Placebo Tests for Causal Inference

(based on joint work with Allan Dafoe and Guadalupe Tunon)

Andy Eggers

5 February 2021

Orientation

A placebo test from epidemiology

The NEW ENGLAND JOURNAL of MEDICINE

SPECIAL ARTICLE

The Spread of Obesity in a Large Social Network over 32 Years

Nicholas A. Christakis, M.D., Ph.D., M.P.H., and James H. Fowler, Ph.D.

ABSTRACT

BACKGROUND

The prevalence of obesity has increased substantially over the past 30 years. We performed a quantitative analysis of the nature and extent of the person-to-person spread of obesity as a possible factor contributing to the obesity epidemic.

A placebo test from epidemiology

Detecting implausible social network effects in acne, height, and headaches: longitudinal analysis

Ethan Cohen-Cole,[1] Jason M Fletcher[2]

ABSTRACT

Objective To investigate whether "network effects" can be detected for health outcomes that are unlikely to be subject to network phenomena.

Design Statistical analysis common in network studies, such as logistic regression analysis, controlled for own and friend's lagged health status. Analyses controlled for environmental confounders.

Setting Subsamples of the National Longitudinal Study of Adolescent Health (Add Health).

Participants 4300 to 5400 male and female adolescents who nominated a friend in the dataset and who were both longitudinally surveyed.

Measurements Health outcomes, including headache severity, acne severity, and height self reported by respondents in 1994-5, 1995-6, and 2000-1.

Results Significant network effects were observed in the acquisition of acne, headaches, and height. A friend's acne problems increased an individual's odds of acne problems (odds ratio 1.47, 95% confidence interval 0.93 to 2.33). The likelihood that an individual had headaches also increased with the presence of a friend with headaches (1.62, 0.91 to 2.89); and an individual's height increased by 20% of his/

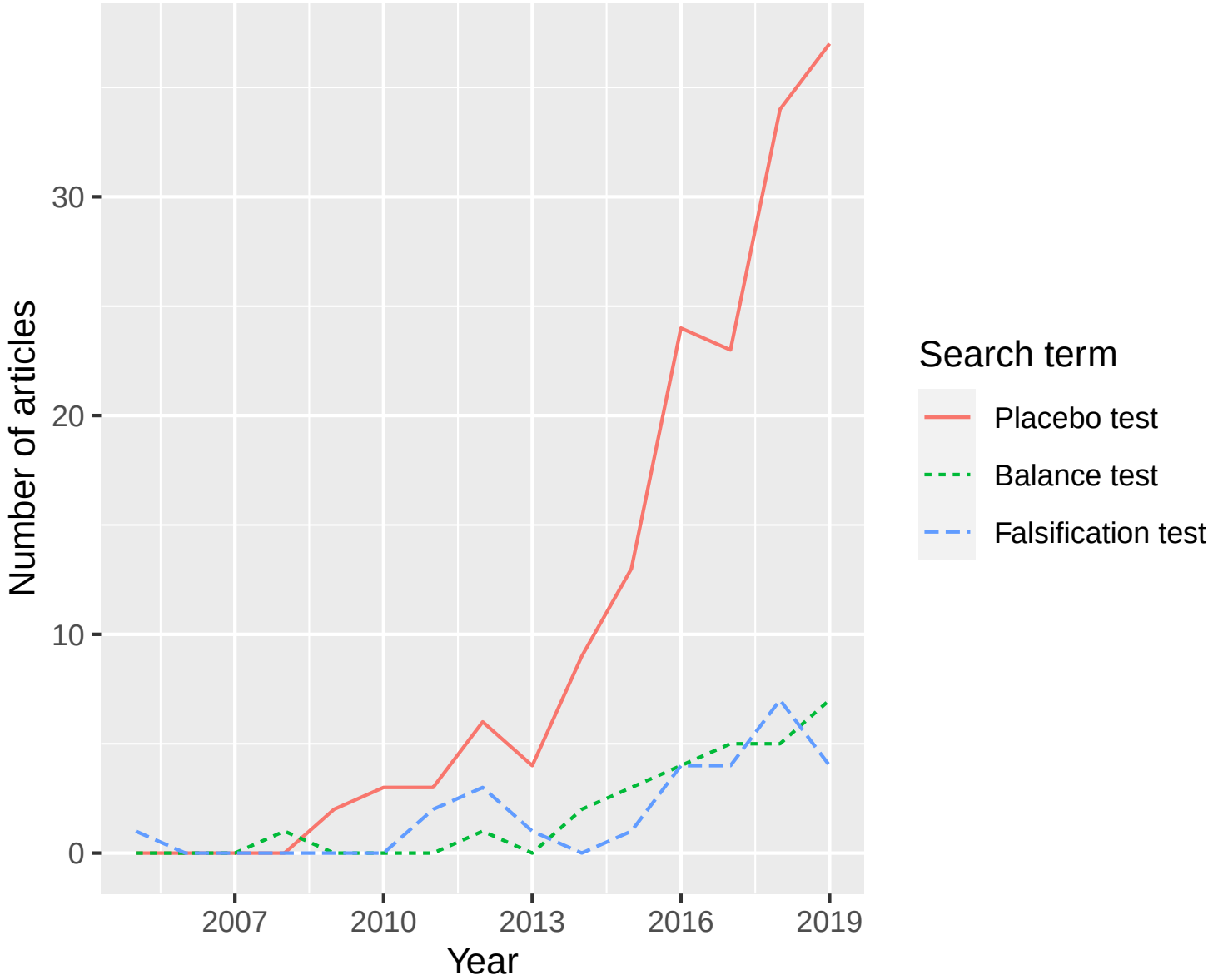
Conclusions Researchers should be cautious in attributing correlations in health outcomes of close friends to social network effects, especially when environmental confounders are not adequately controlled for in the analysis.

INTRODUCTION

Providing credible estimates of the effects of social networks in choices and outcomes in health is important for suggesting policies that could improve health via social networks. For example, Christakis and Fowler have presented evidence of the person to person spread of obesity and quitting smoking among friends.[1 2] Raspe et al proposed that back pain might be a "communicable disease."[3]

Many methods used to estimate social network effects are subject to potentially large biases that result in the increased likelihood of detecting social network effects where none exists. For example, the use of standard econometric methods on peer effects substantially reduces evidence of social network effects in obesity.[4] Previous work that claimed to find social contagion in the diffusion of prescription drugs was confounded by

Prevalence of placebo tests in political science



Some questions I hope you have

- What is a placebo test?
- How can I make one?
- Why would I do a balance test rather than control for something?
- How could you test an IV exclusion restriction with a placebo test?
- What are fake-cutoff placebo tests in RDDs good for?

Placebo tests: basics

Placebo tests are

- a type of **auxiliary analysis** (along with subgroup analysis, robustness tests, sensitivity analysis)
- mainly tests for **bias** in estimates of (causal) effects (cf randomization inference)
- a type of **theory testing**: we locate an association that should be
  - **absent** if the causal interpretation of the main finding is correct
  - **present** if some other interpretation of the main finding (e.g. confounding bias) is correct

Typology

Placebo tests minimally alter the **core analysis**.

Typology

Placebo tests minimally alter the **core analysis**.

They differ in which aspect is altered:

Altered component	Resulting test
Population/sample	Placebo population test
Outcome variable	Placebo outcome test
Treatment variable	Placebo treatment test

Typology

Placebo tests minimally alter the **core analysis**.

They differ in which aspect is altered:

Altered component	Resulting test
Population/sample	Placebo population test
Outcome variable	Placebo outcome test
Treatment variable	Placebo treatment test

Common logic: The alteration shuts down the purported treatment effect while retaining the postulated flaw in the core analysis.

Criteria for evaluating a placebo test for bias

An **informative** placebo test is capable of shaping our beliefs about bias in the core analysis.

Criteria for evaluating a placebo test for bias

An **informative** placebo test is capable of shaping our beliefs about bias in the core analysis.

In NHST framework, this is true if it rejects the null with

- **high** probability when the core analysis is **biased** (high power)
- **low** probability when the core analysis is **unbiased** (low false positive rate).

Criteria for evaluating a placebo test for bias

An **informative** placebo test is capable of shaping our beliefs about bias in the core analysis.

In NHST framework, this is true if it rejects the null with

- **high** probability when the core analysis is **biased** (high power)
- **low** probability when the core analysis is **unbiased** (low false positive rate).

As with any estimate,

$$\text{Estimate in placebo test} \quad = \quad \text{effect} + \text{bias} + \text{noise}$$

Criteria for evaluating a placebo test for bias

An **informative** placebo test is capable of shaping our beliefs about bias in the core analysis.

In NHST framework, this is true if it rejects the null with

- **high** probability when the core analysis is **biased** (high power)
- **low** probability when the core analysis is **unbiased** (low false positive rate).

As with any estimate,

$$\text{Estimate in placebo test} \quad = \quad \text{effect} + \text{bias} + \text{noise}$$

Ideally, effect is zero, bias mirrors bias in core analysis, noise is small.

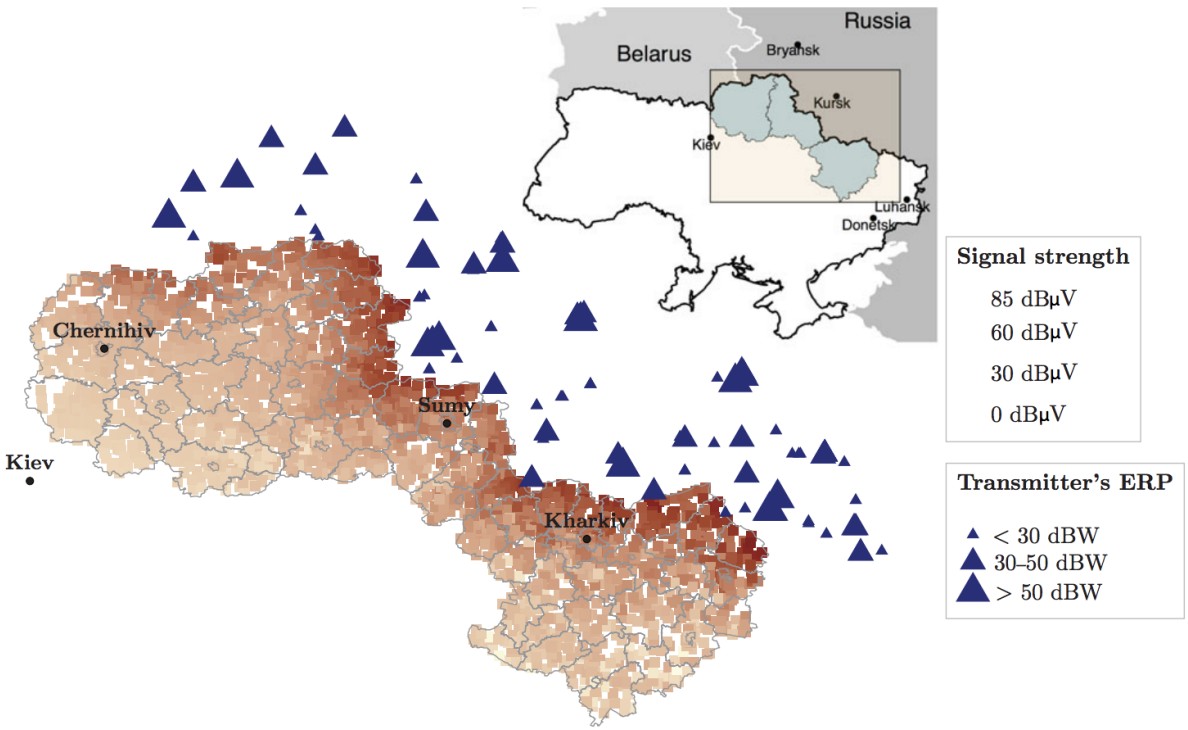
Running example: Peisakhin & Rozenas (2018)

How did pro-Russian TV affect voting/attitudes in Ukraine's 2014 election?

Running example: Peisakhin & Rozenas (2018)

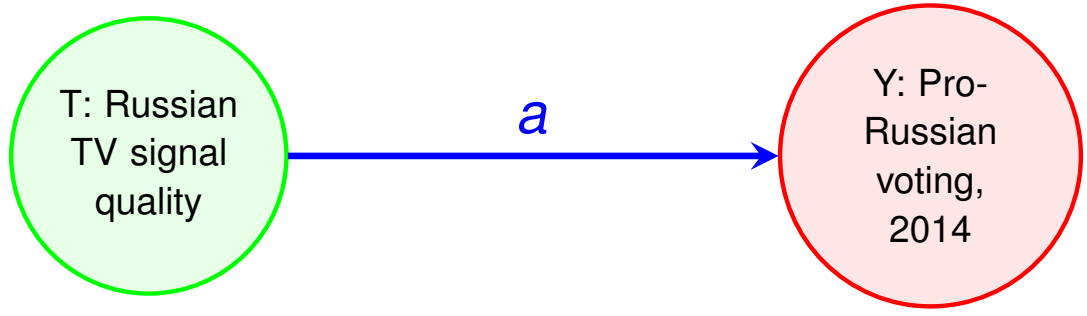
How did pro-Russian TV affect voting/attitudes in Ukraine's 2014 election?

FIGURE 2 Russian Analog TV Transmitters (Triangles) and Predicted Signal Strength at Polling Stations

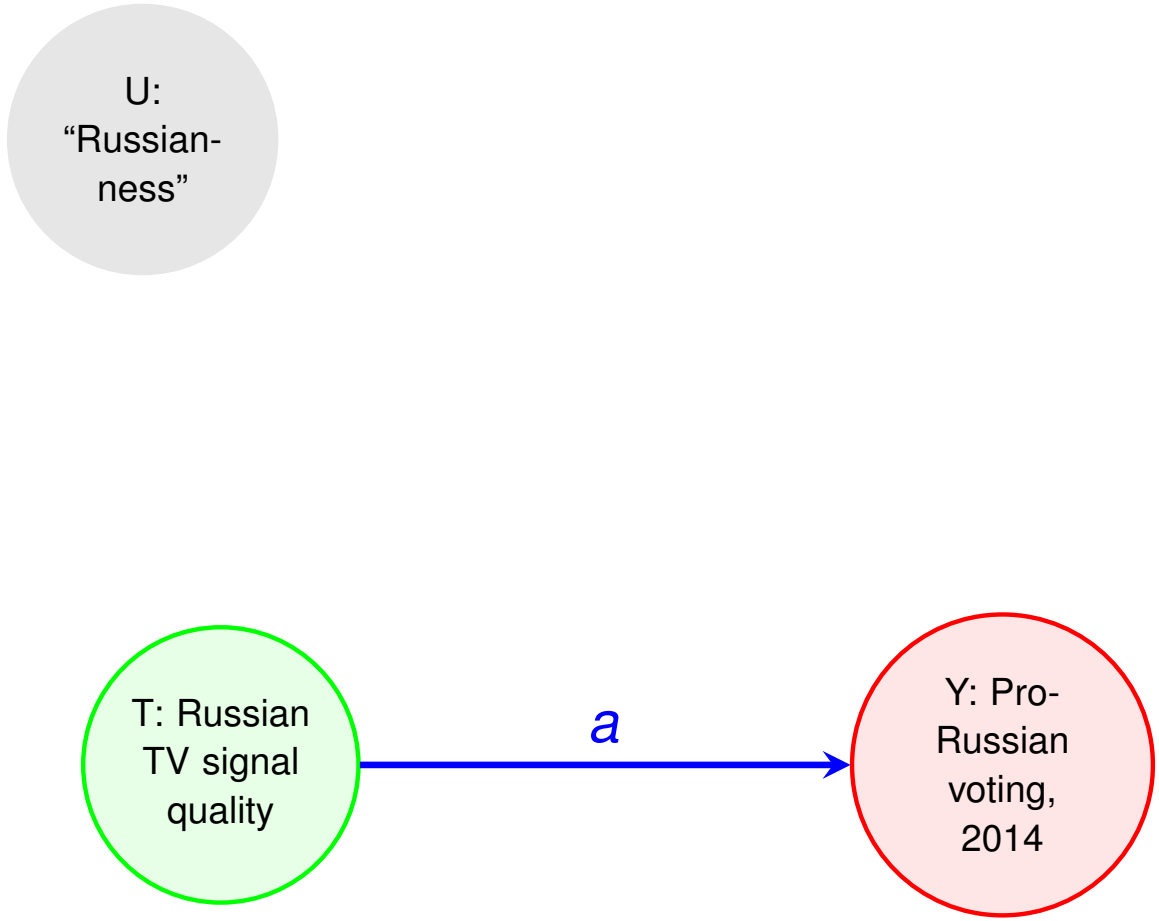


Note: Gray lines are for county borders. ERP = effective radiated power.

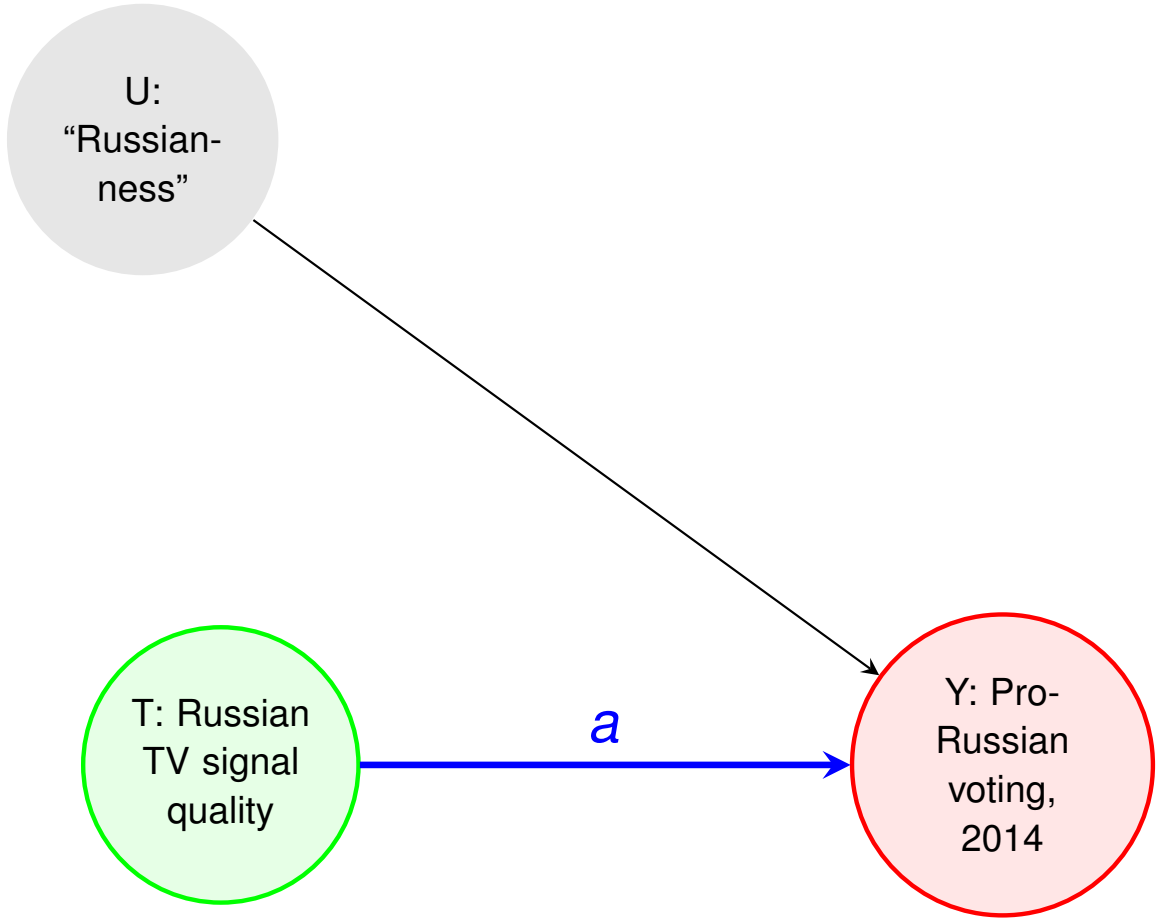
Directed acyclic graph (DAG)



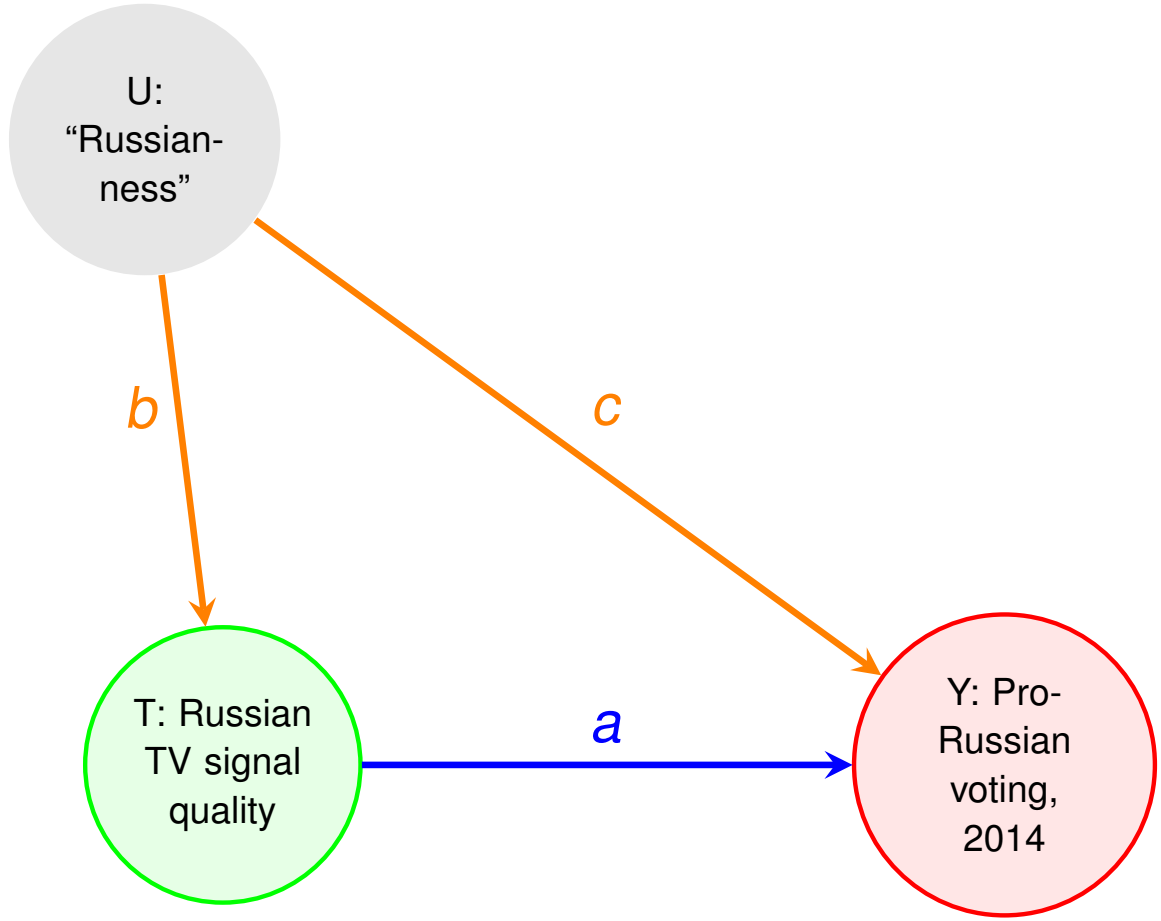
Directed acyclic graph (DAG)



Directed acyclic graph (DAG)



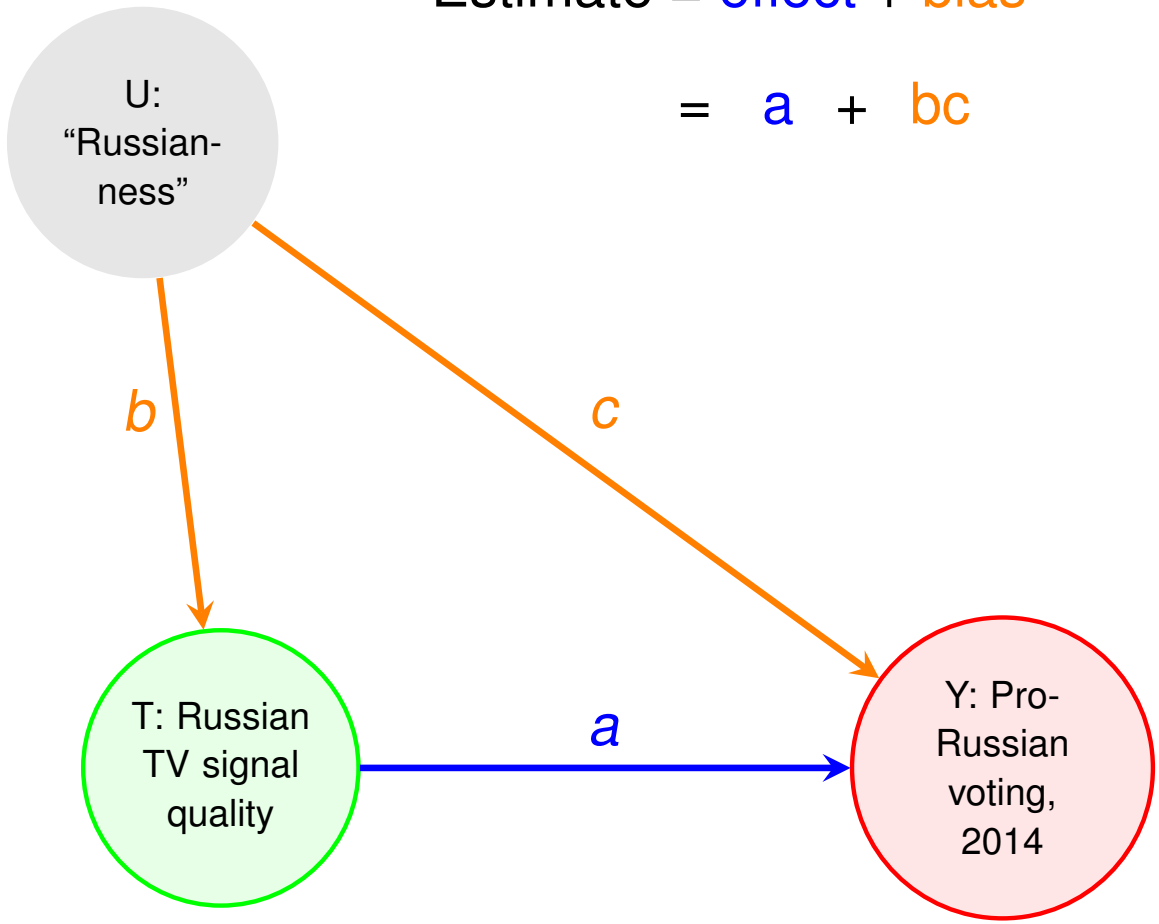
Directed acyclic graph (DAG)



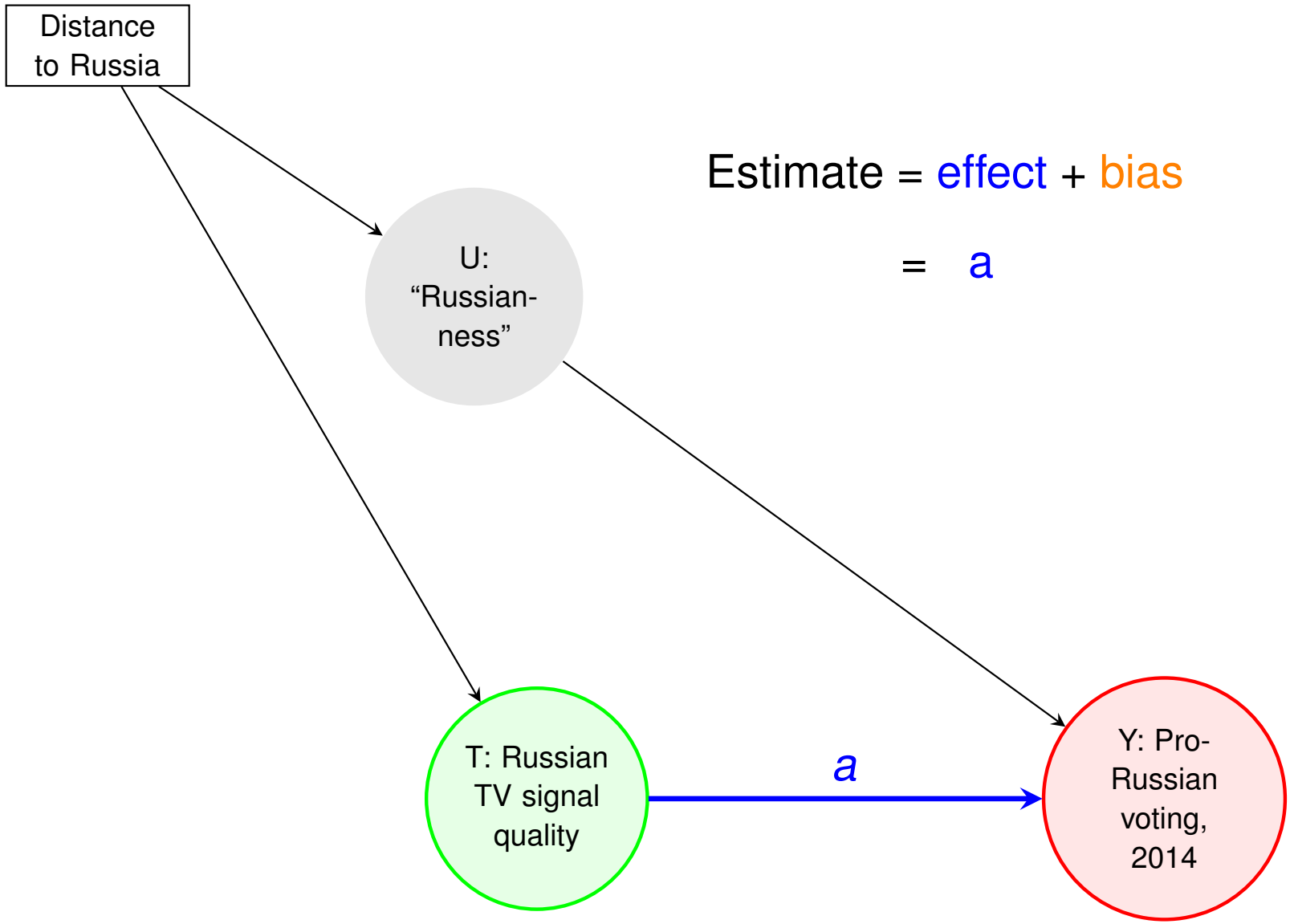
Directed acyclic graph (DAG)

Estimate = effect + bias

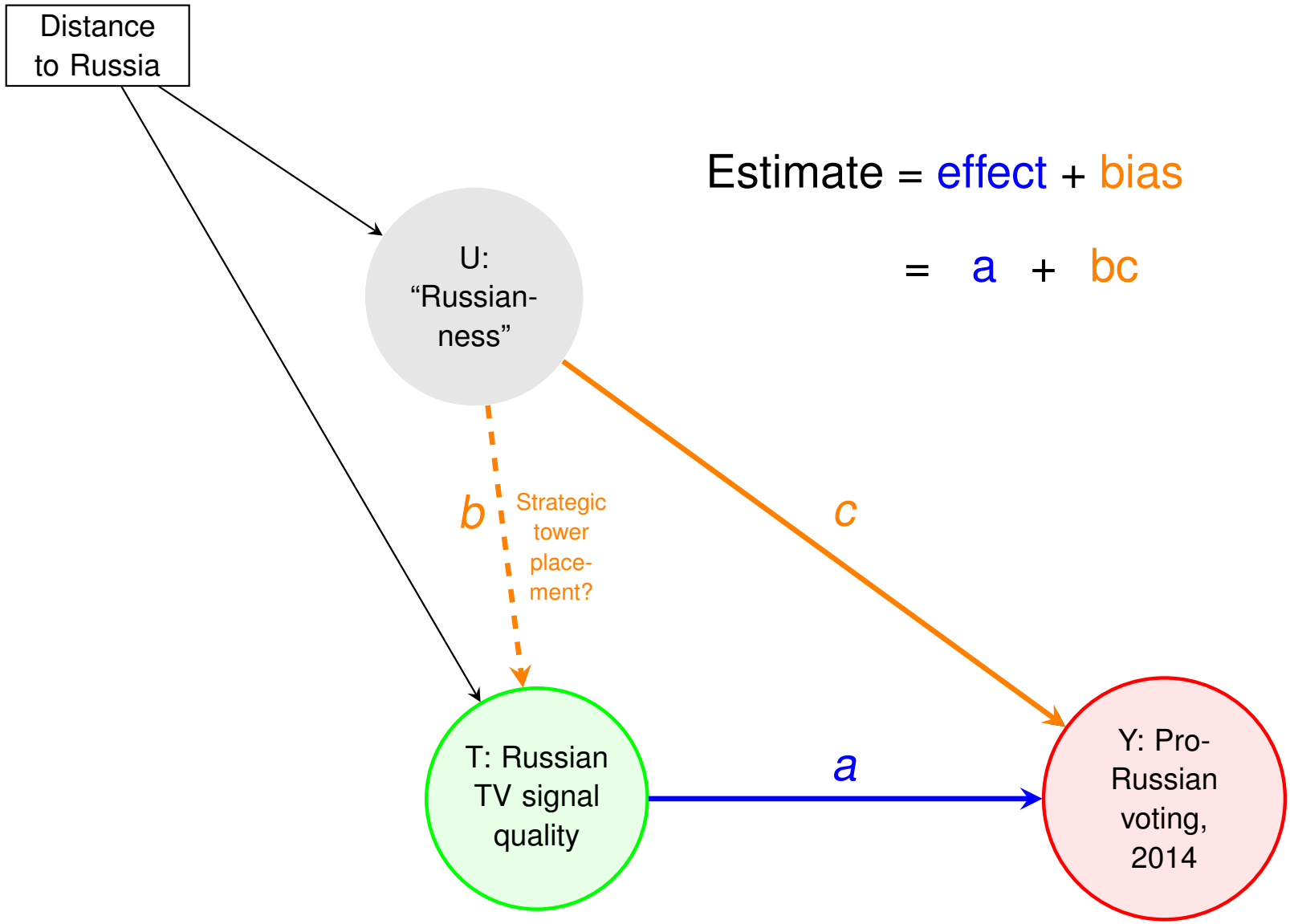
= a + bc



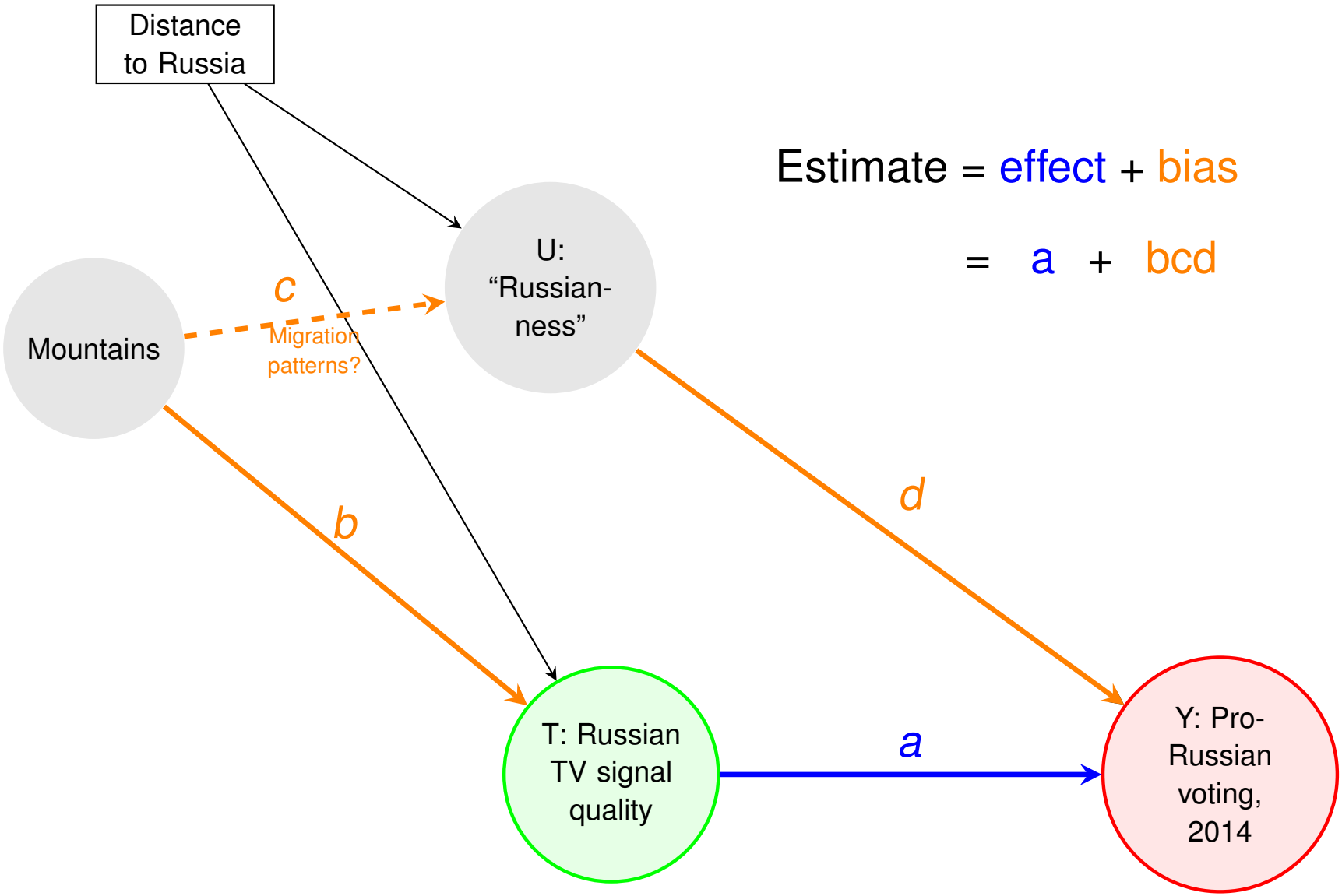
Directed acyclic graph (DAG)



Directed acyclic graph (DAG)



Directed acyclic graph (DAG)



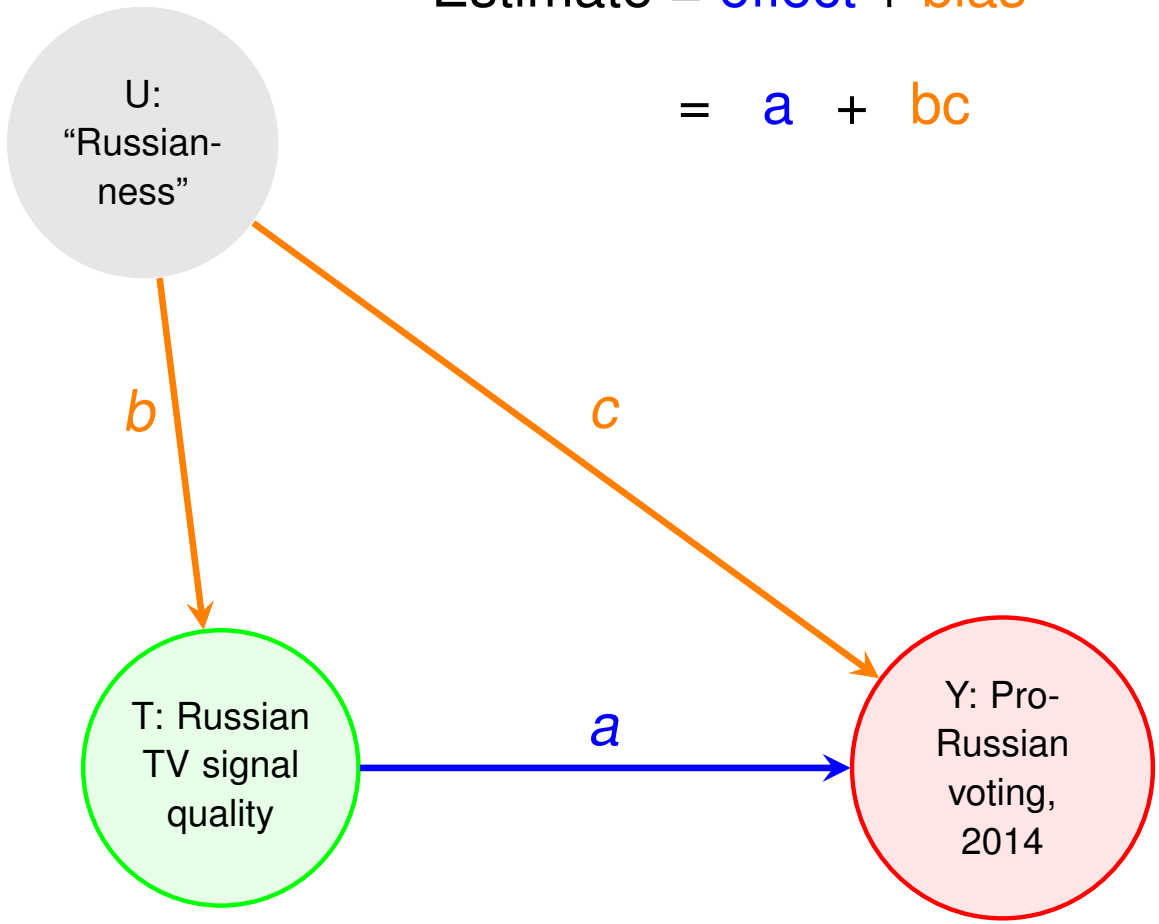
Estimate = effect + bias

= a + bcd

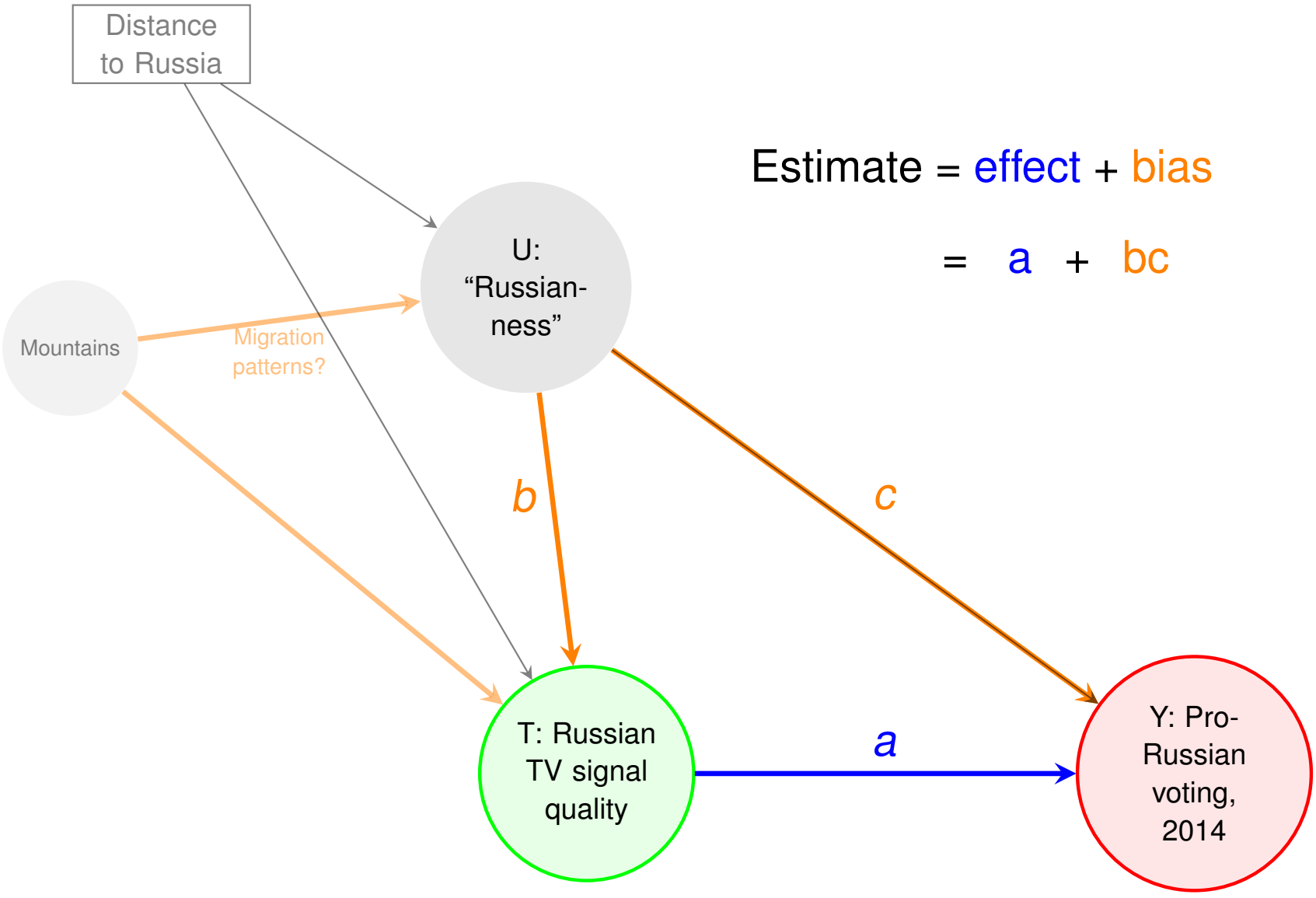
Directed acyclic graph (DAG)

Estimate = effect + bias

= a + bc



Directed acyclic graph (DAG)



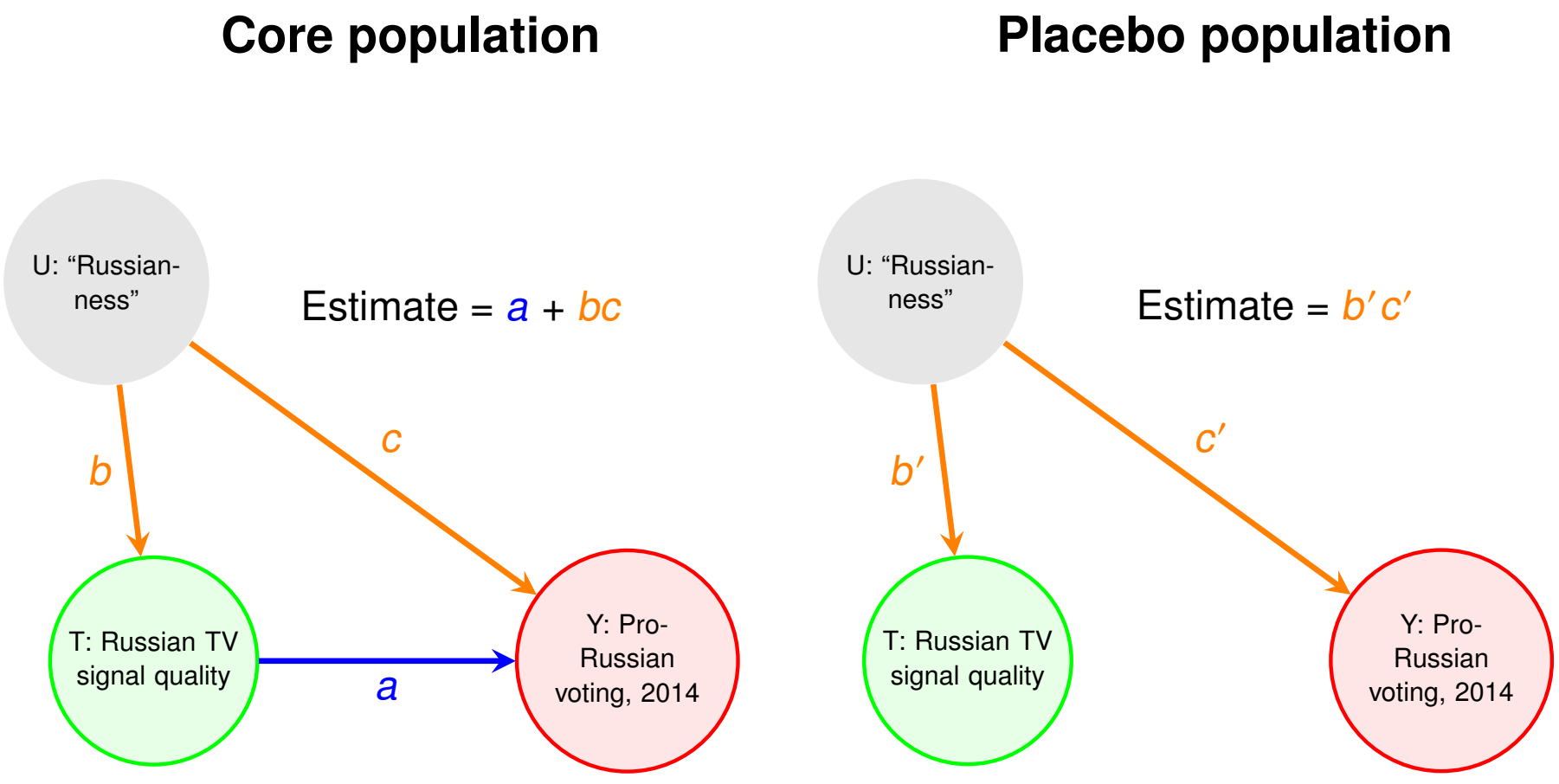
Estimate = effect + bias

= a + bc

Placebo population test

Basic idea: Reproduce core analysis in a sample/population in which

- postulated flaw would operate in a similar way
- treatment effect does not operate

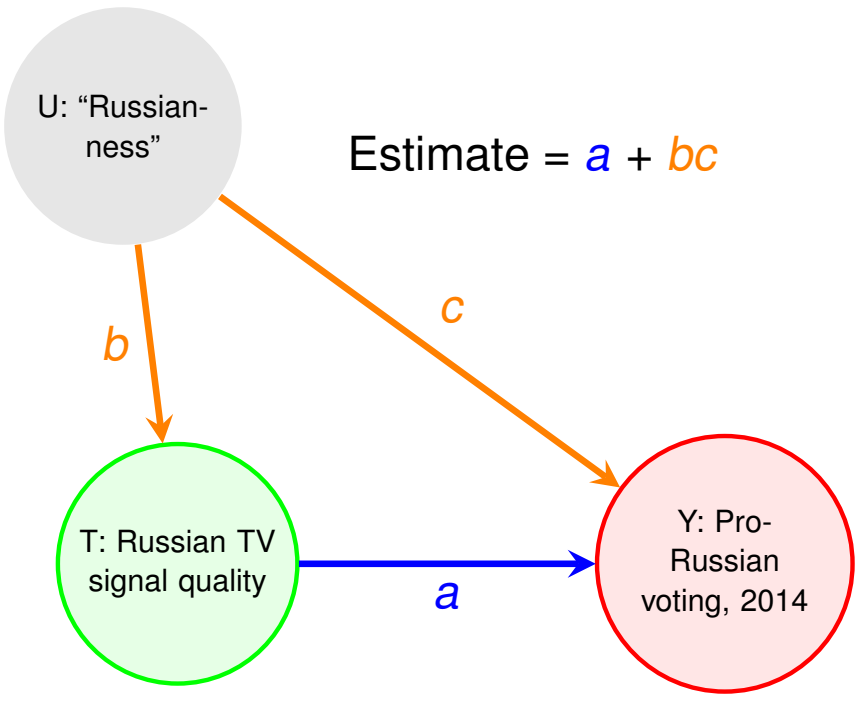


Placebo population test

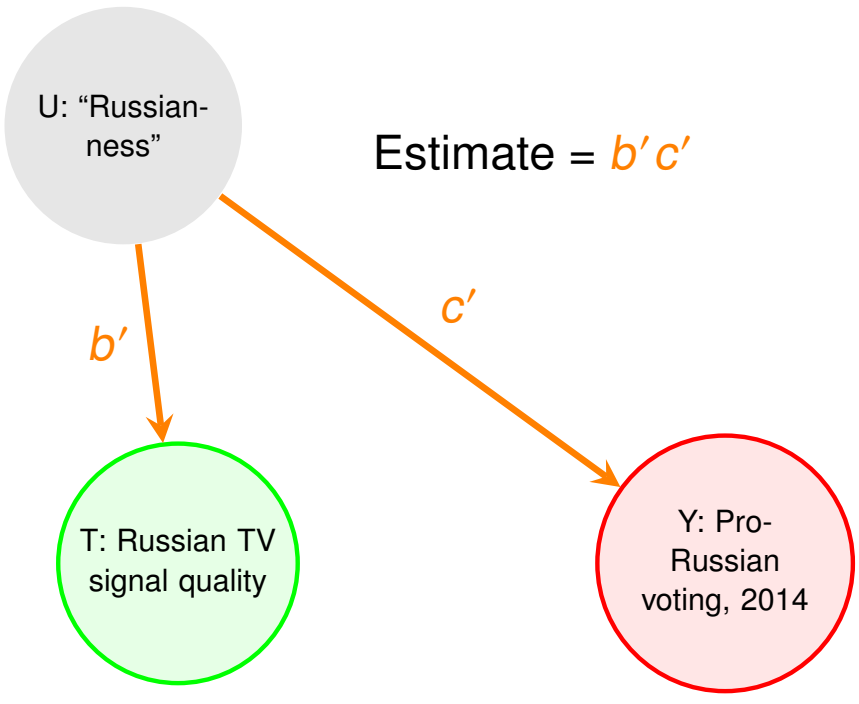
Basic idea: Reproduce core analysis in a sample/population in which

- postulated flaw would operate in a similar way
- treatment effect does not operate

Core population
owners of terrestrial TVs

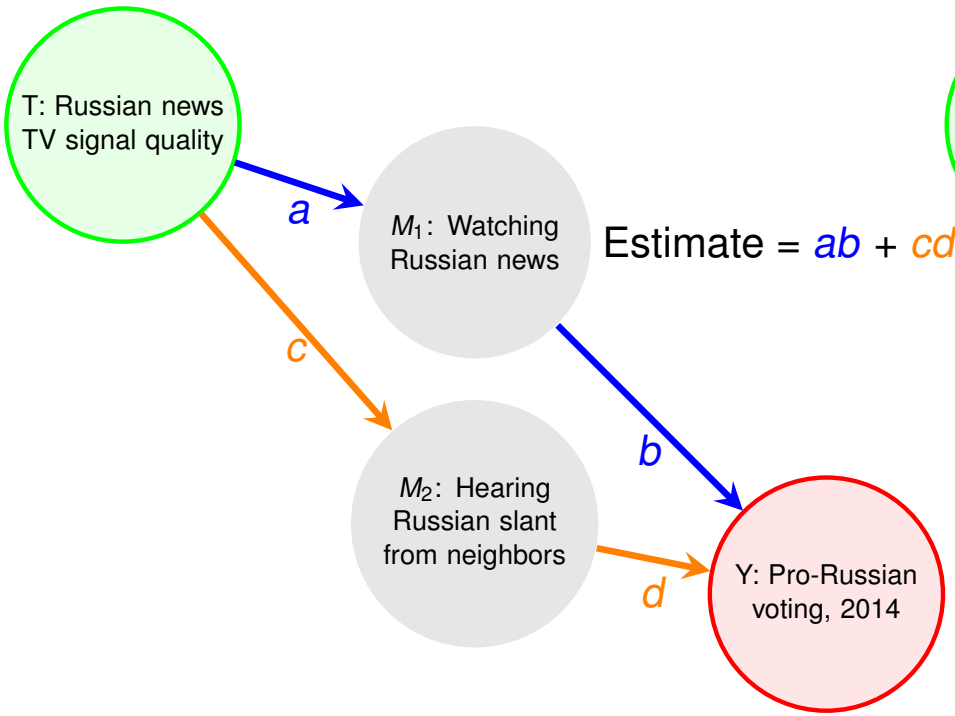


Placebo population
owners of satellite TVs

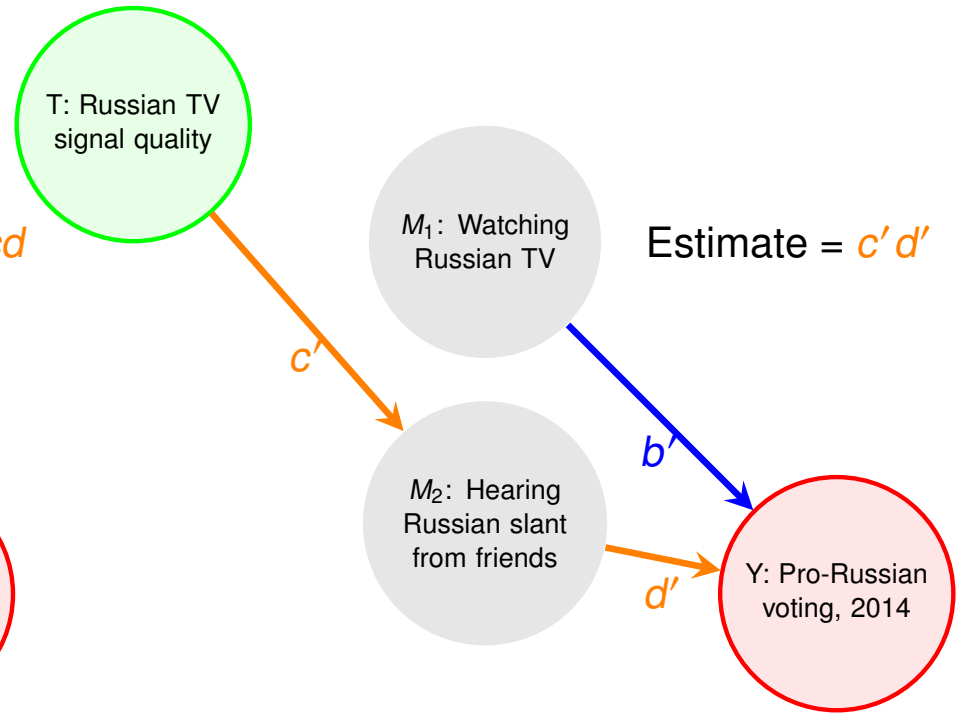


As a test of alternative mechanisms

Core population:
owners of terrestrial TVs



Placebo population:
owners of satellite TVs



Other examples

Paper	Core analysis			Placebo population
	Population	Treatment	Outcome	
Acharya, Blackwell, Sen (2016)	Americans living in the U.S. South	County's suitability for cotton production	Attitudes towards African-Americans today	Americans living in the U.S. North
Erikson & Stoker (2011)	Draft-eligible, college-bound men	Lottery draft number in 1969	Attitude toward Vietnam War in 1973	Non-college bound men; college-bound women
Fowler & Hall (2017) vis-a-vis Achen & Bartels (2016)	Counties in New Jersey in 1916	Having coastal shoreline	Vote for Democratic pres. candidate	Counties in state-years with no shark attacks

Placebo population tests: key considerations

- Would relevant bias operate similarly in placebo population?
- Is there enough power to detect bias? (sample size, variation in treatment, etc)

Placebo population tests: key considerations

- Would relevant bias operate similarly in placebo population?
- Is there enough power to detect bias? (sample size, variation in treatment, etc)

Requires broader scope conditions for theory of bias than for theory of treatment effect.

Placebo outcome test

Two types of placebo outcomes:

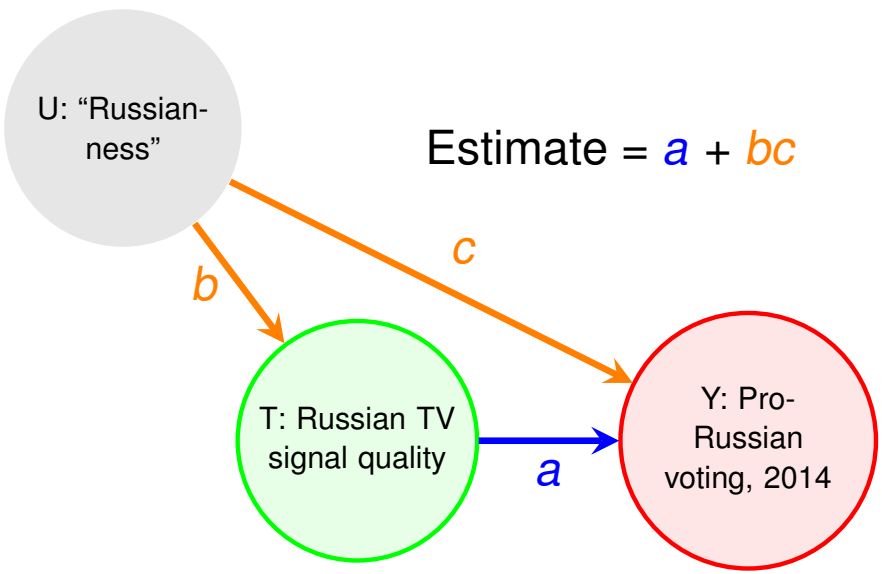
- Pre-treatment ($\rightarrow$ “balance test”)
- Post-treatment

Placebo outcome test

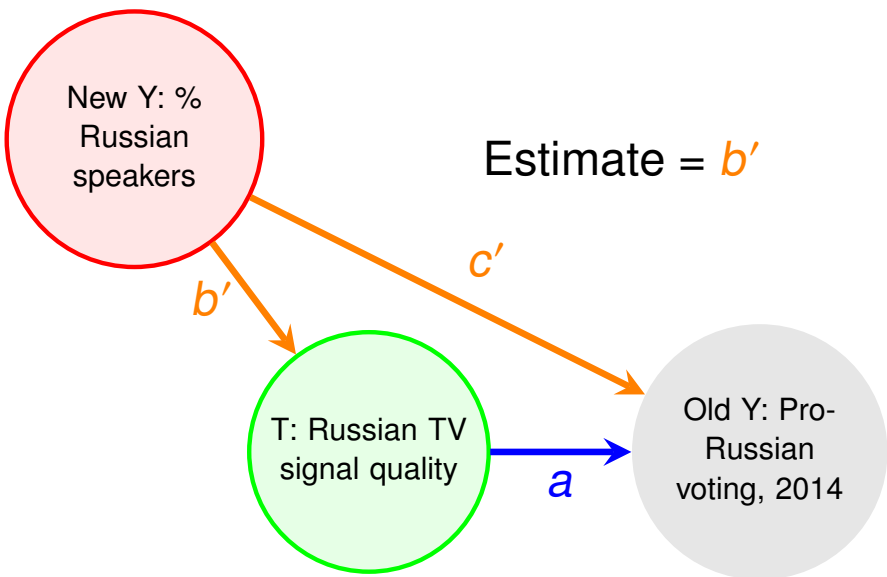
Two types of placebo outcomes:

- Pre-treatment (→ "balance test")
- Post-treatment

Core analysis:
possible confounding



Placebo outcome test
using pre-treatment variable



Placebo outcome or control variable?

Why use potential confounders $Z_1, Z_2, \ldots$ as placebo outcomes rather than adding to controls $\mathbf{X}$?

Placebo outcome or control variable?

Why use potential confounders $Z_1, Z_2, \ldots$ as placebo outcomes rather than adding to controls $\mathbf{X}$?

Recall omitted variable bias (OVB) formula says:

Estimate if Z is omitted from regression	-	Estimate if Z is included in regression	=	**Impact**	$\times$	**Imbalance**

where

- **Impact** is the coefficient on Z when included in the regression
- **Imbalance** is the result of the placebo outcome test for Z

Placebo outcome or control variable?

Why use potential confounders $Z_1, Z_2, \ldots$ as placebo outcomes rather than adding to controls $\mathbf{X}$?

Recall omitted variable bias (OVB) formula says:

$$\begin{array}{ccccccc} \text{Estimate if } Z \text{ is} \\ \text{omitted from} \\ \text{regression} & - & \begin{array}{c}\text{Estimate if } Z \text{ is} \\ \text{included in} \\ \text{regression}\end{array} & = & \textbf{Impact} & \times & \textbf{Imbalance} \end{array}$$

where

- **Impact** is the coefficient on Z when included in the regression
- **Imbalance** is the result of the placebo outcome test for Z

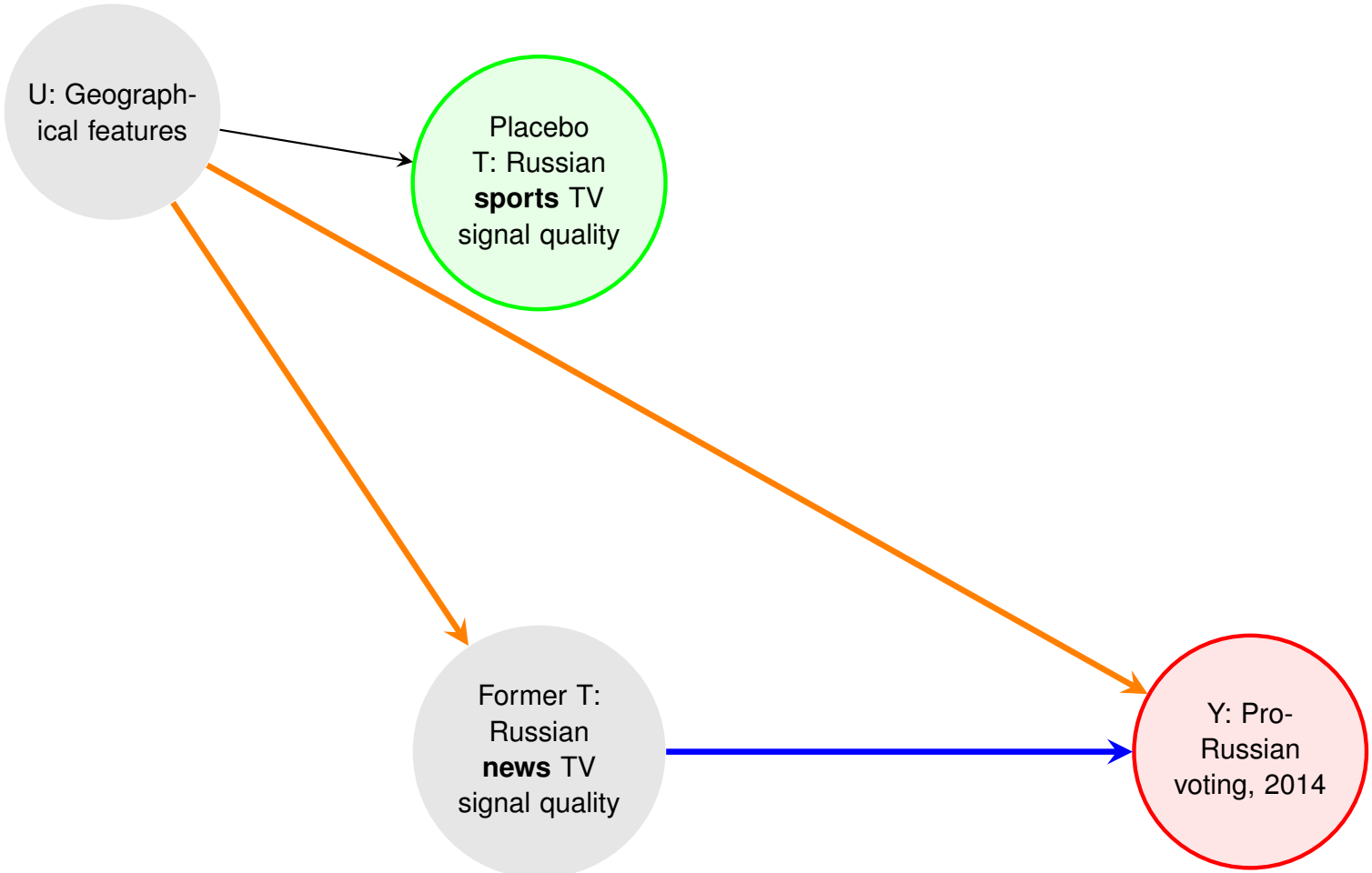
$\Longrightarrow$ using Z as placebo outcome checks **imbalance**; including Z in regression checks **imbalance** $\times$ **impact**.

Examples of post-treatment placebo outcomes

Paper	**Core analysis**			**Placebo outcome**
	Population	Treatment	Outcome	
Cruz & Schneider (2017)	610 Philippines municipalities	Whether or not the municipality participated in the KALAHI aid program	Number of visits to the municipality by local officials	Number of visits to the municipality by midwives
Hainmueller & Hangartner (2015)	1,400 municipalities in Switzerland, 1991-2009	Whether naturalization decisions are made by popular vote	Rate of naturalization through ordinary process	Rate of naturalization through marriage
Dube, Dube & Garcia-Ponce (2013)	Mexican municipalities located close to U.S. border, 2002-2006	Assault weapon availability from neighboring US state (federal ban expires in 2004 but does not affect CA)	Gun-related homicides	Accidents, non-gun homicides, and suicides

Placebo treatment test

Basic idea: Replace treatment with a descendant of a possible confounder (that would not affect the outcome).

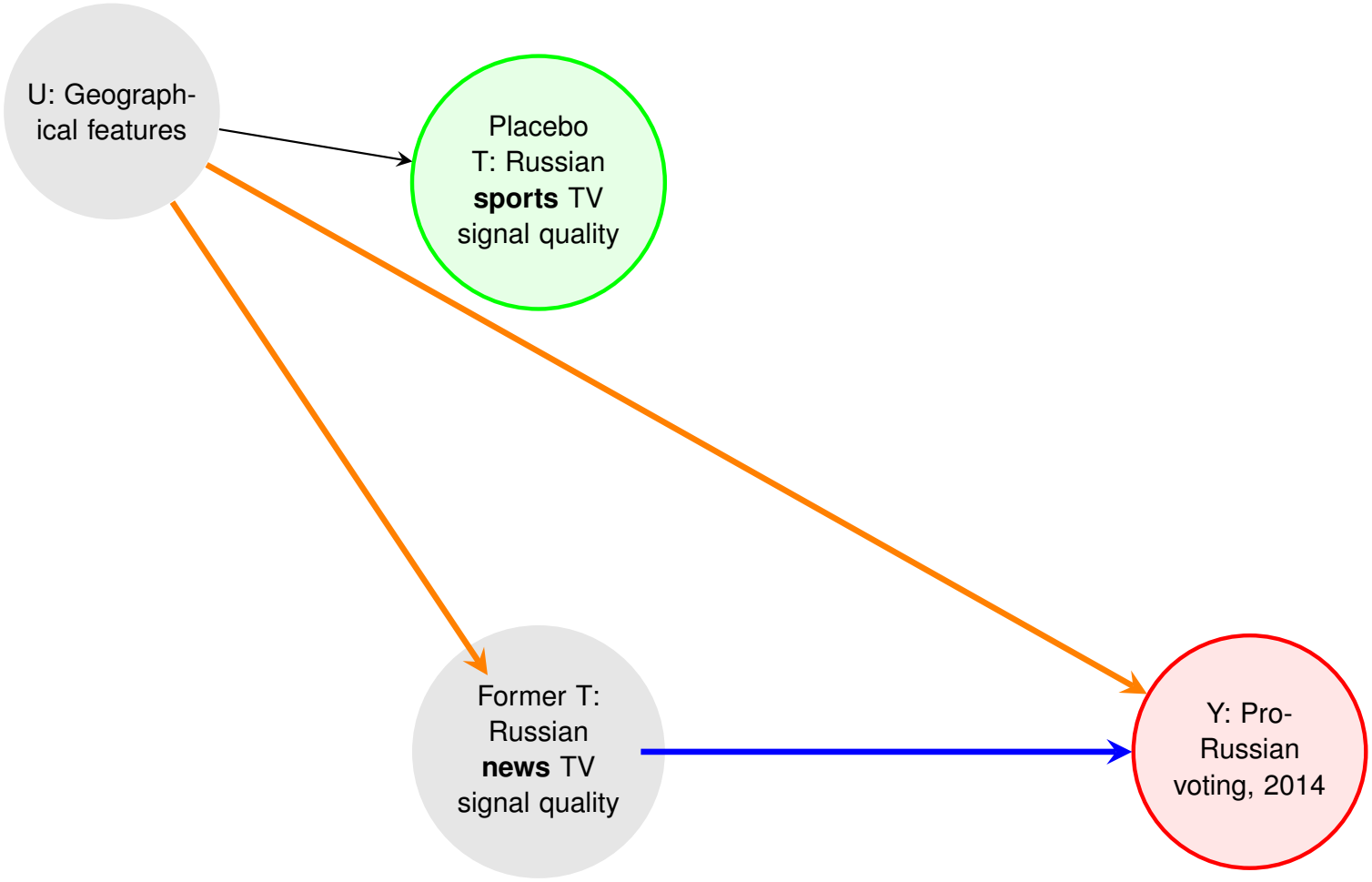


Examples

Paper	Core analysis			Placebo treatment
	Population	Treatment	Outcome	
Jha (2013)	Towns in South Asia proximate to the coast	Whether the town was a medieval trading port	Incidence of Hindu-Muslim riots in 19th and 20th centuries	Whether the town was a colonial overseas port
Burnett & Kogan (2017)	Electoral precincts in San Diego city-wide elections in 2008 and 2010	Citizen pothole complaints before election	Incumbent electoral performance	Pothole complaints in 6 months after election
Brollo & Nannicini 2012	Brazilian municipalities between 1997 and 2008	Partisan alignment between mayor and president (based on election RDD)	Infrastructure transfers from central government	Fake cutoffs

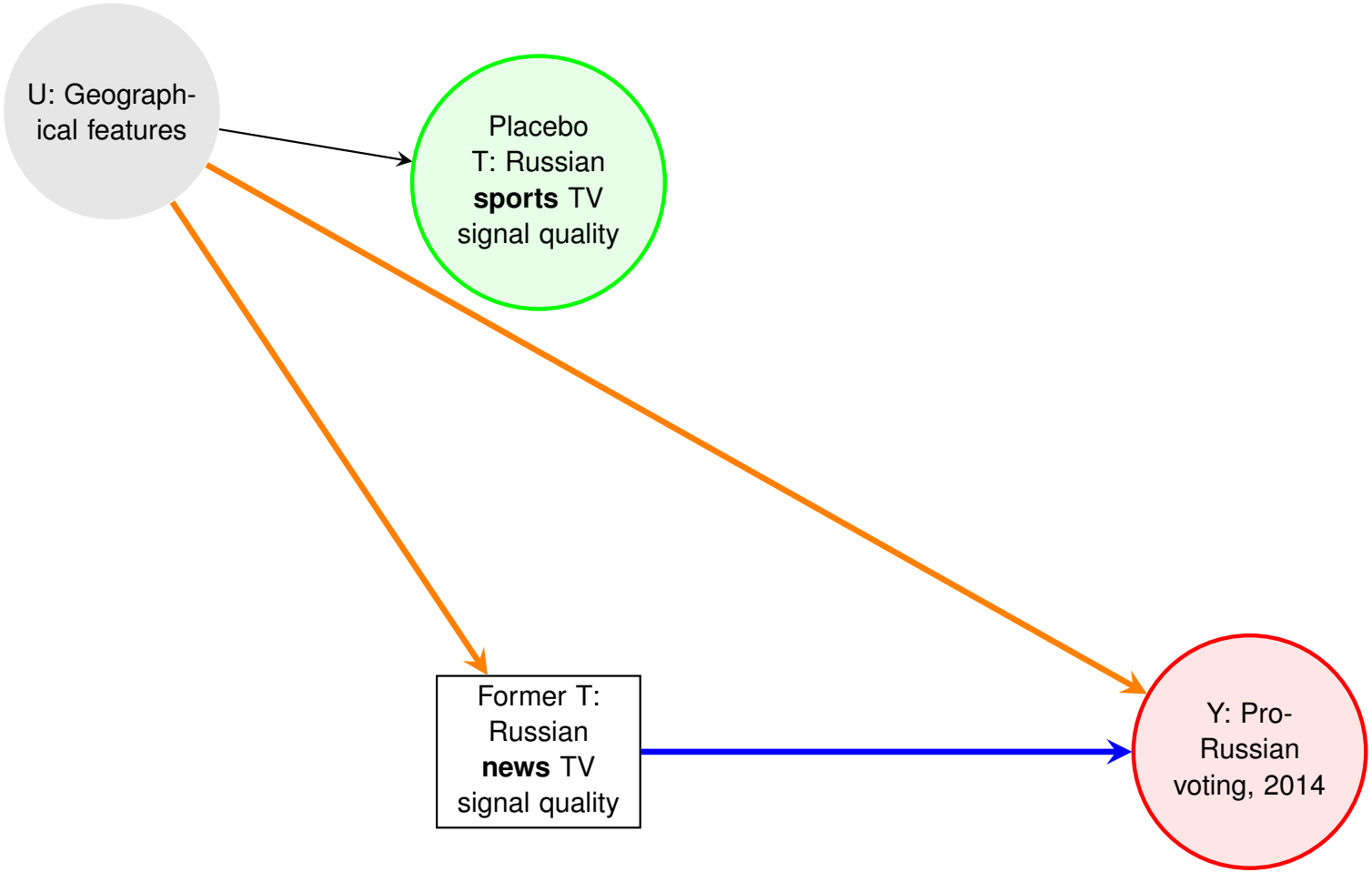
How to run a placebo treatment test

Some authors control for “real” treatment, others don’t. Which is right, the conditional PTT or the unconditional PTT?

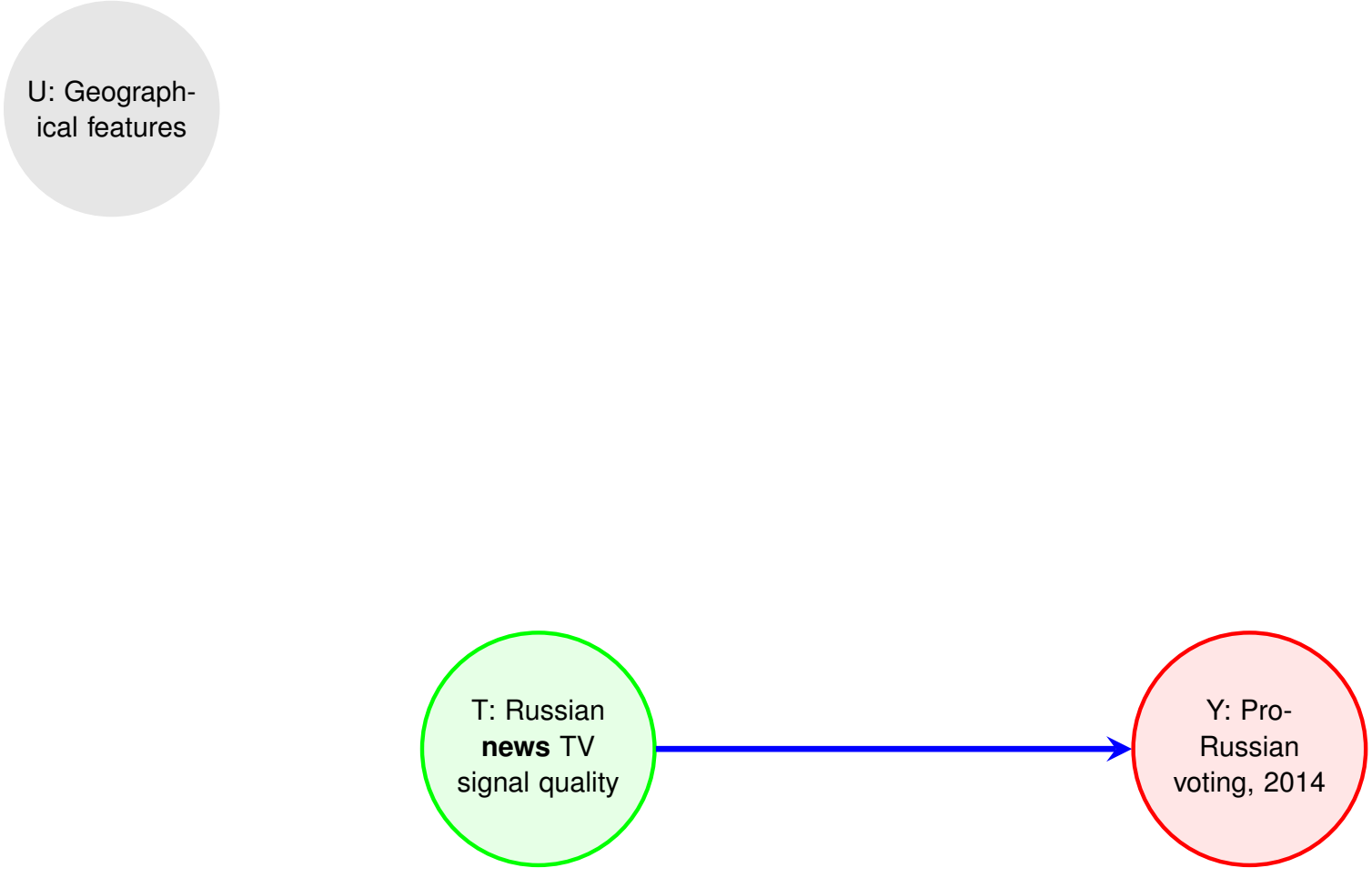


How to run a placebo treatment test

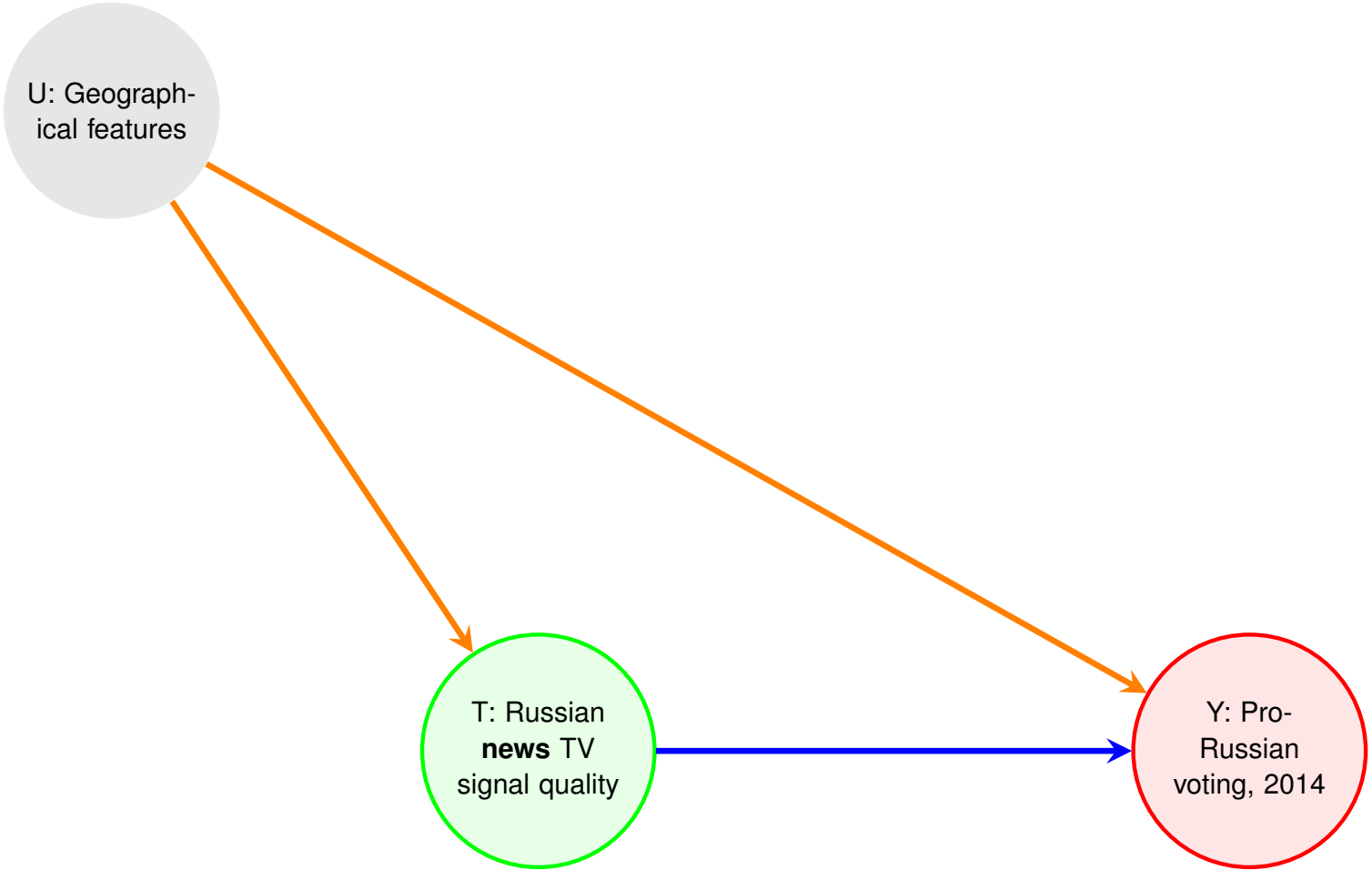
Some authors control for "real" treatment, others don't. Which is right, the conditional PTT or the unconditional PTT?



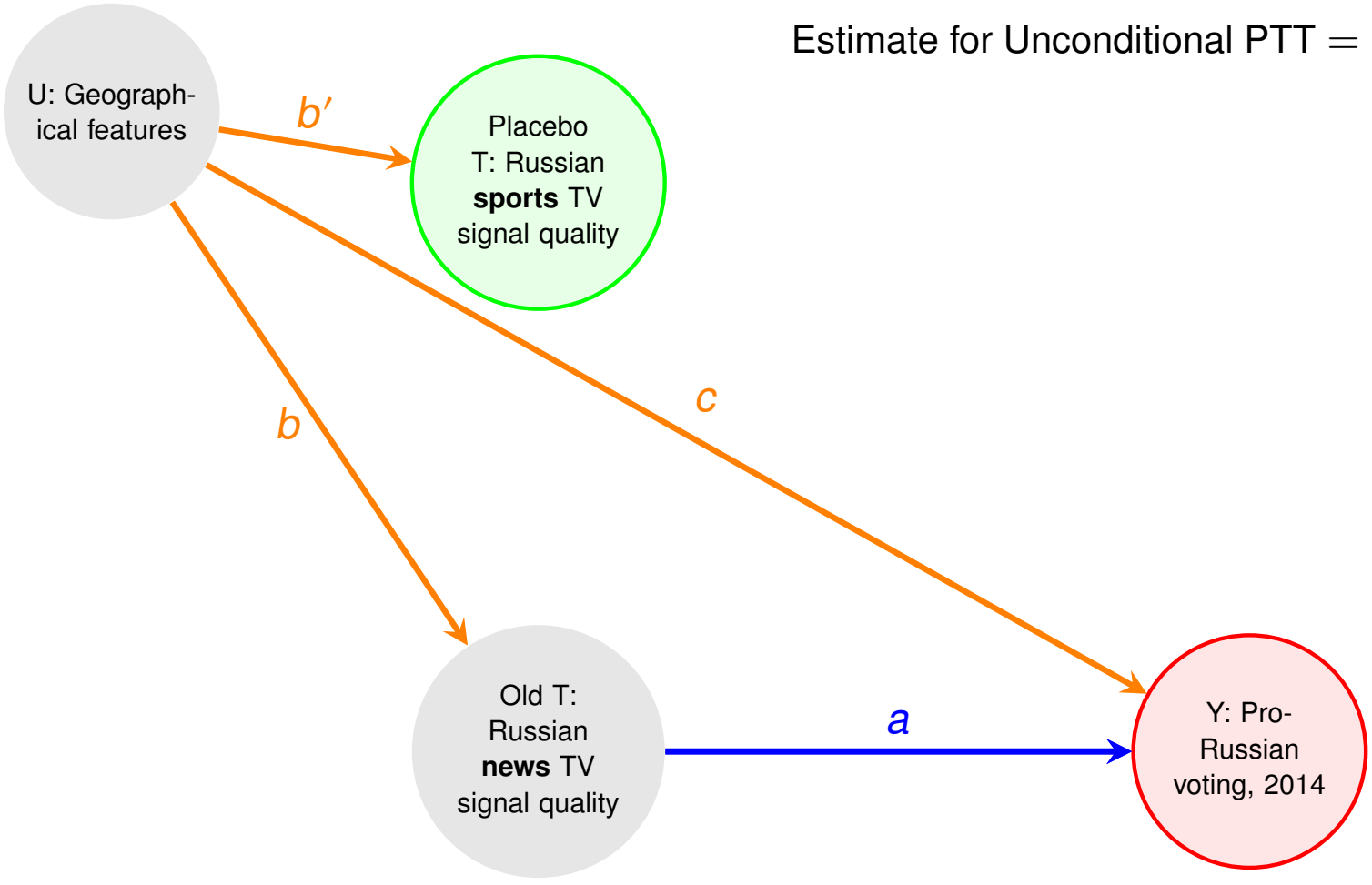
How to run a placebo treatment test (2)



How to run a placebo treatment test (2)

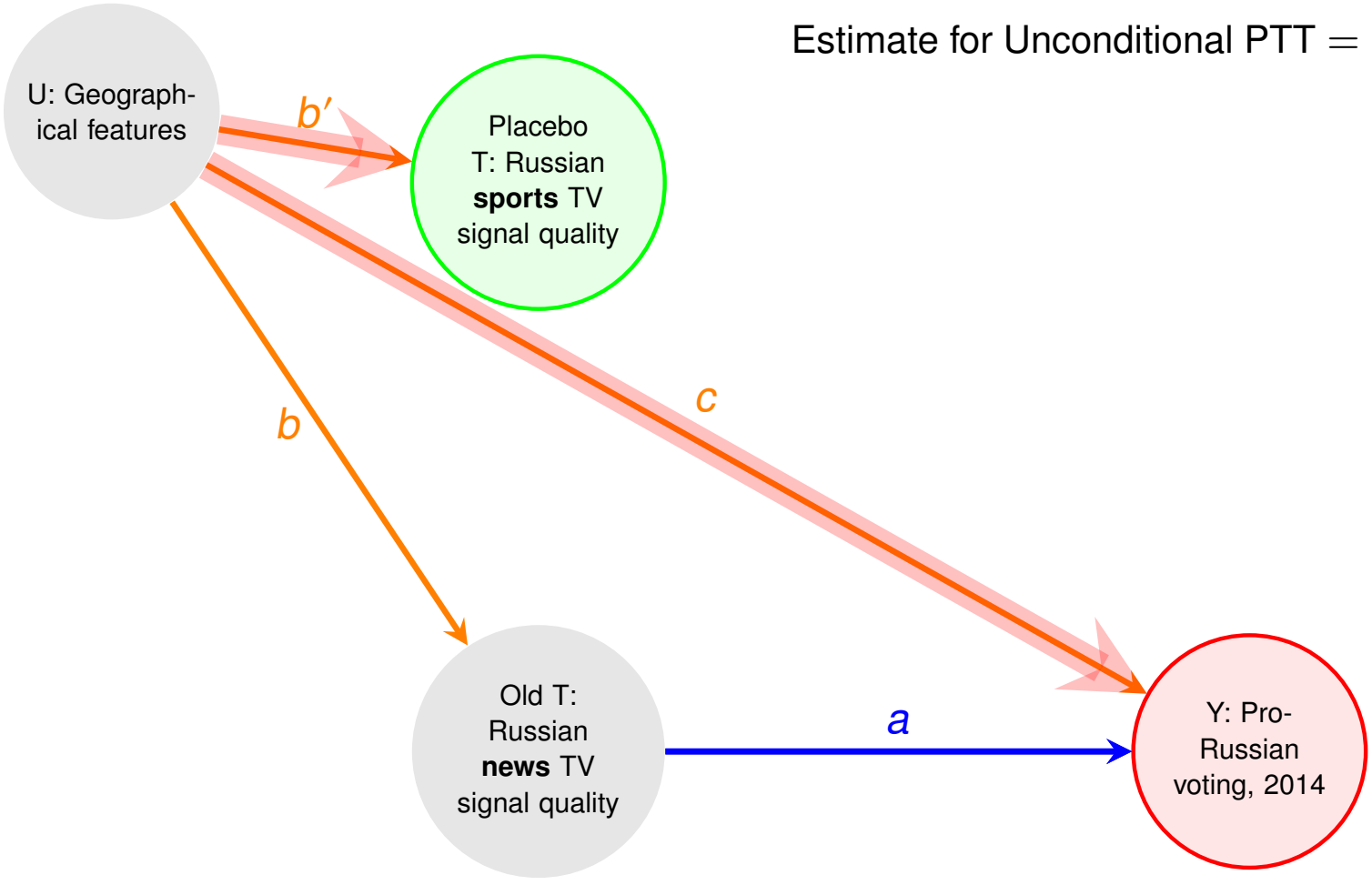


How to run a placebo treatment test (2)



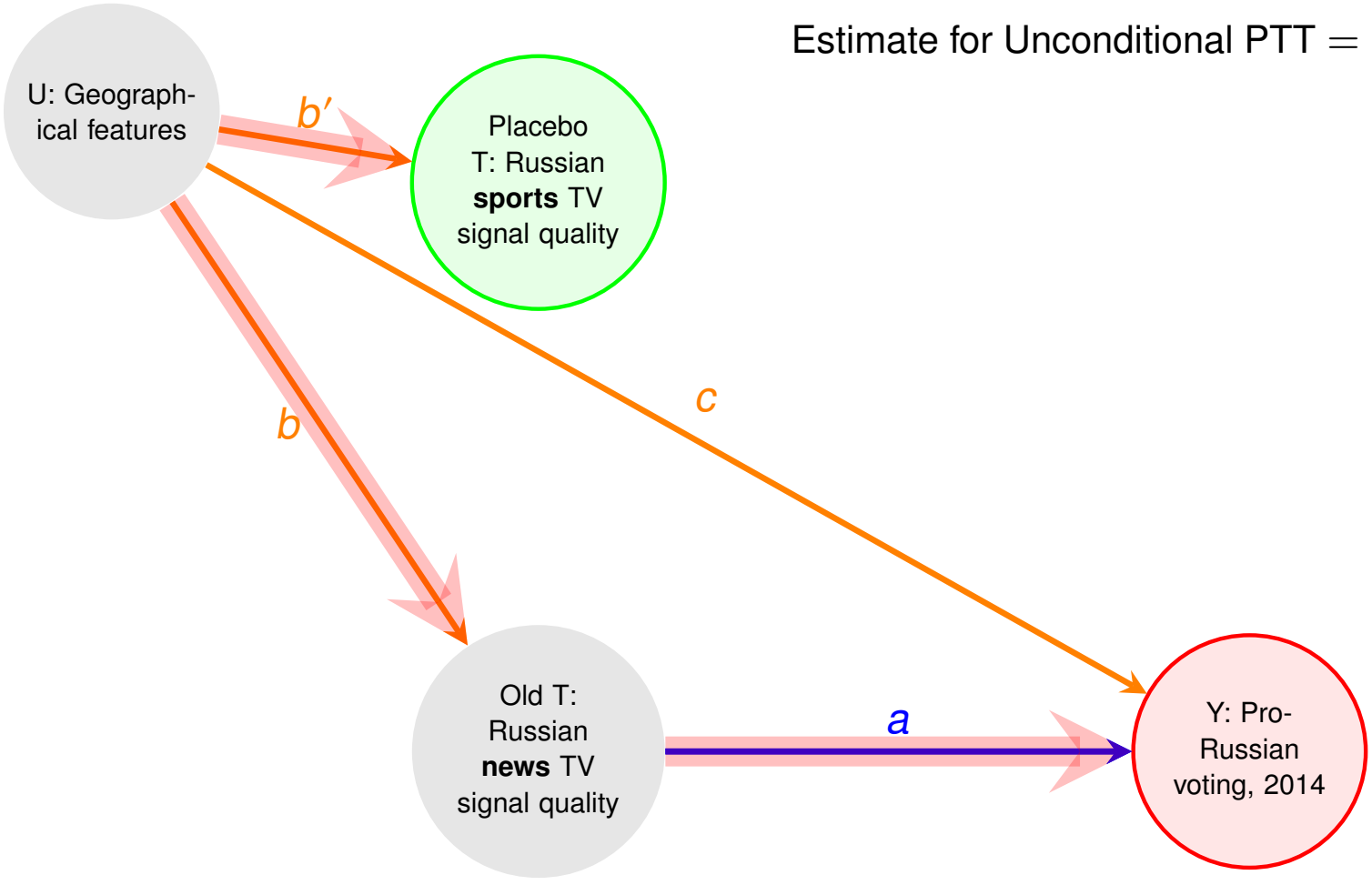
How to run a placebo treatment test (2)

Estimate for Unconditional PTT $= b'c +$



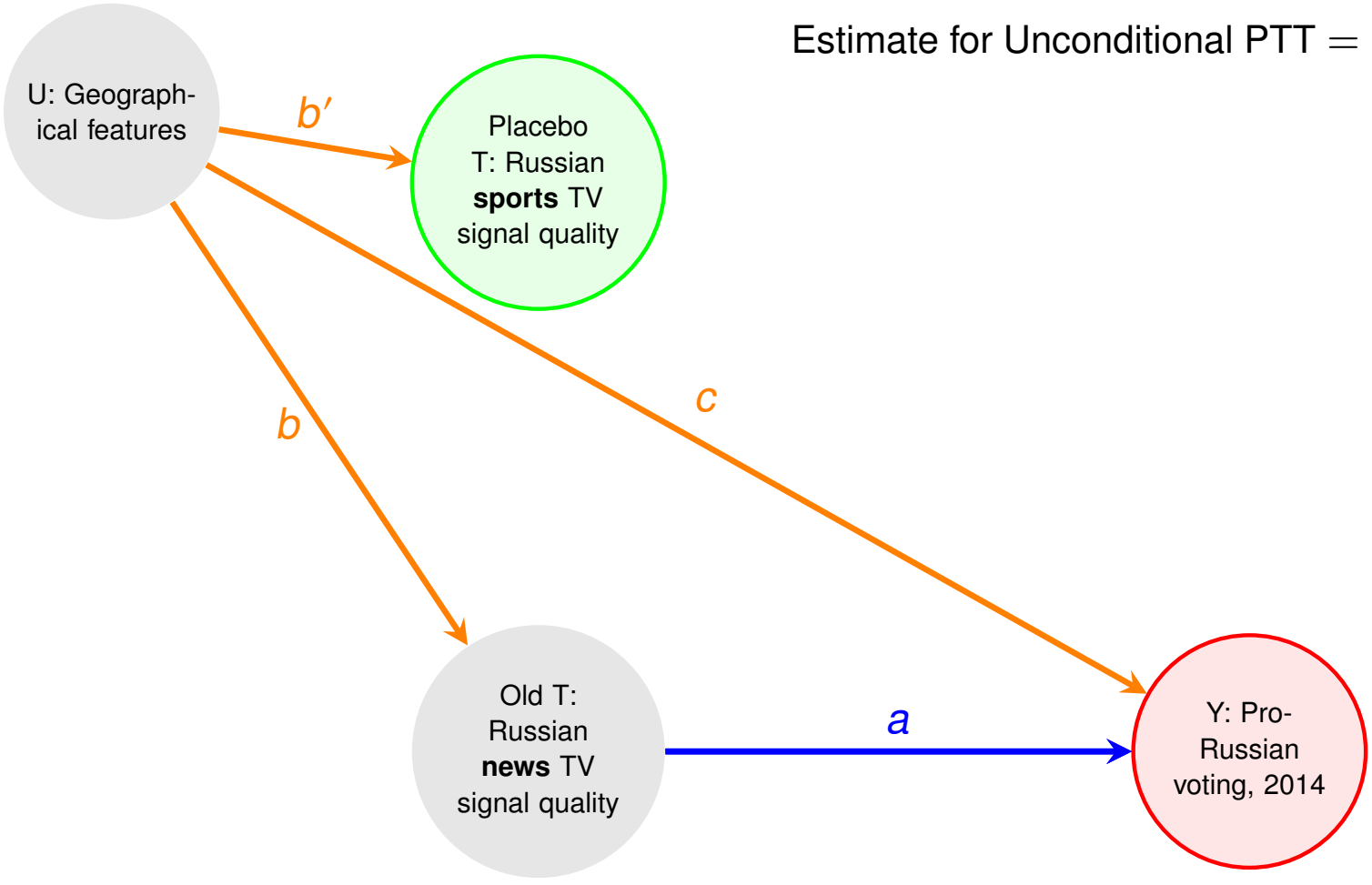
How to run a placebo treatment test (2)

Estimate for Unconditional PTT $= b'c + b'ba$

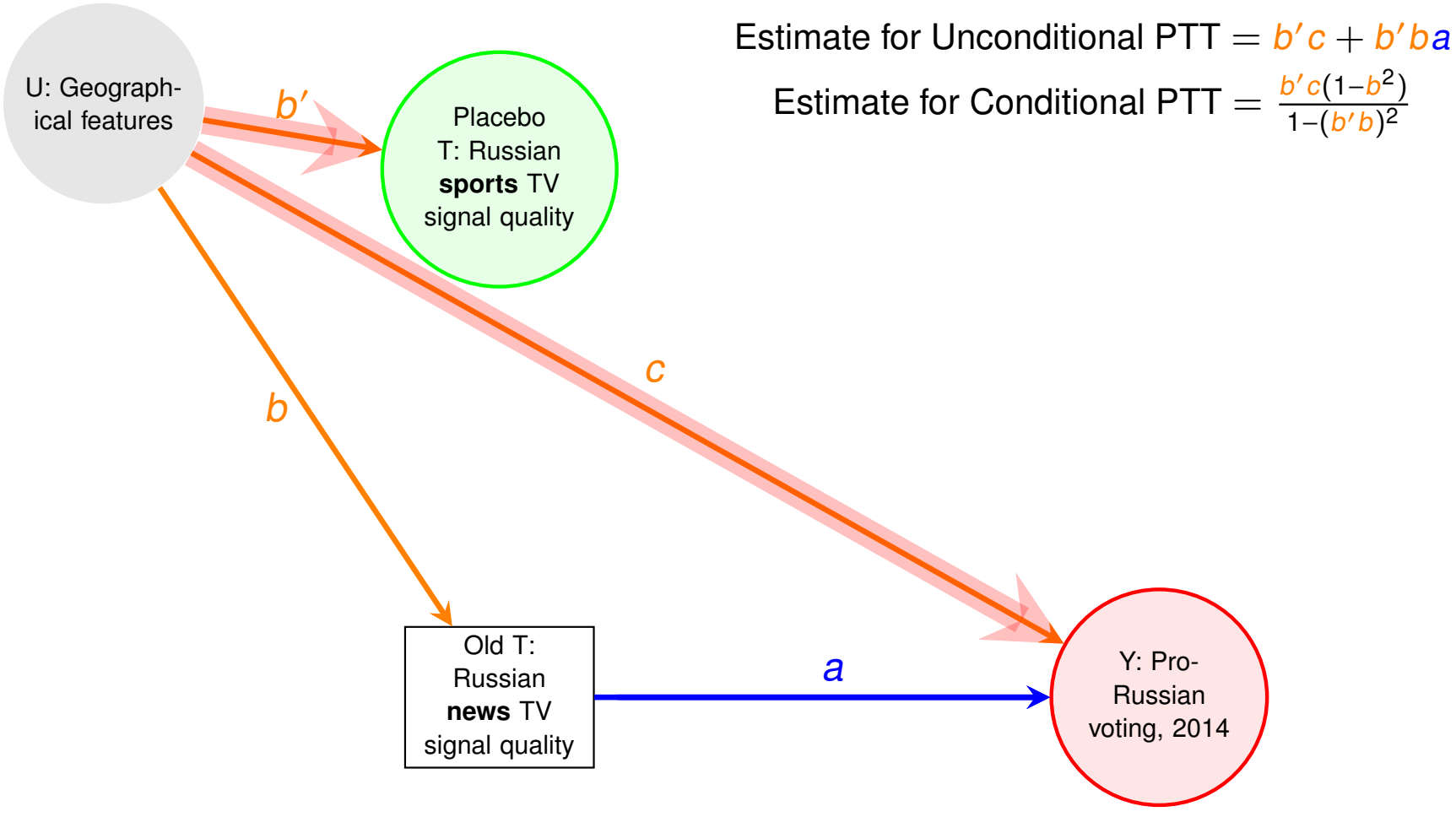


How to run a placebo treatment test (2)

Estimate for Unconditional PTT $= b'c + b'ba$



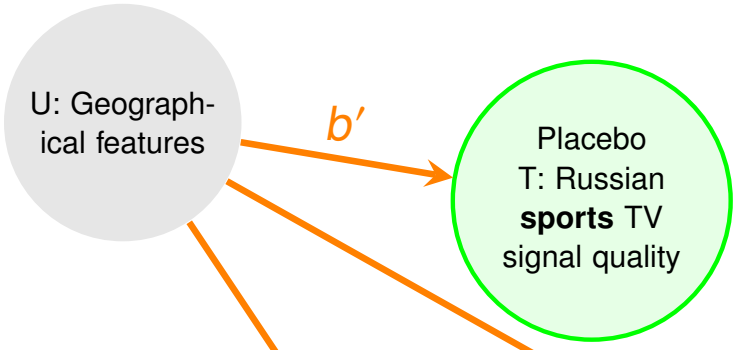
How to run a placebo treatment test (2)

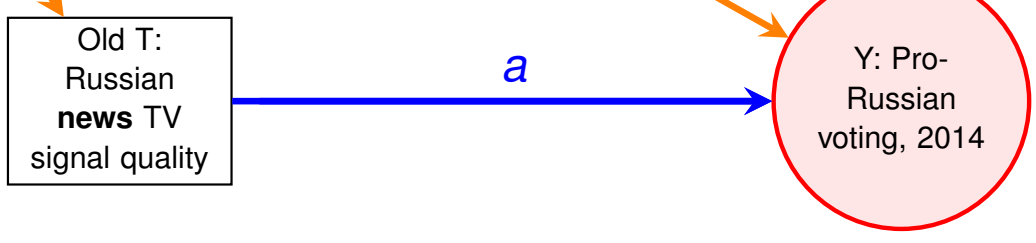


$$\text{Estimate for Unconditional PTT} = b'c + b'ba$$

$$\text{Estimate for Conditional PTT} = \frac{b'c(1-b^2)}{1-(b'b)^2}$$

How to run a placebo treatment test (2)



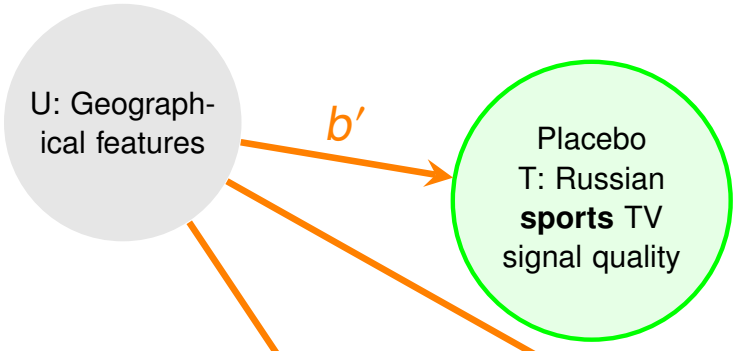


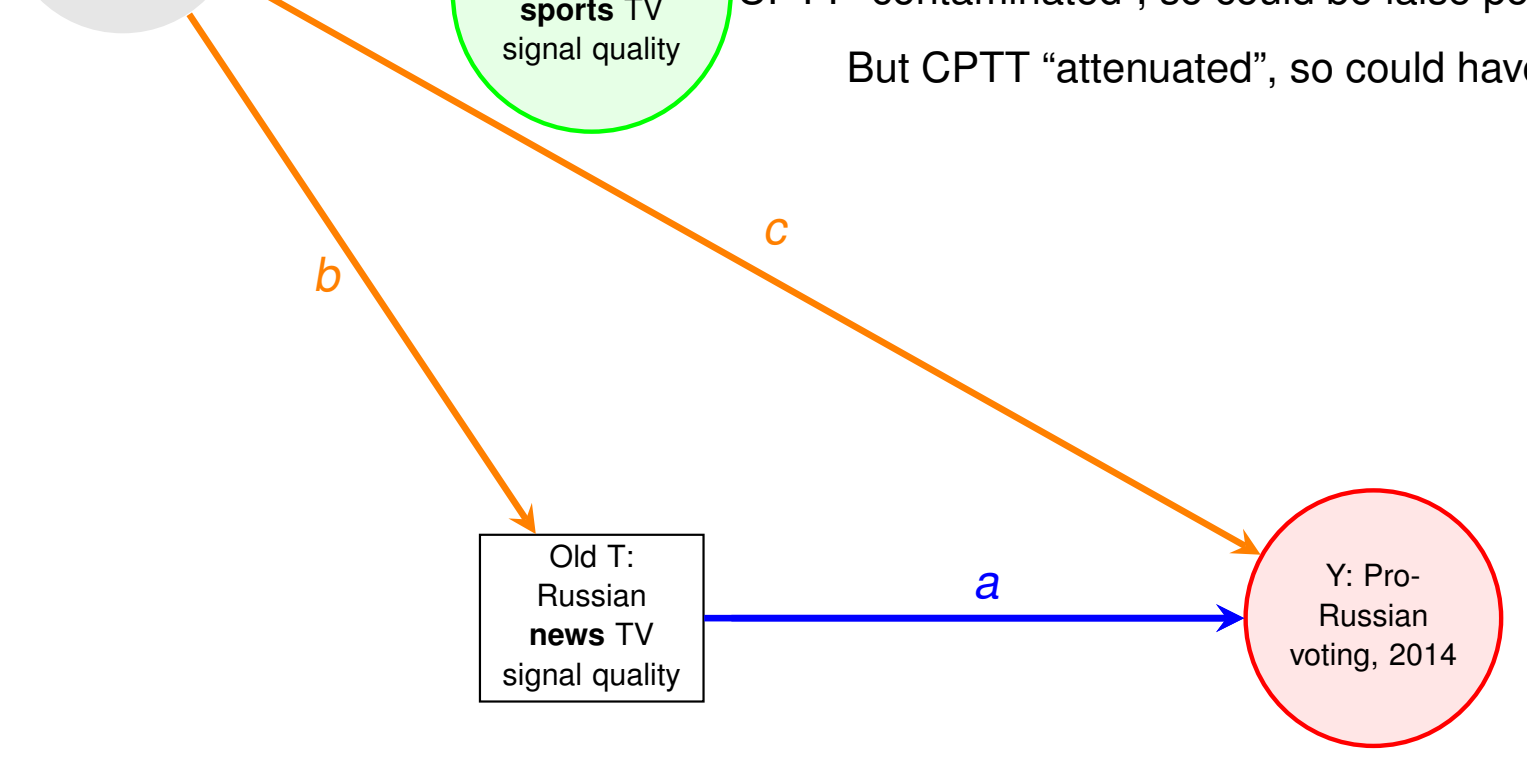
Estimate for Unconditional PTT $= b'c + b'ba$

Estimate for Conditional PTT $= \frac{b'c(1-b^2)}{1-(b'b)^2}$

UPTT "contaminated", so could be false positives, low power

How to run a placebo treatment test (2)





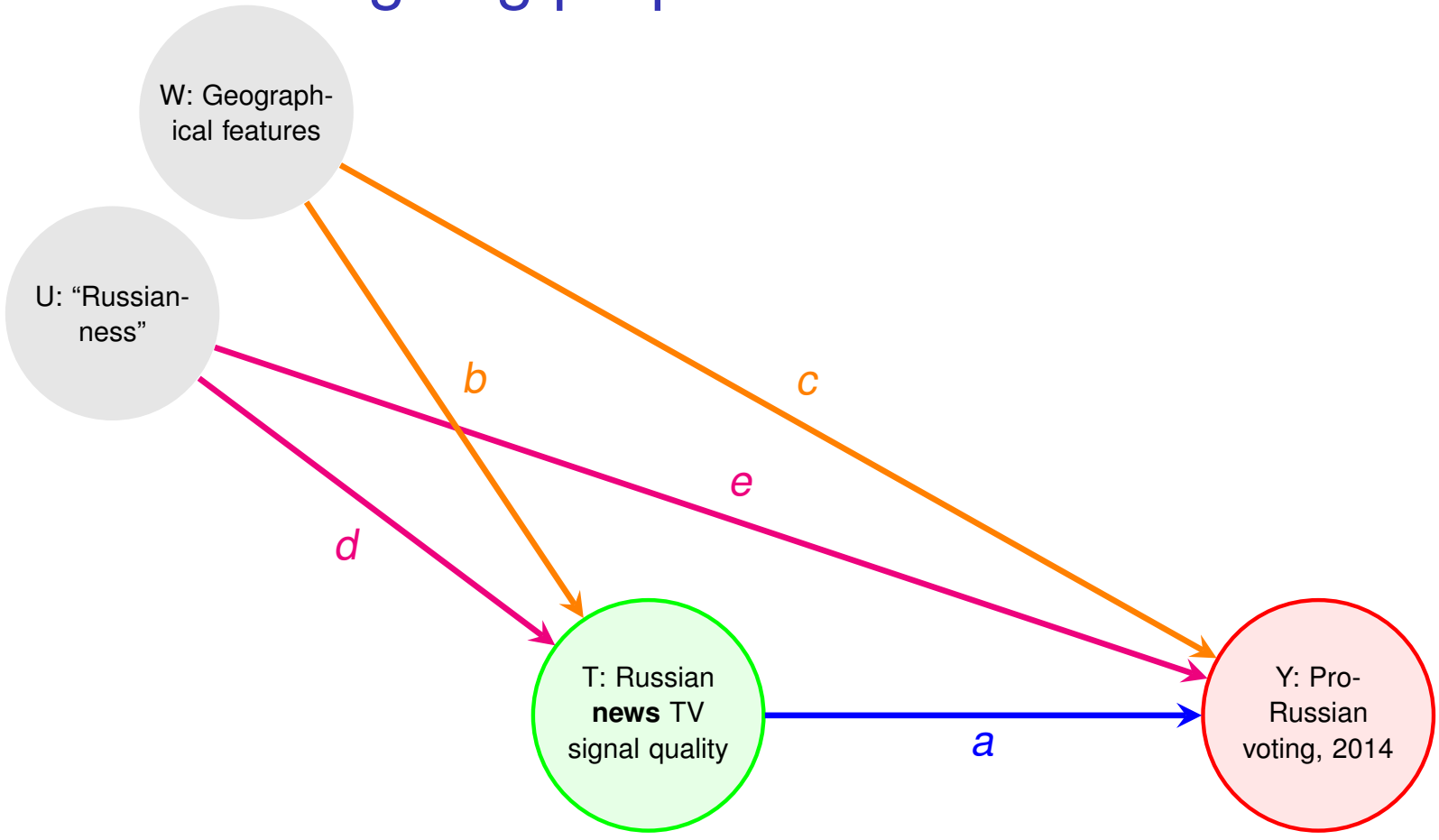
Estimate for Unconditional PTT $= b'c + b'ba$

Estimate for Conditional PTT $= \frac{b'c(1-b^2)}{1-(b'b)^2}$

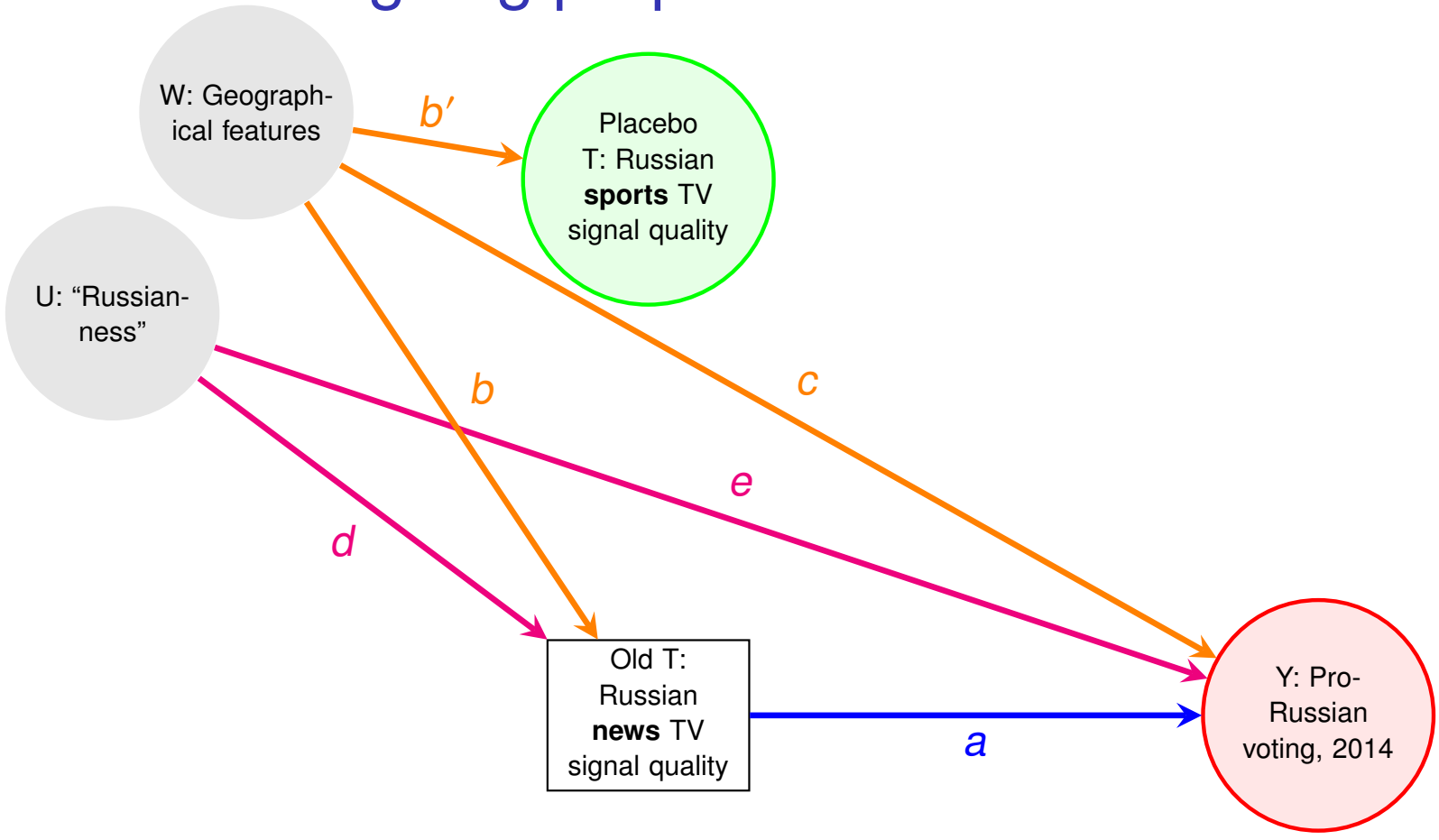
UPTT "contaminated", so could be false positives, low power

But CPTT "attenuated", so could have low power

CPTT: some intriguing properties

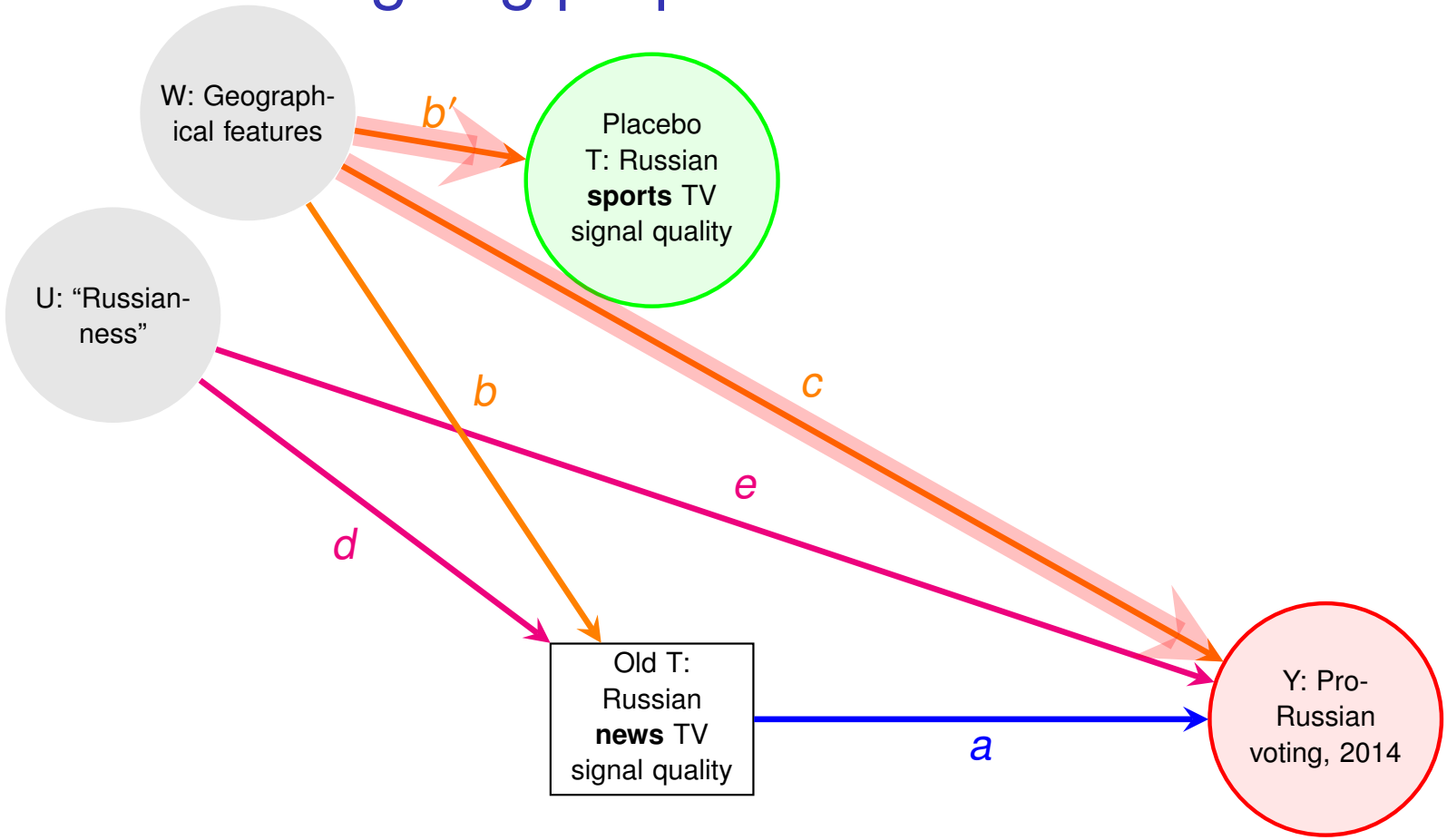


CPTT: some intriguing properties



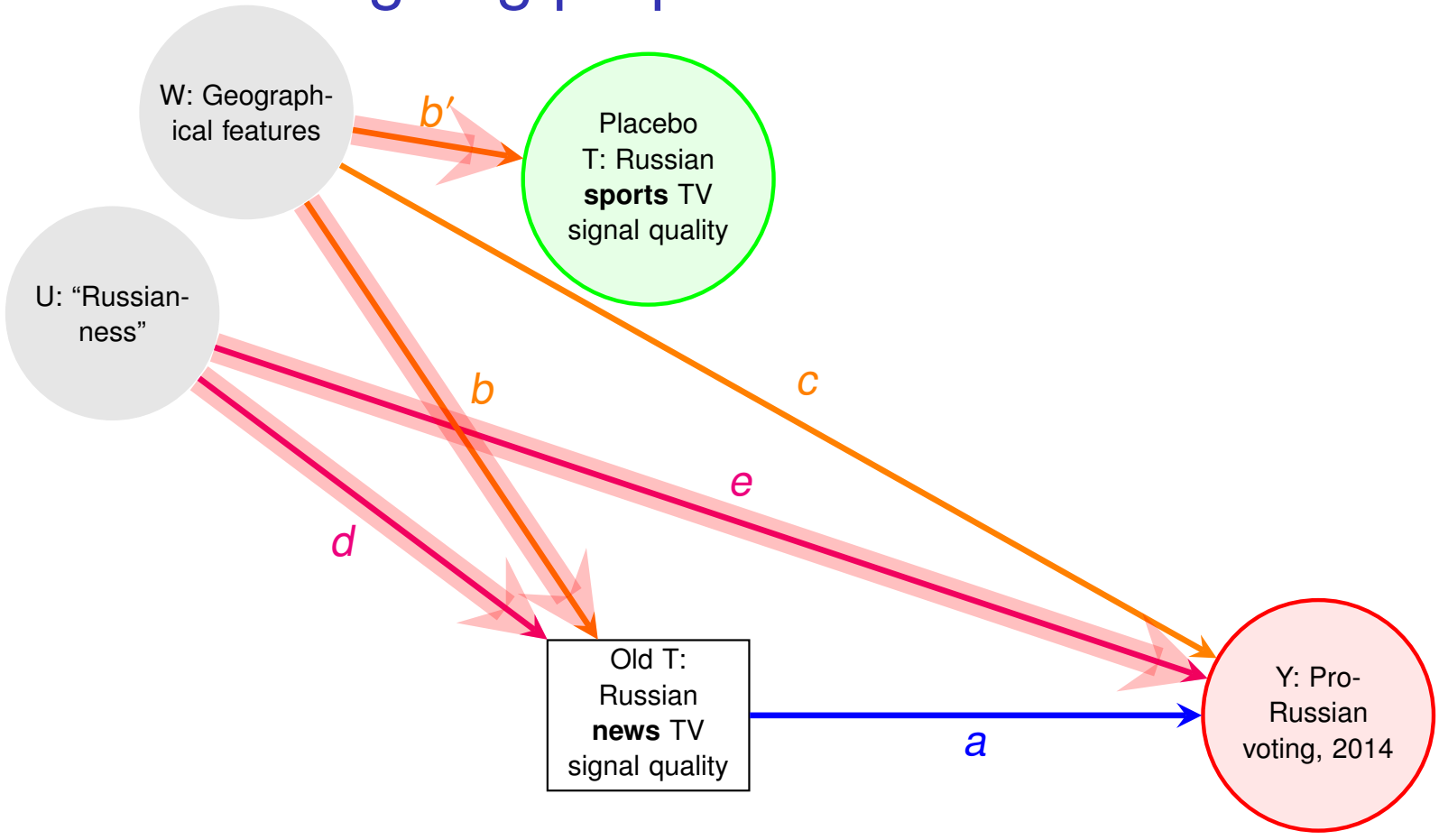
CPTT estimate $=$

CPTT: some intriguing properties



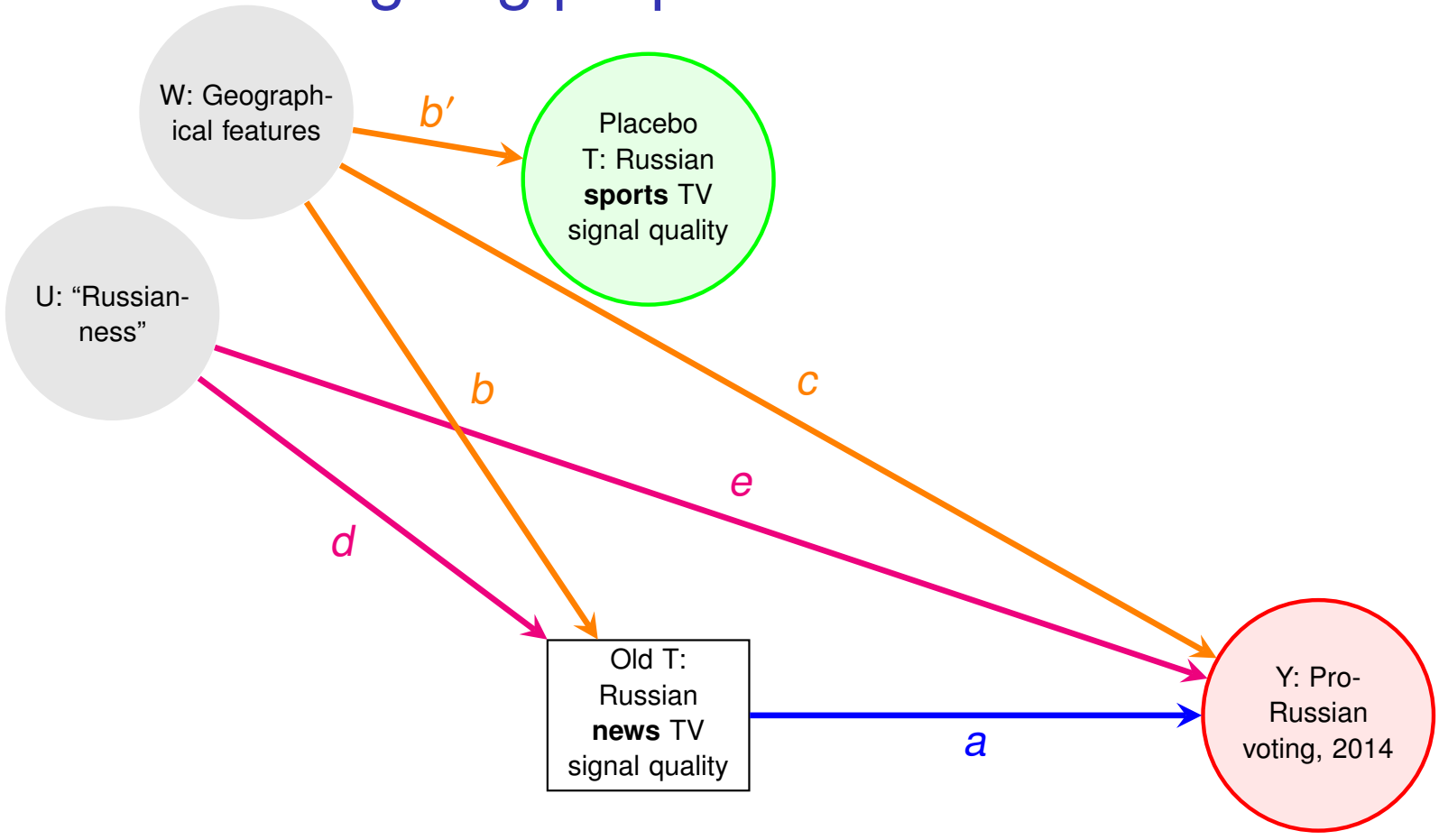
$$\text{CPTT estimate} = \frac{b'c(1-b^2)}{1-(b'b)^2}$$

CPTT: some intriguing properties



$$\text{CPTT estimate} = \frac{b'c(1-b^2)}{1-(b'b)^2} - \frac{b'bde}{1-(b'b)^2}$$

CPTT: some intriguing properties



$$\text{CPTT estimate} = \frac{b'c(1-b^2)}{1-(b'b)^2} - \frac{b'bde}{1-(b'b)^2}$$

Second term is higher-order; perhaps limited relevance (see Ding & Miratrix 2014 on *M*-bias).

Ideally, placebo tests make causal inference more scientific:

"This is my design. I think it works. I hope you like it."	$\Longrightarrow$	"U is a potential confounder; here is a placebo test to check."
"I didn't really buy it. She's making a lot of assumptions."	$\Longrightarrow$	"U is a potential confounder; here is a placebo test to check."

Ideally, placebo tests make causal inference more scientific:

"This is my design. I think it works. I hope you like it."	$\Longrightarrow$	"U is a potential confounder; here is a placebo test to check."
"I didn't really buy it. She's making a lot of assumptions."	$\Longrightarrow$	"U is a potential confounder; here is a placebo test to check."

But

- some tricky issues
- risks of "null hacking"
- important to explain/ask what bias a placebo test is meant to detect