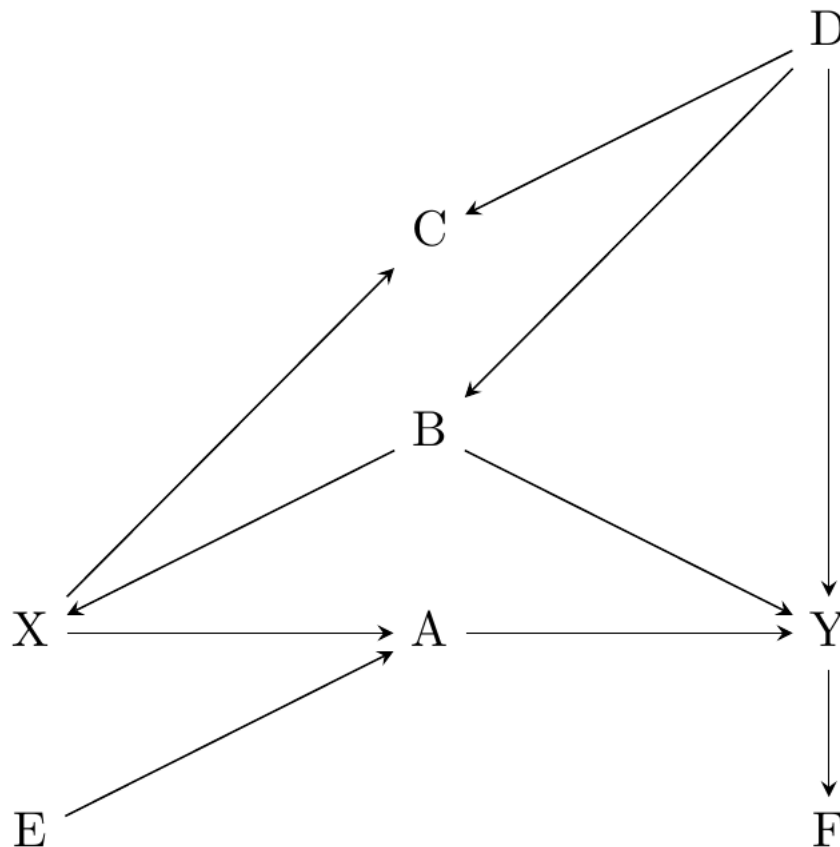


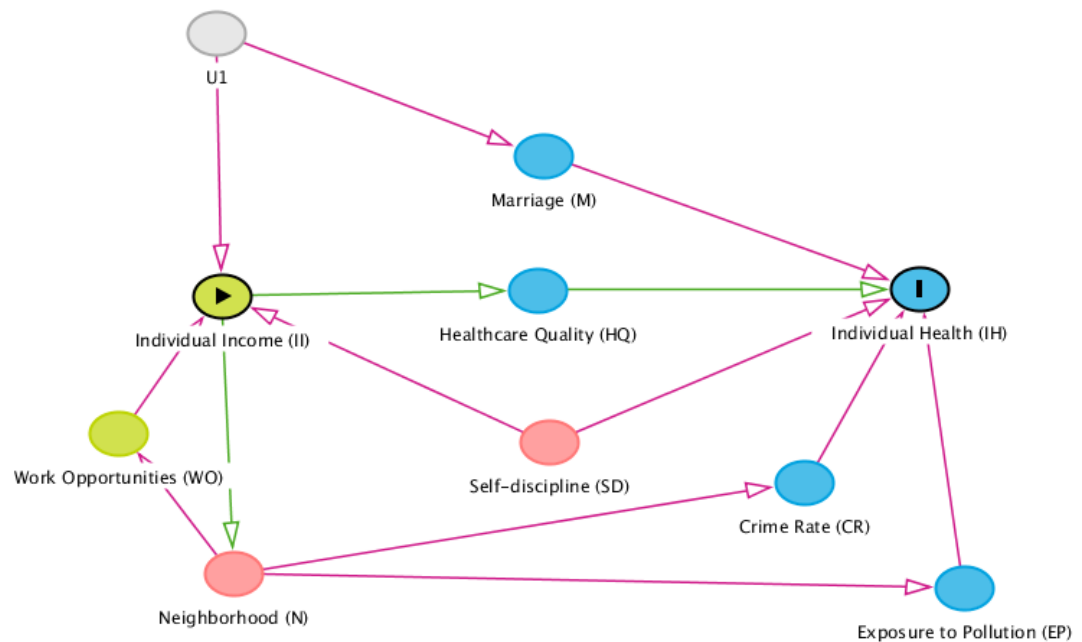
## Homework for Chapter 8: Causal Paths

1. Assuming that a path has no colliders on it, what is the difference between a path being Open and Closed?
  - a. An open path is one where there is variation in all the variables (i.e. we do not condition  $\Pr(Y | X)$  on any of them).
  - b. A closed path is one in which there is either a collider ( and we don't include it in the model) or where one of the variables is conditioned on.
2. Consider the below generic causal diagram.



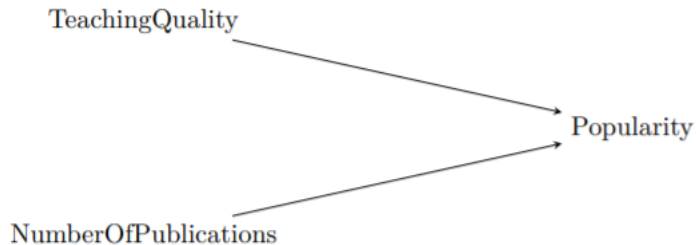
- a. List every path from X to Y.
    - i.  $X \rightarrow A \rightarrow Y$  Open, front door
    - ii.  $X \leftarrow B \rightarrow Y$  Open, back door
    - iii.  $X \rightarrow C \leftarrow D \rightarrow B \rightarrow Y$  Closed, back door
    - iv.  $X \rightarrow C \leftarrow D \rightarrow Y$  Closed, back door
    - v.  $X \leftarrow B \leftarrow D \rightarrow Y$  Open, back door
  - b. Which of the paths are front-door paths? The first one.
  - c. Which of the paths are open back-door paths? The second one and the fifth.
  - d. 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).  $\Pr(Y | X, B)$  (i.e. conditioning on B) give us the causal effect of X on Y.
3. Consider the research question: Does having higher income cause better health?

- a. Draw a causal diagram depicting the data generating process for this relationship with 5-10 variables on it.

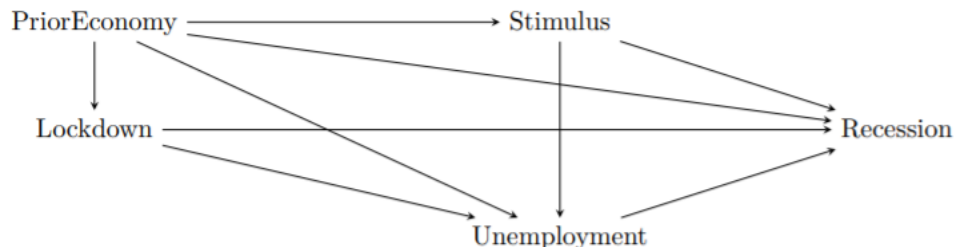


- b. Identify the Front Door paths.
- $II \rightarrow HQ \rightarrow IH$
  - $II \rightarrow N \rightarrow CR \rightarrow IH$
  - $II \rightarrow N \rightarrow EP \rightarrow IH$
- c. Identify the Back Door paths.
- $II \leftarrow U1 \rightarrow M \rightarrow IH$
  - $II \leftarrow SD \rightarrow IA$
  - $II \leftarrow WO \leftarrow N \rightarrow CR \rightarrow IH$
  - $II \leftarrow WO \leftarrow N \rightarrow EP \rightarrow IH$
- d. Identify the paths that represent direct effects.
- There is no direct effect in this DAG.
- e. Identify the Good Paths and the Bad Paths.
- All the back door paths are bad paths in this case, and all the front door paths are good paths.
  - One interesting thing in this DAG is that there is a cycle (between income and neighborhood) that is also part of two of the front doors. So the total effect of income on health cannot be estimated without breaking the cycle.
4. Which of the following describes a causal path where all the arrows point away from the treatment?
- Open Path
  - Closed Path

- c. Front Door Path
- d. Back Door Path
5. 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.



- a. What type of variable is Popularity in one path on this diagram?
  - i. Popularity is a collider.
- b. Discuss what would happen if you controlled for Popularity.
  - i. This is what is called Berkson's paradox. If we control for popularity then teaching quality and number of publications become negatively correlated (assuming they both have a positive effect on popularity). So among all professor there is not relationship between teaching quality and publications, but among popular professors, those that are good at teaching publish little, and those who publish a lot are pretty bad professor.
6. 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. 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.
  - i. Lock  $\rightarrow$  Rec
  - ii. Lock  $\rightarrow$  Un  $\rightarrow$  Rec
  - iii. Lock  $\leftarrow$  PriorE  $\rightarrow$  Un  $\rightarrow$  R
  - iv. Lock  $\leftarrow$  PriorR  $\rightarrow$  R
  - v. Lock  $\rightarrow$  Un  $\leftarrow$  PriorE  $\rightarrow$  Rec
- b. List all of the paths that are Front Door Paths. PAtbs 1 and 2 are front door.
- c. What would happen if we controlled for unemployment? We would close path 2 (where unemployment is a mediator), path 3 (an open backdoor path), and would open path 5 (a backdoor path where unemployment is a collider).

- d. Is it possible to measure each of the variables adequately?
    - i. I don't really understand what "adequately" means here. All measures are imperfect, but they can be more or less valid or reliable. Whether they are adequate or not depends on our goal.
  - e. 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.
    - i. Lockdown ← Type of government (e.g. democratic or authoritarian) → Recession (e.g. authoritarian regimes might be both more inclined to lockdowns and to heavily intervene in the economy to avoid a recession).
7. 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.

Education ← Cognitive Hability → Income