Homework for Chapter 9: Finding Front Doors

1. Which of the following describes when randomization of treatment occurs without a researcher controlling the randomization? b
   1. Exogenous variation
   2. Natural experiment
   3. Instrumental variable
   4. Randomized experiment
2. Necessary conditions:
   1. What needs to be true about the variation we need to isolate when we’re trying to isolate front door paths?

The variation we need to isolate when we’re trying to isolate front door paths are only the variation in treatment that is driven by randomness when any treatment that occurs for other reasons are tossed out.

* 1. How do we know that the variation we get from a randomized controlled experiment fulfills the condition(s) from part a?

The treatment was assigned randomly in a randomized controlled experiment, so for everyone in the experiment, variation in all the variables on all the back doors should be unrelated to whether they got the treatment or not, which closes all the back doors.

1. Describe the four major differences between randomized experiments and natural experiments discussed in the chapter. As a bonus, there’s a fifth difference described in the chapter having to do with sample size and representativeness.
2. Sometimes there will be back doors from the Natural Randomness in natural experiments to the Outcome, which doesn’t happen with pure randomization in randomized experiments.
3. Participants in natural experiments may not realize they’re a part of an experiment which may generate more realistic observations than in randomized experiments.
4. Sample sizes of natural experiments tend to be bigger than randomized experiments and unlike samples of randomized experiments usually made up of a bunch of volunteers, participants in natural experiments are usually not volunteers of the experiment. All this makes samples of natural experiments more representative of the target population than those of randomized experiments.
5. In natural experiments, we are isolating just the variation in treatment that is driven by the NaturalRandomness and tossing out any treatment that occurs for other reasons. So we are seeing the effect only among people who are sensitive to NaturalRandomness. If the effect would be different among another group of people, we won’t learn it for the group we observe with a natural experiment.
6. People believe the exogeneity of pure randomization in randomized experiments. But convincing people that the not-perfectly-random source of exogenous variation in natural experiments is exogenous in the data generating process can be a tall order, given that we’re doing social science where everything is related to everything else.
7. Provide an example of a of research question that is causal in nature but cannot be feasibly answered by a randomized experiment. Explain your reasoning.

The research question that “Does damage to a person’s hippocampus impair the formation of long-term memory traces?” can be about a causal relationship, but it is not possible, feasible, or ethical to manipulate the independent variable or randomly assign participants to conditions (cause damages to a random group of people’ hippocampi and do nothing to another random group of people’ hippocampi) as in a randomized experiment.

1. Define the concept of exogenous variation.

Exogenous variation is a source of variation in treatment that has no open back doors (the assignment of treatment versus control is known to be external to the processes that generate the outcomes that under study). Any path we can walk from the source of exogenous variation to the outcome must be closed or contain the treatment. An ideal source of exogenous variation is not caused by any other variable that belongs on the causal diagram. This means that we can use plenty of things as sources of exogenous variation even if they’re not purely random, so long as they’re as good as random in the context of the data generating process.

1. Provide an example of a natural experiment. It could be an idea of your own, from the book, or from a study you read.

In preparation for the Beijing Olympics in 2008, the Chinese government introduced several measures to improve the city’s poor air quality, including restrictions on vehicle use, the temporary closure of factories and construction projects, and seeding clouds to induce rain. These changes effectively reduced air pollution during the Olympic period, after which things returned to normal, therefore creating a sharply defined window of relatively good air quality. Researchers used this sudden reduction in pollution to compare its effects on the weight of 83,672 babies born in four districts of Beijing during the Olympics (August 8 – September 24) against the weight of babies born during the same period in 2007 and 2009. They found that the babies born in 2008 were 23 grams (0.05 pounds) heavier on average, suggesting that air pollution does interfere with fetal development.

* 1. Draw a causal diagram using 5-10 variables.

Use of Air Filtration

Vehicle Use

AntiAirPolutionPolicy

Factories & Construction Projects

Birth Weight

Air Quality

Artifical Rainfall

Maternal Occupation

Pregnancy Care

* 1. List the paths from the source of exogenous variation to the outcome. AntiAirPolutionPolicy-> Vehicle Use-> Air Quality-> Birth Weight

AntiAirPolutionPolicy-> Factories & Construction Projects-> Air Quality-> Birth Weight

AntiAirPolutionPolicy-> Artificial Rainfall -> Air Quality-> Birth Weight

* 1. Which paths need to be closed?

Air Quality<-Use of Air Filtration<-Maternal Occupation -> Pregnancy Care->Birth Weight

* 1. Is it believable in this context that you haven’t missed any paths and that you can close all the paths you need? Why or why not?

No. Not only Maternal Occupation will affect the pregnancy care a parent can afford to give and the measures to filtering air a parent can adopts under working conditions, Maternal education level may both affect a parent’s knowledge of pregnancy care and awareness of air filtration indoors and outdoors which also can affect a parent’s occupation.

1. Which of the following best explains how “using exogenous variation to predict treatment, and then using the prediction instead of the treatment itself” lets you isolate just the front door from treatment to outcome? c
   1. Because it allows you to control for treatment assignment
   2. Because the exogenous variation has no back doors, so nothing it predicts can have back doors either
   3. Because it takes the relationship between treatment and outcome and, by controlling for the exogenous variation, removes the back doors
   4. It can’t unless the treatment itself is also exogenously assigned
2. Consider the following natural experiment design: in 2019, United States president Donald Trump increased tariffs on Chinese imports considerably. This made United States consumers want to switch to non-Chinese sources of imports. If you are an exporter to the US, like Brazil, then demand for your products just went up *if* China is a competitor of yours. So Brazilian manufacturers of machinery saw their exports increase by a lot, but Brazilian pop musicians didn’t. You want to know whether exporting to a country makes you like it more. You plan to isolate the variation in Brazilian exports driven by the tariffs to see if industries with more exports have a higher opinion of the United States.
   1. What potential back doors might remain between the tariffs and the opinion of the United States?

The Trump’s government trade policy of protectionism might affect both U.S. tariffs on Chinese imports and U.S. tariffs on Brazilian imports, which impacts Brazilian exports to the US and opinion of the United States.

The global economic environment (e.g., prosperous or not) can affect U.S. tariffs on Chinese imports, Brazilian exports, and the opinion of the United States since the US is the dominant player and influencer of the global economy.

* 1. If someone told you that they performed this study as described, and found that increased exports had no effect on opinion of the US, would you believe the result? Why or why not?

I would not believe the result based on the study they performed as described. By isolating just the variation in the export driven by the tariffs, they were tossing out any variation in the export that occurs for other reasons. So they could only see the effect among industries who were sensitive to the tariffs, those who were competitors of certain Chinese exporting industries. The effect would be different among another industry (which was not a competitor of certain Chinese exporting industries faced with increased U.S. tariffs). Since the treatment effect varies from industry to industry, they might not see it for the industries they sampled.