Homework for Chapter 19: Instrumental Variables

*How Does It Work?*

1. Which of the following is the best definition of instrument validity?
   1. The instrument is strongly related to the treatment variable
   2. It is possible to causally identify the effect of the instrument on the outcome by closing all back door paths
   3. There is no path between the instrument and the outcome variable
   4. All paths between the instrument and the outcome variable are either closed or travel through the treatment variable.
2. Assume that we have an instrument that is valid, but which is irrelevant. Which of the following is likely to happen? Select all that apply.
   1. As long as validity holds perfectly (which, granted, does not occur much in real life), there will be no issues
   2. The estimate will become very noisy
   3. The estimation software will return an error when it tries to divide by 0
   4. There is a decent chance of getting an enormous estimate even if the true effect is fairly small
3. You want to know whether better ventilation systems in homes can prevent asthma in the children who live there. You have data on two housing developments that are very similar in price and location, but which happened to use different contractors who tended to use different ventilation system qualities in the homes. Contractor A was more likely to use Good systems, and Contractor B was less likely. You have data on which contractor a home had, whether the system was Good or not, and whether a child in the house has Asthma or not.
   1. What four means would you calculate to produce an instrumental variables estimate of the effect of Good systems on Asthma rates, using Contractor as an instrument?
   2. How would you calculate that instrumental variables estimate from the four means?
4. The chapter discussed how variation over time in the amount of rainfall had been used as an instrument for a number of different treatments. If we assume that there are no back doors between rainfall and whatever our outcome of interest is, why would this still be a problem?
5. You are curious about the impact of airplane flights on the level of air pollution. You have measurements of the number of flights to and from each airport each day for many years, and also air pollution measures in the airport surrounding areas for those same days. You suspect there may be back doors between flights and pollution (both generally trend upwards over time anyway, higher economic activity in general may encourage both to be high at once, less environmentally-conscious people or governments in other areas may also encourage more air travel) so you want an instrumental variable. For each of the following potential IVs, discuss in a few sentences (i) whether you think it would be valid, (ii) why or why not, and (iii) what tests (using only observable data) you might want to perform when thinking about using the IV.
   1. The 9/11 attacks led the US government to ground nearly all aircraft for a short period of time. The instrument here would be “it’s the time when we’d expect pollution from flights just after 9/11 to show up in pollution measurements”.
   2. Every year, the number of flights that people take jumps up significantly around Christmas. The instrument here would be “it’s the last few weeks before Christmas.”
   3. As time goes on, airplane manufacturers develop planes that are faster, cheaper to fly, and safer. The instrument(s) here would be the quality and safety of flights at a given time.
   4. Airplane flights are heavily subsidized, and the amount of subsidy varies by city and over time. The instrument here would be “amount of subsidy in a given place/at a given time.”

*How is it Performed?*

1. Which of the following would make it an especially good idea to estimate your IV model using GMM instead of 2SLS? Select all that apply.
   1. When your errors are heteroskedastic
   2. When the instrument is weak
   3. When there are more instruments than treatments/endogenous variables
   4. When you are worried about a potential validity violation that you can’t test for
2. Which of the following provides the best intuitive explanation of why 2SLS produces an instrumental variables estimate of the effect of treatment?
   1. It uses only the part of the treatment predicted by the instrument, which by the IV validity assumption has no open back doors
   2. It starts from the assumption that the covariance moment between the instrument and the second-stage error is zero and then chooses coefficient to match that assumption, which is an IV assumption
   3. Because a 2SLS model with controls simply produces an estimate of the covariance between the outcome and instrument divided by the covariance between treatment and instrument, which is what IV is
   4. 2SLS forces the correlation between the second-stage error term and the instrument to be zero, which is a necessary assumption for IV
3. Consider the Vitamin E study by Oster way back from Chapter 4. In her study, she found that when Vitamin E became recommended by health professionals, the people who responded most strongly to the recommendation were already-healthy people. We want to do a study based on hers, using the Vitamin E recommendation as an instrument for taking Vitamin E, so we can see the effect of taking Vitamin E on blood pressure.
   1. Assume that the timing of the Vitamin E recommendation is totally random (and there’s no general trend in blood pressure). Is the fact that more-healthy people will respond more strongly to the recommendation a violation of *validity*?
   2. Does the fact that more-healthy people will respond more strongly to the recommendation suggest that the instrument is likely to be *not relevant*?
   3. What does the fact that more-healthy people will respond more strongly say for the kind of *local average treatment effect* we will estimate?
4. Consider the below table of hypothetical outcomes for four people, where Z is the instrument being considered for the treatment D.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Person | Treatment if Z = 0 | Treatment if Z = 1 | Outcome if D = 0 | Outcome if D = 1 |
| A | 0 | 0 | 5 | 6 |
| B | 1 | 1 | 5 | 7 |
| C | 0 | 1 | 5 | 8 |
| D | 1 | 0 | 5 | 11 |

* 1. What is the average treatment effect of D on Y among these four?
  2. What is the local average treatment effect when using Z as an instrument among these four, if we have two C’s and one of each other type of person? (bonus: why am I writing this question so as to include two C’s?)
  3. If we expect that the instrument *should* make treatment more likely, what would we call person D?
  4. What is the local average treatment effect we get if we exclude person D, and then use Z as an instrument among the remaining three?

1. When an instrument is weak, we say that we get weak-instrument bias. But surely, in a random sample, wouldn’t a weak instrument be as likely to overstate an effect as understate it? Why do we call this “bias” rather than just “additional noise”?
   1. Because if we only proceed with analysis if the F-statistic is big enough, this will “bias” results towards too-large effects
   2. Because, in a given sample, validity may fail by random chance, and this becomes a bigger driver of results relative to the intended effect the weaker the instrument is. Validity violations cause bias, so this is “bias.”
   3. Because weaker instruments isolate a smaller portion of the variance in X, favoring low variance over high variance and “bias”ing the estimate in a given direction as a result.
   4. Because the weaker an instrument is (the closer its average effect is to 0), the more likely it is that individuals with unusually-low effects will be defiers. The presence of these defiers will put negative weights in the treatment effect averages, biasing the effect downwards if it’s positive, and upwards if it’s negative.

Coding (which includes any how-the-pros-do-it questions)