

Econ 174, Section 101/103

Week 10

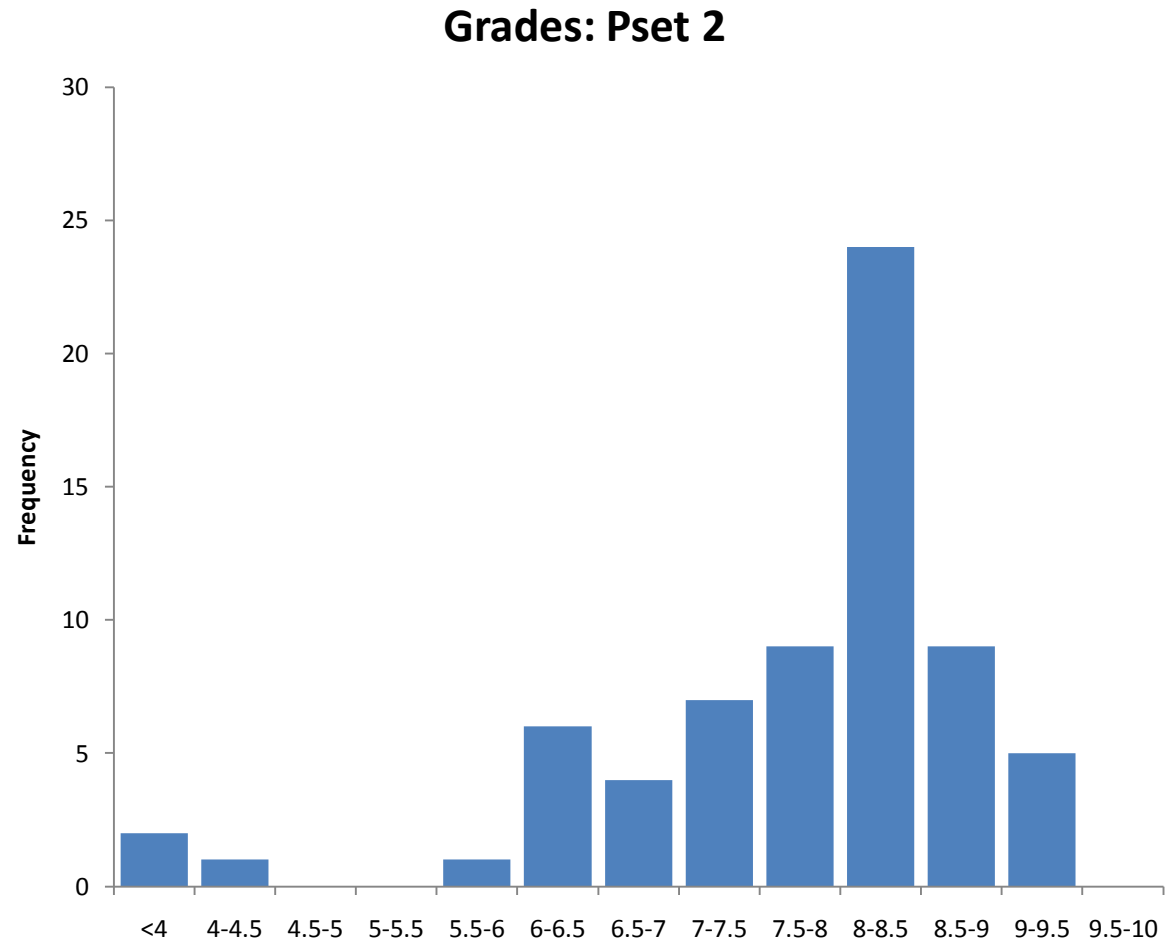
Joshua Blumenstock

# Today

- General questions?
- PSET 2 review
- Reading journal articles
- Propensity score matching
  
- Next week
  - Centering interaction terms?

# PSET 2 review

- Mean: 7.78
- SD: 1.45
- Median: 8.25
- Mode: 8.25



# PSET 2: common errors

- Robust standard errors
- Misinterpreting coefficients on log regressions
- Question 2C: what is the bias?
- Testing for significance of difference
- Centering interaction terms
- Baird et al (2010)

# How to read a journal article

- Abstract
- Introduction
- Conclusion
- Figures & Tables
- Skim Text
- By the end, you should know
  - Main results
  - Ancillary results
  - Context & data (roughly)
  - Methods (including estimation/identification strategy)
  - How authors address limitations/robustness
  - Weaknesses and criticism

# Propensity score matching

- Jalan & Ravallion (2003)
  - What is the main result of the paper?
    - the prevalence and duration of diarrhea among children under five in rural India are significantly lower on average for families with piped water than for observationally identical households without it
  - What ancillary results?
    - Health gains largely by-pass children in poor families, particularly when the mother is poorly educated.
  - What is the data used?
    - cross-sectional survey for rural India implemented in 93–94
  - What methods?
  - Limitations & robustness

# Propensity score matching

- Different types of matching
  - Randomization
  - Exact matching
  - “Fuzzy matching”
  - Composite matching (e.g. weighted average)
  - Propensity score matching

# Propensity score matching

- Each treated person is matched with an observationally similar control, then the average difference in outcomes across the two groups is compared to get treatment effect
- If outcomes are independent of treatment given  $X_i$ , then outcomes are also independent of treatment given  $P(X_i)$ , just as they would be if treatment were assigned randomly.



# PSM: The Big Assumption

- “Unconfoundedness” aka “conditional independence” aka “selection on observables”
- In notation:  $(Y_i^T, Y_i^C) \perp T_i | X_i$
- In Fred’s words: given a set of observable covariates  $X$  that are not affected by treatment, potential outcomes  $Y$  are independent of treatment assignment  $T$ .
- In Josh’s words: to the extent that there is systematic selection into treatment, this selection is only a function of observable variables.

# PSM: allows for matching

- Nearest neighbor
- Radius/caliper
- Stratification/interval
- Kernel weighting

# PSM

- Assuming selection on observables, constant & linear treatment effects, can estimate

$$Y_i = \beta_0 + \beta_1 T_i + \beta_2 \hat{P}(X) + u_i$$

- If  $E[T_i | X_i]$  is linear, just include the  $X_i$  in the regression, but more flexible  $E[T_i | X_i]$  is propensity score.
- Adding interaction term allows for heterogeneity of treatment effect:

$$Y_i = \beta_0 + \beta_1 T_i + \beta_2 \hat{P}(X) + \beta_3 (T_1 \times (\hat{P}(X) - E[\hat{P}(X)])) + u_i$$

# End week 10

- Next week
  - Centering interaction terms
- After break
  - PSM in stata