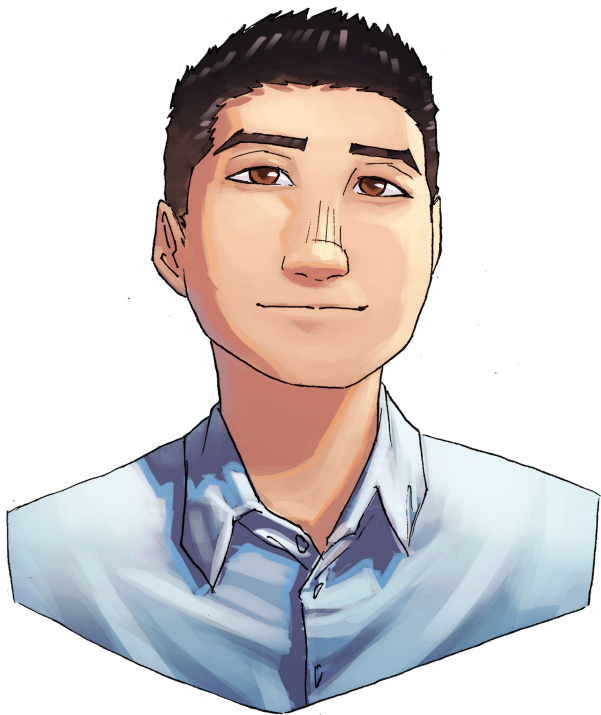CAUSAL BEAST

D.H. Kim

HELLO!

I am D.H. Kim

- I love causal inference
- I am a data scientist
- You can find me at @dhexonian

CAUSAL BEAST

Introduction

Purpose of this course

Universal Dislike for Uncertainty

▷ We're neurologically wired to perceive anything uncertain as a potential threat

▷ Our brains dislike uncertainty so much that they create certainty in **four key ways**

CAUSAL BEAST

(1) Attending to the Familiar

▷ Choosing among the most popular

▷ Doing nothing or choosing the default

CAUSAL BEAST
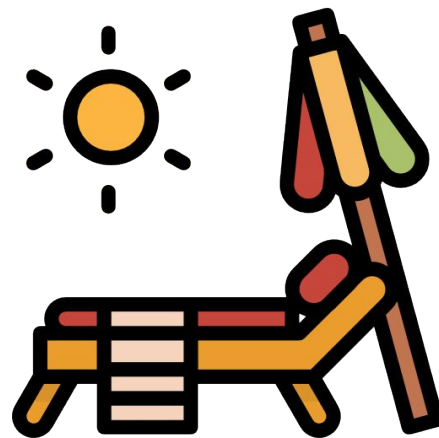
(2) Making Stuff Up

▷ Apophenia
  ▶ "Unmotivated seeing of connections [accompanied by] a specific feeling of abnormal meaningfulness"

▷ E.g. seeing intricate things in clouds like dogs

CAUSAL BEAST

(3) Roll the Dice

▷ Any decision is better than no decision at all because there is certainty in action

▷ Relief comes even when the decision is suboptimal or chosen at random

CAUSAL BEAST

(4) Using Data

▷ We always feel it in our hearts that we can do better than #1, #2 and #3—we prefer using data

▷ Data = Information

▷ But what kind of information? What is the kind of evidence called?

CAUSAL BEAST

Causal Information

The most preferred type of information to reduce uncertainties

CAUSAL BEAST

Rules of Thumb to Guide Desirable Effects

Health

More healthy is better than less healthy

Income

Higher income is better than less income

Productivity

More productive is better than less productive

Employment

More employment is better than less employment

CAUSAL BEAST

Essence of Social Science Research

▷ Answering causal questions
  ▶ Does public healthcare improve health outcomes?
  ▶ Does a sugary drink tax reduce obesity?
  ▶ Are addiction clinics effective in accelerating recovery?

Causal Knowledge in Our Personal Lives

▷ Personal uncertainties are also centered around issues of causality
  ▶ Should I attend college?
    ◆ Does college make people more financially well off?
  ▶ Should my partner move in with me?
    ◆ Does cohabitation decrease chances of divorce?

CAUSAL BEAST

Causality in Everyday Language

- ▷ Why?
- ▷ Cause
- ▷ Make
- ▷ Create
- ▷ Effect
- ▷ Force
- ▷ Lead to

CAUSAL BEAST

Potentially Misinforming News Sources

Trump linked video games and gun violence - CNBC.com
https://www.cnbc.com/.../trump-unlikely-to-change-policy-on-violent-video-games.ht... ▼
Mar 9, 2018 - Trump **linked video games** and gun violence – but don't expect him or Congress to do anything about it. ... Following the high school massacre last month in Parkland, Florida, Trump has suggested that a **link** exists between violent **video games** and violent behavior. ... If President Donald

Soda boosts violence among teens, study finds - The Boston Globe
https://www.bostonglobe.com/news/nation/2011/.../soda-boosts-violence.../story.html
Washington Post October 26, 2011. WASHINGTON - Teenagers who drink **soda** are more likely carry a weapon and act violently, according to new research.

Will Essential Oils Like Lavender And Tea Tree Make Your Breasts ...
https://www.forbes.com/.../will-essential-oils-like-lavender-and-tea-tree-make-your-br... ▼
Mar 18, 2018 - Can **essential oils** like lavender and tea tree oil give you breasts? ... **evidence** that lavender and tea tree oil may give you **more** than just good ...

Marijuana linked to 'unbearable' sickness across US as use grows ...
https://www.independent.co.uk › News › Health ▼
Apr 9, 2018 - By the time Thomas Hodorowski made the connection between his **marijuana** habit and the bouts of pain and vomiting that left him ...

CAUSAL BEAST

What is Correlation?

Correlation vs. Causation

CAUSAL BEAST

Defining Correlation

- ▷ Measure of linear association between two variables
- ▷ The value of correlation coefficient, r denotes both direction and association of the association
- ▷ 0 for no relationship
- ▷ +1 for strong positive relationship
- ▷ -1 for strong negative relationship

Strong Positive

Strong Negative

Weak Positive

Weak Negative

None

None

CAUSAL BEAST
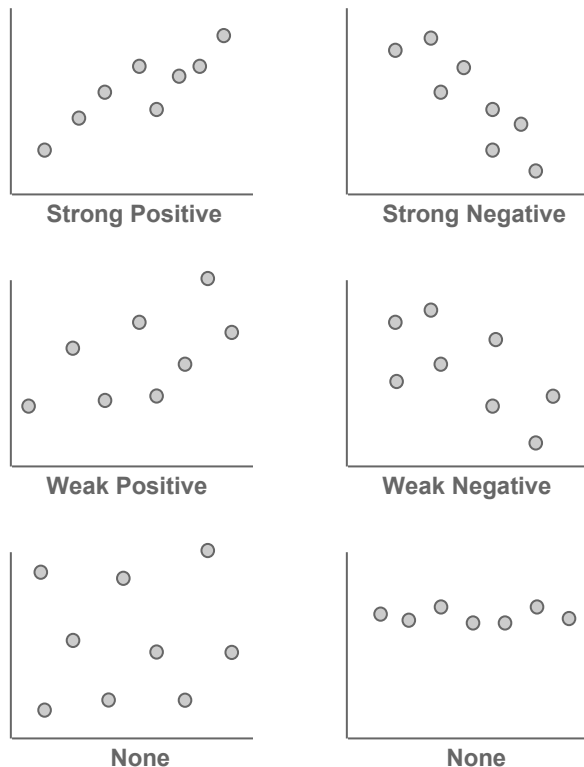
Correlation ≠ Causation

▷ Correlation simply denotes an association, does not denote a causal relationship

▷ Correlation between ice cream sales and sunglasses sold

▷ What is the confounder?

Sunglass vs. Ice Cream Sales
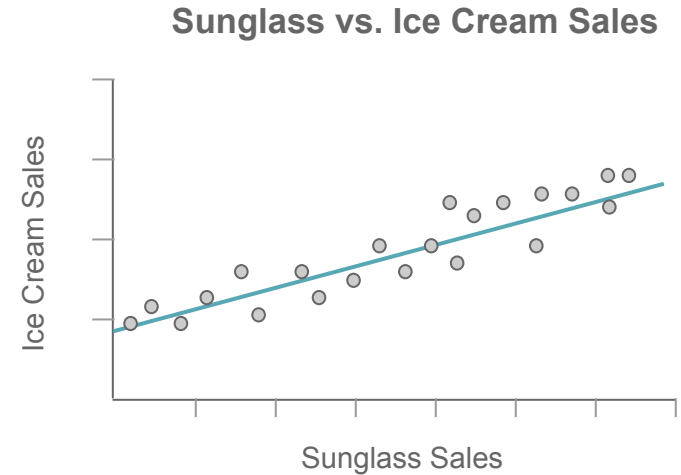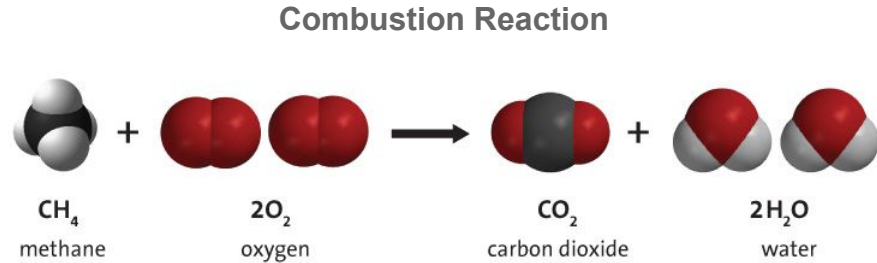
Chocolate Consumption Produces Nobel Laureates?

▷ Nobel Laureate production correlated with chocolate consumption

▷ Is this causal?

▷ What is the confounder?

Chocolate Consumption and Nobel Laureates
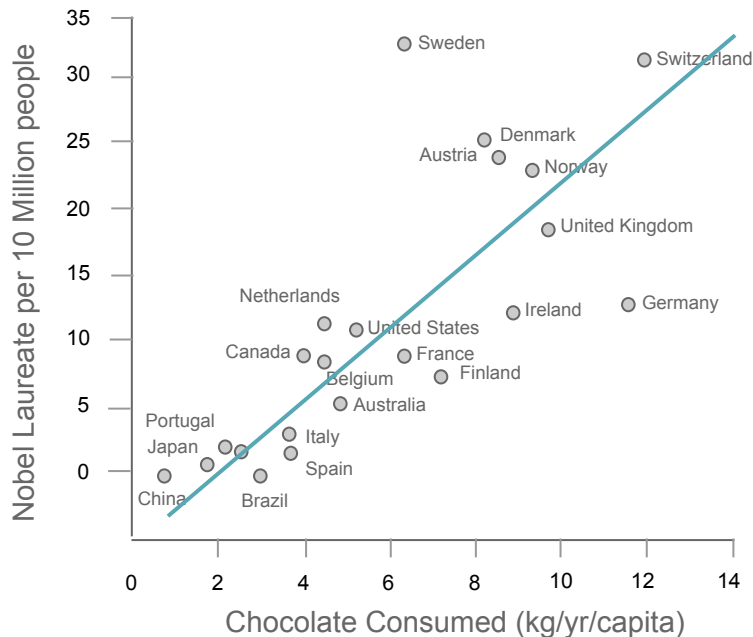
Some Correlations are Causal...

▷ All causal relationships have correlation

▷ Not all correlations are causal
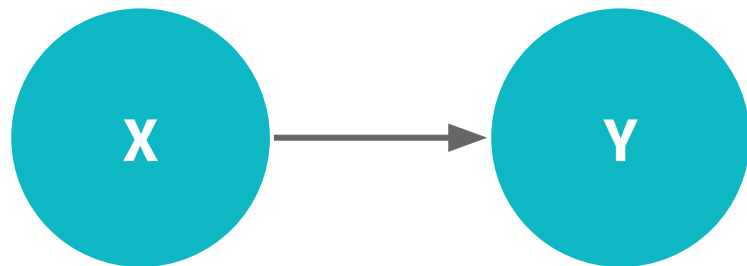
▷ Burning of fossil fuel causes emission of carbon dioxide

Combustion Reaction

CH_4 methane $+$ $2O_2$ oxygen $\rightarrow$ CO_2 carbon dioxide $+$ $2H_2O$ water

CAUSAL BEAST

Types of Correlation

Correlation comes in all shapes and sizes

CAUSAL BEAST

(1) X ⟹ Y
(Direct Causation)

▷ Education (X)

▷ Wages (Y)

(2) X ⇐ Y
(Reverse Causation)

▷ Debt (X)

▷ Economic Growth (Y)
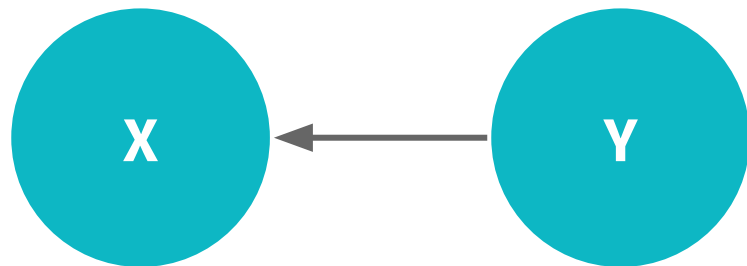
(3) X ⇔ Y
(Cyclic Causation)

▷ Motivation (X)

▷ Learning (Y)
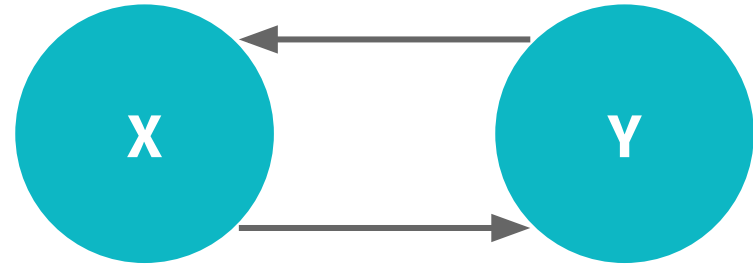
(4) X ⇒ Y ⇒ Z (Chain)
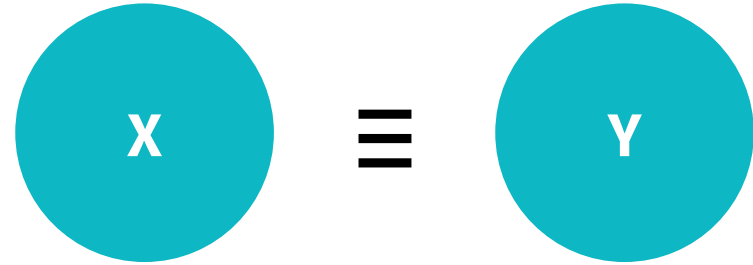
▷ Burglar (X)

▷ Alarm (Y)

▷ Sally (Z)

▷ X ⊥ Z | Y
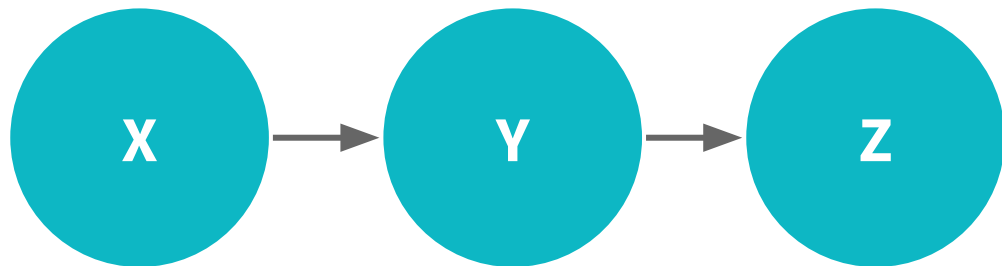
(5) $X \Leftarrow Z \Rightarrow Y$
(Fork/Confounding)

▷ Aircraft tire sales (X)

▷ Rail part sales (Y)

▷ Tariffs (Z)

▷ $X \perp Y | Z$
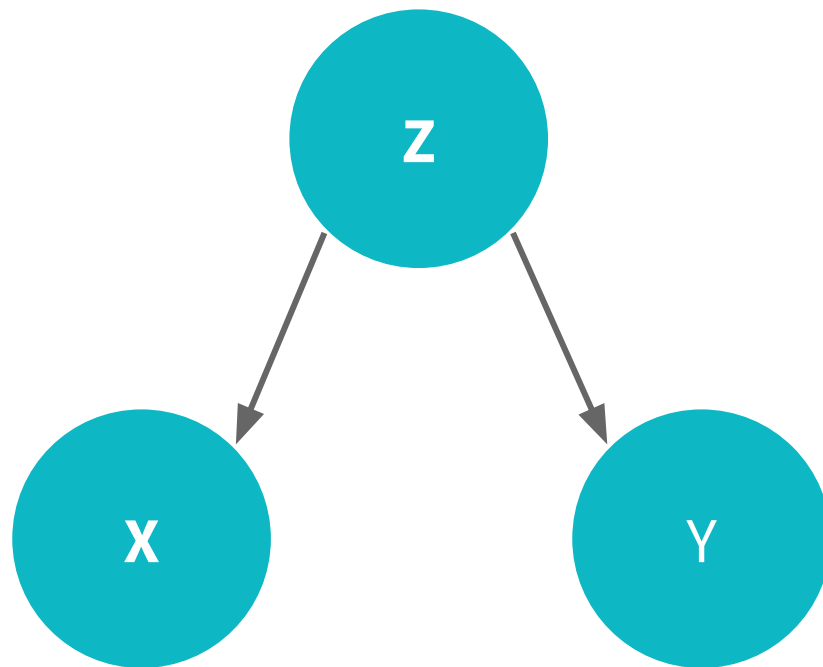
(6) X ≡ Y
(Tautology)

▷ Fahrenheit (X)

▷ Celsius (Y)

▷ X ≡ Y

$$X \equiv Y$$
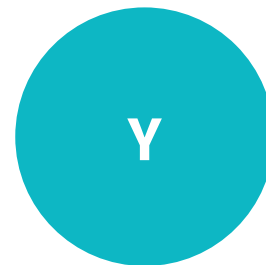
(7) X ⟹ Z ⟸ Y
(Collider)

▷ Battery (X)

▷ Gas tank (Y)

▷ Car running? (Z)

▷ X ⊥/Y | Z
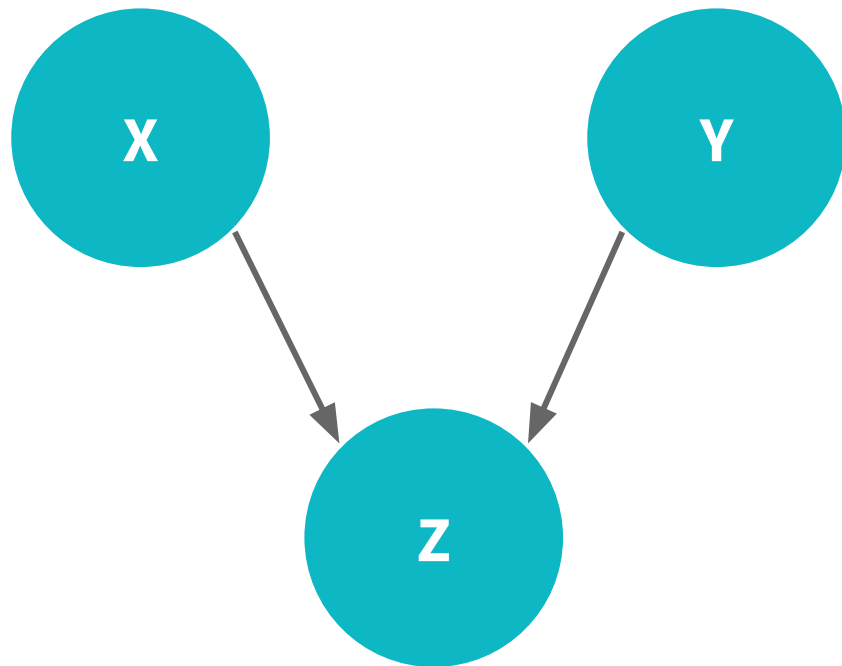
(8) Coincidence

▷ Radish sales (X)

▷ # of forest fires in California (Y)

X

Y
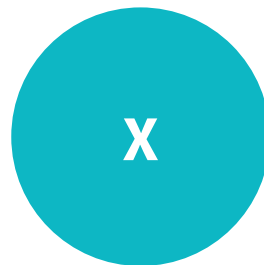
Coincidental Association
Example 1

Coincidental Association
Example 2

Letters in Winning Word of Scripps National Spelling Bee
correlates with
Number of people killed by venomous spiders

Correlation: 80.57% (r=0.8057)

Data sources: National Spelling Bee and Centers for Disease Control & Prevention

tylervigen.com
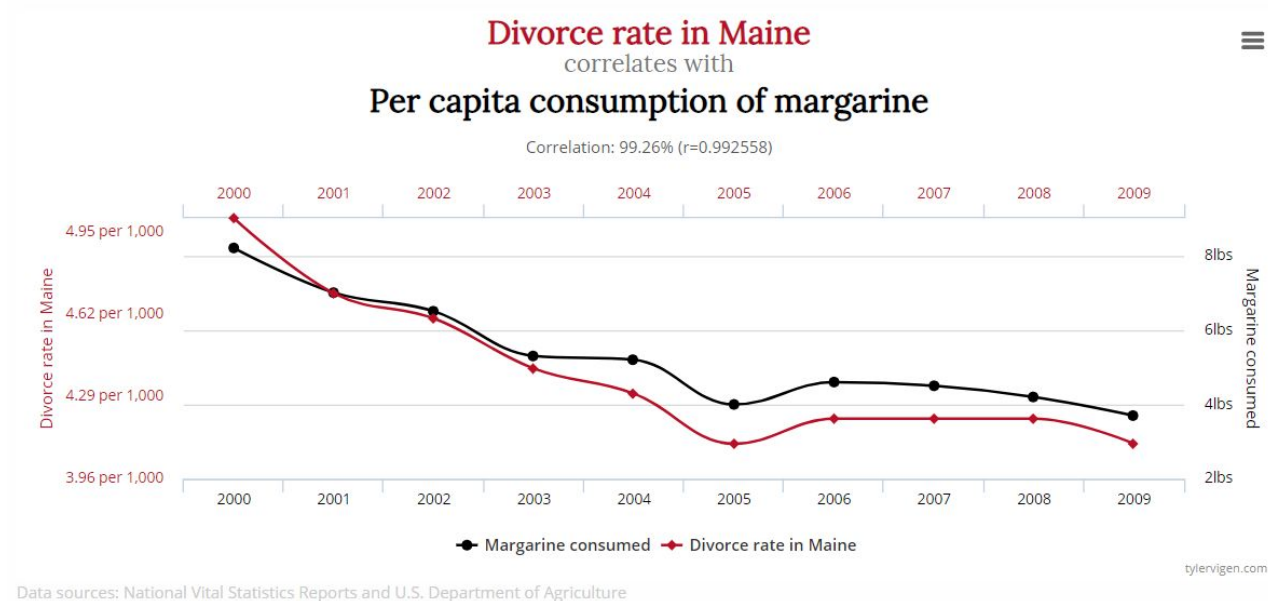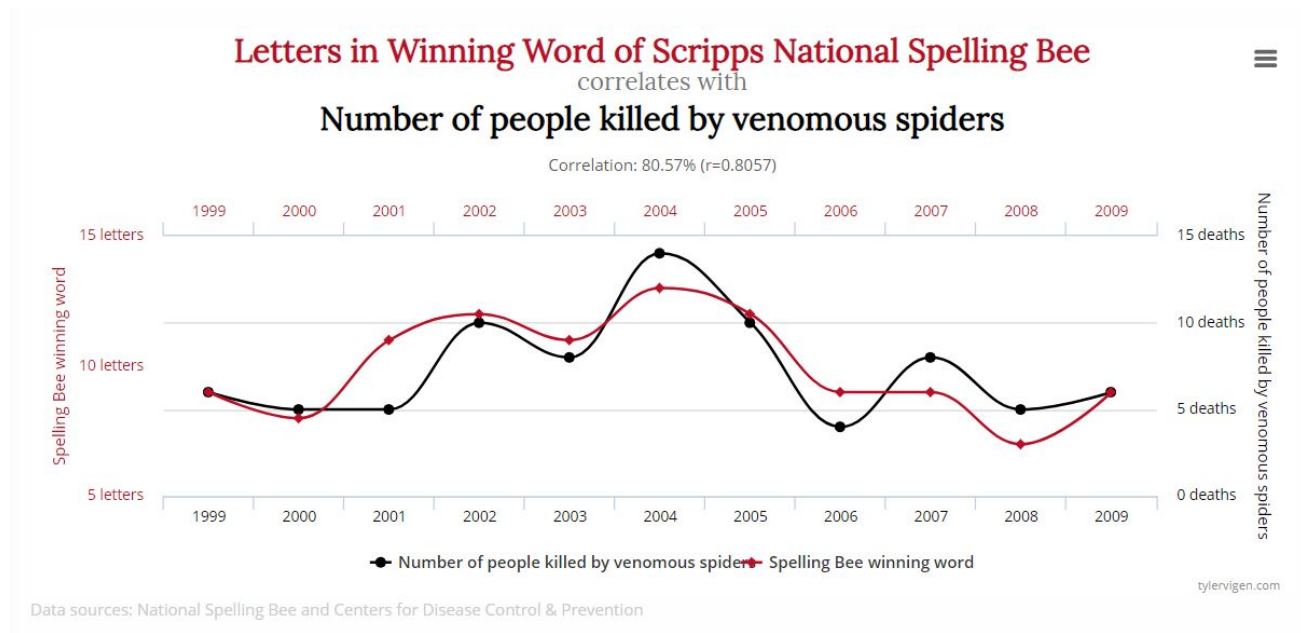
Causality Checklist

Making sure all the boxes are ticked off

CAUSAL BEAST

Causality Checklist

▷ Time order

▷ Co-variation

▷ Significance

▷ Rationale

▷ Non-spuriousness
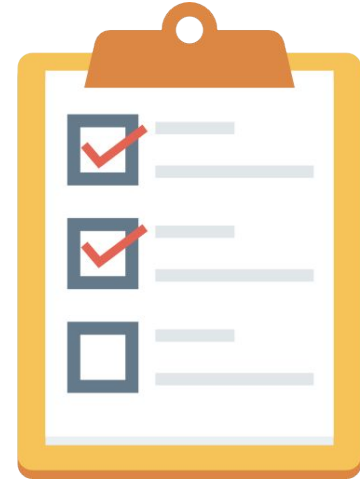
(1) Time Order

▷ Cause must have occurred before the effect

▷ For X to cause Y, it must be the case that X occurs before Y

(2) Covariation

▷ Statistical association (need to have at least some correlation)

▷ Remember, covariation is a necessary but insufficient condition for causality
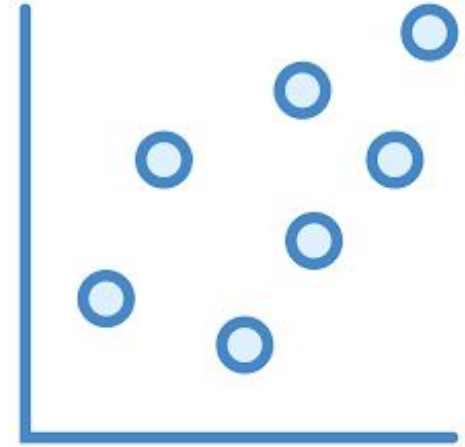
(3) Significance

▷ Are the effects observed due to random chance?

▷ Imprecise estimates

(4) Rationale

▷ Logical, compelling explanation for causal relationship

▷ "Just because" does not suffice

(5) Non-spuriousness

▷ Must be established that X, and only X, was the cause of changes in Y

▷ Rival explanations must be ruled out

Ceteris Paribus

The foundation of causal inference

CAUSAL BEAST

Bivariate Case

▷ Most causal theories are bivariate; they ask whether X causes Y

▷ But crucially, even though our theory might be bivariate, the real world is is multivariate—every interesting phenomenon has several causes

CAUSAL BEAST
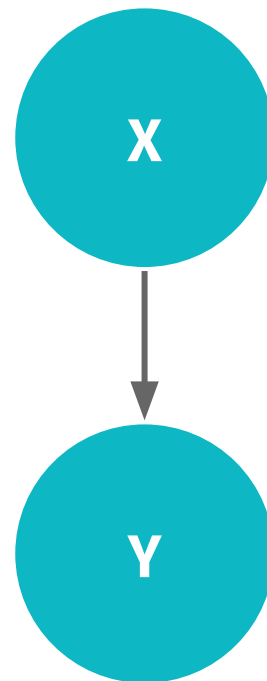
Multivariate World

▷ Theories are bivariate but the world is multivariate

▷ Need to control for all other factors W

▷ We call controlling for all other factors ceteris paribus
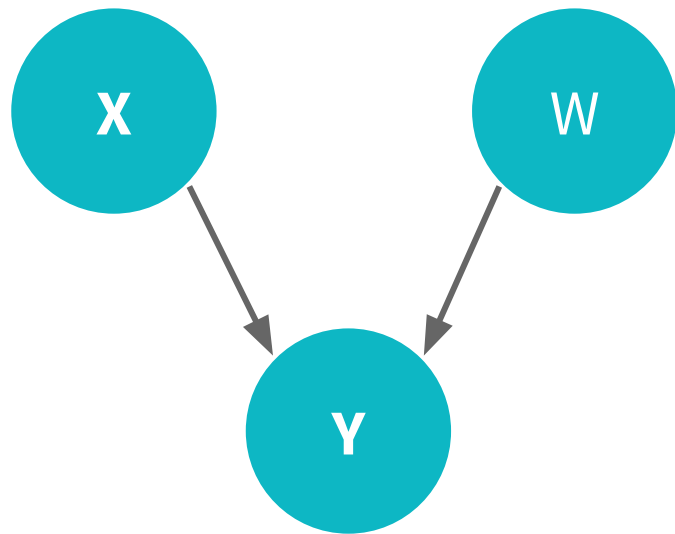
Controlled (Scientific) Experiment

▷ Establish ceteris paribus with tight control of all relevant variables—means that treatment and control group are near replicas

▷ Because they are otherwise equal in every way, difference in outcomes will be attributed to the fact that only one group has received the treatment

X

Treatment Group	Control Group
treatment	control

Y

Treatment Outcome	Control Outcome

CAUSAL BEAST
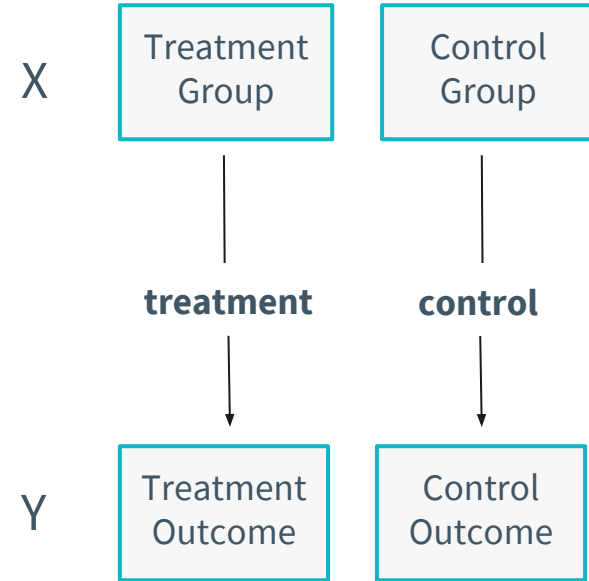
The 8th Grade Science Fair

▷ Does sunlight make plants grow plants?

▷ Clone "new" plants by clipping the plant and taking the cuttings to grow it elsewhere

▷ Then put one clone under darkness and another under artificial light

CAUSAL BEAST

Does it Check the Boxes?

▷ Time order

▷ Co-variation

▷ Significance

▷ Rationale

▷ Non-spuriousness

CAUSAL BEAST
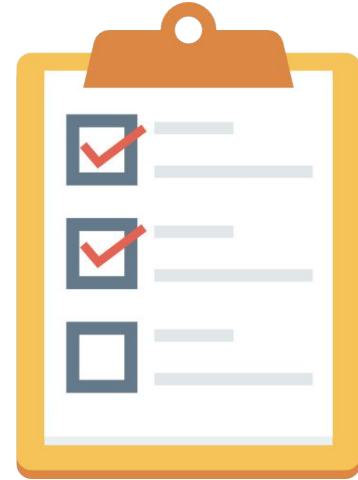
Impracticability of Controlled Experiments in Social Science

▷ Controlled experiments require sameness of subjects (e.g. identical plants, compounds, climates, etc.)

▷ Even if possible, experimentation on people is morally questionable

CAUSAL BEAST

Measurement

What are we trying to measure?

CAUSAL BEAST

Two roads diverged in a yellow wood, and sorry I could not travel both and be one traveler, long I stood and looked down one as far as I could to where it bent in the undergrowth;

…

Two roads diverged in a wood, and I— I took the one less traveled by, and that has made all the difference.

Road Not Taken (Robert Frost)

CAUSAL BEAST

Unit Level Effect of Treatment

▷ $Y_{i,j}$ is the outcome of the jth subject given treatment i where if i=0 the subject is not treated and if i=1 the subject is treated.

▷ Then the unit level effects for the first subject:

$$Y_{1,1} - Y_{0,1}$$

Observing all the Unit Level Treatment Effects

▷ $Y_{i,j}$ is the outcome of the jth subject given treatment i where if i=0 the subject is not treated and if i=1 the subject is treated.

▷ Then the unit level effects are as follows for each subject:

$$Y_{1,1} - Y_{0,1}$$

$$Y_{1,2} - Y_{0,2}$$

$$...$$

$$Y_{1,n} - Y_{0,n}$$

Average Treatment Effect

▷ $Y_{i,j}$ is the outcome of the jth subject given treatment i where if i=0 the subject is not treated and if i=1 the subject is treated.

▷   Then the unit level effects are as follows for each subject:

$$Y_{1,1} - Y_{0,1}$$
$$Y_{1,2} - Y_{0,2}$$
$$\ldots$$
$$Y_{1,n} - Y_{0,n}$$

$$\frac{1}{n}\sum_{j=1}^{n}(Y_{1,j} - Y_{0,j}) = \frac{1}{n}\sum_{i=1}^{n}(Y_{1,j}) - \sum_{i=1}^{n}(Y_{0,j})$$

$$= avg_n(Y_{1,j} - Y_{0,j})$$

$$= k$$

CAUSAL BEAST

Example of Calculating Average Treatment Effect

Subject	w/ Treatment	w/o Treatment	Unit Level Treatment Effect
1	5	4	1
2	4	2	2
3	3	4	-1
4	5	3	2

- Total Treatment Effect = 1+2-1+2 = 4

- Average Treatment Effect = 4/4 = 1

CAUSAL BEAST

Difference Between Treatment and Control Group Means

Subject	w/ Treatment	w/o Treatment	Unit Level Treatment Effect
1	5	?	?
2	?	2	?
3	3	?	?
4	?	3	?

- Average of w/treatment = 4

- Average of w/o treatment = 2.5

- Diff = 1.5

CAUSAL BEAST

Difference Between Group Means is Biased

Assume constant level treatment effect $Y_{1,j} = Y_{0,j} + k$

$$\frac{1}{n}\sum_{j=1}^{n}(Y_{1,j} - Y_{0,j}) = \frac{1}{n}\sum_{j=1}^{n}(Y_{1,j}) - \sum_{j=1}^{n}(Y_{0,j})$$

$$= avg_n(Y_{1,j}|D_j = 1) - avg_n(Y_{0,j}|D_j = 0)$$

$$= k + avg_n(Y_{0,j}|D_j = 1) - avg_n(Y_{0,j}|D_j = 0)$$

Average Y0 of insured group (unobservable)

Average Y0 of uninsured group (observable)

	$D_j = 1$	$D_j = 0$
Treatment Outcome $Y_{0,j}$	$avg_n(Y_{1,j}|D_j = 1)$	$avg_n(Y_{0,j}|D_j = 0)$
Control Outcome $Y_{1,j}$	$avg_n(Y_{0,j}|D_j = 1)$	$avg_n(Y_{0,j}|D_j = 1)$

CAUSAL BEAST

Insurance Example is Not Ceteris Paribus

| Insured	Uninsured
- **Has insurance**	- **Doesn't have insurance**
- Majority above poverty line	- Majority below poverty line
- Majority White	- Majority Hispanic or Black
- Majority employed	- Majority unemployed
- Majority university educated	- Majority high school educated or less

CAUSAL BEAST

Randomized Experiment

The best way to remove selection bias

Toolbox

**Randomized
Experiment**
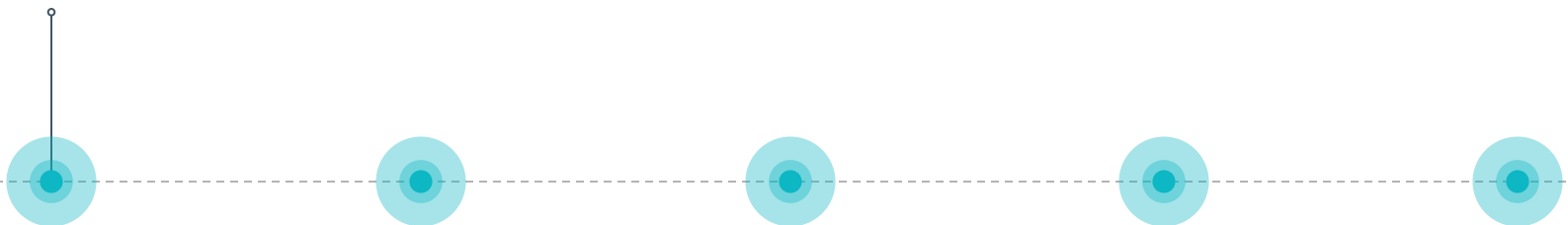
Review from Last Module

$$\frac{1}{n}\sum_{j=1}^{n}(Y_{1,j} - Y_{0,j}) = \frac{1}{n}\sum_{j=1}^{n}(Y_{1,j}) - \sum_{j=1}^{n}(Y_{0,j})$$

$$= avg_n(Y_{1,j}|D_j = 1) - avg_n(Y_{0,j}|D_j = 0)$$

$$= k + avg_n(Y_{0,j}|D_j = 1) - avg_n(Y_{0,j}|D_j = 0)$$

$$= k + selection\ bias$$

$$= selection\ bias\ if\ k = 0$$

$$avg_n(Y_{1,j}|D_j = 1) - avg_n(Y_{0,j}|D_j = 1) = 0$$

$$avg_n(Y_{1,j}|D_j = 1) = avg_n(Y_{0,j}|D_j = 1)$$

	$D_j = 1$	$D_j = 0$
Treatment Outcome $Y_{0,j}$	$avg_n(Y_{1,j}|D_j = 1)$	$avg_n(Y_{0,j}|D_j = 0)$
Control Outcome $Y_{1,j}$	$avg_n(Y_{0,j}|D_j = 1)$	$avg_n(Y_{0,j}|D_j = 1)$

CAUSAL BEAST

Randomized Experiment Design

▷ Treatment group and control group created by random assignment differ only in treatment assignment and any consequences that follow

▷ Pre-treatment outcomes are equivalent on average
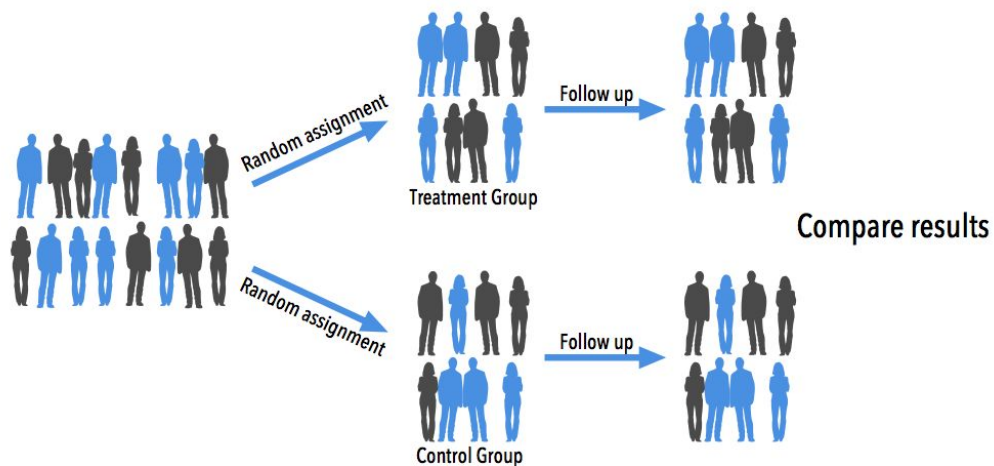
Perry Preschool Project Experiment

▷ 123 children in Ypsilanti, Michigan

▷ Half of them (drawn at random) entered the Perry Preschool Program at 3 or 4 years old

▷ Education for 30 weeks by skilled professionals in nurseries and kindergarten
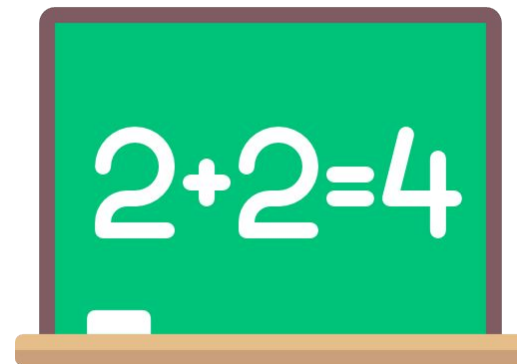
Nores, Milagros et al. (2005): Updating the Economic Impacts of the High/Scope Perry Preschool Program

	Size (n=123)	Treatment Group	Control Group	p-value
Cognitive skills at 15	95	122.2	94.5	< 0.001
University enrollment	121	38%	21%	0.029
Jailed or arrested at least once	121	31%	51%	0.022
Welfare	120	18%	32%	0.044
Employed at 19 years old	121	50%	32%	0.032

CAUSAL BEAST

Because children were randomly assigned to the program or a control group, differences in outcomes are probably attributable to program status… The data shows strong advantages for the treatment group in terms of higher lifetime earnings and lower criminal activity.

"

Nores et al. (2005)

CAUSAL BEAST

Snow, John (1855): On the Mode of Communication of Cholera

▷ Cholera outbreak in London

▷ English physician John Snow identifies a natural experiment
  ▶ Londoners are subscribed to water companies in an as-if random way

CAUSAL BEAST

Snow's Findings

Water Supplier	# Houses	Cholera Deaths	# Cholera Deaths / 10,000 Houses
Southwark & Vauxhall	40,046	1,263	315
Lambeth	26,107	98	37
Others	256,423	1,422	59

CAUSAL BEAST

When a Random Experiment is Not Possible

▷ Find situations that approximate randomized trials

▷ Identification strategy: as if natural experiments

CAUSAL BEAST

Potential Pitfalls of Randomized Experiments

Hawthorne/John Henry Effect

Control or treatment react to being part of an experiment

Ethics

The rationale for conducting a randomized experiment is based on the fact that researchers are genuinely uncertain over the benefits of the treatment

Spillover Effect

Control group is affected by existence of a treatment group nearby

Heterogeneous and Implementer Effect

Same program may have different effects across contexts and target populations

Same program may have different effects depending on the implementer

Compliance Issue

Not all members of the treatment group end up receiving the treatment and benefiting

CAUSAL BEAST

Further Studies on Randomized Experiment

Miracle of Microfinance? Evidence from a Randomized Evaluation

Banerjee et al. (2013)

The Oregon Health Insurance Experiment: Evidence from the First Year

Finkelstein et al. (2011)

The Experimental Approach to Development Economics

Banerjee et Duflo (2008)

Health Insurance and the Demand for Medical Care: Evidence from a Randomized Experiment

Manning et al. (1987)

Regression

Trying to stay in control

CAUSAL BEAST

Toolbox

Randomized Experiment
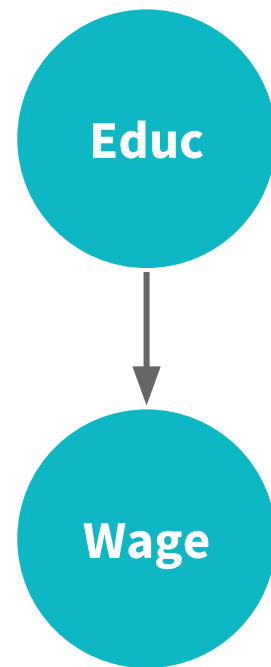
Regression (Controlled)

CAUSAL BEAST
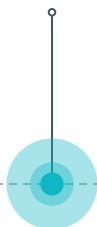
Returns to Education

▷ Regression starts by specifying a model

▷ Start with base model: years of education determining wage

Educ

Wage

CAUSAL BEAST

Gender as Key Omitted Variable

▷ Jeannine
  ▶ 8 years of schooling
  ▶ $40,000 salary

▷ Craig
  ▶ 12 years of schooling
  ▶ $68,000 salary

▷ Apples to apples?

CAUSAL BEAST

Comparing Within Same Gender

▷ Jeannine (Female)
  ▶ 8 years of schooling
  ▶ $40,000 salary

▷ Hannah (Female)
  ▶ 12 years of schooling
  ▶ $62,000 salary

▷ Now apples to apples?

CAUSAL BEAST

Regression of Salary on Years of Education for Females

▷ Coefficient of 2,758

▷ For every additional year of education, a woman earns $2,758 more in wages all else constant

Salary vs. Years of Education

● $ Salary
▪ 2758*x + 20169

Y-axis: $ in Thousands (0, 20, 40, 60, 80)

X-axis: Years of Education Since Middle School (0 1 2 3 4 5 6 7 8 9 10 11 12 13)

CAUSAL BEAST

Separating on Gender for Within Comparisons

Years of Education	Female Group	Male Group
0	24,000	22,000
1	24,111	22,008
2	24,535	22,658
3	25,683	24,483
4	26,983	25,748
...	...	...
13	67,466	57,500
education on wage	Coeff 1	Coeff 2

CAUSAL BEAST

Regression Conceptually as a Weighted Average Effect

Years of Education	Female Group	Male Group
0	24,000	22,000
1	24,111	22,008
2	24,535	22,658
3	25,683	24,483
4	26,983	25,748
…	…	…
13	67,466	57,500
education on wage	Coeff 1	Coeff2

Causal effect of education on wages

CAUSAL BEAST

Accounting for the Unaccounted

▷ Regression model is another representation of graphical model

▷ Many unaccounted factors lumped into u
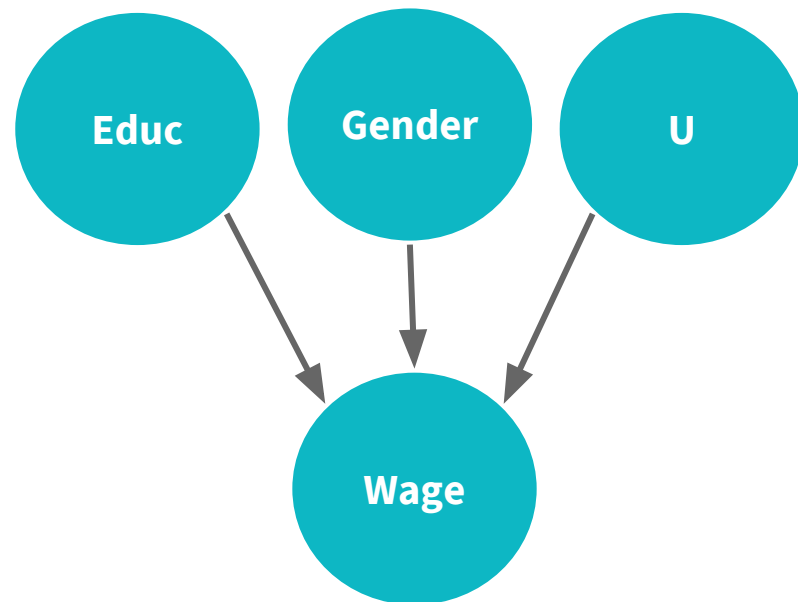
$$wage = B_0 + B_1 * educ + B_2 * gender + u$$

Wage

Effect of education on wages

Effect of gender

Effect of unaccounted factors

Progression of Controlled Regression

$$wage = B_0 + B_1 * educ + u$$

$$\downarrow$$

$$wage = B_0 + B_1 * educ + B_2 * gender + u$$
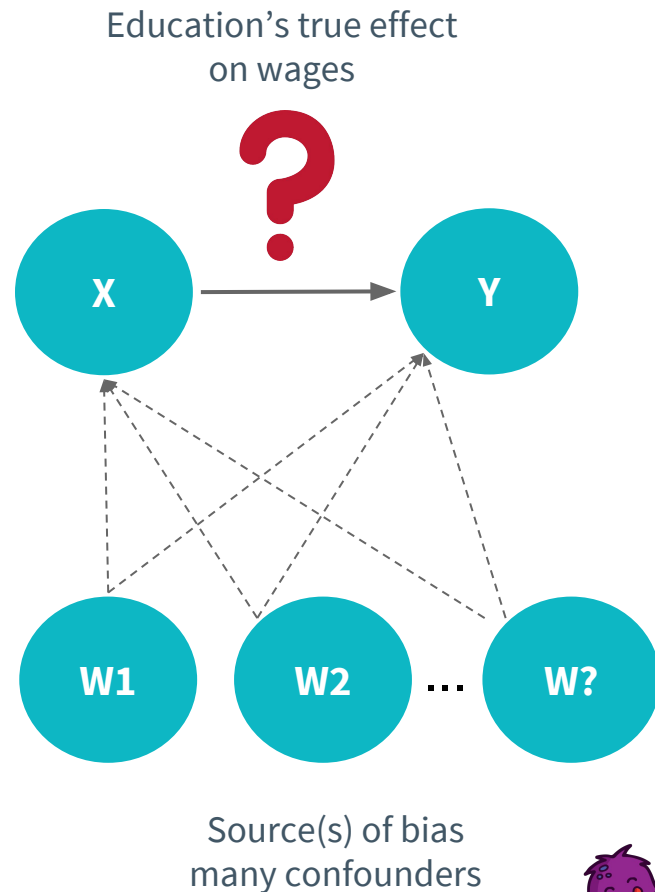
$$\downarrow$$

$$wage = B_0 + B_1 * educ + B_2 * gender + B_3 * score + B_4 * pincome + u$$

CAUSAL BEAST

# How Much to Control for?

▷ Impossible to control for everything

▷ Need to do the "best we can"

▷ Admittedly, fairly subjective

Education's true effect on wages

Source(s) of bias many confounders

CAUSAL BEAST

Black, Sandra (1999): Do Better Schools Matter? Parental Valuation of Elementary Education

▷ Answering the question of whether better schools affect house prices?

▷ Are parents willing to pay more for quality education and if so by how much?

▷ Focus on suburbs of Boston
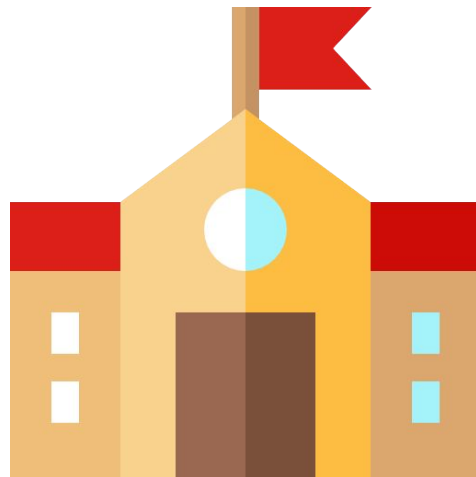
If the estimation strategy does not correct for observable and unobservable neighborhood characteristics, potentially correlated both with housing prices and school quality, then the estimation of the marginal willingness to pay for a better school might suffer from severe biases.

"

Black (1999)

First Attempt at Problem: Standard OLS Regression

Coefficient on school test scores is the main focus

Black first uses a "basic hedonic regression" as follows:

$$(1) \qquad \ln\left(price_{iaj}\right) = \alpha + X'_{iaj}\beta + Z'_j\delta + \gamma test_{aj} + \epsilon_{iaj},$$

where p_{iaj} is the price of house i in attendance district a in school district j. The vector X_{iaj} includes characteristics of house i such as the number of bedrooms and bathrooms in the house, Z_j is a vector of neighborhood and school district characteristics. The regressor of primary interest, $test_{aj}$, is the average test score in the school that children living in attendance district a in school district j would attend.

CAUSAL BEAST

Results from "Basic Hedonic Regression" from Table 2

	(1) Basic hedonic regression[d]
Coefficient on elementary school test score[b]	.035 (.004)
Magnitude of effect (percent change in house price as a result of a 5% change in test scores)[c]	4.9%
$ Value (at mean tax-adjusted house price of $188,000 in $1993)	$9212
$ Value (at median tax-adjusted house price of $158,000 in $1993)	$7742

Causal effect of school quality on home prices

	All houses[d]
Elementary school test score[c]	.035 (.004)
Bedrooms	.033 (.004)
Bathrooms	.147 (.014)
Bathrooms squared	−.013 (.003)
Lot size (1000s)	.003 (.0003)
Internal square footage (1000s)	.207 (.007)
Age of building	−.002 (.0003)
Age squared	.000003 (.000001)
Boundary fixed effects	NO
Census variables	Yes
N	22,679

CAUSAL BEAST

Is this Good Enough?

▷ Controlled regressions require good argumentation on the part of the researchers to show that the unconfoundedness assumption holds

▷ Do we have enough controls to ensure that observable and unobservable neighborhood characteristics are controlled for?

CAUSAL BEAST
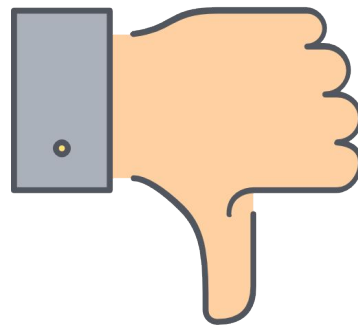
A drawback to these approaches is that all relevant house or neighborhood characteristics cannot be observed; hence results are biased because of omitted variables.

"

Black (1999)

Potential Pitfalls of Controlled Regression

Unavailable Features

Features are measurable in theory but not available or hard to quantify

Need to come up with clever proxy variables

Unconfoundedness Assumption

Patently unprovable assumption

Need to come up good argumentation based on established theory

Included Variable Bias

Function of bad multicollinearity with treatment variable and also carelessness

Model Specification

Extent to which omitted variable W is correlated with X will affect how much multicollinearity will act to increase the standard errors on the regression coefficient for X.

CAUSAL BEAST

Further Studies on Regression

Using Social Security Data on Military Applicants to Estimate the Effect of Voluntary Military Service on Earnings

Angrist (1998)

Estimating the Payoff to Attending a More Selective College: An Application of Selection on Observables and Unobservables

Dale et Krueger (1999)

The Causal Effect of Education on Earnings

Ashenfelter et Card (1999)

Introductory Econometrics: A Modern Approach

Wooldridge (2015)

Mostly Harmless Econometrics

Angrist et Pischke (2009)

CAUSAL BEAST

Instrumental Variables

Finding the source of exogeneity to resolve OVB

CAUSAL BEAST

Toolbox

Randomized Experiment

Instrumental Variables

Regression (Controlled)

CAUSAL BEAST

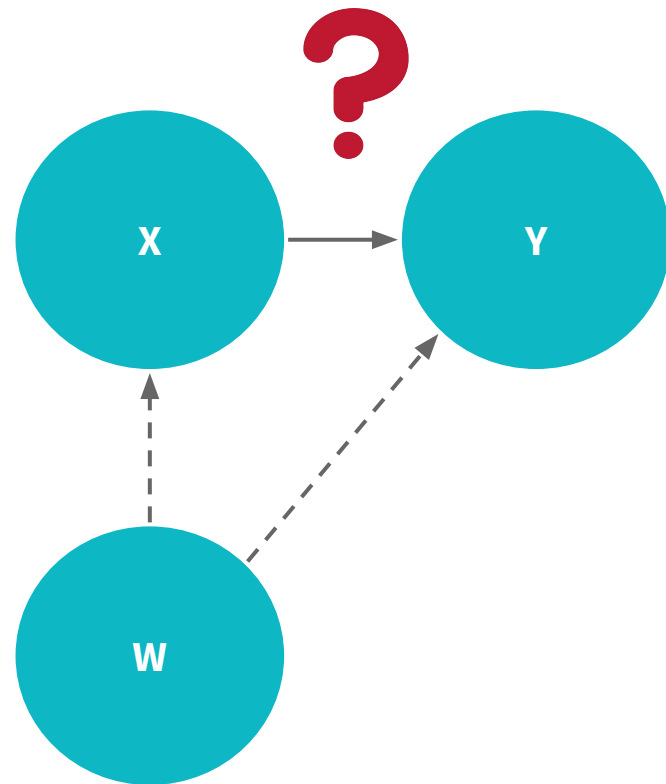# Four Scenarios

▷ If W is related to both X and Y?

▷ If W is unrelated to both X and Y?

▷ If W is related to X but unrelated to Y?

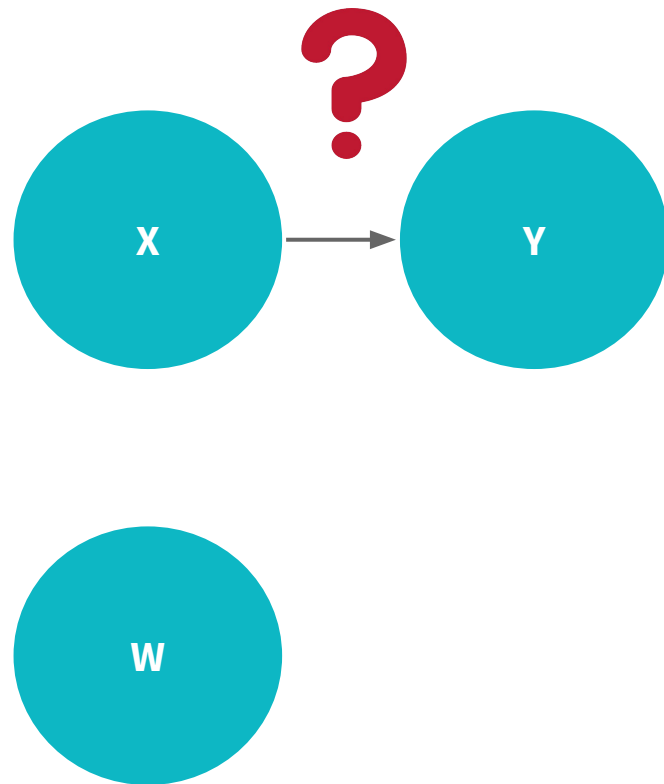▷ If W is unrelated to X but related to Y?

CAUSAL BEAST

W is Related to Both X and Y

▷ W causes Y and W has a correlation with X

▷ The very definition of omitted variable bias

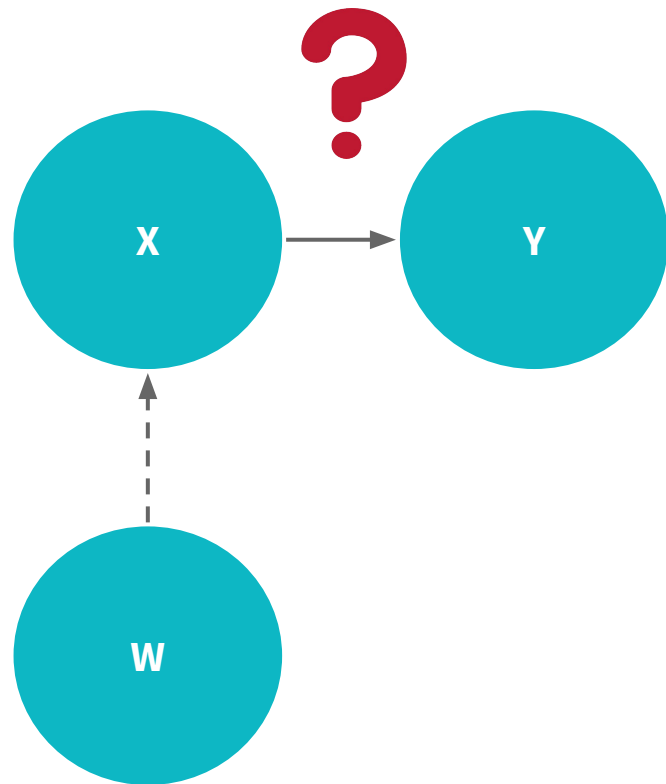# W is Unrelated to Both X and Y

▷ W is independent of the process between X and Y

▷ Keep W omitted (don't include in regression)

W is Related to X but Unrelated to Y

▷ Bad case of multicollinearity

▷ X contains all the information we need to explain Y.

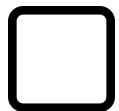▷ W is redundant as it is independent of Y conditioned on X (chain)

W is Unrelated to X but Related to Y

▷ Want to include W if we have data on it for better prediction

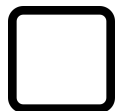▷ Won't reduce any bias of the causal relation between X and Y (note: no OVB)
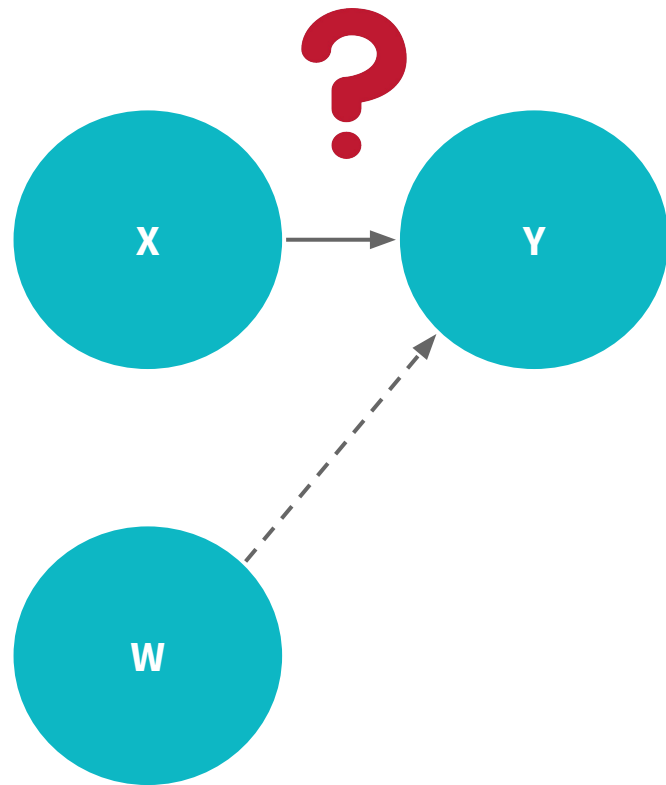
Ideal Scenarios in Regression

☐ If W is related to both X and Y?

☑ If W is unrelated to both X and Y?

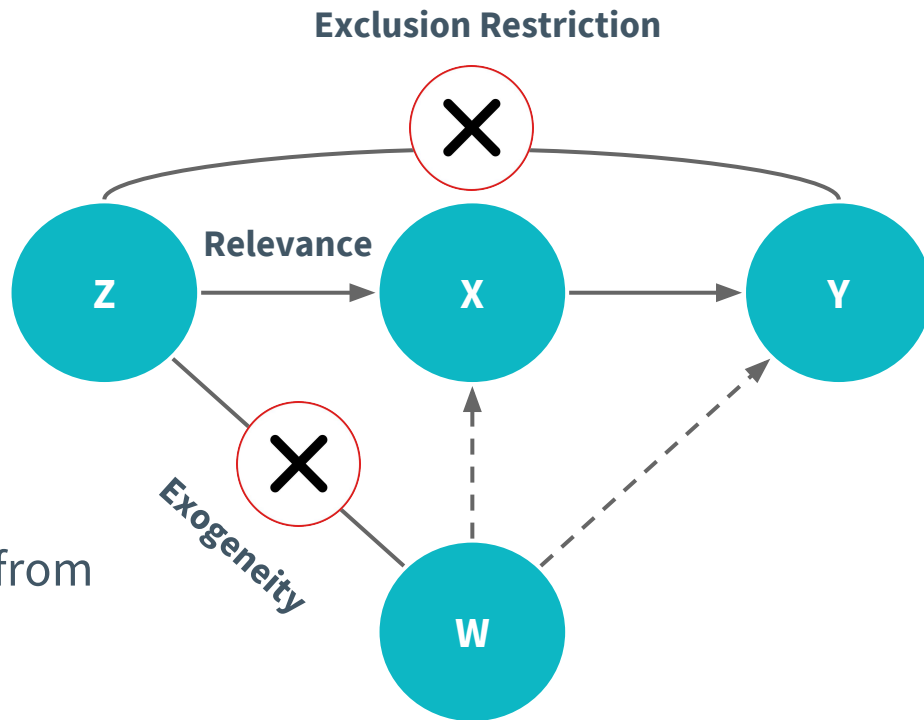☐ If W is related to X but unrelated to Y?

☑ If W is unrelated to X but related to Y?

CAUSAL BEAST

What Makes a Proper Instrumental Variable?

- ▷ **Exogeneity**
  - ▸ Cov(Z,W) = 0
- ▷ **Relevance** (Strong first stage)
  - ▸ Cov(Z,X) ≠ 0
- ▷ **Exclusion Restriction** (follows from exogeneity and relevance)
  - ▸ Z ⇒ X ⇒ Y
  - ▸ Z is conditionally independent on Y given X (chain)

Exclusion Restriction

Relevance

Exogeneity

Z   X   Y

W

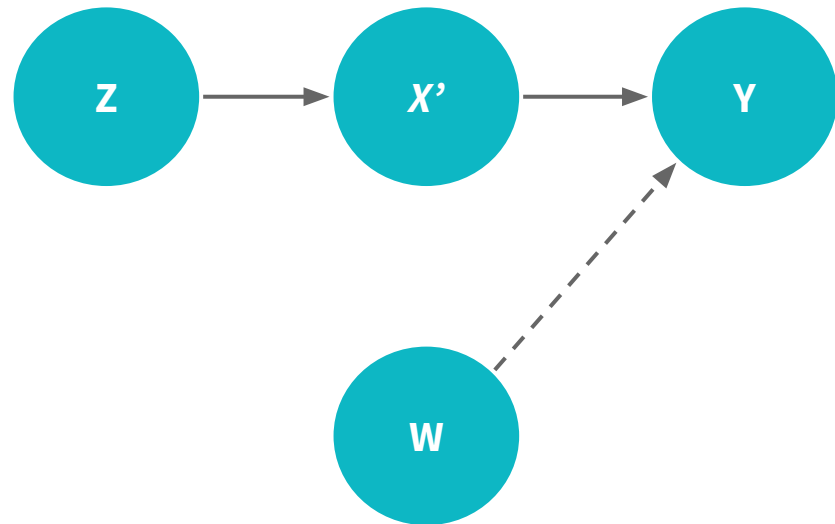# How to Make This More Familiar to Our "Ideal Scenario"

▷ The scenario we want to rid of OVB: W is unrelated to X but related to Y

▷ Solution: Capture the portion of X that is only explained by Z

▶ Denote this portion as X'
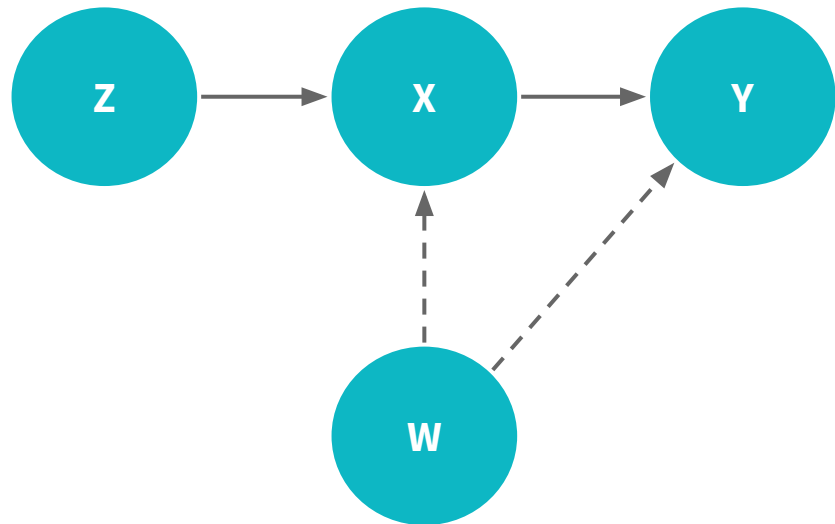
How IV Analysis Addresses the OVB Problem
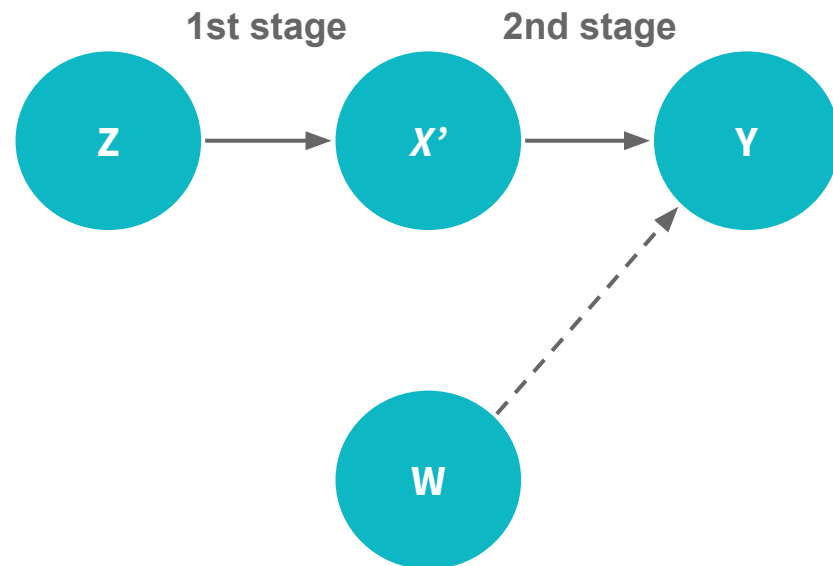
▷ *X'* is the portion of X that is independent of W (no relationship)

▷ *X'* is the predicted values of X provided just Z

Methodology: Two-Stage Least Squares (2SLS) Regression

▷ Stage 1: Regress X on Z and obtain predicted values of X

▷ Stage 2: Regress Y on predicted values of X, and obtain the 2SLS estimator

1st stage   **2nd stage**

Z → X' → Y

W ⇢ Y

CAUSAL BEAST

Acemoglu, Daron et al. (2001): The Colonial Origins of Comparative Development

▷ Acemoglu and colleagues wanted to estimate the effect of institutions on economic performance

▷ Institutions are not exogenous (randomly assigned)

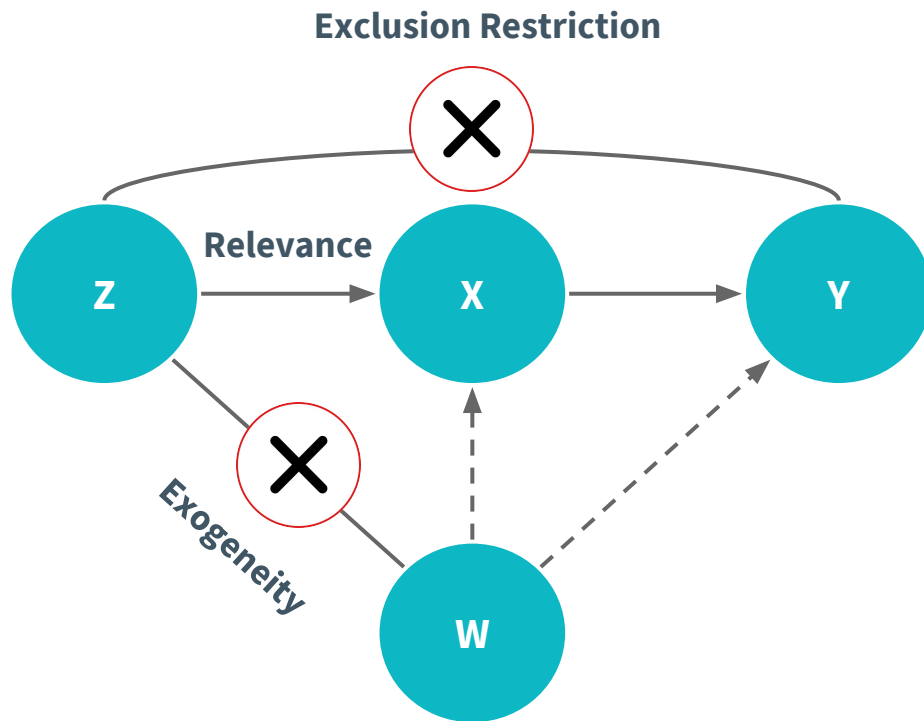# Underlying Assumptions of the Study

▷ Potential settler mortality

↓

▷ Settlements

↓

▷ Early Institutions (17th, 18th century)

↓

▷ Current Institutions

↓

▷ Current Economic Performance

CAUSAL BEAST

Settler Mortality as an IV for Institutions

▷ Settler mortality (Z)

▷ Institutional quality (X)

▷ Economic performance (Y)

▷ Assumption check?

2SLS Procedure

$$(1) \quad R_i = a_0 + a_1 + log(M_i) + a_2 X_i + s_i$$

M_i - Settler mortality

2SLS estimator (causal effect of institutions on economic performance)

$$(2) \quad log(GDP_{pc_i}) = b_0 + \boxed{b_1} R_i + b_2 X_i + e_i$$

X_i - Latitude, continent, temperature, humidity, natural resources...

R_i - Protection against expropriation measure

2SLS Results

	Base sample (1)	Base sample (2)	Base sample without Neo-Europes (3)	Base sample without Neo-Europes (4)	Base sample without Africa (5)	Base sample without Africa (6)	Base sample with continent dummies (7)	Base sample with continent dummies (8)	Base sample, dependent variable is log output per worker (9)
				Panel A: Two-Stage Least Squares					
Average protection against expropriation risk 1985–1995	0.94 (0.16)	1.00 (0.22)	1.28 (0.36)	1.21 (0.35)	0.58 (0.10)	0.58 (0.12)	0.98 (0.30)	1.10 (0.46)	0.98 (0.17)
Latitude		−0.65 (1.34)		0.94 (1.46)		0.04 (0.84)		−1.20 (1.8)	
Asia dummy							−0.92 (0.40)	−1.10 (0.52)	
Africa dummy							−0.46 (0.36)	−0.44 (0.42)	
"Other" continent dummy							−0.94 (0.85)	−0.99 (1.0)	
				Panel B: First Stage for Average Protection Against Expropriation Risk in 1985–1995					
Log European settler mortality	−0.61 (0.13)	−0.51 (0.14)	−0.39 (0.13)	−0.39 (0.14)	−1.20 (0.22)	−1.10 (0.24)	−0.43 (0.17)	−0.34 (0.18)	−0.63 (0.13)
Latitude		2.00 (1.34)		−0.11 (1.50)		0.99 (1.43)		2.00 (1.40)	
Asia dummy							0.33 (0.49)	0.47 (0.50)	
Africa dummy							−0.27 (0.41)	−0.26 (0.41)	
"Other" continent dummy							1.24 (0.84)	1.1 (0.84)	
R^2	0.27	0.30	0.13	0.13	0.47	0.47	0.30	0.33	0.28

>0 and statistically significant

The validity of our 2SLS results in Table 4 depends on the assumption that settler mortality in the past has no direct effect on current economic performance...

Acemoglu et al. (2001)

CAUSAL BEAST

Is this Good Enough?

▷ Results are contentious; arguments for or against the results are centered on whether the IV assumptions are met

▷ Most of the paper is devoted to telling the story of why the authors think the assumptions are valid

CAUSAL BEAST

Potential Pitfalls of IV Analysis

Unprovable Assumptions

Arguing for the exogeneity and exclusion restriction requires solid argumentation and theory (no empirical work will do)

Strong Relevance

Needs to be strong correlation in the first stage; otherwise potentially a lot of bias

Interpretation

If the instrument only assigns a narrow group in the population to treatment and control, then estimator is LATE (local average treatment effect)

Estimates are biased but consistent

Bias from IV estimation can actually be worse than bias from OLS in finite samples (need "a lot" of data)

Especially the case if the estimator is weak (weak first stage)

CAUSAL BEAST

Further Studies on Instrumental Variables

How Large are Human-Capital Externalities? Evidence From Compulsory Schooling Laws

Acemoglu et Angrist (2001)

Instrumental Variables Methods in Experimental Criminological Research: What, Why, and How?

Angrist (2005)

The Oregon Health Insurance Experiment: Evidence From the First Year

Finkelstein et al. (2011)

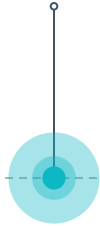**Do Institutions Cause Growth?**

Glaeser et al. (2004)

CAUSAL BEAST

Regression Discontinuity

Looking near a cutoff for a local randomized experiment
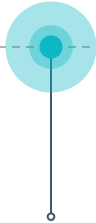
CAUSAL BEAST

Toolbox

Randomized
Experiment

Instrumental
Variables

Regression
(Controlled)

Regression
Discontinuity

CAUSAL
BEAST

Looking at the Cutoff

▷ Basic idea is that policy rules vary at some cutoff point

- ▶ Whether you are eligible for a program
- ▶ Which school district you reside in
- ▶ Whether you are of legal age to drink
- ▶ Whether you pass a university course

Pass Fail Cutoff in University

- ▷ <59 is a fail in university

- ▷ Person A who scored 58.9 fails the course

- ▷ Person B who scored 59.1 passes the course

- ▷ Person A ≈ Person B (ceteris paribus?)
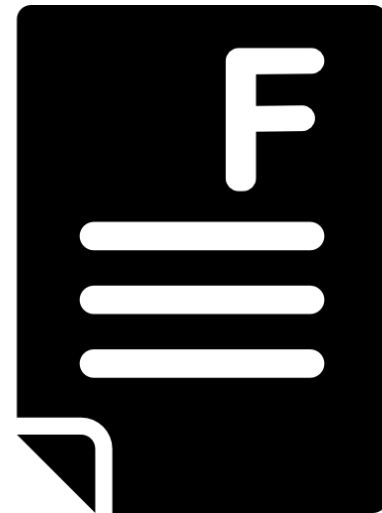
Thistlewaite, Donald, Donald Campbell (1960): Regression–Discontinuity Analysis: An Alternative to the Ex–Post Facto Experiment

▷ Paper introduced regression discontinuity design into mainstream research

▷ Does scholarship improves student achievement after school?

CAUSAL BEAST
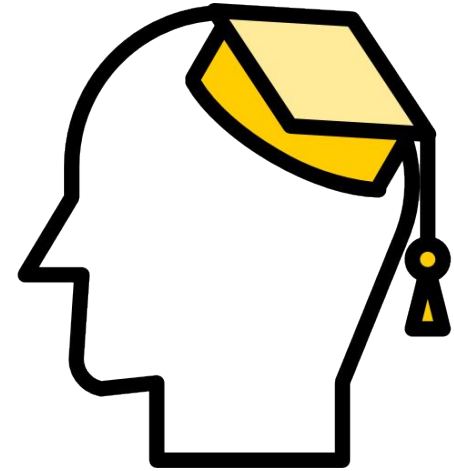
Regression of Study and Career Plans on Test Score
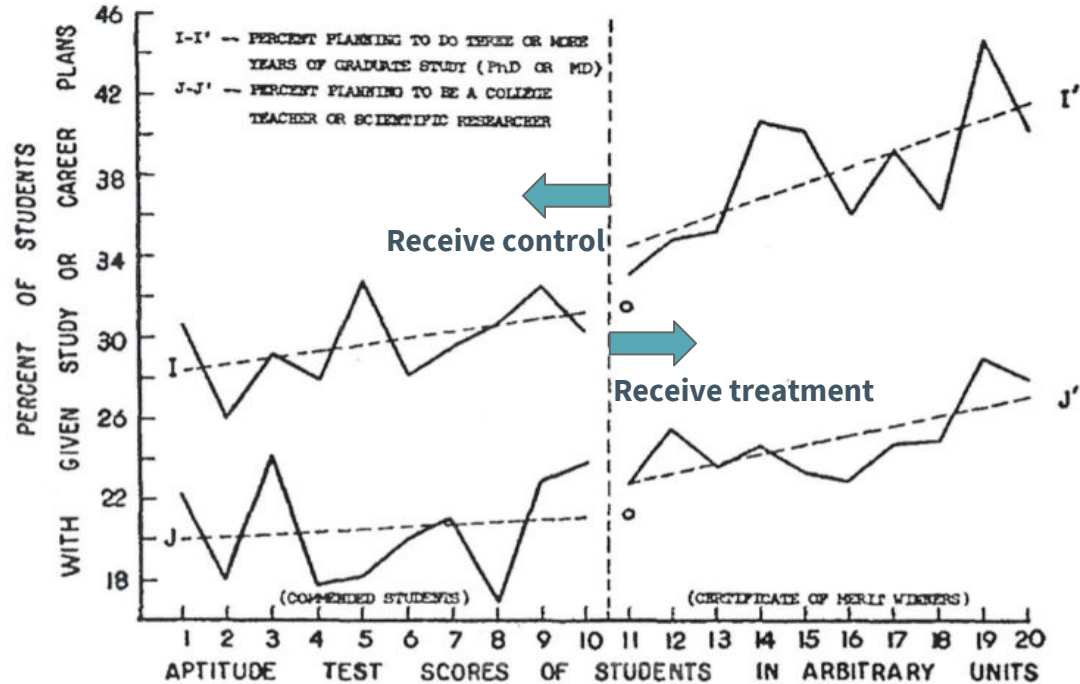

FIG. 3. Regression of study and career plans on exposure determiner.

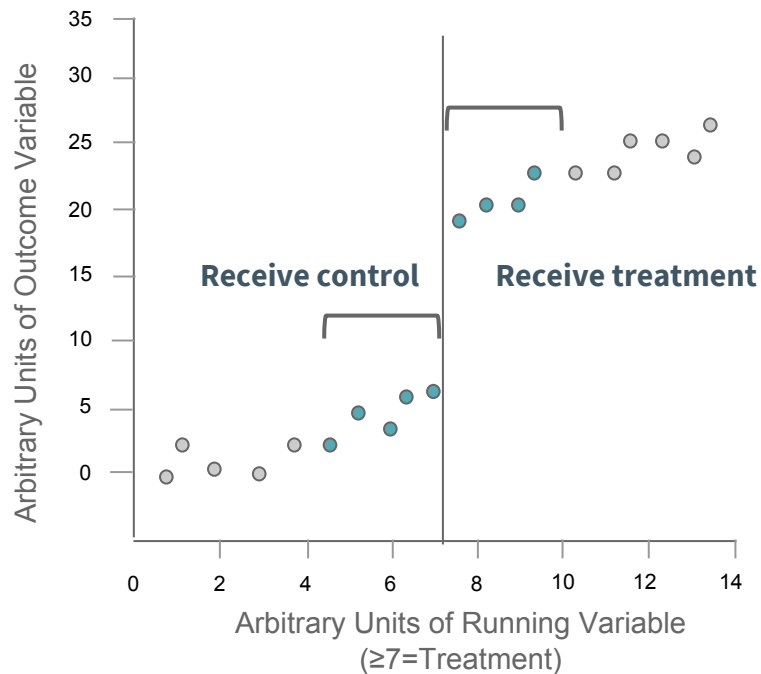# Regression Discontinuity Design (RDD) Options

▷ Nearest neighbors (fixed bandwidth)

▷ Distance kernel

▷ Interpolation (regression)

CAUSAL BEAST

Nearest Neighbors

▷ Create a fixed bandwidth of size M and average the outcomes of those K neighbors that fall within the bandwidth
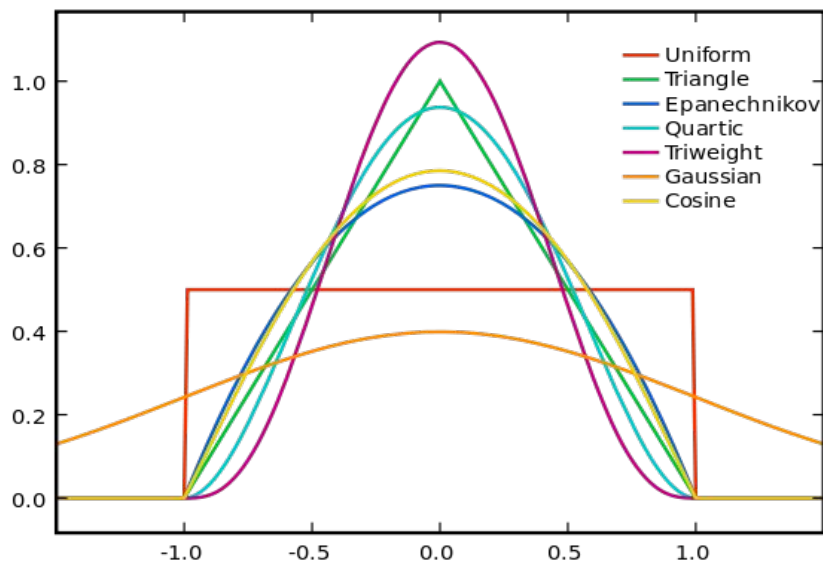
RDD w/ Fixed Bandwidth of Size 3

Distance Kernels

▷ **The weight of each point falls off as as a function of its distance from the cutoff**

▷ **Kernels specify how the weights decay**

▷ **Simple average uses a uniform kernel**

Popular Kernels in Empirical Methods

Interpolation (Regression)

▷ Regress on the left and right hand side of the cut off

▷ Then calculate the height difference between the two trends
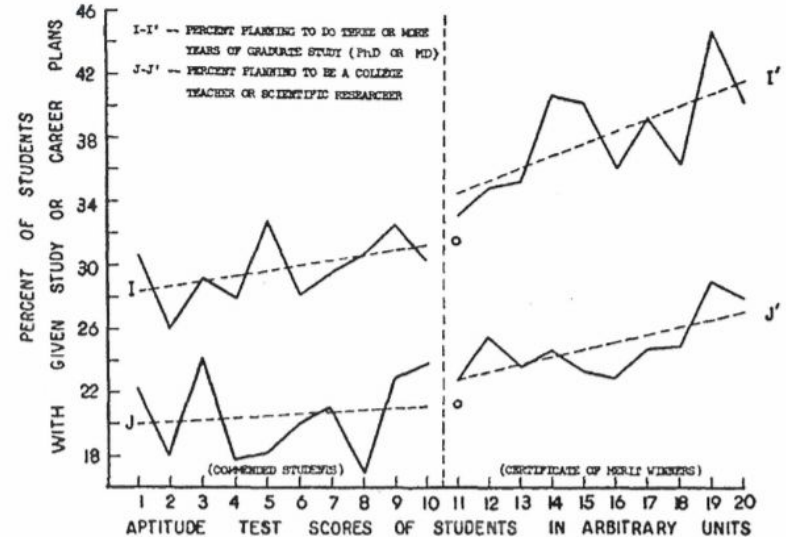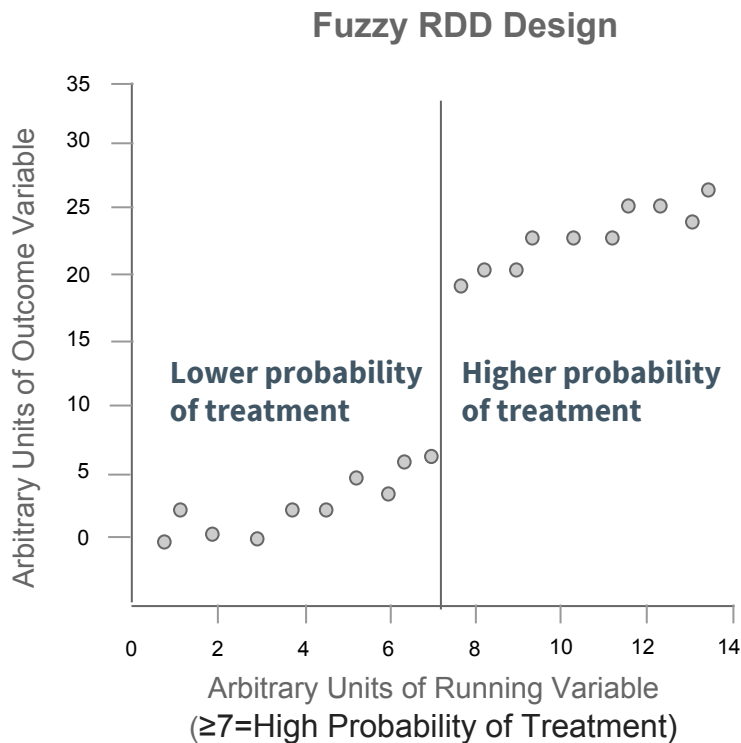
Figure 3. From Thistlewaite et Campbell (1960)

FIG. 3. Regression of study and career plans on exposure determiner.

Sharp RDD vs. Fuzzy RDD

▷ Sharp RDD
  ▸ Discontinuity of treatment (an all or nothing situation)

▷ Fuzzy RDD
  ▸ Discontinuity of the probability of treatment (higher probability above cutoff)

Fuzzy RDD Design

(≥7=High Probability of Treatment)

Black, Sandra (1999): Do Better Schools Matter? Parental Valuation of Elementary Education

▷ Remember this study from the regression module?

▷ Trying to measure the effect of school quality on home values

▷ The main results of the paper actually come from implementing RDD

CAUSAL BEAST

RDD Strategy: Look at Houses Near the District Boundaries

▷ Look at schools near the school district boundaries

▷ Why? Homes near the boundaries are going to be very similar but zoned to go to different schools

Figure 1: Example of District Boundary

Streets
Attendance Districts

FIGURE I
Example of Data Collection for One City: Melrose Streets, and Attendance District Boundaries

CAUSAL BEAST

If neighborhoods change continuously over space, by looking at houses very close to attendance district boundaries-where there is a discrete change in school quality—I can avoid the pitfalls associated with omitted neighborhood characteristics.

"

Black (1999)

Confirming that Houses Near the Boundaries are Similar

Distance from boundary:	Full sample		0.35 mile		0.20 mile		0.15 mile	
	Difference in means	T-statistic	Ratio of 0.35 to full sample[d]	T-statistic	Ratio of 0.20 to full sample[d]	T-statistic	Ratio of 0.15 to full sample[d]	T-statistic
ln (house price)	.045	3.82	0.85	3.32	0.85	3.17	0.93	3.17
Test score (sum of reading and math)	1.0	32.90	1.03	27.28	1.06	24.44	1.06	22.57
House characteristics								
Bedrooms	0.02	1.68	0.90	0.91	−0.35	−0.30	0.25	0.18
Bathrooms	0.03	2.98	0.23	0.52	−0.02	−0.05	−0.07	−0.12
Lot size	2011	11.39	0.22	2.14	0.24	1.95	0.12	0.83
Internal square footage	31	2.93	0.61	1.32	0.61	1.07	0.84	1.17
Age of building	−3.13	−6.92	0.75	−3.71	0.94	−3.76	1.09	−3.52
Neighborhood characteristics[c]								
Percent Hispanic	−.0008	−0.79	2.50	−1.35	2.50	−1.21	2.50	−1.26
Percent non-Hispanic black	−.0007	−1.50	0.43	−0.54	0.00	−0.07	−0.14	0.16
Percent 0–9 years old	.005	3.30	0.16	0.63	−0.08	−0.31	−0.30	−1.21
Percent 65+ years old	−.01	−2.04	0.40	−0.72	0.67	−1.28	0.60	−0.95
Percent female-headed households with children	−.001	−3.67	1.00	−3.17	1.20	−2.53	1.00	−2.38
Percent with bachelor's degree	.002	1.06	0.75	0.64	1.00	0.74	0.75	0.67
Percent with graduate degree	.008	3.32	0.88	2.77	0.88	3.02	0.88	3.31
Percent with less than high school diploma	−.005	−2.19	1.20	−2.02	0.80	−1.57	0.34	−0.64
Median household income	2,135	2.87	0.60	1.90	0.65	2.11	0.52	1.61

CAUSAL BEAST

Methodology: Calculate Means on Opposite Sides of the Border

Black uses a regression based approach

Focus of the study

$$(2) \qquad \ln(price_{iab}) = \alpha + X'_{iab}\beta + K'_b\phi + \boxed{\gamma}test_a + \epsilon_{iab},$$

where K_b is a vector of boundary dummies. Conceptually, this methodology is equivalent to calculating differences in mean house prices on opposite sides of attendance district boundaries (controlling for house characteristics) and relating this to differences in test scores.[3] Boundary dummies account for any unobserved characteristics shared by houses on either side of the boundary.

CAUSAL BEAST

What the Results Mean in Dollars

▷ 3.1% without RDD (not comparing houses on opposite sides of the school district line)

▷ 1.5% treatment effect with boundary fixed effect (RDD)

TABLE IV
MAGNITUDE OF RESULTS[a]

	(1) Basic hedonic regression[d]	(2) 0.35 sample boundary fixed effects	(3) 0.20 sample boundary fixed effects	(4) 0.15 sample boundary fixed effects
Coefficient on elementary school test score[b]	.035 (.004)	.016 (.007)	.013 (.0065)	.015 (.007)
Magnitude of effect (percent change in house price as a result of a 5% change in test scores)[c]	4.9%	2.3%	1.8%	2.1%
$ Value (at mean tax-adjusted house price of $188,000 in 1993)	$9212	$4324	$3384	$3948
$ Value (at median tax-adjusted house price of $158,000 in 1993)	$7742	$3634	$2844	$3318

CAUSAL BEAST

Using an approach that compares houses that are close to each other but are associated with different elementary schools, I find that parents do care about school peers and other unmeasured components of school quality. As such they are willing to pay about 2.1 percent—or $3948—more for houses that are 5 percent higher at the mean.

"

Black (1999)

Potential Pitfalls of RDD Studies

Sample Size & Bandwidth

Sample size is usually small near the cutoff, and it's sort of ambiguous how to choose the bandwidth

Confounding Discontinuities

Other rules changed at the cutoff we are unwary of

Manipulation

Subjects if they know they can receive a benefit if they reach above a certain threshold can manipulate into being treated

External Validity

LATE estimator that pertains to people in a narrow or local range around the cutoff

CAUSAL BEAST

Further Studies on Regression Discontinuity

Evaluating the Worker Profiling and Reemployment Services System Using a Regression Discontinuity Design

Black et al. (2007)

Economic Impacts of New Unionization on Private Sector Employers

DiNardo et Lee (2004)

Using Maimonides' Rule to Estimate the Effect of Class Size on Student Achievement
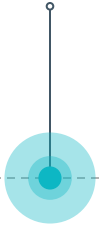
Angrist et Lavy (1997)

Waiting for Life to Arrive: A History of the Regression Discontinuity Design in Psychology, Statistics and Economics

Cook (2008)

Regression Discontinuity Designs: A Guide to Practice

Imbens (2007)

CAUSAL BEAST

Difference-in-Differences

Finding causality using parallel trends
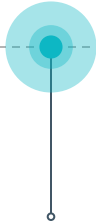
CAUSAL BEAST

Toolbox

Randomized
Experiment

Instrumental
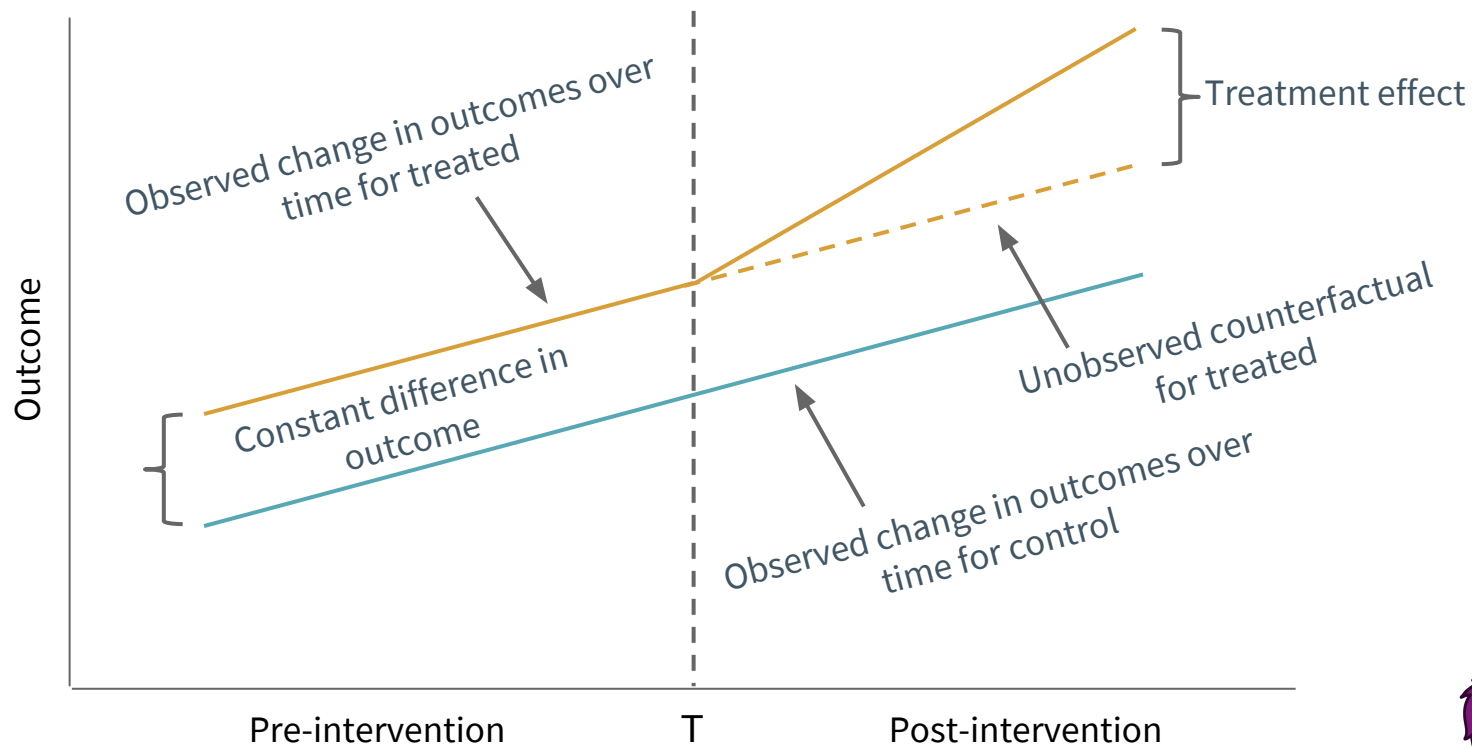Variables

Difference-in-
Differences

Regression
(Controlled)

Regression
Discontinuity

CAUSAL
BEAST

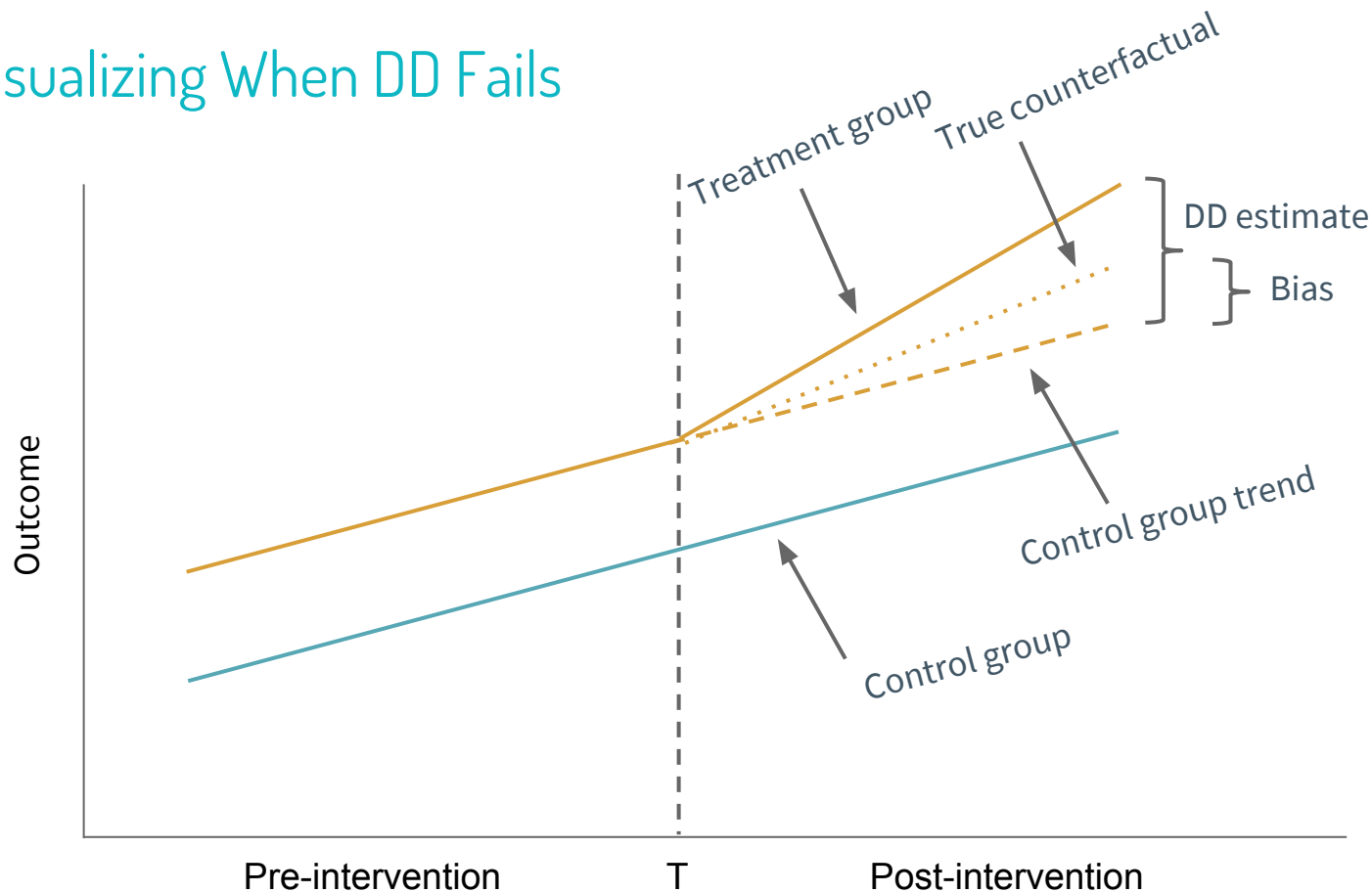# Intuition of Difference-in-Differences (DD)

Example of DD Calculation
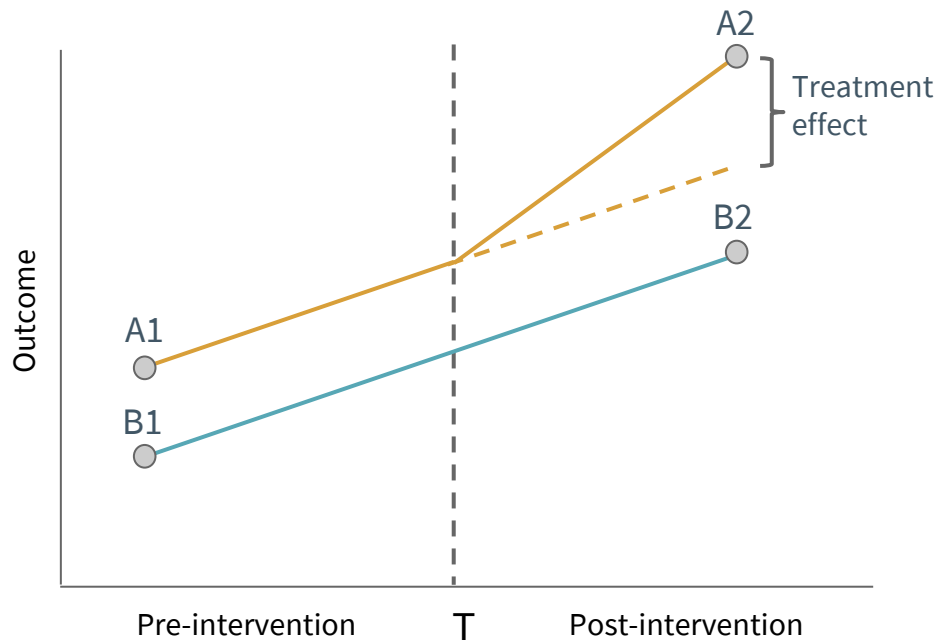
DD Table

	After	**Before**	**Difference**
Treatment	**A2**	**A1**	**A2-A1**
Control	**B2**	**B1**	**B2-B1**
Difference	**A2-B2**	**A1-B1**	**(A2-A1)-(B2-B1)**

DD estimate of causal effect

CAUSAL BEAST

Visualizing When DD Fails

Intuiting the Point of DD With Simple Policy Analysis

▷ Policy analysts often use pre-post analysis to determine the effect of a policy (e.g. What's the effect of raising the minimum wage?)

► Randomly sample low-wage firms and measure employment before minimum wage increase, then again after minimum wage

▷ What's wrong with this?

Card, David, Alan B. Krueger (1994): Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania

▷ Intervention: New Jersey's minimum wage increase in 1992

▷ Treatment group: fast food restaurants in New Jersey

▷ Control group: fast food restaurants in Eastern Pennsylvania
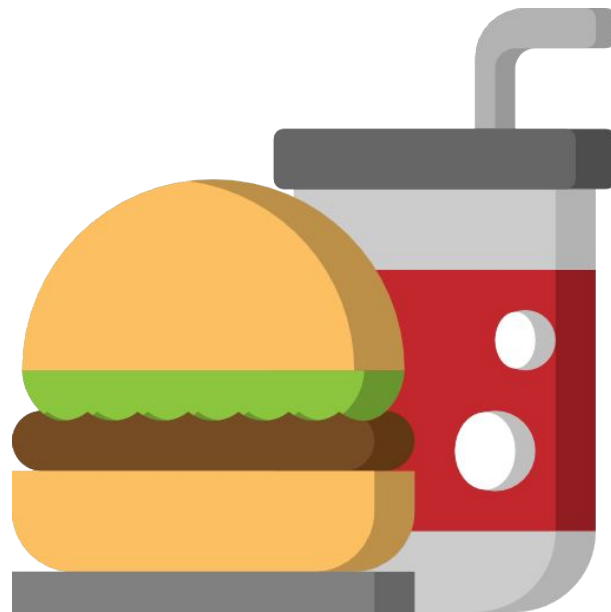
Results of DD Study

▷ Evidence that NJ's minimum wage increase $4.25 to $5.05 per hour actually increased employment

▷ Is this conclusive?
  ▸ Theoretically seems unlikely
  ▸ Much of the paper is dedicated on corroborating the finding since it runs counter to theory

Table 3 - Average Employment Per Store Before and After the Rise in NJ Minimum Wage Law

	NJ	Penn	Difference
Before	23.33	20.44	-2.89
After	21.17	21.03	-0.14
Difference	-2.16	0.59	2.76

DD estimate of causal effect

CAUSAL BEAST

The relative gain (the "difference in differences" of the changes in employment) is 2.76 FTE employees.... Contrary to the central prediction of the textbook model of the minimum wage,...we find no evidence that the rise in New Jersey's minimum wage reduced employment at fast-food restaurants in the state.
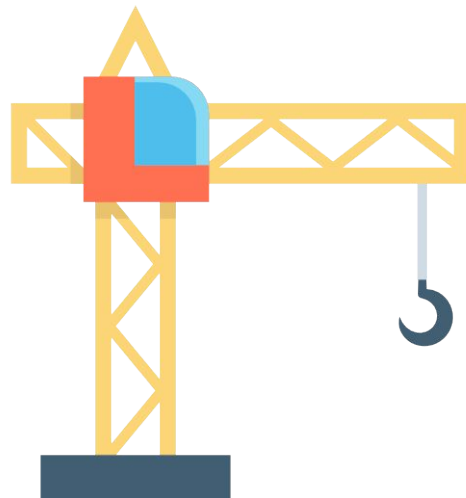
Card et Krueger (1994)

CAUSAL BEAST

▷ Sekolah Dasar INPRES program (1973-1979)

▷ ~61,000 primary schools constructed to increase educational attainment, especially in low-education areas

What are the Axes in the 2x2 DD Table in the Study?

▷ **Region of birth** (control or treatment)
  ▸ More schools built in lesser educated areas
  ▸ Simplify the intensity of the program in a cohort's region as *high* or *low*

▷ **Date of birth** (age at time of program)
  ▸ Young cohort who benefitted
  ▸ Older cohort who did not benefit

CAUSAL BEAST
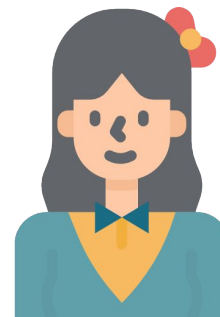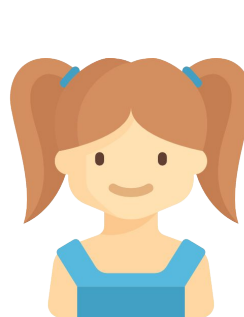
Program Effect on Education

Higher education in low intensity areas as part of INPRES mandate

TABLE 3—MEANS OF EDUCATION AND LOG(WAGE) BY COHORT AND LEVEL OF PROGRAM CELLS

	Years of education			Log(wages)		
	Level of program in region of birth			Level of program in region of birth		
	High (1)	Low (2)	Difference (3)	High (4)	Low (5)	Difference (6)
Panel A: Experiment of Interest						
Aged 2 to 6 in 1974	8.49	9.76	−1.27	6.61	6.73	−0.12
	(0.043)	(0.037)	(0.057)	(0.0078)	(0.0064)	(0.010)
Aged 12 to 17 in 1974	8.02	9.40	−1.39	6.87	7.02	−0.15
	(0.053)	(0.042)	(0.067)	(0.0085)	(0.0069)	(0.011)
Difference	0.47	0.36	0.12	−0.26	−0.29	0.026
	(0.070)	(0.038)	(0.089)	(0.011)	(0.0096)	(0.015)

Education grew more for those in high-intensity areas where more schools were built (Note: std error implies insignificant)

CAUSAL BEAST

In both types of regions, average educational attainment increased over time. However, it increased more in regions that received more schools. The difference in these differences can be interpreted as the causal effect of the program, under the assumption that, in the absence of the program, the increase in educational attainment would not have been systematically different in low and high regions.

"

Duflo (2001)

CAUSAL BEAST

Potential Pitfalls of Difference-in-Differences

Parallel Trends Assumption

DD does not identify the treatment effect if treatment and control groups were on different trajectories prior to intervention

Differential trends in subsection of the population (unbiased for just the local group)

Concomitant Treatments

Confounding events (e.g. other laws and programs implemented)

Pre-treatment Trends

Policy makers intervene to improve a bad condition; therefore treatment effect can be biased (also see Ashenfelter dip)

Manipulate Treatment Assignment

People may migrate such that the population resident in the treatment group before and after the policy is different

CAUSAL BEAST

Further Studies on Difference-in-Differences

Semiparametric Difference-in-Differences Estimators

Abadie (2005)

Water for Life: The Impact of the Privatization of Water Services on Child Mortality

Galiani et al. (2005)

How much should we trust differences-in-differences Estimates?

Bertrand et al. (2004)

Do Conditional Cash Transfers Improve Child Health? Evidence from PROGRESA's Control Randomized Experiment

Gertier (2004)

Estimating the Effect of Training Programs on Earnings

Ashenfelter (1978)

CAUSAL BEAST

Final

Concluding remarks and congratulations

5 Techniques for Causal Inference

You've got some amazing tools there, buddy!

CAUSAL BEAST

The Tools

▷ Randomized Experiment

▷ Controlled Regression

▷ Instrumental Variables

▷ Regression Discontinuity

▷ Difference-in-Differences

The Moral Lesson?

▷ Causal inference is tough because it takes some detective work

▷ In assessing the literature, we will want to read empirical work with a critical eye

▷ Don't be the lazy analyst who just nods

Thank You!

Any questions?

You can find me at

▷   @dhexonian
▷   dhexonian@outlook.com

CAUSAL BEAST

Credits

Many thanks to all the people who made creating this course an enjoyable process

▷ Grace McCormack (economist/advisor)

▷ Sung-Min Kim (copy editor)

▷ Antonela Carpena (illustrator)

▷ Matt Polly (text-to-speech, voice)

▷ Emily Glassberg Sands (inspiration)

▷ Matthew Masten (inspiration)

▷ Andrew Ng (inspiration)

▷ Elizabeth Choe (inspiration)

▷ Katharine Kormanik (inspiration)

CAUSAL BEAST

References

▷ Acemoglu, Daron, Simon Johnson, and James A. Robinson (2001): The Colonial Origins of Comparative Development: An Empirical Investigation, The American Economic Review, 91, 1369-1401.

▷ Sandra E. Black; Do Better Schools Matter? Parental Valuation of Elementary Education, The Quarterly Journal of Economics, Volume 114, Issue 2, 1 May 1999, Pages 577–599, https://doi.org/10.1162/003355399556070

▷ Angrist, Joshua David., and Jörn-Steffen Pischke. Mostly Harmless Econometrics: An Empiricist's Companion. Princeton: Princeton University Press, 2009.

▷ A Finkelstein, S Taubman, B Wright, M Bernstein, J Gruber, JP Newhouse, The Quarterly Journal of Economics 127 (3), 1057-1106

▷ Banerjee, Abhijit, Esther Duflo, Rachel Glennerster, and Cynthia Kinnan. 2015. "The Miracle of Microfinance? Evidence from a Randomized Evaluation." American Economic Journal: Applied Economics, 7(1): 22-53.

▷ Wooldridge, Jeffrey M. Econometric Analysis of Cross Section and Panel Data. Cambridge, Mass: MIT Press, 2002. Print.

▷ Rubin, Donald B. (1973): Matching to Remove Bias in Observational Studies,Biometrics, 29, 159-83.

▷ Judea Pearl; Causal diagrams for empirical research, Biometrika, Volume 82, Issue 4, 1 December 1995, Pages 669–688, https://doi.org/10.1093/biomet/82.4.669

▷ Imbens, Guido, and Joshua Angrist (1994): Identification and Estimation of Local Average Treatment Effects, Econometrica, 62, 467-476.

▷ Imbens, Guido, and Thomas Lemieux (2008): Regression Discontinuity Designs: A Guide to Practice, Journal of Econometrics, 142, 615-635.

▷ Angrist, Joshua D., and Alan B. Krueger (1991): Does Compulsory Schooling Attendance Affect Schooling and Earnings?,Quarterly Journal of Economics, 106, 976-1014.

▷ Duflo, Esther (2001). "Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence From an Unusual Policy Experiment," American Economic Review, Sept 2001

CAUSAL BEAST

Creative Commons

CAUSAL BEAST