

Lecture 4 - Controls and Fixed Effects

Topics in Econometrics - M2 ENS Lyon

Vincent Bagilet

2025-09-24

Outline of the course

1. Introduction and Fundamentals
2. Simulations
3. Design: Identification and Beyond
4. **Controls and Fixed Effects**
5. IV and RDD
6. Modelling
7. Analysis
8. Communication

Goals of the session

- Controls: what do controls do to our regression?
- **Fixed effects** are extremely common in applied economics
- What are they really doing?
- More generally, what are we really estimating in a specific model?
- What are we comparing to what?
- Where does the identifying variation come from?

Controls and projections

Regression as a projection

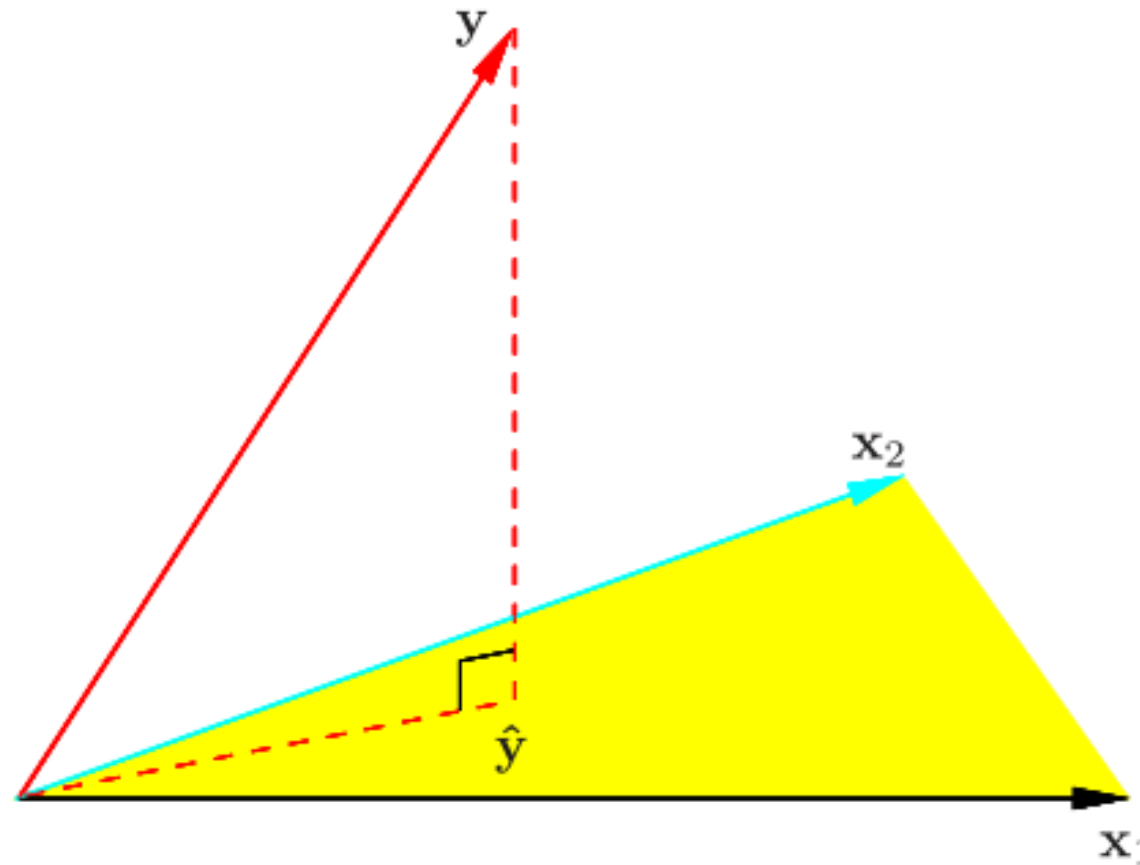


FIGURE 3.2. *The N -dimensional geometry of least squares regression with two predictors. The outcome vector $\mathbf{y}$ is orthogonally projected onto the hyperplane spanned by the input vectors $\mathbf{x}_1$ and $\mathbf{x}_2$. The projection $\hat{\mathbf{y}}$ represents the vector of the least squares predictions*

Frisch-Waugh-Lovell (FWL) Theorem

$$Y = X\beta + W\delta + U$$

- The estimate of β is the same as the estimate of $\tilde{\beta}$ in:

$$Y^{\perp W} = X^{\perp W} \tilde{\beta} + U^{\perp W}$$

- where $^{\perp W}$ denotes each variable where W has been **residualized**
- ie its projection onto the **orthogonal space to W**
- Obtained using:
 - The **projection matrix** $P_W = W(W'W)^{-1}W'$
 - The **residual-maker matrix** $M_W = I - P_W$
- eg $X^{\perp W} = M_W X$
- Fixed effects regression = regression on variables after partialling out the fixed effects

In practice

- To compute the partialled out version of a regression:
 1. Compute the residualized version of y and x : regress them on controls/FE
 2. Regress the **residuals** on one another
- Exercise. Using the data bellow, run two regressions and compare the estimates obtained:
 1. Regress $\ln_murder$ on $\ln_pris$, $male$, $population$ and $income$
 2. Regress their residualized versions of $\ln_murder$ and $\ln_pris$ on one another (partialling out $male$, $population$ and $income$)

```
1 library(AER)
2 data("Guns")
3
4 guns <- Guns |>
5   as_tibble() |>
6   mutate(
7     ln_pris = log(prisoners),
8     ln_murder = log(murder)
9   )
```

Visualizing the raw data

Graph levels

Graph logs

Code

Relationship between incarceration and murder rates

Variables in level: need to transform it

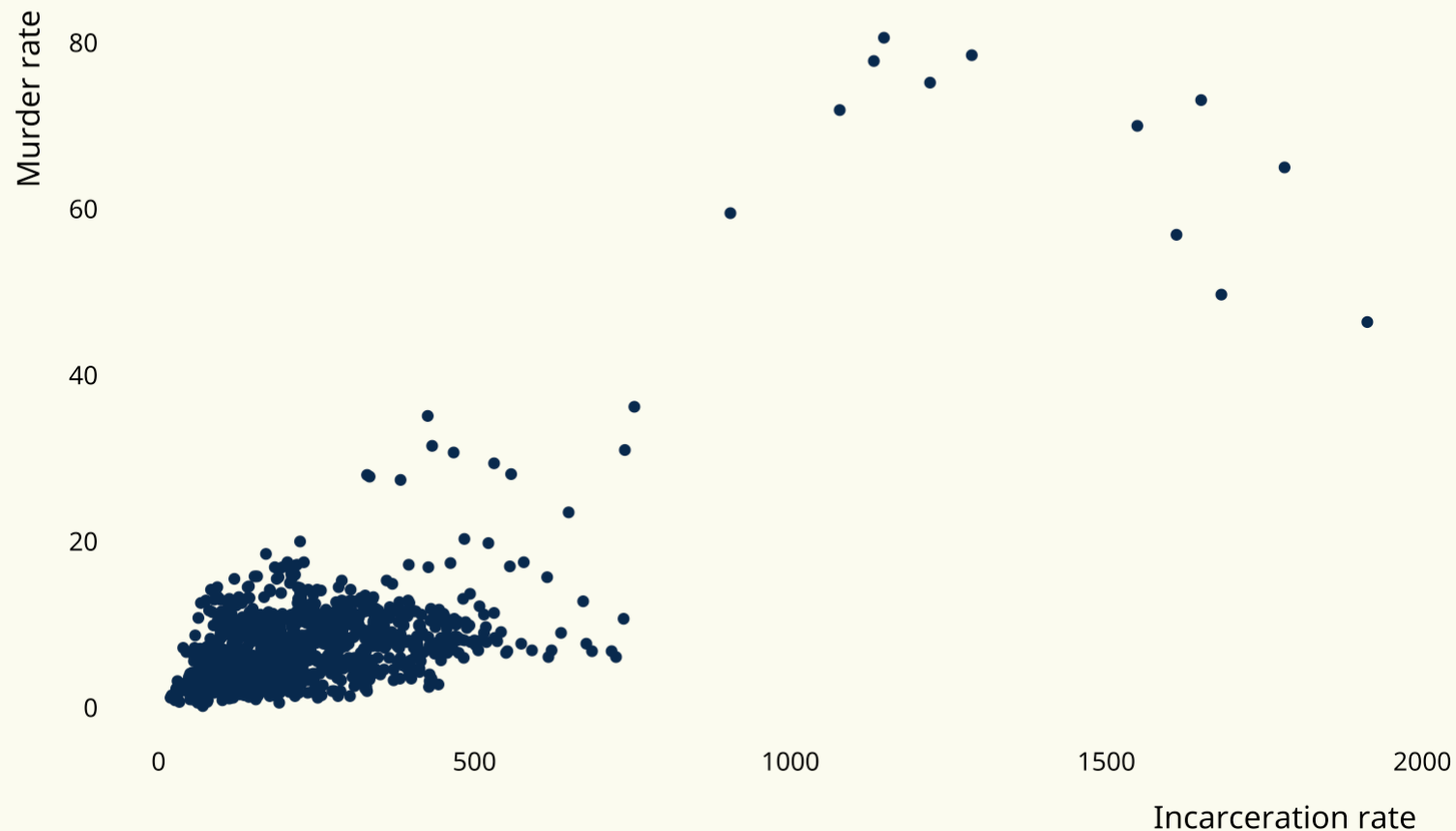


Illustration of the FWL theorem

```
1 #demeaning and showing that equal to residuals
2 guns_demean <- guns |>
3   mutate(
4     l_murder_res = lm(data = guns, formula = l_murder ~ male + population + income) |>
5       residuals(),
6     l_pris_res = lm(data = guns, formula = l_pris ~ male + population + income) |>
7       residuals()
8   )
9
10 reg <- guns |>
11   lm(formula = l_murder ~ l_pris + male + population + income) |>
12   broom::tidy() |>
13   mutate(reg = "raw", .before = 1)
14
15 reg_res <- guns_demean |>
16   lm(formula = l_murder_res ~ l_pris_res - 1) |>
17   broom::tidy() |>
18   mutate(reg = "residualized", .before = 1)
19
20 rbind(reg, reg_res) |>
21   filter(str_starts(term, "l_pris")) |>
22   kable()
```

reg	term	estimate	std.error	statistic	p.value
raw	l_pris	0.8702664	0.0251012	34.67037	0
residualized	l_pris_res	0.8702664	0.0250583	34.72968	0

Why does that matter?

- In economics, we are typically interested in one coefficient, that of the treatment
- The true relationship of interest is the **partialled out** one
- Can bring back a complex linear model to a bivariate regression
- What matters is the variation in the treatment **after partialling out**
- Where does the variation in the treatment comes from?
- Interpretation of the coefficient of interest: “**Comparing ...**”

ATE as a weighted average

- The estimate of the treatment coefficient is in fact **a weighted average of individual treatment effects** τ_i
 - See Aronow and Samii (2016) and Angrist and Pischke (2009) section 3.3.1
 - $\forall i, \hat{\beta} \xrightarrow{p} \frac{\mathbb{E}[w_i \tau_i]}{\mathbb{E}[w_i]}$
- Weight: $w_i = (T_i - \mathbb{E}[T_i | X_i])^2$
- The weight represents:
 - How well the controls explain the treatment status
 - The conditional variance of the treatment, given X_i
- Actually **equivalent to leverage** in the residualized regression

Implications

- Observations whose treatment status is largely explained by covariates therefore **contribute little, if at all, to estimation**
- For FE: if for some groups there is little within variation, these groups do not contribute to identification
- Implications for **external validity** and **representativity**
- May weigh more some characteristics
- Implications for statistical power: the **effective sample** might be much smaller than the nominal sample
- If need a **CIA**, weights are not equal

Effective sample vs nominal sample

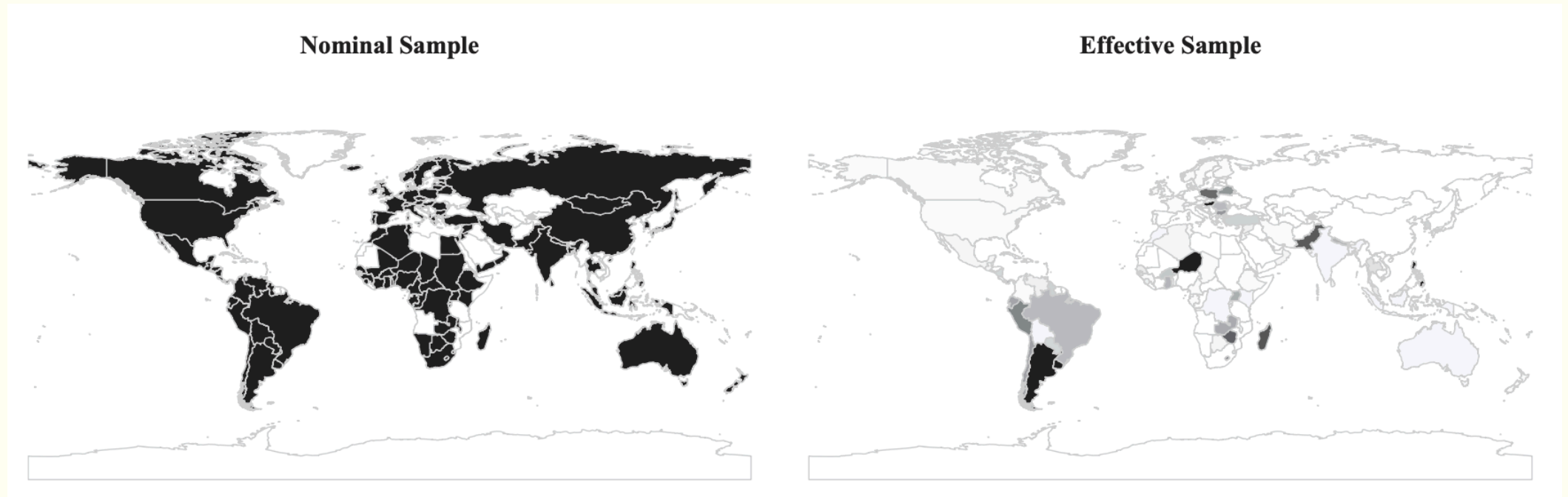


Figure from Aronow and Samii (2016)

Two types of controls

- Can distinguish controls between those that are **identification-related** and those that are not
- Identification-related controls are those that are necessary for the **CIA** to hold
- Other controls are here to **improve precision** of the estimator
 - Explain $y \Rightarrow \searrow \sigma_u^2$ (and $\nearrow R^2$) and therefore $\searrow \mathbb{V}_{\hat{\beta}}$

Identification based on repeated observations

Adjusting for non-varying factors

- Repeated observations over some dimension allow **adjusting for all the unobserved characteristics that are constant across that dimension**
- Transform each variable into its **deviation from the group mean**
- Only keep **within variation** (discards the *between*)
- Two approaches to do that:
 - Manual demeaning
 - Including fixed effects
- Basically build a **counterfactual**

Event studies, DiD, and TWFEs

- *Objective*: estimate the impact of some treatment at a certain time
- Leverages repeated observations, typically **panel data**
- Builds a **counterfactual** that can be explicit or more implicit (eg TWFE):
 - Unit's outcome had the event not occurred

Event study

- Potentially, all units are treated
- Assumed counterfactual: group's past value
- *Within* variation only
- + Flexible, allows looking at whether effects are **dynamic**
- – Difficult to rule out **other things changing at the same time**
 - *The rooster concluding the sun rises because of his crowing?*

$$Y_{it} = \sum_{t=-K}^{\tau-2} [\beta_t \mathbb{1}\{t\}] + \beta_{\tau} \mathbb{1}\{\tau\} + \sum_{t=\tau+1}^L [\beta_t \mathbb{1}\{t\}] + e_{it}$$

DiD, DiDiD, TWFE

- Some units never get treated
- Assumed counterfactual: **parallel trends** of treated and untreated are parallel
- *Within* and *between* variation
- + Pre-trends not a problem (unlike event studies) as long as trends of the groups are parallel
- — Issues when go beyond simple binary DiD (*we discuss that later*)

$$Y_{it} = \beta G_i P_t + \lambda_G + \lambda_P + e_{it}$$

Nuts and bolts of fixed effects

Interpreting fixed effects

- **Group FEs**: compare individuals within the group
- **Time FEs**: compare individuals within a time period
- **TWFEs**:
 - Average of TEs identified from variation within group **and** variation within period
 - $\neq$ variation within “that group that year” (this would be group-year FEs)
- Including **FEs changes the estimand**: we compare observation within a group or within a time period

Illustration of pooled estimate

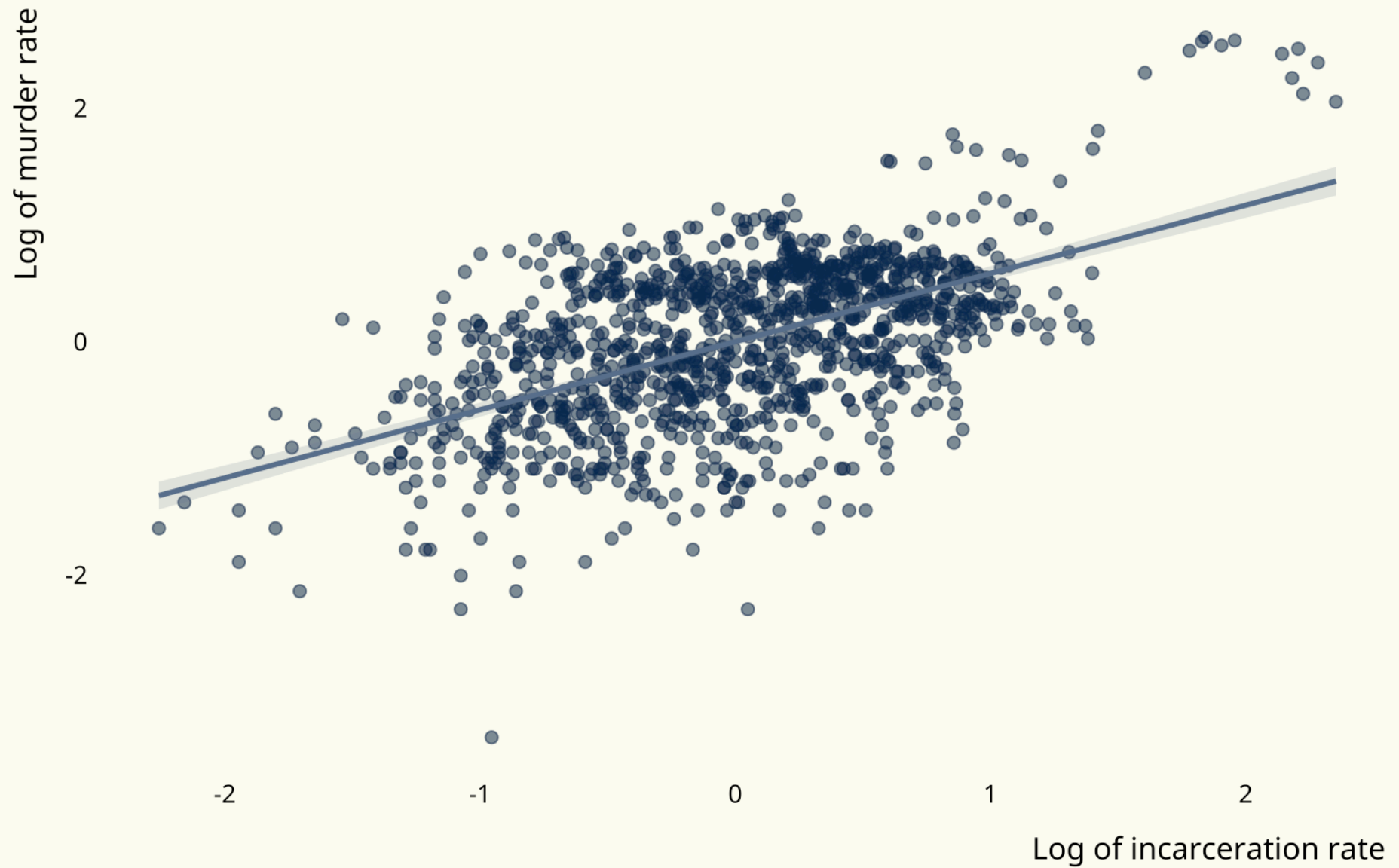
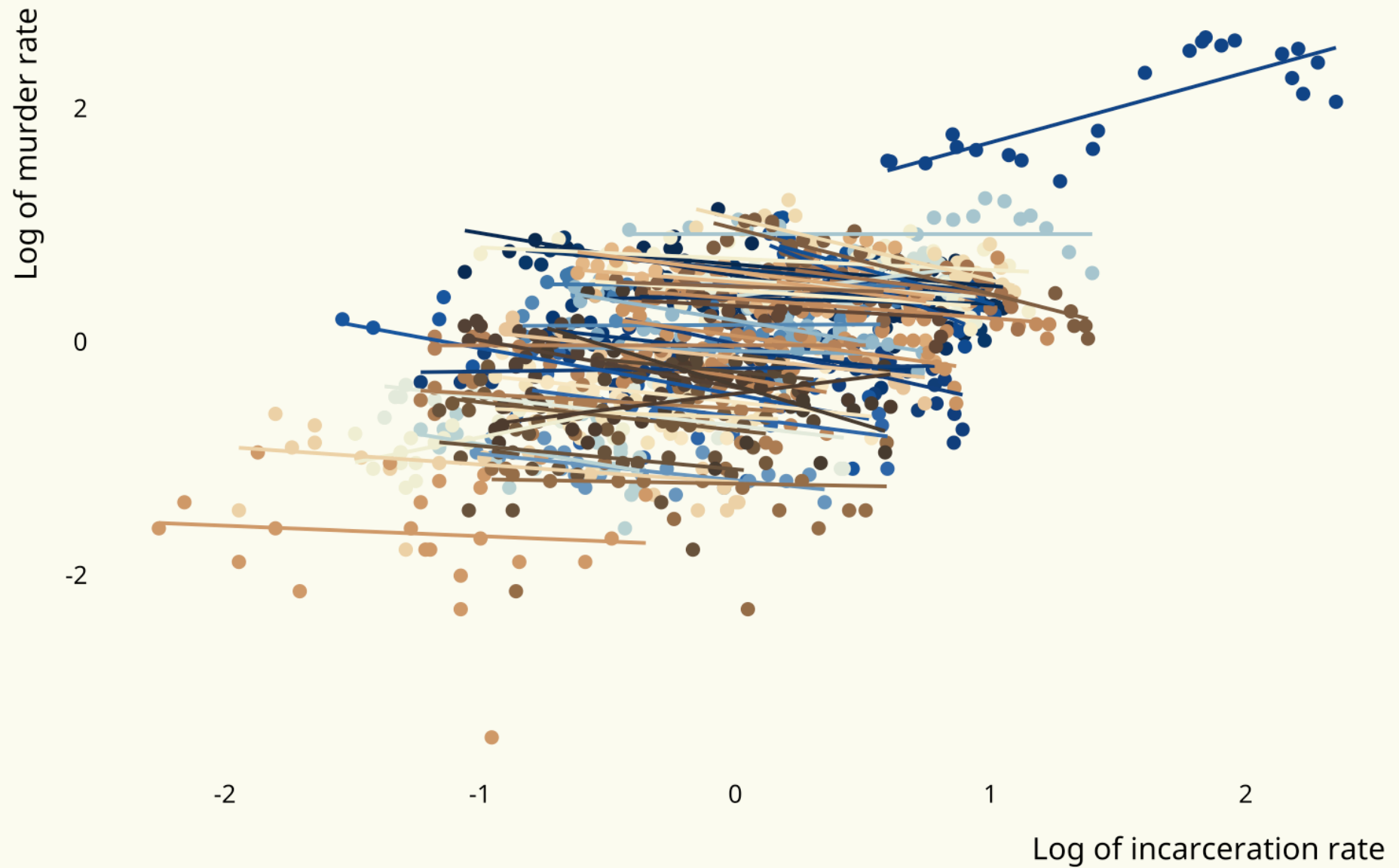
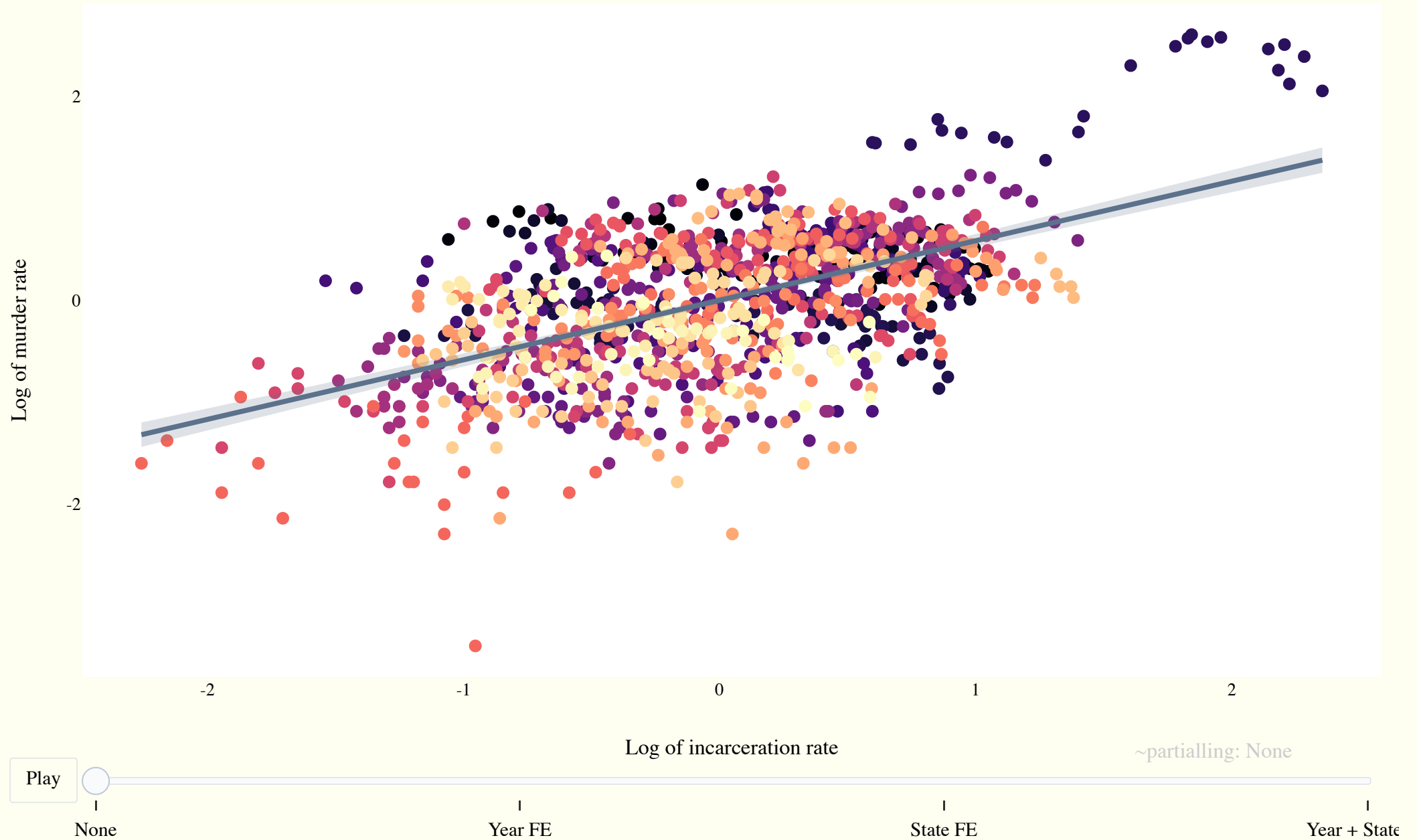


Illustration of within state relationship



Impact of Fixed Effects on the Estimand



Equivalence residual vs manual demean

```
1 #demeaning and showing that equal to residuals
2 sample_demean <- guns |>
3   mutate(
4     l_murder_res = feols(data = guns, fml = l_murder ~ 1 | state) |>
5       residuals()
6   ) |>
7   group_by(state) |>
8   mutate(mean_l_murder = mean(l_murder)) |>
9   ungroup() |>
10  mutate(
11    l_murder_demean = l_murder - mean_l_murder
12  ) |>
13  select(l_murder_res, l_murder_demean) |>
14  head(10)
```

l_murder_res	l_murder_demean
0.2824963	0.2824963
0.2170183	0.2170183
0.2094711	0.2094711
0.2094711	0.2094711
0.1057927	0.1057927
-0.0098917	-0.0098917
-0.1515422	-0.1515422
-0.1300360	-0.1300360
-0.0883633	-0.0883633
-0.0582103	-0.0582103

FWL with fixed effects

```
1 library(fixest)
2
3 #demeaning and showing that equal to residuals
4 guns_demean <- guns |>
5   mutate(
6     l_murder_res = feols(data = guns, fml = l_murder ~ 1 | state) |>
7       residuals(),
8     l_pris_res = feols(data = guns, fml = l_pris ~ 1 | state) |>
9       residuals()
10  )
11
12 reg_fe <- guns |>
13   fixest::feols(fml = l_murder ~ l_pris | state, cluster = "state") |>
14   broom::tidy() |>
15   mutate(reg = "fixed_effects", .before = 1)
16
17 reg_res <- guns_demean |>
18   feols(fml = l_murder_res ~ l_pris_res - 1, cluster = "state") |>
19   broom::tidy() |>
20   mutate(reg = "residualized", .before = 1)
21
22 rbind(reg_fe, reg_res) |>
23   kable()
```

reg	term	estimate	std.error	statistic	p.value
fixed_effects	l_pris	-0.15834	0.0365294	-4.334587	7.05e-05
residualized	l_pris_res	-0.15834	0.0365138	-4.336438	7.01e-05

Playing around

Instructions

Year FE

State FE

- Let's run some R code together to understand how fixed effects work
- Let's still will use the `guns` dataset
- Let's consider several regressions, with various sets of fixed effects
- Build graphs and interpret the coefficients

FE and Identifying Variation

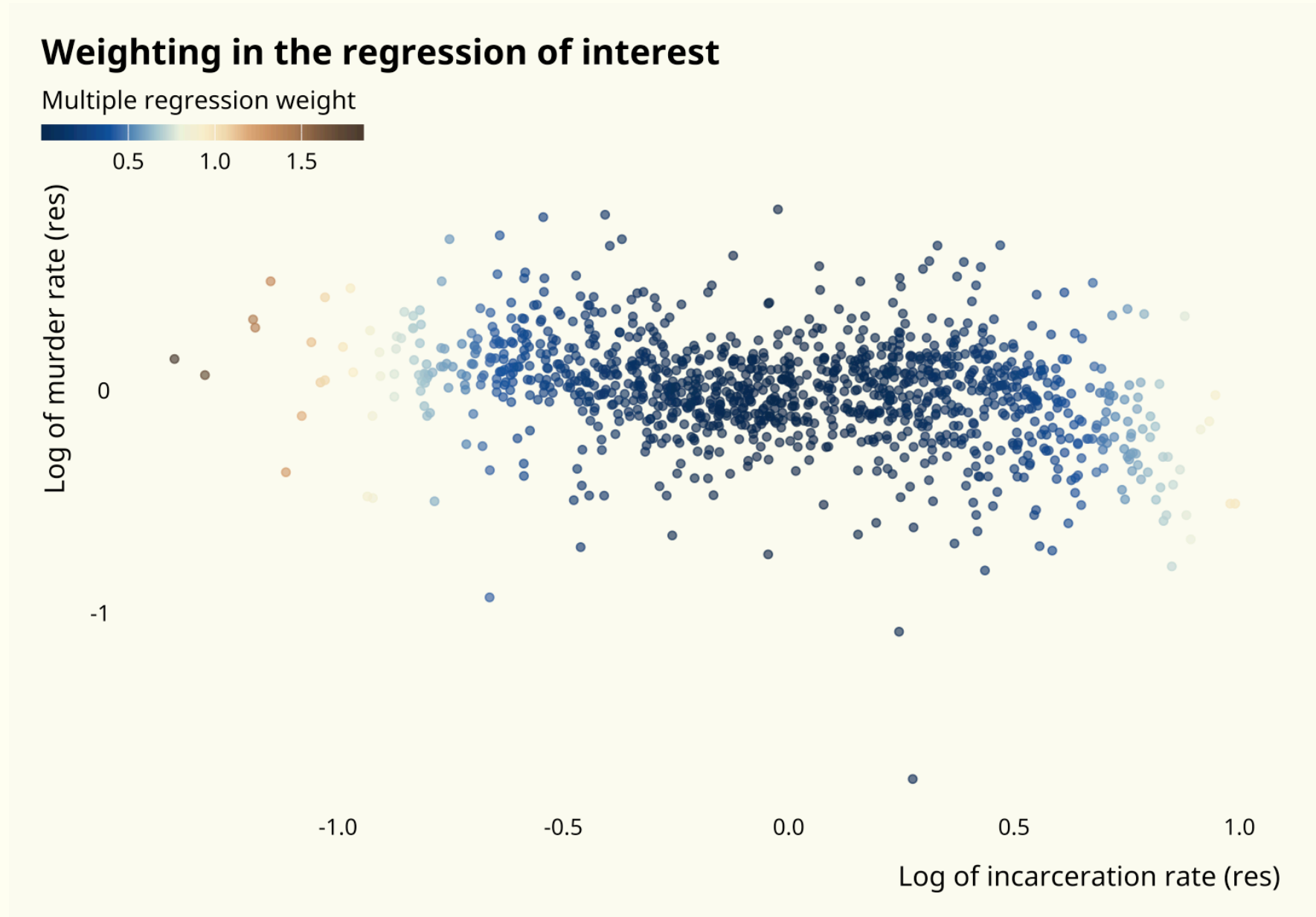
- When adding FE (or controlling in general), we partial out or absorb some of the variation
- We **throw out variation**
- Good if throw out variation that:
 - Is endogenous
 - Explains some of the variance of y $\left(\text{since } \mathbb{V}_{\beta^{\wedge}} = \frac{\sigma_u^2}{n\sigma_x^2} \right)$
- Bad if throw out **identifying variation**, ie variation that allows you to identify the effect of interest

Implicit weighting

Implicit weighting

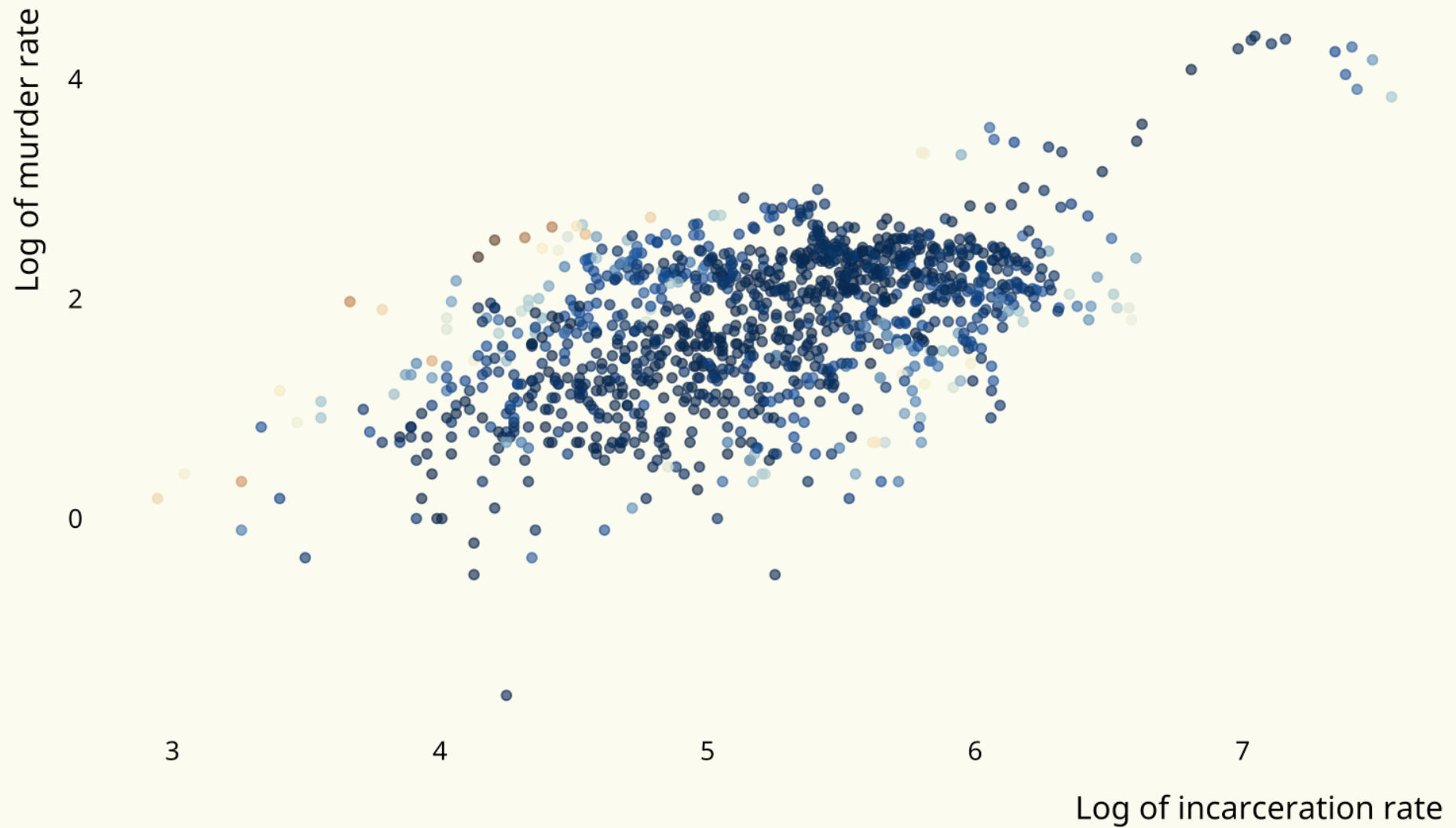
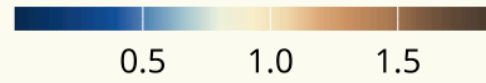
- Each unit does not contribute equally: there are many implicit weights
- At least two different objects can be weighted:
 - **Outcomes**: weighted difference in mean outcomes
 - **Treatment effects**: weighted mean of treatment effects
- Often important when effects are **heterogenous**
- Also important: **which comparisons** are used (ie who counts as “control”)

Aronow and Samii (2016) weights



Weighting in the raw regression

Multiple regression weights



Weighting outcomes

- Different angle: write $\hat{\beta}$ **directly as a weighted sum of outcomes**, not of τ_i

$$\hat{\beta} = \sum_i w_i Y_i$$

- Chattopadhyay and Zubizarreta (2023): characterizes these **implied weights** of linear regression (FE or not)
- Need not be **non-negative**, nor look like the weights of a randomized experiment
- $\Rightarrow$ possible **extrapolation**: $\hat{\beta}$ may not represent any real population
- Useful as a **diagnostic**: inspect the weights before trusting $\hat{\beta}$ (`lmw` package)

TWFE and 2x2 comparisons

- When treatment staggered, TWFE $\hat{\beta}$ is weighted average of many pairwise DiDs (Goodman-Bacon 2021)
- Each compares a newly treated group to a “control”:
 - Never-treated units, **or**
 - Not-yet-treated units
- Issue: some “controls” are **already-treated** units
 - Their own treatment effect gets implicitly subtracted out

Negative weights

- These “forbidden comparisons” can assign some $ATT_{g,t}$ a **negative weight**

$$\beta^{TWFE} = \sum_{g,t} w_{gt} ATT_{g,t}, \quad \sum_{g,t} w_{gt} = 1, \quad w_{gt} \text{ can be } < 0$$

- $\Rightarrow \beta^{TWFE}$ can have the **wrong sign**
- A simple simulation easily illustrates this
- Especially problematic when effects vary in time or are **heterogeneous**
 - Goodman-Bacon (2021); de Chaisemartin and D’Haultfœuille (2020)
- New estimators **reweight** to recover an interpretable, convex average
 - Callaway and Sant’Anna (2021), Sun and Abraham (2021), de Chaisemartin and D’Haultfœuille (2020), Borusyak, Jaravel, and Spiess (2024)

de Chaisemartin and D'Haultfœuille (2023)

- Can anyone **summarize** the idea?
- What do you get out of it?
- What did you think of it?

de Chaisemartin and D'Haultfœuille (2023)

- TWFE $\hat{\beta}$ = weighted sum of $ATT_{g,t}$'s, weights can be **negative**
- Arises in many standard staggered-adoption designs
- Reviews for binary/discrete/continuous treatments, staggered or not:
 - **Diagnostics**: how much of $\hat{\beta}$ comes from negative weights
 - **Heterogeneity-robust estimators**
- Revisits Wolfers (2006) on divorce laws: conclusions **change** across estimators

Calibrating simulations

Why calibrating?

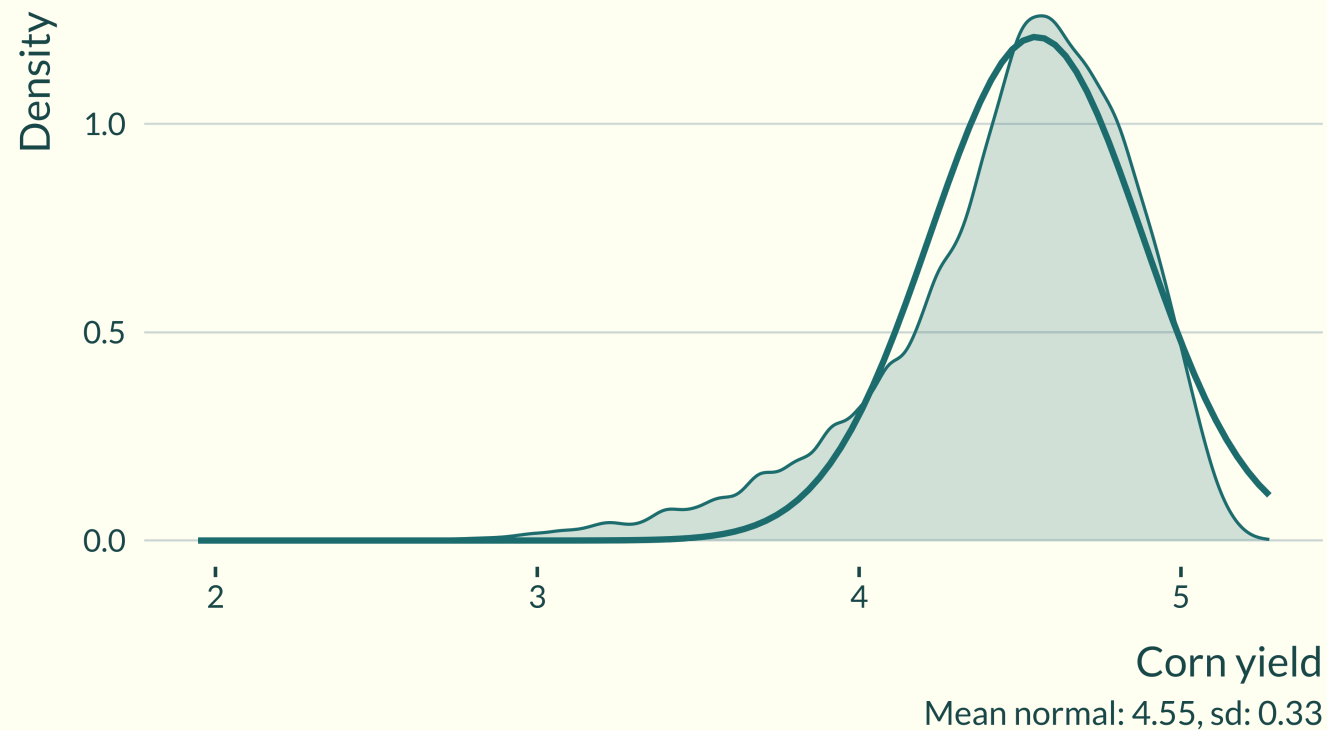
- So far, we considered very simple simulations, with “naive” distributions
- Calibrating can help make simulations more realistic
- But simulations will never be truly realistic
- Yet can still allow to run some sort of **robustness check** on the ability of your design to retrieve the effects of interest
- Also allows you to **think** about the DGP, your identification strategy, and so on

Fake data simulations

Distributions of the variables

- Emulate the distribution of variables in existing data sets

Distribution of corn yields
Tentative normal fit



Fake data simulations

Relationships between variables

- Read the literature
- Get a sense of **typical effect sizes and of relationships** between variables
- Make assumptions on those relationships. Acknowledge them.
- Complexify later if needed. You choose when you stop.
-  Varying parameters values might change the distribution of some “variables”
 - eg of the error term
 - Difficult to work *ceteris paribus*

Real data simulations

General approach

- Start from an existing data set
- Not yours. At least not the subset you are interested in
- Try to pick a subset where there is not already a treatment effect
- Define a treatment allocation mechanism
- **Add an artificial treatment effect** to the outcome variable in your initial data set, eg

$$Y_i(1) = Y_i(0) + \beta_i T_i$$

- Run your analysis and try to recover it

Real data simulations

Complexifying

- There is only one artificial aspect in such simulations: the treatment
- We can play on only 2 components:
 - **Who is treated?** *Treatment allocation*
 - Everyone
 - Only a subset of the population
 - **How?** *Treatment effect*
 - Homogenous
 - Heterogenous but random
 - Some specific correlation structure

Summary

Summary

- Today we reviewed:
 - The projection/geometric interpretation of regression
 - Identification strategies based on repeated observations
 - How fixed effects work, under the hood
 - **Implicit weighting**: of outcomes, of treatment effects
 - Issues with TWFE under heterogeneous effects
- Hopefully you have a better understanding of:
 - Causal inference, from a bird's view
 - **How fixed effects really work**
 - Many details and **intuitions**

Take away messages

- The choice of FE is crucial and **affects the estimand**
- **FE can remove a lot of variation:**
 - Great if removes endogenous variation
 - Problematic if there is too little variation left
- Regressions (with or without FE) **implicitly weight** units
- With heterogeneous or dynamic effects, plain TWFE can be misleading

References

- Angrist, Joshua D., and Jörn-Steffen Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion*. 1 edition. Princeton: Princeton University Press.
- Aronow, Peter M., and Cyrus Samii. 2016. "Does Regression Produce Representative Estimates of Causal Effects?" *American Journal of Political Science* 60 (1): 250–67. <https://doi.org/10.1111/ajps.12185>.
- Borusyak, Kirill, Xavier Jaravel, and Jann Spiess. 2024. "Revisiting Event Study Designs: Robust and Efficient Estimation." *Review of Economic Studies* 91 (6): 3253–85. <https://doi.org/10.1093/restud/rdae007>.
- Callaway, Brantly, and Pedro H. C. Sant'Anna. 2021. "Difference-in-Differences with Multiple Time Periods." *Journal of Econometrics* 225 (2): 200–230. <https://doi.org/10.1016/j.jeconom.2020.12.001>.
- Chattopadhyay, Ambarish, and José R. Zubizarreta. 2023. "On the Implied Weights of Linear Regression for Causal Inference." *Biometrika* 110 (3): 615–29. <https://doi.org/10.1093/biomet/asac058>.
- de Chaisemartin, Clément, and Xavier D'Haultfœuille. 2020. "Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects." *American Economic Review* 110 (9): 2964–96. <https://doi.org/10.1257/aer.20181169>.
- . 2023. "Two-Way Fixed Effects and Differences-in-Differences with Heterogeneous Treatment Effects: A Survey." *The Econometrics Journal* 26 (3): C1–30. <https://doi.org/10.1093/ectj/utac017>.
- Goodman-Bacon, Andrew. 2021. "Difference-in-Differences with Variation in Treatment Timing." *Journal of Econometrics* 225 (2): 254–77. <https://doi.org/10.1016/j.jeconom.2021.03.014>.
- Sun, Liyang, and Sarah Abraham. 2021. "Estimating Dynamic Treatment Effects in Event Studies with Heterogeneous Treatment Effects." *Journal of Econometrics* 225 (2): 175–99. <https://doi.org/10.1016/j.jeconom.2020.09.006>.