Reliability and Validity, Expressed as Random & Non-Random Measurement Error

May 26, 2010

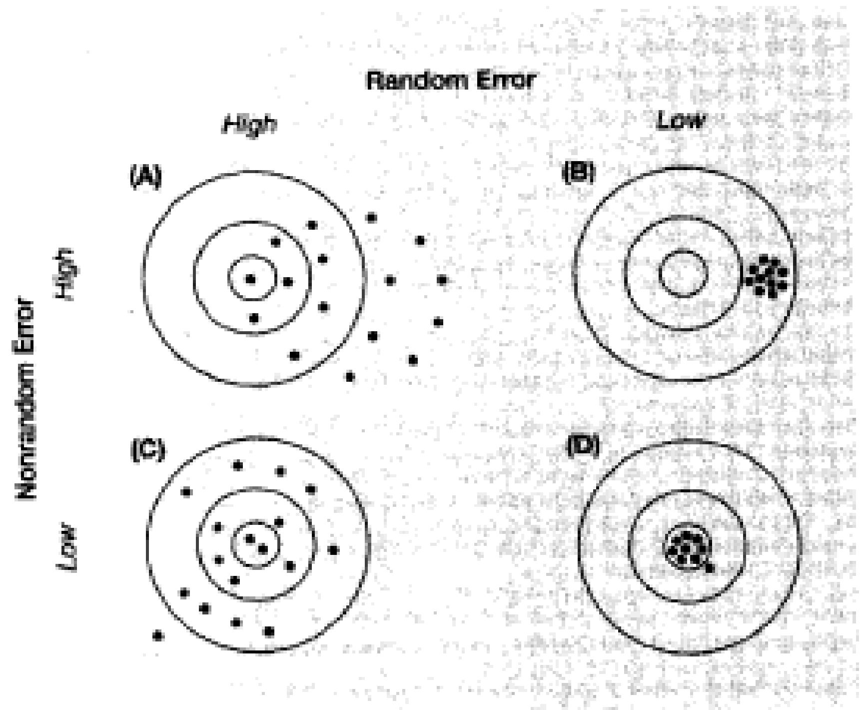
Dr. Kenneth M. Coleman (Ken.Coleman@marketstrategies.com) - University of Michigan

Reliability and Validity, Expressed as Random & Non-Random Measurement Error

as presented by Kenneth M. Coleman,

Figure 4–3 Random and Nonrandom Error

Neither Reliable Nor Valid



Reliable, But Not Valid

Valid, But Not Reliable

Reliable and Valid

Illustration from Shively, <u>The Craft of Political Research</u>, 6th edition, Pearson Prentice Hall, 2005, p. 49.

Causal Inference and Its Limits

May 26, 2010

Dr. Allen Hicken (ahicken@umich.edu) – University of Michigan

What is Causality?

- Causality versus causal inference.
- Causality: X is a cause of Y; Y is an effect of X
- Causal Inference: Can we infer X→Y from our sample?

Causal Inference and Its Limits

Allen Hicken

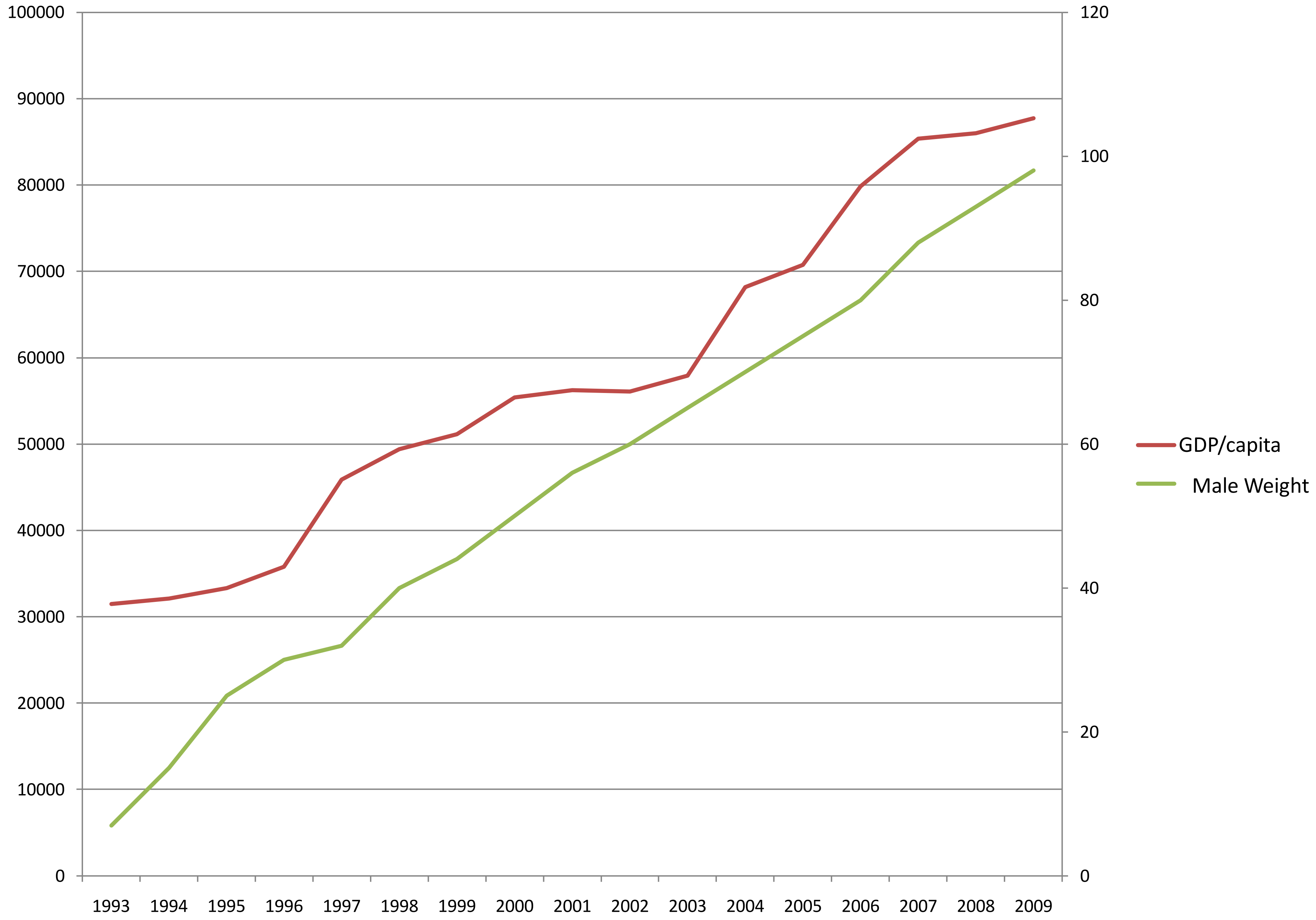
Prepared for Presentation at SESRI, May 2010

Criteria for Establishing Causality

1. Correlation (Association)
2. Temporal Ordering
3. Theory (Causal Mechanism)
4. Isolation (Rule out Confounds)

Criterion #1. Correlation

- Two variables are "correlated" when changes in one variable occur together with changes in the other (Louise White)
  - Correlation is roughly synonymous with association and co-variance.
  - A correlation between two variables can be positive or negative.


100000
90000
80000
70000
60000
50000
40000
30000
20000
10000
0
120
100
80
60
40
20
0
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
GDP/capita
Male Weight

Establishing Causality

- We observe: X correlates with Y

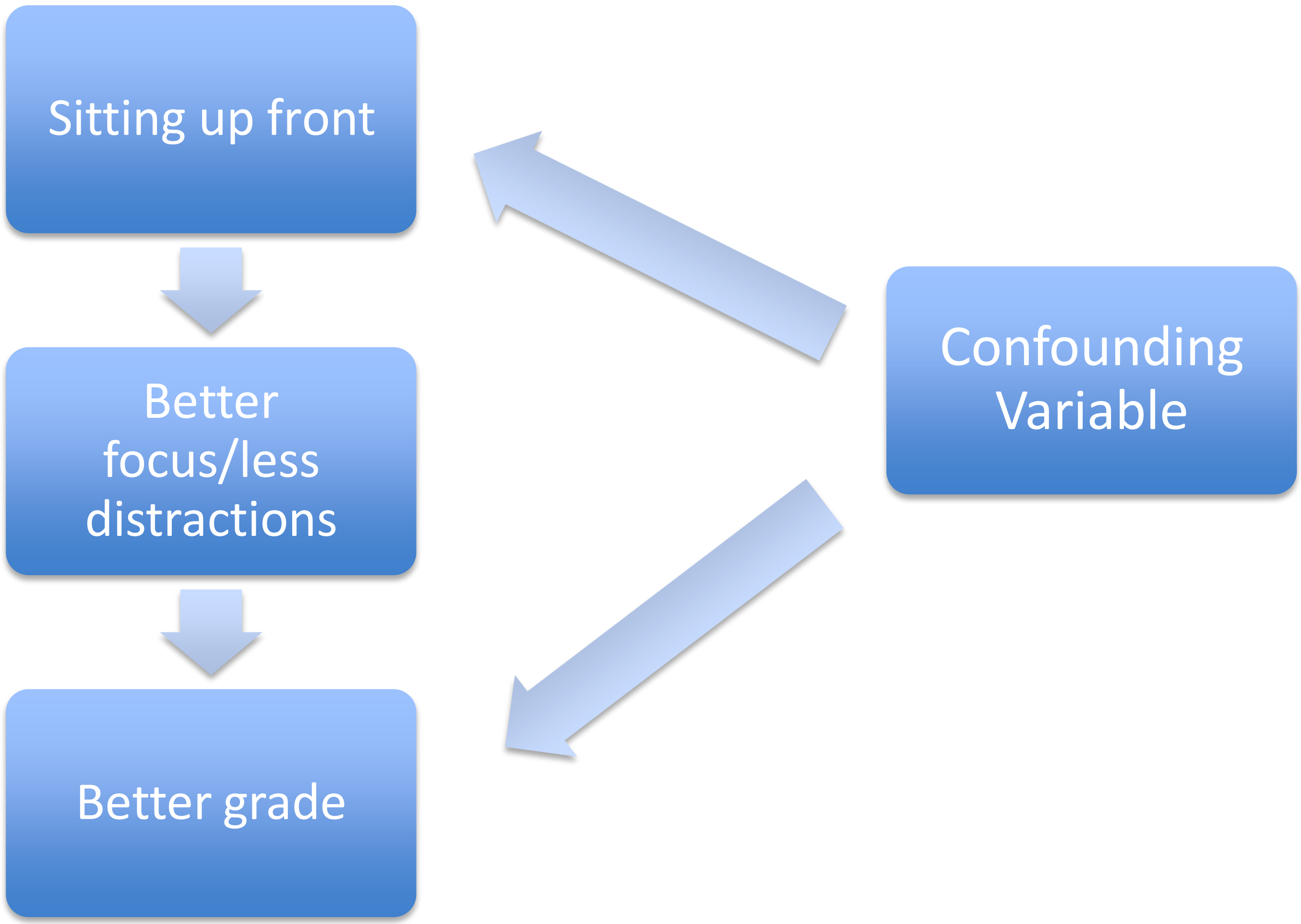
1. Causation is not involved at all
2. There is a causal link
3. Confounding (omitted) variable (Z) causes both X and Y

Criterion #2 Temporal Ordering

- The hypothesized cause (IV) must come before the effect (DV).
  - Students decide whether or not to sit in the front of class before the get their final grade.
    - Or do they?
  - Social science has lots of tricky “chicken-and-egg” situations.

Criterion #3 Causal Mechanism

- You have to be able to tell a plausible story that connects the cause (IV) to the effect (DV)
  - This story often includes an “intervening variable” that gets us from the cause to the effect
  - Students who sit up front are able to hear better, see better, better comprehend the lecture, and are less tempted by distractions (plausible story)
  - Students who sit up front of the class bask in my aura and absorb more of my genius by just being close to me (not plausible)


Sitting up front
Better focus/less distractions
Better grade
Confounding Variable

Criterion #4 Isolation (Rule Out Confounds)

- If there is a confounding variable that is causally prior to both a cause (IV) and an effect (DV), then the correlation we observe between the cause and the effect may be spurious.

Criterion #4 Isolation (Rule Out Confounds)

- If there is a confounding variable that is causally prior to both an cause and an effect, then the correlation we observe between the cause and the effect may be spurious.

- When it comes to causal inference this is perhaps the biggest challenge for non-experimental researchers.

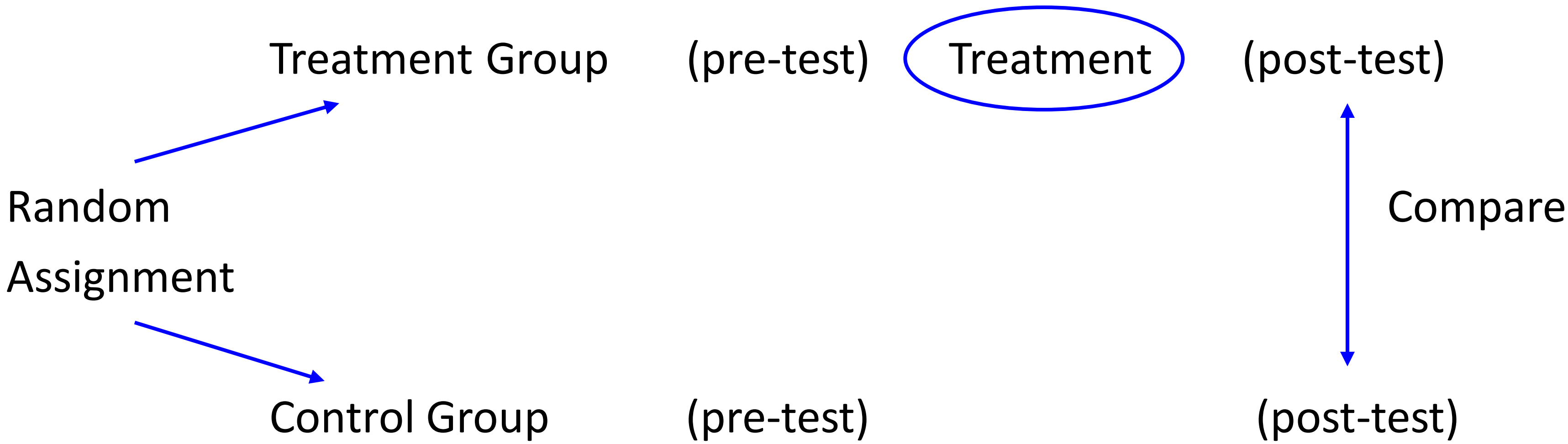
The Beauty of Random Assignment

- Problem: In non-experimental studies, what determines the values that an independent variable takes on? Often, a confounding variable determines these values, *and affects the DV*.
  - For instance, the confound of “Are you a serious student” may determine where you will sit in a class.

The Beauty of Random Assignment

- Solution: Interrupt the causal path that leads from the confound to the independent variable by “randomly assigning” the values that the IV takes on in each case.
  - Randomly assign seats so that there are just as many serious students and slackers in each part of the lecture hall.

Schematic of an Experiment



A Thought Experiment

- An observational study could in principle have been an experiment but for ethical concerns or logistical issues.
- You are probably not estimating a causal effect if you can’t answer Dorn’s (1953) Question: “what experiment would you have run if you were dictator and had infinite resources?”

The Fundamental Problem of Causal Inference

- Problem. We cannot rerun history to see whether changing the value of an independent variable would have changed the value of the dependent variable.

- Solution #1. Give up.

The Fundamental Problem of Causal Inference

- Solution #2. Design your research in a way that comes as close as possible to rerunning history.
  - Observe the effects of changes in one independent variable when all other independent variables remain the same, or
  - Measure other independent variables, then use statistical techniques to hold them constant.

Solution

- We know $Y = \beta X$ is usually wrong
- Progress can be made if we assume that the groups are comparable once we condition on observable covariates denoted by Z.
- $Y = \beta_1 X + \beta_2 Z + \varepsilon$ is the more realistic specification
- Our ability to make causal inferences depends on the quality of the Z variables

Exercises

Each of following conclusions is based on a relationship between X and Y that could be spurious. For each one: (i) identify a plausible confounding variable (Z) for which you would ideally control, (ii) Briefly describe how Z might be affecting the relationship between X and Y.

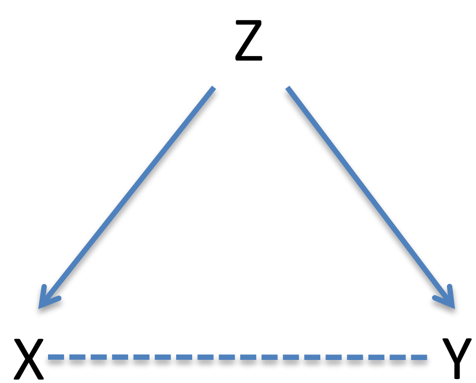
1. In Great Britain, the level of ice cream sales (X) and drowning deaths (Y) are strongly related; as sales go up, so do deaths from drowning. Conclusion: To save lives we should prohibit ice cream sales.
2. Car color (X) and accident rates (Y) are linked: Red cars are more likely to be involved in accidents than are non-red cars. Conclusion: If red cars are banned, the accident rate will drop.
3. Women's education (X) and divorce rates (Y) are correlated: more educated women have a higher divorce rate than less-educated women. Conclusion: Education causes divorce.

Adapted from Pollock 2009

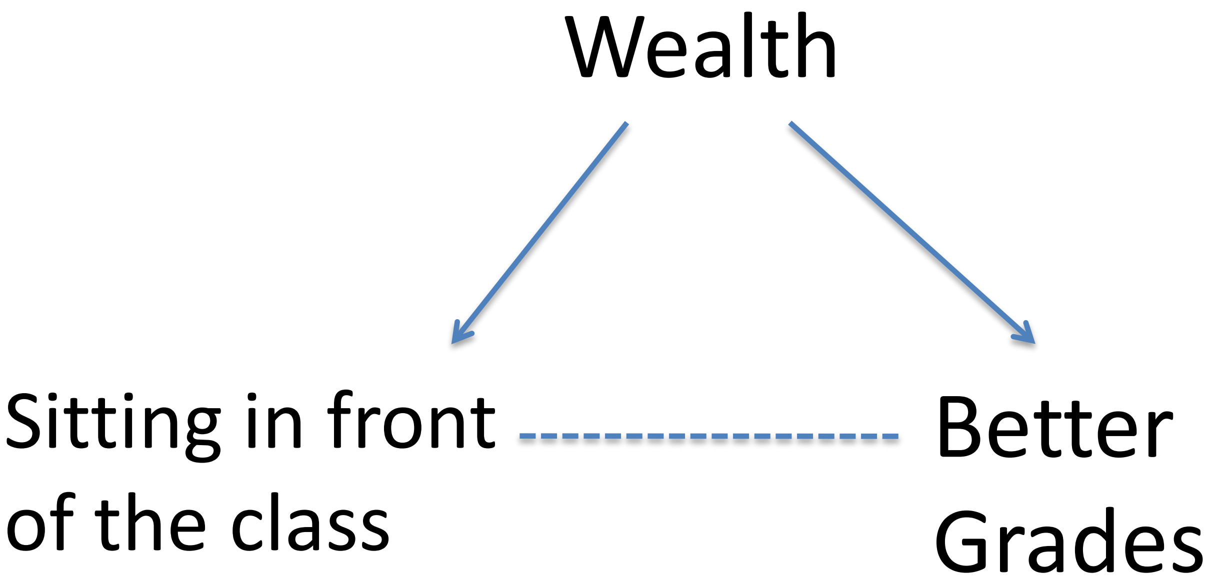
Establishing Causality

- We observe: X correlates with Y

1. Causation is not involved at all
2. There is a causal link
3. Confounding variable causes both X and Y

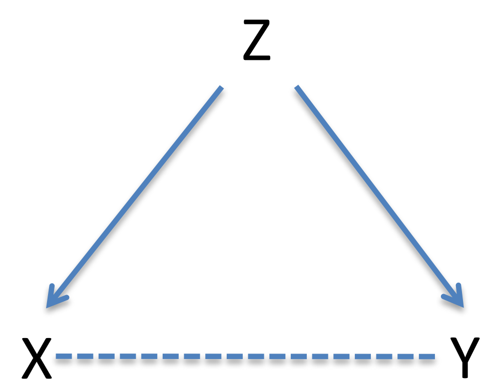


Example



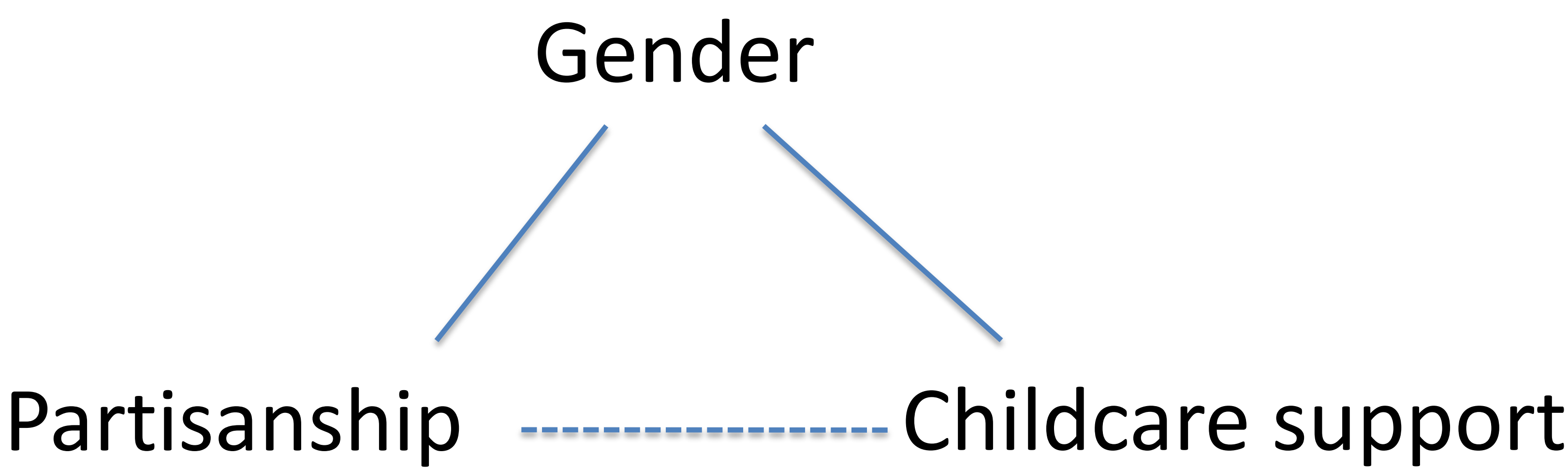
Dealing with Confounding Variables

- Control variables
  - Holding potential confounding variables constant
- 3 possible outcomes when control for Z
  - Spurious relationship
  - Additive relationship
  - Interactive relationship



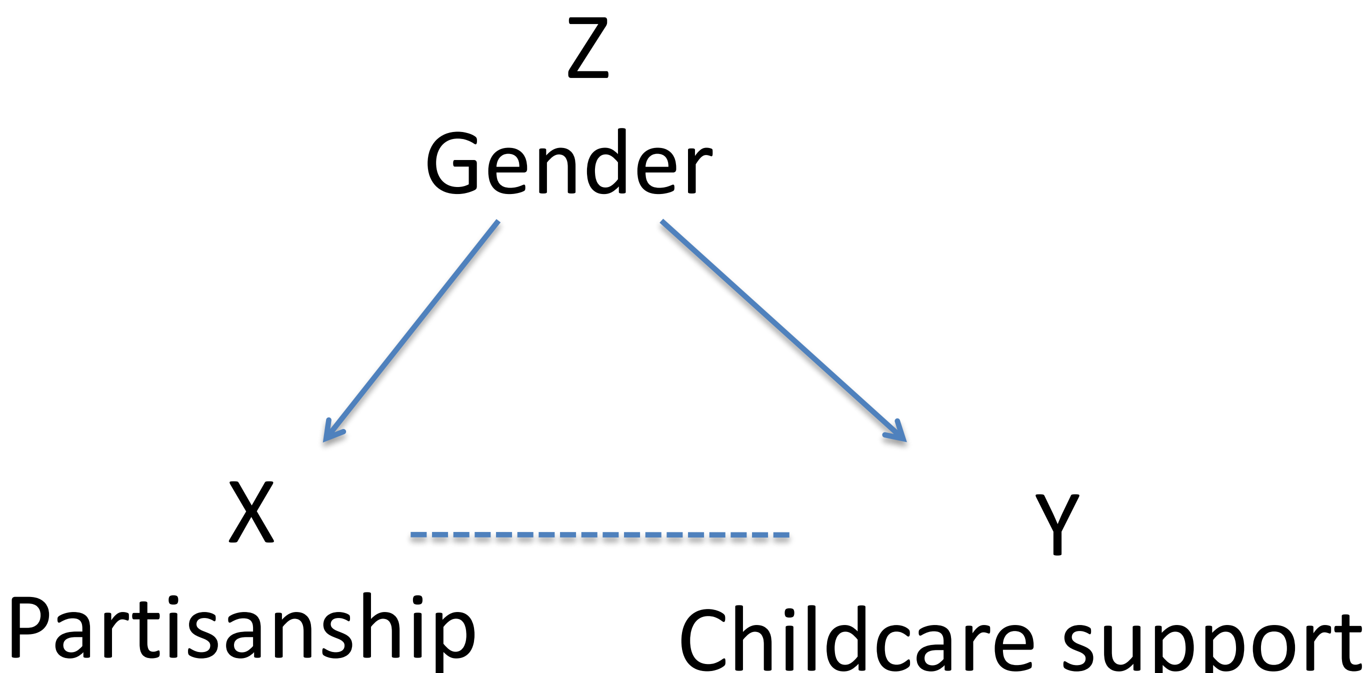
Examples

- Relationship between partisanship and support for government funding for childcare
  - Partisanship → support for childcare
  - What could be a confounding or control variable?

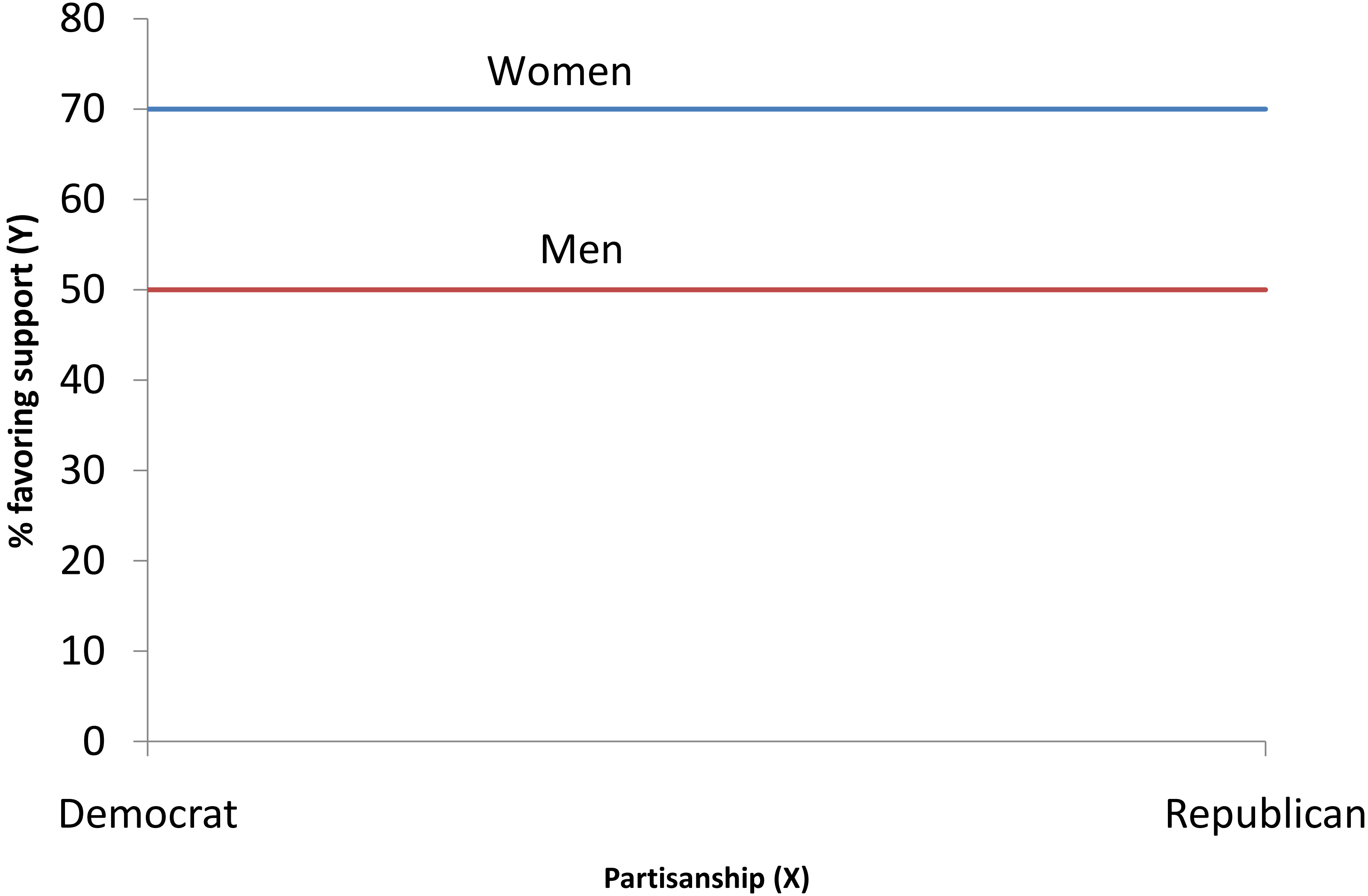


Spurious relationship

- After holding Z constant the causal connection between X and Y disappears

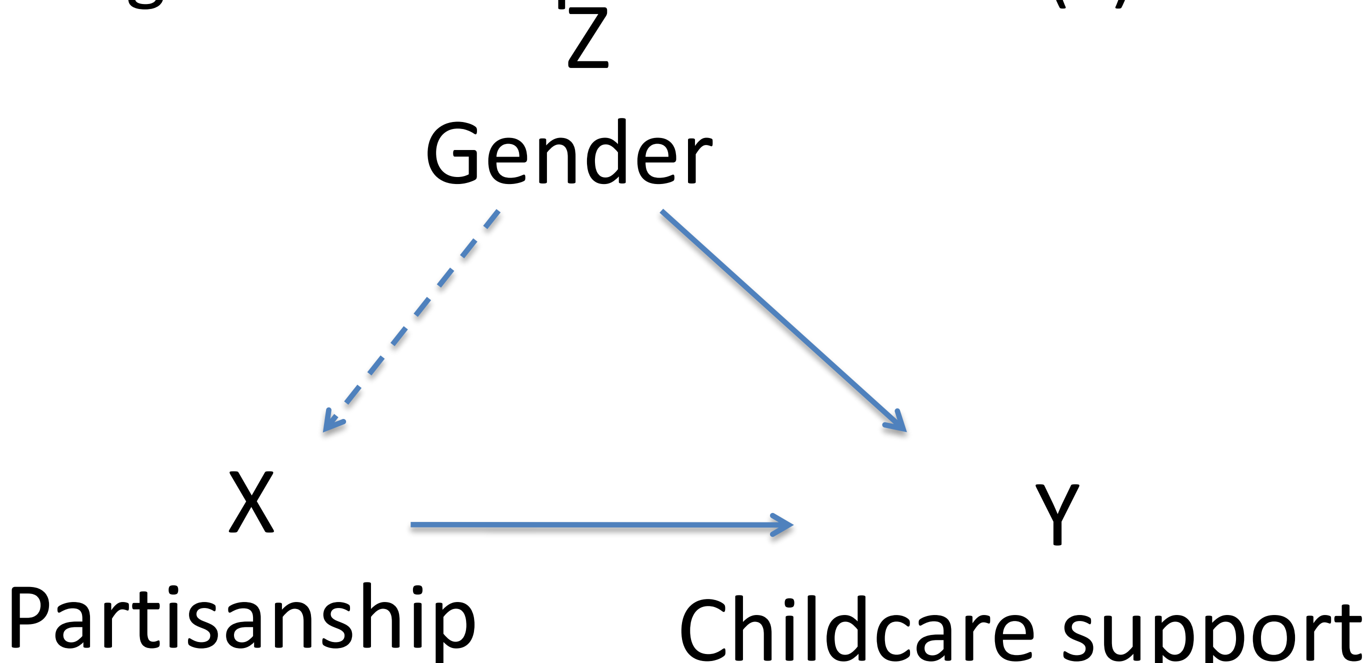


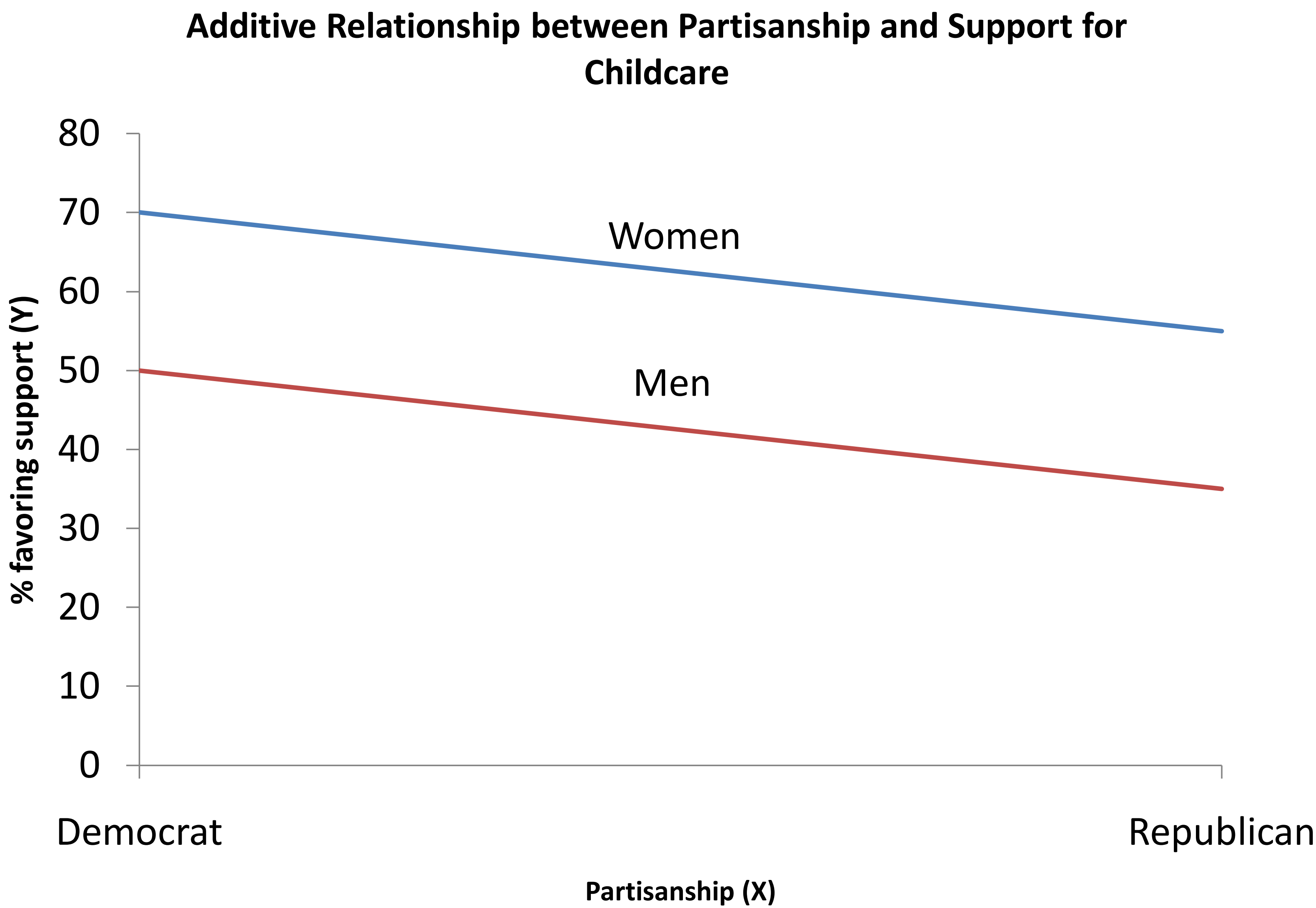
Spurious Relationship between Partisanship and Support for Childcare



Additive Relationship

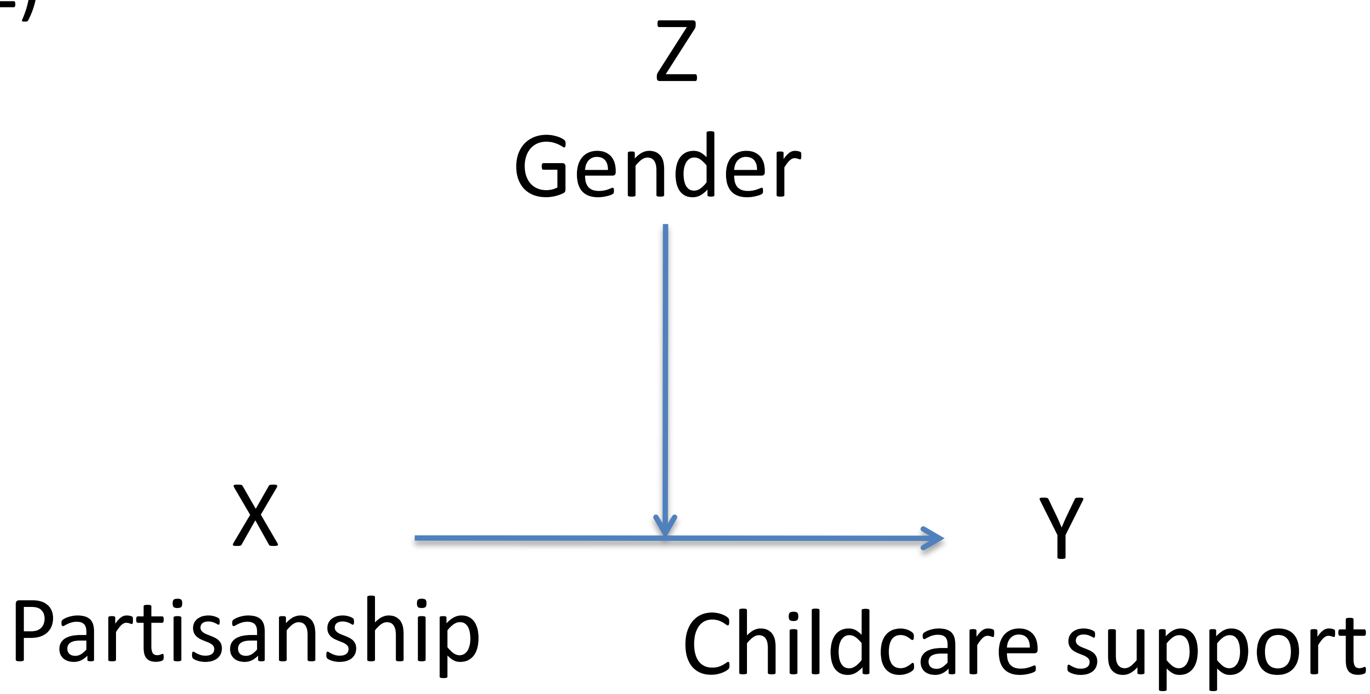
- The control variable (Z) has a weak or non-existent relationship with the IV (X) and a strong relationship with the DV (Z).

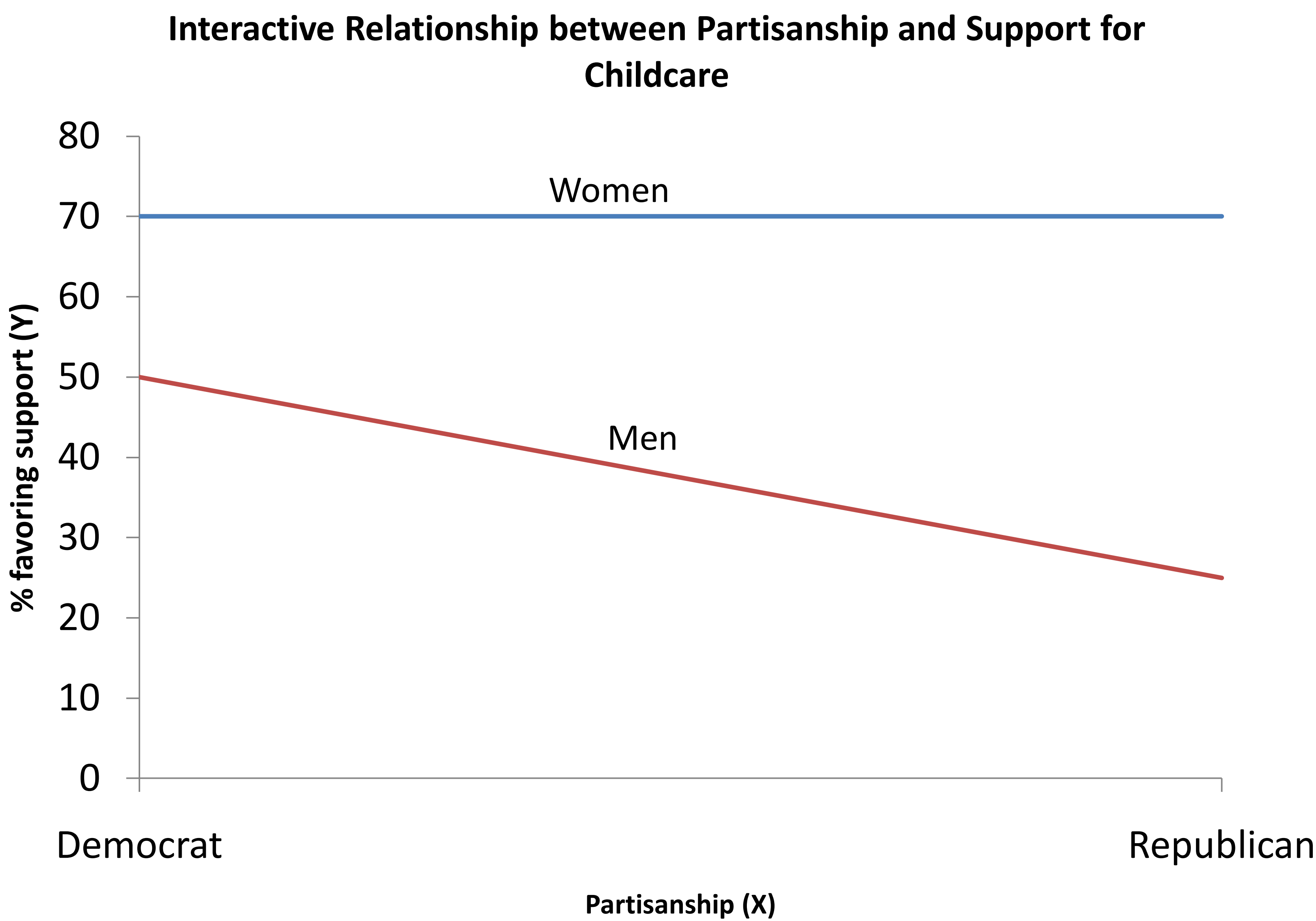



Additive Relationship between Partisanship and Support for Childcare
80
70
60
50
40
30
20
10
0
% favoring support (Y)
Women
Men
Democrat
Republican
Partisanship (X)

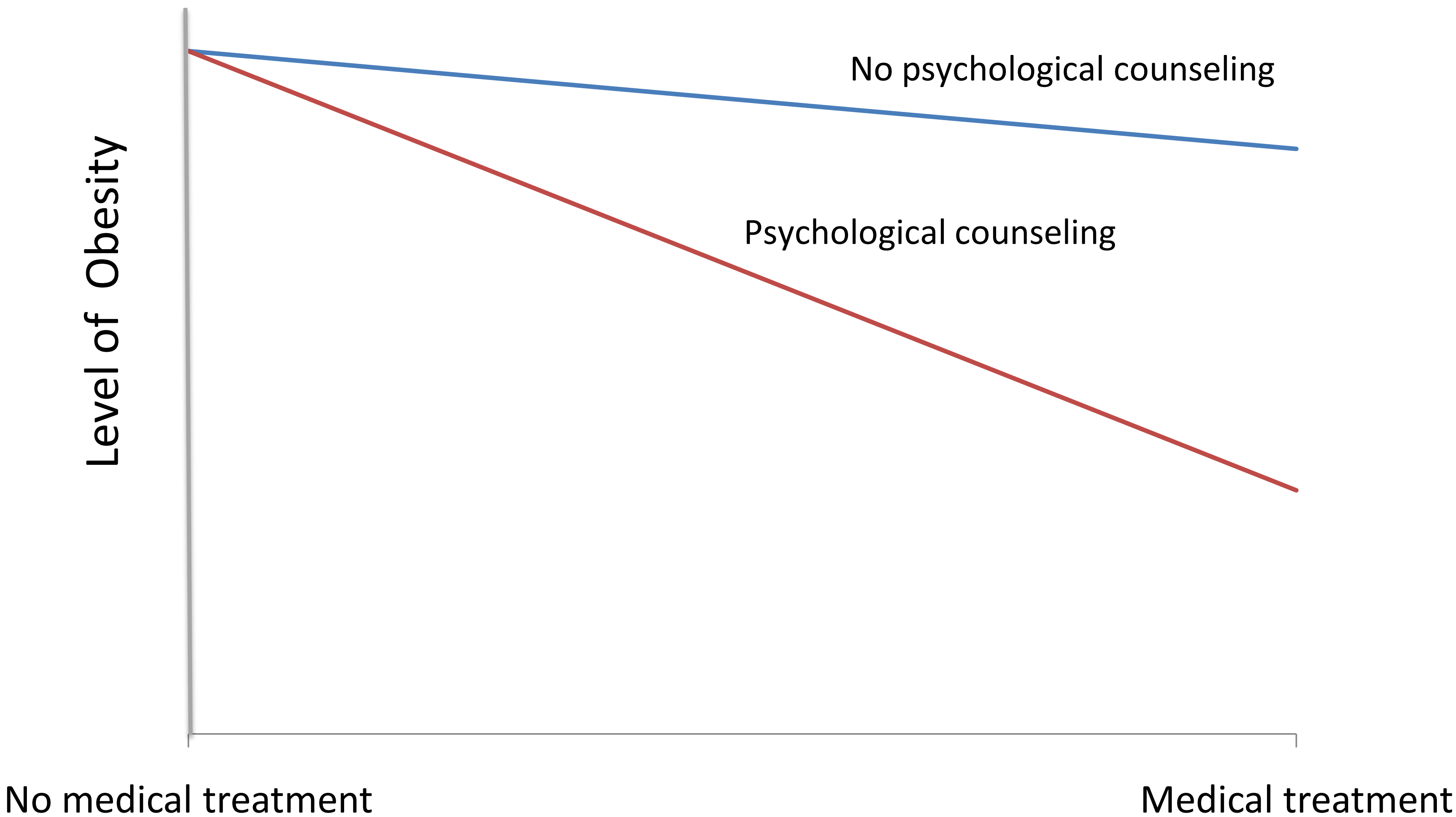
Interactive Relationships

- The relationship between the IV (X) and DV (Y) depends on the value of the control variable (Z)




Interactive Relationship between Partisanship and Support for Childcare
80
70
60
50
40
30
20
10
0
% favoring support (Y)
Women
Men
Democrat
Republican
Partisanship (X)

Interactive Relationship between Medical, Psychological treatment, and Obesity



Sources

- Philip H. Pollock III. 2009. *The Essentials of Political Analysis.* CQ Press.
- W. Philips Shively. *The Craft of Political Research*. Pearson Prentice Hall.