Info 6751. Fall 2022. Problem Set 13. Due on Canvas by 5pm on 29 Nov (Tuesday deadline to avoid the day immediately after break).
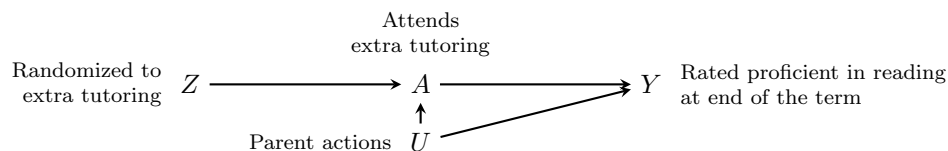
1 (30 points) Instrumental variables in experiments

Suppose you are an elementary school principal. You randomize some students to a new program to receive extra tutoring at an off-site location in the evenings. You randomize other students to a no-tutoring condition.

In many cases, students' treatment assignments Z determines their actual treatments A (when $Z = 1$ then $A = 1$, and when $Z = 0$ then $A = 0$). But there are some difficulties:

a) The parents of some students work evenings and can't drive their children to the tutoring (U). No matter the value of Z, these children do not receive tutoring ($A = 0$).

b) The parents of some students make a huge fuss (U) so that regardless of the value of Z, these parents ensure that their children receive tutoring ($A = 1$).

Using the terms discussed in class,

1.1. (5 points) What is the intent to treat effect (in words)?

1.2. (5 points) Who are the always-takers?

1.3. (5 points) Who are the never-takers?

1.4. (5 points) Who are the compliers?

1.5. (5 points) Although they are not discussed above, describe how someone could be a defier.

1.6. (5 points) What assumption was made credible by randomization of Z?

2 (20 points) Instrumental variables in observational studies

Much of the water supply for the state of California comes from snowmelt in the Sierra Nevada Mountains. Two economists are very excited to notice that some years have much larger snowpacks than others—this could be an instrument!
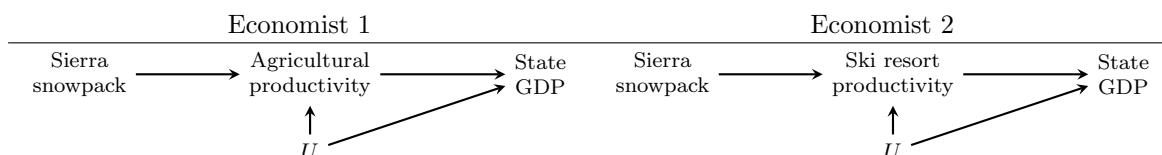
The first economist argues that random differences in the Sierra snowpack create random fluctuations in agricultural productivity, thereby providing an instrumental variable for the effect of agricultural productivity on the state's GDP.

The second economist argues that random difference in the Sierra snowpack create random fluctuations in the quality of skiing at Mammoth Mountain and other Sierra resorts, thereby providing an instrumental variable for the effect of ski resort productivity on the state's GDP.

2.1. (20 points) Both economists argue that their instruments are valid because the snowpack is randomly assigned. Can both economists be right that their instrument is valid? Why or why not?