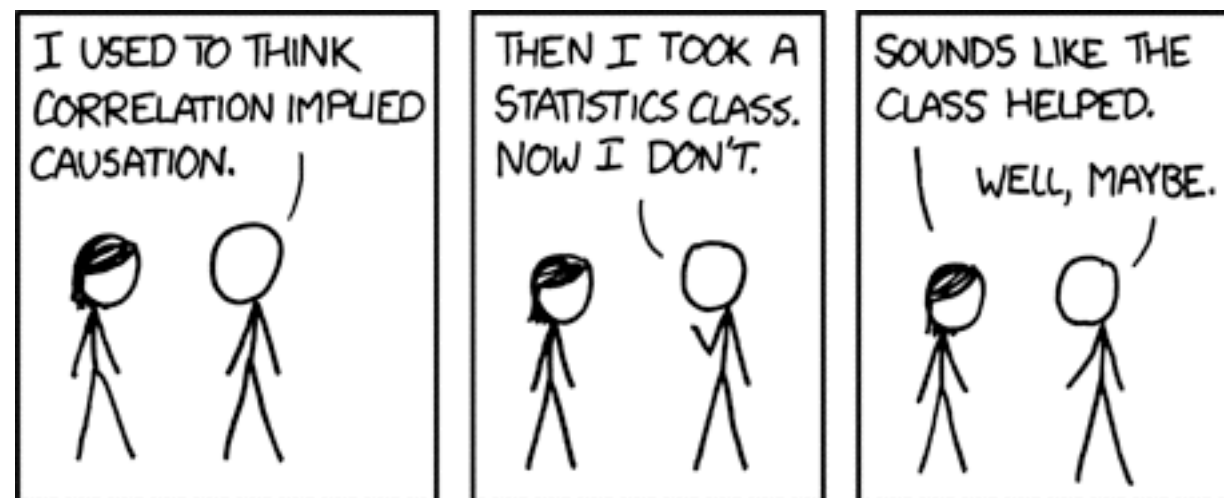


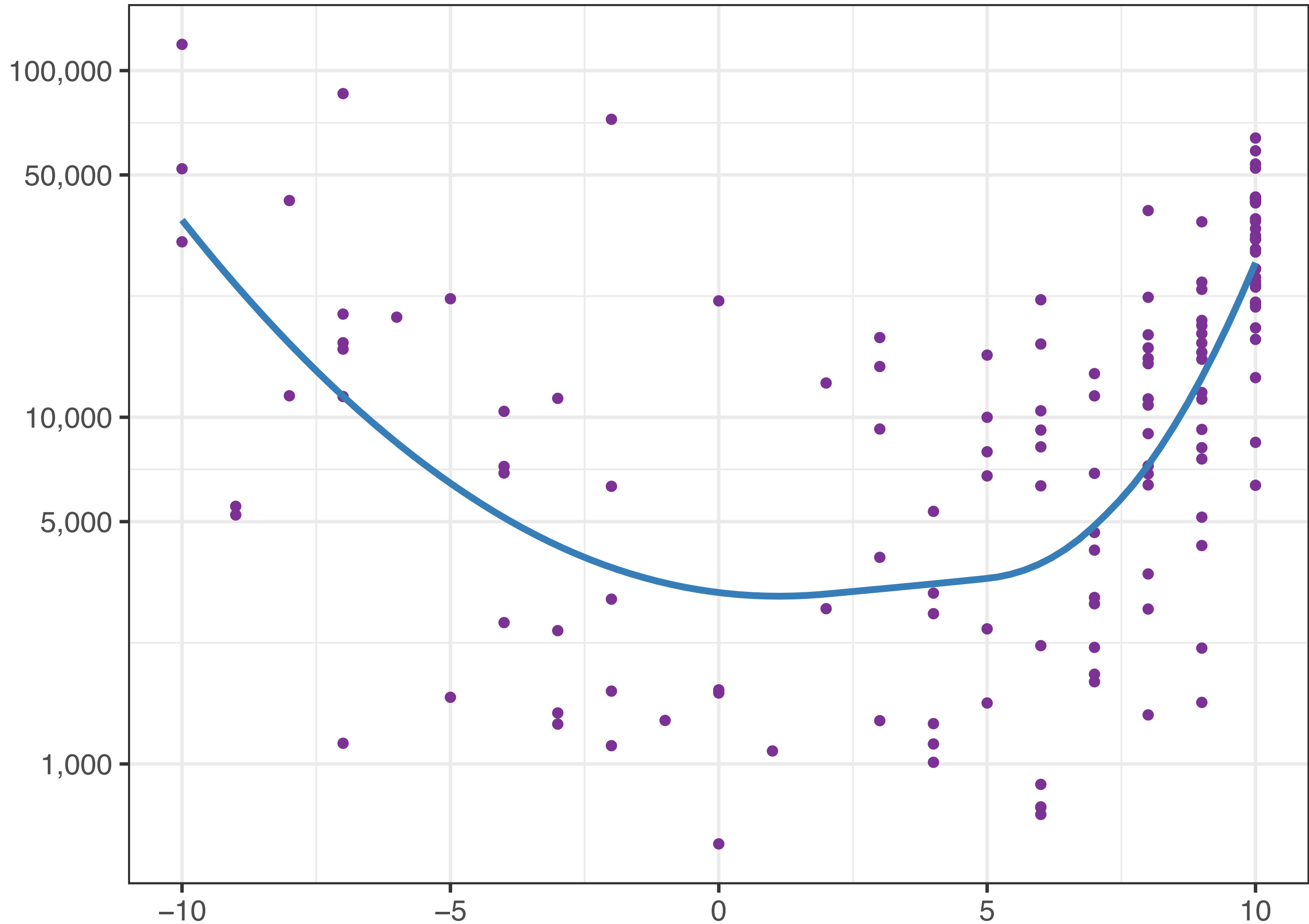
# A Partial Solution

To the Fundamental Problem of Causal Inference



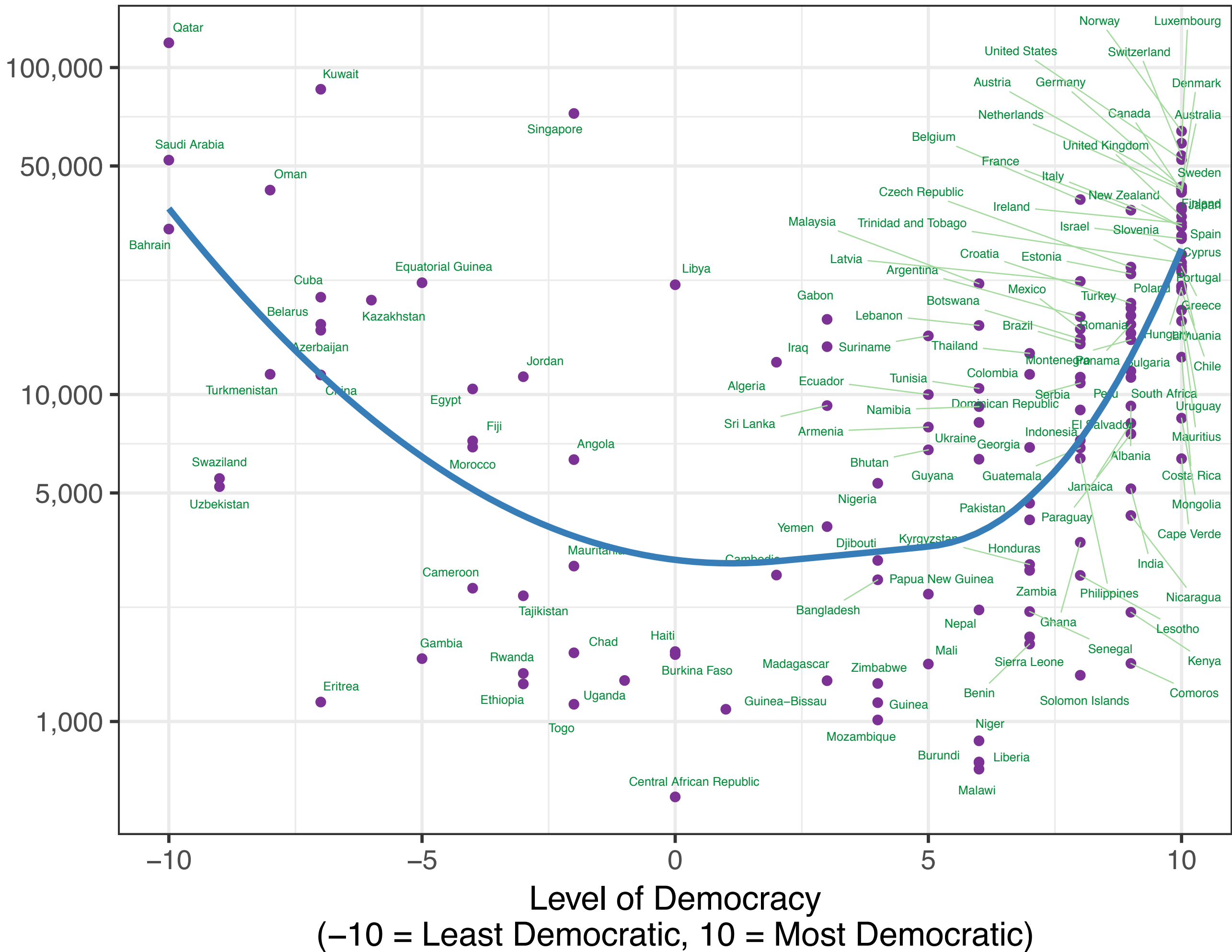
Some of our most  
important questions are  
**causal questions.**

Gross National Income Per Capita

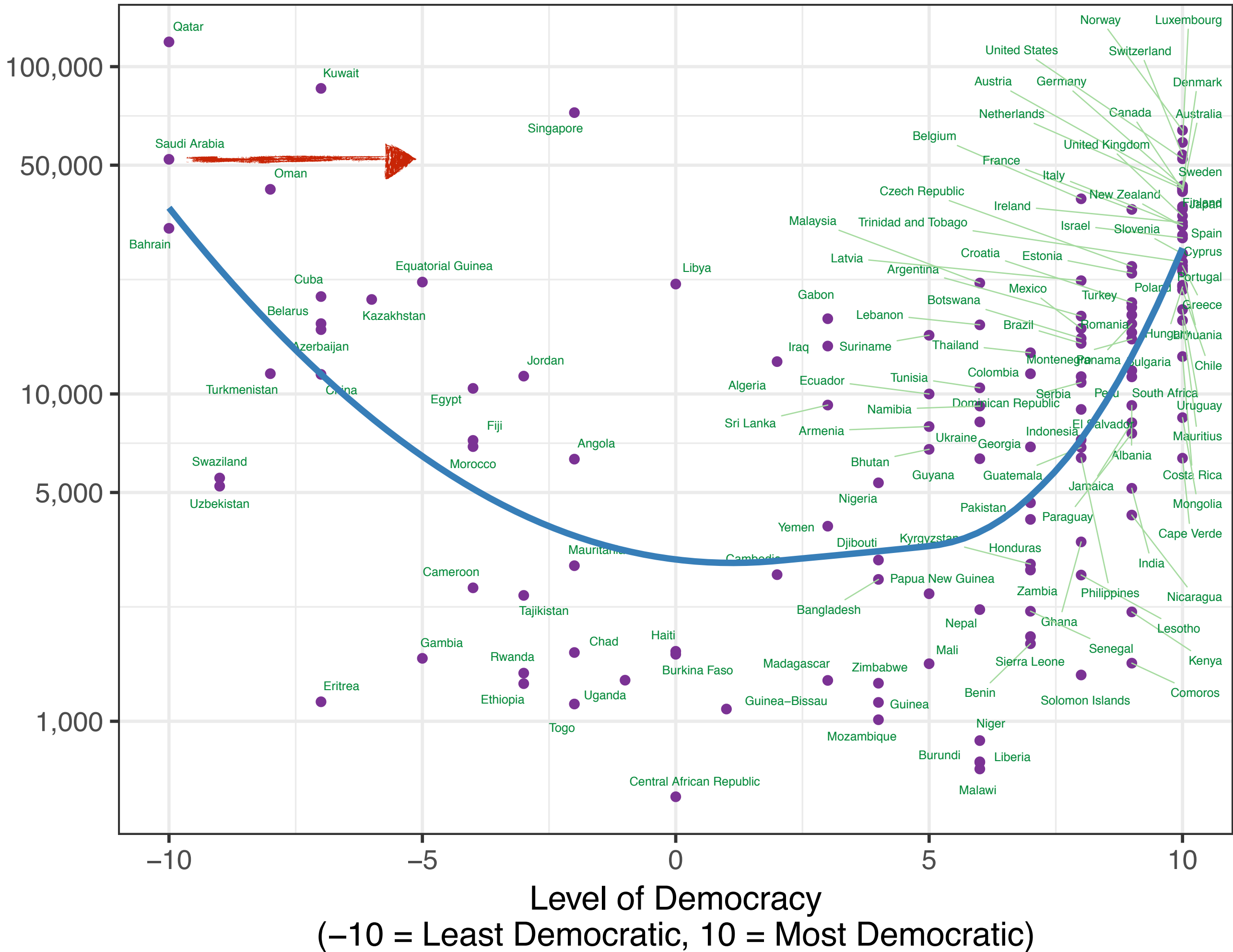


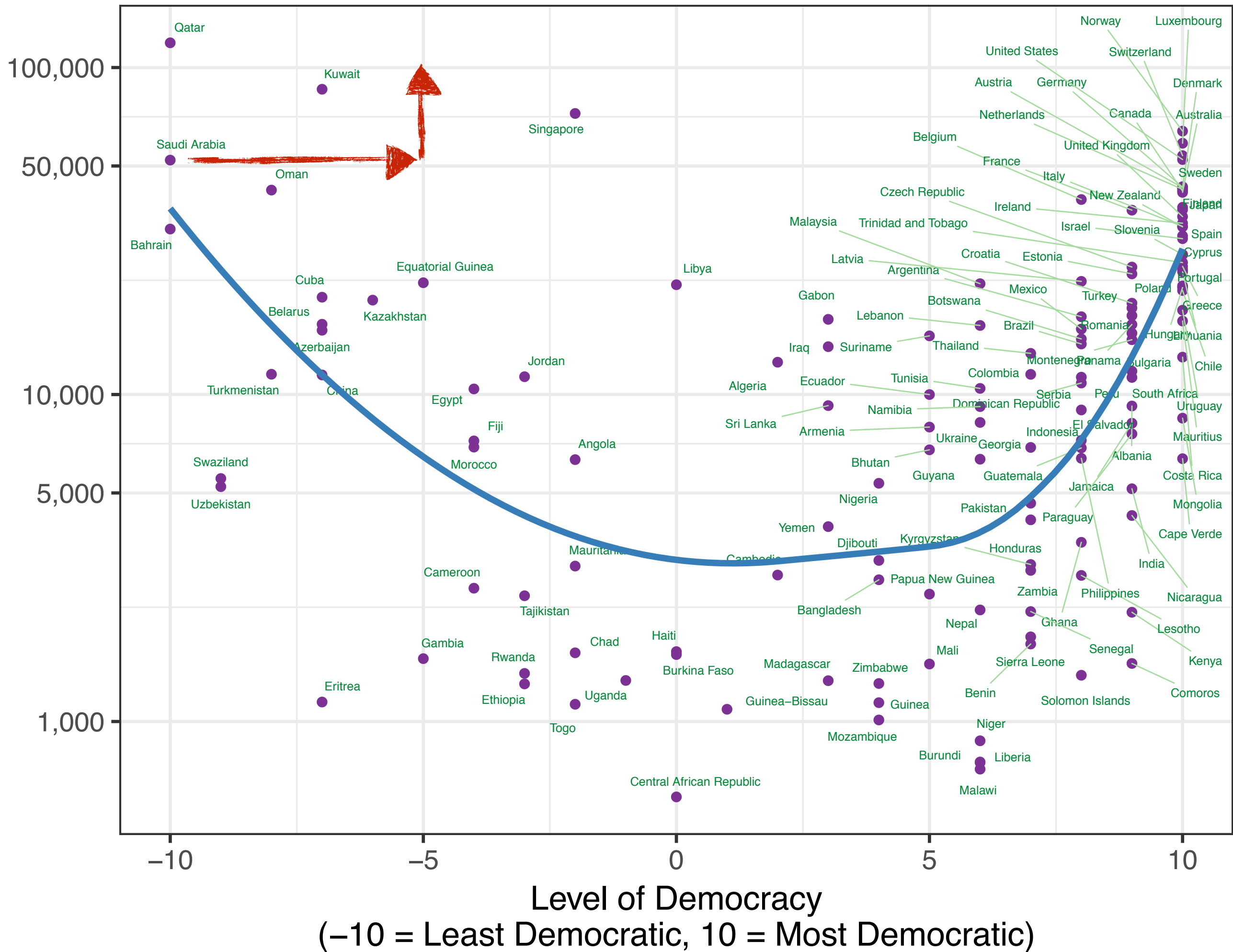
Level of Democracy  
(-10 = Least Democratic, 10 = Most Democratic)

# Gross National Income Per Capita

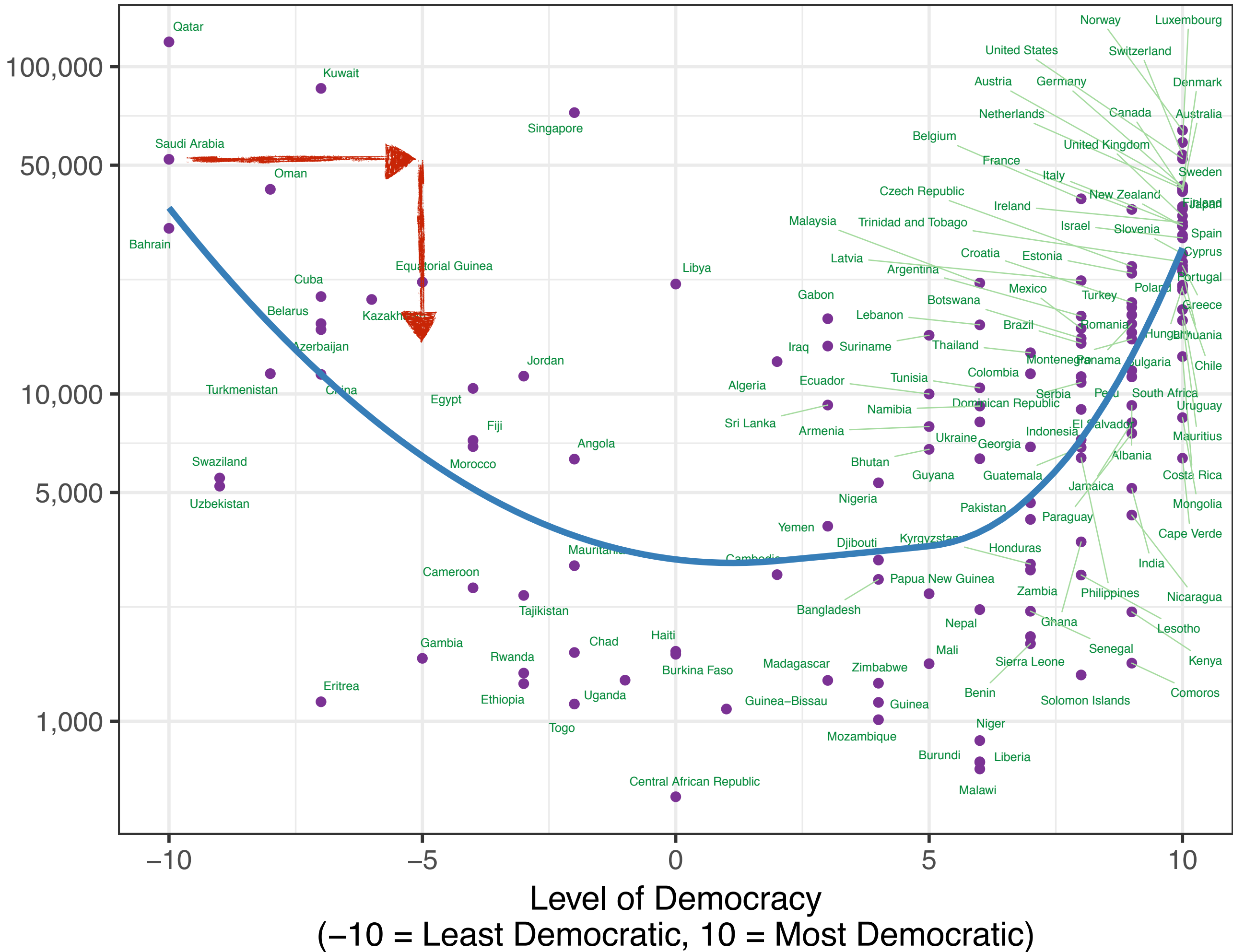


# Gross National Income Per Capita

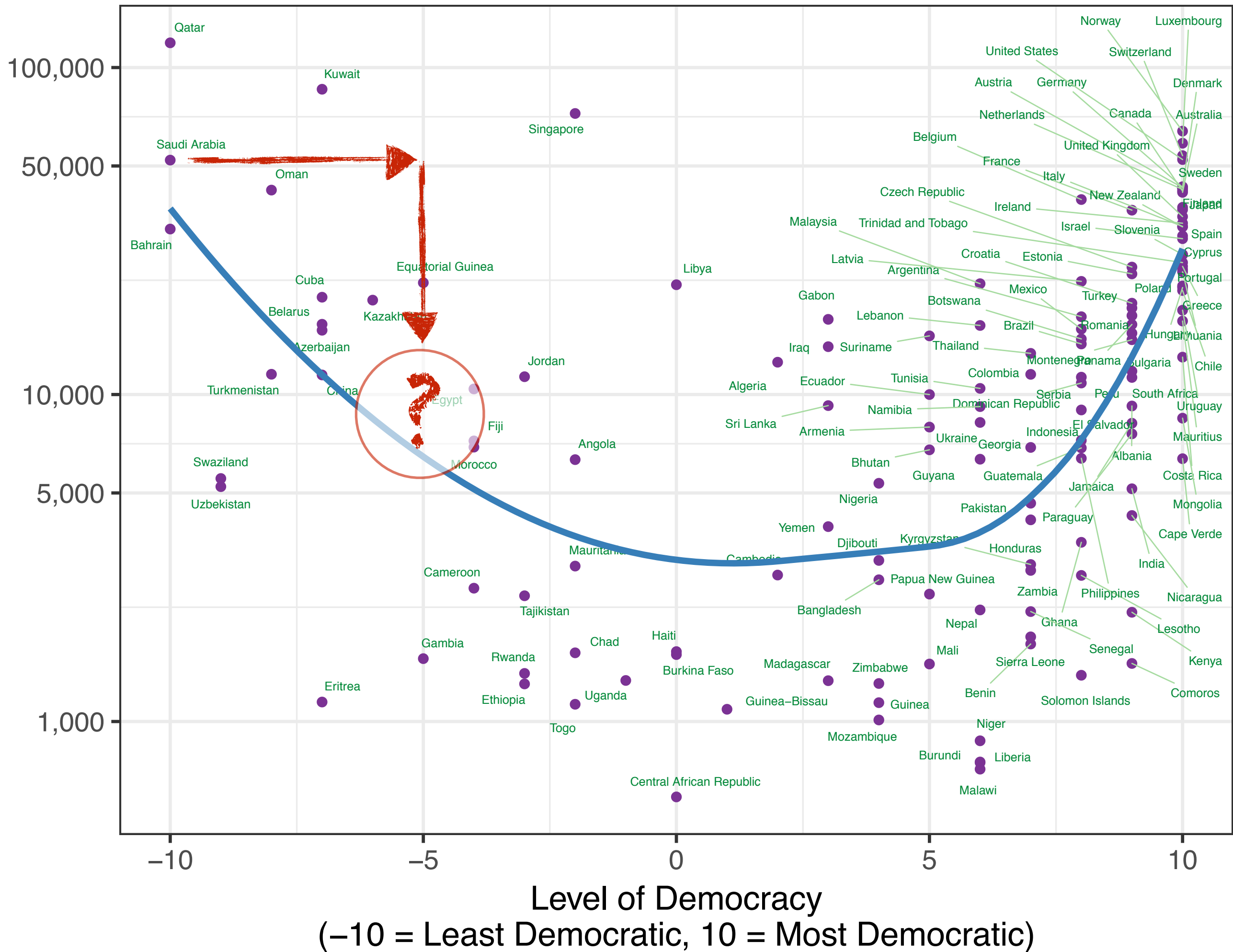




Gross National Income Per Capita

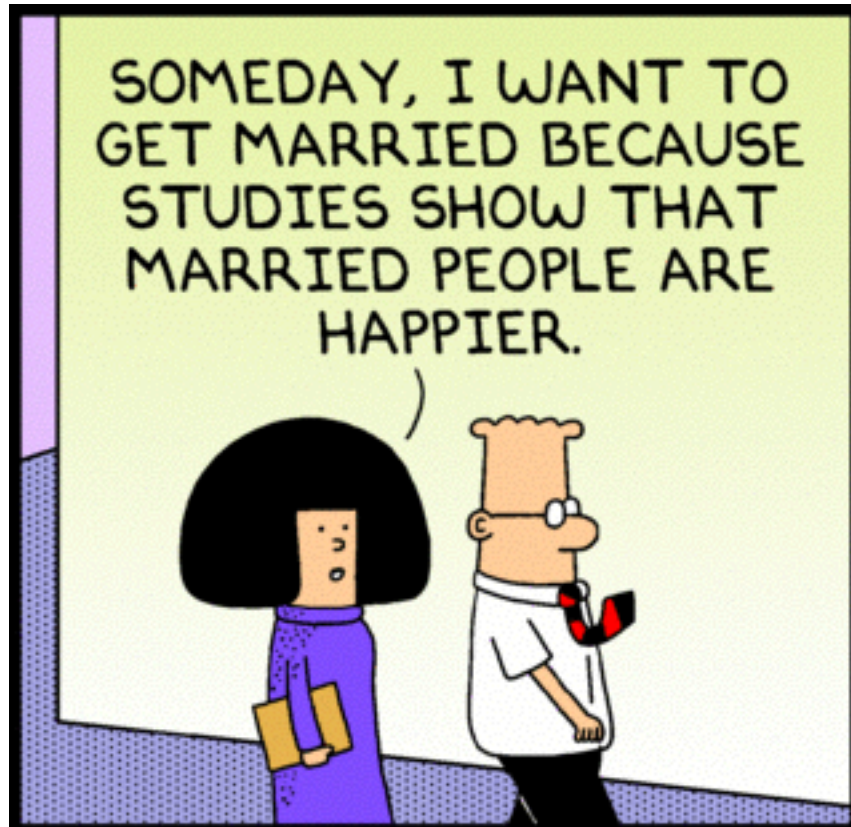




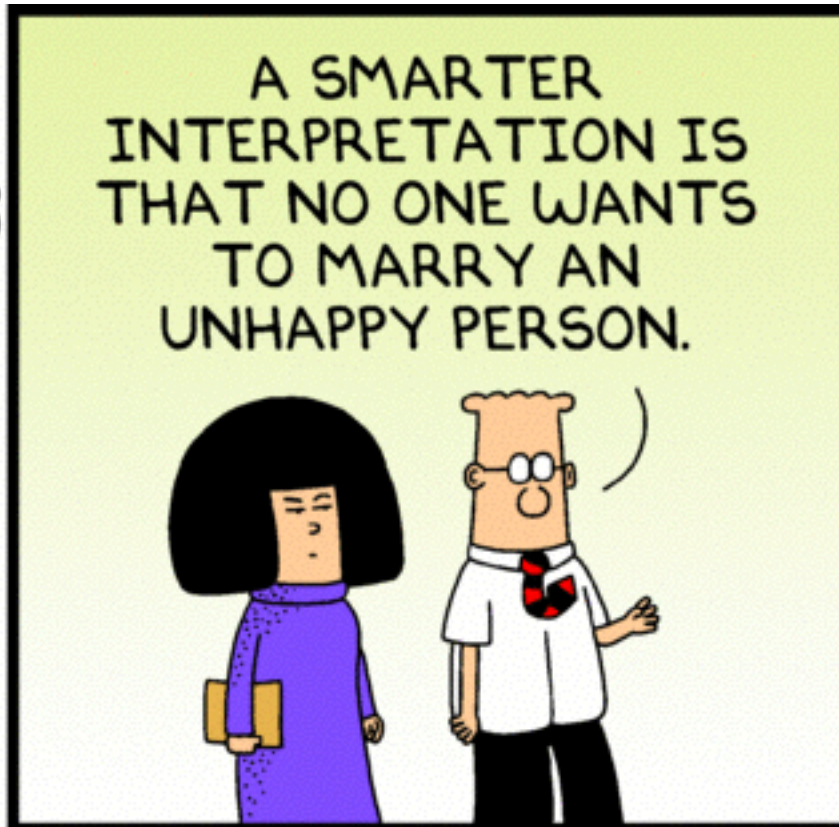




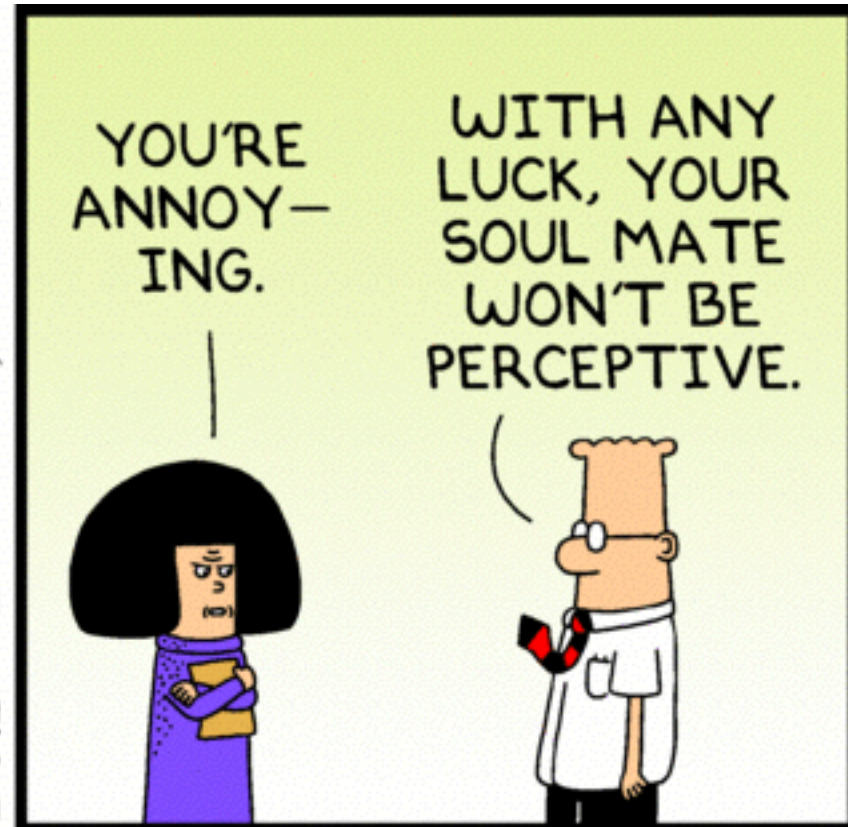
correlation  $\nrightarrow$  correlation



Dilbert.com DilbertCartoonist@gmail.com



2-10-12 ©2012 Scott Adams, Inc./Dist. by Universal Uclick



# Four Ways

to Get a Correlation

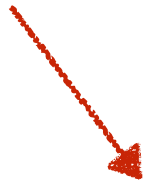
causation

causation



# causation

key explanatory variable



**X**



**Y**

# causation

key explanatory variable

outcome variable

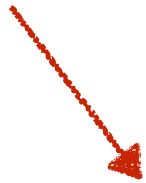
**X**



**Y**

# causation

key explanatory variable



**X**

causes



outcome variable

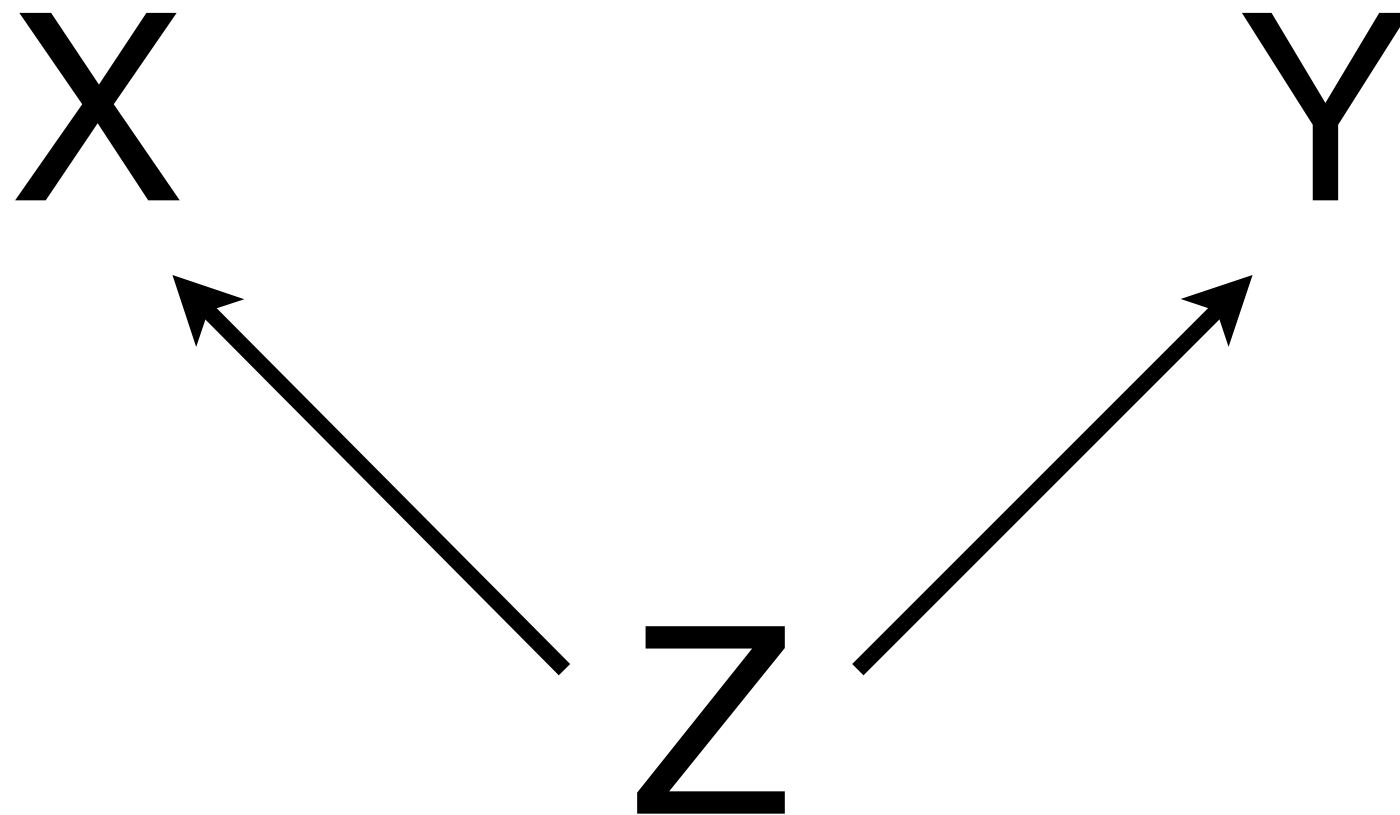


**Y**



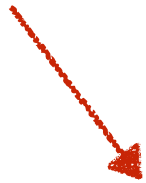
spuriousness

# spuriousness



# spuriousness

key explanatory variable

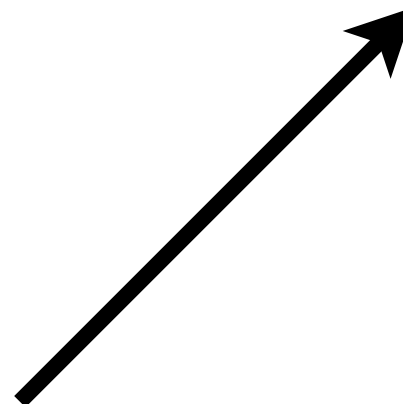


**X**

**Y**



**Z**



# spuriousness

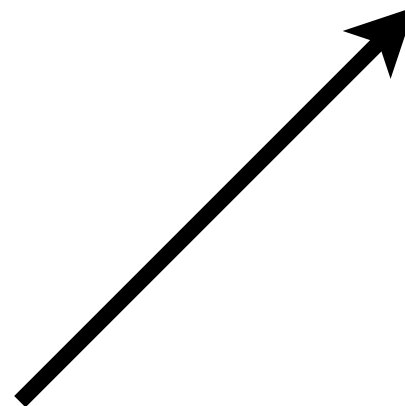
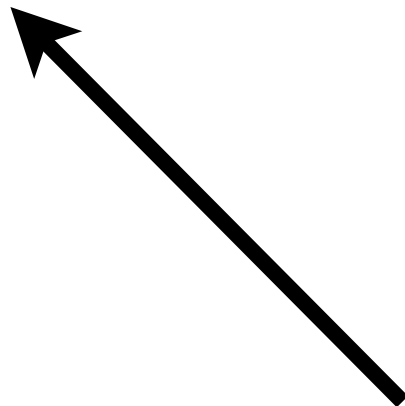
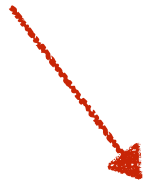
key explanatory variable

outcome variable

X

Y

Z



# spuriousness

key explanatory variable

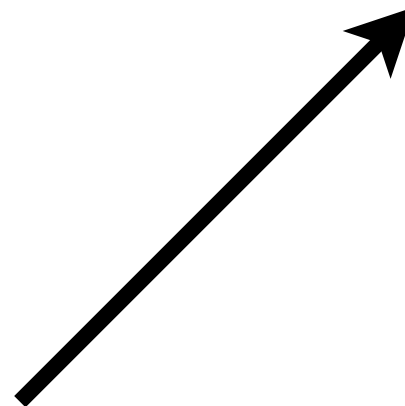
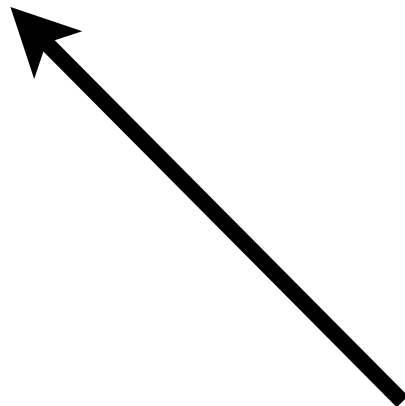
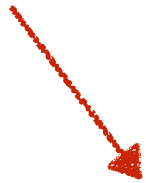
outcome variable

X

Y

Z

confounder



# spuriousness

key explanatory variable

outcome variable

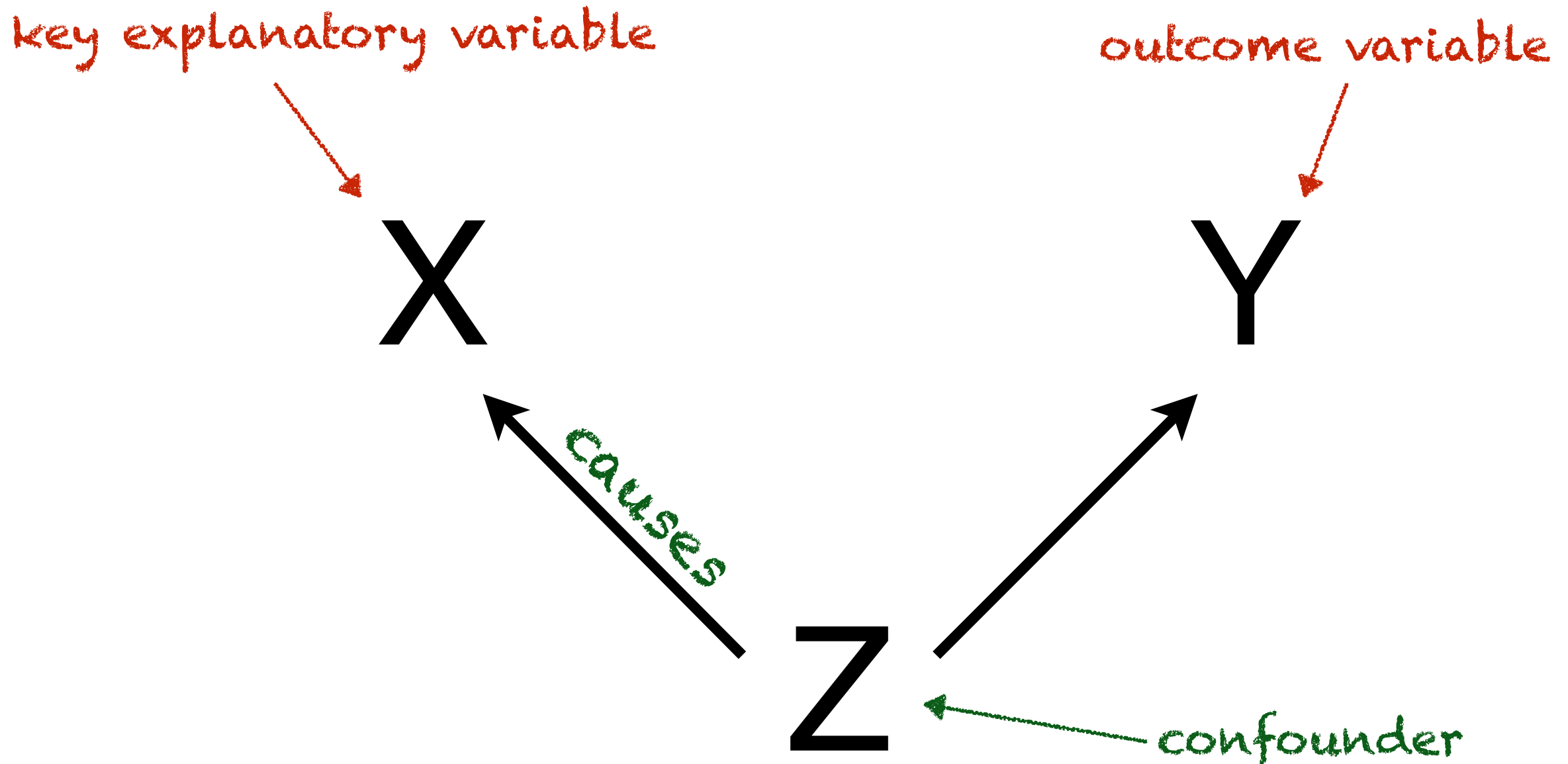
X

Y

causes

Z

confounder



# spuriousness

key explanatory variable

outcome variable

X

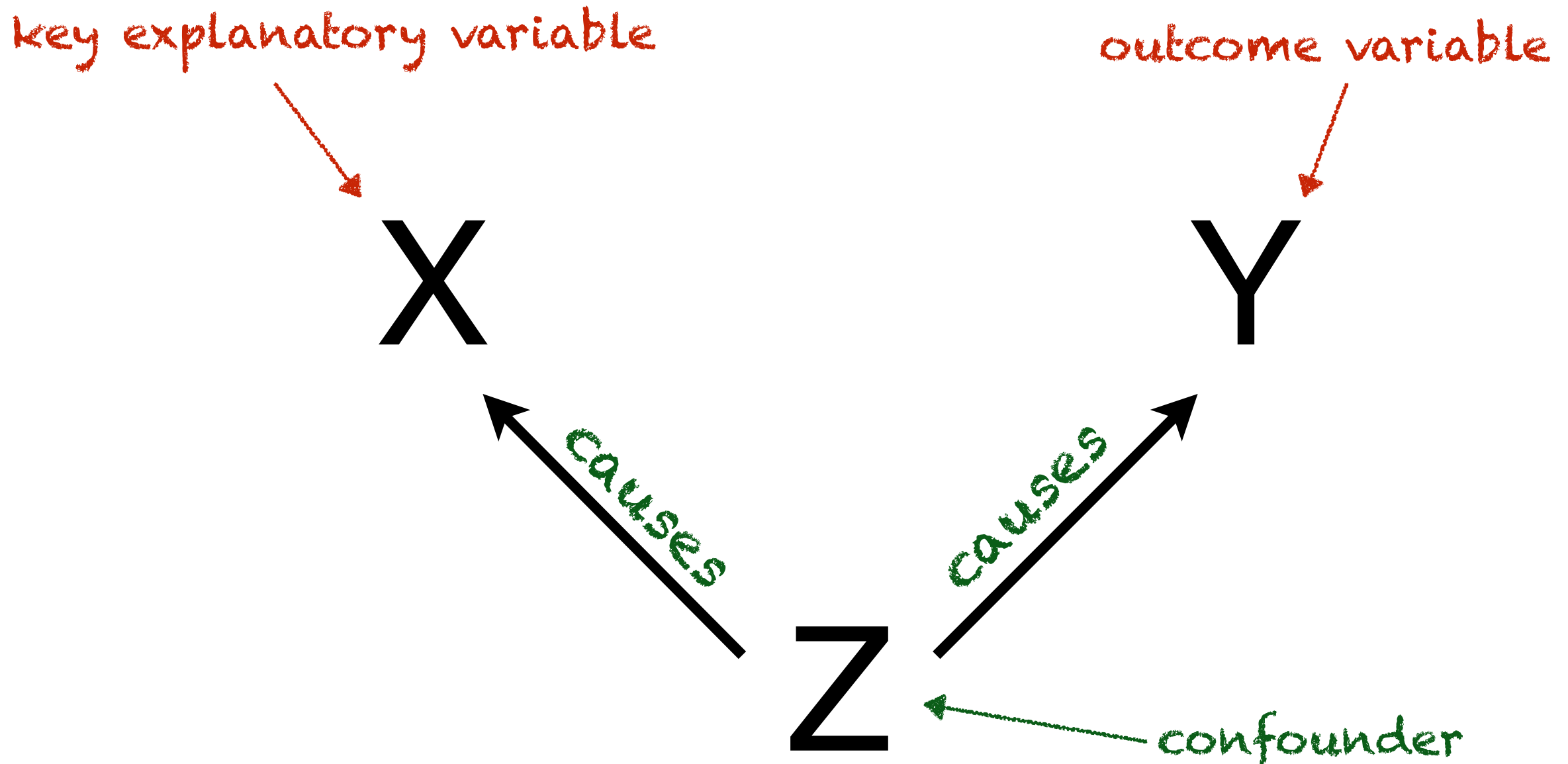
Y

causes

causes

Z

confounder





Note: a confounder is a variable that causes both X and Y.

spuriousness

key explanatory variable

outcome variable

X

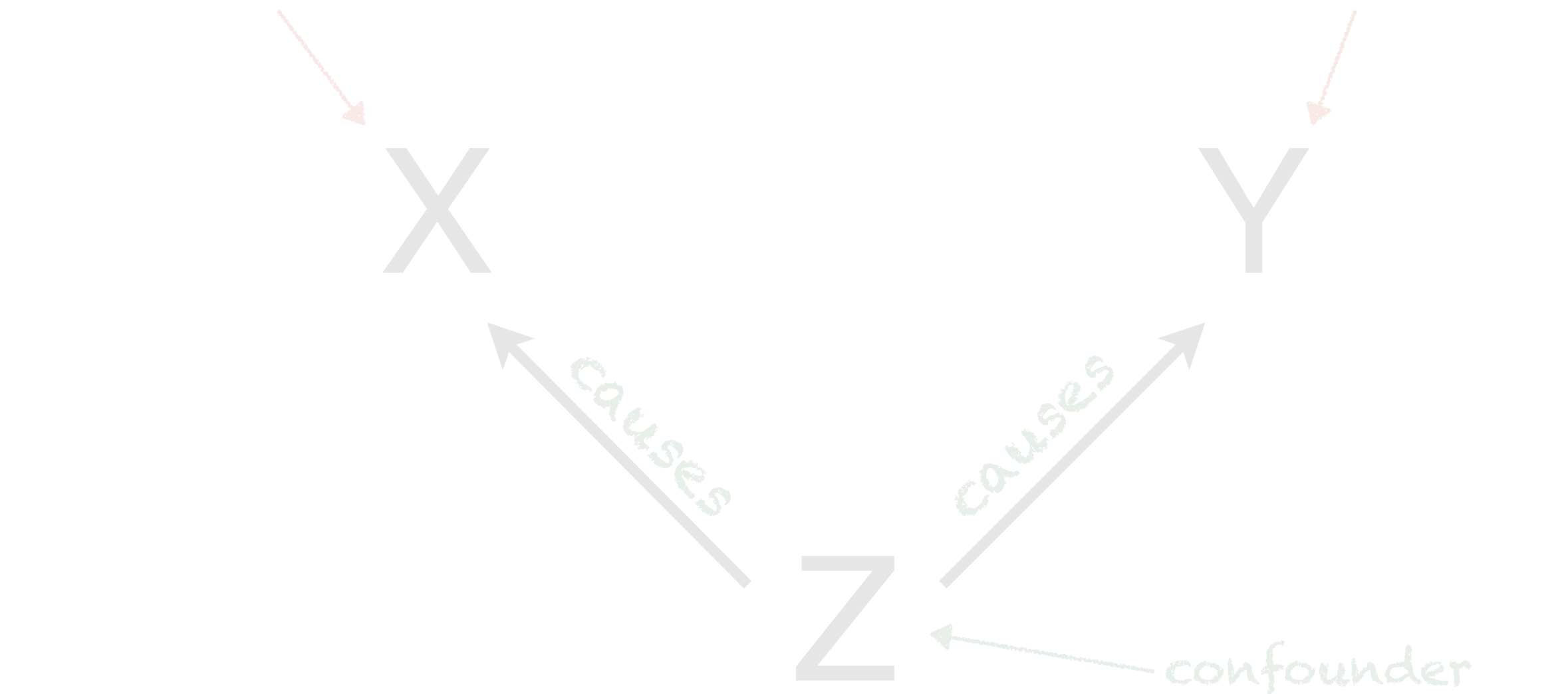
Y

causes

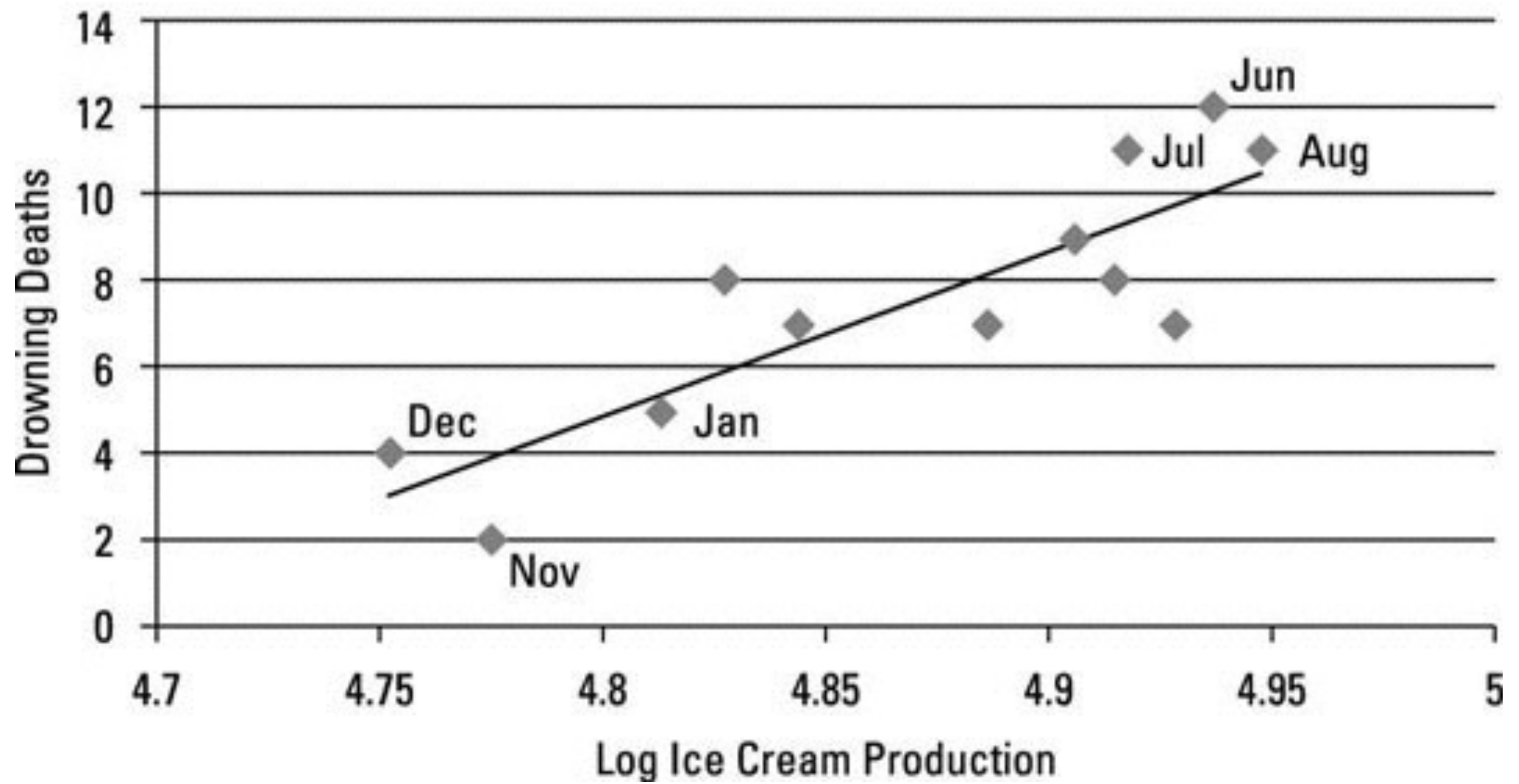
causes

Z

confounder



Ice Cream and Drowning Scatter, 2006



reverse causation

# reverse causation



# reverse causation

key explanatory variable

outcome variable

X

causes

Y



# reverse causation

key explanatory variable

outcome variable

X

causes

Y

Notice the the arrow goes  
the wrong direction!



## THE FAMILY CIRCUS



8-5

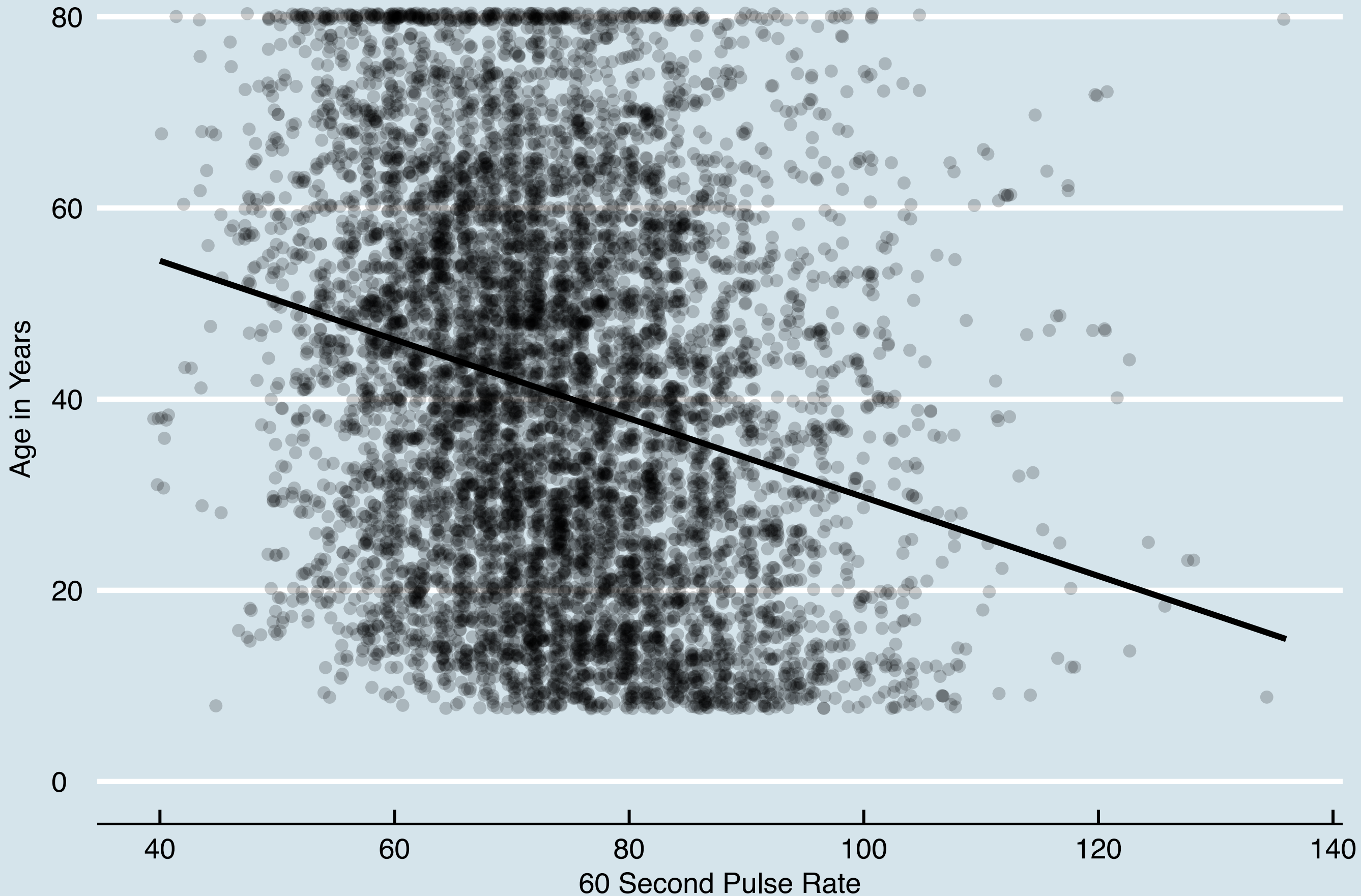
© 1988 by Disney, Inc.  
All Rights Reserved.

KEANE

"I wish they didn't turn on that seatbelt sign so much! Every time they do, it gets bumpy."



# Can Running Just Few Minutes Make You Years Younger?



chance

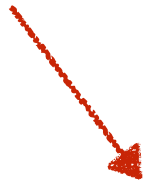
chance

**X**

**Y**

# chance

key explanatory variable



**X**

outcome variable



**Y**

Sometimes, X and Y will be correlated just by chance, even when there is no systematic relationship between the two.

chance

key explanatory variable

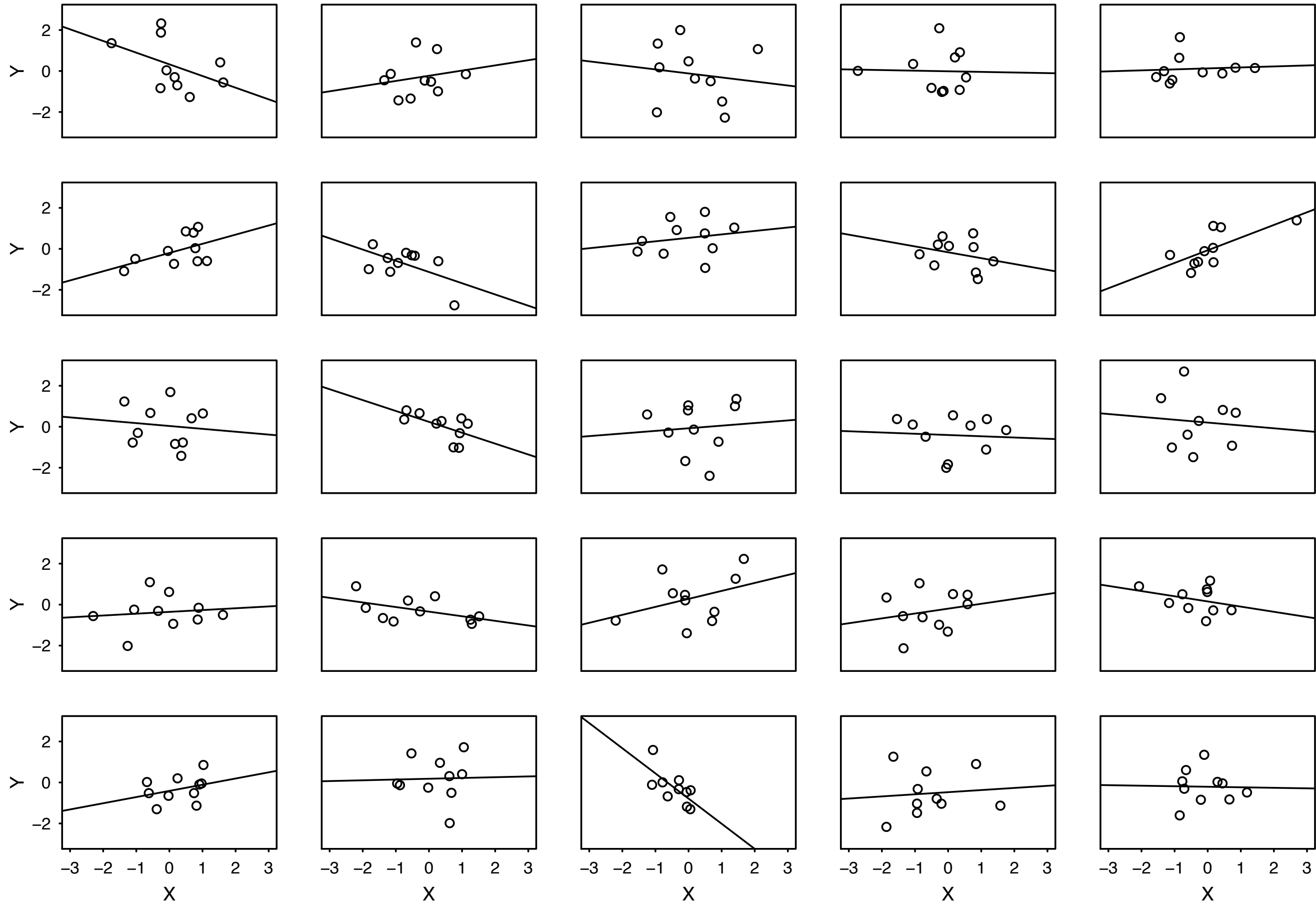
X

outcome variable

Y

Notice there is no  
causal arrow!

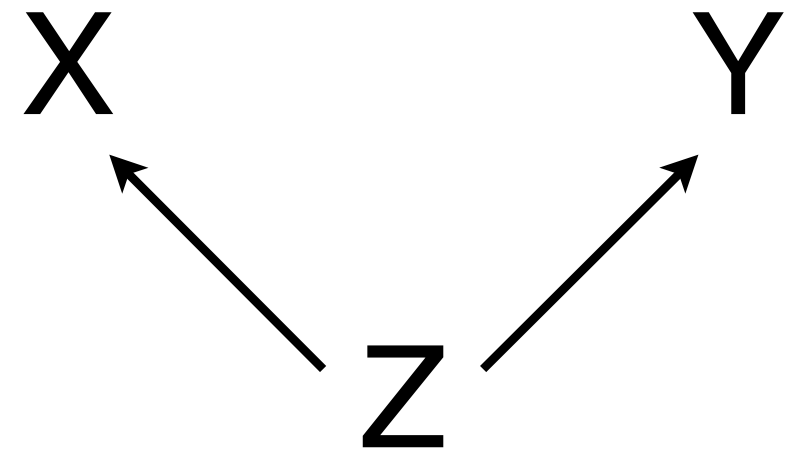
# Pure Noise Generated by a Computer



causation



spuriousness



reverse causation



chance

no systematic relationship;  
correlation simply due to chance



# Ruling Out the Alternatives

spuriousness and reverse causation

- a compelling theoretical model
- randomization

spuriousness

- controlling for confounders

chance

- box model and statistical inference

# Ruling Out the Alternatives

spuriousness and reverse causation

- a compelling theoretical model
- randomization

spuriousness

- controlling for confounders

chance

- box model and statistical inference

save for later...

# Ruling Out the Alternatives

spuriousness and reverse causation

- a compelling theoretical model
- randomization

**Discuss Today**

spuriousness

- controlling for confounders

*save for later...*

chance

- box model and statistical inference

# A Compelling Theoretical Model

# A Compelling Theoretical Model

Simply explain why spuriousness and reverse causation make little theoretical sense.


# A Compelling Theoretical Model

Simply explain why spuriousness and reverse causation make little theoretical sense.

- Could it be that some other variable causes both democratic institutions and GNI? What might this be?
- Could it be that GNI causes democratic institutions?

# Randomization


# What is the effect of a campaign mailer on a citizen's decision to turn out and vote?



\* Last Day to Register to VOTE is **MARCH 28<sup>th</sup>**  
\* PA Primary Election is **April 26<sup>th</sup>**  
\* PA General Election is **November 8, 2016**


**In 2016 YOUR Vote Will Help To Elect:**

- THE PRESIDENT OF THE UNITED STATES
- PA UNITED STATES SENATOR
- PA ATTORNEY GENERAL
- PA AUDITOR GENERAL
- PA STATE TREASURER
- PA REPRESENTATIVES IN CONGRESS (DISTRICTS 2, 6, 7, 8, 13)
- PA SENATOR IN GENERAL ASSEMBLY (DISTRICTS .., ..)
- PA REPRESENTATIVES IN THE GENERAL ASSEMBLY (DISTRICTS 26, 53, 61, 70, 131, 146, 147, 148, 149, 150, 151, 152, 153, 154, 157, 166, 172, 194)
- DELEGATE TO THE NATIONAL CONVENTION (DEMOCRATIC & REPUBLICAN)
- ALTERNATIVE TO THE NATIONAL CONVENTION (DEMOCRATIC & REPUBLICAN)



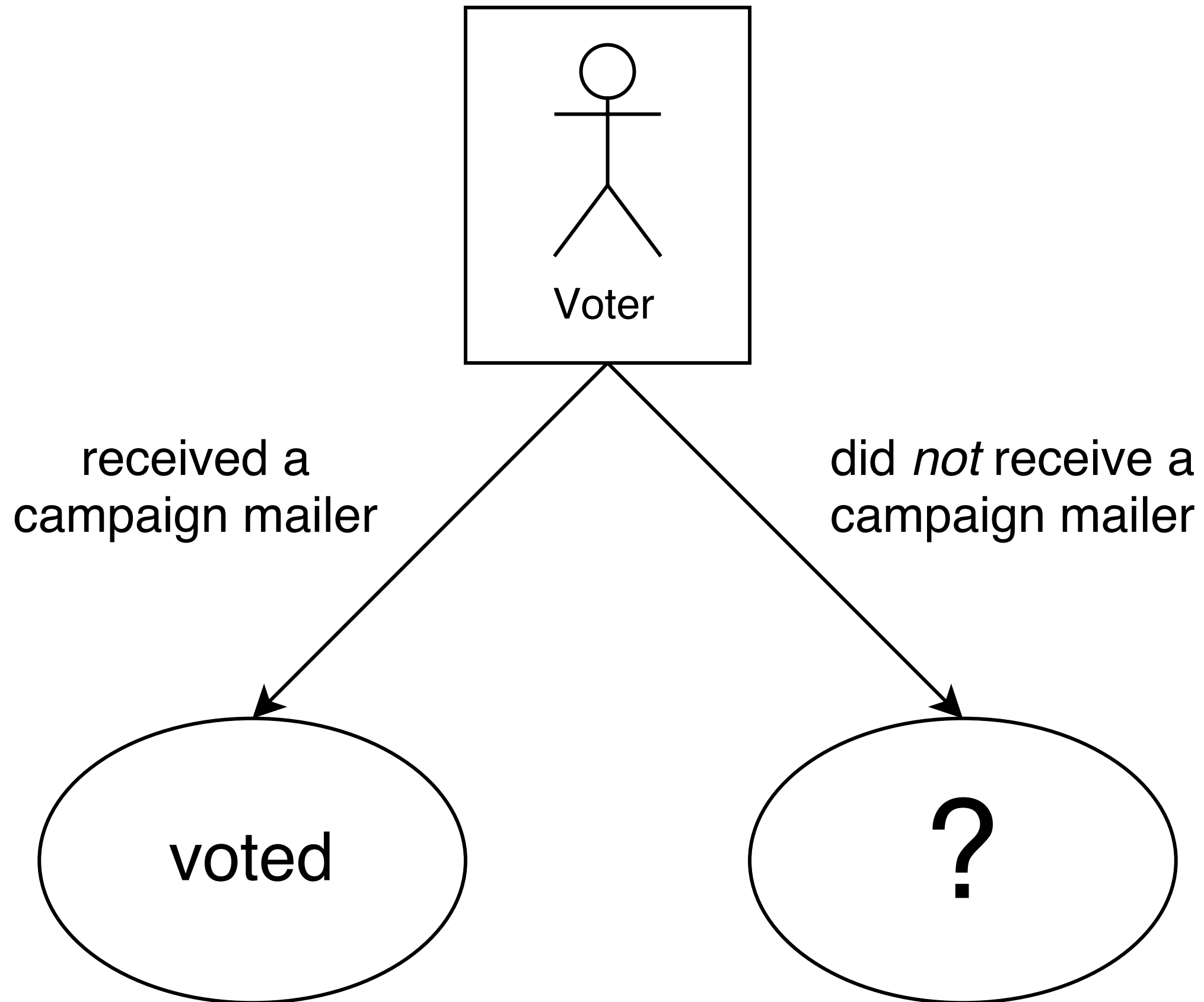
**NOT SURE WHERE TO VOTE?** Go to  
<https://www.pavoterservices.state.pa.us/Pages/PollingPlaceInfo.aspx>

**TO VIEW THE OFFICIAL CANDIDATE LIST** Go to  
<https://www.pavoterservices.state.pa.us/ElectionInfo/electioninfo.aspx>



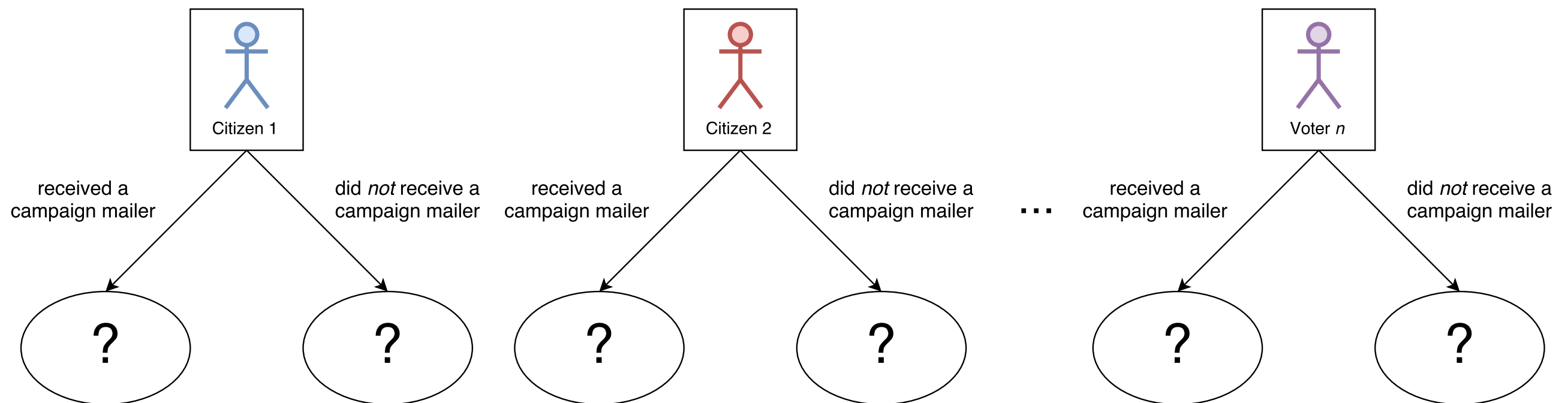
This reminder brought to you by the  
**Willow Grove Branch of the NAACP**  
Montgomery County, PA  
**Valerie Ward/Branch President**





Imagine we're in the following **ideal situation**:

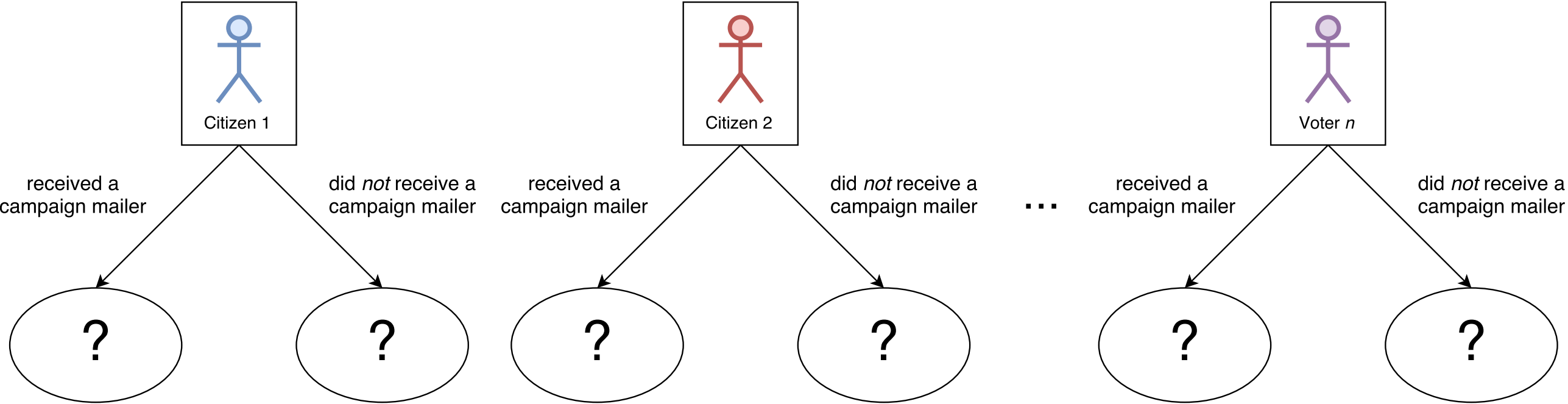
- A. we have  $n$  potential voters,
- B. the election hasn't yet happened, and
- C. we can control the assignment of the treatment.



$R_T^{hyp}$ : The hypothetical turnout Rate if everyone was in the Treatment group.

$R_C^{hyp}$ : The hypothetical turnout Rate if everyone was in the Control group.

$R_T^{hyp} - R_C^{hyp}$ : average treatment effect (ATE)



$R_T^{obs}$ : The observed turnout Rate in the Treatment group.

$R_C^{obs}$ : The observed turnout Rate in the Control group.

$$R_T^{obs} - R_C^{obs} \approx R_T^{hyp} - R_C^{hyp}$$

$R_T^{obs}$ : The observed Rate in the Treatment group.

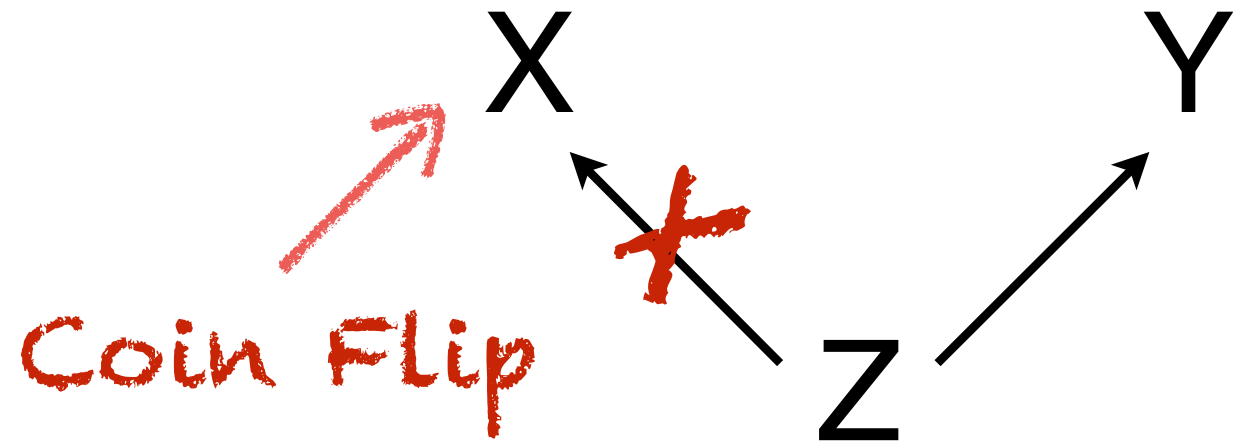
$R_C^{obs}$ : The observed Rate in the Control group.

$$\underbrace{R_T^{obs} - R_C^{obs}}_{\text{estimate}} \approx \overbrace{R_T^{hyp} - R_C^{hyp}}^{\text{ATE}}$$

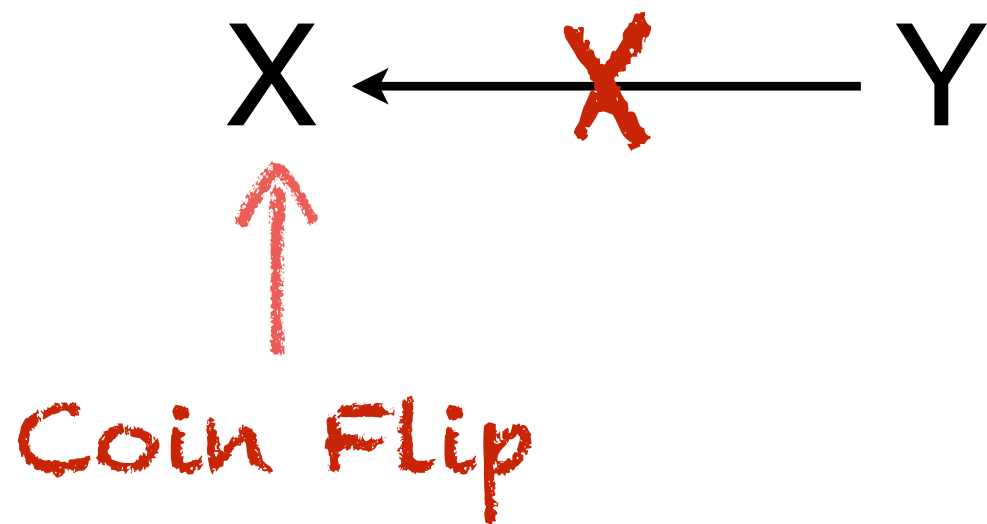
causation



spuriousness



reverse causation



chance

no systematic relationship;  
correlation simply due to chance

**TABLE 2. Effects of Four Mail Treatments on Voter Turnout in the August 2006 Primary Election**

	Experimental Group				
	Control	Civic Duty	Hawthorne	Self	Neighbors
Percentage Voting					
N of Individuals					



**TABLE 2. Effects of Four Mail Treatments on Voter Turnout in the August 2006 Primary Election**

	Experimental Group				
	Control	Civic Duty	Hawthorne	Self	Neighbors
Percentage Voting	29.7%				
N of Individuals	191,243				

Dear Registered Voter:

DO YOUR CIVIC DUTY AND VOTE!

Why do so many people fail to vote? We've been talking about this problem for years, but it only seems to get worse.

The whole point of democracy is that citizens are active participants in government; that we have a voice in government. Your voice starts with your vote. On August 8, remember your rights and responsibilities as a citizen. Remember to vote.

DO YOUR CIVIC DUTY — VOTE!

**TABLE 2. Effects of Four Mail Treatments on Voter Turnout in the August 2006 Primary Election**

	Experimental Group				
	Control	Civic Duty	Hawthorne	Self	Neighbors
Percentage Voting	29.7%	31.5%			
N of Individuals	191,243	38,218			

Dear Registered Voter:

**YOU ARE BEING STUDIED!**

Why do so many people fail to vote? We've been talking about this problem for years, but it only seems to get worse.

This year, we're trying to figure out why people do or do not vote. We'll be studying voter turnout in the August 8 primary election.

Our analysis will be based on public records, so you will not be contacted again or disturbed in any way. Anything we learn about your voting or not voting will remain confidential and will not be disclosed to anyone else.

**DO YOUR CIVIC DUTY — VOTE!**

**TABLE 2. Effects of Four Mail Treatments on Voter Turnout in the August 2006 Primary Election**

	Experimental Group				
	Control	Civic Duty	Hawthorne	Self	Neighbors
Percentage Voting	29.7%	31.5%	32.2%		
N of Individuals	191,243	38,218	38,204		

Dear Registered Voter:

## WHO VOTES IS PUBLIC INFORMATION!

Why do so many people fail to vote? We've been talking about the problem for years, but it only seems to get worse.

This year, we're taking a different approach. We are reminding people that who votes is a matter of public record.

The chart shows your name from the list of registered voters, showing past votes, as well as an empty box which we will fill in to show whether you vote in the August 8 primary election. We intend to mail you an updated chart when we have that information.

We will leave the box blank if you do not vote.

DO YOUR CIVIC DUTY—VOTE!

OAK ST  
9999 ROBERT WAYNE  
9999 LAURA WAYNE

Aug 04	Nov 04	Aug 06
	Voted	_____
Voted	Voted	_____

**TABLE 2. Effects of Four Mail Treatments on Voter Turnout in the August 2006 Primary Election**

	Experimental Group				
	Control	Civic Duty	Hawthorne	Self	Neighbors
Percentage Voting	29.7%	31.5%	32.2%	34.5%	
N of Individuals	191,243	38,218	38,204	38,218	

Dear Registered Voter:

## WHAT IF YOUR NEIGHBORS KNEW WHETHER YOU VOTED?

Why do so many people fail to vote? We've been talking about the problem for years, but it only seems to get worse. This year, we're taking a new approach. We're sending this mailing to you and your neighbors to publicize who does and does not vote.

The chart shows the names of some of your neighbors, showing which have voted in the past. After the August 8 election, we intend to mail an updated chart. You and your neighbors will all know who voted and who did not.

## DO YOUR CIVIC DUTY — VOTE!

MAPLE DR	Aug 04	Nov 04	Aug 06
9995 JOSEPH JAMES SMITH	Voted	Voted	_____
9995 JENNIFER KAY SMITH		Voted	_____
9997 RICHARD B JACKSON		Voted	_____
9999 KATHY MARIE JACKSON		Voted	_____
9999 BRIAN JOSEPH JACKSON		Voted	_____
9991 JENNIFER KAY THOMPSON		Voted	_____
9991 BOB R THOMPSON		Voted	_____
9993 BILL S SMITH			_____
9989 WILLIAM LUKE CASPER		Voted	_____
9989 JENNIFER SUE CASPER		Voted	_____
9987 MARIA S JOHNSON	Voted	Voted	_____
9987 TOM JACK JOHNSON	Voted	Voted	_____
9987 RICHARD TOM JOHNSON		Voted	_____
9985 ROSEMARY S SUE		Voted	_____
9985 KATHRYN L SUE		Voted	_____
9985 HOWARD BEN SUE		Voted	_____
9983 NATHAN CHAD BERG		Voted	_____
9983 CARRIE ANN BERG		Voted	_____
9981 EARL JOEL SMITH			_____
9979 DEBORAH KAY WAYNE		Voted	_____
9979 JOEL R WAYNE		Voted	_____



**TABLE 2. Effects of Four Mail Treatments on Voter Turnout in the August 2006 Primary Election**

	Experimental Group				
	Control	Civic Duty	Hawthorne	Self	Neighbors
Percentage Voting	29.7%	31.5%	32.2%	34.5%	37.8%
N of Individuals	191,243	38,218	38,204	38,218	38,201