ECON 1123 Section 8

Slides at github.com/cjleggett/1123-section

Outline

- Name Circle
- Problem Set Feedback
- Lecture Recap / Questions + Practice Problems
    - Weak Instruments
    - Intrinsic Heterogeneity
    - J Statistics

Name Circle

Name Circle

- Name
- Where you're from

Problem Set Feedback

Problem Set 7 Feedback

- OVB Problems: Omitted Variable should be measured before treatment.
- Exogeneity: Randomness is not sufficient
- Reduced form does not give any insights into exogeneity

Lecture Recap

Weak Instruments

Weak Instruments

- If our instruments are weak (π_1 close to 0) then our results are biased toward OLS results
- We can test against this using “effective First-Stage F-statistic” or F^{MOP}
- For just one instrument, we already practiced this.
- For more than one instrument, we need to use code for calculation
- But same idea, compare F^{MOP} with critical value

Weak Instruments

- What if we have some weak instruments?
  - Throw them away
  - Use tools robust to weak instruments

Anderson and Ruben CI

- Works even in the case of weak instruments!
- We have code to help you build this

Anderson and Ruben CI: Math

- For each potential β_1:
  - Use OLS with X and Y to find residuals using β_1
  - Use OLS with residuals on left and instruments on the right
  - Use an F test to see if the instruments are significant
  - If they are significant, then assuming exogeneity, we have the wrong β_1
  - If they are not significant, then include β_1 in the CI

Many Weak Instruments

- If we have a lot of weak instruments, sometimes the results will be misleading
- Solve this by not having a lot of weak instruments
- We can use LASSO to eliminate instruments
- You won’t have to do this in class

Nonlinear IV Regression

- Sometimes we want to introduce nonlinearity into our models, but we have to be careful in IV regressions
- DO NOT use probit/logit in first-stage
- If you want effect of X and X^2, use instruments Z and Z^2
- If you want effect of X and X*I, use instruments Z and Z*I

Intrinsic Heterogeneity

ATE vs. LATE

- ATE: Average Treatment Effect
- LATE: Local Average Treatment Effect
- Treatments have different effects on different people. (eg. College)
  $\beta_i = Y_i(1) - Y_i(0)$ vs. $\beta_j = Y_j(1) - Y_j(0)$
- We can't measure individual effects, so we hope to measure ATE
  $E[\beta_i]$

ATE vs. LATE

- In IV Regressions, we actually estimate the LATE:

$$LATE = \frac{E[\beta_{1i} \times \pi_{1i}]}{E[\pi_{1i}]} = E[\beta_{1i}] + \frac{cov(\beta_{1i}, \pi_{1i})}{E[\pi_{1i}]}$$

- LATE is a weighted average of treatment effects, where the weight is $\frac{\pi_{1i}}{E[_{1i}]}$
- Who is this higher for?
  - People effected most by the instrument

ATE vs. LATE

- Why is this important?
- IV Results only apply to those who respond to instrument (Compliers) (Grasshoppers)
- So it’s as if we’re only studying the compliers

Which is Bigger, LATE or ATE?

- New problem type for psets + exams!
- Solve with three steps:
  1. Explain why ATE is not equal to LATE
  2. Use intuition to think of who the compliant people (grasshoppers) are
  3. Use intuition to decide whether treatment effect is higher or lower for compliant people (grasshoppers) and form conclusion

Step 1: Eliminate Three Cases

- π_{1i} does not vary
- β_{1i} does not vary
- $cov(\pi_{1i}, \beta_{1i}) = 0$
- Very Rare!

Step 2: Identify Compliant Subjects

- Who responds more strongly to the **instrument**
- Explain verbally

Step 3: Compare compliant to population

- LATE puts more weight on compliant
- Who will have a larger β_{1i}? Use intuition.
- Use this to determine whether LATE is bigger than or smaller than ATE **in absolute value.**

ATE vs. LATE: Which is bigger?

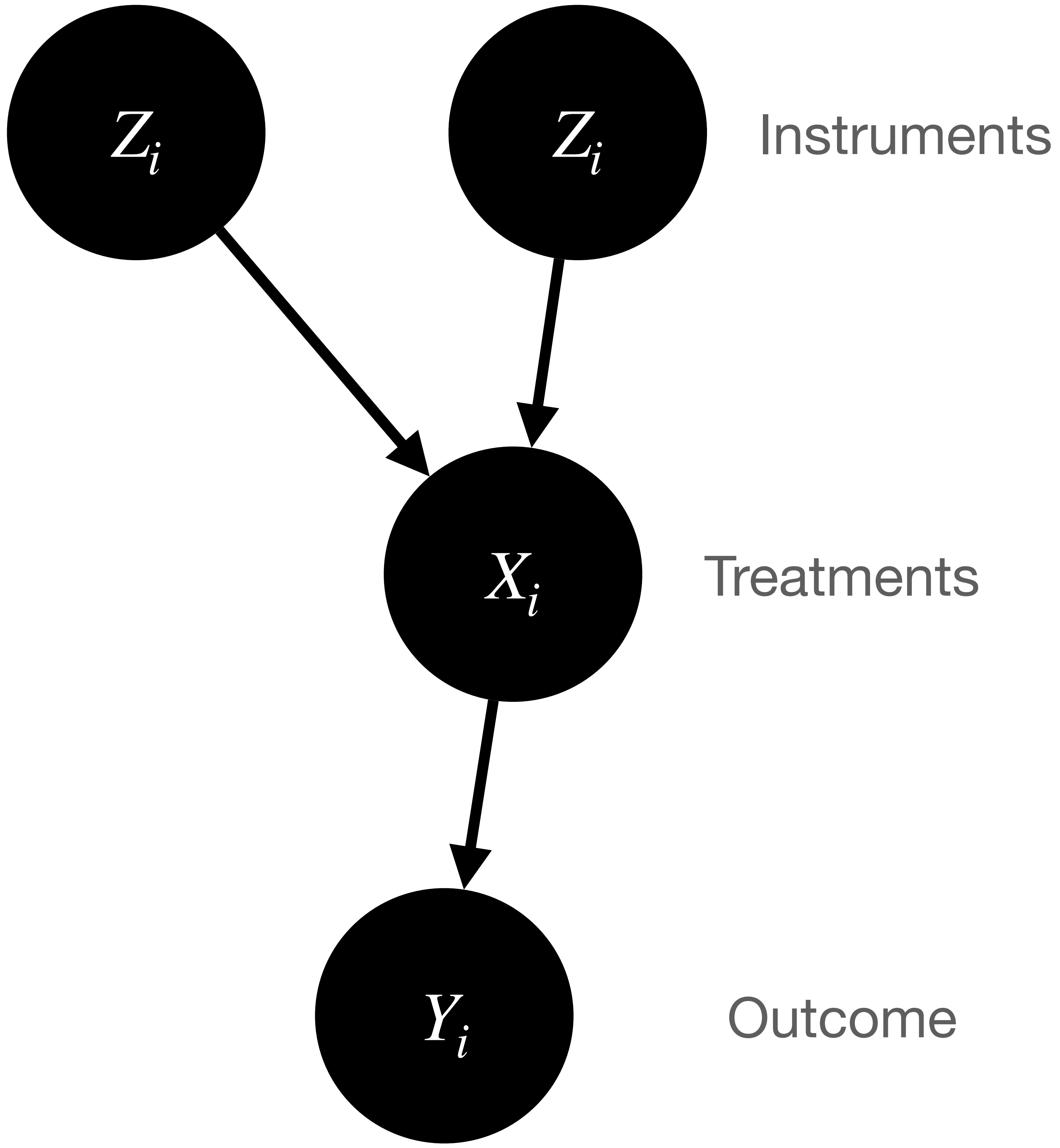
$\beta^{Compliers}$ bigger than $\beta^{Non-Compliers}$	LATE **bigger than** ATE
$\beta^{Compliers}$ smaller than $\beta^{Non-Compliers}$	LATE **smaller than** ATE

J Statistic

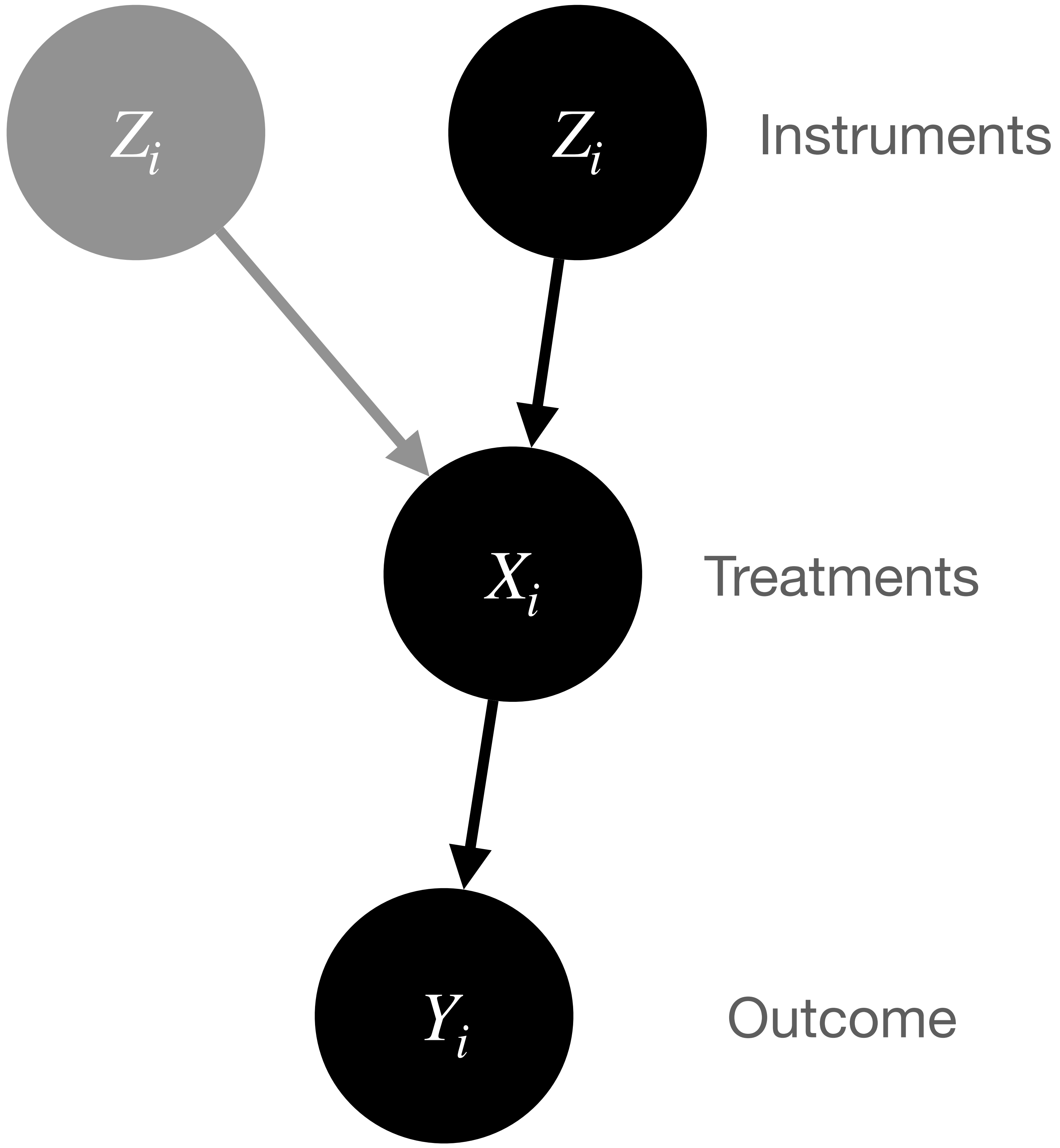
J Test of over identifying restrictions

- Sometimes we have more instruments than X variables
- J Test checks whether or not each instrument individually identifies the same treatment effect
- If we reject that they are the same:
    - Something may be wrong: one or more variables not exogenous
    - Everything is fine, but the two variables identify different LATEs

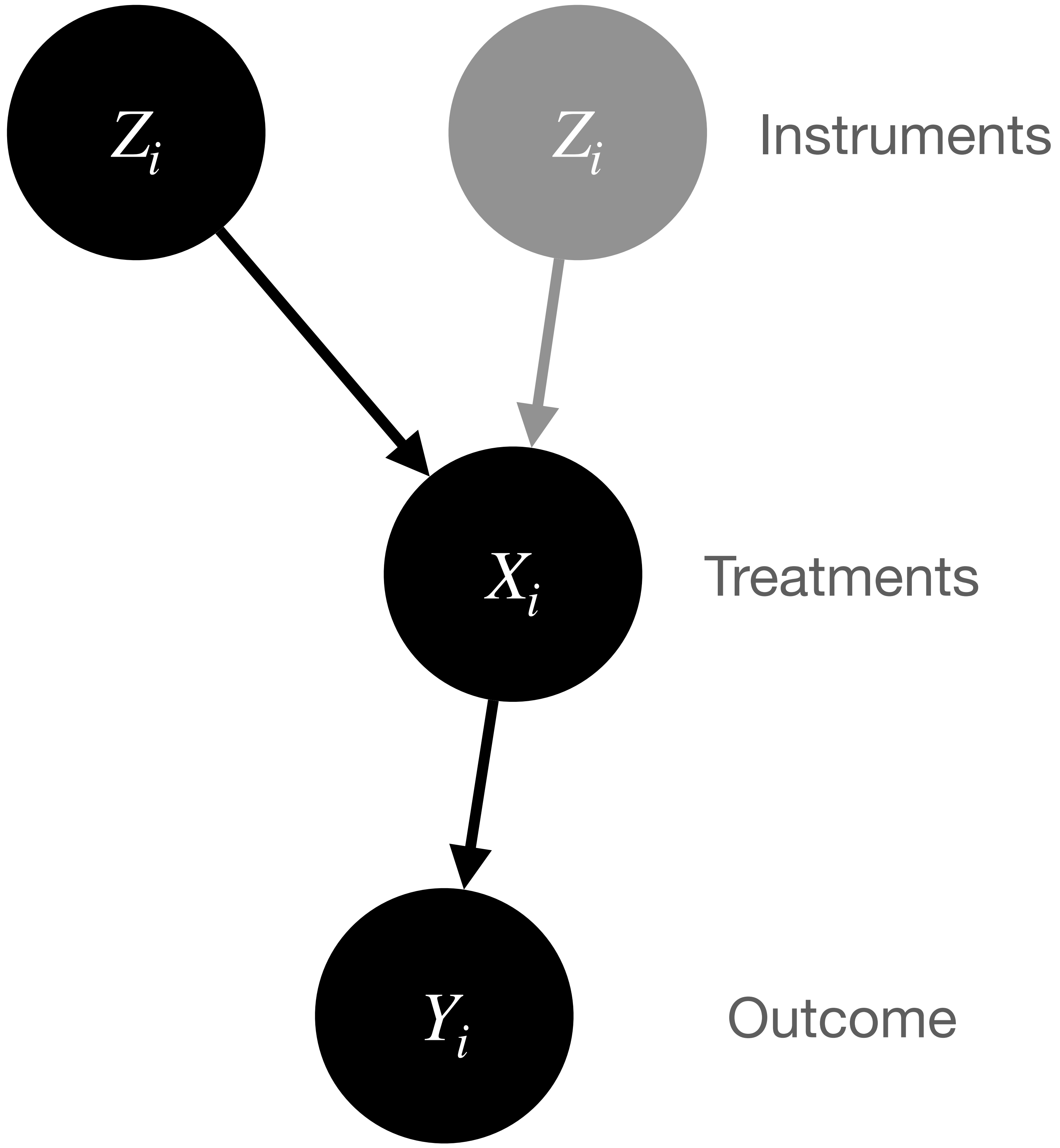
J Test Interpretation



J Test Interpretation



J Test Interpretation



J Test Math

- First step is to estimate 2SLS using both instruments
- Then calculate residuals from 2sls using the actual X_i

$$\hat{u}_i \equiv Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_{1i} - \hat{\beta}_2 W_i$$

- Now run OLS with homoskedastic standard errors (no `, robust`):

$$\hat{u}_i = \lambda_0 + \lambda_1 Z_{1i} + \lambda_2 Z_{2i} + \lambda_3 W_i + v_i$$

- Test hypothesis that H_0: $\lambda_1 = 0$ and $\lambda_2 = 0$
- The *J* statistic = # instruments × *F* stat
- Distributed as χ^2_k with dof = # instruments– # endogenous X_i's

Exercises!