A Crash Course in Good and Bad Controls

Carlos Cinelli (UW), Andrew Forney (LMU), and Judea Pearl (UCLA)

A Crash Course in Good and Bad Controls

30 Pages • Posted: 29 Oct 2020 • Last revised: 21 Mar 2022

Carlos Cinelli
University of Washington - Department of Statistics

Andrew Forney
Loyola Marymount University

Judea Pearl
University of California, Los Angeles (UCLA) - Computer Science Department

A Crash Course in Good and Bad Controls

Carlos Cinelli (UW), Andrew Forney (LMU), and Judea Pearl (UCLA)

Have you controlled for…

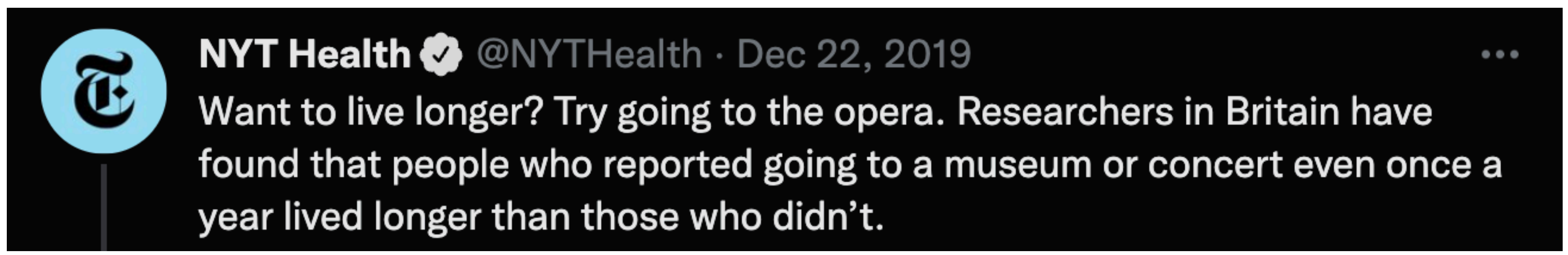
Have you “controlled” for…

Have you "controlled" for…

Many (if not most) regression studies aim to identify a causal effect between one variable in another (even if implicitly). For example:

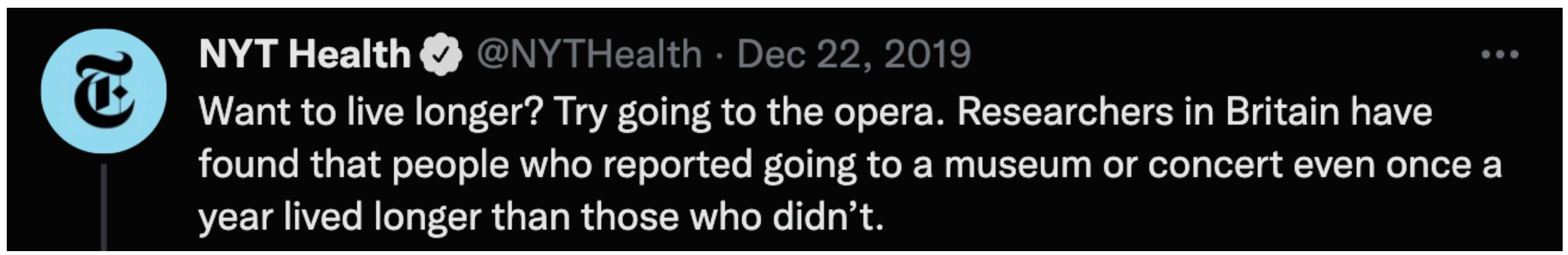
Have you "controlled" for…

Many (if not most) regression studies aim to identify a causal effect between one variable in another (even if implicitly). For example:



Have you "controlled" for…

Many (if not most) regression studies aim to identify a causal effect between one variable in another (even if implicitly). For example:



However, we all know that association is not causation, and observational studies are always under the threat of unobserved "confounders."

Have you "controlled" for…

Many (if not most) regression studies aim to identify a causal effect between one variable in another (even if implicitly). For example:

> **NYT Health** @NYTHealth · Dec 22, 2019
> Want to live longer? Try going to the opera. Researchers in Britain have found that people who reported going to a museum or concert even once a year lived longer than those who didn't.

However, we all know that association is not causation, and observational studies are always under the threat of unobserved "confounders."

Thus, one of the first things applied researchers learn in traditional regression courses is that you should add "controls" to your regression equation, in order to avoid *omitted variable bias.*

Have you "controlled" for…

Many (if not most) regression studies aim to identify a causal effect between one variable in another (even if implicitly). For example:

> **NYT Health** @NYTHealth · Dec 22, 2019
> Want to live longer? Try going to the opera. Researchers in Britain have found that people who reported going to a museum or concert even once a year lived longer than those who didn't.

However, we all know that association is not causation, and observational studies are always under the threat of unobserved "confounders."

Thus, one of the first things applied researchers learn in traditional regression courses is that you should add "controls" to your regression equation, in order to avoid *omitted variable bias.*

> The study controlled for socioeconomic factors like a participant's income, education level and mobility, said Andrew Steptoe, a co-author of the study and the head of University College London's research department of behavioral science and health.

Not all controls are created equal…

Researchers trained only in a traditional regression course may thus be left under the impression that we should always control for everything we measure. However not all controls are created equal.

Not all controls are created equal…

Researchers trained only in a traditional regression course may thus be left under the impression that we should always control for everything we measure. However not all controls are created equal.

Analysts have long known that some variables, when added to the regression equation, can in fact *increase bias*. Such variables have are known as “bad controls” in the econometrics literature.

Not all controls are created equal…

Researchers trained only in a traditional regression course may thus be left under the impression that we should always control for everything we measure. However not all controls are created equal.

Analysts have long known that some variables, when added to the regression equation, can in fact *increase bias*. Such variables have are known as “bad controls” in the econometrics literature.

The typical warning you find in textbooks is that “bad controls” are variables that could be “affected by the treatment” (i.e, “post-treatment” variables) [MHE, p.64]:

Not all controls are created equal…

Researchers trained only in a traditional regression course may thus be left under the impression that we should always control for everything we measure. However not all controls are created equal.

Analysts have long known that some variables, when added to the regression equation, can in fact *increase bias*. Such variables have are known as "bad controls" in the econometrics literature.

The typical warning you find in textbooks is that "bad controls" are variables that could be "affected by the treatment" (i.e, "post-treatment" variables) [MHE, p.64]:

> Some variables are bad controls and should not be included in a regression model, even when their inclusion might be expected to change the short regression coefficients. Bad controls are variables that are themselves outcome variables in the notional experiment at hand. That is, bad controls might just as well be dependent variables too. Good controls are variables that we can think of having been fixed at the time the regressor of interest was determined.

Not all controls are created equal…

Researchers trained only in a traditional regression course may thus be left under the impression that we should always control for everything we measure. However not all controls are created equal.

Analysts have long known that some variables, when added to the regression equation, can in fact *increase bias*. Such variables have are known as "bad controls" in the econometrics literature.

The typical warning you find in textbooks is that "bad controls" are variables that could be "affected by the treatment" (i.e, "post-treatment" variables) [MHE, p.64]:

> Some variables are bad controls and should not be included in a regression model, even when their inclusion might be expected to change the short regression coefficients. Bad controls are variables that are themselves outcome variables in the notional experiment at hand. That is, bad controls might just as well be dependent variables too. Good controls are variables that we can think of having been fixed at the time the regressor of interest was determined.

Although an improvement, it turns out these conditions are *neither necessary nor sufficient* for deciding whether a variable is a good or bad control.

Not all controls are created equal…

Researchers trained only in a traditional regression course may thus be left under the impression that we should always control for everything we measure. However not all controls are created equal.

Analysts have long known that some variables, when added to the regression equation, can in fact *increase bias*. Such variables have are known as "bad controls" in the econometrics literature.

The typical warning you find in textbooks is that "bad controls" are variables that could be "affected by the treatment" (i.e, "post-treatment" variables) [MHE, p.64]:

> Some variables are bad controls and should not be included in a regression model, even when their inclusion might be expected to change the short regression coefficients. Bad controls are variables that are themselves outcome variables in the notional experiment at hand. That is, bad controls might just as well be dependent variables too. Good controls are variables that we can think of having been fixed at the time the regressor of interest was determined.

Although an improvement, it turns out these conditions are *neither necessary nor sufficient* for deciding whether a variable is a good or bad control.

What now?

Making sense of "controls," with graphs.

Recent advances in graphical models have produced simple criteria to distinguish "good" from "bad" controls.

Making sense of "controls," with graphs.

Recent advances in graphical models have produced simple criteria to distinguish "good" from "bad" controls.

These criteria allow us to:

Making sense of "controls," with graphs.

Recent advances in graphical models have produced simple criteria to distinguish "good" from "bad" controls.

These criteria allow us to:

1. Decide which variables should be included in the regression equation in order to *identify* the causal effect of interest;

Making sense of "controls," with graphs.

Recent advances in graphical models have produced simple criteria to distinguish "good" from "bad" controls.

These criteria allow us to:

1. Decide which variables should be included in the regression equation in order to *identify* the causal effect of interest;

2. Decide, among a set of valid adjustment sets, which variables would yield more *precise* estimates.

Making sense of "controls," with graphs.

Recent advances in graphical models have produced simple criteria to distinguish "good" from "bad" controls.

These criteria allow us to:

1. Decide which variables should be included in the regression equation in order to *identify* the causal effect of interest;

2. Decide, among a set of valid adjustment sets, which variables would yield more *precise* estimates.

Our goal today is to learn the main lessons of these general graphical criteria *through simple examples.*

Making sense of "controls," with graphs.

Recent advances in graphical models have produced simple criteria to distinguish "good" from "bad" controls.

These criteria allow us to:

1. Decide which variables should be included in the regression equation in order to *identify* the causal effect of interest;

2. Decide, among a set of valid adjustment sets, which variables would yield more *precise* estimates.

Our goal today is to learn the main lessons of these general graphical criteria *through simple examples.*

We will see how causal diagrams can make otherwise difficult problems very easy to solve, *by – literally – simple inspection of a diagram.*

Can you give plausible explanations to these paradoxes?

Can you give plausible explanations to these paradoxes?

Birth-weight paradox: Infants born to smokers were found to have higher risks of mortality than infants born to non-smokers. However, among infants with low birth-weight (LBW), this relationship was ***reversed***. This reversal of effects has created many controversies in epidemiology—does it mean that maternal smoking is beneficial for LBW infants?

Can you give plausible explanations to these paradoxes?

Birth-weight paradox: Infants born to smokers were found to have higher risks of mortality than infants born to non-smokers. However, among infants with low birth-weight (LBW), this relationship was ***reversed***. This reversal of effects has created many controversies in epidemiology—does it mean that maternal smoking is beneficial for LBW infants?

Antebellum puzzle: An interesting puzzle of economic history is the fact that, during the nineteenth century in Britain and the United States, the average height of adult men fell even though the economic conditions of these countries improved alongside childhood nutrition. Does that mean better nutrition reduce the heights of adult men?

Can you give plausible explanations to these paradoxes?

Birth-weight paradox: Infants born to smokers were found to have higher risks of mortality than infants born to non-smokers. However, among infants with low birth-weight (LBW), this relationship was ***reversed***. This reversal of effects has created many controversies in epidemiology—does it mean that maternal smoking is beneficial for LBW infants?

Antebellum puzzle: An interesting puzzle of economic history is the fact that, during the nineteenth century in Britain and the United States, the average height of adult men fell even though the economic conditions of these countries improved alongside childhood nutrition. Does that mean better nutrition reduce the heights of adult men?

What is happening here?

Can you give plausible explanations to these paradoxes?

Birth-weight paradox: Infants born to smokers were found to have higher risks of mortality than infants born to non-smokers. However, among infants with low birth-weight (LBW), this relationship was ***reversed***. This reversal of effects has created many controversies in epidemiology—does it mean that maternal smoking is beneficial for LBW infants?

Antebellum puzzle: An interesting puzzle of economic history is the fact that, during the nineteenth century in Britain and the United States, the average height of adult men fell even though the economic conditions of these countries improved alongside childhood nutrition. Does that mean better nutrition reduce the heights of adult men?

What is happening here?

These are all related to "bad controls."

Preliminaries – Causal Diagrams

Causal Diagrams

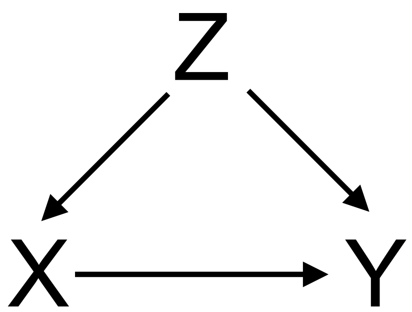
Causal Diagrams

Causal diagrams have become popular in the social and health sciences for explaining and resolving difficult problems in causal inference in a rigorous, yet accessible manner.

Causal Diagrams

Causal diagrams have become popular in the social and health sciences for explaining and resolving difficult problems in causal inference in a rigorous, yet accessible manner.

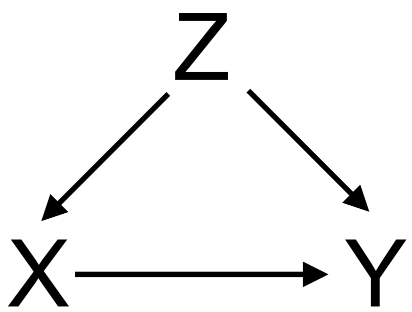
Briefly causal diagrams provide a *parsimonious representation* of the qualitative aspects of the data generating process. For example,



Causal Diagrams

Causal diagrams have become popular in the social and health sciences for explaining and resolving difficult problems in causal inference in a rigorous, yet accessible manner.

Briefly causal diagrams provide a *parsimonious representation* of the qualitative aspects of the data generating process. For example,

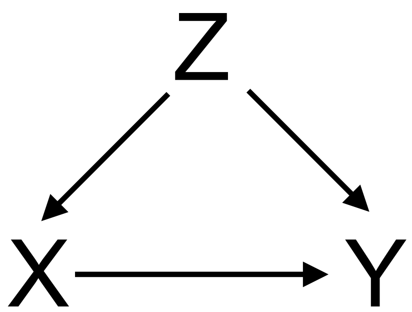


1. Here letters represent random variables (eg., X=drug, Z=income, Y=health).

Causal Diagrams

Causal diagrams have become popular in the social and health sciences for explaining and resolving difficult problems in causal inference in a rigorous, yet accessible manner.

Briefly causal diagrams provide a *parsimonious representation* of the qualitative aspects of the data generating process. For example,



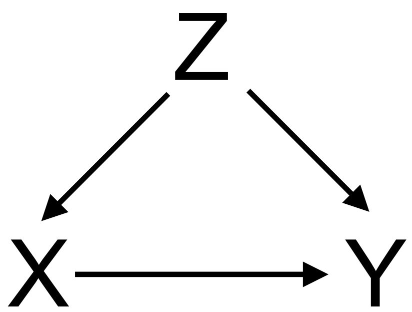
1. Here letters represent random variables (eg., X=drug, Z=income, Y=health).

2. Arrows denote a (possible) direct causal effect between one variable on another. For instance, the arrows X→Y and Z→Y state that both the drug and income could possibly affect health.

Causal Diagrams

Causal diagrams have become popular in the social and health sciences for explaining and resolving difficult problems in causal inference in a rigorous, yet accessible manner.

Briefly causal diagrams provide a *parsimonious representation* of the qualitative aspects of the data generating process. For example,



1. Here letters represent random variables (eg., X=drug, Z=income, Y=health).

2. Arrows denote a (possible) direct causal effect between one variable on another. For instance, the arrows X→Y and Z→Y state that both the drug and income could possibly affect health.

Note that *no parametric assumptions* need to be made regarding the functional form of the causal relationships, nor the distribution of variables.

Building blocks of a causal diagram

Building blocks of a causal diagram

Any causal diagram, no matter how complicated, can be understood in terms of three main sources of association: (1) mediators, (2) common causes and (3) common effects. They form the *building blocks* of any causal model.

Building blocks of a causal diagram

Any causal diagram, no matter how complicated, can be understood in terms of three main sources of association: (1) mediators, (2) common causes and (3) common effects. They form the *building blocks* of any causal model.

1. **Mediators:** X causally affects Z which causally affects Y.

$$X \longrightarrow Z \longrightarrow Y$$

Building blocks of a causal diagram

Any causal diagram, no matter how complicated, can be understood in terms of three main sources of association: (1) <u>mediators</u>, (2) <u>common causes</u> and (3) <u>common effects</u>. They form the *building blocks* of any causal model.

1. **<u>Mediators:</u>** X causally affects Z which causally affects Y.

$$X \longrightarrow Z \longrightarrow Y$$

2. **<u>Common causes:</u>** X and Y share a common cause Z (aka confounder).

$$X \longleftarrow Z \longrightarrow Y$$

Building blocks of a causal diagram

Any causal diagram, no matter how complicated, can be understood in terms of three main sources of association: (1) mediators, (2) common causes and (3) common effects. They form the *building blocks* of any causal model.

1. **Mediators:** X causally affects Z which causally affects Y.

$$X \longrightarrow Z \longrightarrow Y$$

2. **Common causes:** X and Y share a common cause Z (aka confounder).

$$X \longleftarrow Z \longrightarrow Y$$

3. **Common effects:** X and Y share a common effect Z (aka a collider).

$$X \longrightarrow Z \longleftarrow Y$$

Building blocks of a causal diagram

Any causal diagram, no matter how complicated, can be understood in terms of three main sources of association: (1) mediators, (2) common causes and (3) common effects. They form the *building blocks* of any causal model.

1. **Mediators:** X causally affects Z which causally affects Y.

$$X \longrightarrow Z \longrightarrow Y$$

2. **Common causes:** X and Y share a common cause Z (aka confounder).

$$X \longleftarrow Z \longrightarrow Y$$

3. **Common effects:** X and Y share a common effect Z (aka a collider).

$$X \longrightarrow Z \longleftarrow Y$$

Let us understand better each of these forms of association, and when they are closed or opened.

Building blocks of a causal diagram

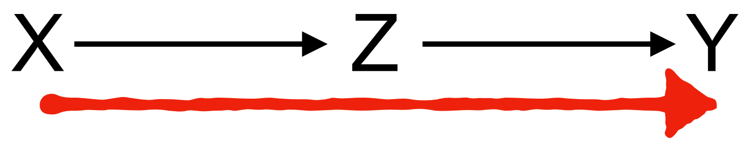
1. <u>Mediators:</u> X *causally* affects Y through Z. This is a <u>*causal*</u> path from X to Y. If left untouched, the path is open.

Building blocks of a causal diagram

1. <u>Mediators:</u> X *causally* affects Y through Z. This is a <u>*causal*</u> path from X to Y. If left untouched, the path is open.

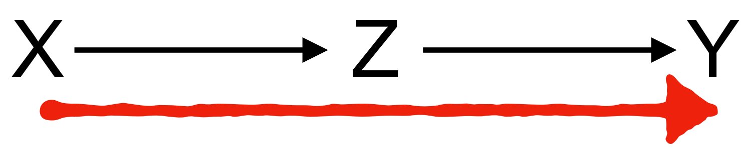
Building blocks of a causal diagram

1. <u>Mediators:</u> X *causally* affects Y through Z. This is a <u>*causal*</u> path from X to Y. If left untouched, the path is open.

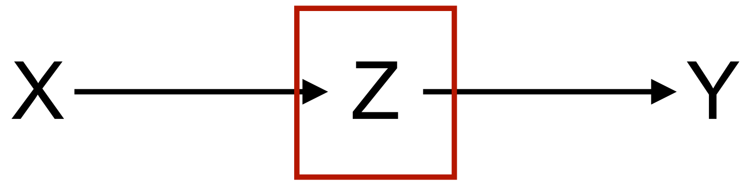


Building blocks of a causal diagram

1. <u>Mediators:</u> X *causally* affects Y through Z. This is a <u>*causal*</u> path from X to Y. If left untouched, the path is open.

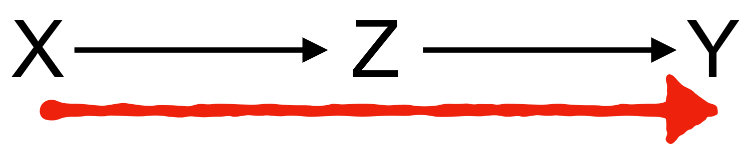


Conditioning on the mediator Z <u>*blocks*</u> the flow of association.

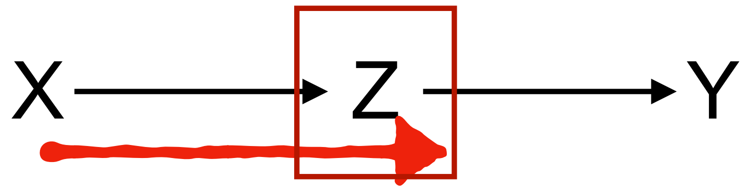


Building blocks of a causal diagram

1. <u>Mediators:</u> X *causally* affects Y through Z. This is a <u>*causal*</u> path from X to Y. If left untouched, the path is open.

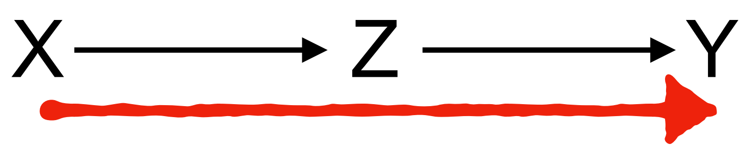


Conditioning on the mediator Z <u>*blocks*</u> the flow of association.

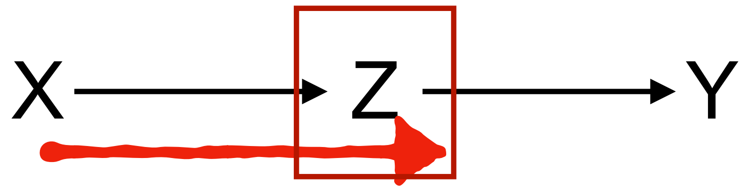


Building blocks of a causal diagram

1. Mediators: X *causally* affects Y through Z. This is a *causal* path from X to Y. If left untouched, the path is open.



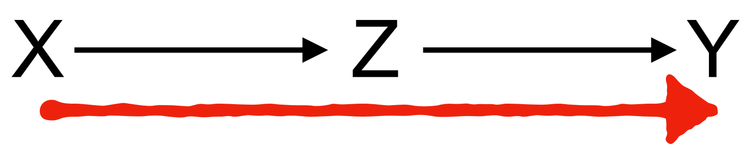
Conditioning on the mediator Z *blocks* the flow of association.



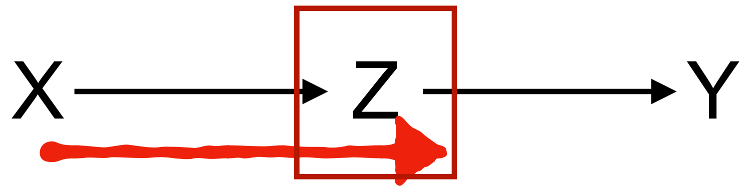
Example: Consider a drug (X) that affects a health outcome (Y) by lowering blood pressure (Z).

Building blocks of a causal diagram

1. <u>Mediators:</u> X *causally* affects Y through Z. This is a <u>*causal*</u> path from X to Y. If left untouched, the path is open.



Conditioning on the mediator Z <u>*blocks*</u> the flow of association.



<u>Example</u>: Consider a drug (X) that affects a health outcome (Y) by lowering blood pressure (Z).

Conditioning on blood pressure <u>blocks</u> the mechanism through which the drug affects health. Thus you will not see any association between drug use and health status among those with the same level of blood pressure.

Building blocks of a causal diagram

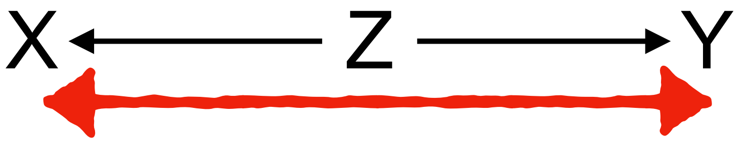
2. <u>Common causes:</u> X and Y share a common cause Z (aka confounder, or "back-door" path). Left on its own, it is open, and it induces a <u>*non-causal*</u> (spurious) association between X and Y.

Building blocks of a causal diagram

2. **<u>Common causes:</u>** X and Y share a common cause Z (aka confounder, or "back-door" path). Left on its own, it is open, and it induces a ***<u>non-causal</u>*** (spurious) association between X and Y.

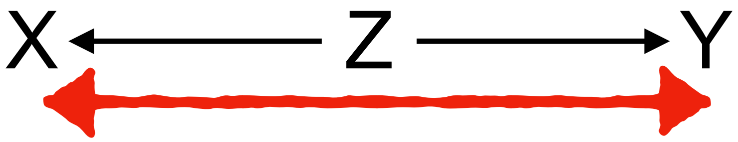
Building blocks of a causal diagram

2. <u>Common causes:</u> X and Y share a common cause Z (aka confounder, or "back-door" path). Left on its own, it is open, and it induces a *<u>non-causal</u>* (spurious) association between X and Y.

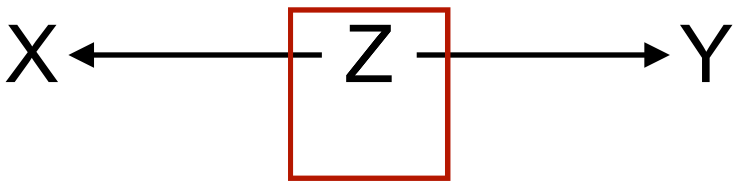


Building blocks of a causal diagram

2. **Common causes:** X and Y share a common cause Z (aka confounder, or "back-door" path). Left on its own, it is open, and it induces a *non-causal* (spurious) association between X and Y.

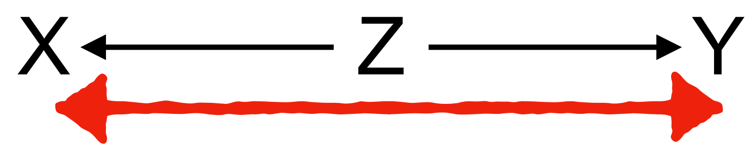


Conditioning on the common cause Z *blocks* the flow of association.

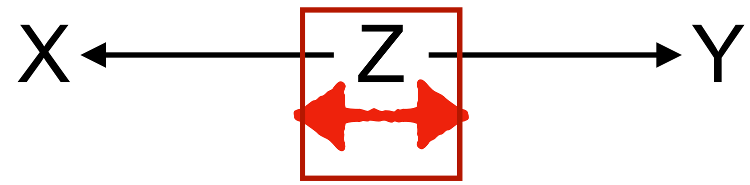


Building blocks of a causal diagram

2. <u>Common causes:</u> X and Y share a common cause Z (aka confounder, or "back-door" path). Left on its own, it is open, and it induces a *<u>non-causal</u>* (spurious) association between X and Y.

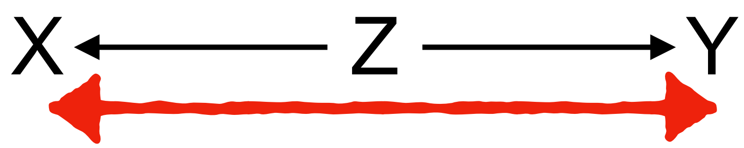


Conditioning on the common cause Z *<u>blocks</u>* the flow of association.

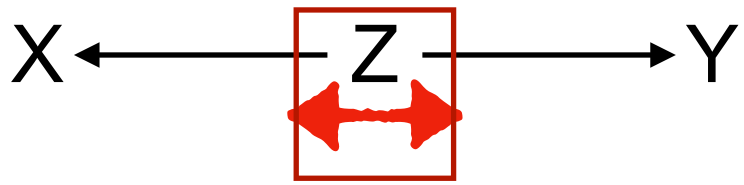


Building blocks of a causal diagram

2. <u>Common causes:</u> X and Y share a common cause Z (aka confounder, or "back-door" path). Left on its own, it is open, and it induces a *<u>non-causal</u>* (spurious) association between X and Y.



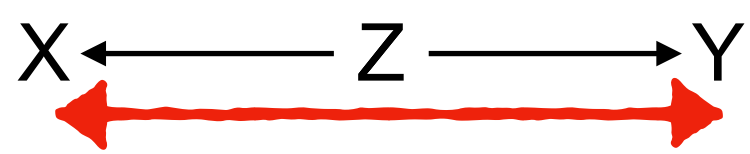
Conditioning on the common cause Z *<u>blocks</u>* the flow of association.



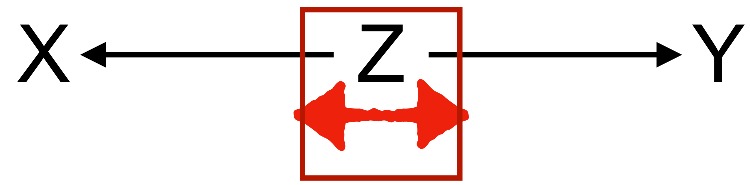
<u>Example</u>: Suppose that going to the opera (X) *does not* affect mortality (Y).

Building blocks of a causal diagram

2. <u>Common causes:</u> X and Y share a common cause Z (aka confounder, or "back-door" path). Left on its own, it is open, and it induces a *<u>non-causal</u>* (spurious) association between X and Y.



Conditioning on the common cause Z *<u>blocks</u>* the flow of association.

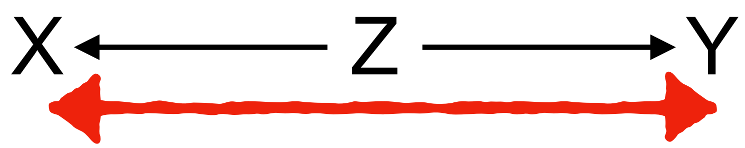


<u>Example</u>: Suppose that going to the opera (X) *does not* affect mortality (Y).

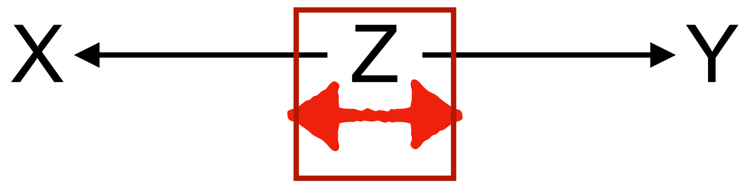
However, rich people (Z = income) are both more likely to go to the opera and also more likely to be healthy. This will induce a spurious association between X and Y.

Building blocks of a causal diagram

2. <u>Common causes:</u> X and Y share a common cause Z (aka confounder, or "back-door" path). Left on its own, it is open, and it induces a <u>*non-causal*</u> (spurious) association between X and Y.



Conditioning on the common cause Z <u>*blocks*</u> the flow of association.



<u>Example</u>: Suppose that going to the opera (X) *does not* affect mortality (Y).

However, rich people (Z = income) are both more likely to go to the opera and also more likely to be healthy. This will induce a spurious association between X and Y.

Conditioning on income (Z) <u>blocks</u> this spurious association.

Building blocks of a causal diagram

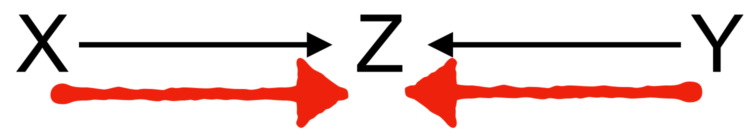
3. **<u>Common effects:</u>** X and Y share a common effect Z (aka a collider). Left alone, a common effect does not induce any association between its causes, and the path is *<u>closed</u>*.

Building blocks of a causal diagram

3. Common effects: X and Y share a common effect Z (aka a collider). Left alone, a common effect does not induce any association between its causes, and the path is *closed*.

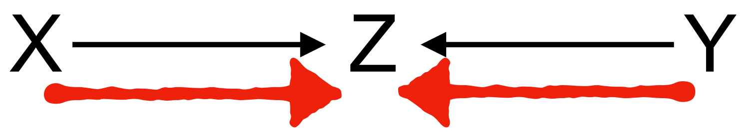
Building blocks of a causal diagram

3. **Common effects:** X and Y share a common effect Z (aka a collider). Left alone, a common effect does not induce any association between its causes, and the path is *closed*.

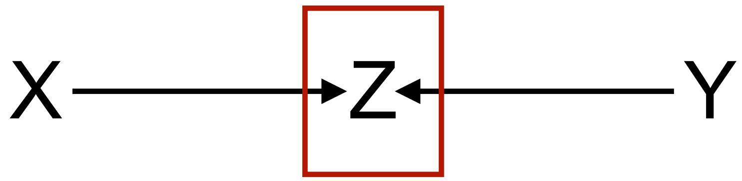


Building blocks of a causal diagram

3. **Common effects:** X and Y share a common effect Z (aka a collider). Left alone, a common effect does not induce any association between its causes, and the path is *closed*.

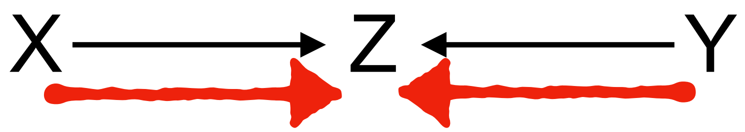


However, conditioning on the common effect *opens* the path, and it induces a *non-causal* association between X and Y.

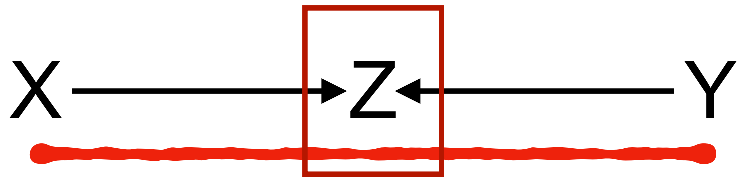


Building blocks of a causal diagram

3. **<u>Common effects:</u>** X and Y share a common effect Z (aka a collider). Left alone, a common effect does not induce any association between its causes, and the path is *<u>closed</u>*.

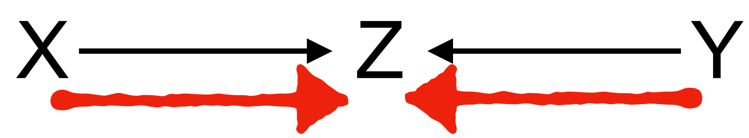


However, conditioning on the common effect *<u>opens</u>* the path, and it induces a *<u>non-causal</u>* association between X and Y.

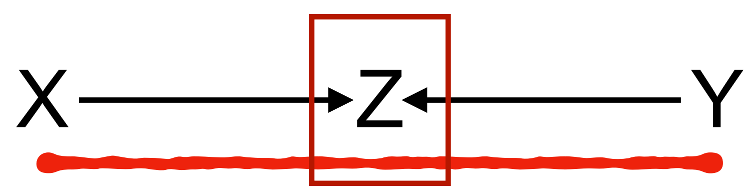


Building blocks of a causal diagram

3. **Common effects:** X and Y share a common effect Z (aka a collider). Left alone, a common effect does not induce any association between its causes, and the path is *closed*.



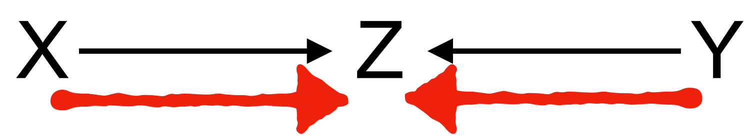
However, conditioning on the common effect *opens* the path, and it induces a *non-causal* association between X and Y.



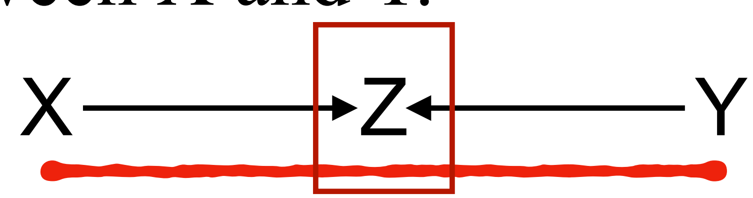
Example: Imagine beauty (X) and talent (Y) are independent in the general population. However, suppose that movie agencies only hire (Z=1) actors whose beauty + talent exceed a certain threshold.

Building blocks of a causal diagram

3. **Common effects:** X and Y share a common effect Z (aka a collider). Left alone, a common effect does not induce any association between its causes, and the path is *closed*.



However, conditioning on the common effect *opens* the path, and it induces a *non-causal* association between X and Y.



Example: Imagine beauty (X) and talent (Y) are independent in the general population. However, suppose that movie agencies only hire (Z=1) actors whose beauty + talent exceed a certain threshold.

Even though there's no causal relationship between beauty and talent in the general population, you will see a *negative* association between beauty and talent *both among hired actors (Z=1) and not hired actors (Z=0)*.

Building blocks of a causal diagram

Since this is the most unintuitive form of association, let's see the example in practice via simulation.

Building blocks of a causal diagram

Since this is the most unintuitive form of association, let's see the example in practice via simulation.

$$X \longrightarrow Z \longleftarrow Y$$

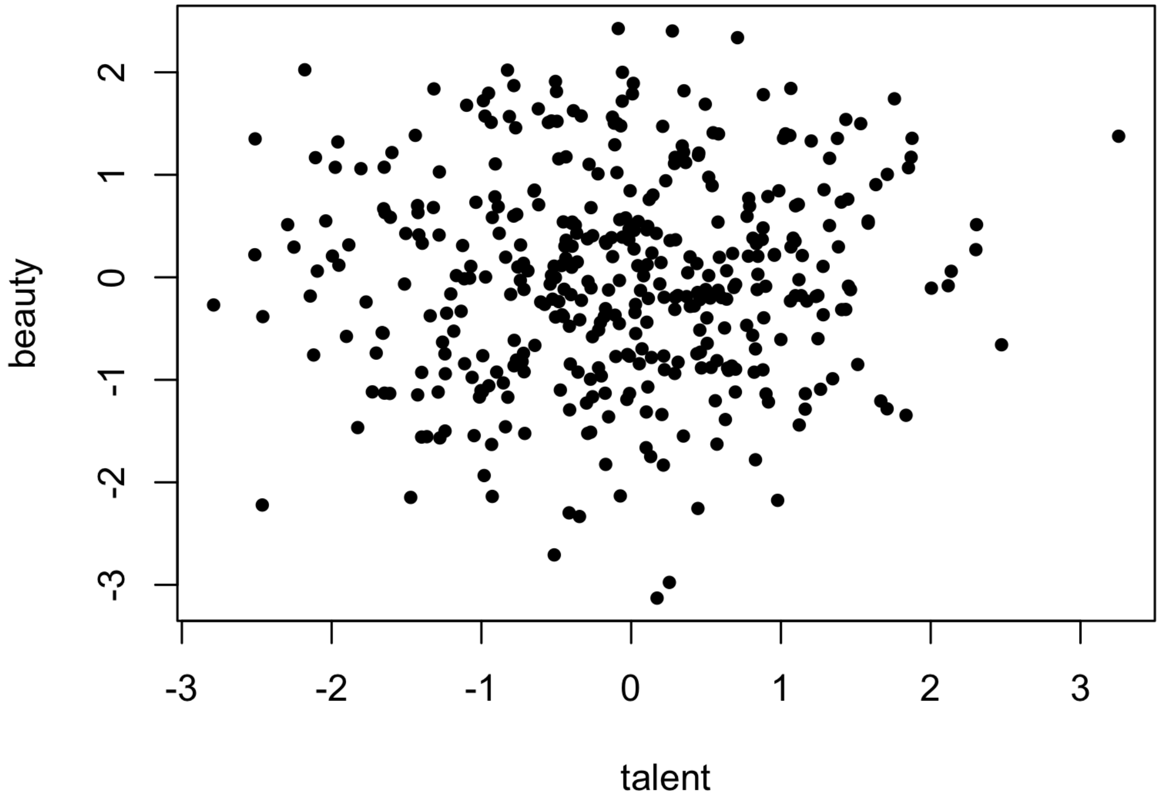
```
beauty <- rnorm(n)
talent <- rnorm(n)
hire <- (beauty + talent > 0)
```

Building blocks of a causal diagram

Since this is the most unintuitive form of association, let's see the example in practice via simulation.

$$X \longrightarrow Z \longleftarrow Y$$

```r
beauty <- rnorm(n)
talent <- rnorm(n)
hire <- (beauty + talent > 0)
```



General Population

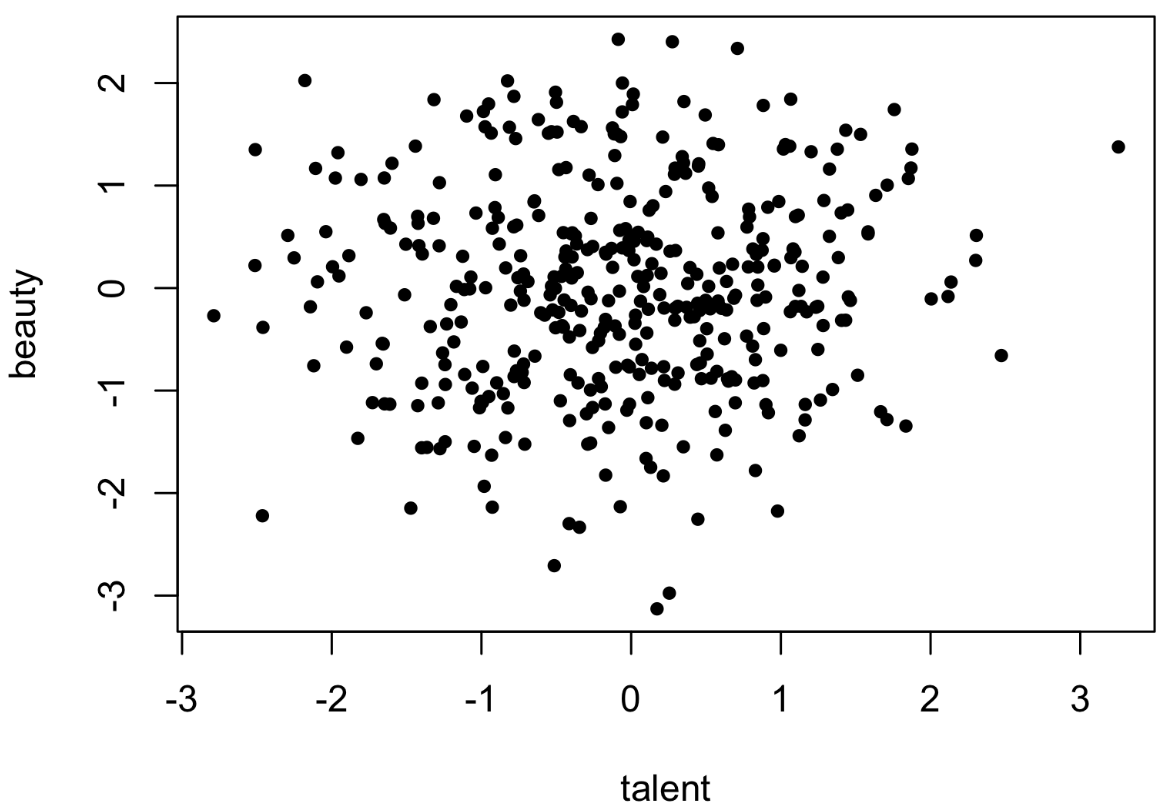
(no association)

Building blocks of a causal diagram

Since this is the most unintuitive form of association, let's see the example in practice via simulation.

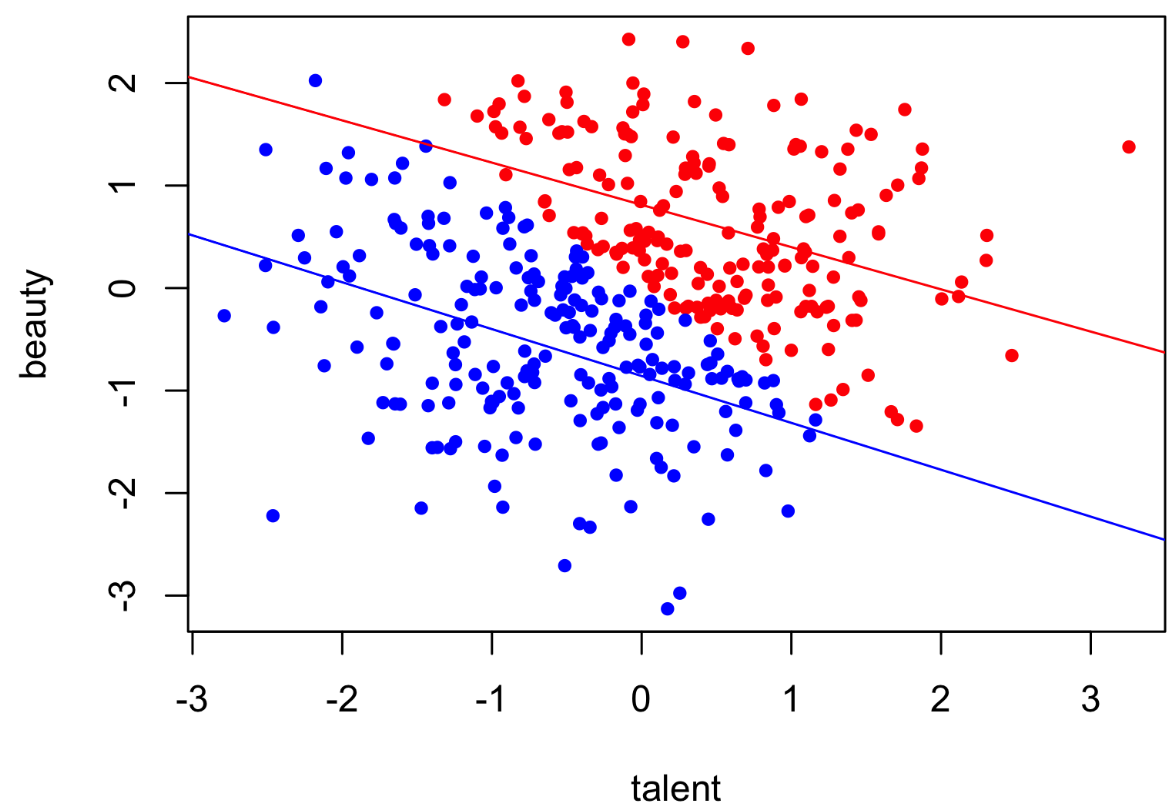
X ⟶ Z ⟵ Y

```
beauty <- rnorm(n)
talent <- rnorm(n)
hire <- (beauty + talent > 0)
```



General Population

(no association)



Within Z=1 or Z=0

(negative association)

Building blocks of a causal diagram

To summarize:

Path		Type	Not conditioning on Z	Conditioning on Z
Mediators	X ⟶ Z ⟶ Y	Causal	Open	Closed
Common Causes (Confounders)	X ⟵ Z ⟶ Y	Non-Causal (Spurious)	Open	Closed
Common Effects (Colliders)	X ⟶ Z ⟵ Y	Non-Causal (Spurious)	Closed	Open

Building blocks of a causal diagram

To summarize:

Path		Type	Not conditioning on Z	Conditioning on Z
Mediators	$X \rightarrow Z \rightarrow Y$	<u>Causal</u>	Open	Closed
Common Causes (Confounders)	$X \leftarrow Z \rightarrow Y$	<u>Non-Causal</u> (Spurious)	Open	Closed
Common Effects (Colliders)	$X \rightarrow Z \leftarrow Y$	<u>Non-Causal</u> (Spurious)	Closed	Open

Building blocks of a causal diagram

To summarize:

Path		Type	Not conditioning on Z	Conditioning on Z
Mediators	X ⟶ Z ⟶ Y	Causal	Open	Closed
Common Causes (Confounders)	X ⟵ Z ⟶ Y	Non-Causal (Spurious)	Open	Closed
Common Effects (Colliders)	X ⟶ Z ⟵ Y	Non-Causal (Spurious)	Closed	Open

Building blocks of a causal diagram

To summarize:

Path		Type	Not conditioning on Z	Conditioning on Z
Mediators	X ⟶ Z ⟶ Y	Causal	Open	Closed
Common Causes (Confounders)	X ⟵ Z ⟶ Y	Non-Causal (Spurious)	Open	Closed
Common Effects (Colliders)	X ⟶ Z ⟵ Y	Non-Causal (Spurious)	Closed	Open

Building blocks of a causal diagram

To summarize:

Path		Type	Not conditioning on Z	Conditioning on Z
Mediators	X ⟶ Z ⟶ Y	Causal	Open	Closed
Common Causes (Confounders)	X ⟵ Z ⟶ Y	Non-Causal (Spurious)	Open	Closed
Common Effects (Colliders)	X ⟶ Z ⟵ Y	Non-Causal (Spurious)	Closed	Open

Building blocks of a causal diagram

To summarize:

Path		Type	Not conditioning on Z	Conditioning on Z
Mediators	X ⟶ Z ⟶ Y	Causal	Open	Closed
Common Causes (Confounders)	X ⟵ Z ⟶ Y	Non-Causal (Spurious)	Open	Closed
Common Effects (Colliders)	X ⟶ Z ⟵ Y	Non-Causal (Spurious)	Closed	Open

Building blocks of a causal diagram

To summarize:

Path		Type	Not conditioning on Z	Conditioning on Z
Mediators	X ⟶ Z ⟶ Y	Causal	Open	Closed
Common Causes (Confounders)	X ⟵ Z ⟶ Y	Non-Causal (Spurious)	Open	Closed
Common Effects (Colliders)	X ⟶ Z ⟵ Y	Non-Causal (Spurious)	Closed	Open

Building blocks of a causal diagram

To summarize:

Path		Type	Not conditioning on Z	Conditioning on Z
Mediators	X ⟶ Z ⟶ Y	Causal	Open	Closed
Common Causes (Confounders)	X ⟵ Z ⟶ Y	Non-Causal (Spurious)	Open	Closed
Common Effects (Colliders)	X ⟶ Z ⟵ Y	Non-Causal (Spurious)	Closed	Open

Building blocks of a causal diagram

To summarize:

Path		Type	Not conditioning on Z	Conditioning on Z
Mediators	X ⟶ Z ⟶ Y	Causal	Open	Closed
Common Causes (Confounders)	X ⟵ Z ⟶ Y	Non-Causal (Spurious)	Open	Closed
Common Effects (Colliders)	X ⟶ Z ⟵ Y	Non-Causal (Spurious)	Closed	Open

Building blocks of a causal diagram

To summarize:

Path		Type	Not conditioning on Z	Conditioning on Z
Mediators	X ⟶ Z ⟶ Y	Causal	Open	Closed
Common Causes (Confounders)	X ⟵ Z ⟶ Y	Non-Causal (Spurious)	Open	Closed
Common Effects (Colliders)	X ⟶ Z ⟵ Y	Non-Causal (Spurious)	Closed	Open

Building blocks of a causal diagram

To summarize:

Path		Type	Not conditioning on Z	Conditioning on Z
Mediators	X ⟶ Z ⟶ Y	Causal	Open	Closed
Common Causes (Confounders)	X ⟵ Z ⟶ Y	Non-Causal (Spurious)	Open	Closed
Common Effects (Colliders)	X ⟶ Z ⟵ Y	Non-Causal (Spurious)	Closed	Open

Building blocks of a causal diagram

To summarize:

Path		Type	Not conditioning on Z	Conditioning on Z
Mediators	X ⟶ Z ⟶ Y	<u>Causal</u>	Open	Closed
Common Causes (Confounders)	X ⟵ Z ⟶ Y	<u>Non-Causal</u> (Spurious)	Open	Closed
Common Effects (Colliders)	X ⟶ Z ⟵ Y	<u>Non-Causal</u> (Spurious)	Closed	Open

Building blocks of a causal diagram

To summarize:

Path		Type	Not conditioning on Z	Conditioning on Z
Mediators	X ⟶ Z ⟶ Y	Causal	Open	Closed
Common Causes (Confounders)	X ⟵ Z ⟶ Y	Non-Causal (Spurious)	Open	Closed
Common Effects (Colliders)	X ⟶ Z ⟵ Y	Non-Causal (Spurious)	Closed	Open

Building blocks of a causal diagram

To summarize:

Path		Type	Not conditioning on Z	Conditioning on Z
Mediators	X → Z → Y	Causal	Open	Closed
Common Causes (Confounders)	X ← Z → Y	Non-Causal (Spurious)	Open	Closed
Common Effects (Colliders)	X → Z ← Y	Non-Causal (Spurious)	Closed	Open

One final rule to keep in mind:

Building blocks of a causal diagram

To summarize:

Path		Type	Not conditioning on Z	Conditioning on Z
Mediators	X ⟶ Z ⟶ Y	Causal	Open	Closed
Common Causes (Confounders)	X ⟵ Z ⟶ Y	Non-Causal (Spurious)	Open	Closed
Common Effects (Colliders)	X ⟶ Z ⟵ Y	Non-Causal (Spurious)	Closed	Open

One final rule to keep in mind:

Controlling for the effects of a variable is equivalent to partially controlling for that variable.

Building blocks of a causal diagram

To summarize:

Path		Type	Not conditioning on Z	Conditioning on Z
Mediators	$X \rightarrow Z \rightarrow Y$	<u>Causal</u>	Open	Closed
Common Causes (Confounders)	$X \leftarrow Z \rightarrow Y$	<u>Non-Causal</u> (Spurious)	Open	Closed
Common Effects (Colliders)	$X \rightarrow Z \leftarrow Y$	<u>Non-Causal</u> (Spurious)	Closed	Open

One final rule to keep in mind:

Controlling for the effects of a variable is equivalent to <u>partially controlling</u> for that variable.

Most importantly for us:

Building blocks of a causal diagram

To summarize:

Path		Type	Not conditioning on Z	Conditioning on Z
Mediators	X ⟶ Z ⟶ Y	Causal	Open	Closed
Common Causes (Confounders)	X ⟵ Z ⟶ Y	Non-Causal (Spurious)	Open	Closed
Common Effects (Colliders)	X ⟶ Z ⟵ Y	Non-Causal (Spurious)	Closed	Open

One final rule to keep in mind:

Controlling for the effects of a variable is equivalent to partially controlling for that variable.

Most importantly for us:

(i) conditioning on the effects of a mediator partially closes a causal path;

Building blocks of a causal diagram

To summarize:

Path		Type	Not conditioning on Z	Conditioning on Z
Mediators	$X \rightarrow Z \rightarrow Y$	Causal	Open	Closed
Common Causes (Confounders)	$X \leftarrow Z \rightarrow Y$	Non-Causal (Spurious)	Open	Closed
Common Effects (Colliders)	$X \rightarrow Z \leftarrow Y$	Non-Causal (Spurious)	Closed	Open

One final rule to keep in mind:

Controlling for the effects of a variable is equivalent to partially controlling for that variable.

Most importantly for us:

(i) conditioning on the effects of a mediator partially closes a causal path;

(ii) conditioning on the effects of a collider partially opens a non-causal path.

Opening and closing paths.

We can now judge whether any path on a causal diagram, no matter how complicated, is opened or closed.

Opening and closing paths.

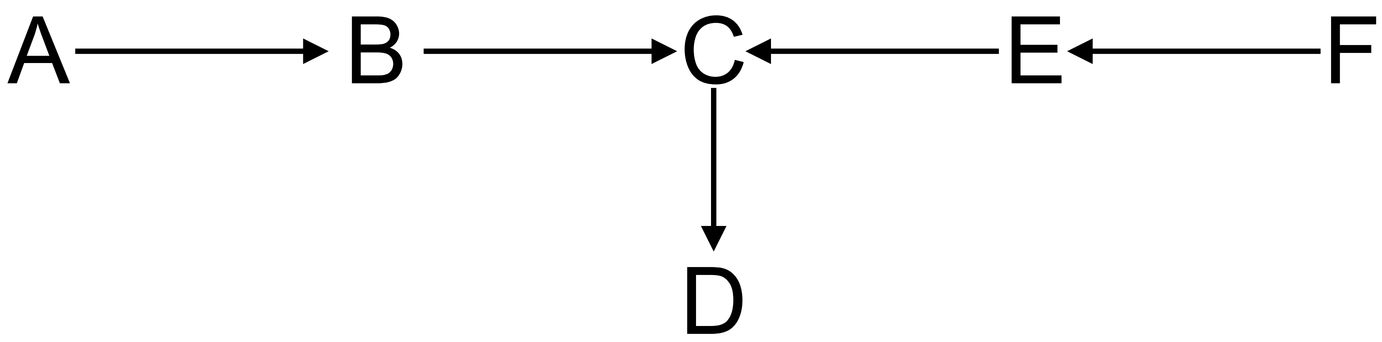
We can now judge whether any path on a causal diagram, no matter how complicated, is opened or closed.

A path is simply a sequence of mediators, common causes and common effects. If any of these elements is closed, the full path is closed.

Opening and closing paths.

We can now judge whether any path on a causal diagram, no matter how complicated, is opened or closed.

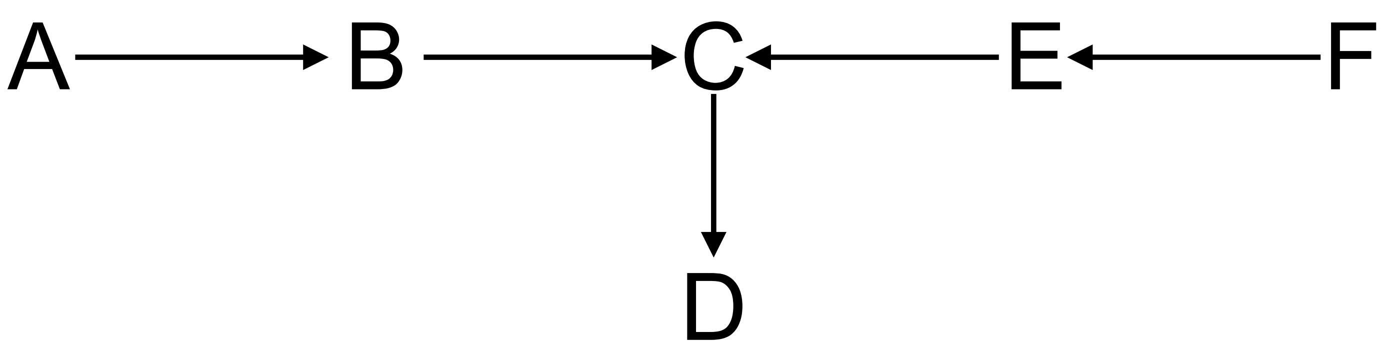
A path is simply a sequence of mediators, common causes and common effects. If any of these elements is closed, the full path is closed.



Opening and closing paths.

We can now judge whether any path on a causal diagram, no matter how complicated, is opened or closed.

A path is simply a sequence of mediators, common causes and common effects. If any of these elements is closed, the full path is closed.

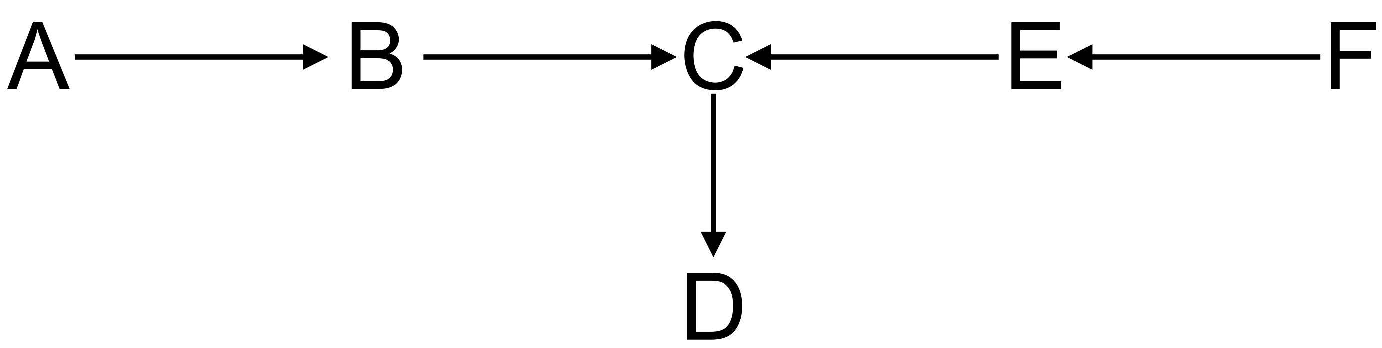


Is the path from A to F opened or closed:

Opening and closing paths.

We can now judge whether any path on a causal diagram, no matter how complicated, is opened or closed.

A path is simply a sequence of mediators, common causes and common effects. If any of these elements is closed, the full path is closed.



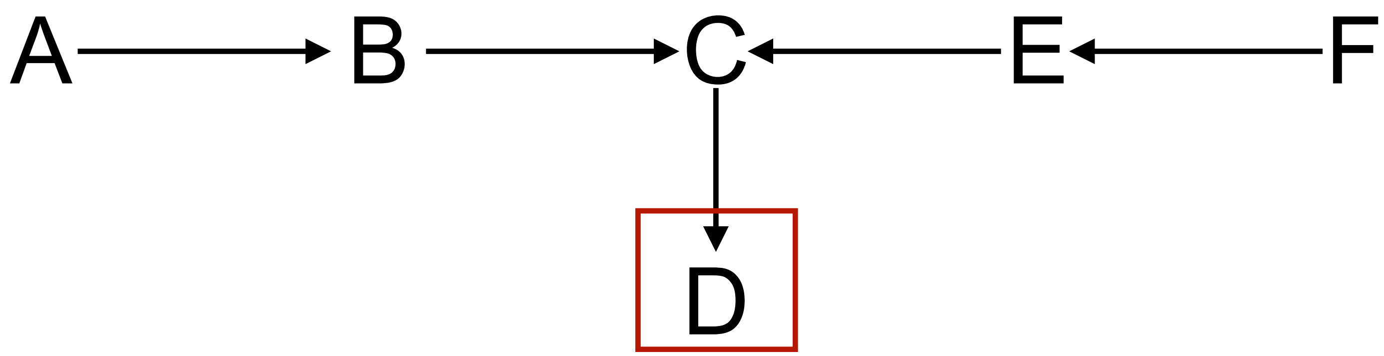
Is the path from A to F opened or closed:

a) conditioning on nothing?

Opening and closing paths.

We can now judge whether any path on a causal diagram, no matter how complicated, is opened or closed.

A path is simply a sequence of mediators, common causes and common effects. If any of these elements is closed, the full path is closed.



Is the path from A to F opened or closed:

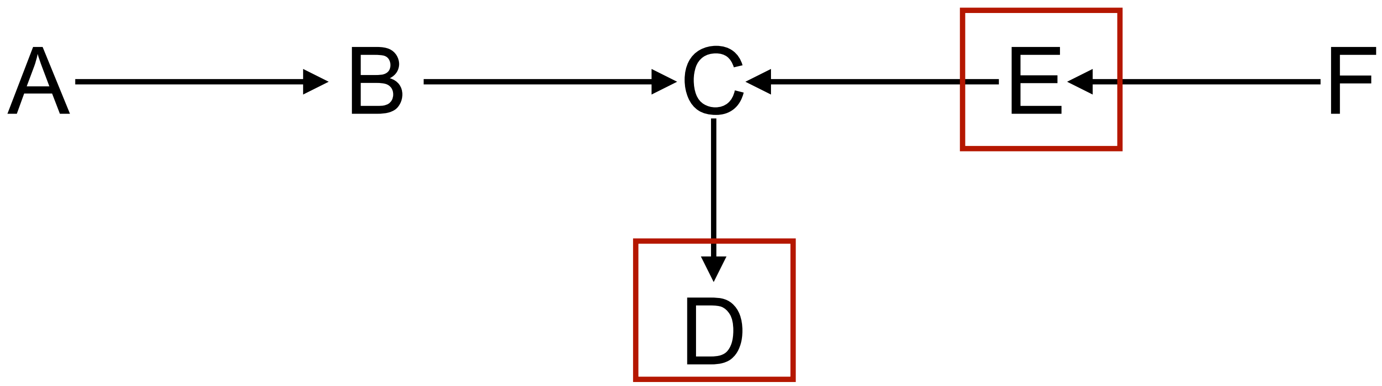
a) conditioning on nothing?

b) conditioning on D?

Opening and closing paths.

We can now judge whether any path on a causal diagram, no matter how complicated, is opened or closed.

A path is simply a sequence of mediators, common causes and common effects. If any of these elements is closed, the full path is closed.



Is the path from A to F opened or closed:

a) conditioning on nothing?

b) conditioning on D?

c) conditioning on D and E?

Blocking spurious paths, and only spurious paths.

Blocking spurious paths, and only spurious paths.

We now have all the tools we need to decide which variables to adjust for to identify the (total) causal effect of X on Y.

Blocking spurious paths, and only spurious paths.

We now have all the tools we need to decide which variables to adjust for to identify the (total) causal effect of X on Y.

First, note that causal paths are paths consisting of a sequence of mediators. All other paths are non-causal.

Blocking spurious paths, and only spurious paths.

We now have all the tools we need to decide which variables to adjust for to identify the (total) causal effect of X on Y.

First, note that causal paths are paths consisting of a sequence of mediators. All other paths are non-causal.

If we are interested, then, in estimating the causal effect of X on Y, our task is conceptually simple:

Blocking spurious paths, and only spurious paths.

We now have all the tools we need to decide which variables to adjust for to identify the (total) causal effect of X on Y.

First, note that causal paths are paths consisting of a sequence of mediators. All other paths are non-causal.

If we are interested, then, in estimating the causal effect of X on Y, our task is conceptually simple:

1. We must block all spurious paths between X and Y;
2. We must not perturb any of the (relevant) causal paths.

Blocking spurious paths, and only spurious paths.

We now have all the tools we need to decide which variables to adjust for to identify the (total) causal effect of X on Y.

First, note that causal paths are paths consisting of a sequence of mediators. All other paths are non-causal.

If we are interested, then, in estimating the causal effect of X on Y, our task is conceptually simple:

1. We must block all spurious paths between X and Y;
2. We must not perturb any of the (relevant) causal paths.

This is the essence of the graphical conditions known as the *back-door criterion* and the *adjustment criterion.* This will be our guiding principle to decide whether to include a variable to a regression equation.

Blocking spurious paths, and only spurious paths.

We now have all the tools we need to decide which variables to adjust for to identify the (total) causal effect of X on Y.

First, note that <u>causal paths</u> are paths consisting of a <u>sequence of mediators. All other paths are non-causal.</u>

If we are interested, then, in estimating the causal effect of X on Y, our task is conceptually simple:

1. We must <u>block all spurious paths</u> between X and Y;
2. We must <u>not perturb any of the (relevant) causal paths</u>.

This is the essence of the graphical conditions known as the *back-door criterion* and the *adjustment criterion.* This will be our guiding principle to decide whether to include a variable to a regression equation.

Although simple, mastering this does require some practice. So let's apply those principles in very simple examples.

Illustrative Examples

Illustrative examples

In the following set of models, the target of analysis is the *average treatment effect (ATE) of X on Y:*

Illustrative examples

In the following set of models, the target of analysis is the *average treatment effect (ATE) of X on Y:*

$$ATE = E[Y_1] - E[Y_0]$$

Illustrative examples

In the following set of models, the target of analysis is the *average treatment effect (ATE) of X on Y:*

$$ATE = E[Y_1] - E[Y_0]$$

Observed variables we be denoted by *black* dots, and *unobserved* variables by *white* empty circles.

Illustrative examples

In the following set of models, the target of analysis is the *average treatment effect (ATE) of X on Y:*

$$ATE = E[Y_1] - E[Y_0]$$

Observed variables we be denoted by *black* dots, and *unobserved* variables by *white* empty circles.

Variable Z, highlighted in **red**, will represent the variable whose inclusion in the regression equation is to be decided.

Illustrative examples

In the following set of models, the target of analysis is the *average treatment effect (ATE) of X on Y:*

$$ATE = E[Y_1] - E[Y_0]$$

Observed variables we be denoted by *black* dots, and *unobserved* variables by *white* empty circles.

Variable Z, highlighted in **red**, will represent the variable whose inclusion in the regression equation is to be decided.

We will say that Z is a *"good control"* if it helps reducing (asymptotic) bias, *"bad control"* if increases bias, and *"neutral control"* if the addition of Z neither reduces or increases bias.

Illustrative examples

In the following set of models, the target of analysis is the *average treatment effect (ATE) of X on Y:*

$$ATE = E[Y_1] - E[Y_0]$$

Observed variables we be denoted by *black* dots, and *unobserved* variables by *white* empty circles.

Variable Z, highlighted in **red**, will represent the variable whose inclusion in the regression equation is to be decided.

We will say that Z is a *"good control"* if it helps reducing (asymptotic) bias, *"bad control"* if increases bias, and *"neutral control"* if the addition of Z neither reduces or increases bias.

For this last case, we will also make brief remarks about how Z affects the precision of the ATE estimate.

Illustrative examples

In the following set of models, the target of analysis is the *average treatment effect (ATE) of X on Y:*

$$ATE = E[Y_1] - E[Y_0]$$

Observed variables we be denoted by *black* dots, and *unobserved* variables by *white* empty circles.

Variable Z, highlighted in **red**, will represent the variable whose inclusion in the regression equation is to be decided.

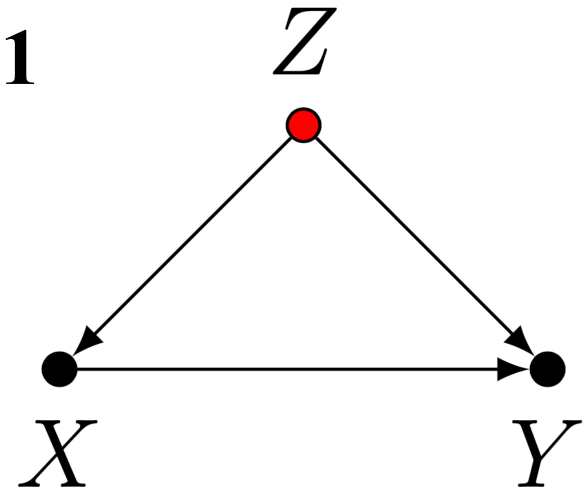
We will say that Z is a *"good control"* if it helps reducing (asymptotic) bias, *"bad control"* if increases bias, and *"neutral control"* if the addition of Z neither reduces or increases bias.

For this last case, we will also make brief remarks about how Z affects the precision of the ATE estimate.

Here will focus on *practicing our graphical skills*. Later we will see how these very simple models can help you make sense of real world scenarios.

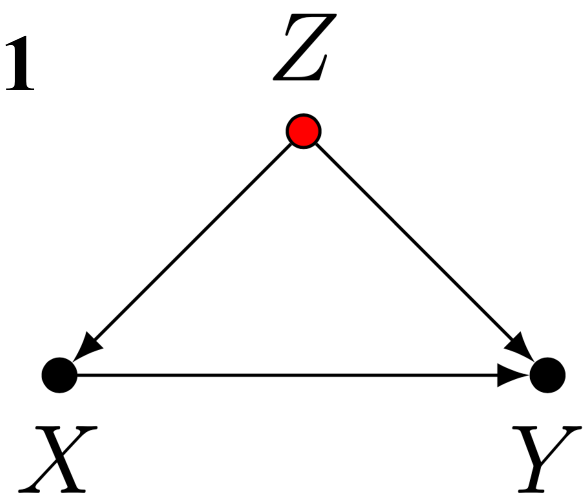
“Good” Controls – Blocking backdoor paths

Model 1



“Good” Controls – Blocking backdoor paths

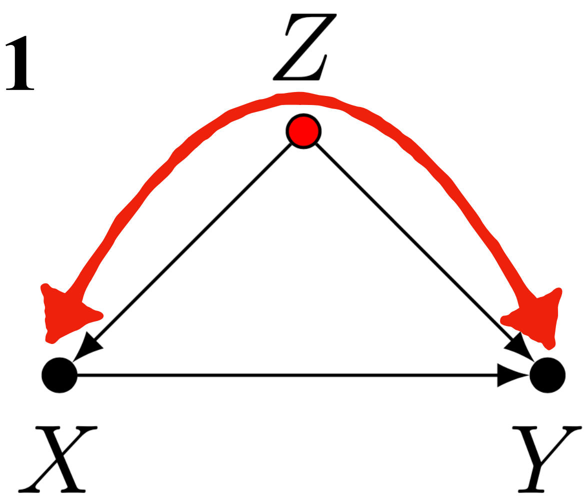
Model 1



Z is a confounder, and creates a spurious association between X and Y.

"Good" Controls – Blocking backdoor paths

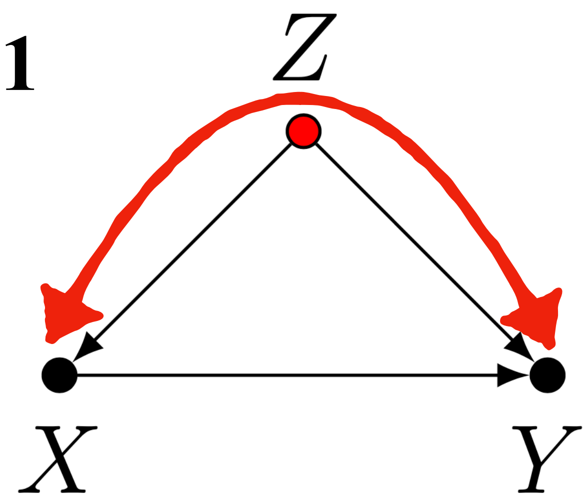
Model 1



Z is a confounder, and creates a spurious association between X and Y.

"Good" Controls – Blocking backdoor paths

Model 1

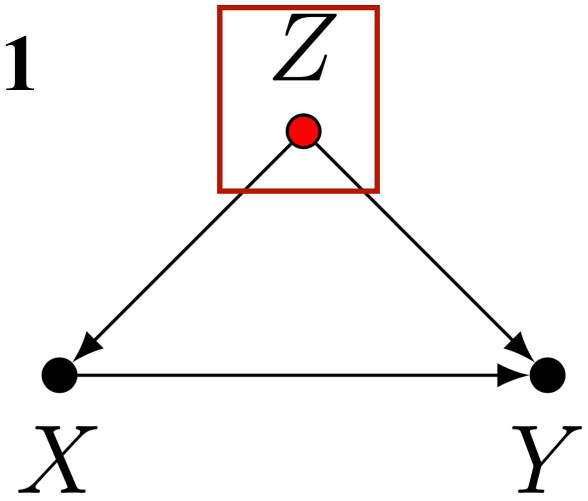


Z is a confounder, and creates a spurious association between X and Y.

Once we control for Z we block the backdoor path, producing an unbiased estimate of the ATE.

"Good" Controls – Blocking backdoor paths

Model 1

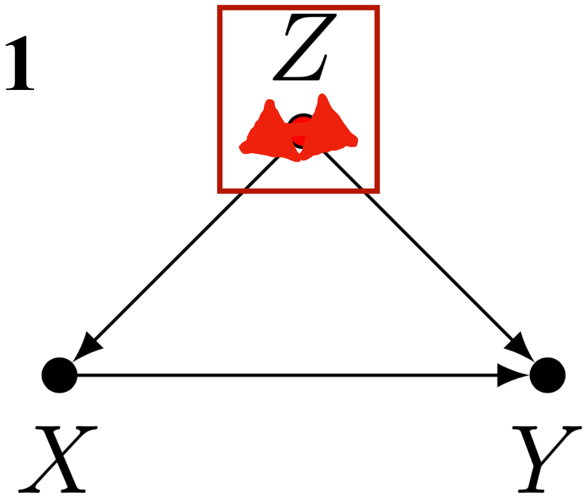


Z is a confounder, and creates a spurious association between X and Y.

Once we control for Z we block the backdoor path, producing an unbiased estimate of the ATE.

"Good" Controls – Blocking backdoor paths

Model 1

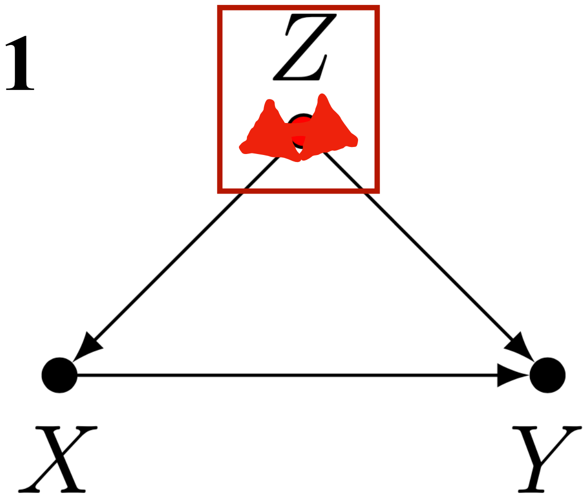


Z is a confounder, and creates a spurious association between X and Y.

Once we control for Z we block the backdoor path, producing an unbiased estimate of the ATE.

“Good” Controls – Blocking backdoor paths

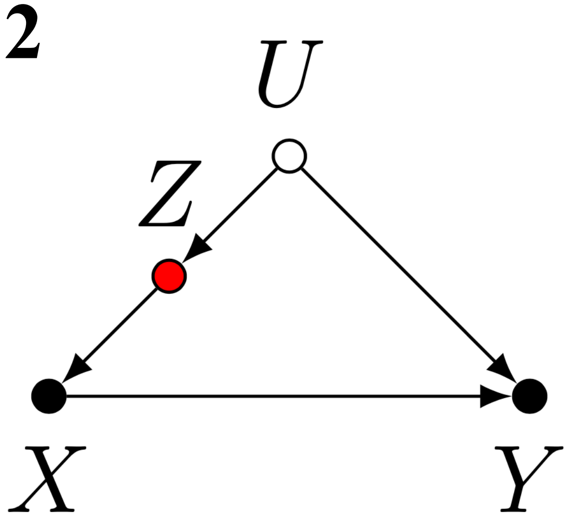
Model 1



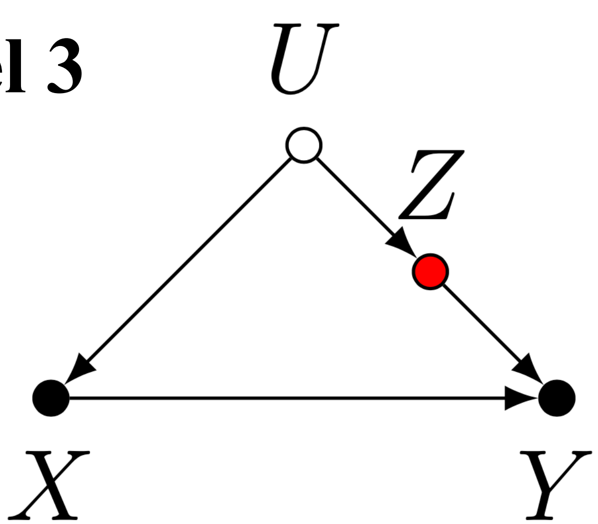
Z is a confounder, and creates a spurious association between X and Y.

Once we control for Z we block the backdoor path, producing an unbiased estimate of the ATE.

Model 2

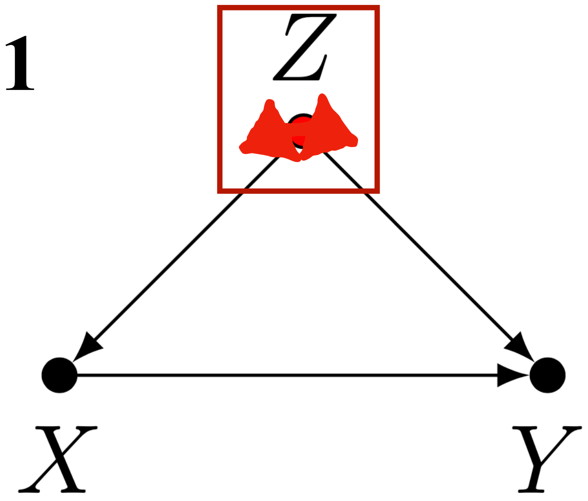


Model 3



“Good” Controls – Blocking backdoor paths

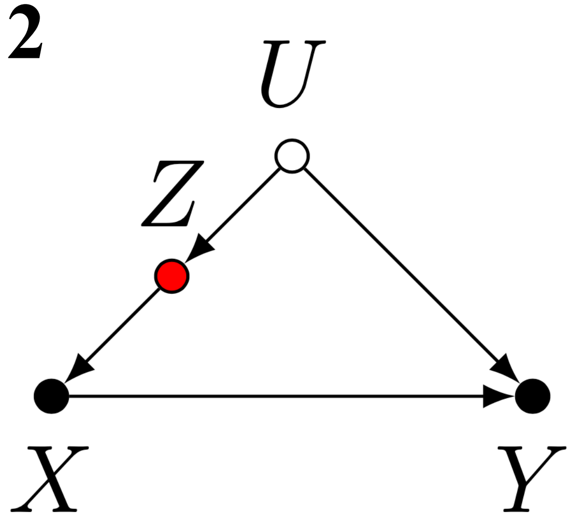
Model 1



Z is a confounder, and creates a spurious association between X and Y.

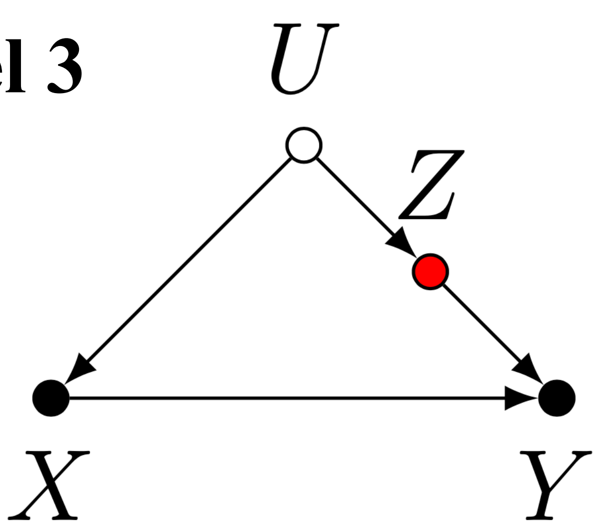
Once we control for Z we block the backdoor path, producing an unbiased estimate of the ATE.

Model 2



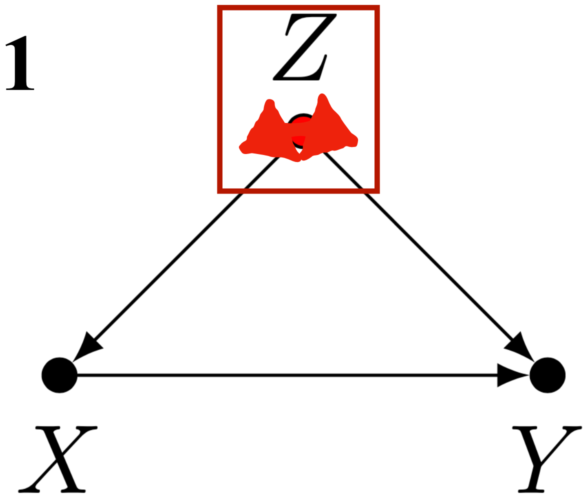
In Models 2 and 3 we have a backdoor path due to the unobserved confounder U.

Model 3



"Good" Controls – Blocking backdoor paths

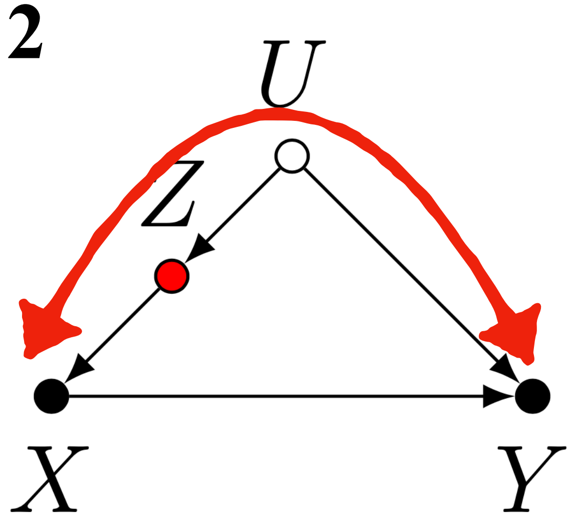
Model 1



Z is a confounder, and creates a spurious association between X and Y.

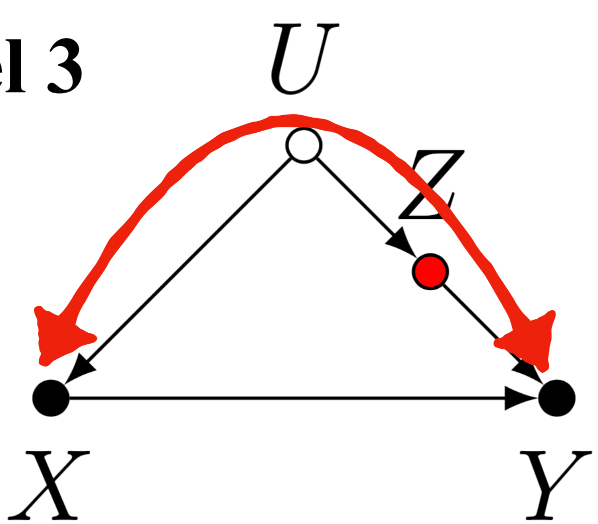
Once we control for Z we block the backdoor path, producing an unbiased estimate of the ATE.

Model 2



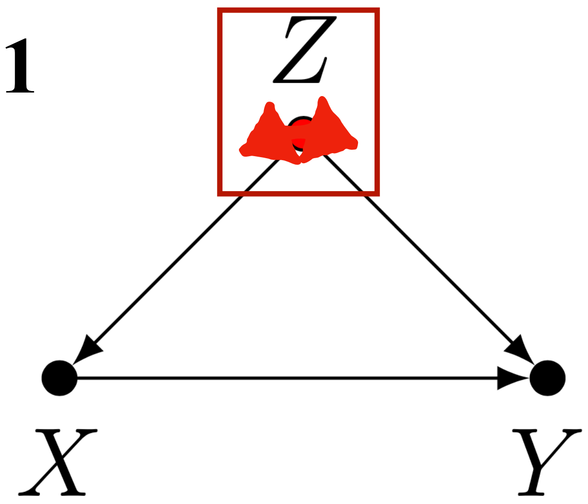
In Models 2 and 3 we have a backdoor path due to the unobserved confounder U.

Model 3



"Good" Controls – Blocking backdoor paths

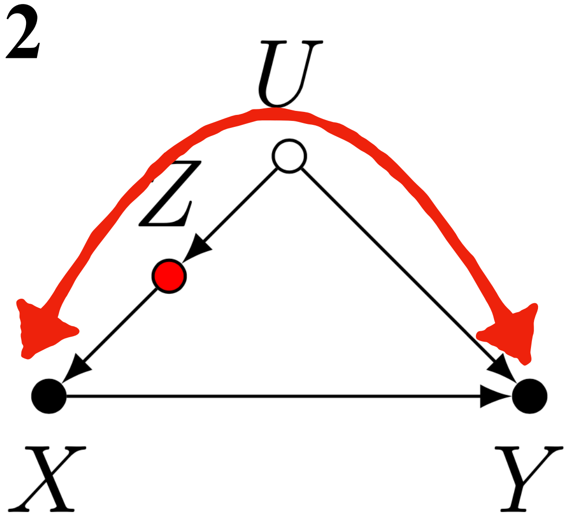
Model 1



Z is a confounder, and creates a spurious association between X and Y.

Once we control for Z we block the backdoor path, producing an unbiased estimate of the ATE.

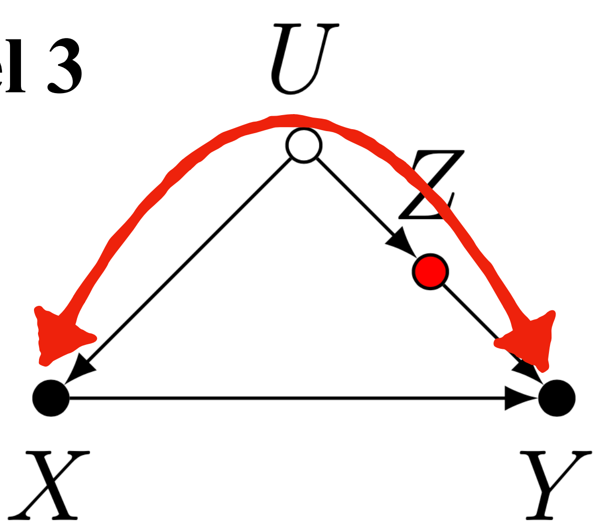
Model 2



In Models 2 and 3 we have a backdoor path due to the unobserved confounder U.

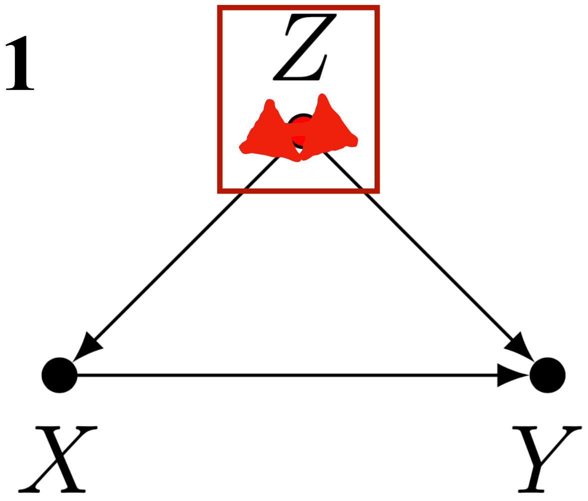
Note that Z is not a common cause of X and Y. Thus Z is not a traditional "confounder" as before.

Model 3



“Good” Controls – Blocking backdoor paths

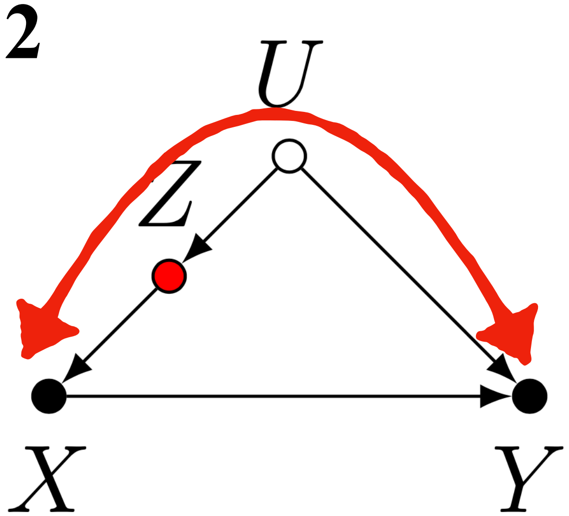
Model 1



Z is a confounder, and creates a spurious association between X and Y.

Once we control for Z we block the backdoor path, producing an unbiased estimate of the ATE.

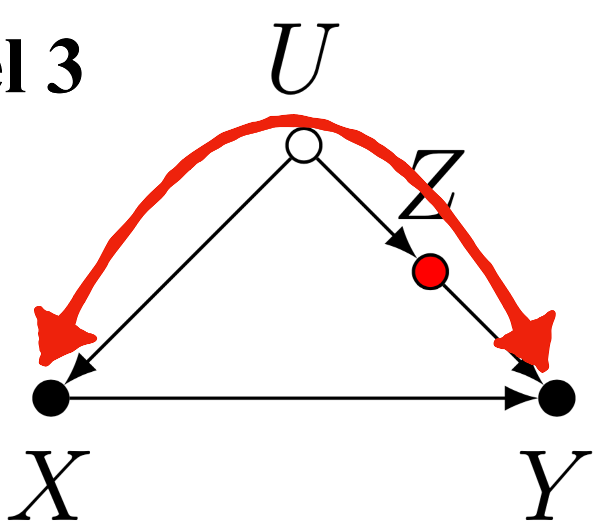
Model 2



In Models 2 and 3 we have a backdoor path due to the unobserved confounder U.

Note that Z is not a common cause of X and Y. Thus Z is not a traditional “confounder” as before.

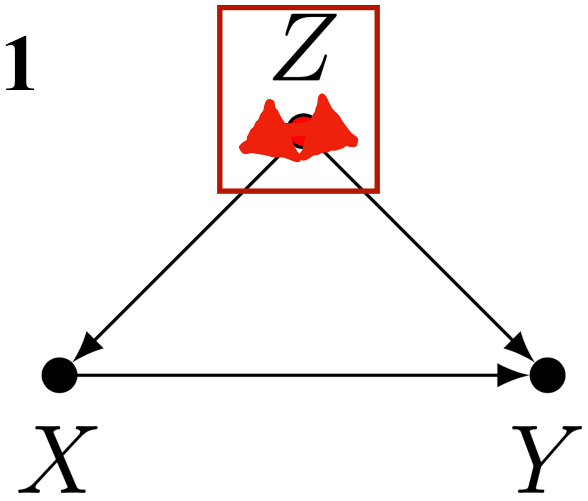
Model 3



However controlling for Z does block the backdoor path due to U, and again, we obtain an unbiased estimate of the ATE.

“Good” Controls – Blocking backdoor paths

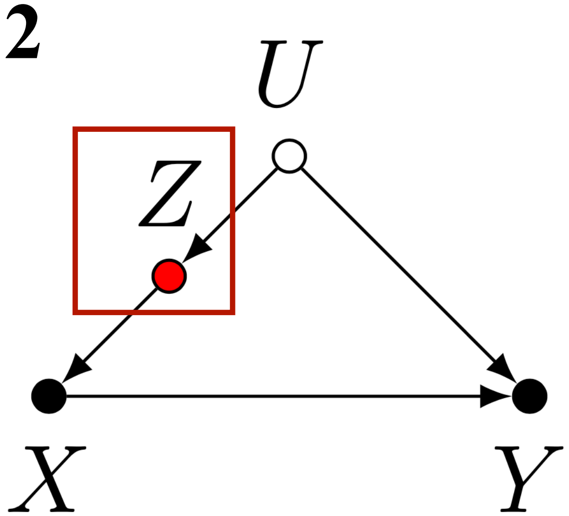
Model 1



Z is a confounder, and creates a spurious association between X and Y.

Once we control for Z we block the backdoor path, producing an unbiased estimate of the ATE.

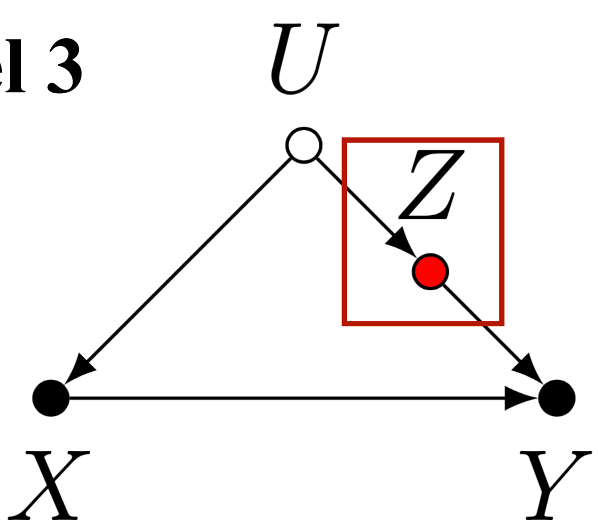
Model 2



In Models 2 and 3 we have a backdoor path due to the unobserved confounder U.

Note that Z is not a common cause of X and Y. Thus Z is not a traditional “confounder” as before.

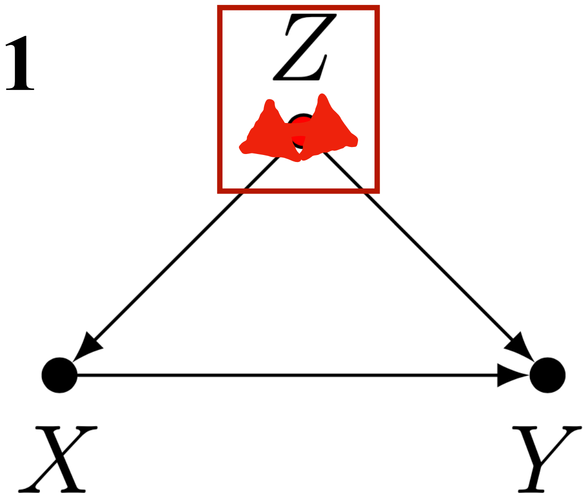
Model 3



However controlling for Z does block the backdoor path due to U, and again, we obtain an unbiased estimate of the ATE.

"Good" Controls – Blocking backdoor paths

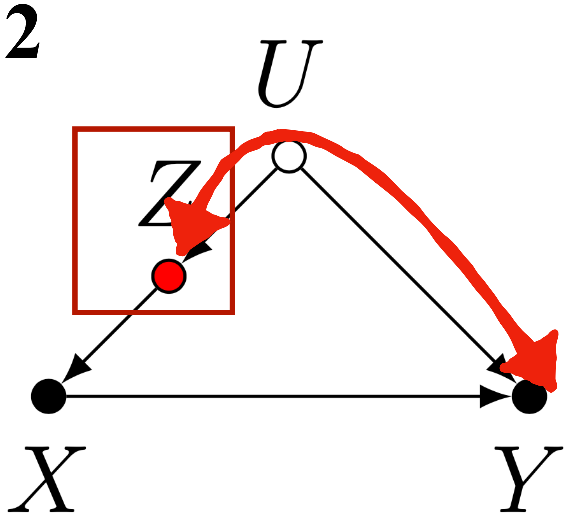
Model 1



Z is a confounder, and creates a spurious association between X and Y.

Once we control for Z we block the backdoor path, producing an unbiased estimate of the ATE.

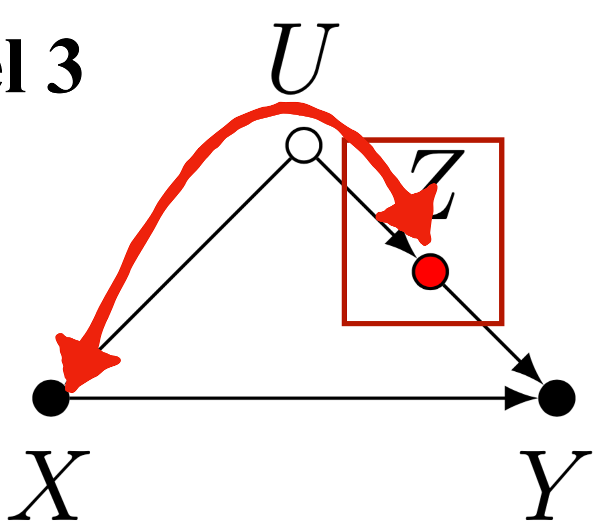
Model 2



In Models 2 and 3 we have a backdoor path due to the unobserved confounder U.

Note that Z is not a common cause of X and Y. Thus Z is not a traditional "confounder" as before.

Model 3



However controlling for Z does block the backdoor path due to U, and again, we obtain an unbiased estimate of the ATE.

“Good” Controls – Blocking backdoor paths

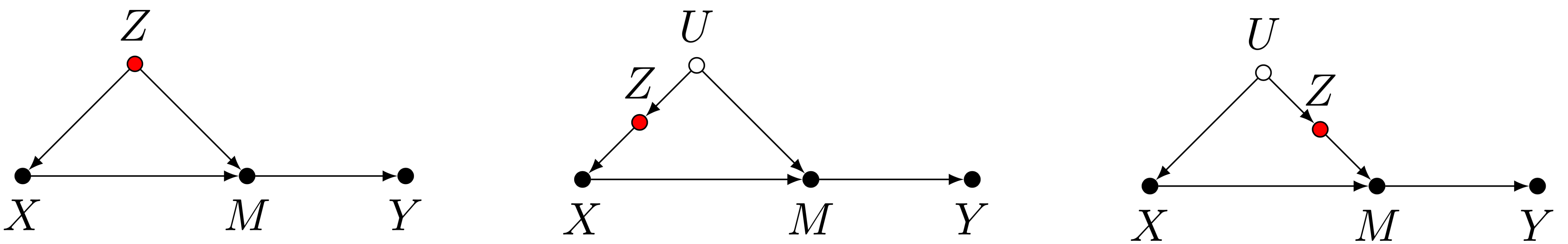
Modelers need to keep in mind that common causes of X and any mediator (between X and Y) also confound the effect of X and Y.

For instance consider models 4, 5 and 6 below:

"Good" Controls – Blocking backdoor paths

Modelers need to keep in mind that common causes of X and any mediator (between X and Y) also confound the effect of X and Y.

For instance consider models 4, 5 and 6 below:

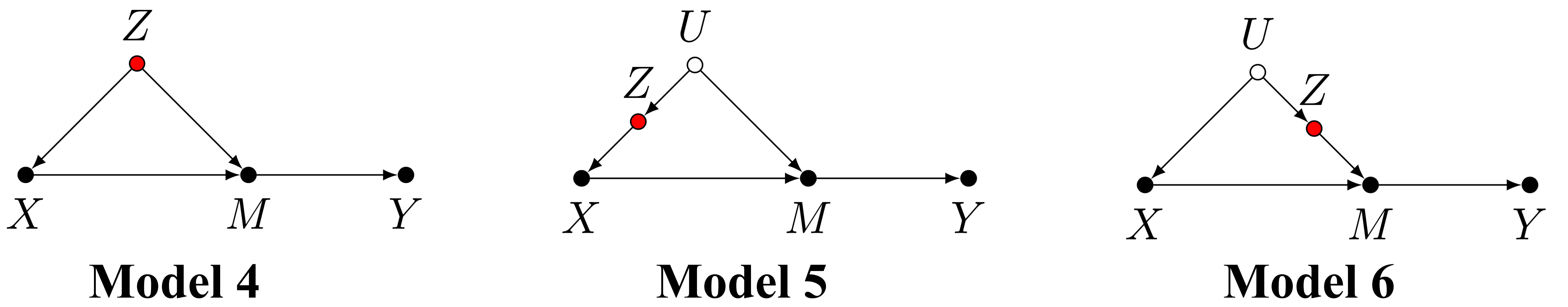


Model 4 **Model 5** **Model 6**

"Good" Controls – Blocking backdoor paths

Modelers need to keep in mind that common causes of X and any mediator (between X and Y) also confound the effect of X and Y.

For instance consider models 4, 5 and 6 below:

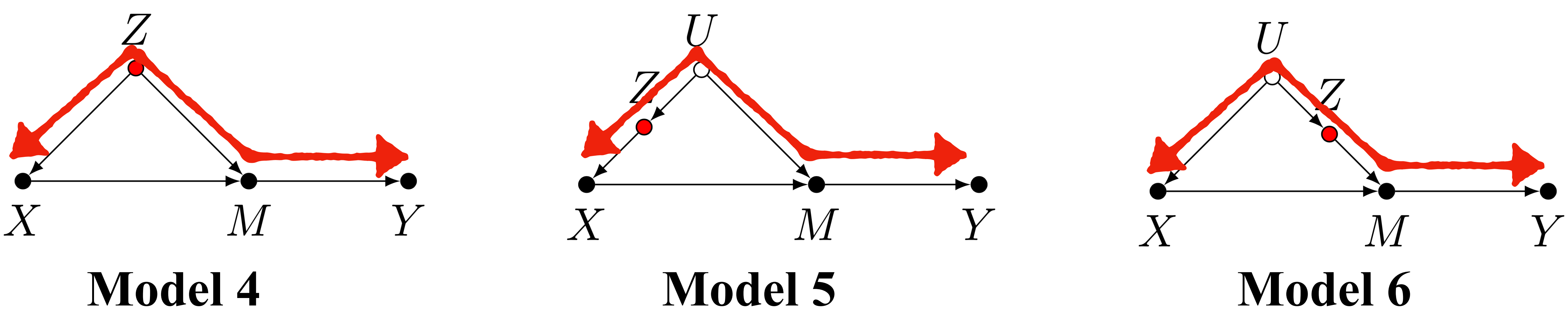


Note these are analogous to Models 1, 2 and 3, and there is an open backdoor path from X to Y.

"Good" Controls – Blocking backdoor paths

Modelers need to keep in mind that common causes of X and any mediator (between X and Y) also confound the effect of X and Y.

For instance consider models 4, 5 and 6 below:

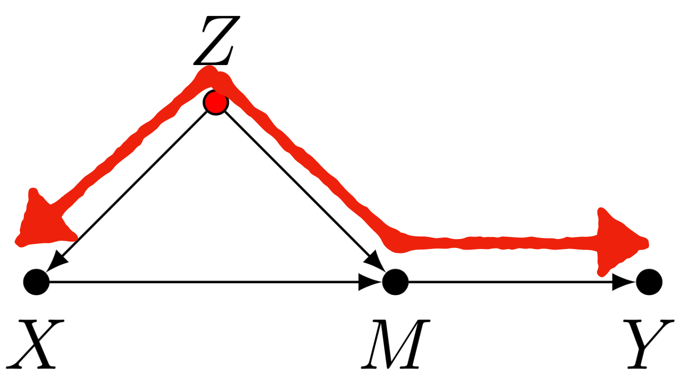


Note these are analogous to Models 1, 2 and 3, and there is an open backdoor path from X to Y.

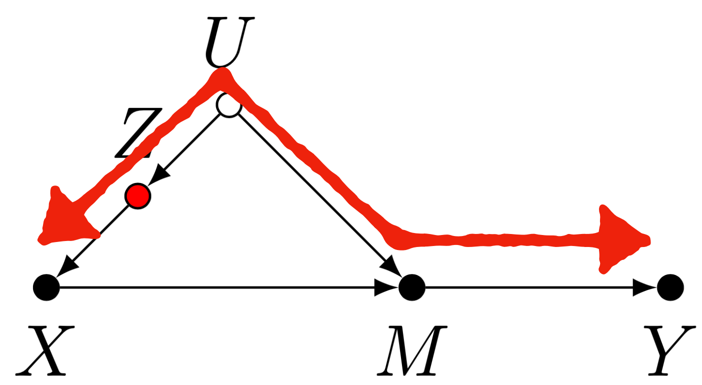
“Good” Controls – Blocking backdoor paths

Modelers need to keep in mind that common causes of X and any mediator (between X and Y) also confound the effect of X and Y.

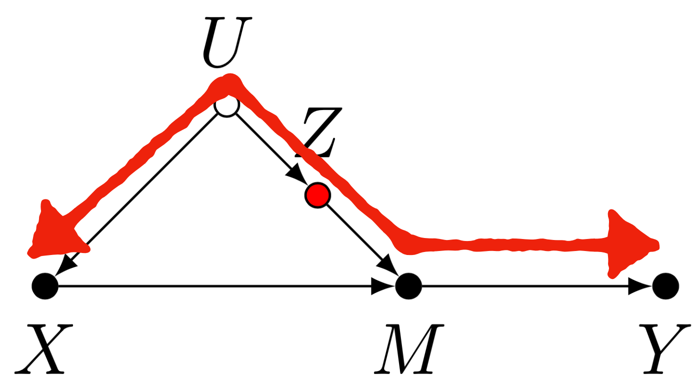
For instance consider models 4, 5 and 6 below:



Model 4



Model 5



Model 6

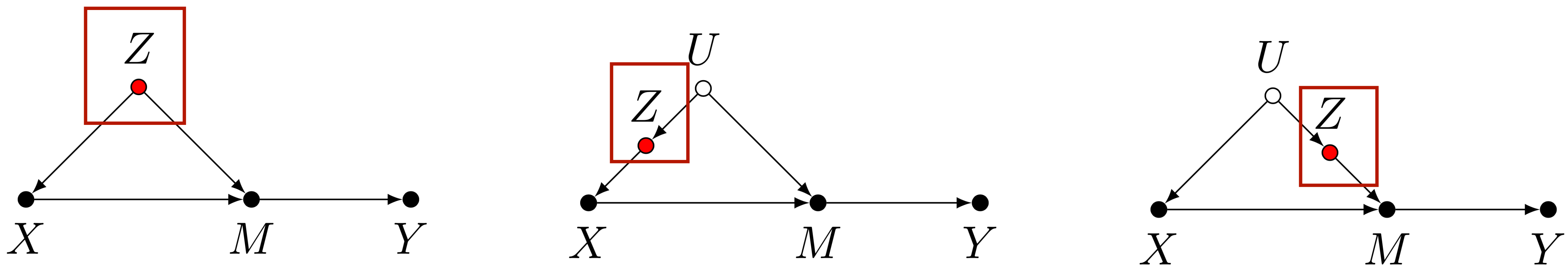
Note these are analogous to Models 1, 2 and 3, and there is an open backdoor path from X to Y.

Here, as before, controlling for Z <u>blocks</u> the backdoor path from X to Y, and produces an unbiased estimate of the ATE.

“Good” Controls – Blocking backdoor paths

Modelers need to keep in mind that common causes of X and any mediator (between X and Y) also confound the effect of X and Y.

For instance consider models 4, 5 and 6 below:



Model 4 **Model 5** **Model 6**

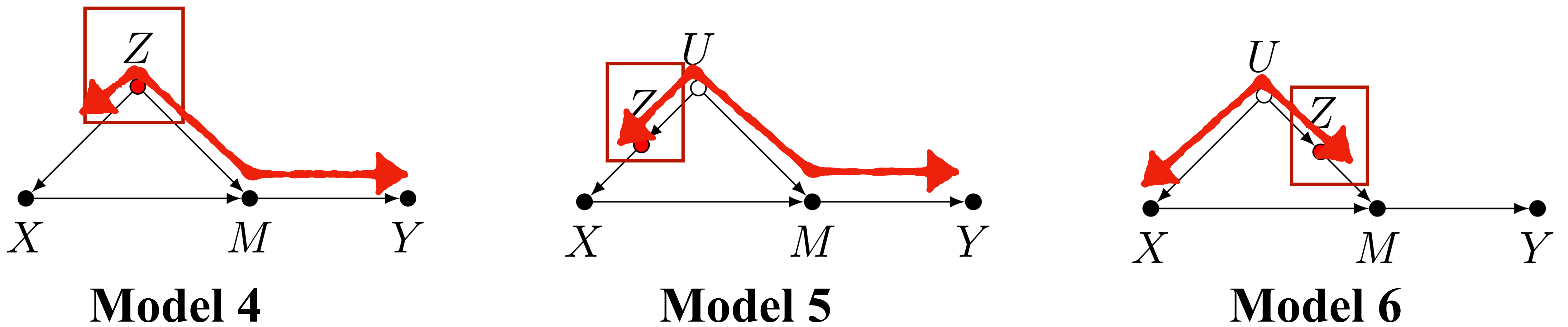
Note these are analogous to Models 1, 2 and 3, and there is an open backdoor path from X to Y.

Here, as before, controlling for Z blocks the backdoor path from X to Y, and produces an unbiased estimate of the ATE.

“Good” Controls – Blocking backdoor paths

Modelers need to keep in mind that common causes of X and any mediator (between X and Y) also confound the effect of X and Y.

For instance consider models 4, 5 and 6 below:

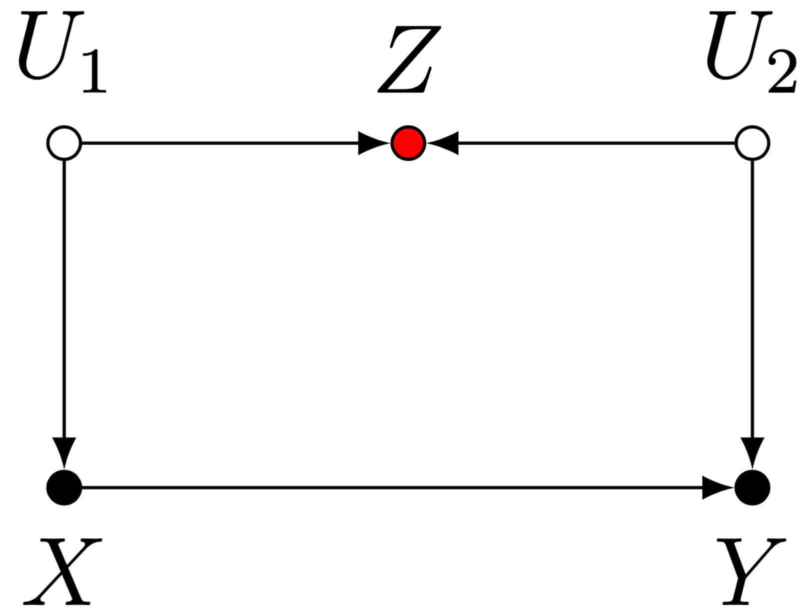


Note these are analogous to Models 1, 2 and 3, and there is an open backdoor path from X to Y.

Here, as before, controlling for Z blocks the backdoor path from X to Y, and produces an unbiased estimate of the ATE.

“Bad” Control – M-Bias

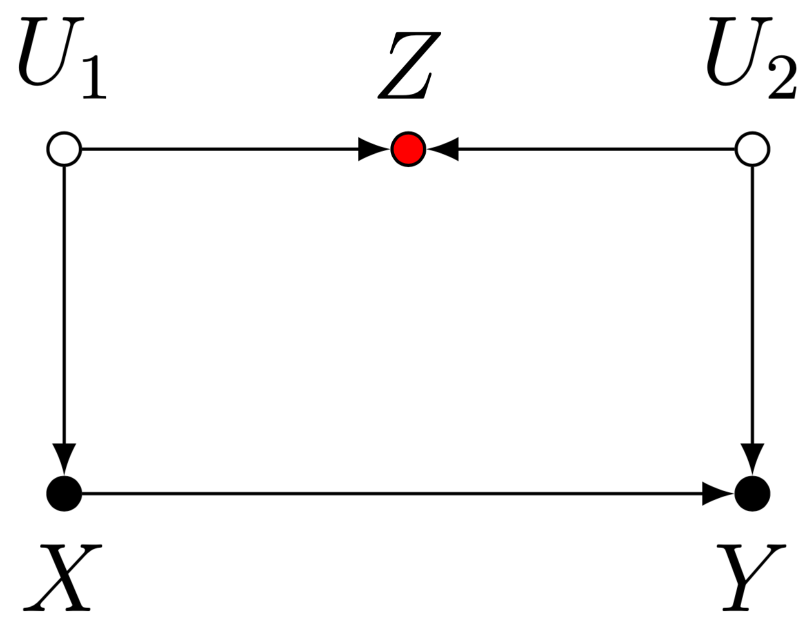
We now encounter our first bad control.



Model 7

“Bad” Control – M-Bias

We now encounter our first bad control.

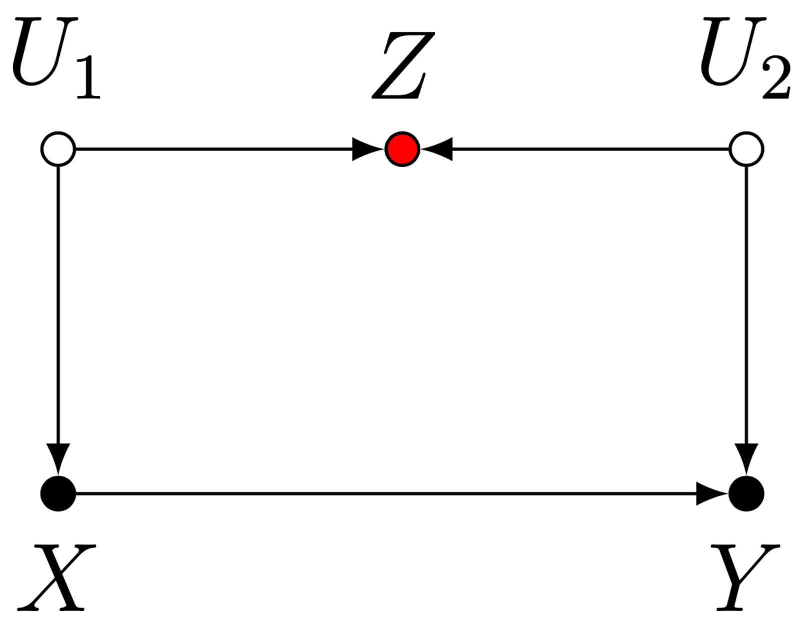


Model 7

Note that: (i) Z is correlated with the treatment X; (ii) Z is correlated with the outcome Y; (iii) Z is “pre-treatment.” Traditional analysis would thus consider Z a confounder, and perhaps deem Z a good control.

“Bad” Control – M-Bias

We now encounter our first bad control.



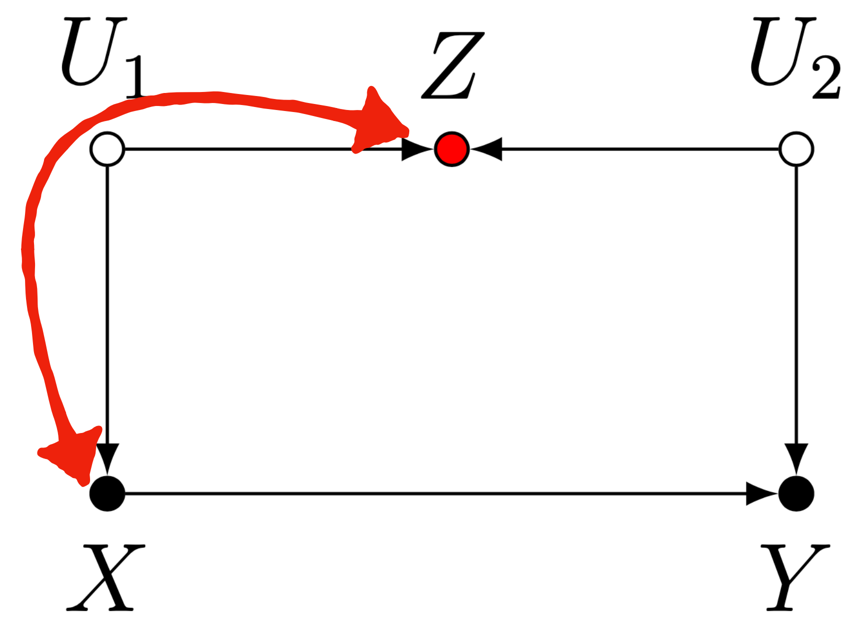
Model 7

Note that: (i) Z is correlated with the treatment X; (ii) Z is correlated with the outcome Y; (iii) Z is “pre-treatment.” Traditional analysis would thus consider Z a confounder, and perhaps deem Z a good control.

However, first note that there is <u>no open confounding path</u> from X to Y. The unadjusted estimate is thus unbiased!

“Bad” Control – M-Bias

We now encounter our first bad control.



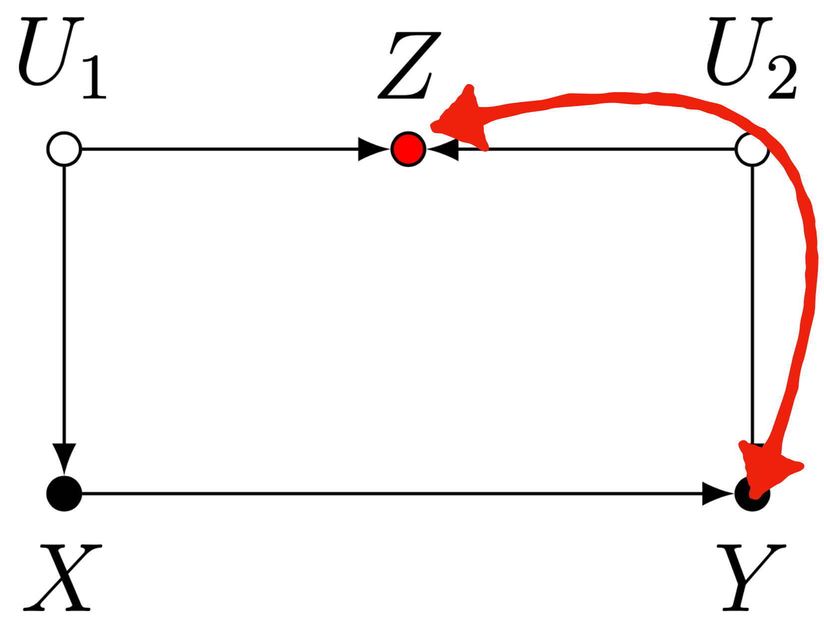
Model 7

Note that: (i) Z is correlated with the treatment X; (ii) Z is correlated with the outcome Y; (iii) Z is “pre-treatment.” Traditional analysis would thus consider Z a confounder, and perhaps deem Z a good control.

However, first note that there is <u>no open confounding path</u> from X to Y. The unadjusted estimate is thus unbiased!

"Bad" Control – M-Bias

We now encounter our first bad control.



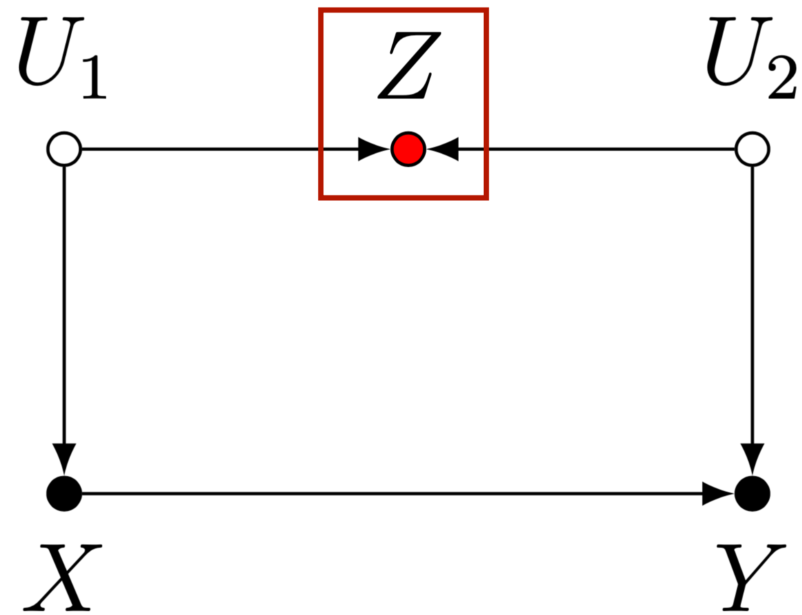
Model 7

Note that: (i) Z is correlated with the treatment X; (ii) Z is correlated with the outcome Y; (iii) Z is "pre-treatment." Traditional analysis would thus consider Z a confounder, and perhaps deem Z a good control.

However, first note that there is <u>no open confounding path</u> from X to Y. The unadjusted estimate is thus unbiased!

"Bad" Control – M-Bias

We now encounter our first bad control.



Model 7

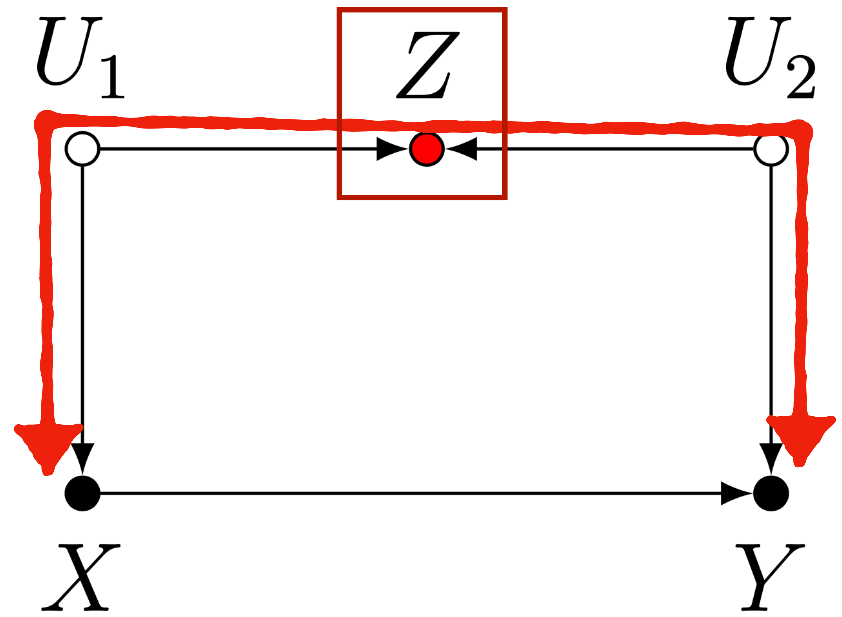
Note that: (i) Z is correlated with the treatment X; (ii) Z is correlated with the outcome Y; (iii) Z is "pre-treatment." Traditional analysis would thus consider Z a confounder, and perhaps deem Z a good control.

However, first note that there is <u>no open confounding path</u> from X to Y. The unadjusted estimate is thus unbiased!

Conditioning on Z, however, *opens* the path X←U1→Z←U2→Y, and thus spoils a previously unbiased estimate.

“Bad” Control – M-Bias

We now encounter our first bad control.



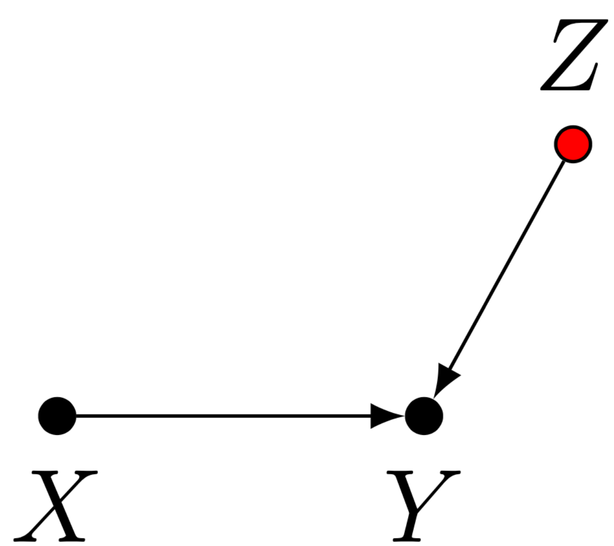
Model 7

Note that: (i) Z is correlated with the treatment X; (ii) Z is correlated with the outcome Y; (iii) Z is “pre-treatment.” Traditional analysis would thus consider Z a confounder, and perhaps deem Z a good control.

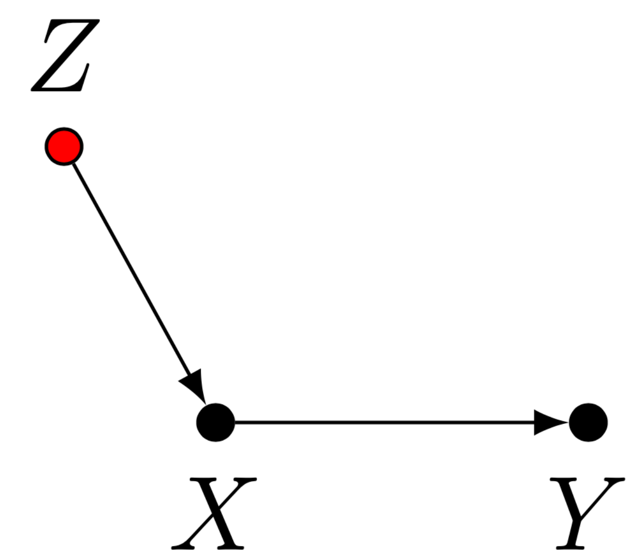
However, first note that there is <u>no open confounding path</u> from X to Y. The unadjusted estimate is thus unbiased!

Conditioning on Z, however, *opens* the path X←U1→Z←U2→Y, and thus spoils a previously unbiased estimate.

“Neutral” Controls (but different impacts on precision)

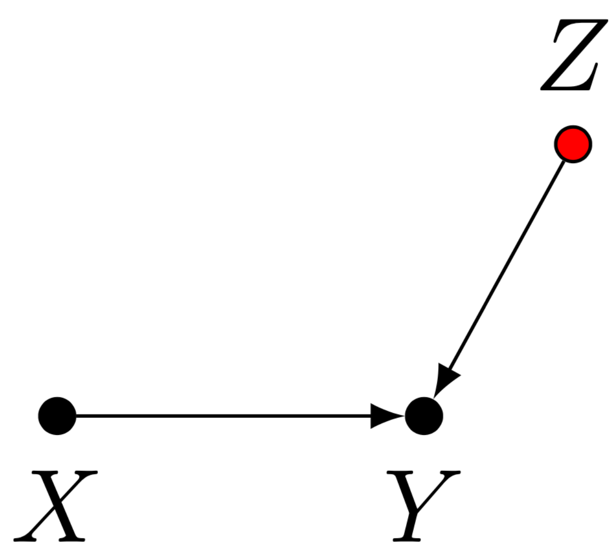


Model 8

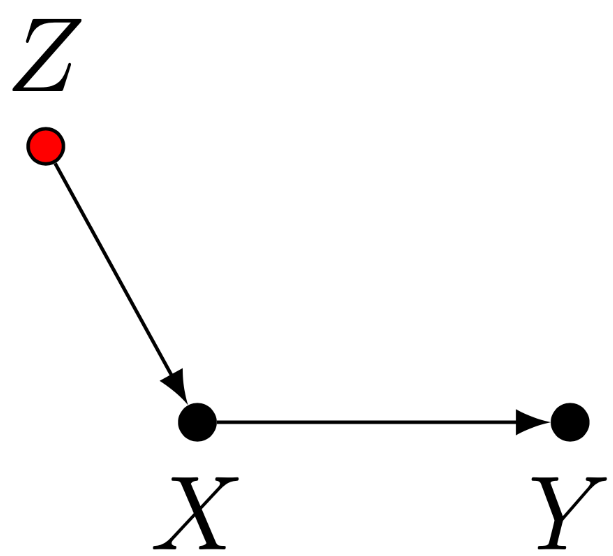


Model 9

"Neutral" Controls (but different impacts on precision)



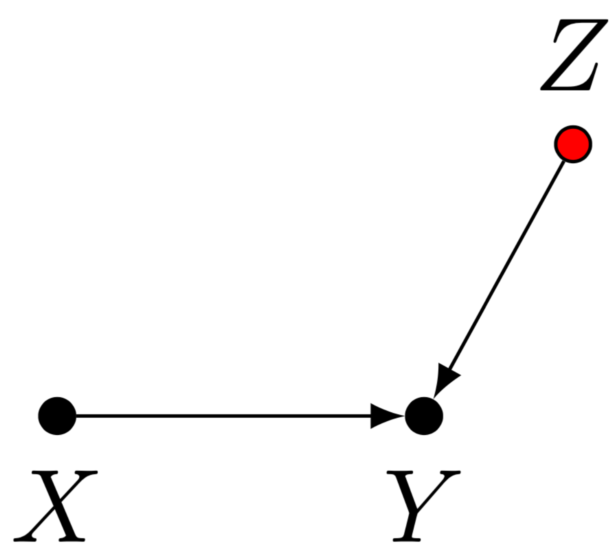
Model 8



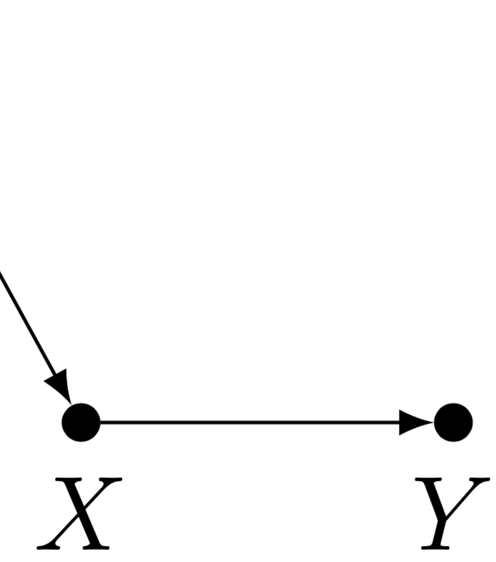
Model 9

In Models 8 and 9, Z is not a confounder, nor does Z block any backdoor paths. Likewise, controlling for Z does not open any spurious paths from X to Y.

"Neutral" Controls (but different impacts on precision)



Model 8

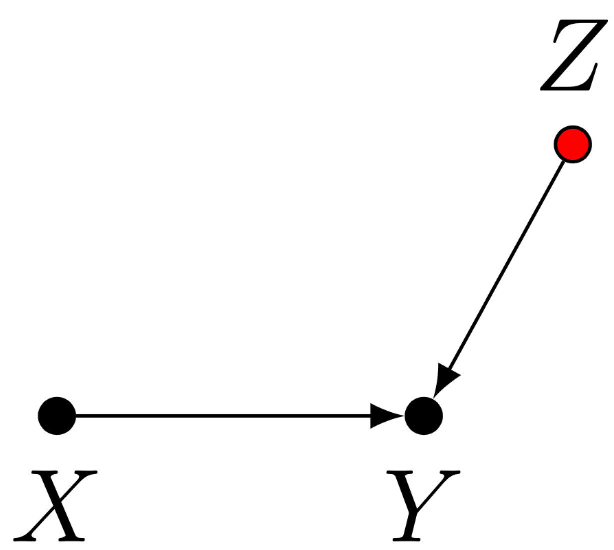


Model 9

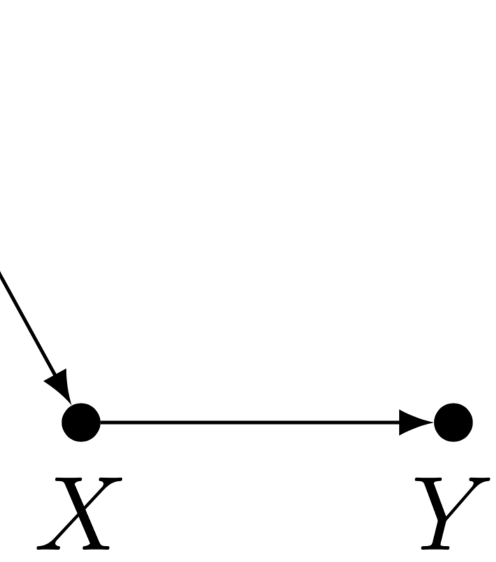
In Models 8 and 9, Z is not a confounder, nor does Z block any backdoor paths. Likewise, controlling for Z does not open any spurious paths from X to Y.

Thus, in terms of asymptotic bias, Z is thus a "neutral control."

"Neutral" Controls (but different impacts on precision)



Model 8



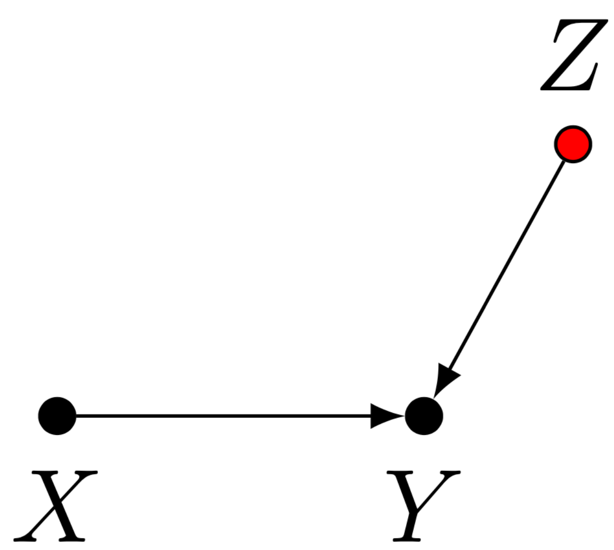
Model 9

In Models 8 and 9, Z is not a confounder, nor does Z block any backdoor paths. Likewise, controlling for Z does not open any spurious paths from X to Y.

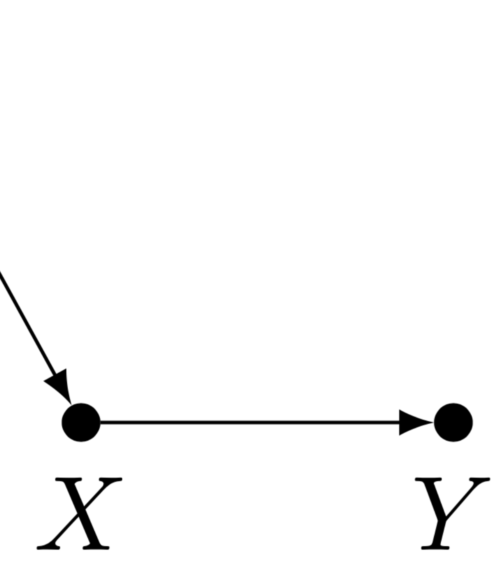
Thus, in terms of asymptotic bias, Z is thus a "neutral control."

As a general rule-of-thumb, however, in order to obtain more precise estimates of the ATE, we want to reduce the variation of the outcome (due to sources other than the treatment), and not reduce the variation of the treatment itself.

"Neutral" Controls (but different impacts on precision)



Model 8



Model 9

In Models 8 and 9, Z is not a confounder, nor does Z block any backdoor paths. Likewise, controlling for Z does not open any spurious paths from X to Y.

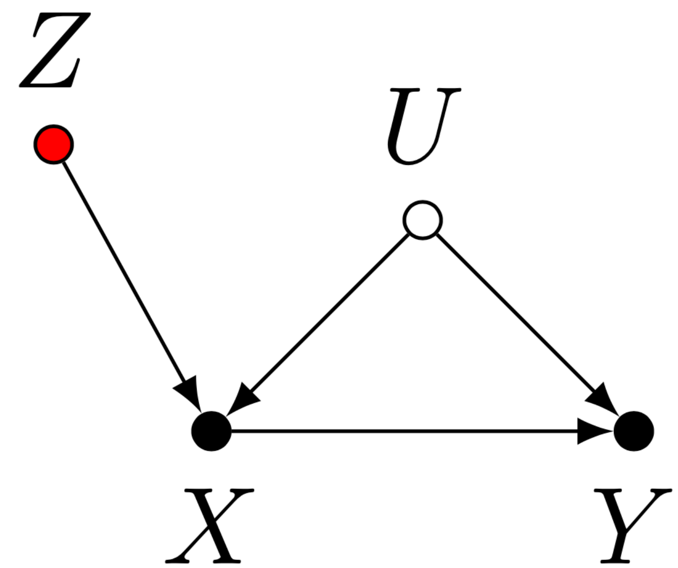
Thus, in terms of asymptotic bias, Z is thus a "neutral control."

As a general rule-of-thumb, however, in order to obtain more precise estimates of the ATE, we want to <u>reduce the variation of the outcome (due to sources other than the treatment),</u> and <u>not reduce the variation of the treatment itself.</u>

Thus, in Model 8, Z improves the precision of the ATE estimate;

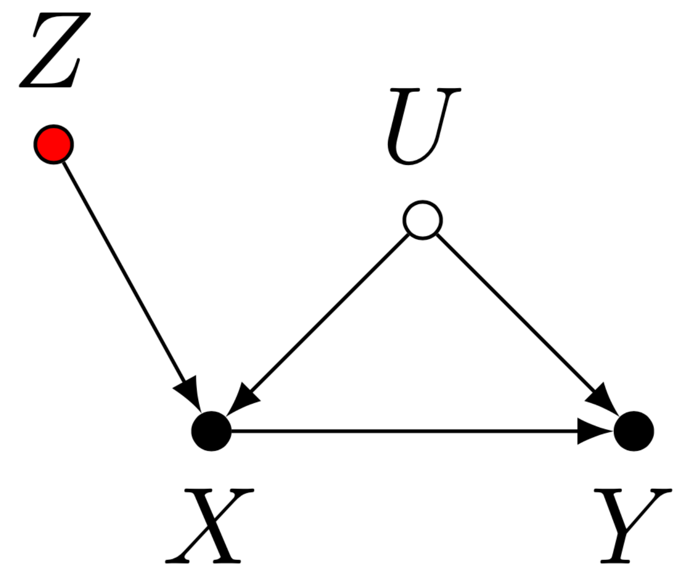
Whereas in Model 9 Z hurts the precision of the ATE estimate.

"Bad" Control (bias amplification)



Model 10

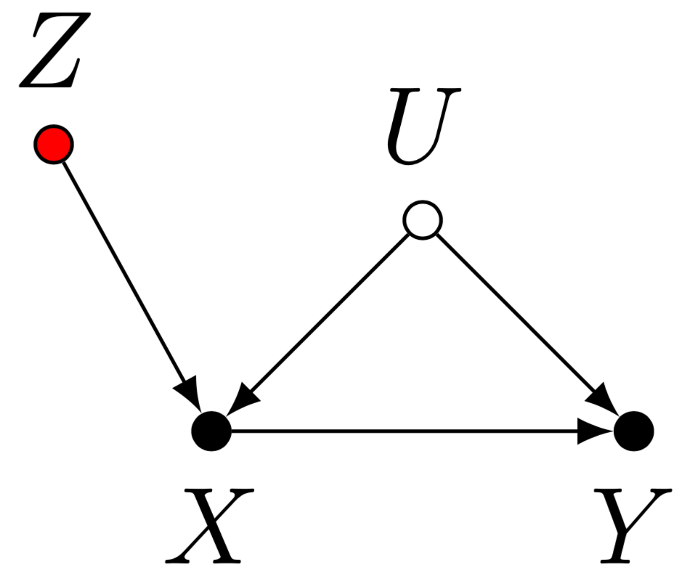
“Bad” Control (bias amplification)



Model 10

We now encounter our second “pre-treatment” bad control.

“Bad” Control (bias amplification)

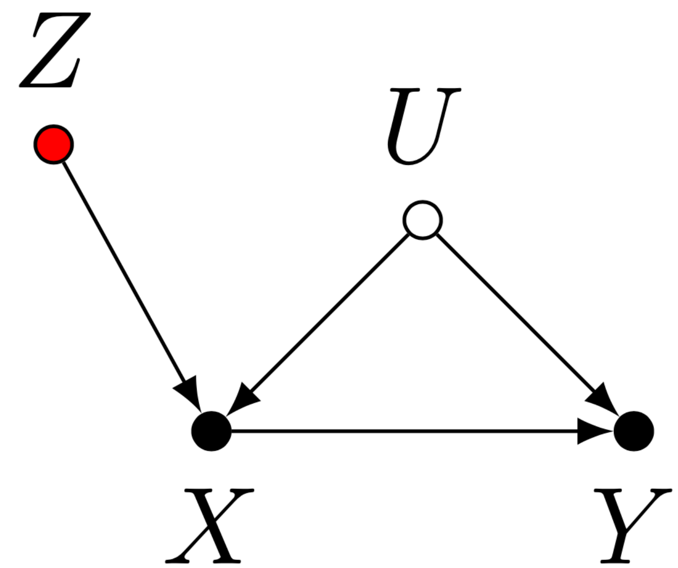


Model 10

We now encounter our second “pre-treatment” bad control.

Note here that: (i) Z is “pre-treatment;”; (ii) Z is associated with X (causally); (ii) Z is associated with Y; and (iii) Z is associated with Y conditional on X.

“Bad” Control (bias amplification)



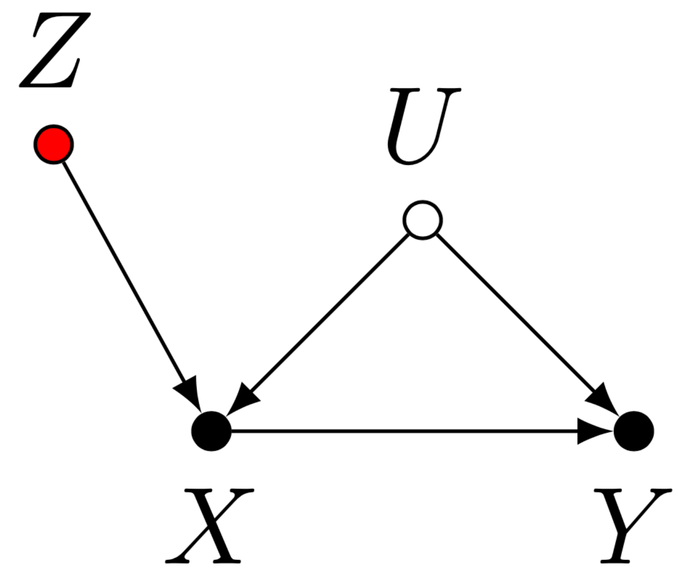
Model 10

We now encounter our second “pre-treatment” bad control.

Note here that: (i) Z is “pre-treatment;”; (ii) Z is associated with X (causally); (ii) Z is associated with Y; and (iii) Z is associated with Y conditional on X.

Thus, Z seems like an ordinary confounder begging to be controlled.

“Bad” Control (bias amplification)



Model 10

We now encounter our second “pre-treatment” bad control.

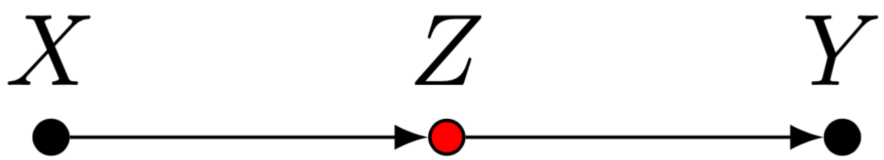
Note here that: (i) Z is “pre-treatment;”; (ii) Z is associated with X (causally); (ii) Z is associated with Y; and (iii) Z is associated with Y conditional on X.

Thus, Z seems like an ordinary confounder begging to be controlled.

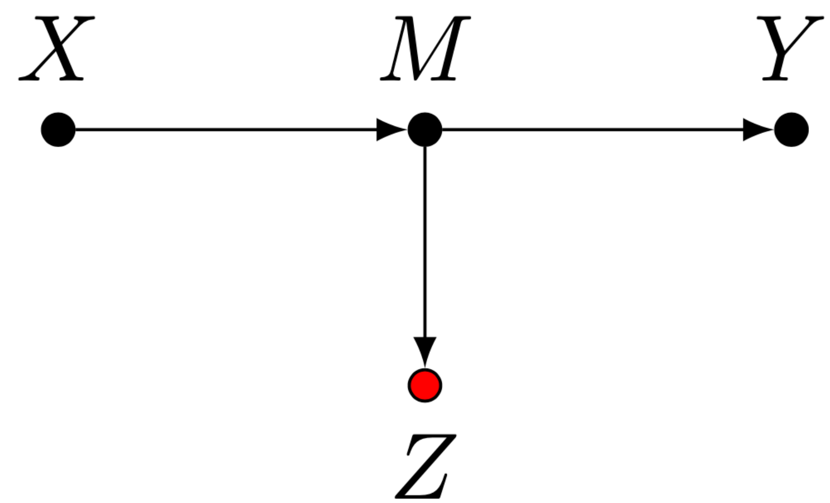
However, analysis shows that adjusting for Z will not only fail to deconfound the effect of X on Y, but, in linear models, it will *amplify* any existing bias.

"Bad" Controls (overcontrol bias)

Model 11

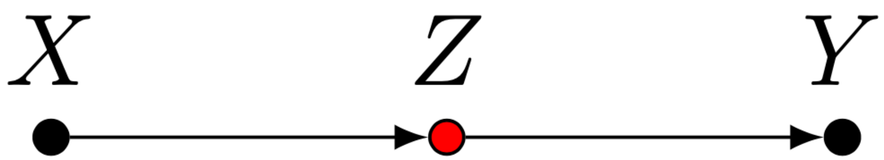


Model 12

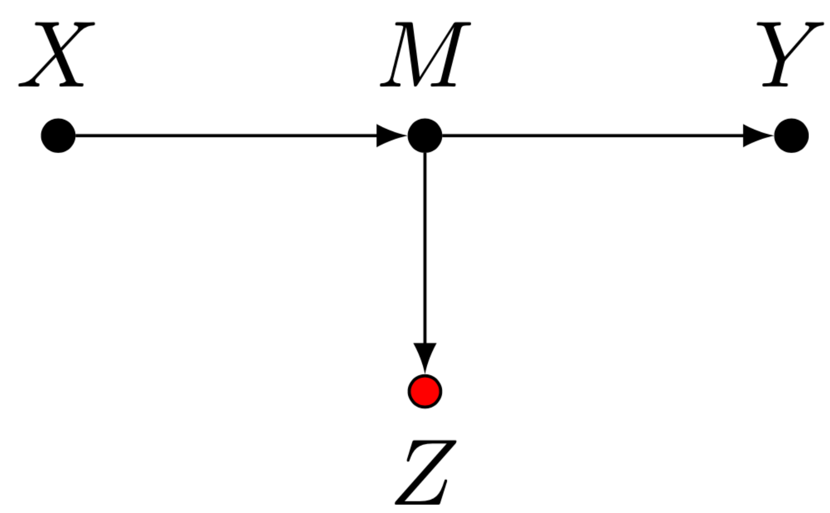


"Bad" Controls (overcontrol bias)

Model 11



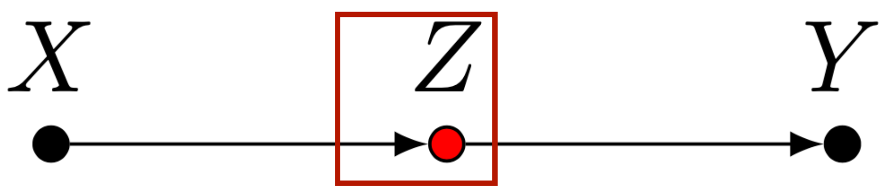
Model 12



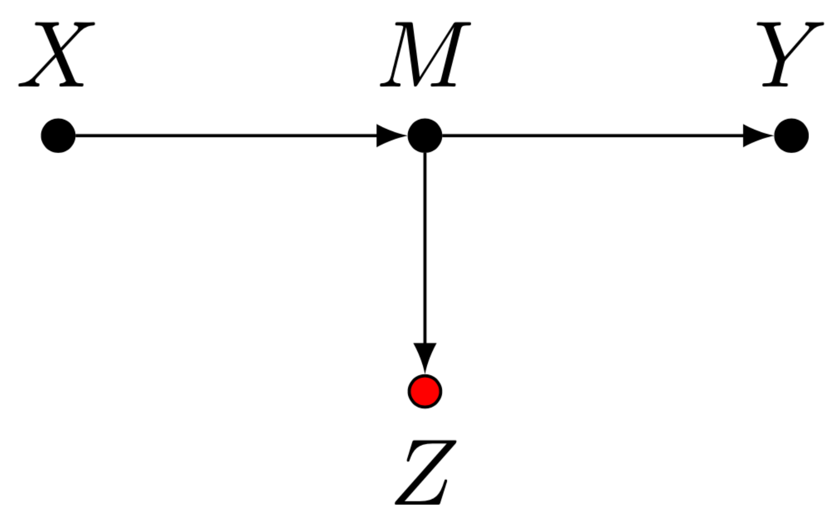
If our target quantity is the ATE, we want to leave all channels through which the causal effect flows "untouched."

"Bad" Controls (overcontrol bias)

Model 11



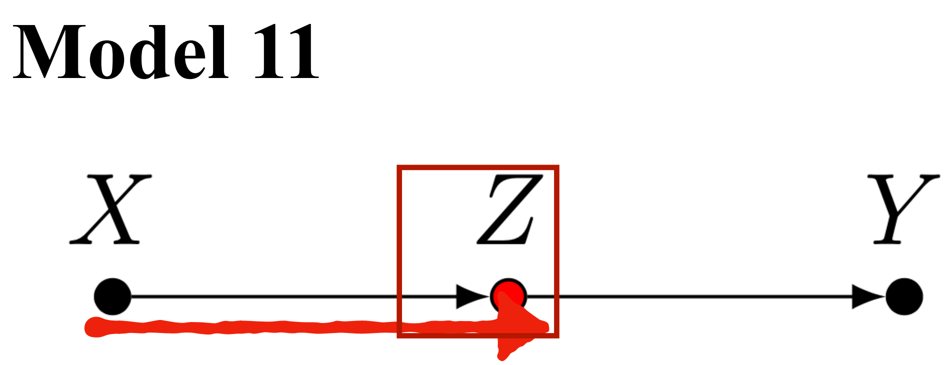
Model 12

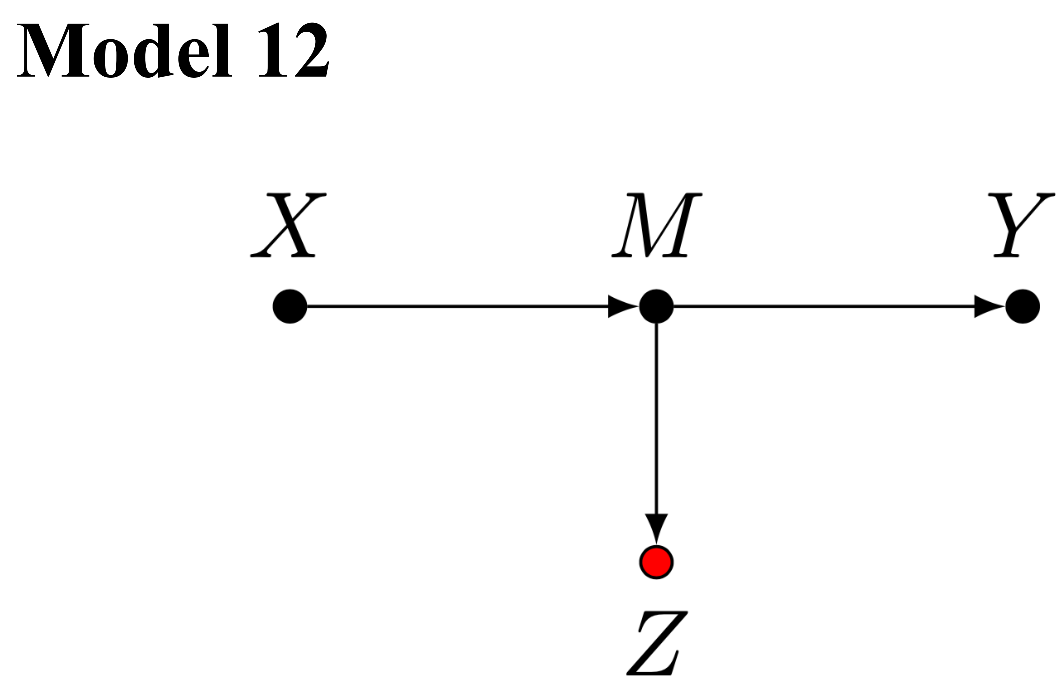


If our target quantity is the ATE, we want to leave all channels through which the causal effect flows "untouched."

In Model 11, Z is a *mediator* of the causal effect of X on Y . Controlling for Z will block the very effect we want to estimate (the total effect of X on Y), thus biasing our estimates (this is usually known as "overcontrol bias").

“Bad” Controls (overcontrol bias)

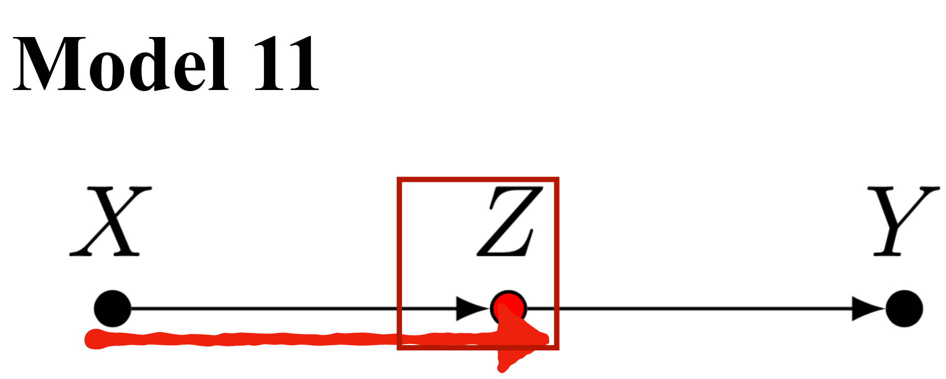


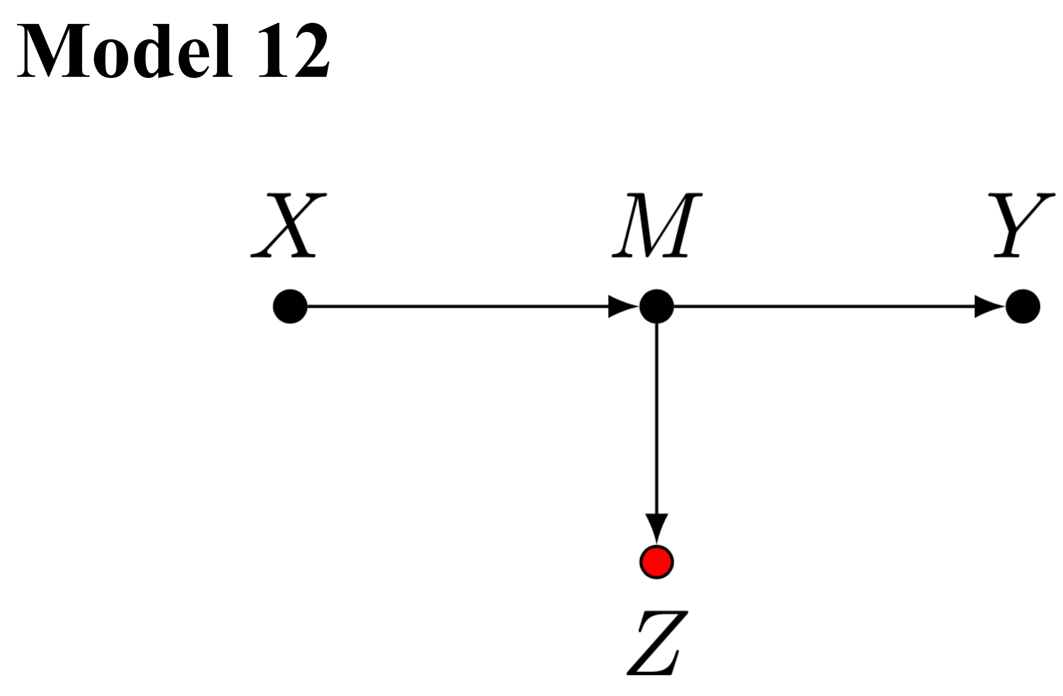


If our target quantity is the ATE, we want to leave all channels through which the causal effect flows “untouched.”

In Model 11, Z is a *mediator* of the causal effect of X on Y . Controlling for Z will block the very effect we want to estimate (the total effect of X on Y), thus biasing our estimates (this is usually known as “overcontrol bias”).

“Bad” Controls (overcontrol bias)



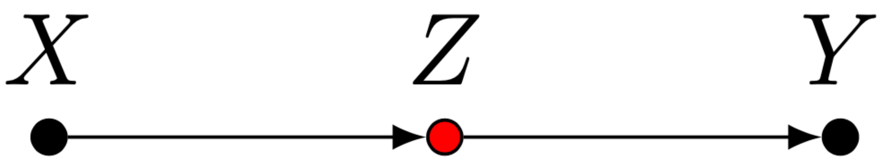


If our target quantity is the ATE, we want to leave all channels through which the causal effect flows “untouched.”

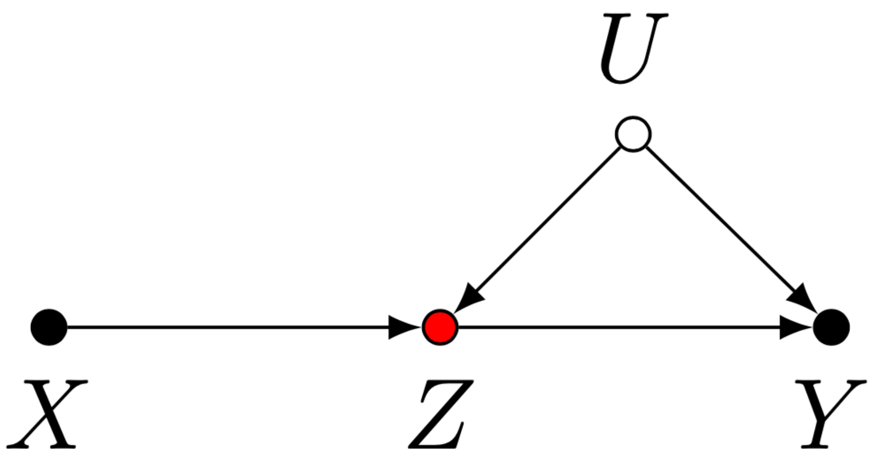
In Model 11, Z is a *mediator* of the causal effect of X on Y . Controlling for Z will block the very effect we want to estimate (the total effect of X on Y), thus biasing our estimates (this is usually known as “overcontrol bias”).

In Model 12, although Z is not itself a mediator of the causal effect of X on Y , *controlling for Z is equivalent to partially controlling for the mediator M*, and will thus bias our estimates.

Total versus direct effect

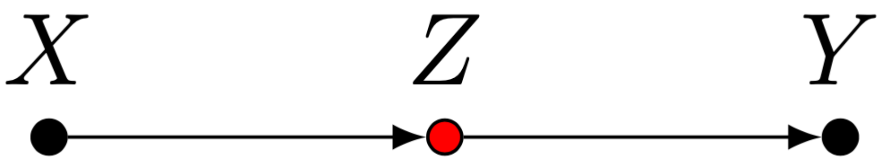


Model 11

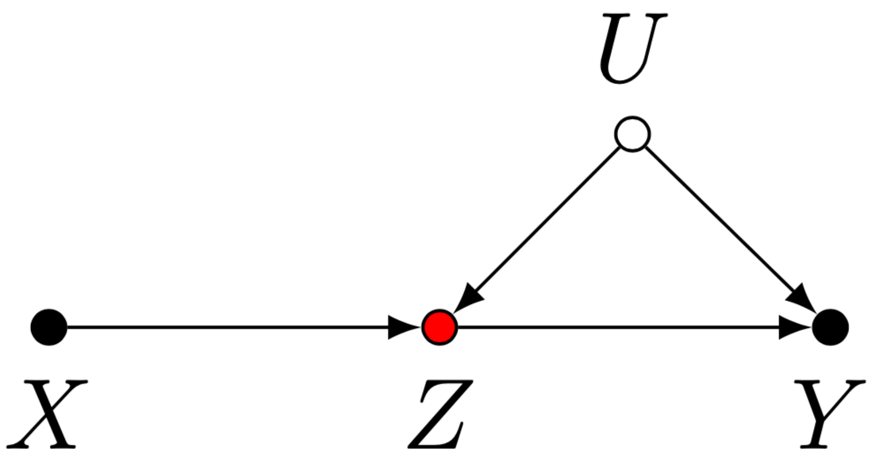


Variation of Model 11

Total versus direct effect



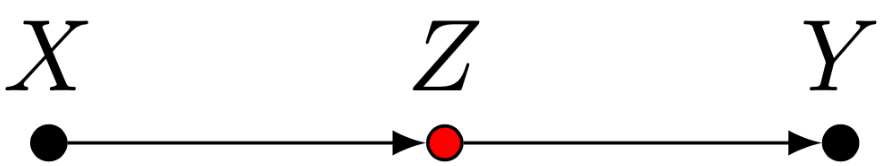
Model 11



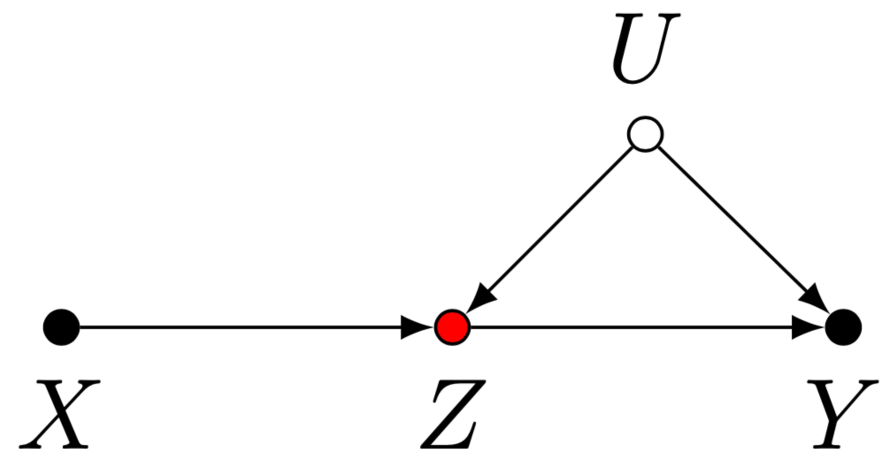
Variation of Model 11

The previous considerations assumed we were interested in the *total* effect of X on Y. If we were interested in the *controlled direct effect (CDE)* of X on Y, then adjusting for Z in Model 11 would *indeed be appropriate.*

Total versus direct effect



Model 11

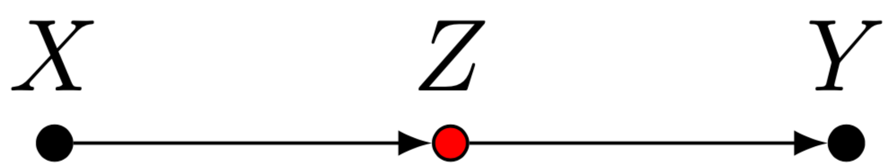


Variation of Model 11

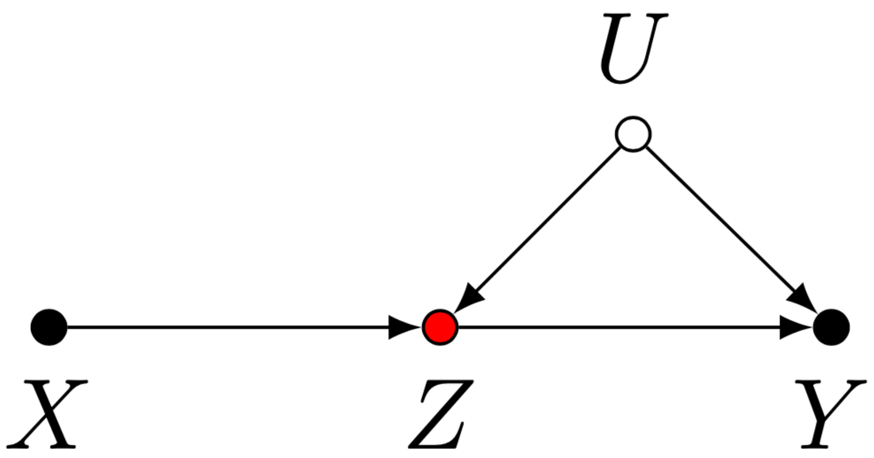
The previous considerations assumed we were interested in the *total* effect of X on Y. If we were interested in the *controlled direct effect (CDE)* of X on Y, then adjusting for Z in Model 11 would *indeed be appropriate.*

However, consider a variation of Model 11 with an unobserved confounder of Z and Y , denoted by U.

Total versus direct effect



Model 11



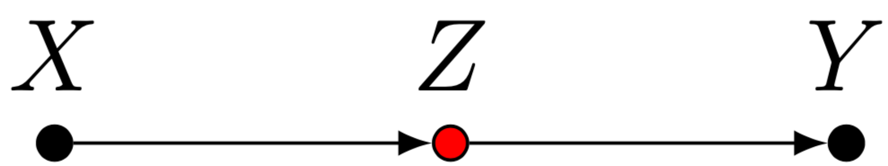
Variation of Model 11

The previous considerations assumed we were interested in the *total* effect of X on Y. If we were interested in the *controlled direct effect (CDE)* of X on Y, then adjusting for Z in Model 11 would *indeed be appropriate.*

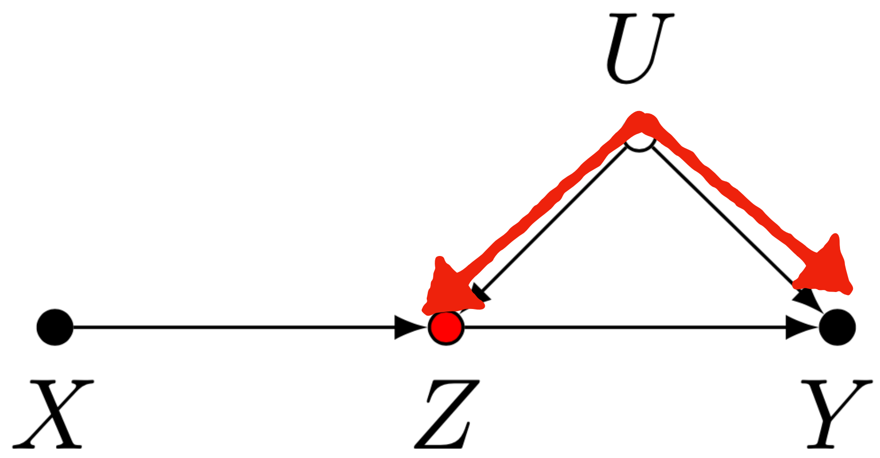
However, consider a variation of Model 11 with an unobserved confounder of Z and Y , denoted by U.

First notice that U does not confound the effect of X on Y. Thus the total effect remains unbiased, as it were in Model 11, *so long as we do not adjust for Z.*

Total versus direct effect



Model 11



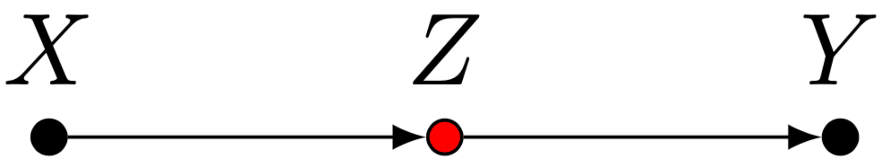
Variation of Model 11

The previous considerations assumed we were interested in the *total* effect of X on Y. If we were interested in the *controlled direct effect (CDE)* of X on Y, then adjusting for Z in Model 11 would *indeed be appropriate.*

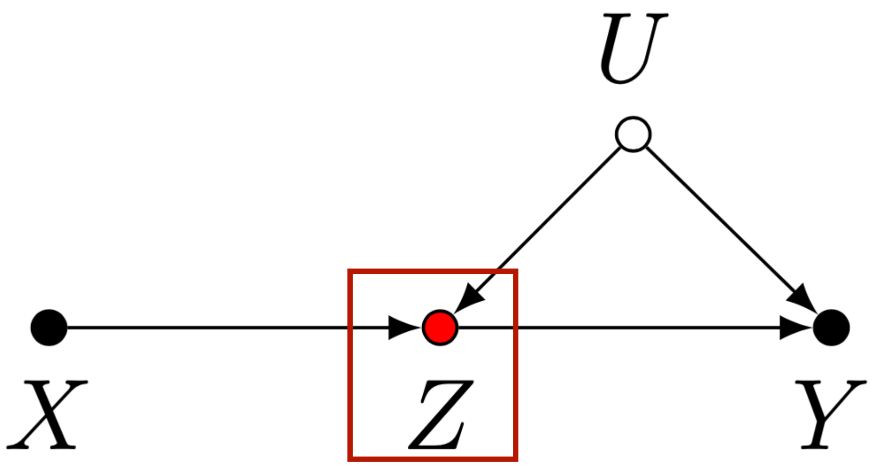
However, consider a variation of Model 11 with an unobserved confounder of Z and Y , denoted by U.

First notice that U does not confound the effect of X on Y. Thus the total effect remains unbiased, as it were in Model 11, *so long as we do not adjust for Z.*

Total versus direct effect



Model 11



Variation of Model 11

The previous considerations assumed we were interested in the *total* effect of X on Y. If we were interested in the *controlled direct effect (CDE)* of X on Y, then adjusting for Z in Model 11 would *indeed be appropriate.*

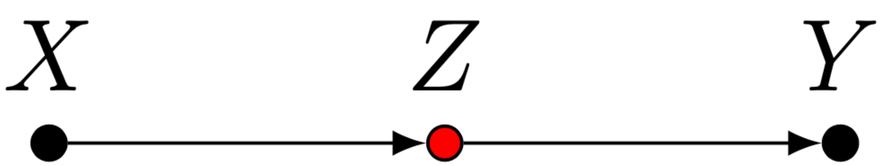
However, consider a variation of Model 11 with an unobserved confounder of Z and Y , denoted by U.

First notice that U does not confound the effect of X on Y. Thus the total effect remains unbiased, as it were in Model 11, *so long as we do not adjust for Z.*

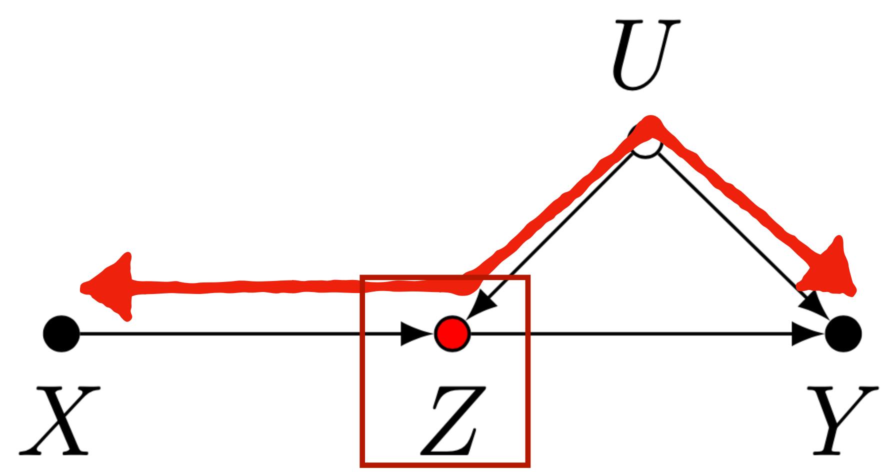
On the other hand, here adjusting for Z now opens the colliding path

$X \rightarrow Z \leftarrow U \rightarrow Y$, thus biasing the CDE estimate.

Total versus direct effect



Model 11



Variation of Model 11

The previous considerations assumed we were interested in the *total* effect of X on Y. If we were interested in the *controlled direct effect (CDE)* of X on Y, then adjusting for Z in Model 11 would *indeed be appropriate.*

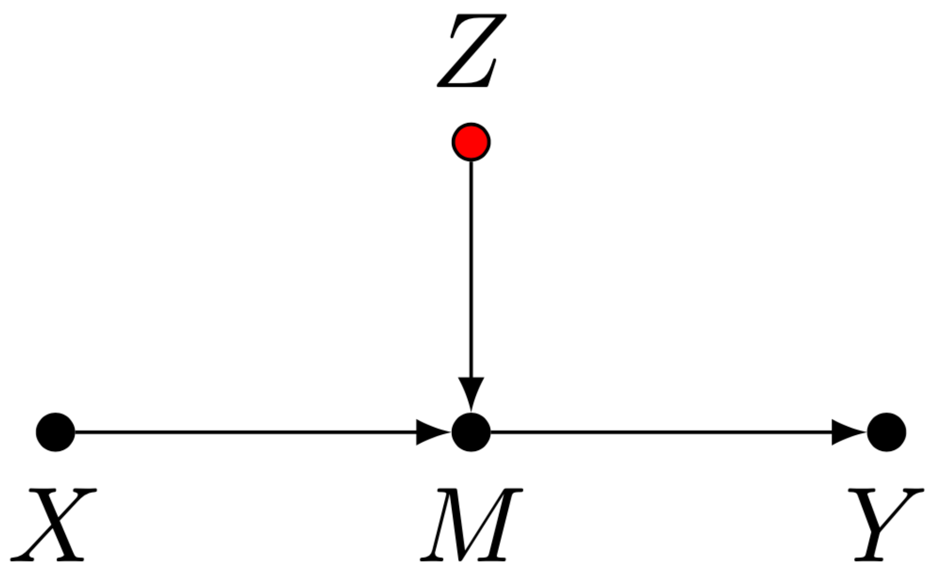
However, consider a variation of Model 11 with an unobserved confounder of Z and Y , denoted by U.

First notice that U does not confound the effect of X on Y. Thus the total effect remains unbiased, as it were in Model 11, *so long as we do not adjust for Z.*

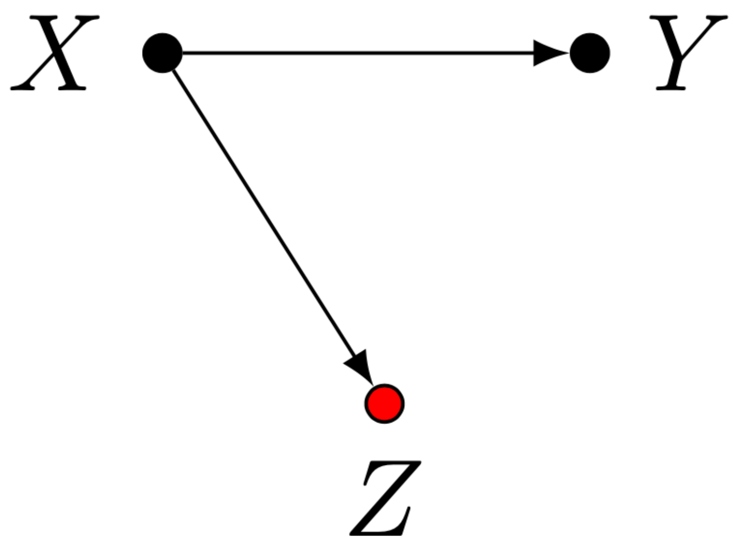
On the other hand, here adjusting for Z now opens the colliding path

$X \rightarrow Z \leftarrow U \rightarrow Y$, thus biasing the CDE estimate.

"Neutral" Controls (but different impacts on precision)

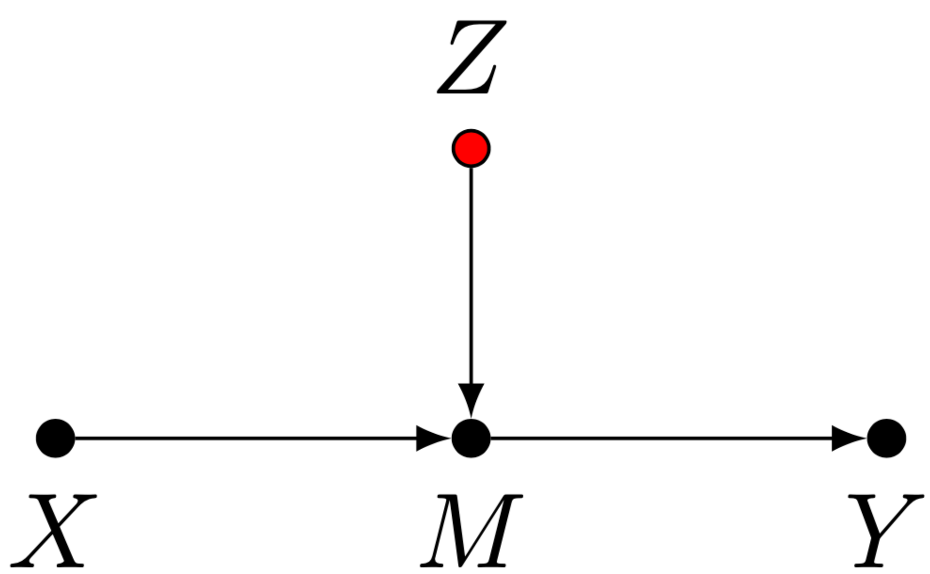


Model 13

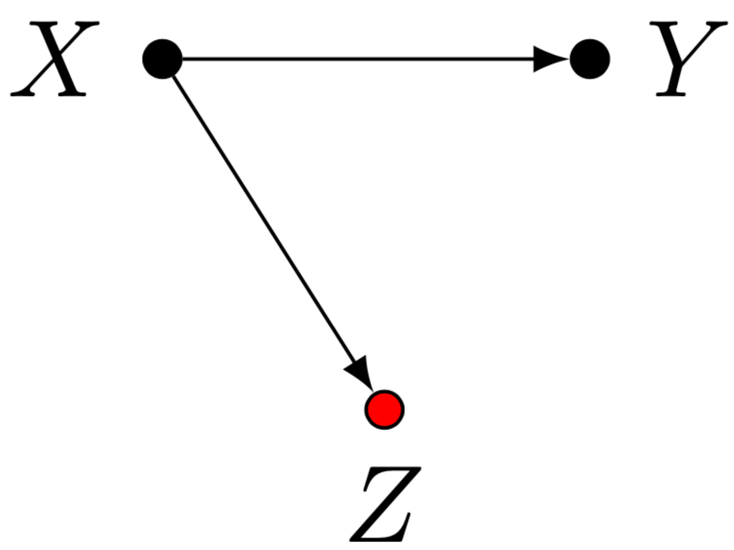


Model 14

“Neutral” Controls (but different impacts on precision)



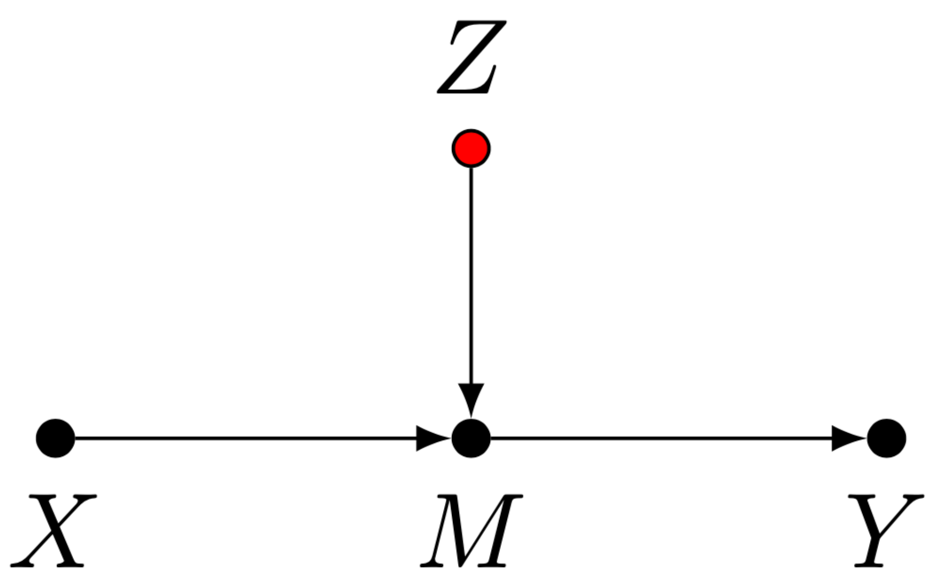
Model 13



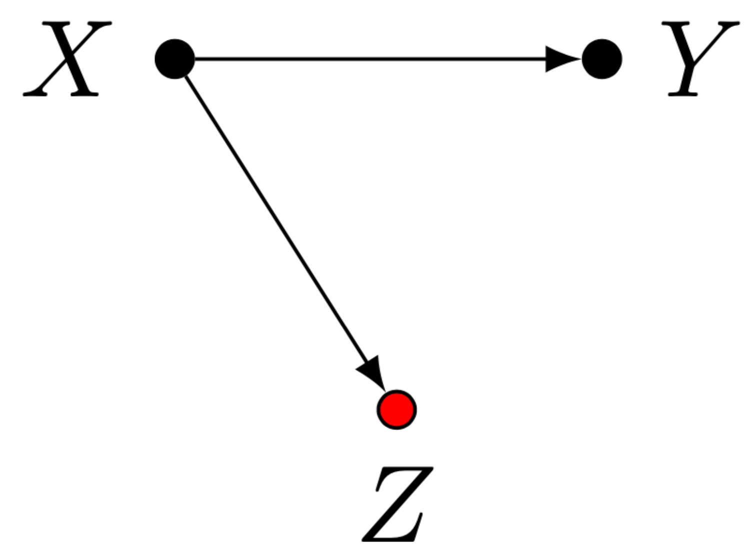
Model 14

At first look, one may think that adjusting for Z in Model 13 would bias the estimate of the ATE, by restricting variations of the mediator M.

"Neutral" Controls (but different impacts on precision)



Model 13

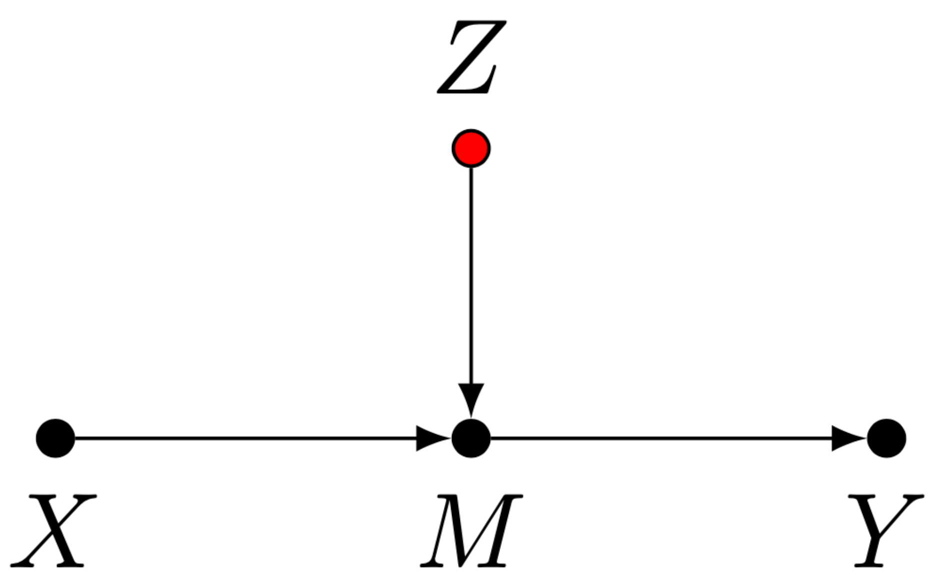


Model 14

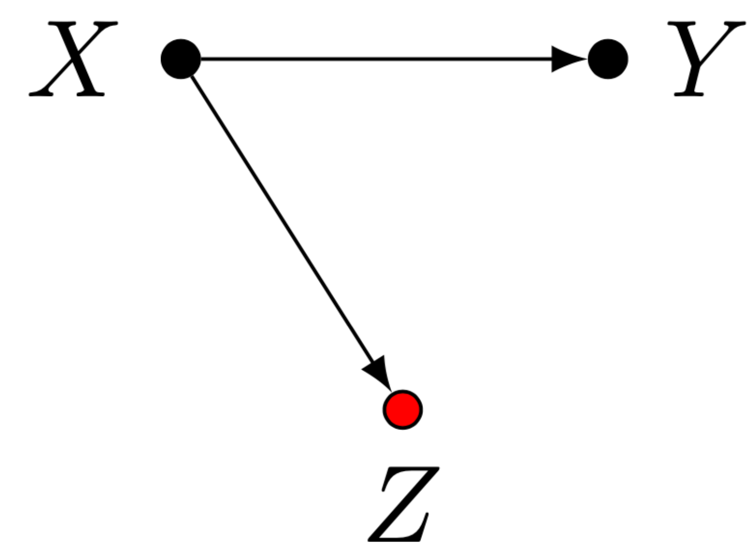
At first look, one may think that adjusting for Z in Model 13 would bias the estimate of the ATE, by restricting variations of the mediator M.

However, the key difference here is that Z is a *cause*, not an *effect*, of M. Thus, Model 13 is analogous to Model 8, and controlling for Z will be neutral in terms of bias, and may *improve* the precision of the ATE estimate in finite samples.

“Neutral” Controls (but different impacts on precision)



Model 13



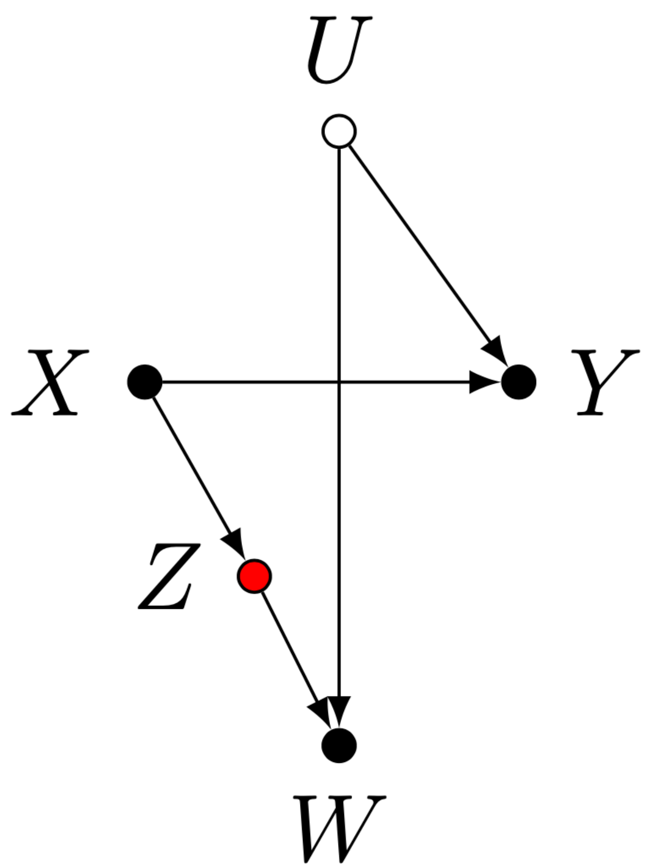
Model 14

At first look, one may think that adjusting for Z in Model 13 would bias the estimate of the ATE, by restricting variations of the mediator M.

However, the key difference here is that Z is a *cause*, not an *effect*, of M. Thus, Model 13 is analogous to Model 8, and controlling for Z will be neutral in terms of bias, and may *improve* the precision of the ATE estimate in finite samples.

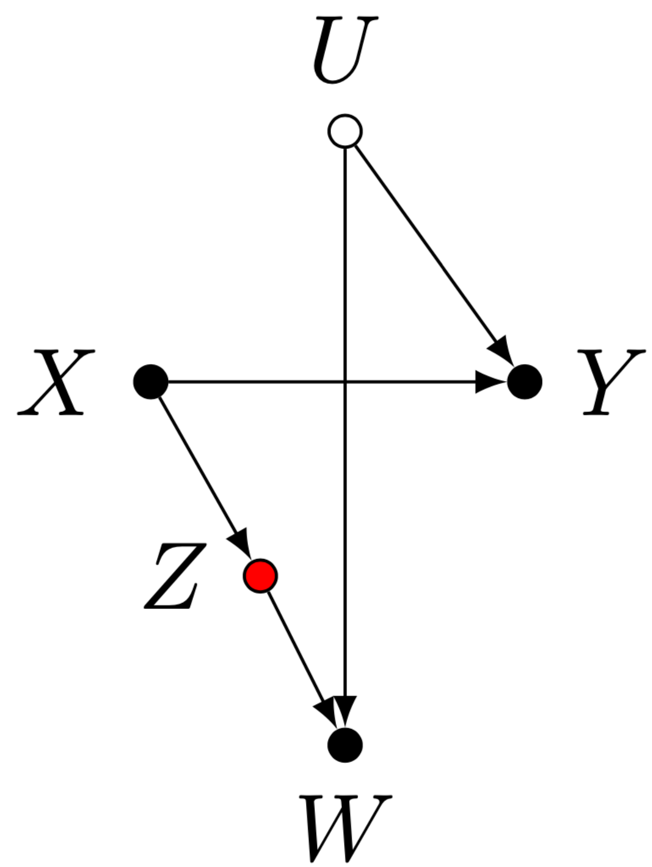
Contrary to folklore, not all “post-treatment” variables are inherently bad controls. In Model 14, Z is post-treatment, and controlling for Z does not open any confounding paths between X and Y. However, as before, controlling for Z may *hurt* the precision of the ACE estimate in finite samples.

“Good” control (in case of selection bias)



Model 15

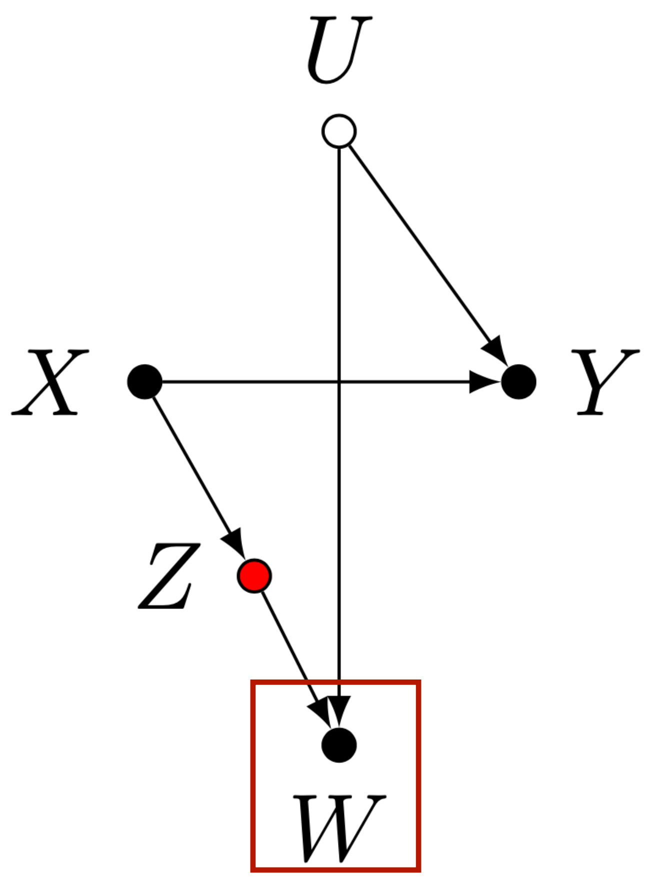
"Good" control (in case of selection bias)



Model 15

We now encounter our first "good" post-treatment variable.

"Good" control (in case of selection bias)

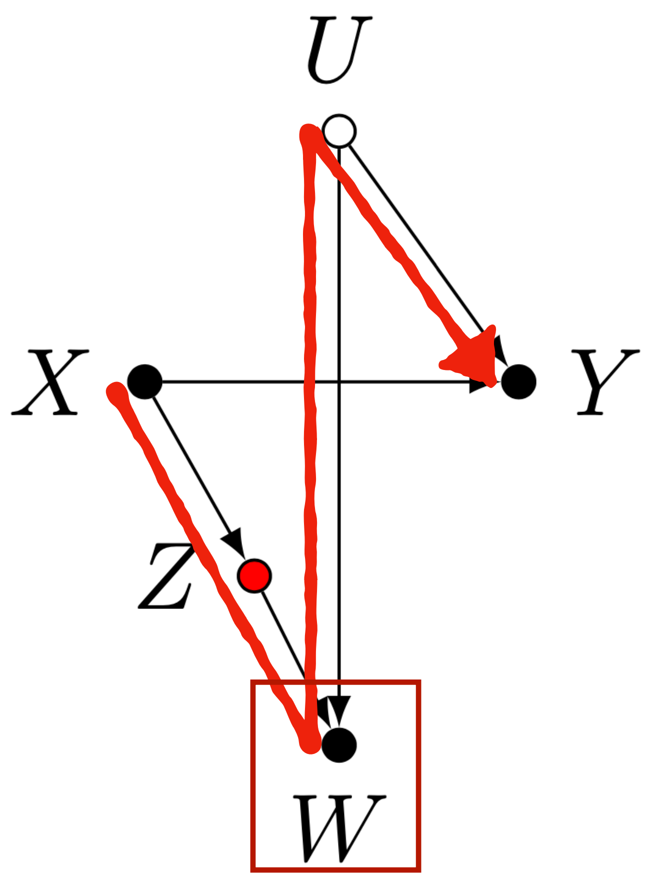


Model 15

We now encounter our first "good" post-treatment variable.

Suppose we only have observations with W=1 recorded. This means the colliding path X → Z →W← U → Y is opened.

"Good" control (in case of selection bias)

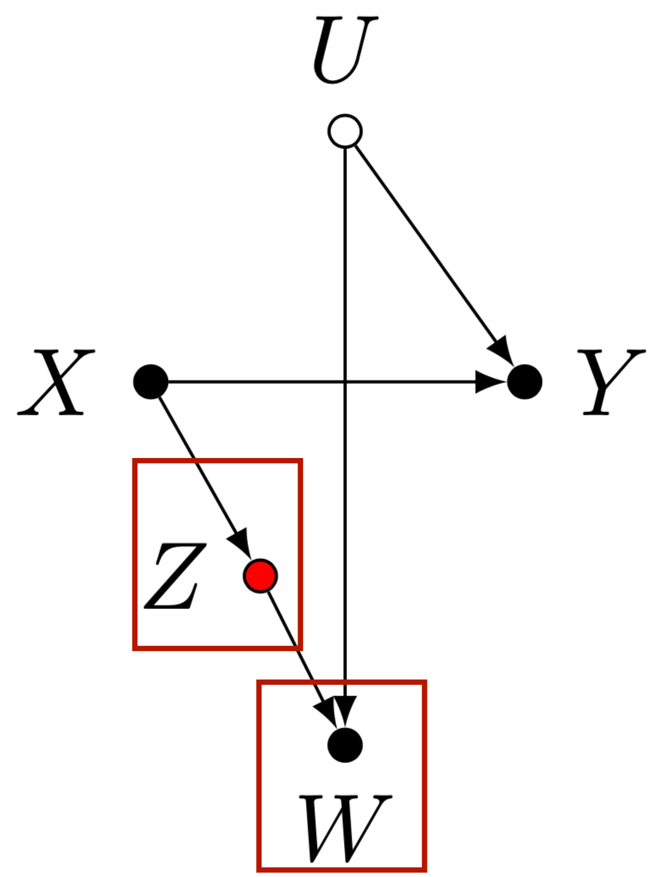


Model 15

We now encounter our first "good" post-treatment variable.

Suppose we only have observations with W=1 recorded. This means the colliding path X → Z →W← U → Y is opened.

“Good” control (in case of selection bias)



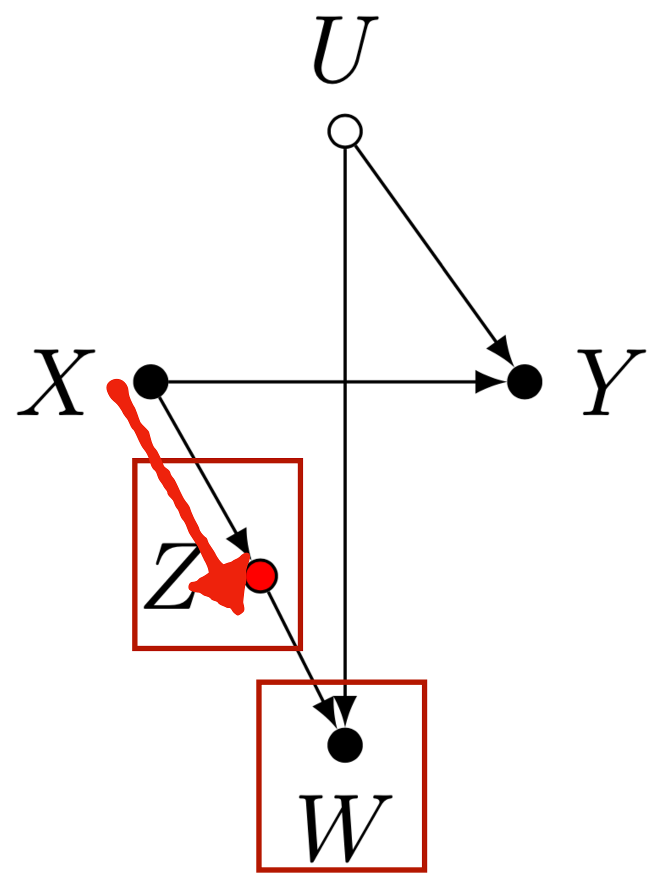
Model 15

We now encounter our first “good” post-treatment variable.

Suppose we only have observations with W=1 recorded. This means the colliding path X → Z →W← U → Y is opened.

However, further adjusting for Z blocks the colliding path. This allows us to estimate the conditional ATE (conditional on W=1). In linear models, we actually recover the full ATE.

"Good" control (in case of selection bias)



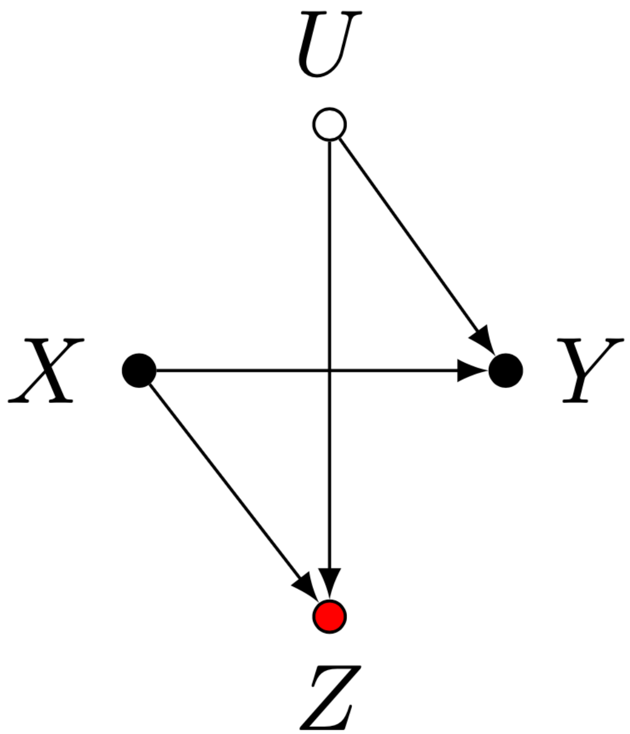
Model 15

We now encounter our first "good" post-treatment variable.

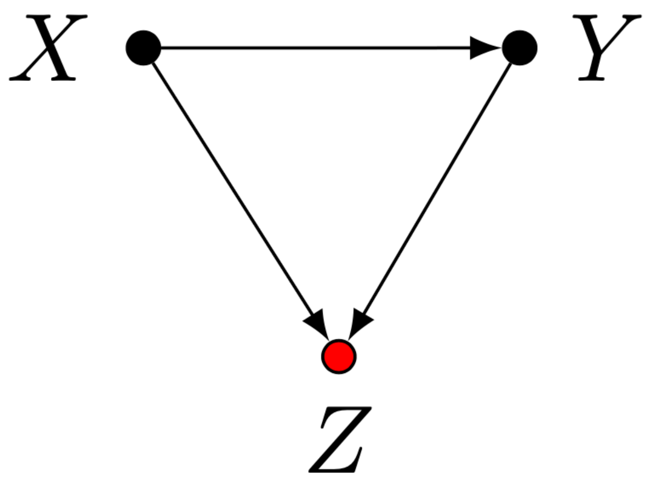
Suppose we only have observations with W=1 recorded. This means the colliding path X → Z →W← U → Y is opened.

However, further adjusting for Z blocks the colliding path. This allows us to estimate the conditional ATE (conditional on W=1). In linear models, we actually recover the full ATE.

"Bad" Controls (selection bias)

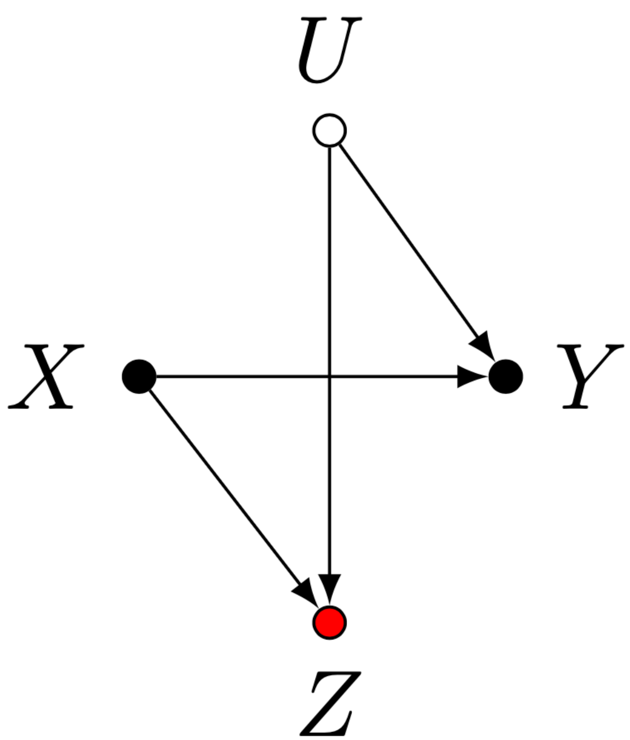


Model 16

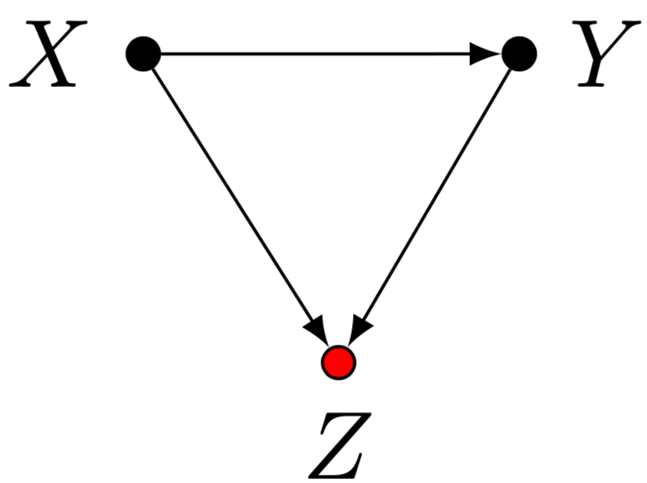


Model 17

"Bad" Controls (selection bias)



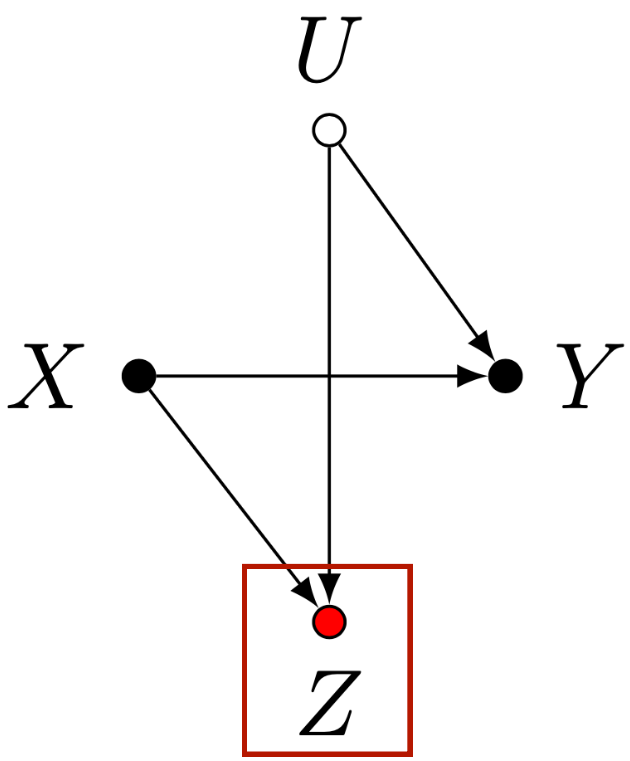
Model 16



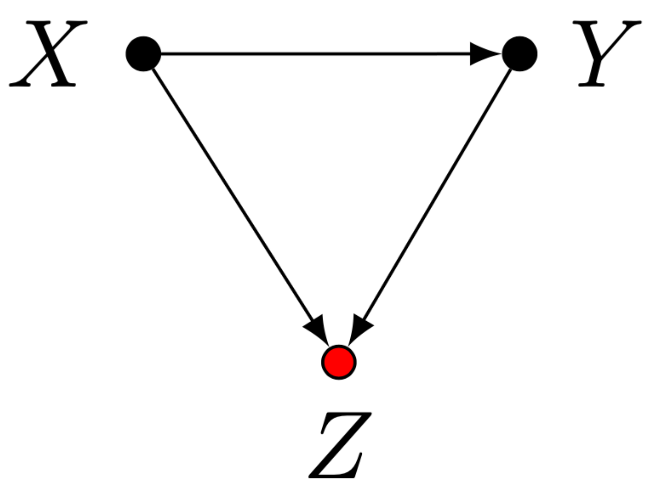
Model 17

Contrary to Models 14 and 15, here controlling for Z is no longer harmless, and induces what is classically known as "selection bias" or "collider bias."

"Bad" Controls (selection bias)



Model 16

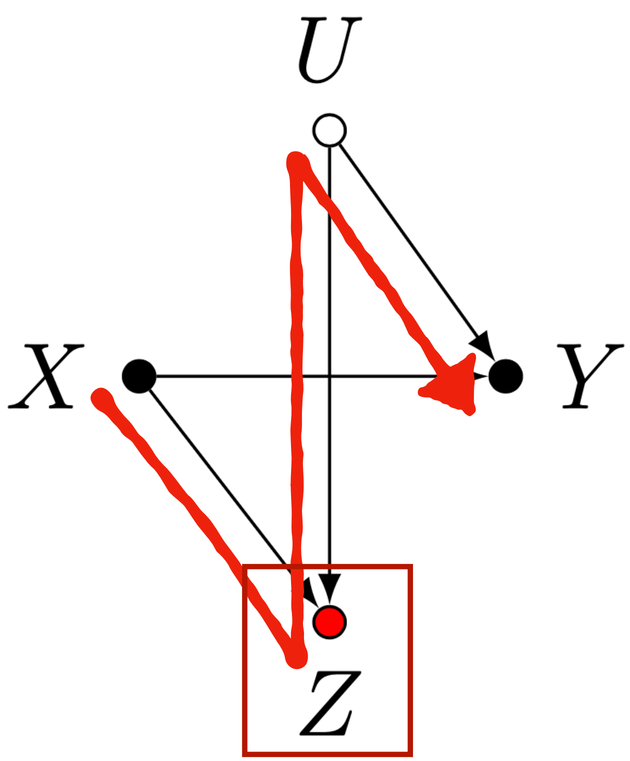


Model 17

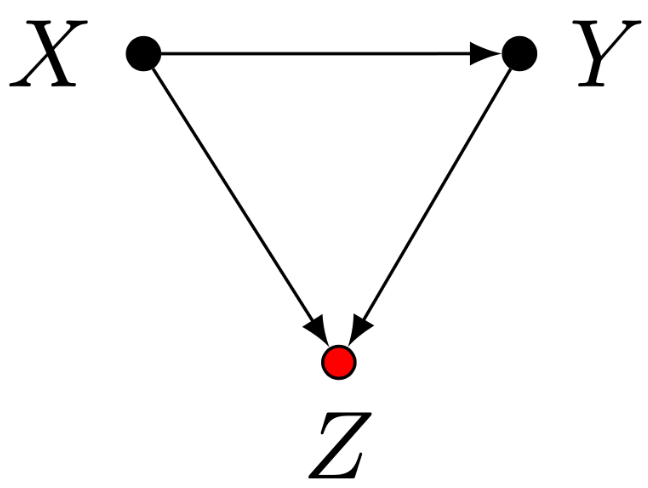
Contrary to Models 14 and 15, here controlling for Z is no longer harmless, and induces what is classically known as "selection bias" or "collider bias."

Adjusting for Z in Model 16 opens the colliding path X→Z←U→Y and so biases the ATE.

“Bad” Controls (selection bias)



Model 16

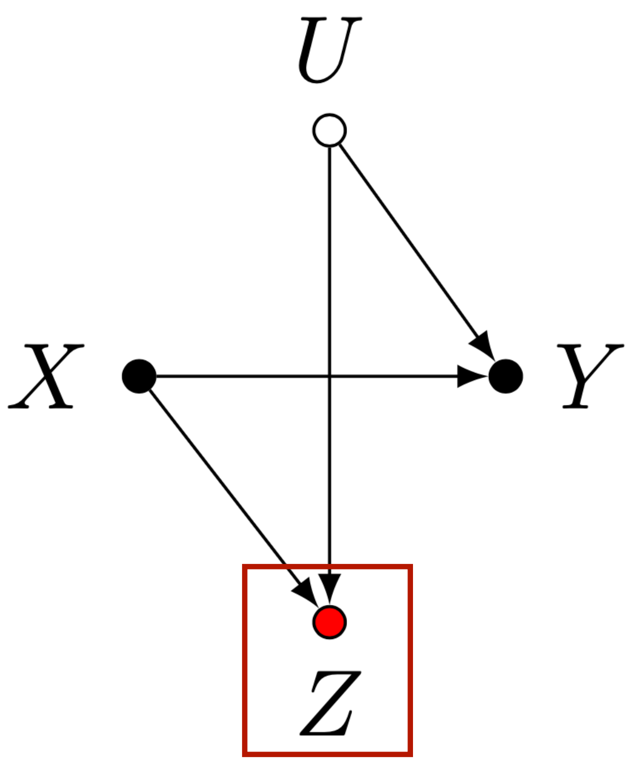


Model 17

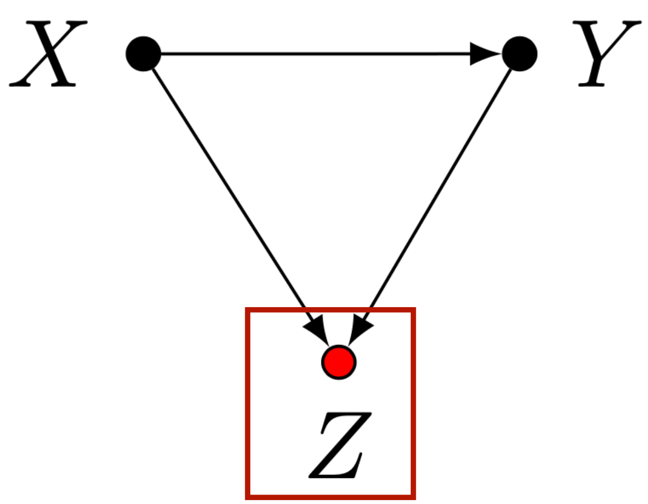
Contrary to Models 14 and 15, here controlling for Z is no longer harmless, and induces what is classically known as “selection bias” or “collider bias.”

Adjusting for Z in Model 16 opens the colliding path X→Z←U→Y and so biases the ATE.

“Bad” Controls (selection bias)



Model 16



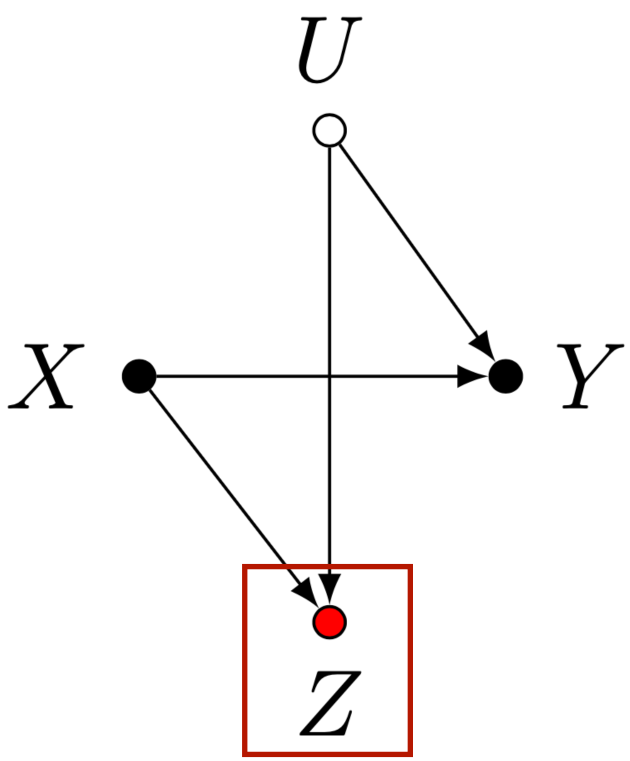
Model 17

Contrary to Models 14 and 15, here controlling for Z is no longer harmless, and induces what is classically known as “selection bias” or “collider bias.”

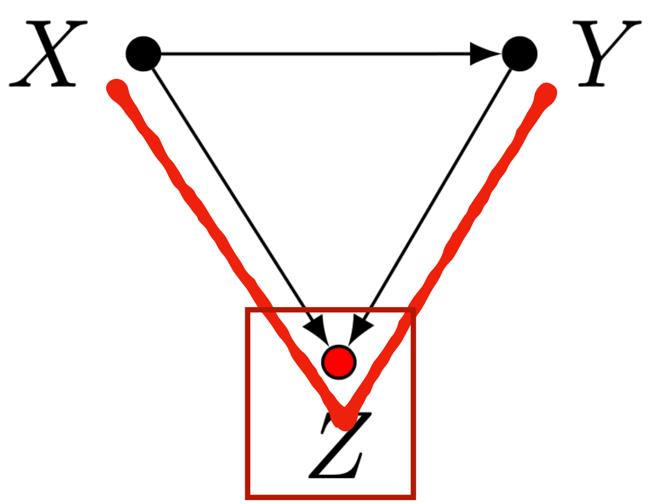
Adjusting for Z in Model 16 opens the colliding path X→Z←U→Y and so biases the ATE.

In Model 17, adjusting for Z not only opens the path X → Z ← Y , but also the colliding path due to the latent parents of Y, thus biasing the ACE and motivating our final example.

“Bad” Controls (selection bias)



Model 16



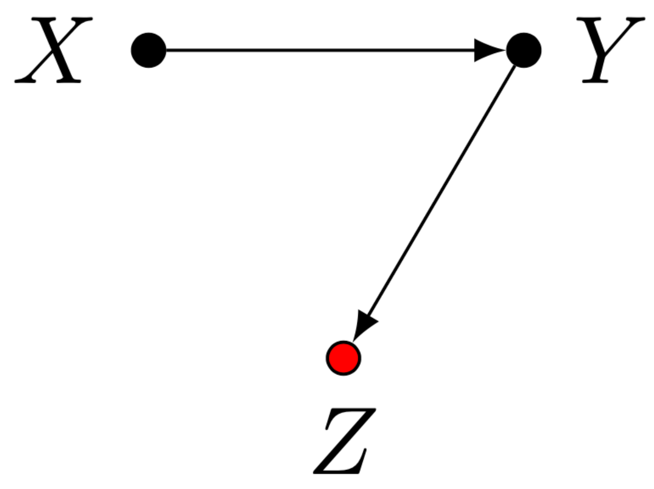
Model 17

Contrary to Models 14 and 15, here controlling for Z is no longer harmless, and induces what is classically known as “selection bias” or “collider bias.”

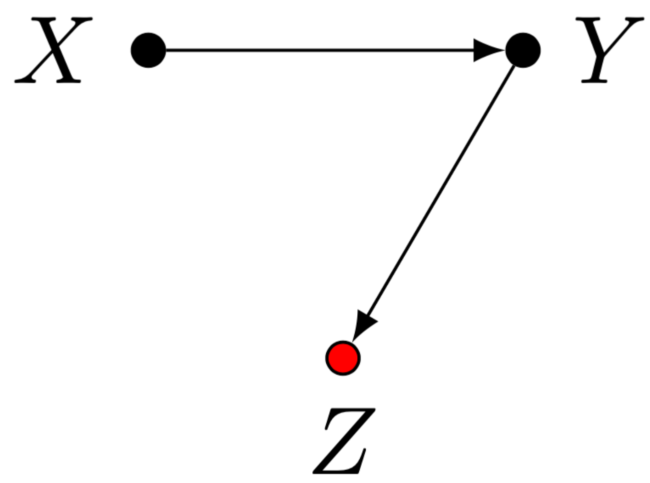
Adjusting for Z in Model 16 opens the colliding path X→Z←U→Y and so biases the ATE.

In Model 17, adjusting for Z not only opens the path X → Z ← Y , but also the colliding path due to the latent parents of Y, thus biasing the ACE and motivating our final example.

“Bad” Controls (case-control bias)



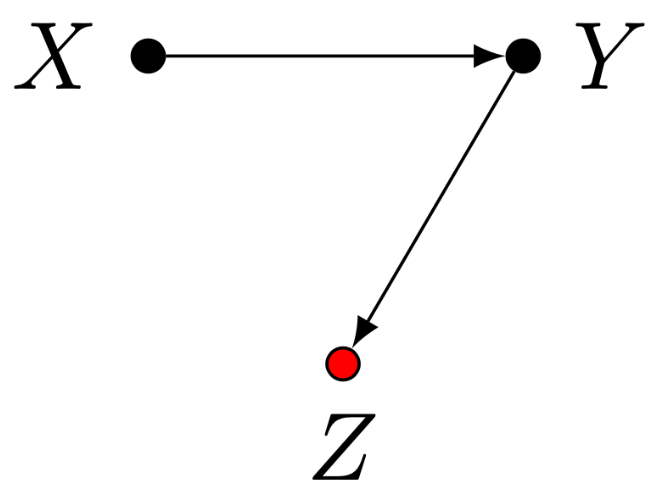
“Bad” Controls (case-control bias)



In our last example, Z is not in the causal pathway from X to Y , Z is not a direct cause of X, and Z is connected to Y .

Thus, one might surmise that, as in Model 8, controlling for Z is harmless for identification, and perhaps beneficial for finite sample efficiency.

“Bad” Controls (case-control bias)

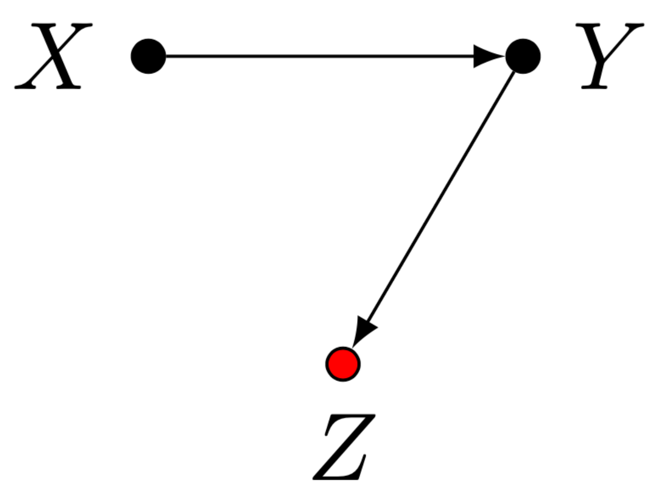


In our last example, Z is not in the causal pathway from X to Y , Z is not a direct cause of X, and Z is connected to Y .

Thus, one might surmise that, as in Model 8, controlling for Z is harmless for identification, and perhaps beneficial for finite sample efficiency.

However, controlling for the effects of the outcome Y will induce bias in the estimate of the ACE, even without the direct arrow X → Z, thus making Z a “bad control.”

“Bad” Controls (case-control bias)



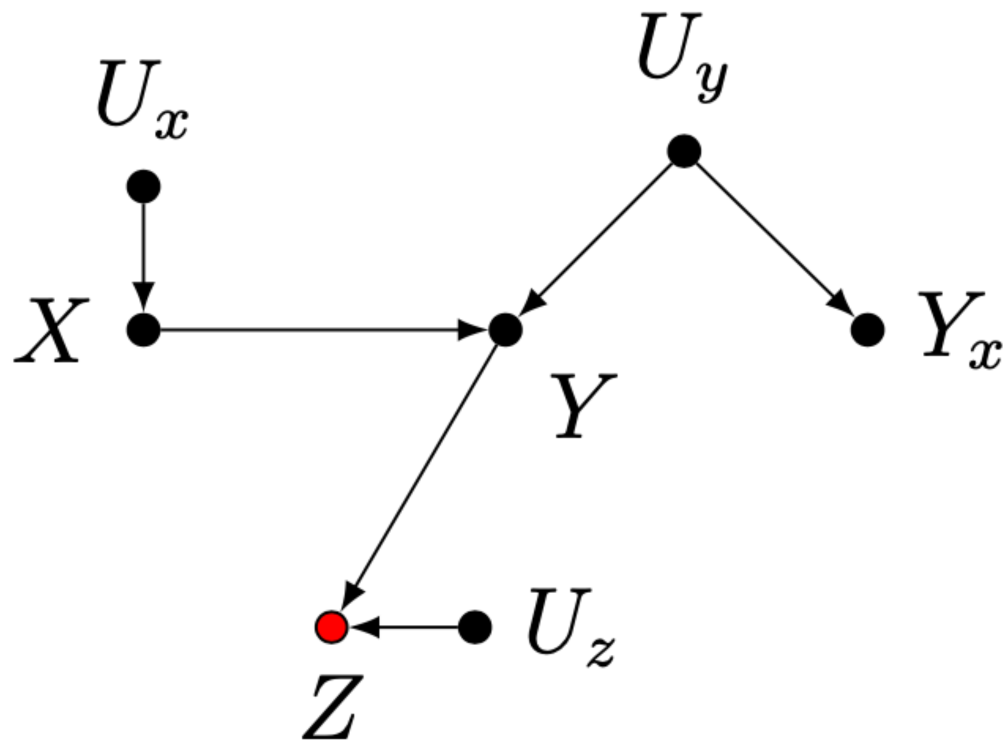
In our last example, Z is not in the causal pathway from X to Y , Z is not a direct cause of X, and Z is connected to Y .

Thus, one might surmise that, as in Model 8, controlling for Z is harmless for identification, and perhaps beneficial for finite sample efficiency.

However, controlling for the effects of the outcome Y will induce bias in the estimate of the ACE, even without the direct arrow X → Z, thus making Z a “bad control.”

One way to see this is drawing the potential outcome explicitly in the DAG, along with the latent parents of Y (see appendix of the Crash Course for details).

“Bad” Controls (case-control bias)



In our last example, Z is not in the causal pathway from X to Y , Z is not a direct cause of X, and Z is connected to Y .

Thus, one might surmise that, as in Model 8, controlling for Z is harmless for identification, and perhaps beneficial for finite sample efficiency.

However, controlling for the effects of the outcome Y will induce bias in the estimate of the ACE, even without the direct arrow $X \to Z$, thus making Z a “bad control.”

One way to see this is drawing the potential outcome explicitly in the DAG, along with the latent parents of Y (see appendix of the Crash Course for details).

Some Real Examples of Bad Controls

The birth-weight paradox

The birth-weight paradox

Infants born to smokers were found to have higher risks of mortality than infants born to non-smokers.

The birth-weight paradox

Infants born to smokers were found to have higher risks of mortality than infants born to non-smokers.

However, among infants with low birth-weight (LBW), this relationship was reversed.

The birth-weight paradox

Infants born to smokers were found to have higher risks of mortality than infants born to non-smokers.

However, among infants with low birth-weight (LBW), this relationship was reversed.

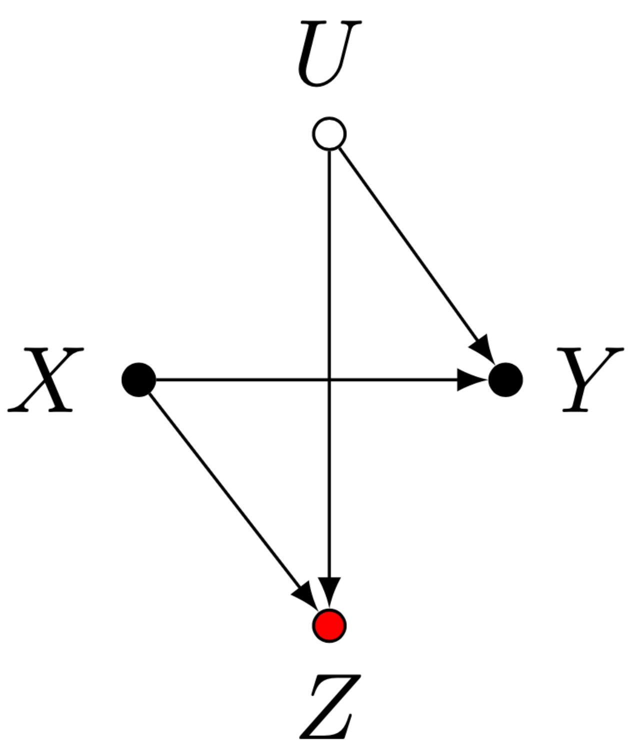
This reversal of effects has created many controversies in epidemiology—does it mean that maternal smoking is beneficial for LBW infants?

The birth-weight paradox

Infants born to smokers were found to have higher risks of mortality than infants born to non-smokers.

However, among infants with low birth-weight (LBW), this relationship was reversed.

This reversal of effects has created many controversies in epidemiology—does it mean that maternal smoking is beneficial for LBW infants?



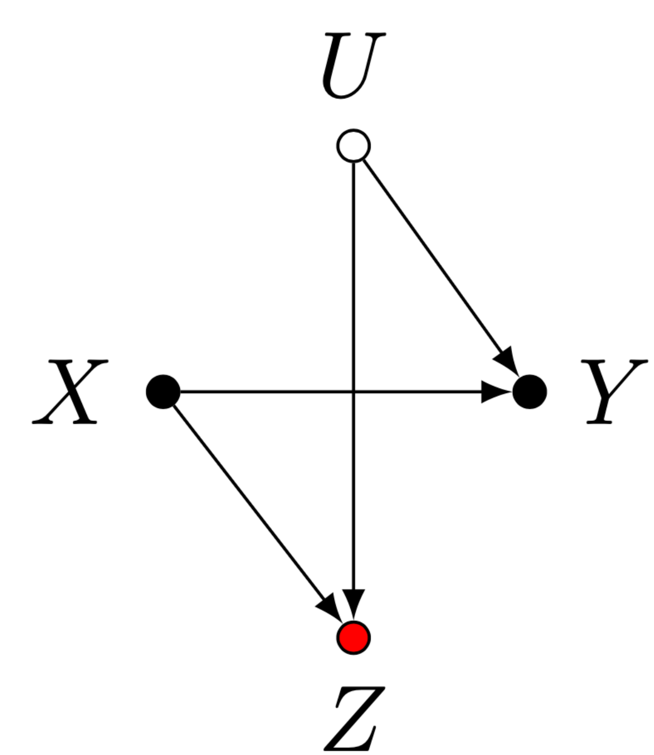
The birth-weight paradox

Infants born to smokers were found to have higher risks of mortality than infants born to non-smokers.

However, among infants with low birth-weight (LBW), this relationship was reversed.

This reversal of effects has created many controversies in epidemiology—does it mean that maternal smoking is beneficial for LBW infants?

Here X is maternal smoking, Y infant mortality, Z birth-weight, and U stands for unobserved risk-factors (such as birth-defects and malnutrition), that could also affect birth-weight.



The birth-weight paradox

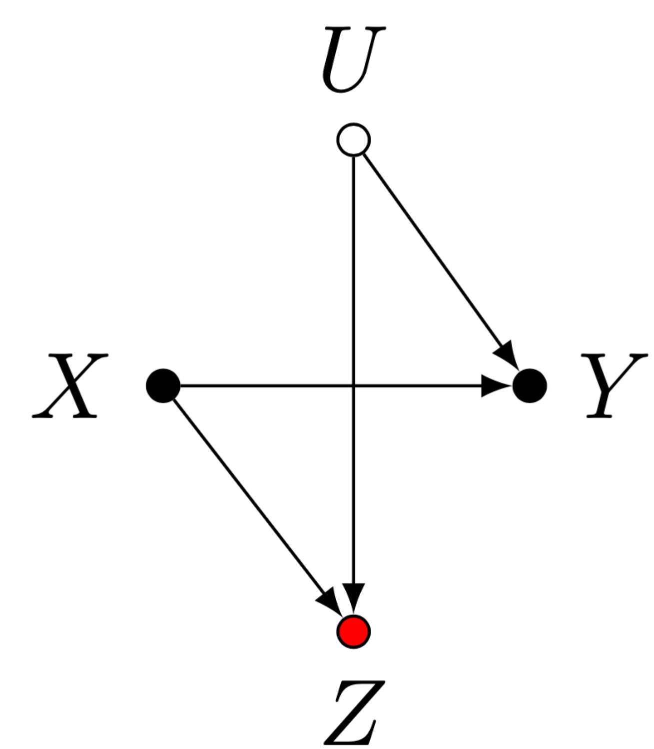
Infants born to smokers were found to have higher risks of mortality than infants born to non-smokers.

However, among infants with low birth-weight (LBW), this relationship was reversed.

This reversal of effects has created many controversies in epidemiology—does it mean that maternal smoking is beneficial for LBW infants?

Here X is maternal smoking, Y infant mortality, Z birth-weight, and U stands for unobserved risk-factors (such as birth-defects and malnutrition), that could also affect birth-weight.

Note that stratifying the analysis by birth-weight would induce a spurious association between smoking and mortality due to the competing risk-factors.



The birth-weight paradox

Infants born to smokers were found to have higher risks of mortality than infants born to non-smokers.

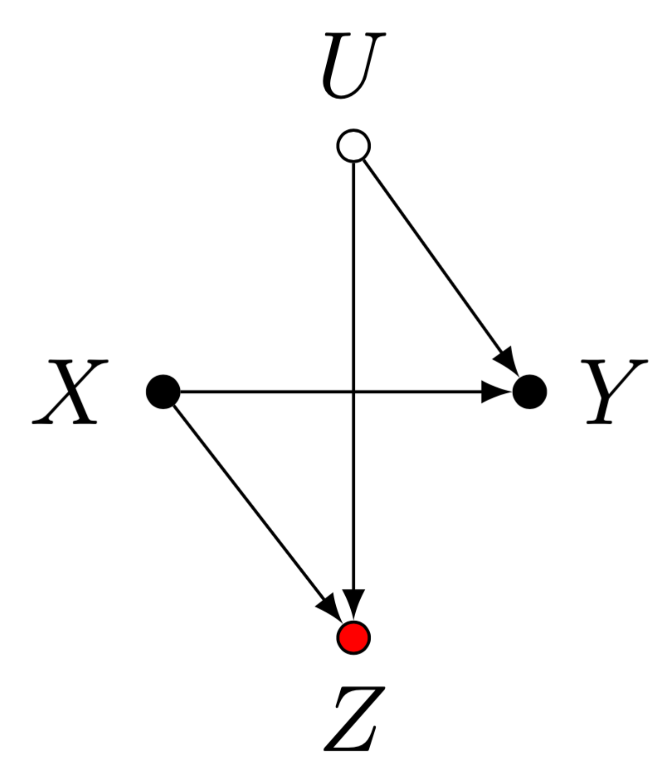
However, among infants with low birth-weight (LBW), this relationship was reversed.

This reversal of effects has created many controversies in epidemiology—does it mean that maternal smoking is beneficial for LBW infants?

Here X is maternal smoking, Y infant mortality, Z birth-weight, and U stands for unobserved risk-factors (such as birth-defects and malnutrition), that could also affect birth-weight.

Note that stratifying the analysis by birth-weight would induce a spurious association between smoking and mortality due to the competing risk-factors.

LBW infants of non-smokers need to have alternative causes for their LBW (such as malnutrition), and such causes could also lead to higher mortality.



The Antebellum Puzzle

The Antebellum Puzzle

An interesting puzzle of economic history is the fact that, during the nineteenth century in Britain and the United States, the average height of adult men fell even though the economic conditions of these countries improved alongside childhood nutrition.

The Antebellum Puzzle

An interesting puzzle of economic history is the fact that, during the nineteenth century in Britain and the United States, the average height of adult men fell even though the economic conditions of these countries improved alongside childhood nutrition.

Does better nutrition reduce the heights of adult men?

The Antebellum Puzzle

An interesting puzzle of economic history is the fact that, during the nineteenth century in Britain and the United States, the average height of adult men fell even though the economic conditions of these countries improved alongside childhood nutrition.

Does better nutrition reduce the heights of adult men?

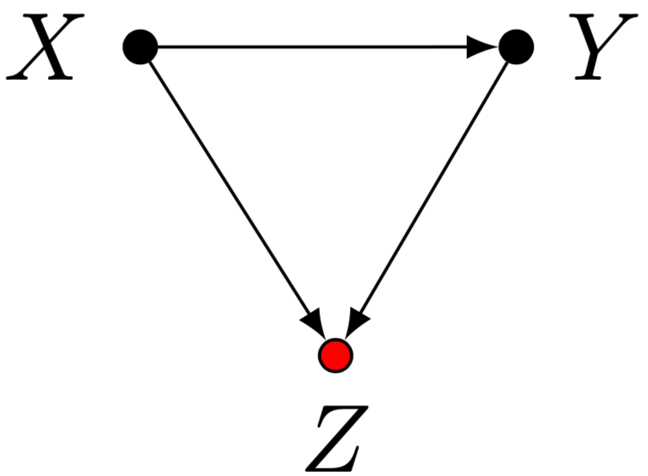
One possible explanation for such a paradoxical finding is selection bias.

The Antebellum Puzzle

An interesting puzzle of economic history is the fact that, during the nineteenth century in Britain and the United States, the average height of adult men fell even though the economic conditions of these countries improved alongside childhood nutrition.

Does better nutrition reduce the heights of adult men?

One possible explanation for such a paradoxical finding is selection bias.



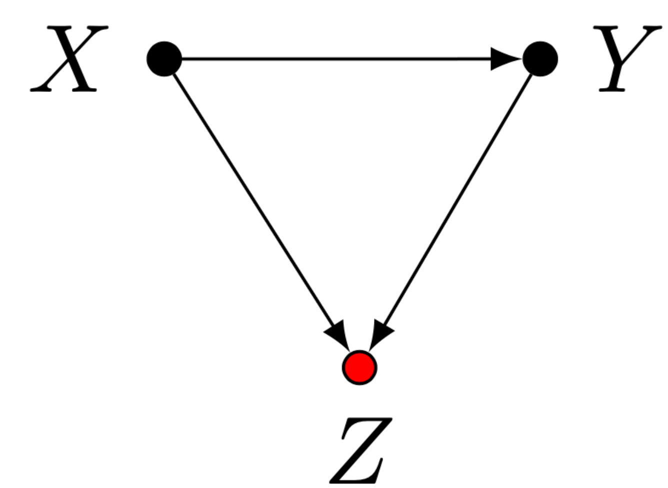
The Antebellum Puzzle

An interesting puzzle of economic history is the fact that, during the nineteenth century in Britain and the United States, the average height of adult men fell even though the economic conditions of these countries improved alongside childhood nutrition.

Does better nutrition reduce the heights of adult men?

One possible explanation for such a paradoxical finding is selection bias.

Researchers do not use data from the whole population, but from individuals enlisted in the military or in prison, which is equivalent to conditioning on colliders.



The Antebellum Puzzle

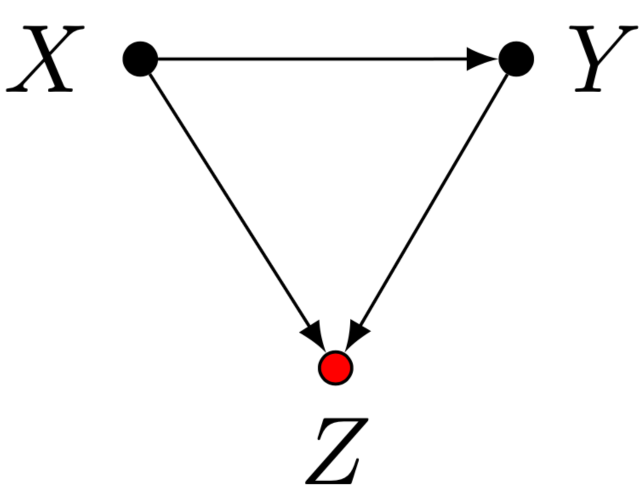
An interesting puzzle of economic history is the fact that, during the nineteenth century in Britain and the United States, the average height of adult men fell even though the economic conditions of these countries improved alongside childhood nutrition.

Does better nutrition reduce the heights of adult men?

One possible explanation for such a paradoxical finding is selection bias.

Researchers do not use data from the whole population, but from individuals enlisted in the military or in prison, which is equivalent to conditioning on colliders.

Consider the case of prison records. Let X be childhood nutrition, Y adult height, and let Z be an indicator of whether the individual was arrested.



The Antebellum Puzzle

An interesting puzzle of economic history is the fact that, during the nineteenth century in Britain and the United States, the average height of adult men fell even though the economic conditions of these countries improved alongside childhood nutrition.

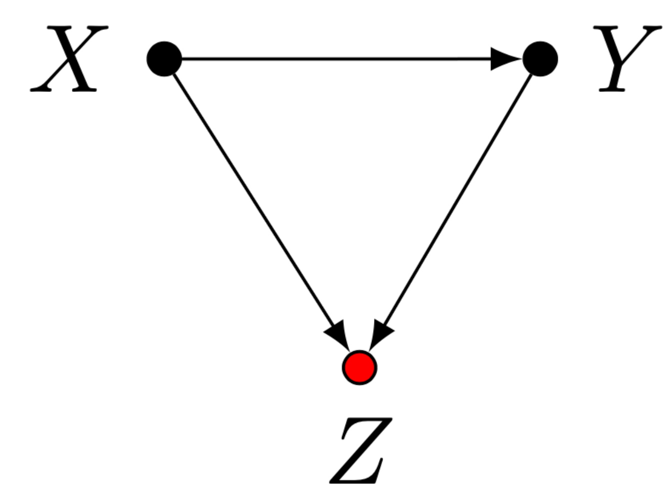
Does better nutrition reduce the heights of adult men?

One possible explanation for such a paradoxical finding is selection bias.

Researchers do not use data from the whole population, but from individuals enlisted in the military or in prison, which is equivalent to conditioning on colliders.

Consider the case of prison records. Let X be childhood nutrition, Y adult height, and let Z be an indicator of whether the individual was arrested.

Here one could argue that both childhood nutrition and adult height have pathways to committing a crime through socioeconomic opportunities, leading to selection bias.



Why you should not adjust for reading grade when estimating the causal effect of class size on math grade?

Jeffrey Wooldridge
@jmwooldridge

Suppose I have two 4th grade test scores, math4 and read4. I want to estimate the causal effect of class size on performance. Assume I have convincing controls. Is there a way to use a DAG to illustrate why I shouldn't include read4 in the equation for math4?

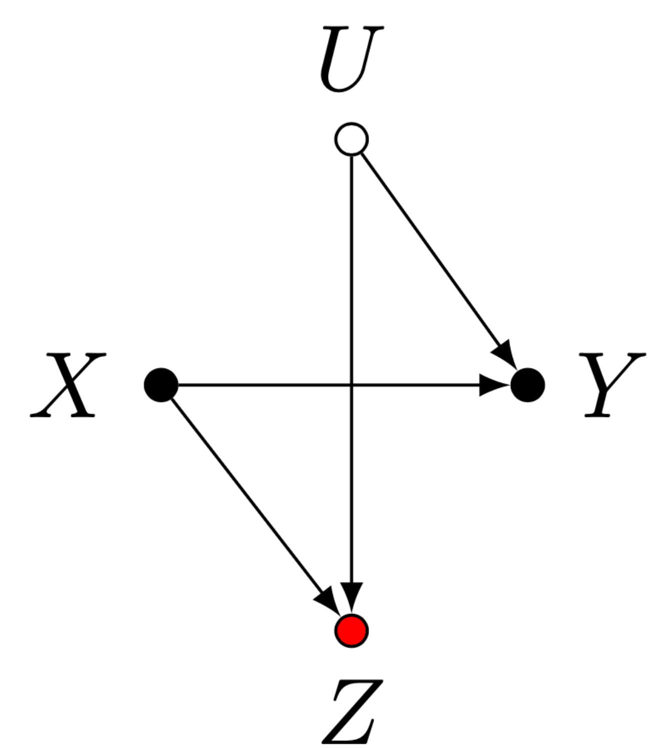
Why you should not adjust for reading grade when estimating the causal effect of class size on math grade?

Jeffrey Wooldridge
@jmwooldridge

Suppose I have two 4th grade test scores, math4 and read4. I want to estimate the causal effect of class size on performance. Assume I have convincing controls. Is there a way to use a DAG to illustrate why I shouldn't include read4 in the equation for math4?

Análise Real
@analisereal

We can illustrate this with Model 16 of the "Crash Course in Good and Bad Controls" (papers.ssrn.com/sol3/papers.cf…). Here X = class size, Y = math4, Z = read4, and U = student's ability. Conditioning on Z opens the path X -> Z <- U -> Y and it is thus a "bad control."



What About Multiple Controls?

Multiple Controls

Multiple Controls

When considering multiple controls, the status of a single control as "good" or "bad" may change depending on the context of the other variables under consideration.

Multiple Controls

When considering multiple controls, the status of a single control as “good” or “bad” may change depending on the context of the other variables under consideration.

Nevertheless, the main lessons from our illustrative examples remain.

Multiple Controls

When considering multiple controls, the status of a single control as "good" or "bad" may change depending on the context of the other variables under consideration.

Nevertheless, the main lessons from our illustrative examples remain.

A set of control variables Z will be "good" if:

Multiple Controls

When considering multiple controls, the status of a single control as "good" or "bad" may change depending on the context of the other variables under consideration.

Nevertheless, the main lessons from our illustrative examples remain.

A set of control variables Z will be "good" if:

(i) it blocks all non-causal paths from the treatment to the outcome;

Multiple Controls

When considering multiple controls, the status of a single control as "good" or "bad" may change depending on the context of the other variables under consideration.

Nevertheless, the main lessons from our illustrative examples remain.

A set of control variables Z will be "good" if:

(i) it blocks all non-causal paths from the treatment to the outcome;

(ii) it leaves any mediating paths from the treatment to the outcome "untouched" (since we are interested in the total effect); and,

Multiple Controls

When considering multiple controls, the status of a single control as "good" or "bad" may change depending on the context of the other variables under consideration.

Nevertheless, the main lessons from our illustrative examples remain.

A set of control variables Z will be "good" if:

(i) it blocks all non-causal paths from the treatment to the outcome;

(ii) it leaves any mediating paths from the treatment to the outcome "untouched" (since we are interested in the total effect); and,

(iii) it does not open new spurious paths between the treatment and the outcome (e.g., due to colliders).

Multiple Controls

When considering multiple controls, the status of a single control as "good" or "bad" may change depending on the context of the other variables under consideration.

Nevertheless, the main lessons from our illustrative examples remain.

A set of control variables Z will be "good" if:

(i) it blocks all non-causal paths from the treatment to the outcome;

(ii) it leaves any mediating paths from the treatment to the outcome "untouched" (since we are interested in the total effect); and,

(iii) it does not open new spurious paths between the treatment and the outcome (e.g., due to colliders).

As to efficiency considerations, we should give preference to those variables "closer" to the outcome, in opposition to those closer to the treatment—so long as, of course, this does not spoil identification

Multiple Controls

When considering multiple controls, the status of a single control as "good" or "bad" may change depending on the context of the other variables under consideration.

Nevertheless, the main lessons from our illustrative examples remain.

A set of control variables Z will be "good" if:

(i) it blocks all non-causal paths from the treatment to the outcome;

(ii) it leaves any mediating paths from the treatment to the outcome "untouched" (since we are interested in the total effect); and,

(iii) it does not open new spurious paths between the treatment and the outcome (e.g., due to colliders).

As to efficiency considerations, we should give preference to those variables "closer" to the outcome, in opposition to those closer to the treatment—so long as, of course, this does not spoil identification

You don't need to do it by hand!

Multiple Controls

When considering multiple controls, the status of a single control as "good" or "bad" may change depending on the context of the other variables under consideration.

Nevertheless, the main lessons from our illustrative examples remain.

A set of control variables Z will be "good" if:

(i) it blocks all non-causal paths from the treatment to the outcome;

(ii) it leaves any mediating paths from the treatment to the outcome "untouched" (since we are interested in the total effect); and,

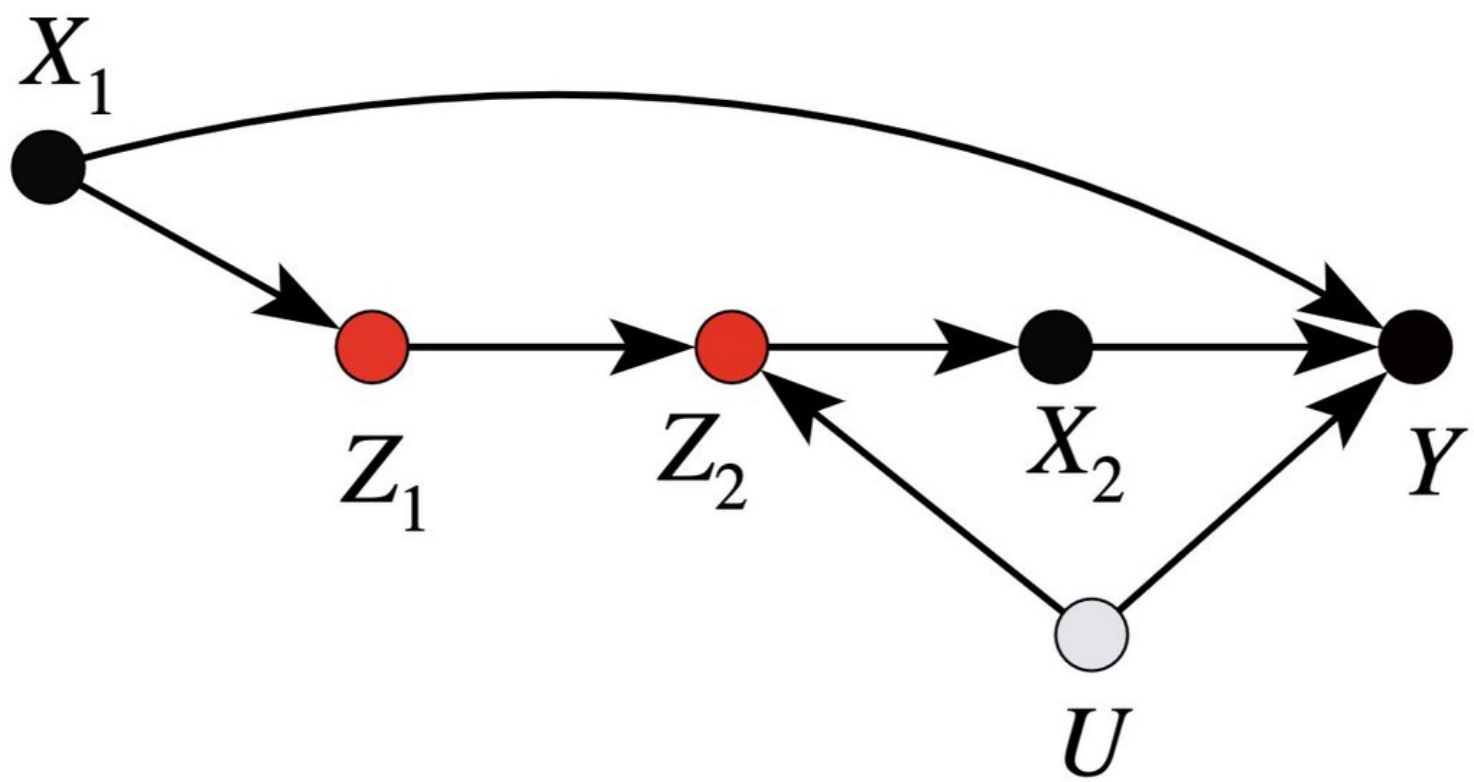
(iii) it does not open new spurious paths between the treatment and the outcome (e.g., due to colliders).

As to efficiency considerations, we should give preference to those variables "closer" to the outcome, in opposition to those closer to the treatment—so long as, of course, this does not spoil identification

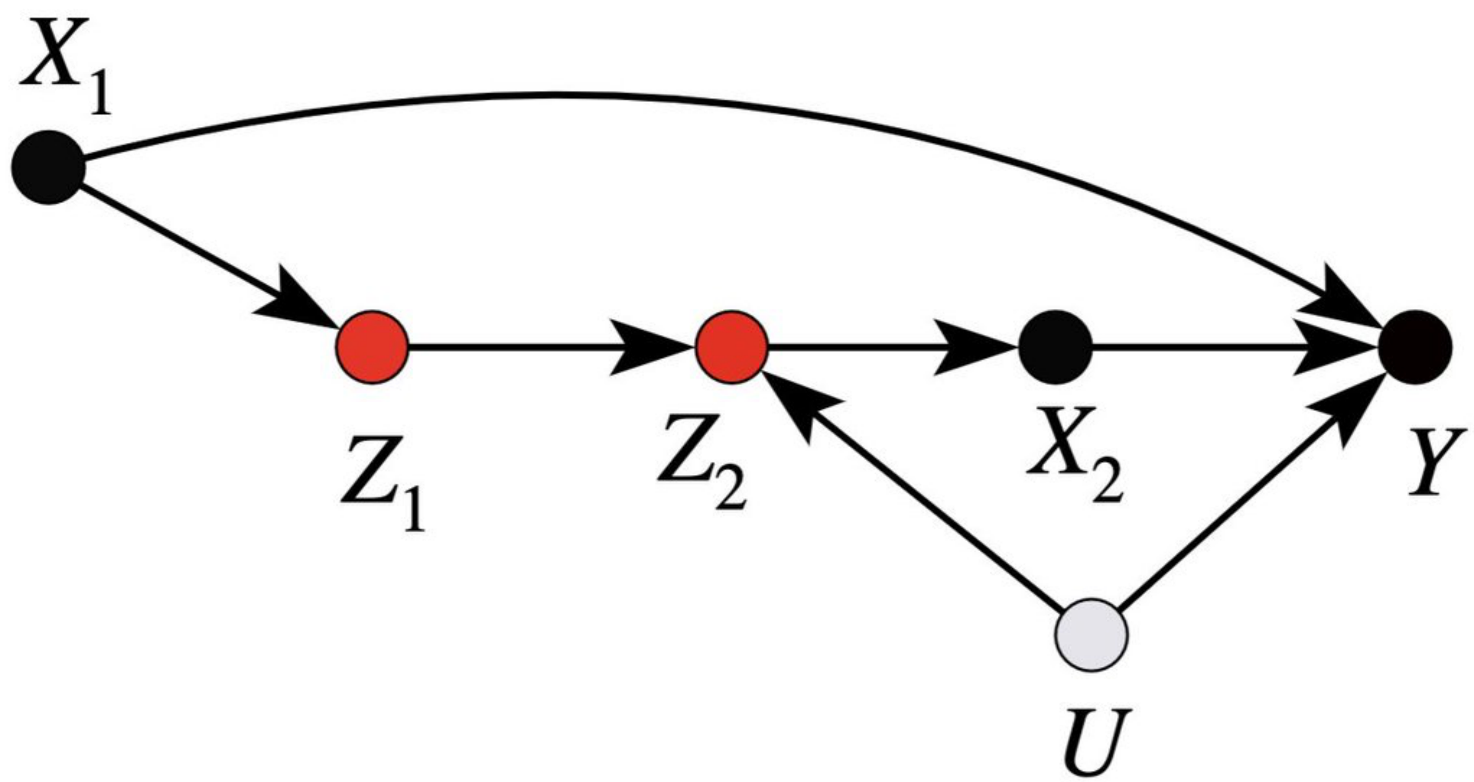
You don't need to do it by hand!

There is open-source software with efficient procedures to identify (optimal) adjustment sets for you (*dagitty, causal fusion, etc*).

Challenge:

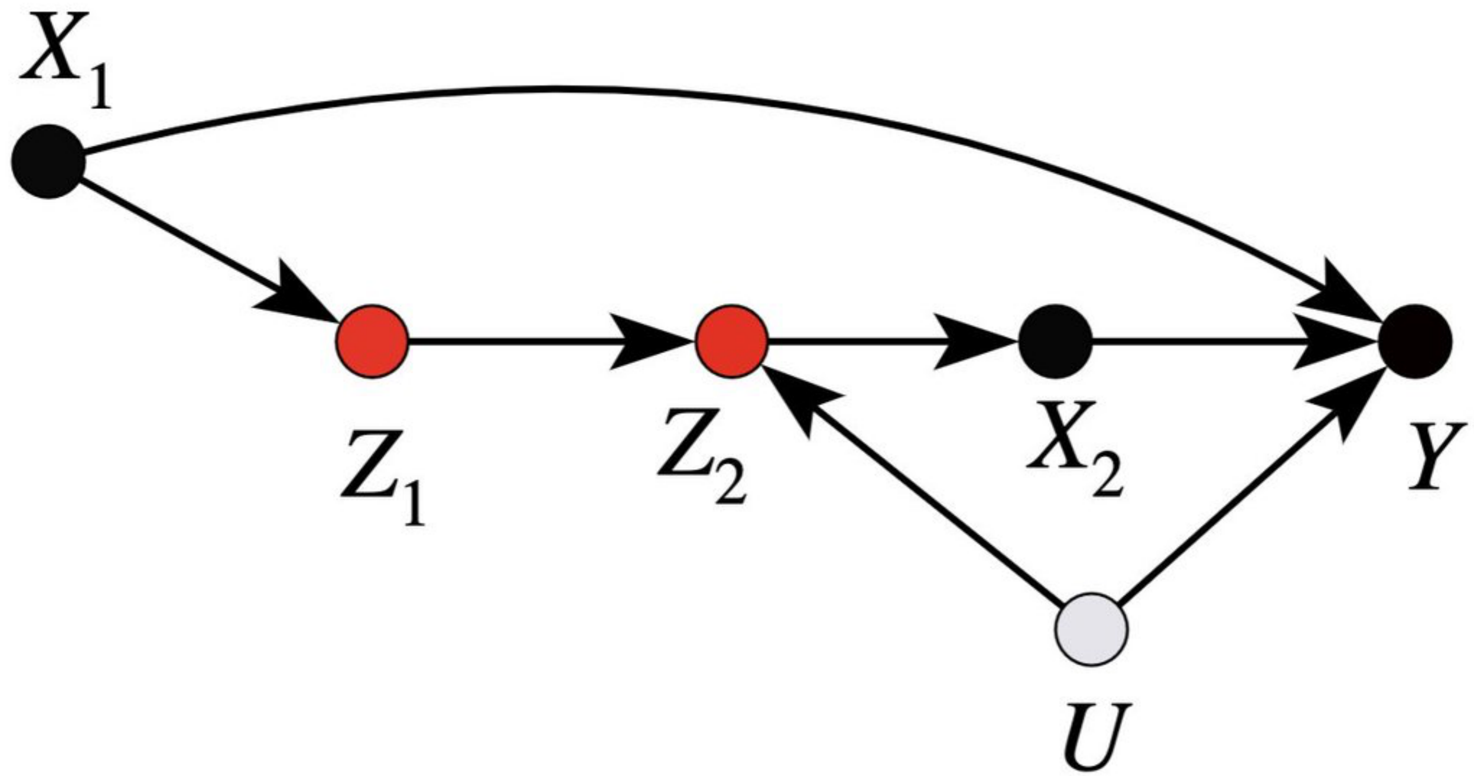


Challenge:



Goal: estimate the *joint effect* of X1 and X2 on Y.

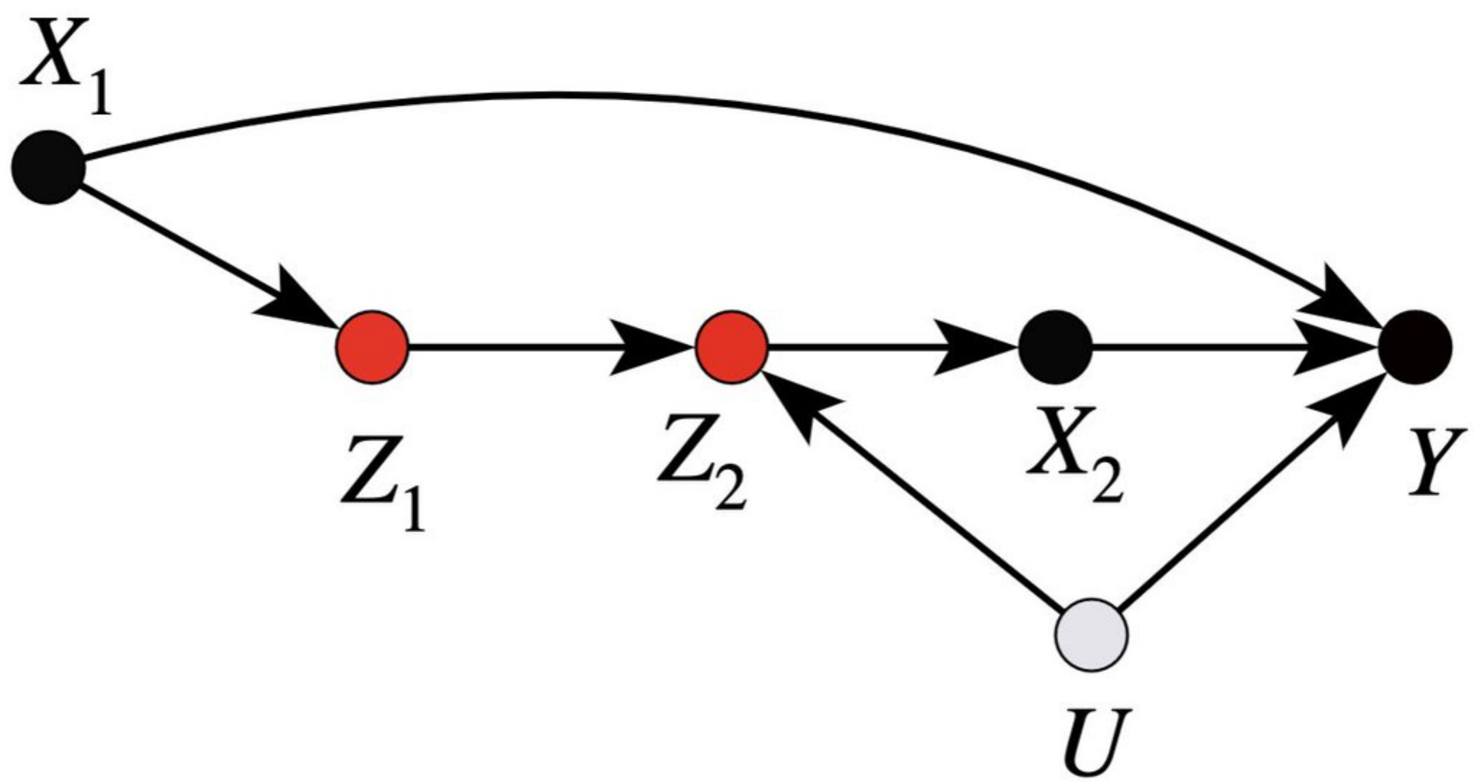
Challenge:



Goal: estimate the *joint effect* of X1 and X2 on Y.

Question: which variables should you include in the regression equation?

Challenge:

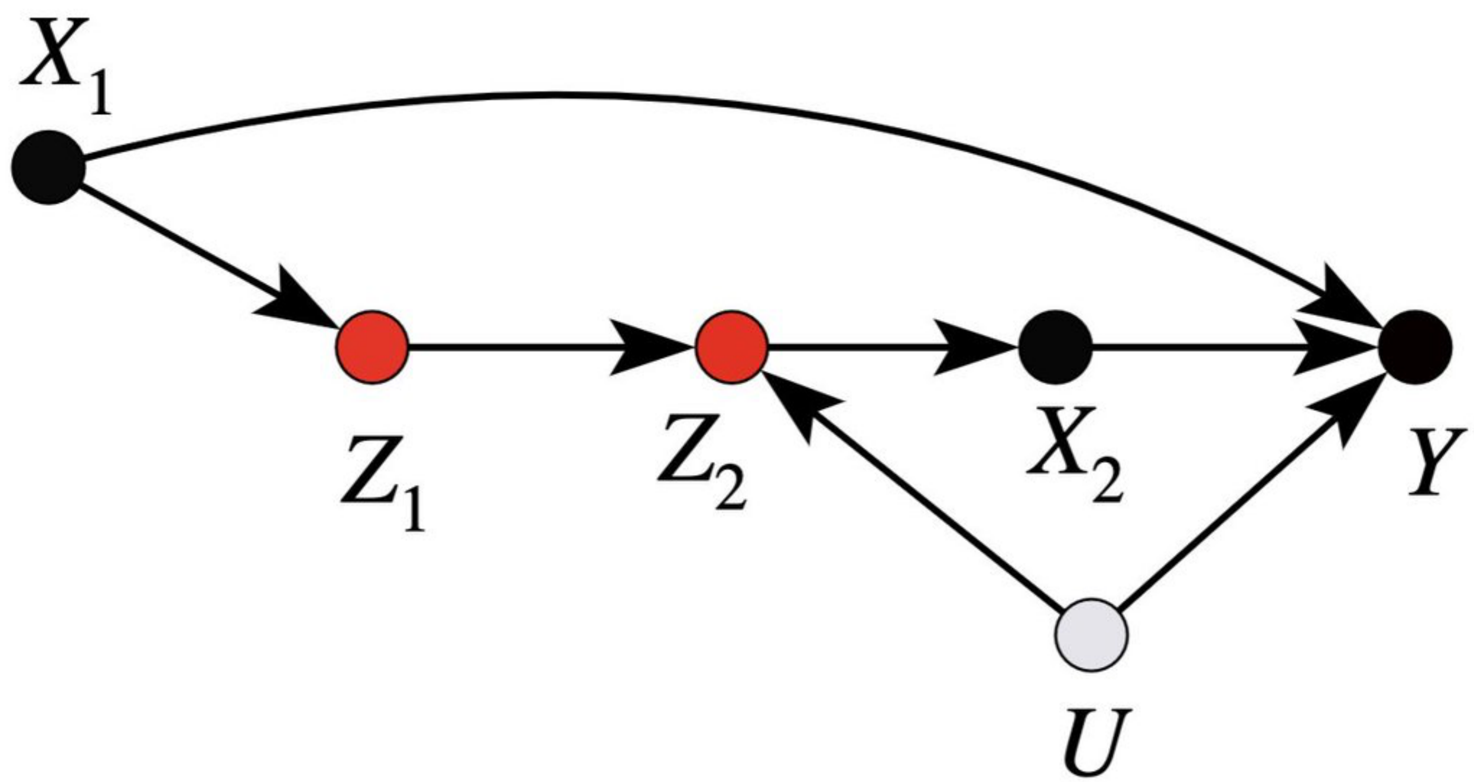


Goal: estimate the *joint effect* of X1 and X2 on Y.

Question: which variables should you include in the regression equation?

Apply the same first principles:

Challenge:



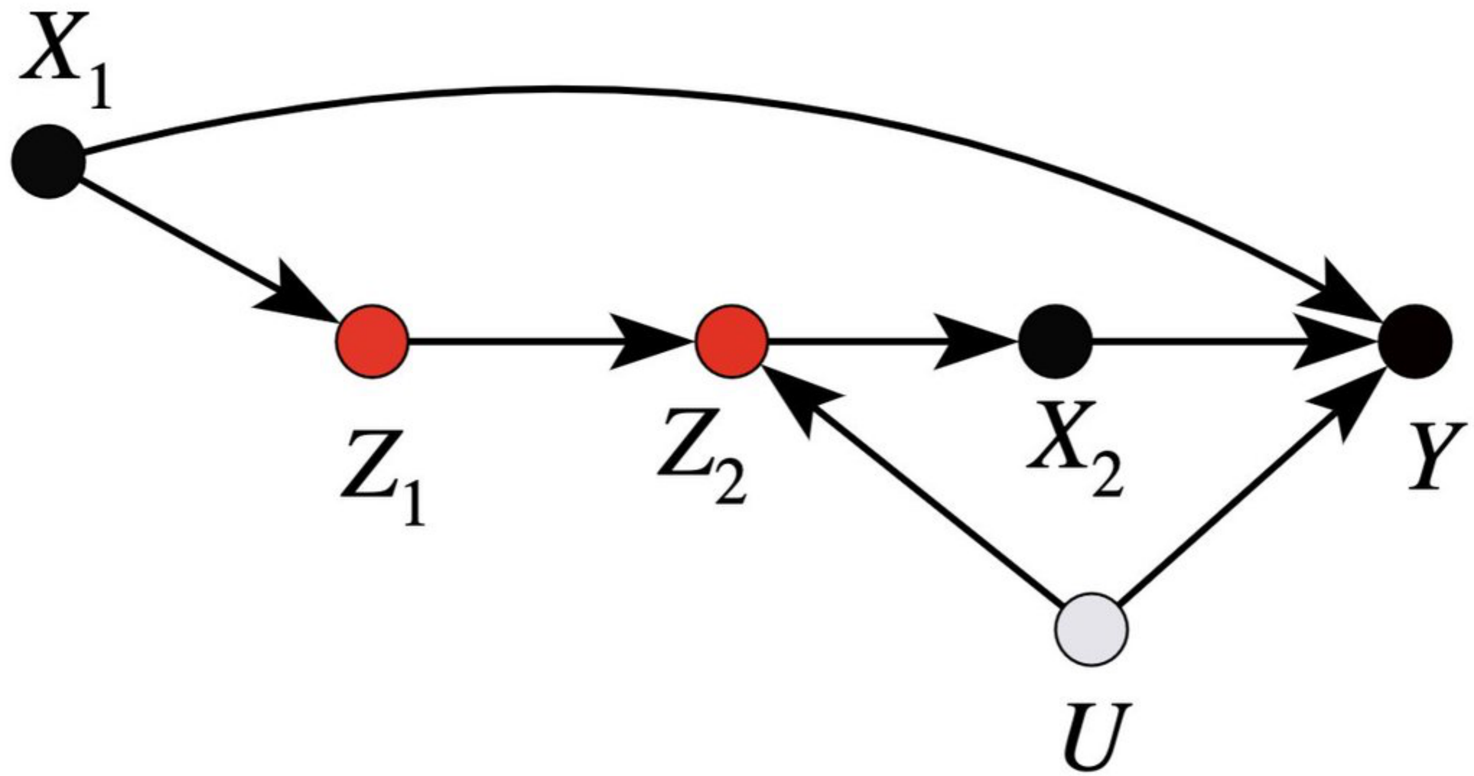
Goal: estimate the *joint effect* of X1 and X2 on Y.

Question: which variables should you include in the regression equation?

Apply the same first principles:

(i) do not close relevant causal paths,

Challenge:



Goal: estimate the *joint effect* of X1 and X2 on Y.

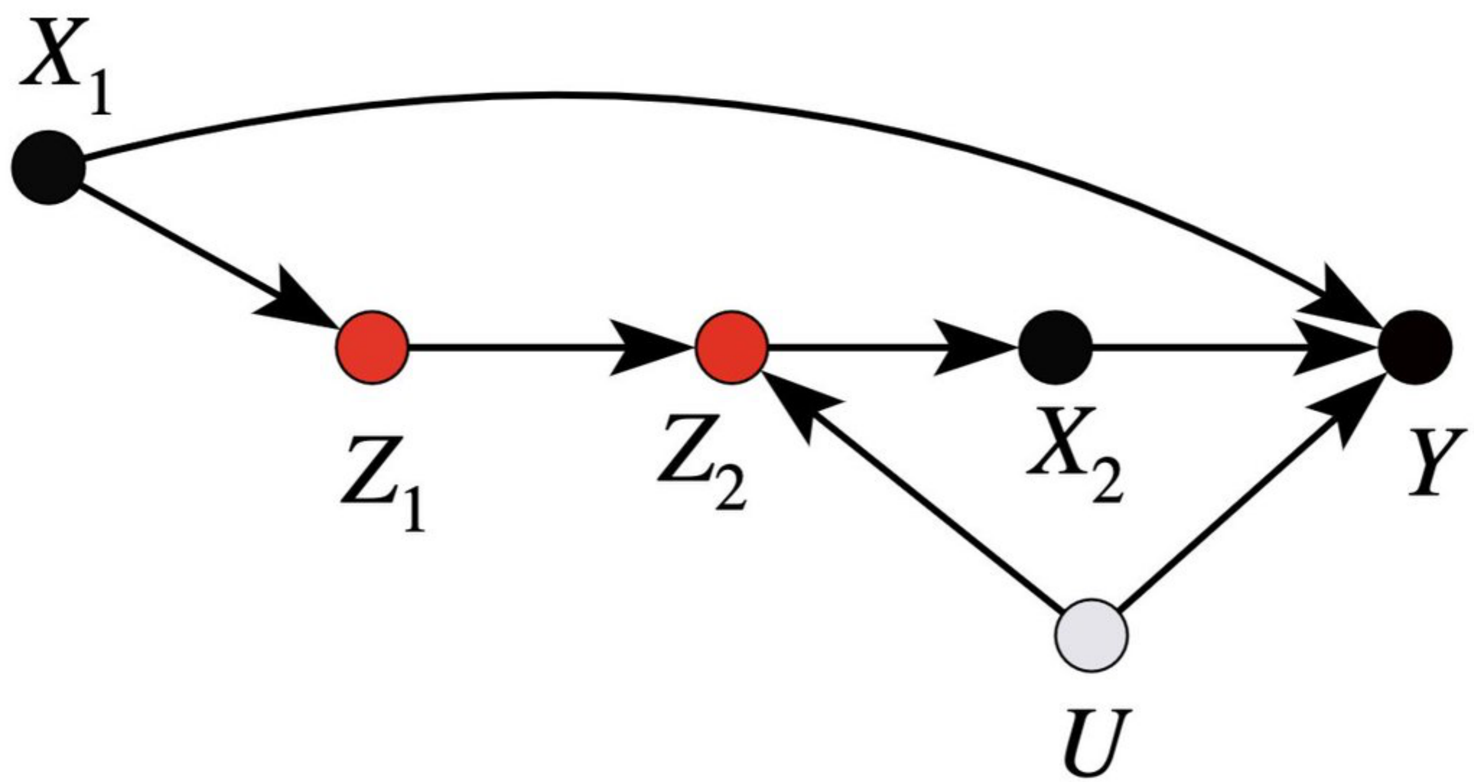
Question: which variables should you include in the regression equation?

Apply the same first principles:

(i) do not close relevant causal paths,

(ii) block spurious paths.

Challenge:



Goal: estimate the *joint effect* of X1 and X2 on Y.

Question: which variables should you include in the regression equation?

Apply the same first principles:

(i) do not close relevant causal paths,

(ii) block spurious paths.

Answer: include both Z1 and Z2. Note again another example where post-treatment variables are *necessary* for identification.

Concluding Remarks

Concluding remarks

Concluding remarks

We have seen through several illustrative examples how simple graphical criteria can be used to decide when a variable should (or should not) be included in a regression equation—and thus whether it can be deemed a “good” or “bad” control.

Concluding remarks

We have seen through several illustrative examples how simple graphical criteria can be used to decide when a variable should (or should not) be included in a regression equation—and thus whether it can be deemed a "good" or "bad" control.

Many of these examples act as cautionary notes against prevailing practices. We have seen, for instance, that:

Concluding remarks

We have seen through several illustrative examples how simple graphical criteria can be used to decide when a variable should (or should not) be included in a regression equation—and thus whether it can be deemed a "good" or "bad" control.

Many of these examples act as cautionary notes against prevailing practices. We have seen, for instance, that:

- Not all pre-treatment variables are "good" controls;

Concluding remarks

We have seen through several illustrative examples how simple graphical criteria can be used to decide when a variable should (or should not) be included in a regression equation—and thus whether it can be deemed a “good” or “bad” control.

Many of these examples act as cautionary notes against prevailing practices. We have seen, for instance, that:

- Not all pre-treatment variables are “good” controls;
- Not all post-treatment variables are “bad” controls; some may even be *necessary* for identification.

Concluding remarks

We have seen through several illustrative examples how simple graphical criteria can be used to decide when a variable should (or should not) be included in a regression equation—and thus whether it can be deemed a “good” or “bad” control.

Many of these examples act as cautionary notes against prevailing practices. We have seen, for instance, that:

- Not all pre-treatment variables are “good” controls;
- Not all post-treatment variables are “bad” controls; some may even be *necessary* for identification.

In all cases, structural knowledge is indispensable for deciding whether a variable is a good or bad control.

Concluding remarks

We have seen through several illustrative examples how simple graphical criteria can be used to decide when a variable should (or should not) be included in a regression equation—and thus whether it can be deemed a "good" or "bad" control.

Many of these examples act as cautionary notes against prevailing practices. We have seen, for instance, that:

- Not all pre-treatment variables are "good" controls;
- Not all post-treatment variables are "bad" controls; some may even be *necessary* for identification.

In all cases, structural knowledge is indispensable for deciding whether a variable is a good or bad control.

Graphical models provide a natural language for articulating such knowledge, as well as efficient tools for examining its logical ramifications.

Thank you!

carloscinelli.com