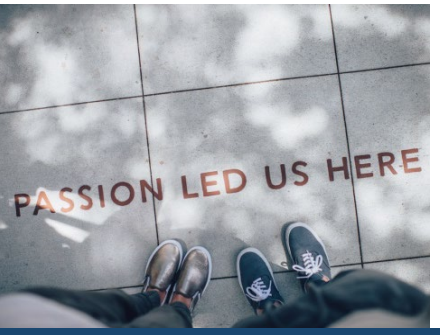


Lab 4: The Tennessee STAR Experiment

Gregory Bruich, Ph.D.

HARVARD UNIVERSITY

OPPORTUNITY INSIGHTS

Tennessee Student-Teacher Achievement Ratio (STAR) Experiment

Source: Mosteller (1995)

Lab 4: The Tennessee STAR Experiment

- In today's lab, we will analyze data from the Tennessee STAR experiment
- Student/Teacher Achievement Ratio (STAR) experiment is one of the most widely studied education interventions
  - Conducted from 1985 to 1988 in Tennessee
  - One cohort of 11,571 children in grades K-3 at 79 schools
  - Students and teachers randomized into classrooms *within* schools
  - Class size differs: small (~15 students) or large (~22 students)
- Coding exercise: estimate the impact of small class on kindergarten test scores, taking into account that this was a *stratified randomized experiment*

Key Lessons from Lab 4

- Substantive question: do small classes improve student achievement?

- Key methodological tools:

  1. Randomization tests, statistical significance, and practical significance

  2. Converting test scores into a standardized index (standard deviation units)

  3. Using histograms to visualize results from an experiment

  4. Using bar graphs to visualize results from an experiment with 95% confidence interval bars to show precision of estimate

  5. Multivariable regression with fixed effects as control variables for treatment effect estimation in a stratified randomized experiment

Quantifying Uncertainty using a 95% Confidence Interval

- For policy decisions, it is critical to accurately quantify uncertainty associated with our estimates

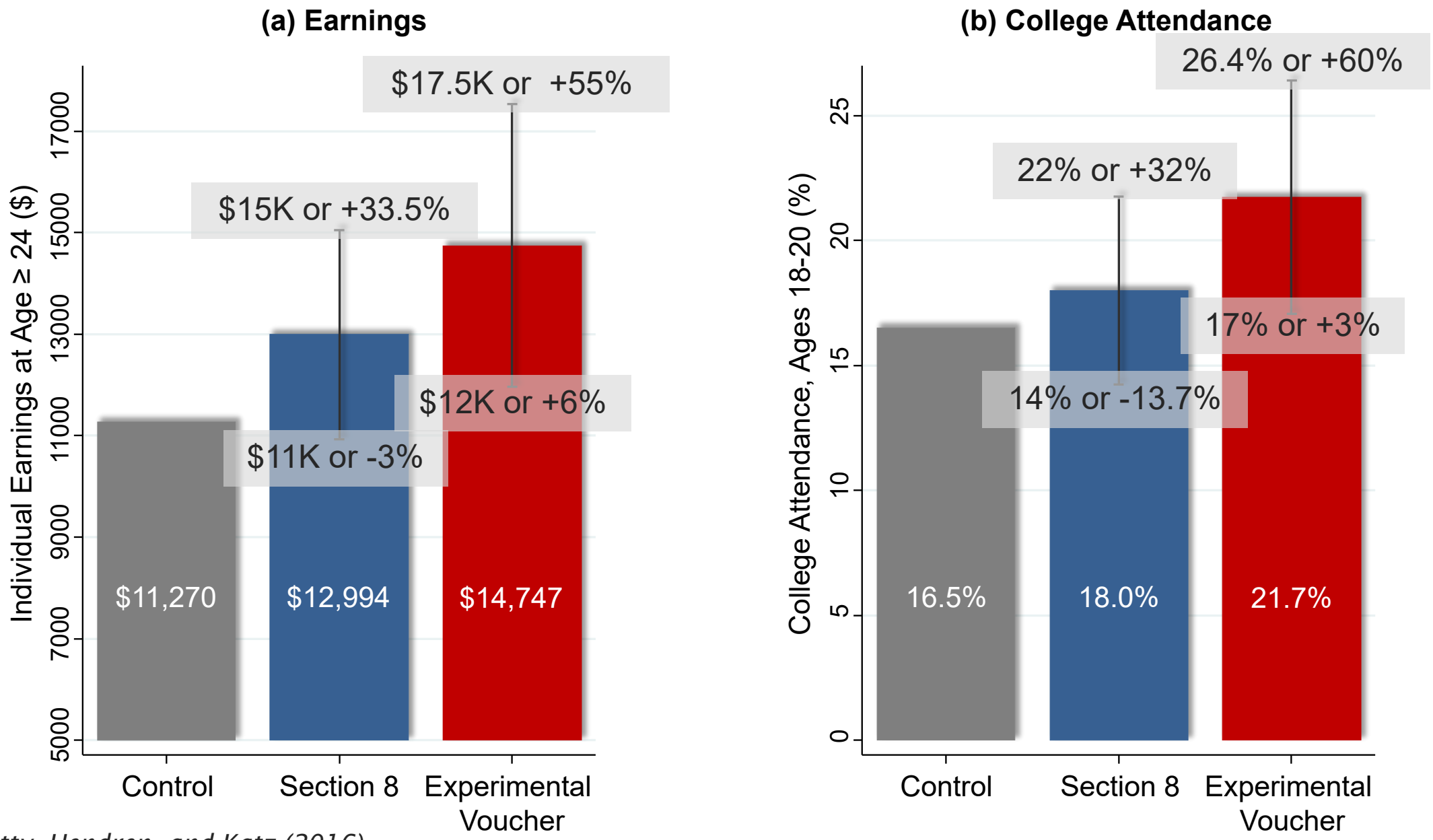
- Source of uncertainty in an experiment: if different subjects had been assigned to different groups, we would observe different potential outcomes

- We do this by constructing 95% confidence intervals:

$$\text{Estimated Difference} \pm 1.96 \times \text{standard error}$$

- The 95% confidence interval gives us the range of values that are consistent with the data that we observe

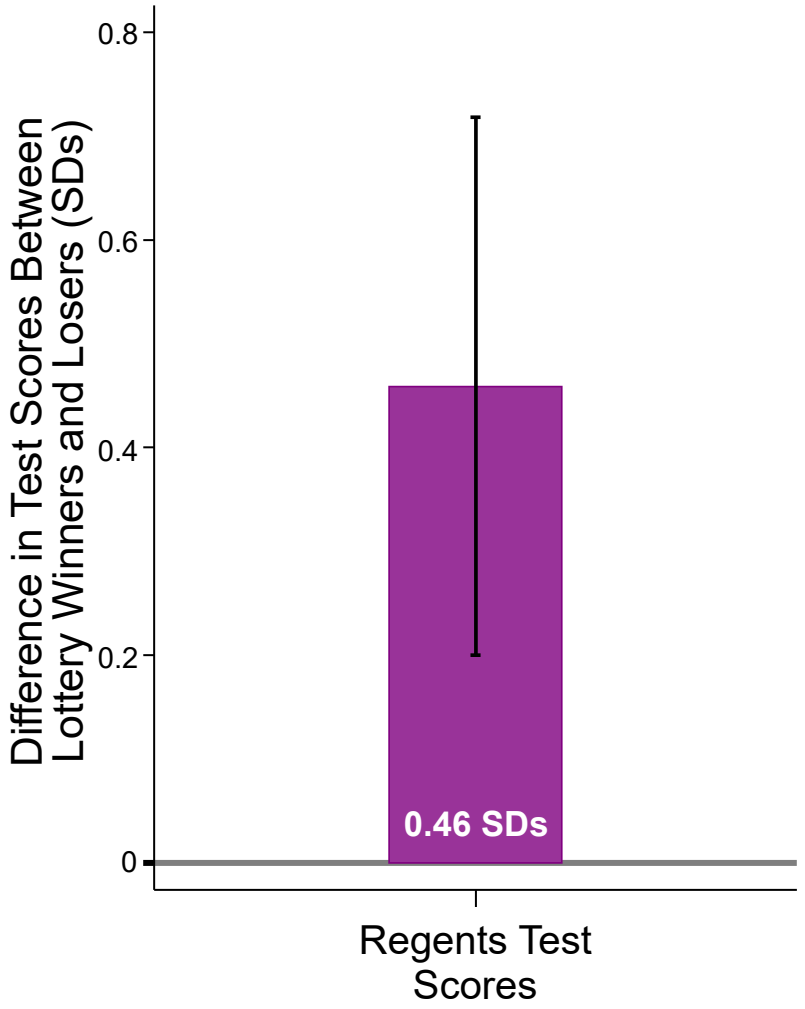
Impacts of MTO on Children Below Age 13 at Random Assignment



Source: Chetty, Hendren, and Katz (2016)

Impacts of Harlem Children's Zone on Test Scores

Promise Academy Lottery Winners vs. Losers' Outcomes



Test scores: z-scores or standard deviation units

- In the analysis of the Harlem Children's Zone Promise Academy, test scores were converted into z-scores or standard deviation units:

$$\text{z-score} = \frac{test\ score - \text{mean}}{\text{standard deviation}}$$

- The estimated impact was 0.46 standard deviations or 0.46σ's

- To help judge the magnitudes, recall from lab 1 that most of the data will usually be within 1 standard deviation of the mean and almost all the data will usually be within 2 standard deviations of the mean

- In an experiment, it is common to subtract the control group's mean and divide by the control group's standard deviation

How to analyze multiple outcomes?

- In the STAR experiment, there were four Stanford Achievement Tests that students took at the end of Kindergarten school year

- In question 3, you will combine all four tests into one index:

$$\text{Index} = \frac{\text{zscore}_1+\text{zscore}_2+\text{zscore}_3+\text{zscore}_4}{4}$$

- Since all the terms in the numerator are in standard deviation units or σ's, the index is also in standard deviation units or σ's

- Forming an index addresses issues related to multiple hypothesis testing and can increase the likelihood of detecting an impact if there is any (called "statistical power")

Are the data consistent with random assignment of teachers to classes?

- In question 4, you'll estimate simple regressions like in Lab 3 to check for balance of pre-determined teacher characteristics:

$$\hat{Y}_i = \alpha_0 + \alpha_1 Small_i$$

- Recall from Lab 3 that α_1 is the estimated difference in means for the treatment group versus the control group

- Judge statistical significance using 95% confidence interval for α_1

- Judge practical significance by thinking about whether α_1 is large in a real world sense: this requires *judgement* but it is not the case that anything goes!

Estimation of Treatment Effects via Multivariable Regressions

- In question 5, you'll estimate *multivariable regressions:*

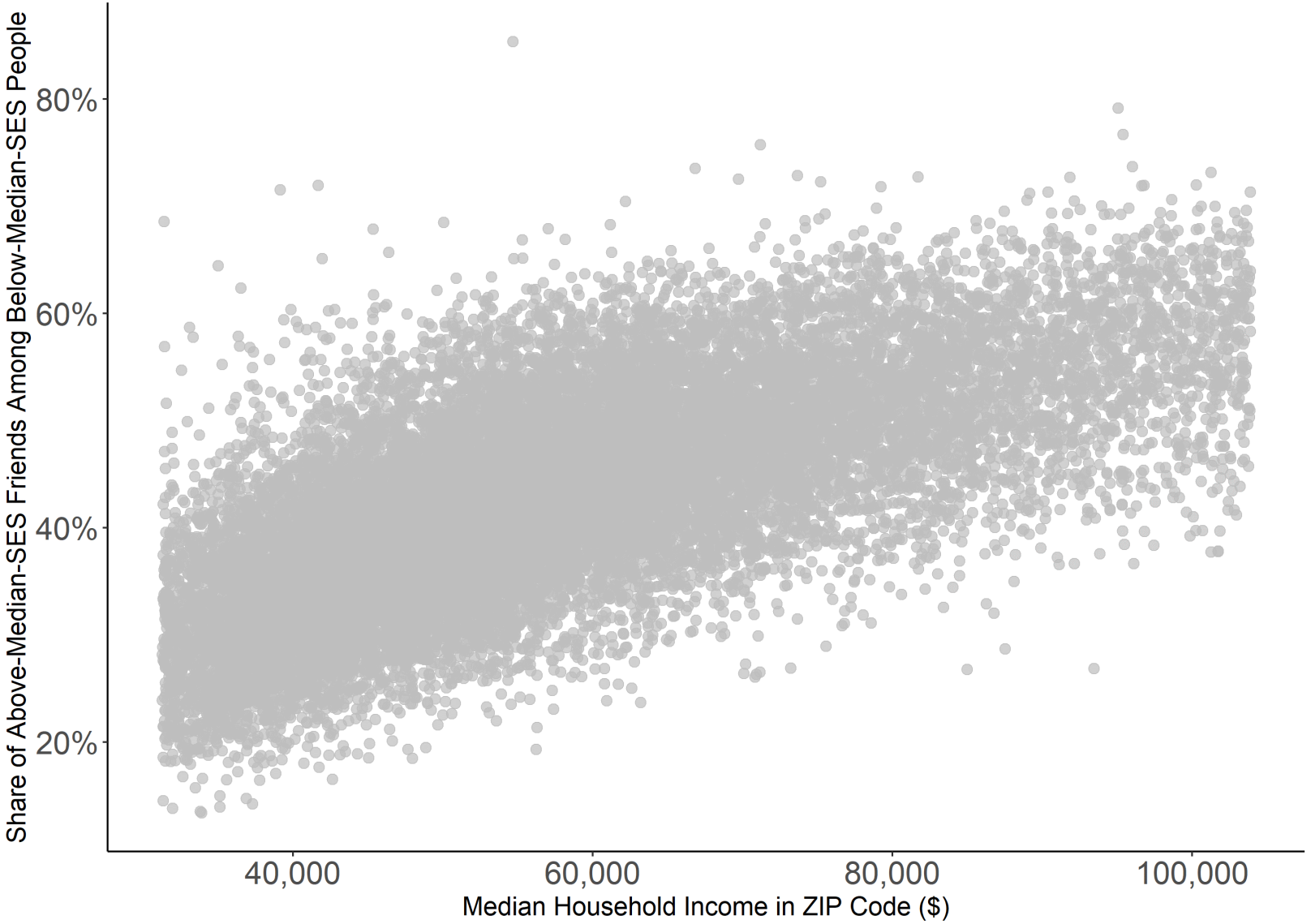
$$\hat{Y}_i = \alpha_0 + \alpha_1 Small_i + \alpha_2 School1_i + \alpha_3 School2_i + \cdots + \alpha_{79} School78_i$$

- The coefficient α_1 is the estimated treatment effect and all the other coefficients can safely be ignored (and don't have useful interpretations)

- Create an Opportunity Insights style bar graph of the control group mean $\bar{Y}_0$ and the adjusted mean for the treatment group $\bar{Y}_0 + \alpha_1$, along with 95% confidence intervals in brackets

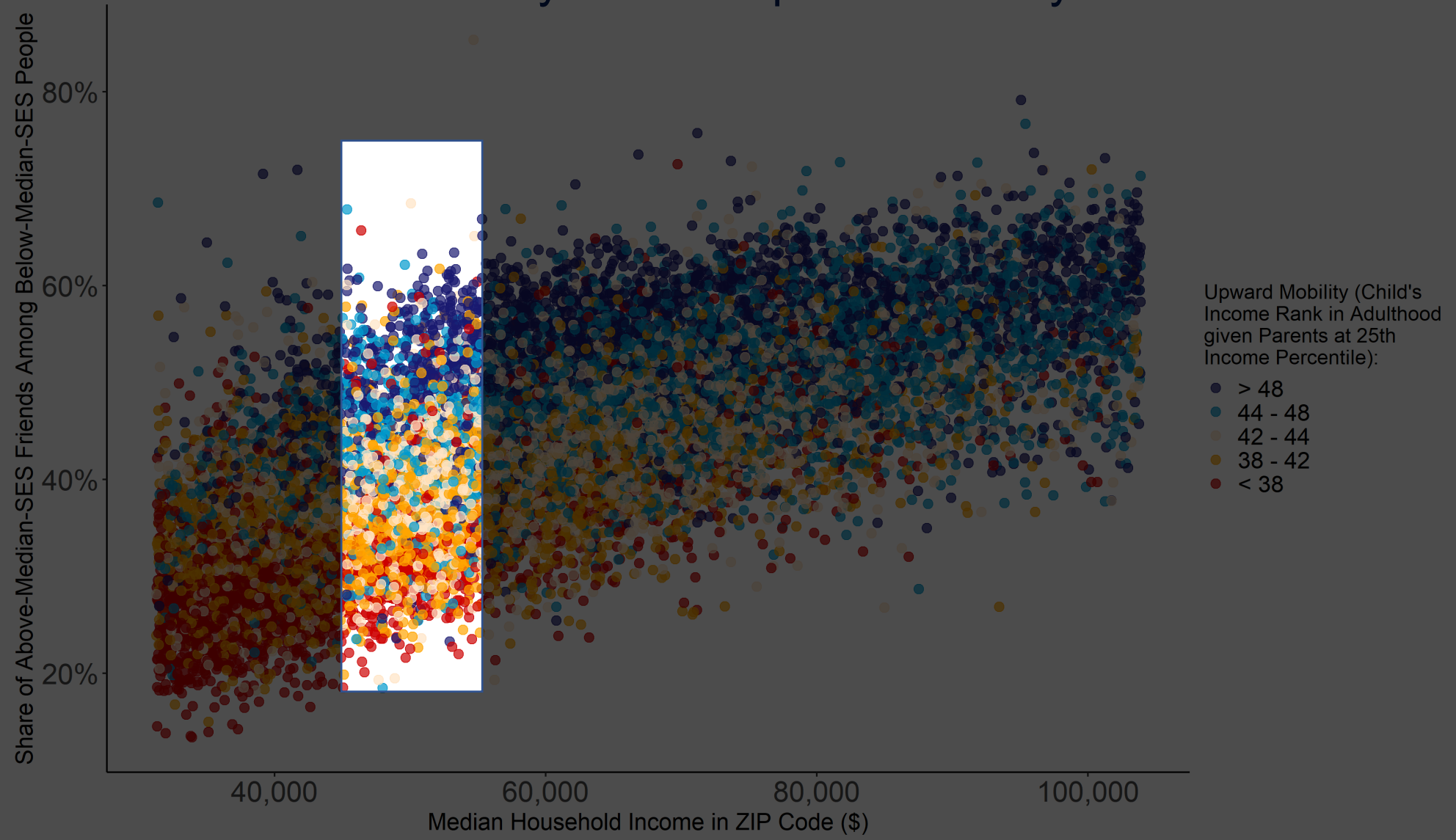
- Where have we seen *multivariable regressions* before? Social capital lecture

Economic Connectedness vs. Household Median Income, by ZIP Code



Economic Connectedness vs. Household Median Income, by ZIP Code

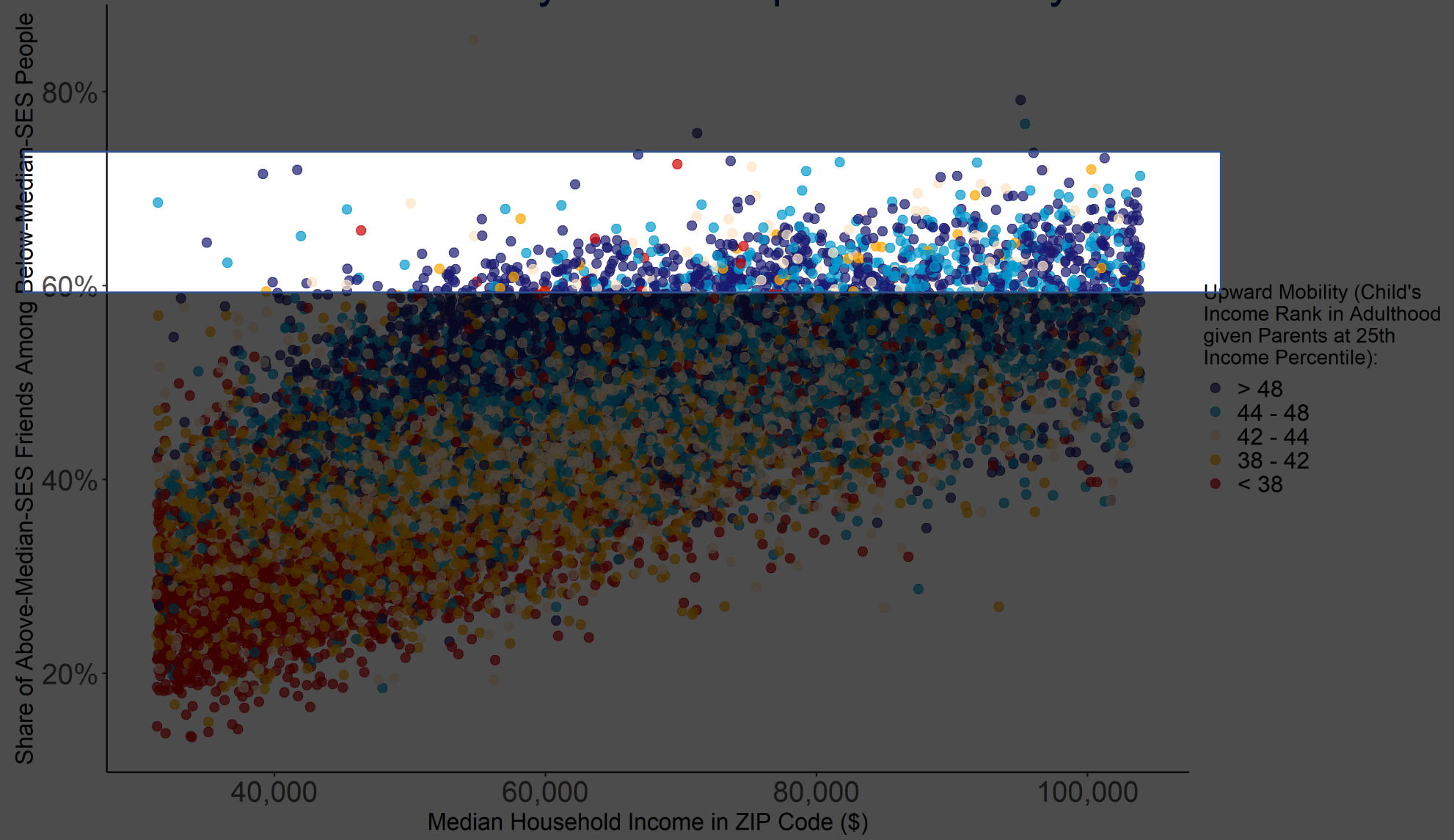
Colored by Rate of Upward Mobility



Source: Chetty, Jackson, Kuchler, Stroebel et al. (Nature 2022a,b)

Economic Connectedness vs. Household Median Income, by ZIP Code

Colored by Rate of Upward Mobility



Source: Chetty, Jackson, Kuchler, Stroebel et al. (Nature 2022a,b)

Upward Mobility vs. Economic Connectedness, Inequality, and Segregation

Regression Estimates, Across Counties

Dependent Variable:	Upward Mobility (Mean Income Rank at Age 35 for Children with Parents at 25th Percentile)	
	Across Counties	
	(1)	(2)
Income Inequality (Gini coefficient)	–0.449***	-0.103
	(-0.084)	(-0.091)
Share Black		
Economic Connectedness		0.577***
		(0.063)
Observations	2,741	2,741
R-squared	0.207	0.424

Connectedness explains the link between inequality and mobility (Great Gatsby Curve) [Corak 2013, Krueger 2016]