

Reading Assignment 8

Andrew Shaw

2025-10-18

1. What is a backdoor path?

A path which can be open or closed is a variable in a model that introduces additional pathways from the treatment variable to the outcome variable potentially introducing complications if left open are akin to omitted variable bias if not controlled for.

2. What is the difference between an open and closed backdoor path?

Open backdoors are confounder variables in which the direction of influence flows from a change in the X confounder affects both the treatment group D and outcome Y . A closed backdoor is a collider variable where the changes in D and Y affect the X collider but that change does not affect the $D \rightarrow Y$ relationship.

3. What is a collider variable?

A variable in which the D and Y causal effects collide in the collider variable which has no bearing on the causal relationship between D and Y .

4. Why shouldn't you control for colliders?

Through controlling for a collider variable and not allowing to freely change with changes in D and Y the model is constrained forcing D and Y to counterbalance while the collider is fixed creating a relationship between X and the treatment and outcome variables that didn't exist before conditioning.

5. Write down a DAG for the following variables on infant mortality using your own intuition:

fertility drugs, number of fetuses in pregnancy (plurality), birth month, low birth weight, use of inducement drugs.

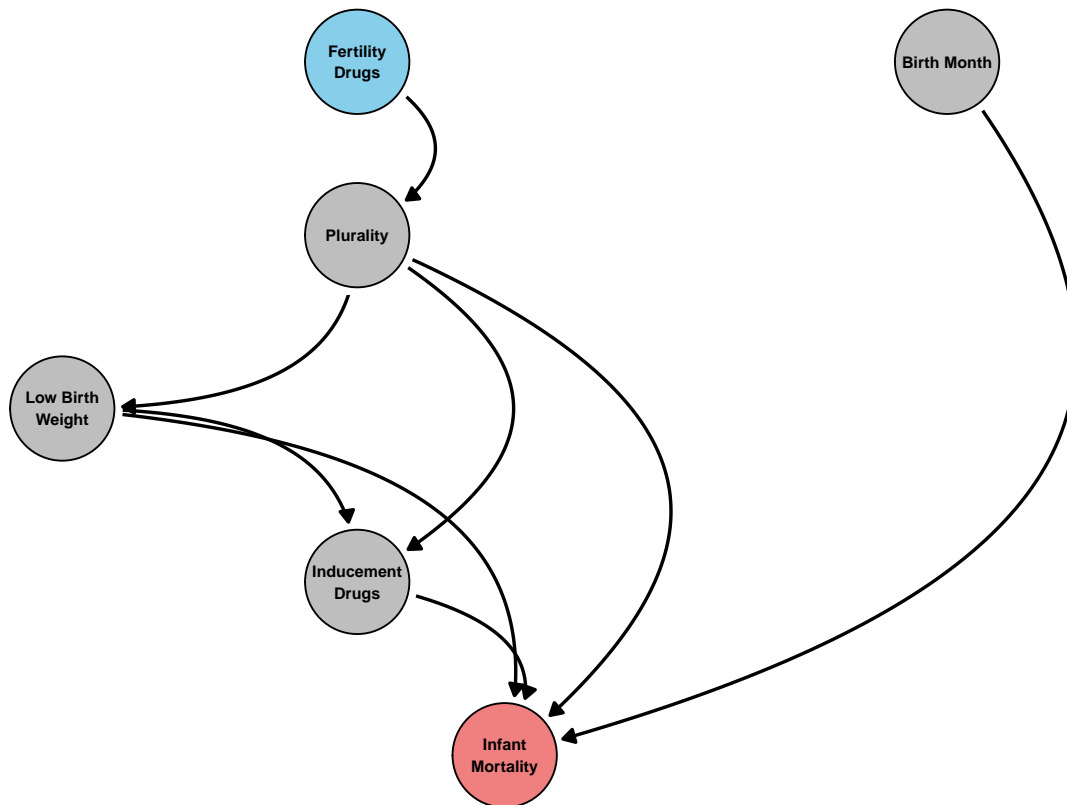
```
mortality <- dagify(y ~ m,
                    y ~ l,
                    l ~ p,
                    p ~ fd,
                    i ~ p,
                    i ~ l,
                    y ~ p,
                    y ~ i,
                    labels = c(fd = "Fertility\nDrugs",
```

```

p = "Plurality",
l = "Low Birth\nWeight",
y = "Infant\nMortality",
m = "Birth Month",
i = "Inducement\nDrugs"))

ggdag(mortality, text = FALSE, edge_type = "arc", layout = "sugiyama") +
  geom_dag_point(aes(fill = ifelse(name == "y", "outcome",
                                   ifelse(name == "fd", "treatment", "other"))),
                shape = 21, size = 18) +
  geom_dag_text(aes(label = label), size = 2, color = "black") +
  scale_fill_manual(values = c(treatment = "skyblue", outcome = "lightcoral", other = "grey"), guide =
  theme_dag()

```



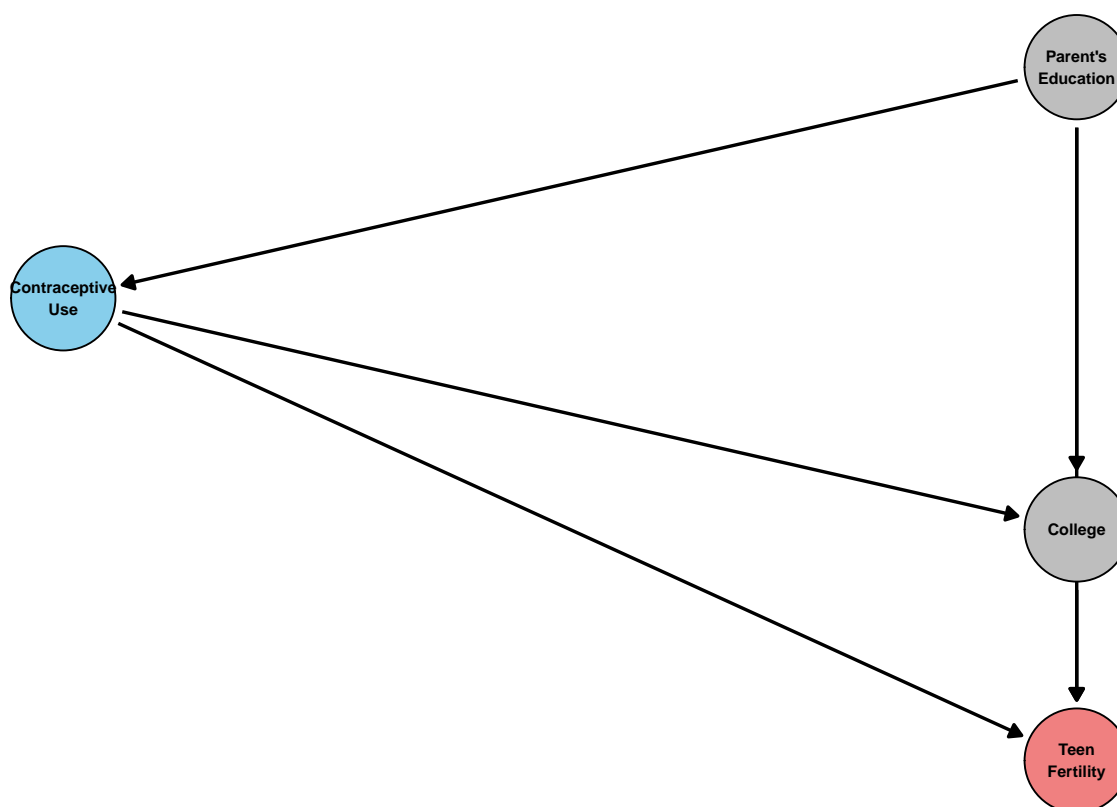
6. Which variables should you control for if you want to know the effect of fertility on infant mortality?

If testing for effect of fertility drugs on infant mortality, there doesn't seem to be a variable confounding the treatment group of fertility drugs.

7. What about inducement drugs?

To see the relationship between inducement drugs on fertility, *Plurarlity* and *LowBirthWeight* need to be controlled for so their fluctuations do not influence the variable's effect on mortality.

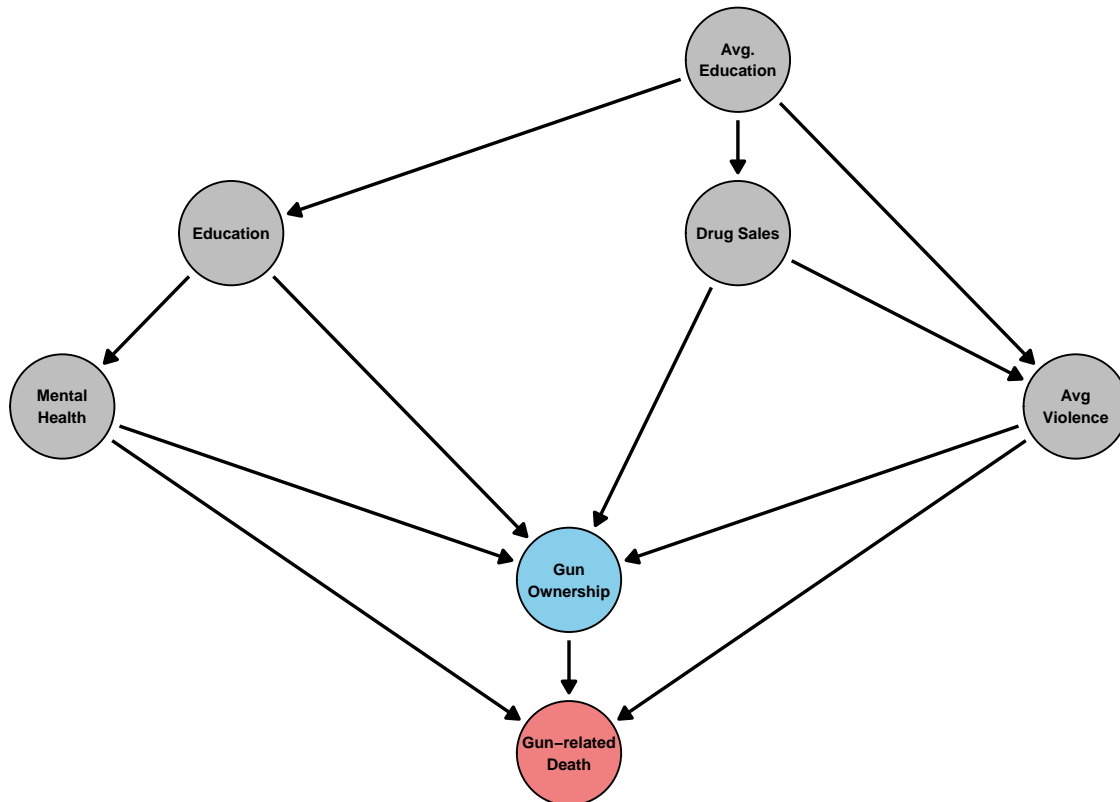
8. Write down a DAG for the following variables on teen fertility, parent's education, contraceptive use, whether the teen goes to college.



9. We want to know to what parent's education effects teen fertility, should we control for whether the teen goes to college?

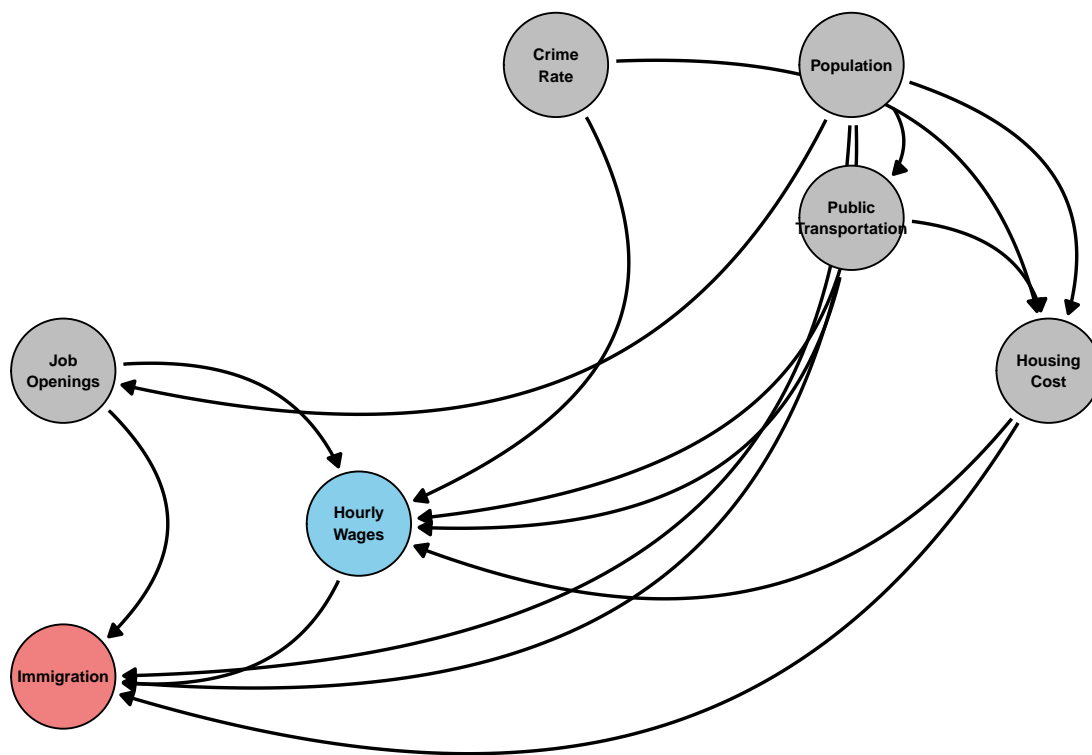
No, because whether teens goes to college or not is a mediator variable that also effects *Teen Fertility*. If the goal is to observe the direct effect of *Parent's Education* on fertility then condition on *College*, otherwise leave untouched.

10. We want see if there's a relationship between gun ownership and gun related deaths. We have the following variables at hand, person i's gun ownership, average violence in person i's community, average drug sales in person's I's community, person I's education level, average education level of person I's community, and person I's mental health. Build a DAG to map out the relationship between gun ownership and person i suffering a gun related death.



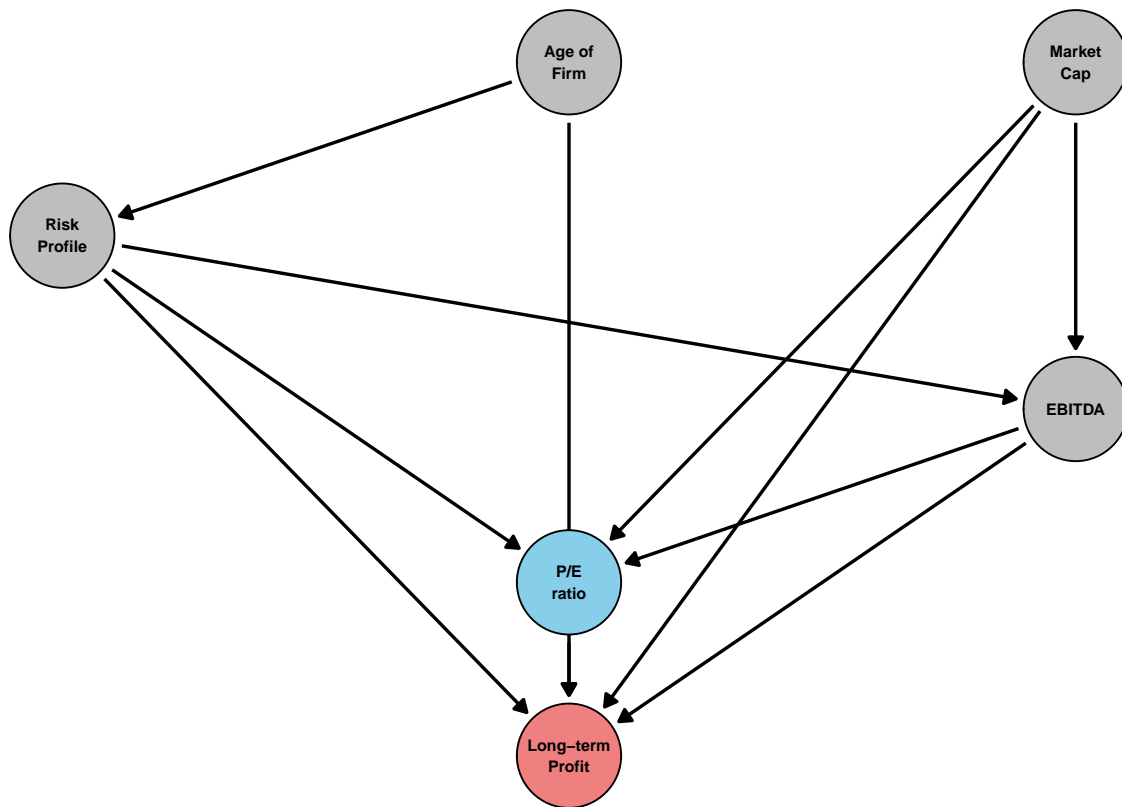
11. We want to see how wages in a city effects whether or not someone immigrates to that city.

What variables would you include in this model? Based on the variables you choose, write a DAG and what controls would you include? Here controls are needed for population, public transportation quality, crime rates, and job openings.



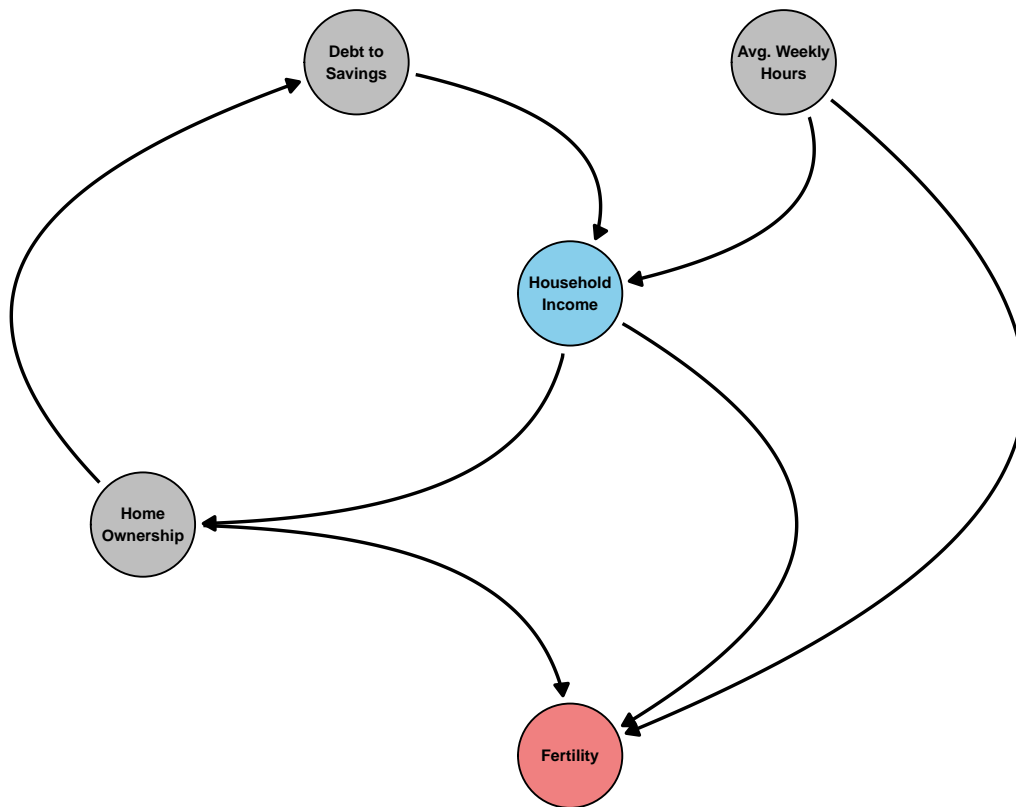
12. You're doing a market analysis about how price/earnings ratio effects long term profits of a company. What variables would you include in this model? Based on the variables you choose, write a DAG and what controls would you include?

Controls are needed for market cap, EBITDA, and the risk profile (beta).



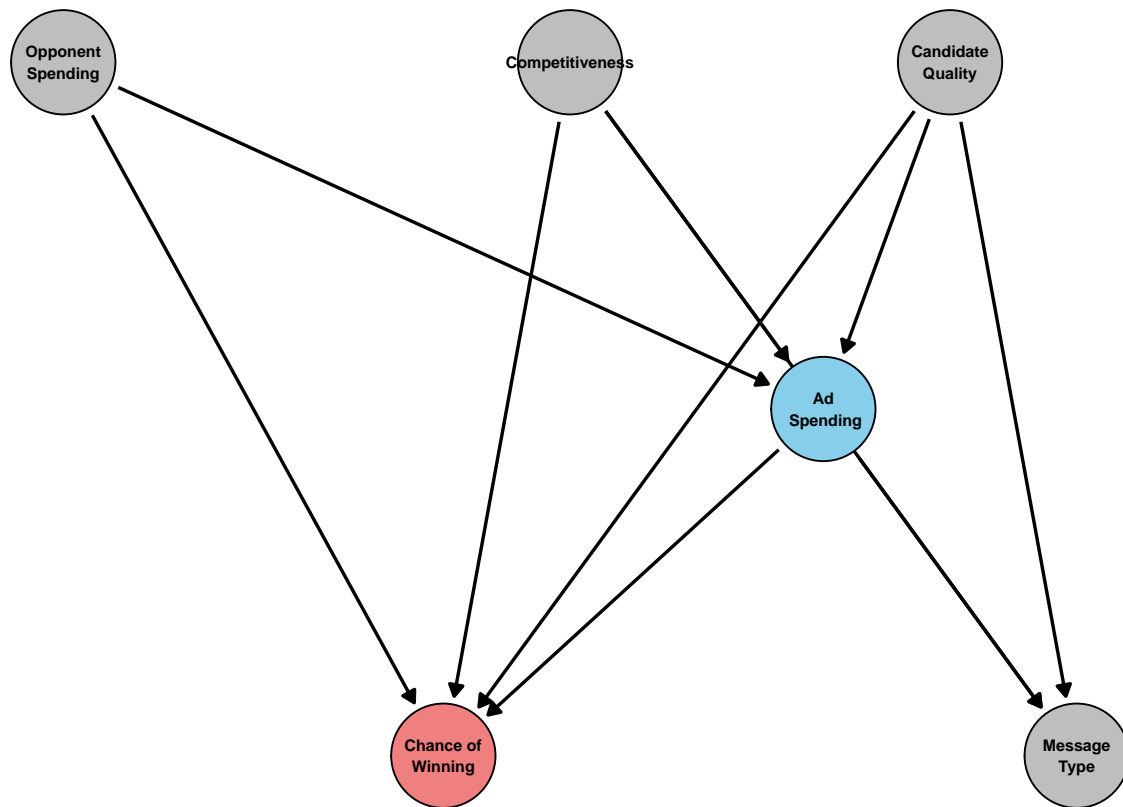
13. You want to test how economic activity effects fertility. What variables would you include in this model? Based on the variables you choose, write a DAG and what controls would you include?

If household income is set as the treatment group, the required controls are for debt-to-savings ratio and average hours worked per week.



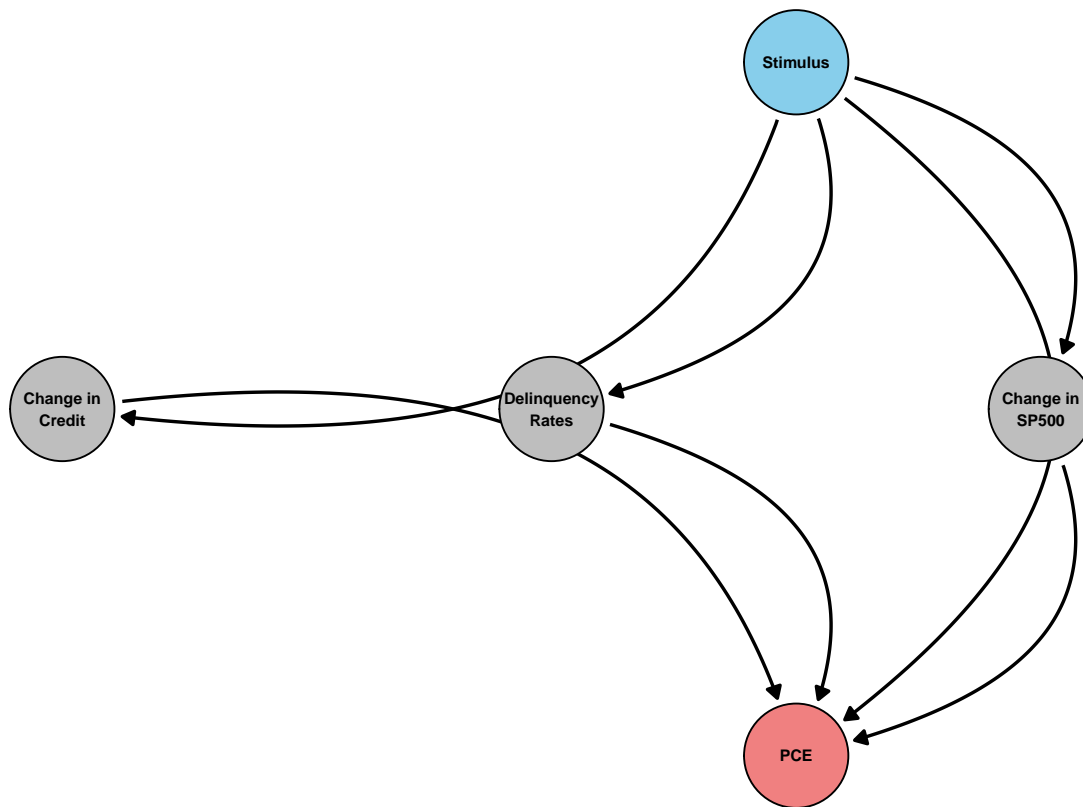
14. You want to know how ad spending effects political candidates' chances of winning. What variables would you include in this model? Based on the variables you choose, write a DAG and what controls would you include?

Candidate quality, competitiveness, and opponent spending require controls.



15. You want to test how stimulus spending effect economic activity. What variables would you include in this model? Based on the variables you choose, write a DAG and what controls would you include?

Using the *PCE*, personal consumption expenditures as the gauge for economic activity, there are no controls required for this model.



16. You want to test how SAT scores effect the ranking of college accepted to. What variables would you include in this model? Based on the variables you choose, write a DAG and what controls would you include?

All three variables require controls to isolate the effects of an individuals SAT scores on a college's rank.

