

# 04 - Graphical Causal Models

Print to PDF

☰ Contents

[Thinking About Causality](#)  
[Crash Course in Graphical Models](#)  
[Confounding Bias](#)  
[Selection Bias](#)  
[Key Ideas](#)  
[References](#)  
[Contribute](#)

## Thinking About Causality

Have you ever noticed how those cooks in YouTube videos are excellent at describing food? “Reduce the sauce until it reaches a velvety consistency”. If you are just starting to learn how to cook, you have no idea what this even means. Just give me the time I should leave this thing on the stove! With causality, it’s the same thing. If you walk into a bar and hear folks discussing causality (probably a bar next to an economics department), you will hear them say how the confounding of income made it challenging to identify the immigration effect on that neighborhood, so they had to use an instrumental variable. And by now, you might not understand what they are talking about. But I’ll fix at least some of this problem right now.

**Graphical models are the language of causality.** They are not only what you use to talk with other brave and true causality aficionados, but also something you use to make your own thoughts clearer.

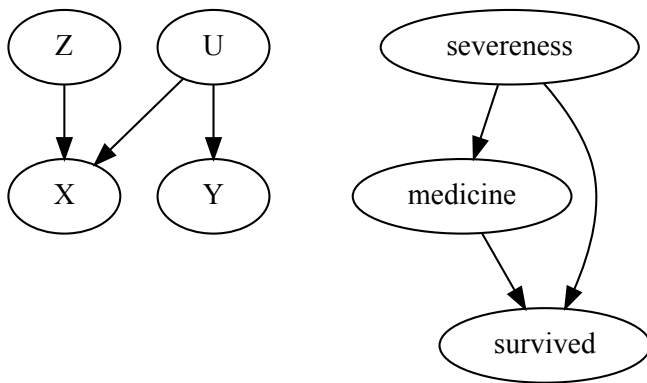
As a starting point, let’s take conditional independence of the potential outcomes, for example. This is one of the main assumptions that we require to be true when doing causal inference:

$$(Y_0, Y_1) \perp T | X$$

**Conditional Independence makes it possible for us to measure an effect on the outcome that is solely due to the treatment, and not any other variable lurking around.** The classic example of this is the effect of a medicine on an ill patient. If only severely ill patients get the drug, it might even look like giving the drug decreases the patients’ health. That is because the effect of the severity is getting mixed up with the effect of the drug. If, however, we break down the patients by severe and not severe cases and analyse the drug impact in each subgroup, we will get a more clear picture of what the true effect is. **This breaking down the population by its features is what we call controlling for or conditioning on X.** By conditioning on the severe cases, the treatment mechanism becomes as good as random. **Patients within the severe group may or may not receive the drug only due to chance, not due a high severity anymore, since all patients are the same on this dimension.** And if treatment is as if randomly assigned within groups, the treatment becomes conditionally independent of the potential outcomes.

Independence and conditional independence are central to causal inference. Yet, it can be quite challenging to wrap our head around them. But this can change if we use the right language to describe this problem. Here is where **causal graphical models** comes in. A causal graphical model is a way to represent how causality works in terms of what causes what.

A graphical model looks like this

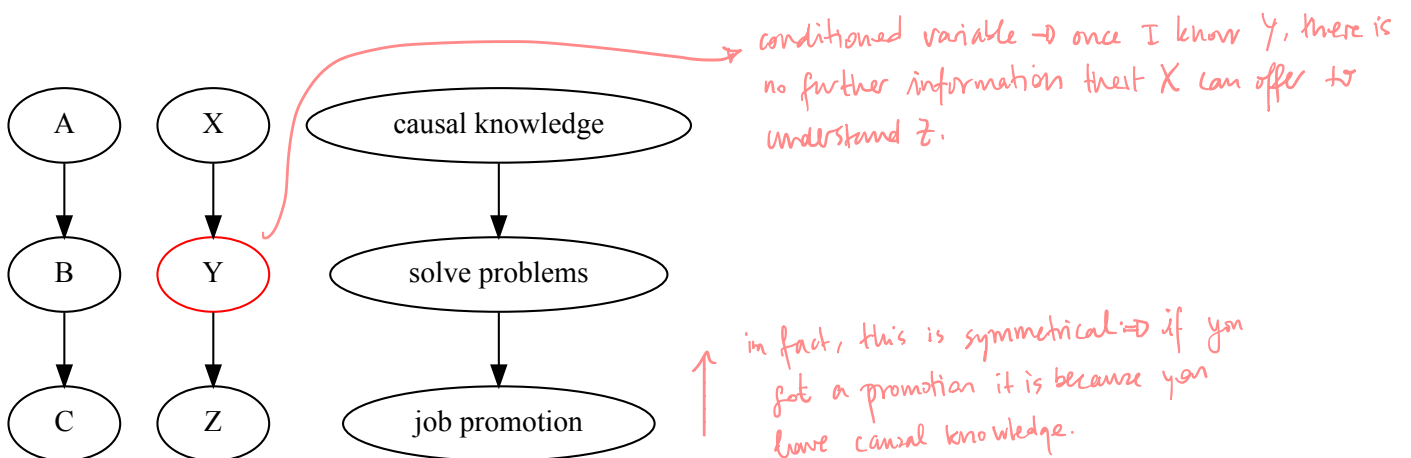


Each node is a random variable. We use arrows, or edges, to show if a variable causes another. In the first graphical model above we are saying that Z causes X and that U causes X and Y. To give a more concrete example, we can translate our thoughts about the impact of the medicine on patient survival as the second graph above. Severeness causes both medicine and survival and medicine also causes survival. As we will see, this causal graphical models language will help us make our thinking about causality more clear, as it makes explicit our beliefs about how the world works.

## Crash Course in Graphical Models

There are [whole semesters on graphical models](#). But, for our purpose, it is just (very) important that we understand what kind of independence and conditional independence assumptions a graphical model entails. As we shall see, independence flows through a graphical model like water flows through a stream. We can stop this flow or we can enable it, depending on how we treat the variables in it. To understand this, let's examine some common graphical structures and examples. They will be quite simple, but they are the sufficient building blocks to understand everything about independence and conditional independence on graphical models.

First, look at this very simple graph. A causes B, B causes C. Or X causes Y which causes Z.



In the first graph, dependence flows in the direction of the arrows. To give a more concrete example, let's say that knowing about causal inference is the only way to solve business problems and solving those problems is the only way to get a job promotion. So causal

knowledge causes problem solving which causes job promotion. We can say here that job promotion is dependent on causal knowledge. The greater the causal knowledge, the greater your chances of getting a promotion. Notice that dependence is symmetric, although it is a little less intuitive. The greater your chances of promotion, the greater the chance you have causal knowledge, otherwise it would be difficult to get a promotion.

Now, let's say I condition on the intermediary variable. In this case, the dependence is blocked. So, X and Z are independent given Y. In the graph above, red indicates that Y is a conditioned variable. By the same token, in our example, if I know that you are good at solving problems, knowing that you know causal inference doesn't give any further information about your chances of getting a job promotion. In mathematical terms,  
 $E[\text{Promotion} | \text{Solve problems, Causal knowledge}] = E[\text{Promotion} | \text{Solve problems}]$   
 . The inverse is also true, once I know how good you are at solving problems, knowing about your job promotion status gives me no further information about how likely you are to know causal inference.

As a general rule, the dependence flow in the direct path from A to C is blocked when we condition on an intermediary variable B. Or,

$$A \not\perp C$$

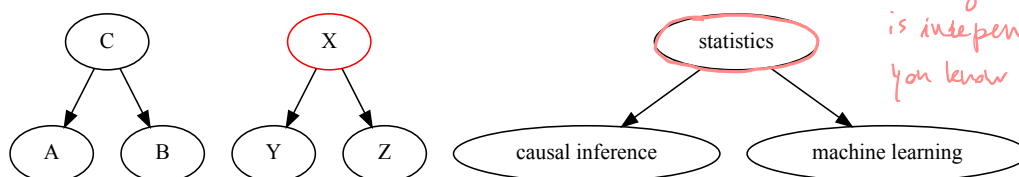
and

$$A \perp C | B$$

*A is independent on C when conditioning on B.*



Now, let's consider a fork structure. In this case, the same variable causes two other variables down the graph. In this case, the dependence flows backward through the arrows and we have what is called a **backdoor path**. We can close the backdoor path and shut down dependence by **conditioning on the common cause**.



*Because we condition on X (stats), knowing how much you know about ML is independent from knowing how much you know about causal inference.*

As an example, let's say your knowledge of statistics causes you to know more of causal inference and machine learning. If I don't know your level of statistical knowledge, then knowing that you are good at causal inference makes it more likely that you are also good at machine learning. That is because even if I don't know your level of statistical knowledge, I can infer it from your causal inference knowledge: if you are good at causal inference you are probably good at statistics, which also makes it more likely that you are good at machine learning.

*↳ 1. If I know you are good at stats → then you are good at ML*

*↳ 2. Because causal inf. and ML are not independent, you might not have knowledge on causal inference, but you might still be very good at ML.*

Now, if I condition on your knowledge about statistics, then how much you know about machine learning becomes independent of how much you know about causal inference. You see, knowing your level of statistics already gives me all the information I need to infer the level of your machine learning skills. Knowing your level of causal inference will give no further information in this case.

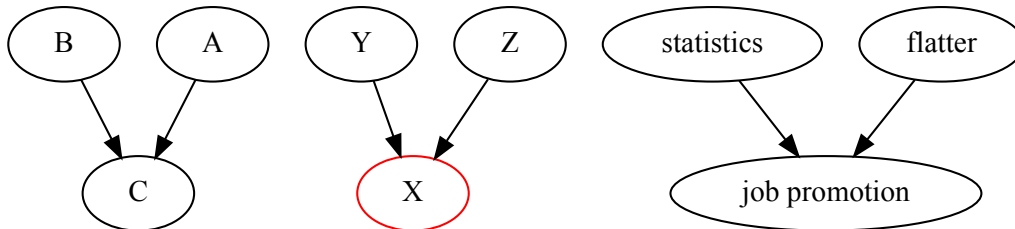
As a general rule, two variables that share a common cause are dependent, but independent when we condition on the common cause. Or

$$A \not\perp B$$

and

$$A \perp B|C$$

The only structure that is missing is the collider. A collider is when two arrows collide on a single variable. We can say that in this case both variables share a common effect.



As an example, consider that there are two ways to get a job promotion. You can either be good at statistics or flatter your boss. If I don't condition on your job promotion, that is, I know nothing if you will or won't get it, then your level of statistics and flattering are independent. In other words, knowing how good you are at statistics tells me nothing about how good you are at flattering your boss. On the other hand, if you did get a job promotion, suddenly, knowing your level of statistics tells me about your level of flattering. If you are bad at statistics and you did get a promotion, it becomes more likely that you know how to flatter, otherwise you wouldn't get a promotion. Conversely, if you are bad at flattering, it must be the case that you are good at statistics. This phenomenon is sometimes called **explaining away**, because one cause already explains the effect, making the other cause less likely.

As a general rule, conditioning on a collider opens the dependence path. Not conditioning on it leaves it closed. Or

$$A \perp B$$

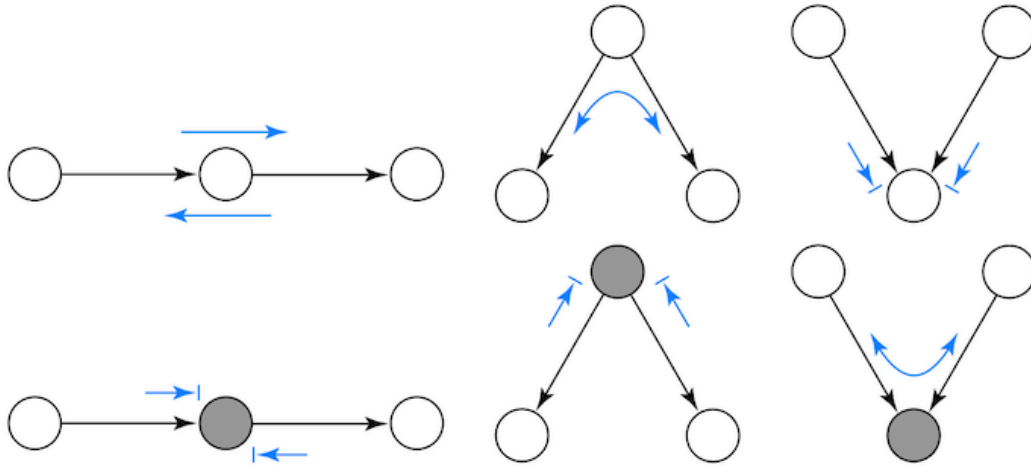
and

$$A \not\perp B|C$$

Knowing the three structures, we can derive an even more general rule. A path is blocked if and only if:

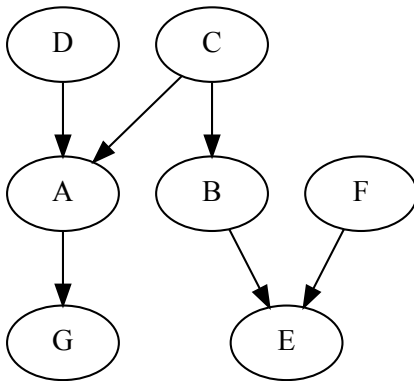
1. It contains a non collider that has been conditioned on
2. It contains a collider that has not been conditioned on and has no descendants that have been conditioned on.

Here is a cheat sheet about how dependence flows in a graph. I've taken from a [Stanford presentation](#) by Mark Paskin. The arrows with lines at their tips signify independence, and the arrows without lines at their tips signify dependence.



As a final example, try to figure out some independence and dependence relationship in the following causal graph.

1. Is  $D \perp C$ ?
2. Is  $D \perp C | A$ ?
3. Is  $D \perp C | G$ ?
4. Is  $A \perp F$ ?
5. Is  $A \perp F | E$ ?
6. Is  $A \perp F | E, C$ ?



**Answers:**

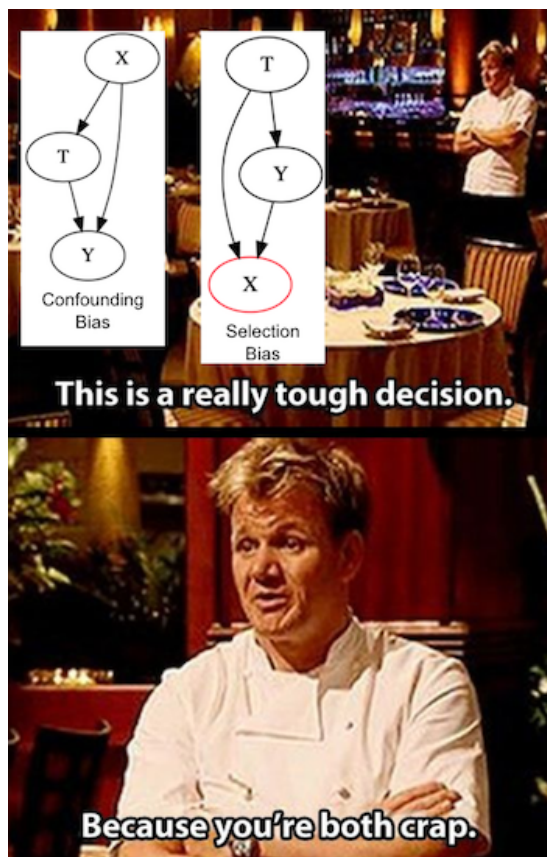
1.  $D \perp C$ . It contains a collider that it has **not** been conditioned on.
2.  $D \perp C | A$ . It contains a collider that it has been conditioned on.
3.  $D \perp C | G$ . It contains the descendent of a collider that has been conditioned on. You can see G as some kind of proxy for A here.
4.  $A \perp F$ . It contains a collider,  $B \rightarrow E \leftarrow F$ , that it has **not** been conditioned on.
5.  $A \perp F | E$ . It contains a collider,  $B \rightarrow E \leftarrow F$ , that it has been conditioned on.
6.  $A \perp F | E, C$ . It contains a collider,  $B \rightarrow E \leftarrow F$ , that it has been conditioned on, but it contains a non collider that has been conditioned on. Conditioning on E opens the path, but conditioning on C closes it again.

Knowing about causal graphical models enables us to understand the problems that arise in causal inference. As we've seen, the problem always boils down to bias.

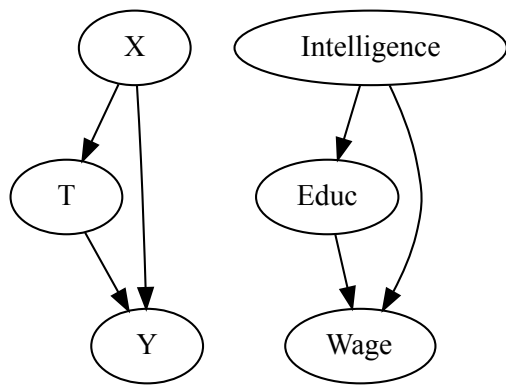
$$E[Y|T = 1] - E[Y|T = 0] = \underbrace{E[Y_1 - Y_0|T = 1]}_{\text{ATE}} + \underbrace{\{E[Y_0|T = 1] - E[Y_0|T = 0]\}}_{\text{BIAS}}$$

Graphical models allow us to diagnose which bias we are dealing with and what are the tools we need to correct for them.

## Confounding Bias

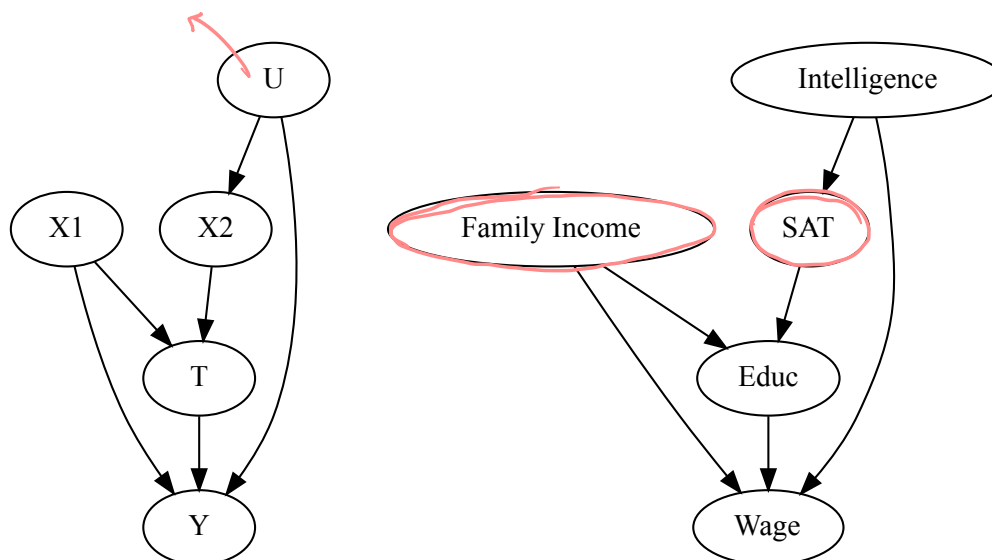


The first big cause of bias is **confounding**. It happens when the treatment and the outcome shares a common cause. For example, let's say that the treatment is education and the outcome is income. It is hard to know the causal effect of education on the wage because both share a common cause: intelligence. So we could make the argument that more educated people earn more money simply because they are more intelligent, not because they have more education. In order to identify the causal effect, we need to close all backdoor paths between the treatment and the outcome. If we do so, the only effect that will be left is the direct effect  $T \rightarrow Y$ . In our example, if we control for intelligence, that is, we compare people with the same level of intelligence but different levels of education, the difference in the outcome will be only due to the difference in education, since intelligence will be the same for everyone. In order to fix confounding bias, we need to control all common causes of the treatment and the outcome.



Unfortunately, it is not always possible to control for all common causes. Sometimes, there are unknown causes or known causes that we can't measure. The case of intelligence is one of the latter. Despite all the effort, scientists haven't yet figured out how to measure intelligence well. I'll use  $U$  to denote unmeasured variables here. Now, assume for a moment that intelligence can't affect your education directly. It just affects how well you do on the SATs, but it is the SATs that determine your level of education, since it opens the possibility of a good college. Even if we can't control for the unmeasurable intelligence, we can control for SAT and close that backdoor path.

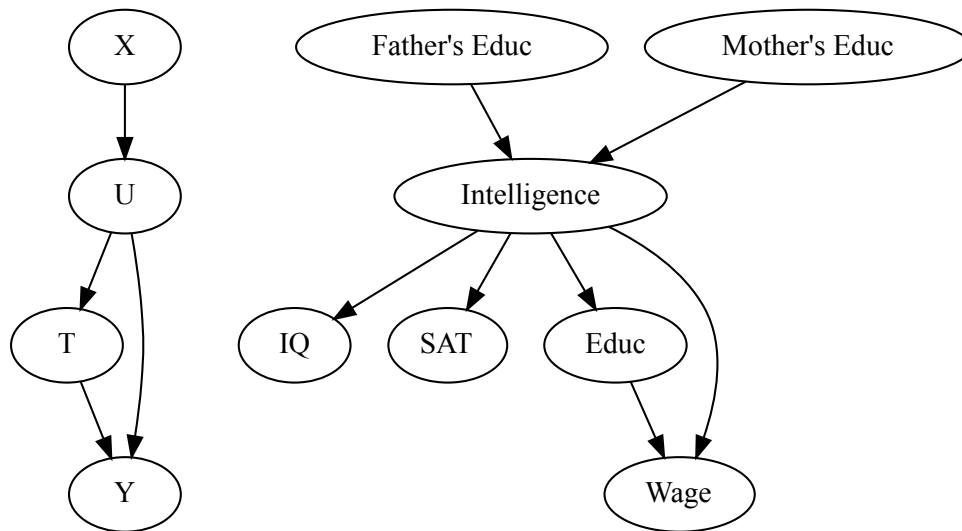
*unmeasured or uncontrolled variables.*



In the following graph, conditioning on  $X_1$  and  $X_2$ , or, SAT and family income, is sufficient to close all backdoor paths between the treatment and the outcome. In other words,  $(Y_0, Y_1) \perp T | X_1, X_2$ . So even if we can't measure all common causes, we can still attain conditional independence if we control for measurable variables that mediate the effect of the unmeasured on the treatment. One quick note, we also have  $(Y_0, Y_1) \perp T | X_1, U$ , but since we can't observe  $U$ , we can't condition on it.

But what if that is not the case? What if the unmeasured variable causes the treatment and the outcome directly? In the following example, intelligence causes education and income directly. So there is confounding on the relationship between the treatment education and the outcome wage. In this case, we can't control the confounder, because it is unmeasurable. However, we have other measured variables that can act as a proxy for the confounder. Those variables are not in the backdoor path, but controlling for them will help lower the bias (but it won't eliminate it). Those variables are sometimes referred to as surrogate confounders.

In our example, we can't measure intelligence, but we can measure some causes of it, like the father's and mother's education, and some effects of it, like IQ or SAT score. Controlling for those surrogate variables is not sufficient to eliminate bias, but it sure helps.



## Selection Bias

You might think that it is a good idea to add everything you can measure to your model just to be sure you don't have confounding bias. Well, think again.

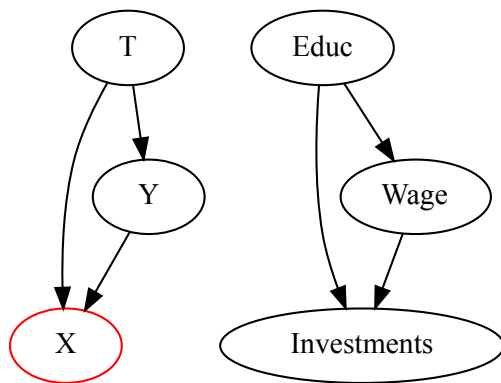


The second big source of bias is what we will call **selection bias**. If confounding bias happens when we don't control for a common cause, selection bias is more related to effects. One word of caution here, economists tend to refer to all sorts of bias as selection bias. Here, I think the distinction between it and confounding bias is very helpful, so I'll stick to it.

More often than not, **selection bias arises when we control for more variables than we should**. It might be the case that the treatment and the potential outcome are marginally independent, but become dependent once we condition on a collider.



Imagine that with the help of some miracle you are finally able to randomize education in order to measure its effect on wage. But just to be sure you won't have confounding, you control for a lot of variables. Among them, you control for investments. But investment is not a common cause of education and wage. Instead, it is a consequence of both. More educated people both earn more and invest more. Also, those who earn more invest more. Since investment is a collider, by conditioning on it, you are opening a second path between the treatment and the outcome, which will make it harder to measure the direct effect. One way to think about this is that by controlling investments, you are looking at small groups of the population where investment is the same and then finding the effect of education on those groups. But by doing so, you are also indirectly and inadvertently not allowing wages to change much. As a result, you won't be able to see how education changes wage, because you are not allowing wages to change as it should.

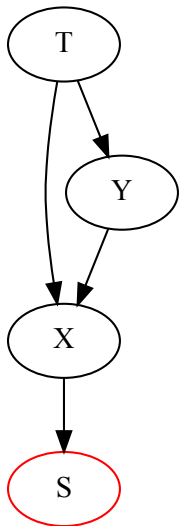


To demonstrate why this is the case, imagine that investments and education take only 2 values. Either people invest or not. They are either educated or not. Initially, when we don't control for investments, the bias term is zero:  $E[Y_0|T = 1] - E[Y_0|T = 0] = 0$  because the education was randomised. This means that the wage people would have in the case they didn't receive education  $Wage_0$  is the same if they do or don't receive the education treatment. But what happens if we condition on investments?

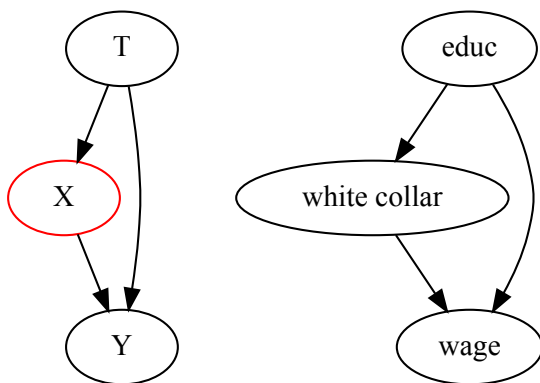
Looking at those that invest, we probably have the case that  $E[Y_0|T = 0, I = 1] > E[Y_0|T = 1, I = 1]$ . In words, among those that invest, those that manage to do so even without education are more independent of education to achieve high earnings. For this reason, the wage those people have,  $Wage_0|T = 0$ , is probably higher than the wage the educated group would have in the case that they didn't have education,  $Wage_0|T = 1$ . A similar reasoning can be applied to those that don't invest, where we also probably have  $E[Y_0|T = 0, I = 0] > E[Y_0|T = 1, I = 0]$ . Those that don't invest even with education, probably would have a lower wage, had they not got the education, than those that didn't invest but also didn't have education.

To use a purely graphical argument, if someone invests, knowing that they have high education explains away the second cause which is wage. Conditioned on investing, higher education is associated with low wages and we have a negative bias  $E[Y_0|T = 0, I = i] > E[Y_0|T = 1, I = i]$ .

Just as a side note, all of this we've discussed is also true if we condition on any descendent of a common effect.



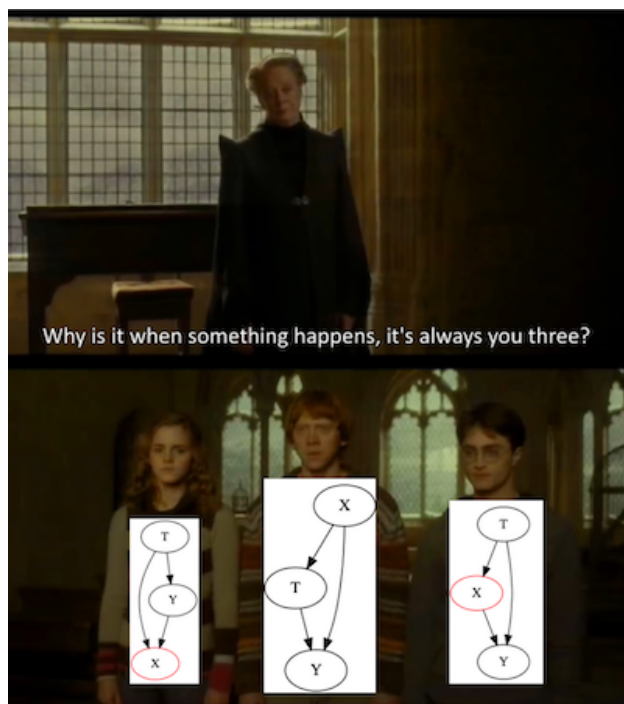
A similar thing happens when we condition on a mediator of the treatment. A mediator is a variable between the treatment and the outcome. It, well, mediates the causal effect. For example, suppose again you are able to randomize education. But, just to be sure, you decide to control whether or not the person had a white collar job. Once again, this conditioning biases the causal effect estimation. This time, not because it opens a front door path with a collider, but because it closes one of the channels through which the treatment operates. In our example, getting a white collar job is one way that more education leads to higher pay. By controlling it, we close this channel and leave open only the direct effect of education on wage.



To give a potential outcome argument, we know that, due to randomisation, the bias is zero  $E[Y_0|T = 0] - E[Y_0|T = 1] = 0$ . However, if we condition on the white collar individuals, we have that  $E[Y_0|T = 0, WC = 1] > E[Y_0|T = 1, WC = 1]$ . That is because those that manage to get a white collar job even without education are probably more hard working than those that required the help of education to get the same job. With the same reasoning,  $E[Y_0|T = 0, WC = 0] > E[Y_0|T = 1, WC = 0]$  because those that didn't get a white collar job even with education are probably less hard working than those that didn't, but also didn't have any education.

In our case, conditioning on the mediator induces a negative bias. It makes the effect of education seem lower than it actually is. This is the case because the causal effect is positive. If the effect were negative, conditioning on a mediator would have a positive bias. In all cases, this sort of conditioning makes the effect look weaker than it is.

To put it in a more prosaic way, suppose that you have to choose between two candidates for a job at your company. Both have equally impressive professional achievements, but one does not have a higher education degree. Which one should you choose? Of course, you should go with the one without the higher education, because he managed to achieve the same things as the other one but had the odds stacked against him.



## Key Ideas

We've studied graphical models as a language to better understand and express causality ideas. We did a quick summary on the rules of conditional independence on a graph. This helped us then explore three structures that can lead to bias.

The first one was confounding, which happens when treatment and outcome have a common cause that we don't account or control for. The second is selection bias due to conditioning on a common effect. This excessive controlling can lead to bias even if the treatment was randomly assigned. The third structure is also a form of selection bias, this time due to excessive controlling of mediator variables. Selection bias can often be fixed by simply doing nothing, which is why it is so dangerous. Since we are biased to action, we tend to see ideas that control for things as clever, when they can be doing more harm than good.

## References

I like to think of this entire book as a tribute to Joshua Angrist, Alberto Abadie and Christopher Walters for their amazing Econometrics class. Most of the ideas here are taken from their classes at the American Economic Association. Watching them is what is keeping me sane during this tough year of 2020.

- [Cross-Section Econometrics](#)
- [Mastering Mostly Harmless Econometrics](#)

I'll also like to reference the amazing books from Angrist. They have shown me that Econometrics, or 'Metrics as they call it, is not only extremely useful but also profoundly fun.

- [Mostly Harmless Econometrics](#)
- [Mastering 'Metrics](#)

My final reference is Miguel Hernan and Jamie Robins' book. It has been my trustworthy companion in the most thorny causal questions I had to answer.

- [Causal Inference Book](#)



## Contribute

Causal Inference for the Brave and True is an open-source material on causal inference, the statistics of science. It uses only free software, based in Python. Its goal is to be accessible monetarily and intellectually. If you found this book valuable and you want to support it, please go to [Patreon](#). If you are not ready to contribute financially, you can also help by fixing typos, suggesting edits or giving feedback on passages you didn't understand. Just go to the book's repository and [open an issue](#). Finally, if you liked this content, please share it with others who might find it useful and give it a [star on GitHub](#).

---

By Matheus Facure Alves

© Copyright 2021.