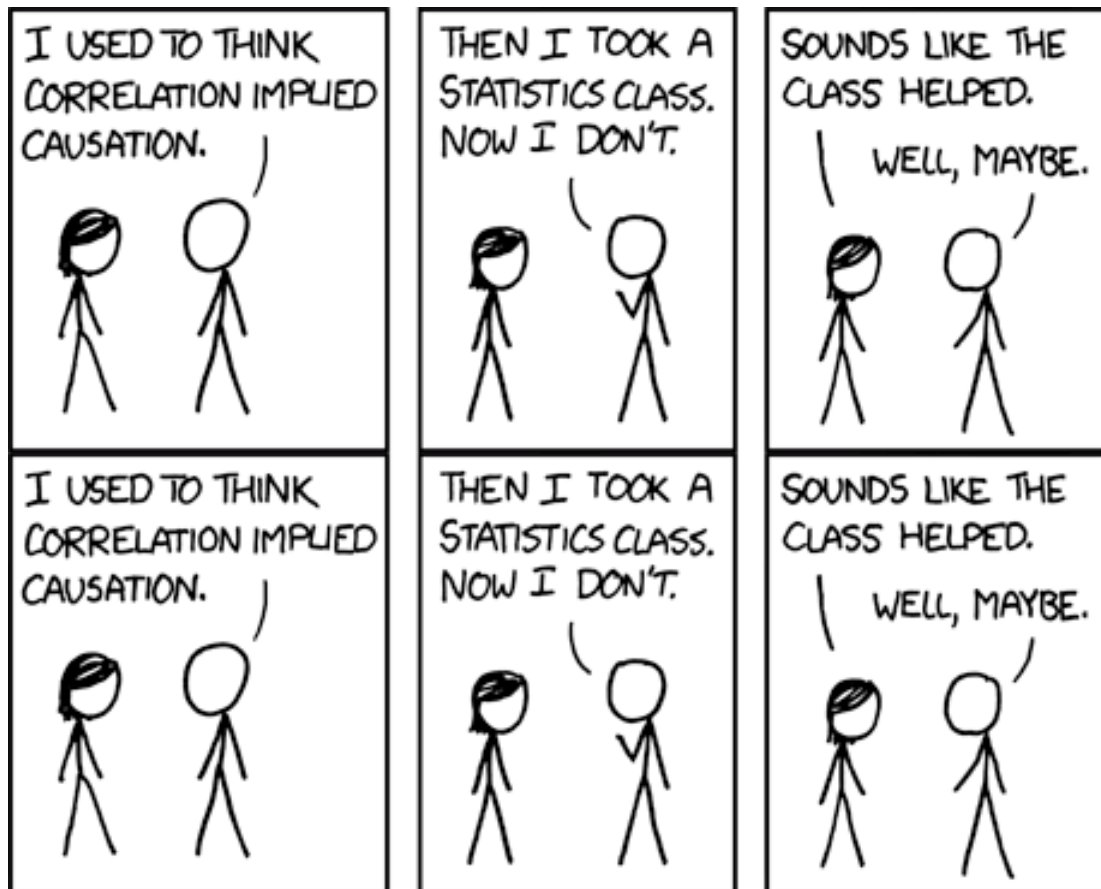


A Primer On Correlation Versus Causation

March 13th, 2009 · 6 Comments · Econ 101 · Just For Fun

The motivating visual, from [xkcd](#):



So what's the deal here? In the economics world, it's hard to swing a dead cat (yes, I swear that is an actual phrase) without hitting the phrase "correlation does not imply causation." What does that mean, really? It means that just because two things happen together doesn't necessarily mean that one of them causes the other. Let's go over some examples:

- I tend to carry an umbrella when it rains, so umbrella and rain are correlated. However, it would be unreasonable to say that my carrying the umbrella caused it to rain. It is probably reasonable instead that the rain caused me to carry the umbrella.
- Students who do better in my class (pre extra credit) do more extra credit. But I cannot conclude that doing the extra credit makes them do better, nor can I conclude that getting better scores makes them do more extra credit. It is in fact likely that an outside factor such as intrinsic motivation is responsible for both the good scores and the extra credit participation.

So, to generalize, if we see that A and B happen together, we don't know whether A caused B, B caused A, or some C caused both A and B. Why is this problematic? One of the main challenges to economists is that of the policy evaluation problem. In other words, economists want to understand the causal impact of changes in policy, environment, etc. Given what we just went over, it should be clear that just saying "Well, policy X was enacted and then change Y occurred" isn't enough. We would instead have to show that nothing else that could affect outcome Y changed. In a perfect world, we would run a controlled experiment using

the [scientific method](#) and have a control and an experimental group, basically like a middle school science project. (Everything I need to know about life I learned in middle school.) This is not usually realistic- it's not like economists can convince policymakers to raise taxes for a randomly selected subset of the population so that they can estimate the causal impact, for example. So what do they do?

It turns out that economists are a pretty resourceful bunch, and they look for what are known as [instrumental variables](#) and [natural experiments](#). (Warning: don't click on those links if you are scared of technical jargon, since the actual terms aren't that important.) Let me illustrate via a very famous example...famous in the economics world at least.

[Josh Angrist](#), a professor of economics at MIT (you want to click the link because the guy totally reminds me of [Ben Stein](#)), wanted to understand the causal impact of enlisting in the military on one's lifetime earnings. Normally, this is a difficult problem, since even if you could control for everything that you could see (age, gender, race, SAT scores, favorite color, whatever), you can't get around the fact that even if you had two otherwise identical people to compare, there was something that made them different, since it is this difference that caused one to enlist in the military and the other choose not to. Maybe the person who enlisted knew that he wouldn't make a very good civilian worker, maybe he knew that he needed discipline, who knows. The point is that there is some difference that the researcher can't see, and the decision of whether to enlist is what is known as a choice variable (an endogenous variable in technical terms, but you don't need to know that). Because of this, the researcher can't just compare lifetime earnings for those who went into the military versus those who didn't (even when controlling for all of those observable characteristics), since it isn't clear whether any difference was due to the military itself or due to the innate differences that influenced the choice of whether to go into the military.

So what is an economist to do? Enter the Vietnam draft lottery as a natural experiment. Draft dodging aside (let's assume there wasn't enough of this going on to materially affect the outcome), the draft lottery provides a handy control and experimental group, since you now do have people that are identical save for the fact that one got lucky and one didn't. Luckily (pun intended), being lucky in the draft lottery in and of itself probably doesn't contribute to lifetime earnings, so the researcher can compare draft winners to draft losers to see the effect on earnings.

What did Angrist find? His paper "Lifetime Earnings and the Vietnam Era Draft Lottery: Evidence From Social Security Administrative Records" shows that, at least for white males, there was a persistent negative impact of military service of about 15% of earnings. ([Click here](#) for a version of the paper)

Now, for the moral of the story...why did I tell you all of this? I told you this so that you could be a skeptical citizen. So, next time someone tries to give you a "A and B happen together, therefore A causes B" type argument, at least you know the right questions to ask. And, like GI Joe says (again, reference intended), knowing is half the battle.