**Samuel Snelson**

**SOC 723 Spring 2023**

**17 Feb 2023**

Homework for Chapter 8: Causal Paths

1. Assuming that a path has no colliders on it, what is the difference between a path being Open and Closed?

McElreath and NHK’s articulation of open v. closed is slightly different but I think the same really. McElreath provided the simple heuristic that a path is closed if there is a collider. NHK qualified this with the fact that closed really means (and I think McElreath means) that causal influence via variability is not transmitted to the end of the path. This can take place either by a non-collider with no variability or when a collider is on the path (because causal influence cannot travel back through an arrow in reverse).

1. Consider the below generic causal diagram.  
   Diagram, shape

   Description automatically generated
   1. List every path from X to Y.
      1. X 🡪 C 🡨 D 🡪 Y (backdoor, closed)
      2. X 🡨 B 🡨 D 🡪 Y (backdoor, tentatively open)
      3. X 🡪 A 🡪 Y (frontdoor, tentatively open)
      4. X 🡪 C 🡨 D 🡪 B 🡪 Y (backdoor, closed)
   2. Which of the paths are front-door paths?
      1. X 🡪 A 🡪 Y
   3. Which of the paths are open back-door paths?
      1. X 🡨 B 🡨 D 🡪 Y
   4. What variables must be controlled for in order to identify the effect of X on Y? (only list what *must* be controlled for, not anything that additionally *could* be controlled for).
      1. Only B needs to be adjusted for to identify the effect of X 🡪 Y.
         1. The path through C is closed.
         2. Adjusting for B closes the path through D to Y and leaves the effect of D on Y independent of X.
         3. A is a mediating path so adjusting may induce post-treatment bias.
2. Consider the research question: Does having higher income cause better health?
   1. Draw a causal diagram depicting the data generating process for this relationship with 5-10 variables on it.

Diagram

Description automatically generated

* 1. Identify the Front Door paths.
     1. Income 🡪 Health
     2. Income 🡪 Stress 🡪 Health
     3. Income 🡪 Medical Treatment 🡪 Health
     4. Income 🡪 Healthy Food 🡪 Health
  2. Identify the Back Door paths.
     1. Income 🡨 ParentSES 🡪 Health
     2. Income 🡨 ParentSES 🡪 Education 🡪 Health
     3. Income 🡨 Education 🡪 Health
  3. Identify the paths that represent direct effects.
     1. Income 🡪 Health
     2. Stress 🡪 Health
     3. MedicalTreatment 🡪 Health
     4. HealthyFood 🡪 Health
     5. ParentSES 🡪 Health
     6. Education 🡪 Health
  4. Identify the Good Paths and the Bad Paths.
     1. Thinking about bad paths as those which are unrelated to our research question, I don’t think that there are any bad paths with this particular DAG structure. The fork structure at the top concerns the extent to which the relationship between Income and Health is attributable to alternative explanations which cause income and health. The lower mediation structure concerns alternative explanations for which income causes health conditional on the three variables.

1. Which of the following describes a causal path where all the arrows point away from the treatment?
   1. Open Path (not this one because of open paths with fork confounders)
   2. Closed Path (not this one because it’s not a causal path)
   3. Front Door Path (this one is the answer)
   4. Back Door Path (not this one because arrows, by definition, point to the treatment)
2. Consider the figure below, which depicts the relationship between teaching quality, number of publications (e.g., articles, books), and popularity among scholars and students in a population of professors.  
   Line chart

   Description automatically generated
   1. What type of variable is Popularity in one path on this diagram?
      1. Collider
   2. Discuss what would happen if you controlled for Popularity.
      1. This may induce an erroneous non-causal influence between number of publications and teaching quality.
3. Consider the figure below, which depicts the relationship between a pandemic-related lockdown and an economic recession. The research question of interest is: Does a pandemic-related lockdown cause recession?   
   A picture containing text, antenna, line

   Description automatically generated
   1. Write down all the paths in the diagram from Lockdown to Recession. To make our lives simpler (there are a lot of paths in this diagram), ignore any path that goes through Stimulus.
      1. Sorry doing stimulus anyway.
      2. Lockdown 🡪 Recession
      3. Lockdown 🡪 Unemployment 🡪 Recession
      4. Lockdown 🡨 PriorEconomy 🡪 Recession
      5. Lockdown 🡨 PriorEconomy 🡪 Unemployment 🡪 Recession
      6. Lockdown 🡪 Unemployment 🡨 PriorEconomy 🡪 Recession
      7. Lockdown 🡨 PriorEconomy 🡪 Stimulus 🡪 Recession
      8. Lockdown 🡨 Prior 🡪 Stimulus 🡪 Unemployment 🡪 Recession
      9. Lockdown 🡪 Unemployment 🡨 Prior 🡪 Stimulus 🡪 Recession
      10. Lockdown 🡨 PriorEconomy 🡪 Unemployment 🡨 Stimulus 🡪 Recession
   2. List all of the paths that are Front Door Paths.
      1. Lockdown 🡪 Recession
      2. Lockdown 🡪 Unemployment 🡪 Recession
   3. What would happen if we controlled for unemployment?
      1. Two things:
         1. Unemployment is a mediator on the front door path and this would induce post-treatment bias if we were estimating the total effect. Direct effect depends on whether it is a collider (it is) and the other paths which tell us whether or not the paths can be closed.
         2. Unemployment is also a collider. Adjusting for unemployment may induce collider bias (unless the paths are closed by adjusting for PriorEconomy which we want to adjust for anyways).
   4. Is it possible to measure each of the variables adequately?
      1. Total effect: We can get this by adjusting for PriorEconomy (this gives us the two front door paths [made of the direct path and indirect path through Unemployment]) .
      2. Direct effect: All of the parameters can be recovered if PriorEconomy and Unemployment is adjusted for (Stimulus also can be estimated by adjusting for it). This is a situation where there is a collider and adjusting for it opens the path, but it can be closed by adjusting for other variables on the path. This isn’t the best situation and out estimates would likely be less efficient, but they can be recovered.
   5. Can you think of any variables and paths not depicted in the diagram that may be relevant to identify the answer to the research question? List at least one and no more than three.
      1. MarketResponse
      2. LaborMarketContraction
      3. OnlineWork
4. Consider the question: Does obtaining a higher education improve income? Think of a couple of examples of Bad Paths in a causal diagram depicting the data generating process for this research question.
   1. HigherEd 🡪 CollegeNetwork 🡪 Income
      1. This would be a bad path if were interested in estimating the total effect of getting an education on income (meaning don’t adjust for it)
   2. HigherEd 🡪 PolViews 🡨 Income
      1. This is a bad path because conditioning on political views of respondents may induce collider bias. Depending on the relationship, we might find that HigherEd decreases liberal views and Income increases liberal views (I’ll run some simulations later to see what actually happens).