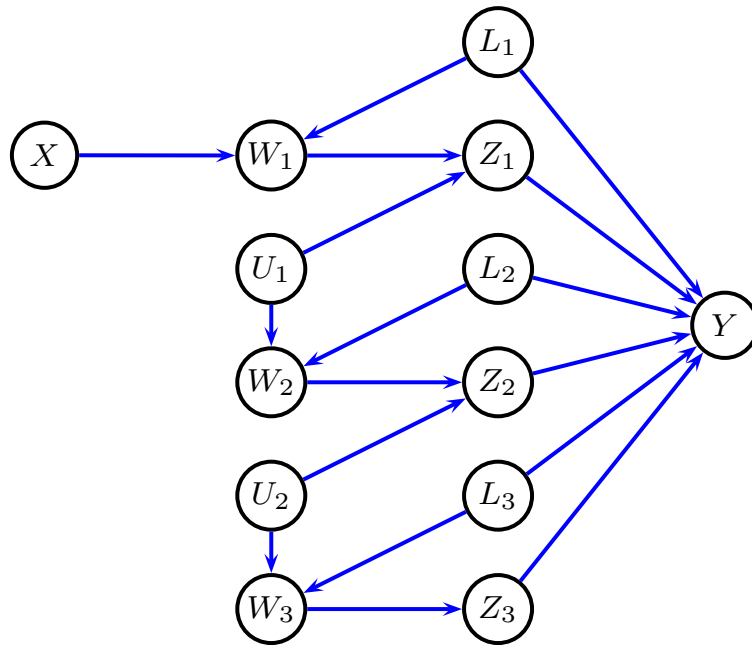


COMH7266: Longitudinal Analysis and Causal Inference

Causal Inference Assignment: Due 24 July 2017

1. Locate website with causal inference materials for this course  
<http://www.bios.unc.edu/~mhudgens/wits/wits.html>
2. Read Chapters 1 - 15 of the Hernan and Robins (HR) book *Causal Inference*. The book is in pdf format and available for free here  
<http://www.hsph.harvard.edu/miguel-hernan/causal-inference-book/>
3. (Chapter 1) As in HR, let the random variable  $A$  denote treatment and the random variable  $Y$  a binary outcome where  $Y = 1$  denotes death and  $Y = 0$  denotes survival. Using the concept of potential outcomes, formally define four strata: immune, helped, harmed, and doomed. Consider the four probabilities of being in each of these strata. In terms of these probabilities, give a necessary and sufficient condition for the average causal effect to be zero. Interpret.
4. (Chapter 4) Prove there is qualitative, additive effect modification if and only if there is qualitative, multiplicative effect modification.
5. (Chapter 6) Consider the following DAG



Which vertices are d-separated from  $Y$  if we do not condition on any variables? Provide a justification for your answer.

- (Chapter 12) Using a statistical software package such as Stata, SAS, or R, read in the NHEFS data set posted on the course website. Using data from 1566 cigarette smokers aged 25-74 years who had a baseline visit and a follow-up visit 10 years later, estimate the average weight gain for the quitters and the average weight gain for the non-quitters. Compute a 95% confidence for the difference in average weight gain between the two groups. The code posted on the course website may be helpful for completing this exercise. Include in your solution the key parts (but not all) of your software code and the corresponding output.