060 (0.002)	—	0.063 (0.002)	—	0.068 (0.002)	—	0.070 (0.002)

Notes: The samples are the same as in Table 2. Standard errors are reported in parentheses.

Wald Estimates (RF/First Stage)

First stage(s)

TABLE 5—WALD ESTIMATES OF LABOR-SUPPLY MODELS

	1980 PUMS			1990 PUMS			1980 PUMS		
		Wald estimate using as covariate:			Wald estimate using as covariate:			Wald estimate using as covariate:	
Variable	Mean difference by *Same sex*	*More than 2 children*	*Number of children*	Mean difference by *Same sex*	*More than 2 children*	*Number of children*	Mean difference by *Twins-2*	*More than 2 children*	*Number of children*
More than 2 children	0.0600 (0.0016)	—	—	0.0628 (0.0016)	—	—	0.6031 (0.0084)	—	—
Number of children	0.0765 (0.0026)	—	—	0.0836 (0.0025)	—	—	0.8094 (0.0139)	—	—
Worked for pay	−0.0080 (0.0016)	−0.133 (0.026)	−0.104 (0.021)	−0.0053 (0.0015)	−0.084 (0.024)	−0.063 (0.018)	−0.0459 (0.0086)	−0.076 (0.014)	−0.057 (0.011)
Weeks worked	−0.3826 (0.0709)	−6.38 (1.17)	−5.00 (0.92)	−0.3233 (0.0743)	−5.15 (1.17)	−3.87 (0.88)	−1.982 (0.386)	−3.28 (0.63)	−2.45 (0.47)
Hours/week	−0.3110 (0.0602)	−5.18 (1.00)	−4.07 (0.78)	−0.2363 (0.0620)	−3.76 (0.98)	−2.83 (0.73)	−1.979 (0.327)	−3.28 (0.54)	−2.44 (0.40)
Labor income	−132.5 (34.4)	−2208.8 (569.2)	−1732.4 (446.3)	−119.4 (42.4)	−1901.4 (670.3)	−1428.0 (502.6)	−570.8 (186.9)	−946.4 (308.6)	−705.2 (229.8)
ln(*Family income*)	−0.0018 (0.0041)	−0.029 (0.068)	−0.023 (0.054)	−0.0085 (0.0047)	−0.136 (0.074)	−0.102 (0.056)	−0.0341 (0.0223)	−0.057 (0.037)	−0.042 (0.027)

Notes: The samples are the same as in Table 2. Standard errors are reported in parentheses.

Wald Estimates (RF/First Stage)

Reduced form(s)

TABLE 5—WALD ESTIMATES OF LABOR-SUPPLY MODELS

Variable	1980 PUMS: Mean difference by *Same sex*	1980 PUMS: Wald estimate using as covariate: *More than 2 children*	1980 PUMS: Wald estimate using as covariate: *Number of children*	1990 PUMS: Mean difference by *Same sex*	1990 PUMS: Wald estimate using as covariate: *More than 2 children*	1990 PUMS: Wald estimate using as covariate: *Number of children*	1980 PUMS: Mean difference by *Twins-2*	1980 PUMS: Wald estimate using as covariate: *More than 2 children*	1980 PUMS: Wald estimate using as covariate: *Number of children*
More than 2 children	0.0600 (0.0016)	—	—	0.0628 (0.0016)	—	—	0.6031 (0.0084)	—	—
Number of children	0.0765 (0.0026)	—	—	0.0836 (0.0025)	—	—	0.8094 (0.0139)	—	—
Worked for pay	−0.0080 (0.0016)	−0.133 (0.026)	−0.104 (0.021)	−0.0053 (0.0015)	−0.084 (0.024)	−0.063 (0.018)	−0.0459 (0.0086)	−0.076 (0.014)	−0.057 (0.011)
Weeks worked	−0.3826 (0.0709)	−6.38 (1.17)	−5.00 (0.92)	−0.3233 (0.0743)	−5.15 (1.17)	−3.87 (0.88)	−1.982 (0.386)	−3.28 (0.63)	−2.45 (0.47)
Hours/week	−0.3110 (0.0602)	−5.18 (1.00)	−4.07 (0.78)	−0.2363 (0.0620)	−3.76 (0.98)	−2.83 (0.73)	−1.979 (0.327)	−3.28 (0.54)	−2.44 (0.40)
Labor income	−132.5 (34.4)	−2208.8 (569.2)	−1732.4 (446.3)	−119.4 (42.4)	−1901.4 (670.3)	−1428.0 (502.6)	−570.8 (186.9)	−946.4 (308.6)	−705.2 (229.8)
ln(*Family income*)	−0.0018 (0.0041)	−0.029 (0.068)	−0.023 (0.054)	−0.0085 (0.0047)	−0.136 (0.074)	−0.102 (0.056)	−0.0341 (0.0223)	−0.057 (0.037)	−0.042 (0.027)

Notes: The samples are the same as in Table 2. Standard errors are reported in parentheses.

Wald Estimates (RF/First Stage)

Wald estimates

TABLE 5—WALD ESTIMATES OF LABOR-SUPPLY MODELS

Variable	1980 PUMS			1990 PUMS			1980 PUMS		
	Mean difference by *Same sex*	Wald estimate using as covariate: *More than 2 children*	Wald estimate using as covariate: *Number of children*	Mean difference by *Same sex*	Wald estimate using as covariate: *More than 2 children*	Wald estimate using as covariate: *Number of children*	Mean difference by *Twins-2*	Wald estimate using as covariate: *More than 2 children*	Wald estimate using as covariate: *Number of children*
More than 2 children	0.0600 (0.0016)	—	—	0.0628 (0.0016)	—	—	0.6031 (0.0084)	—	—
Number of children	0.0765 (0.0026)	—	—	0.0836 (0.0025)	—	—	0.8094 (0.0139)	—	—
Worked for pay	−0.0080 (0.0016)	−0.133 (0.026)	−0.104 (0.021)	−0.0053 (0.0015)	−0.084 (0.024)	−0.063 (0.018)	−0.0459 (0.0086)	−0.076 (0.014)	−0.057 (0.011)
Weeks worked	−0.3826 (0.0709)	−6.38 (1.17)	−5.00 (0.92)	−0.3233 (0.0743)	−5.15 (1.17)	−3.87 (0.88)	−1.982 (0.386)	−3.28 (0.63)	−2.45 (0.47)
Hours/week	−0.3110 (0.0602)	−5.18 (1.00)	−4.07 (0.78)	−0.2363 (0.0620)	−3.76 (0.98)	−2.83 (0.73)	−1.979 (0.327)	−3.28 (0.54)	−2.44 (0.40)
Labor income	−132.5 (34.4)	−2208.8 (569.2)	−1732.4 (446.3)	−119.4 (42.4)	−1901.4 (670.3)	−1428.0 (502.6)	−570.8 (186.9)	−946.4 (308.6)	−705.2 (229.8)
ln(*Family income*)	−0.0018 (0.0041)	−0.029 (0.068)	−0.023 (0.054)	−0.0085 (0.0047)	−0.136 (0.074)	−0.102 (0.056)	−0.0341 (0.0223)	−0.057 (0.037)	−0.042 (0.027)

Notes: The samples are the same as in Table 2. Standard errors are reported in parentheses.

Is the No-Defiers Assumption Reasonable?

- To get to the LATE interpretation, we would assume that there are no defiers. Is this reasonable? Probably not…
- Suppose there are families that want to have three kids (so $T_i = 1$) but they only do so if they have two children of different sex ($Z_i = 0$) because they are worried about having three sons.
- That is, if they had two children of the same sex ($Z_i = 1$), they would choose not to have more children (i.e., just two children, $T_i = 0$) because they do not want to have a third child of the same sex.
- It is easy to imagine that people like this could exist—and they would be defiers!
- We are no longer identifying a LATE! But maybe we are still quite close to it if there are not that many defiers?

LATE

- What is the LATE? When is what we have estimated LATE?
- It is quite simple to show that our IV estimator

$$\hat{\beta} = \frac{E(Y|Z=1) - E(Y|Z=0)}{E(T|Z=1) - E(T|Z=0)}$$

is identified from the set of compliers when the conditions in the previous slide hold;

$$\hat{\beta} = E(Y_{i1} - Y_{i0} | T_{i1} > T_{i0}).$$

LATE

With binary treatment and instrument, RF is difference in means of Y between groups with Z = 1 and Z = 0

- What is the LATE?
- It is quite simple to show that our IV estimator

$$\hat{\beta} = \frac{E(Y|Z=1) - E(Y|Z=0)}{E(T|Z=1) - E(T|Z=0)}$$

is identified from the set of compliers when the conditions in the previous slide hold;

$$\hat{\beta} = E(Y_{i1} - Y_{i0} | T_{i1} > T_{i0}).$$

LATE

With binary treatment and instrument, the first stage is difference in means of T between groups with Z = 1 and Z = 0

- What is the LATE?

- It is quite simple to show that our IV estimator

$$\hat{\beta} = \frac{E(Y|Z=1) - E(Y|Z=0)}{E(T|Z=1) - E(T|Z=0)}$$

is identified from the set of compliers when the conditions in the previous slide hold;

$$\hat{\beta} = E(Y_{i1} - Y_{i0} | T_{i1} > T_{i0}).$$

LATE

This is the group of individuals for whom treatment is 1 when they are assigned 1 and 0 when they are assigned 0

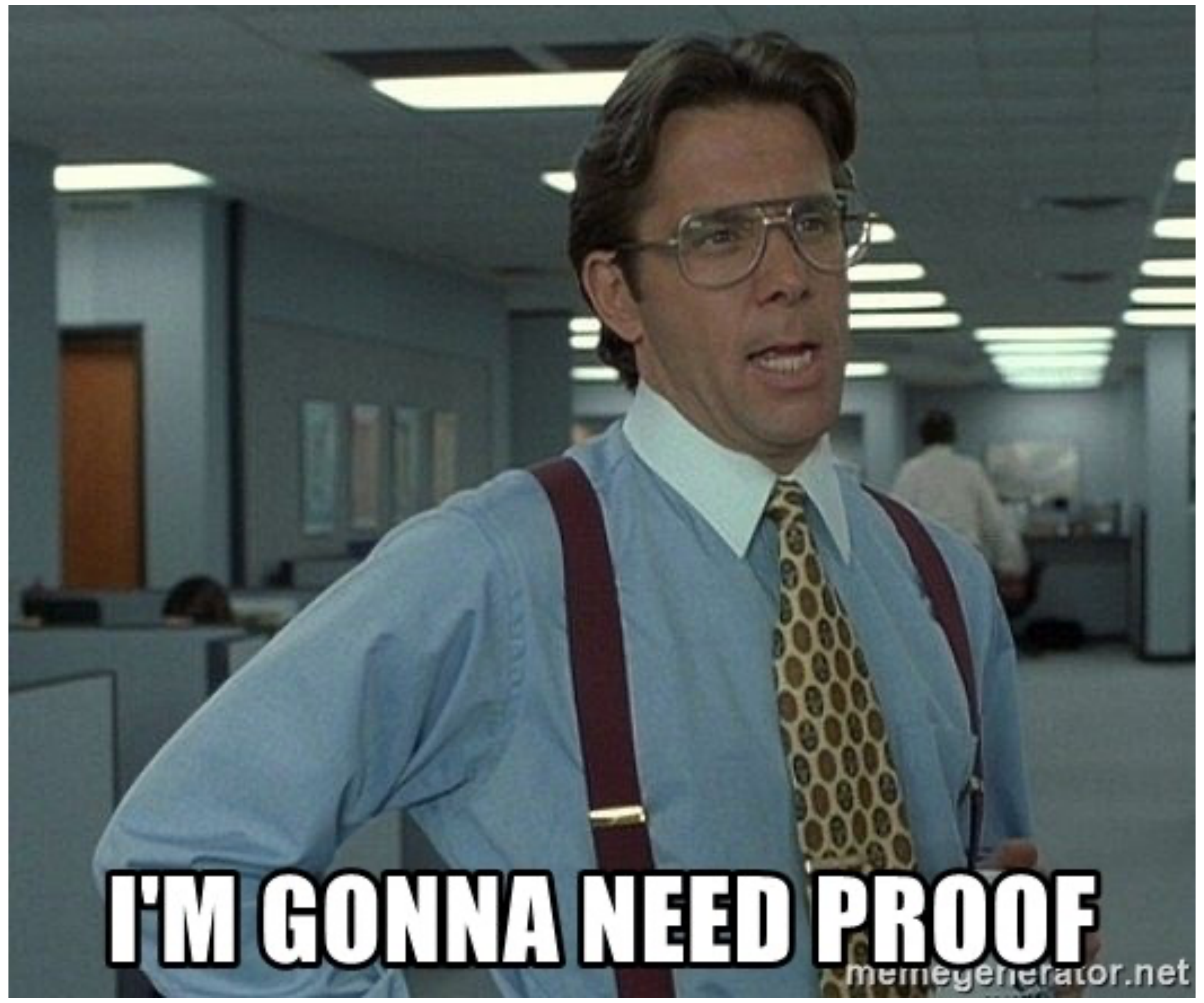
- What is the LATE?

- It is quite simple to show that our IV estimator

$$\hat{\beta} = \frac{E(Y|Z=1) - E(Y|Z=0)}{E(T|Z=1) - E(T|Z=0)}$$

is identified from the set of compliers when the conditions in the previous slide hold;

$$\hat{\beta} = E(Y_{i1} - Y_{i0} | T_{i1} > T_{i0}).$$



We will cover it in the class time permitting… Otherwise, a proof is available in the end of these slides!

Generalizing LATE

- It is easiest to illustrate LATE with a binary treatment and binary instrument.
- But LATE is more general than this.
- It extends to
  - multivalued endogenous regressors
  - multiple instruments
  - IV with covariates
- See Section 4.5 in *Mostly Harmless Econometrics*! There was also some discussion in Lecture 8.

Two Empirical Examples in the Slides

- Lecture 8 contains two empirical examples that you should give a look.
- **Example 1:** Estimates of returns to education across three countries (Oreopoulos 2006).
    - Oreopoulos estimates economic returns to education using compulsory schooling laws in different countries.
    - LATE estimates are very similar across contexts despite very different compliance rates.
    - OLS and IV estimates very different, but perhaps the reason is not that the LATE is being estimated on some peculiar sample.

Two Empirical Examples in the Slides

- **Example 2:** Learning gains from a personalized training program in Delhi (Muralidharan et al. 2019)
  - Lottery-winners were offered free access to the program for ~1/2 a school year.
  - High take-up (among lottery participants) if defined as showing up ever, but variable take-up when looking at number of days actually attended.
  - Muralidharan et al. look at the effect of lottery win on learning outcomes (intent-to-treat estimates) and also instrument attendance with winning the lottery.
  - Higher attendance induced by winning the lottery leads to higher math and Hindi test scores—this is a LATE identified for compliers.
  - Suggestive evidence that LATE ~ ATE in this particular setting.

Proof of the LATE Theorem

$$\hat{\beta} = \frac{E(Y|Z=1) - E(Y|Z=0)}{E(T|Z=1) - E(T|Z=0)}$$

- Let us denote the set of always-takers as A, the set of never-takers as N, the set of compliers as C, and the set of of defiers as D.
- Note first that

$$E(Y|Z=1) = E(Y_1|i \in A) \times P(i \in A) + E(Y_1|i \in C) \times P(i \in C)$$

$$+E(Y_1|i \in N) \times P(i \in N) + E(Y_1|i \in D) \times P(i \in D)$$

Proof of the LATE Theorem

- Recall that for never-takers and defiers, $Y_1 = Y_0$, so

$$E(Y|Z=1) = E(Y_1|i \in A) \times P(i \in A) + E(Y_1|i \in C) \times P(i \in C)$$

$$+E(Y_0|i \in N) \times P(i \in N) + E(Y_0|i \in D) \times P(i \in D) .$$

- Similarly,

$$E(Y|Z=0) = E(Y_0|i \in A) \times P(i \in A) + E(Y_0|i \in C) \times P(i \in C)$$

$$+E(Y_0|i \in N) \times P(i \in N) + E(Y_0|i \in D) \times P(i \in D) .$$

Proof of the LATE Theorem

- Given that for always-takers $Y_0 = Y_1$, this becomes

$$E(Y|Z=0) = E(Y_1|i \in A) \times P(i \in A) + E(Y_0|i \in C) \times P(i \in C)$$

$$+E(Y_0|i \in N) \times P(i \in N) + E(Y_0|i \in D) \times P(i \in D)\,.$$

- It is now easy to see that the numerator can be expressed as

$$E(Y|Z=1) - E(Y|Z=0) = E(Y_1|i \in C) \times P(i \in C) - E(Y_0|i \in C) \times P(i \in C)\,.$$

Proof of the LATE Theorem

- Consider then the denominator, $E(T|Z=1) - E(T|Z=0)$. The first term can be expressed as

$$E(T|Z=1) = P(T=1|Z=1, i \in A) \times P(i \in A) + P(T=1|Z=1, i \in C) \times P(i \in C)$$

$$+P(T=1|Z=1, i \in N) + P(T=1|Z=1, i \in D)\,.$$

- Here, just remember that never-takers never have $T=1$ and defiers do the opposite, whereas always-takers and compliers always have $T=1$. So the expression simply becomes

$$E(T|Z=1) = P(i \in A) + P(i \in C)\,.$$

Proof of the LATE Theorem

- The second part of the denominator can be written as

$$E(T|Z=0) = P(T=1|Z=0, i \in A) \times P(i \in A) + P(T=1|Z=0, i \in C) \times P(i \in C)$$

$$+P(T=1|Z=0, i \in N) + P(T=1|Z=0, i \in D)\,.$$

- Again, think what our compliance groups would do. Always-takers will always choose $T=1$, compliers would choose $T=0$ if $Z=0$, never-takers would choose $T=0$, and defiers would choose $T=1$ since $Z=0$. So, we have

$$E(T|Z=0) = P(i \in A) + P(i \in D)\,.$$

Proof of the LATE Theorem

- Now look at the difference between these two to get the denominator:

$$E(T \mid Z = 1) - E(T \mid Z = 0) = P(i \in A) + P(i \in C) - P(i \in A) - P(i \in D)\,.$$

- We now get to use one of the assumptions that we made, namely that there are no defiers! This means that $E(T \mid Z = 1) - E(T \mid Z = 0) = P(i \in C)$.

Proof of the LATE Theorem

- We can finally put everything together:

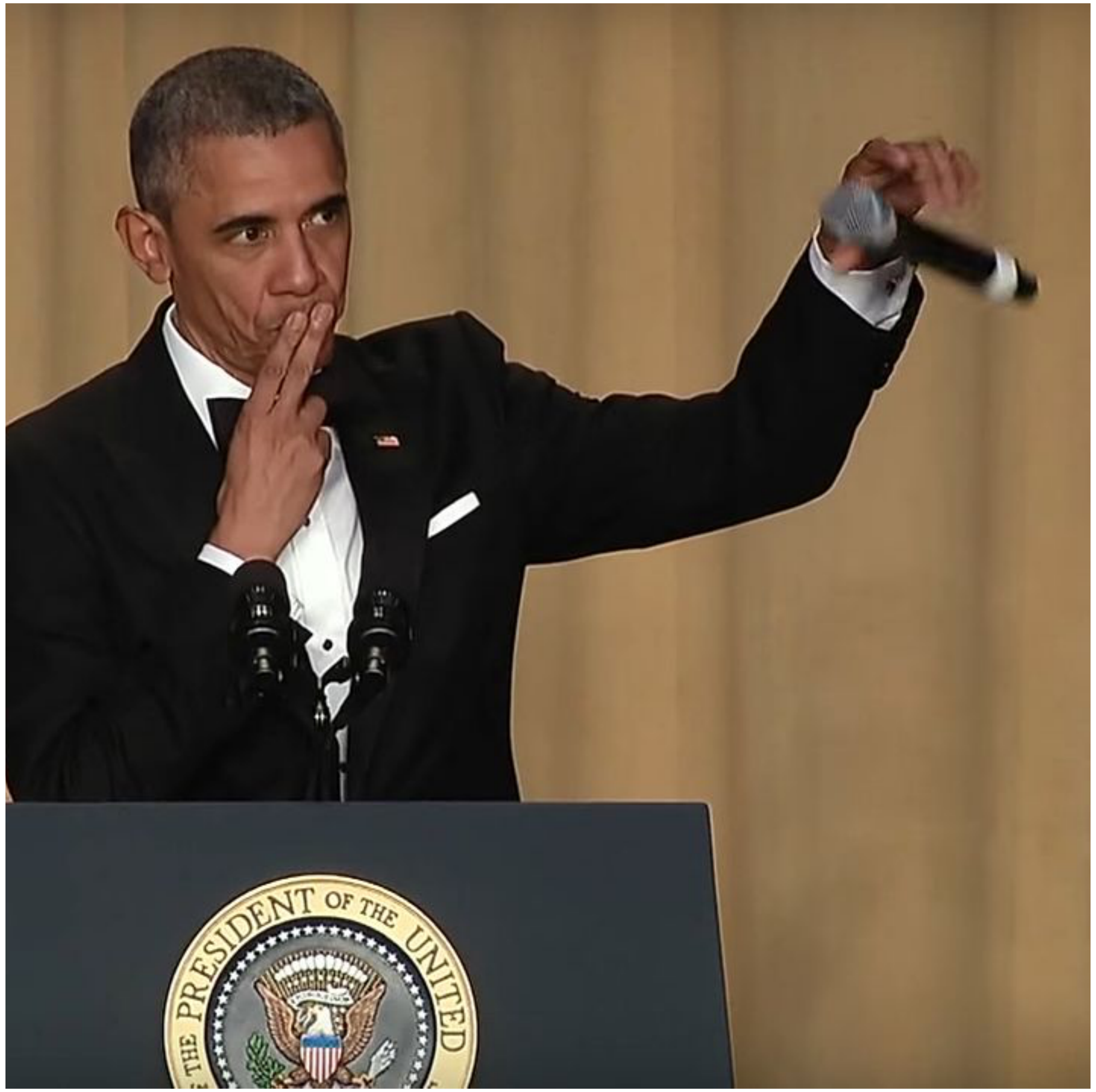
$$\hat{\beta} = \frac{E(Y|Z=1) - E(Y|Z=0)}{E(T|Z=1) - E(T|Z=0)}$$

$$= \frac{E(Y_1|i \in C) \times P(i \in C) - E(Y_0|i \in C) \times P(i \in C)}{P(i \in C)}$$

$$= E(Y_1 - Y_0|i \in C)\,.$$

- So, if our assumptions hold, IV identifies the LATE.

$\square$


THE PRESIDENT OF THE UNITED