Structural Causal Models & Instrumental Variable

Group 22:
Mentor: Shiyu He
Mentee: Anne Zheng & Yifan Cao

Summary

Motivation

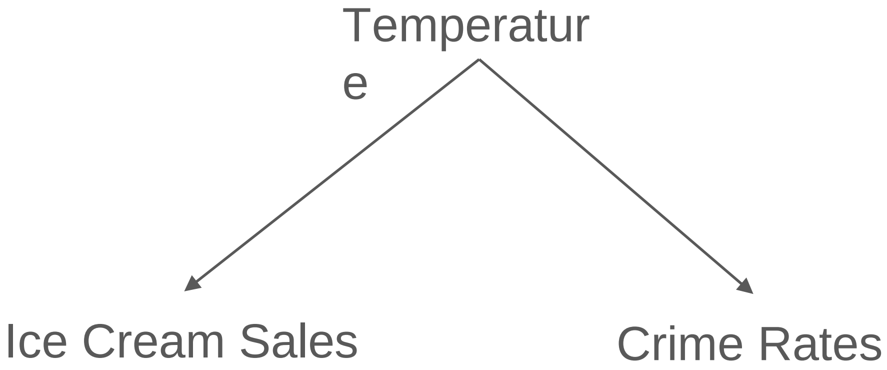
Regression VS Causation

Structural Causal Model Theory by Judea Pearl

Final Project: Instrumental Variable

Motivation - Omitted Variables

An increase in ice cream sales is correlated with an increase in violent crime—not because ice cream causes crime, but because both ice cream sales and violent crime are more common in hot weather.



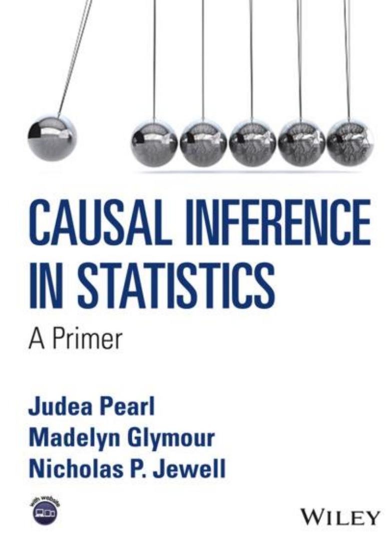
Problems of Correlation (Endogeneity)

Errors in Variables (lie in the surveys)

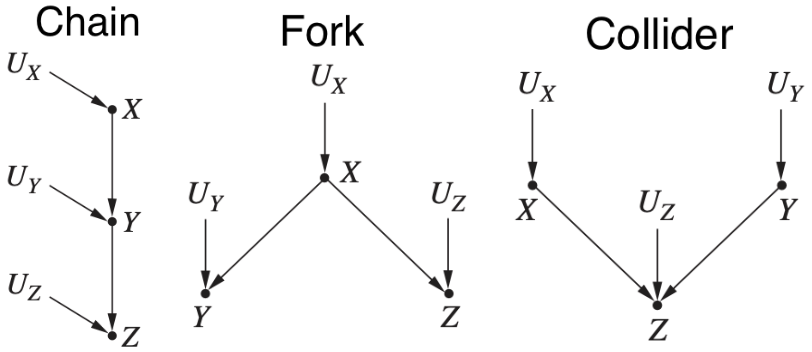
Reverse Causality (ability = a * income)

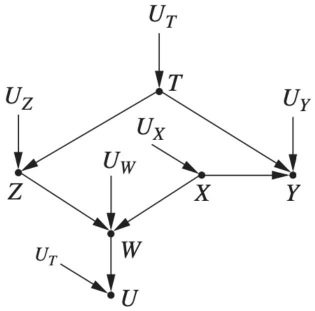
Simultaneity (e.g. Supply and Demand)

Omitted Variables (ice cream sales & crime rate)



Graphical Causal Models Theories





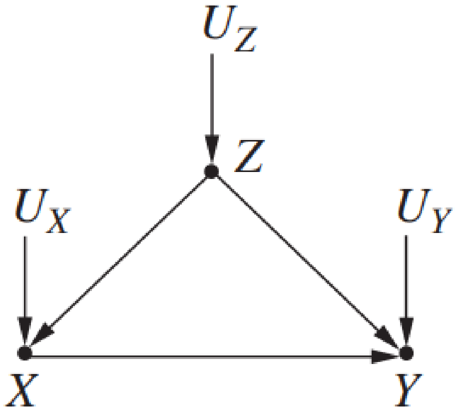
Endogenous and Exogenous Variable

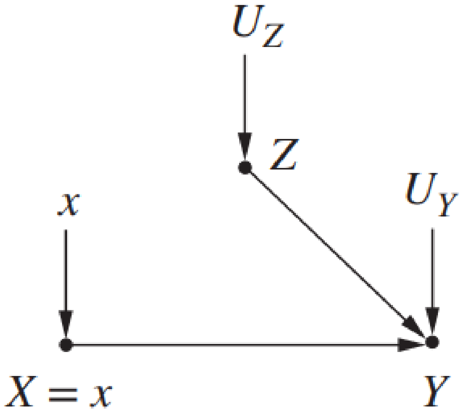
Independence and Conditional Independence

d-connected and d-separated

Intervene

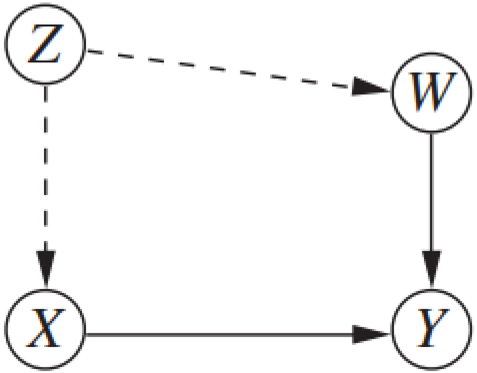
Y=y | do(X=x)





Causual Effect = E(Y | do(X = x+1)) - E(Y | do(X = x))

Back-Door Criterion



$$P(Y = y \mid do(X = x)) = \sum_z P(Y = y \mid X = x, Z = z)P(Z = z)$$

Conditional on Z

1.We block all spurious paths between X and Y.

2. We leave all directed paths from X to Y unperturbed.

3. We create no new spurious paths

Direct Effect in Linear Situation

Assumption: the relationships between variables are linear, and that all error terms have Gaussian distributions. X and all Z are the causal factors

$Y = aX + \beta_1 Z_1 + \beta_2 Z_2 + \ldots + \varepsilon \quad \varepsilon \sim G(0,1)$

under backdoor criteria

Causual Effect = a

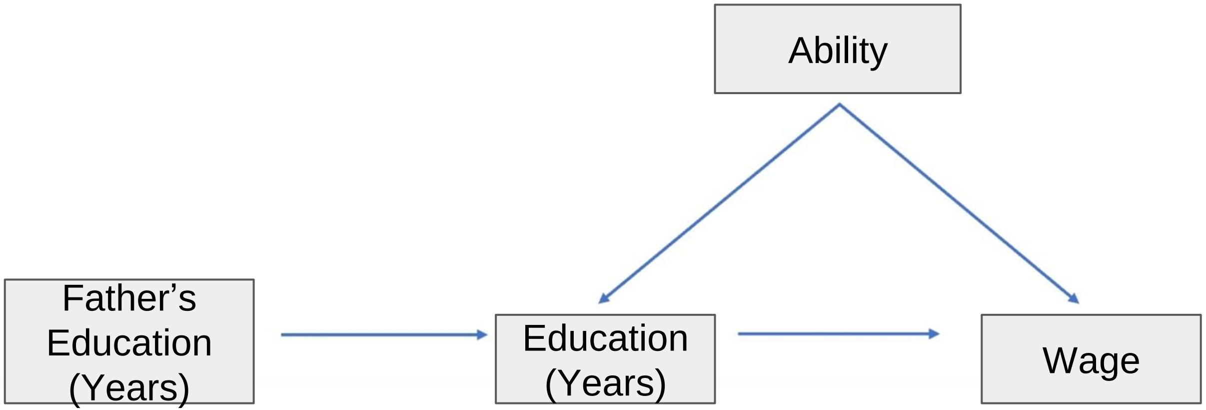
Instrumental Variable

Instrument = “Tool”

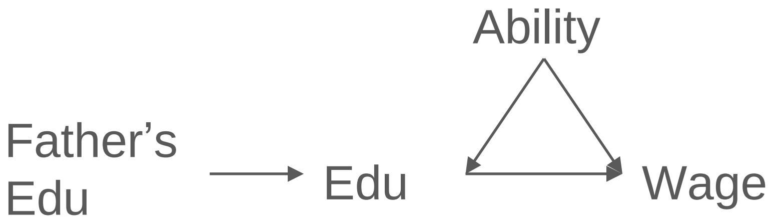
Relevancy Exclusion Exogeneity

Project: Does an extra Year of Eduation causes Increased Wages

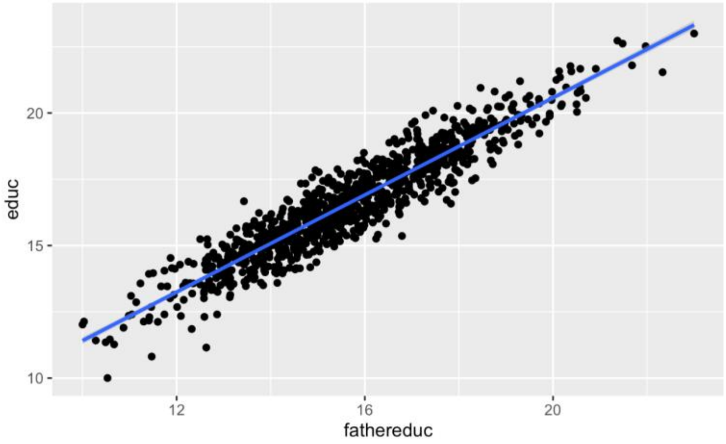
Instrumental Variable Method

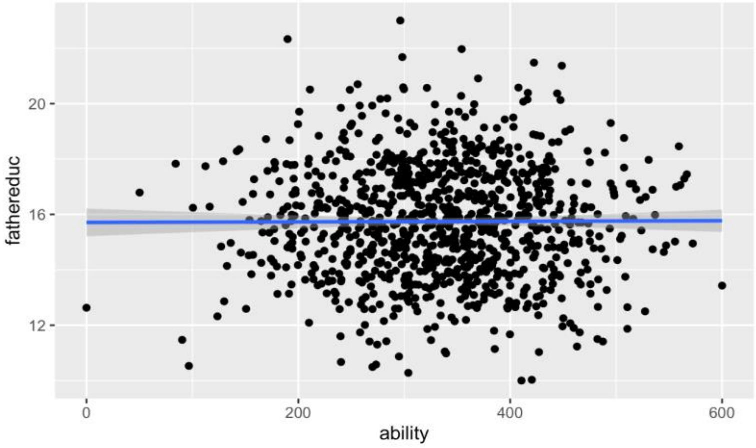


Does an extra Year of Eduation causes Increased Wages?
Part 2: Instrumental Variable Assumptions

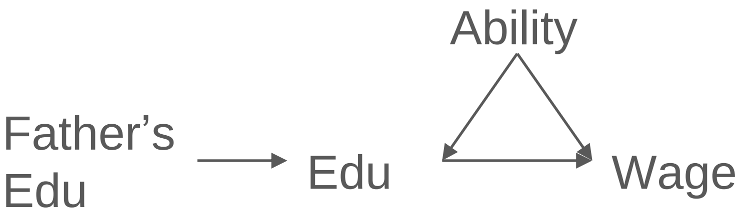


Relevancy, Exclusion, Exogeneity





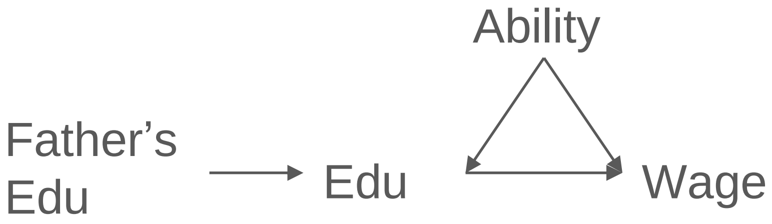
Does an extra Year of Eduation causes Increased Wages?
Part 1: Explore the use of Instrumental Variable



Interested in how **Edu** influences **Wage**
Ability as confounder
Father's Edu as Instrumental Variable

Methodology:
Fake Data -> **Ability** is accessible
Regress **Wage** with **Edu** & **Ability**
Wage = $\boldsymbol{\alpha}$Edu + $\boldsymbol{\beta}$Ability + ε -> reveal true $\boldsymbol{\alpha}$
Regress **Edu** with **Father's Edu**
Edu = a(Father's Edu) + ε
Obtain **Edu_hat**
Regress **Wage** with **Edu_hat**
Wage = $\boldsymbol{\alpha}$Edu_hat + ε -> Theoretically same $\boldsymbol{\alpha}$

Does an extra Year of Eduation causes Increased Wages?
Part 3: Computations



Regress **Wage** with **Edu** & **Ability**

Wage = $\boldsymbol{\alpha}$Edu + $\boldsymbol{\beta}$Ability + ε $\boldsymbol{\alpha}$ = 7.767

Regress **Edu** with **Father's Edu**

Edu = a(Father's Edu) + ε

Regress **Wage** with **Edu_hat**

Wage = $\boldsymbol{\alpha}$Edu_hat + ε $\boldsymbol{\alpha}$ = 7.835

Naive Model:

Regress **Wage** with **Edu** $\boldsymbol{\alpha}$' = 13.1

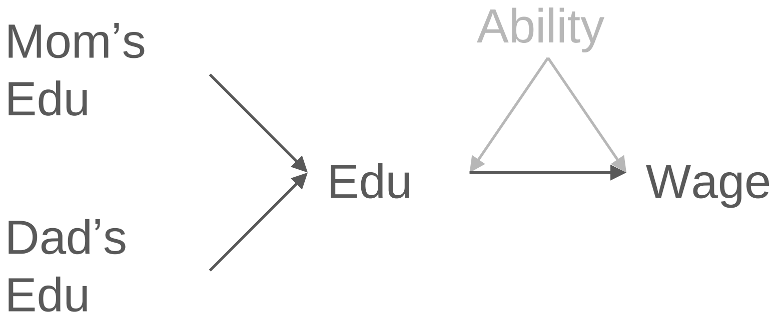
Instrumental Variable Model:

Regress **Edu** with **Father's Edu**

Regress **Wage** with **Edu_hat** $\boldsymbol{\alpha}$ = 7.835

Ready to explore real data without the need of Ability data!

Does an extra Year of Eduation causes Increased Wages?
Part 4: Real Data Analysis



Have both parents’ Education as IVs to better predict children’s Education
No longer have access to confounding variable Ability, but it still exists.
Regress **Edu** with **Mom’s Edu** and **Dad’s Edu**
Obtain **Edu_hat**
Regress **Wage** with **Edu_hat**
Coefficient is the desired α !

A year of education ***causes*** a 15.7% increase in annual wages, on average.

Data retrieved from paper https.//doi.org/10.3386/w3857

Questions?

Structural Causal Models & Instrumental Variable

Group 22:
Mentor: Shiyu He
Mentee: Anne Zheng & Yifan Cao