Problem Set 2 Solutions

11.126/11.249/14.48

Due March 20, 2007 in class

1. (25 points) In 2001, Mexico instituted a program called Programa Escuelas de Calidad (PEC). Primary schools enrolled in PEC received a 5-year grant of $15,000 for supplies and teacher skill development and free training for the principal. PEC also requires school staff and parent associations to become actively involved in drafting a plan for school improvement. In principle, every primary school may participate; in practice, about 10% did by the 2003 – 04 school year. The program specifically targets disadvantaged urban schools through an advertising campaign.

You have a data set that tells you the dropout rate (the proportion of students who did not return to school in the following year) at 74,701 primary schools in each year from 2000 – 01 to 2003 – 04. Of these schools, 1767 became enrolled in PEC at the beginning of 2001 – 02, an additional 7477 became enrolled by the beginning of 2003 – 04, and 65,457 were still not enrolled by the end of the period covered. (You can assume that once a school becomes enrolled, it stays with the program for 5 years.) You want to estimate the effect of participation in PEC on a school's dropout rate.

a. (5 points) Someone suggests that you look at the 1767 schools that were enrolled in 2001 – 02 and compare their dropout rate in 2000 – 01 to their dropout rate in 2003 – 04. Do you think this is a good approach? Explain why or why not.

Answer: No, it is not. This approach confounds the effect of PEC with the effects of any other factors that are changing over time. For example, dropout rates may be changing in response to improvements in the economy or other programs that affect schools.

b. (5 points) Someone else suggests that you compare the 1767 schools that had been receiving PEC from the start to the 65,457 schools that never enrolled, and take the difference in their 2003 – 04 dropout rates as the effect of PEC. Do you think this is a good approach? Explain why or why not.

Answer: No, it is not. The schools that enrolled in PEC are clearly not equivalent to the non-PEC schools in every respect except for their participation in PEC. It is unclear in which direction this comparison is biased, however. On the one hand, the PEC schools had sufficient motivation (from staff and parents) to enroll in the program, and urban schools were targeted (urban schools very likely have lower dropout); these factors suggest that the PEC schools would have lower dropout even without PEC. On the other hand, the PEC schools felt they had some need to enroll and disadvantaged schools were targeted; these factors suggest that the PEC schools would have higher dropout without PEC. Even though we can't be sure in which direction this approach is biased, there is little reason to believe that it correctly estimates the causal effect of PEC.

c. (10 points) Using this data set, devise your own research strategy to estimate the effect of PEC on dropout. Be precise in describing how you will calculate your estimate.

Answer: You likely want to do some kind of difference-in-differences comparison, but there is more than one reasonable answer. The simplest approach is to calculate the change in the dropout rate over $2000-01$ to $2003-04$ for the 1767 schools that enrolled initially and the $65,457$ schools that never enrolled. The difference between these changes would be our estimate of the causal effect of PEC. Alternatively, you could use the 7477 schools that enrolled later as a control group and calculate changes in the dropout rate over $2000-01$ to $2001-02$. The advantage of this strategy is that the control group likely corresponds more closely to the treatment group, since both eventually enroll in PEC; the disadvantages are that the sample size is smaller (so standard errors will be larger) and we have to measure the change in dropout over a single year (which might make PEC look less useful than it really is if some of its effects take a while to kick in).

d. (5 points) Even though you believe your strategy in (c) to be a good one, it probably requires some assumptions in order to be valid. Explain the conditions under which your approach in (c) would give you a misleading estimate of the causal effect of PEC.

Answer: The key assumption, if you answered difference-in-differences in (c), is that the underlying trend in dropout must be the same across treatment and control schools. That is, absent PEC, the schools that enrolled in PEC would have had the same *change* in dropout rates as the schools that didn't enroll.

You do NOT need to assume that assignment to PEC is essentially random. If we had a true experiment here, then there would be nothing wrong with the method outlined in part (b); random assignment means that, absent PEC, the *level* of dropout would be the same in treatment and control schools.

2. (21 points) You want to estimate the return to an additional year of schooling, but you are concerned about ability bias. Consider each of the following empirical approaches, and explain why or why not you think it would solve the ability bias problem and correctly estimate the return to schooling:

a. (7 points) You have data on each person's state of birth. You regress average earnings of people born in state s on average years of schooling for people born in state s.

Answer: This removes ability bias if we believe that average ability is the same across high-education and low-education states. This could be false if high-ability parents tend to move to places like Massachusetts or New York, where they have high ability children.

Even if we believe that average ability is the same across states, this approach probably does not correctly estimate the return to education. High education states tend to be much more urban and have higher costs of living; these considerations will push up earnings there.

Some people read this question as proposing to estimate a standard earnings regression with data on individuals, but including only individuals born within one state (s). I'm not sure how the references to "average earnings" can be interpreted in that context, but I didn't dock marks for interpreting the question that way as long as you had a good explanation.

b. (7 points) You find a data set on twins. You regress an individual's earnings on his or her years of schooling, controlling for the individual's age and the twin's years of schooling.

Answer: This certainly controls for ability to some extent, especially if the twins are identical and grew up in the same household. However, consider exactly what this approach is doing. Suppose we have individuals A and B with 18 and 12 years of schooling respectively (assume they are the same age). We know that $A's$ twin (A^T) and $B's$ twin (B^T) each have 16 years of schooling. This approach is saying that the earnings difference between A and B can be attributed entirely to the schooling difference between A and B, because the fact that A^T and B^T have the same years of education means that we have controlled for ability. However, there are at least two problems with this inference:

i. On average, $A's$ family has higher years of schooling than $B's$ family (17 versus 14 years). It is natural to infer that genetic and family background characteristics are "better" in $A's$ family.

ii. The fact that A achieved a higher level of schooling than his twin suggests that he might be higher ability than his twin. Obviously if the twins are identical and grew up together, they must have the same genes and family background, but A may have had life experiences that made him more motivated or more interested in pursuing a high-profile career.

The upshot is that this method likely controls for ability to some extent, but not fully. This was a hard question, and grading was largely based on the justification you gave for your answer.

c. (7 points) You have data on the Socio-Economic Status (SES) of each person's parents. You regress earnings on years of schooling and age, and you instrument years of schooling with parental SES. (If you're confused about instrumental variables, try looking at the econometrics notes for recitation again.)

Answer: This is a horrible instrument! Parental SES very likely has a direct effect on earnings apart from its effect on education: e.g., high SES parents may teach their kids more outside the formal school system or may give their kids access to better job networks. This violates one of the two key assumptions necessary for the IV strategy.

3. (24 points) Read "How Experts Differ From Novices." The author lists six characteristics of experts.

a. (12 points) To what extent do you think changes in technology have increased demand for these expert characteristics? (One paragraph is sufficient. The last two characteristics may not be relevant.)

Answer: Answers will vary. A 12/12 answer made specific reference to some of the expert characteristics (not just experts in a general sense) and mentioned specific ways in which technology has (or has not) increased demand for workers with this characteristic. A 9 or 10 means a "good" answer.

For example, experts' knowledge of a subject area is organized around a set of "big ideas" that they use to determine how to attack a problem (characteristic #2). Novices may be able to solve a problem once they have slotted it into a category, but they have a harder time determining which category is correct (characteristic #3). Computers are similar to novices in this way. Once the universe of possible answers is well defined, computers are very efficient at searching this universe for the right answer. However, it is very difficult to design a simple algorithm for choosing the appropriate universe in the first place; experts use a large number of clues in the statement of a problem, some of which may conflict with one another, and aggregate them to determine the problem's most likely category. The problem is made worse by the fact that experts cannot even articulate what their procedure is (characteristic #5). Thus the development of computers has substituted for novice workers and increased the relative demand for experts.

b. (12 points) Consider the teaching of algebra in 8th or 9th grade, a subject most people would classify as problem solving. Does all of this subject matter qualify as "expert problem solving" in the sense defined by the chapter? Does any part of the subject matter qualify? Explain why teaching the kinds of skills involved in "expert problem solving" may be harder than teaching basic algebra. (Again, one paragraph is sufficient and the last two characteristics may not be relevant.)

Answer: Once a problem has been transformed into equations, the equations can be solved by applying simple, articulable rules; this does not qualify as expert problem solving. The transformation of a problem in words into a set of equations is certainly closer to expert problem solving, since there is often no easy algorithm that tells a student whether two numbers should be added, subtracted, multiplied, or divided. Poorly designed algebra problems, however, may allow students to set up the equations by application of rote rules (e.g., if the problem mentions two numbers and uses the word "sum" in the final question).

It is relatively difficult to teach expert problem solving skills precisely because the skills involved cannot be enunciated in a set of simple and articulable rules. Teachers cannot tell students how to solve algebra problems like an expert; instead, they must know what experiences are most likely to give students an intuitive understanding of which method to apply to a given context.

4. (30 points) Consider the following model: there is a very large number of workers, and they vary uniformly in their lifetime productivity between 0 and 1. If we label each worker by her productivity y, this means that the average level of productivity for workers in the interval $[\underline{y}, \bar{y}]$ is

$$\frac{1}{\bar{y} - \underline{y}} \int_{\underline{y}}^{\bar{y}} y dy$$

Each worker decides whether or not to go to college before entering the workforce. A worker chooses to go to college if the increase in lifetime earnings is greater than her monetary and psychological cost; the cost for a worker of productivity y is $c(y) = 1 - y$. That is, more productive workers have a lower cost of going to college. Going to college doesn't affect productivity for a given individual.

Employers can't observe anything directly about a person's productivity. In fact, for the purposes of the model they never observe an individual's productivity until that person retires. All the employer can observe is whether a person went to college. Thus they offer a wage w_1 to people with a college degree and a wage w_0 to people without a college degree. Because of competition, the wage for a college-educated worker must equal the average lifetime productivity of people with a college degree, and the wage for people without college must equal their average lifetime productivity.

a. (8 points) There will be some cutoff $\hat{y}$ such that workers above $\hat{y}$ go to college and workers below $\hat{y}$ don't. Find w_1 and w_0 as a function of $\hat{y}$.

Answer: We know that w_1 is the average productivity of workers above $\hat{y}$, so

$$\begin{aligned} w_1 &= \frac{1}{1-\hat{y}} \int_{\hat{y}}^{1} y dy \\ &= \frac{1}{1-\hat{y}} \cdot \frac{1}{2} y^2 |_{\hat{y}}^{1} \\ &= \frac{1/2}{1-\hat{y}} (1-\hat{y}^2) \\ &= \frac{1+\hat{y}}{2} \end{aligned}$$

Similarly,

$$\begin{aligned} w_0 &= \frac{1}{\hat{y}} \int_{0}^{\hat{y}} y dy \\ &= \frac{1/2}{\hat{y}} \hat{y}^2 \\ &= \frac{\hat{y}}{2} \end{aligned}$$

b. (8 points) In making their schooling decisions, workers all want to maximize the value of their wage minus their cost of going to college. Using this information, which workers will go to college in equilibrium? What is the wage premium for going to college $(w_1 - w_0)$?

Answer: We have $w_1 - w_0 = 1/2$ no matter the value of $\hat{y}$. Since the worker at $\hat{y}$ must be indifferent between going to college or not, we can conclude that $\hat{y} = 1/2$. That is, all workers with $y > 1/2$ go to college, all workers with $y < 1/2$ do not, and the worker with $y = 1/2$ is indifferent.

c. (4 points) Suppose that we had $c(y) = \frac{1}{2}$ for all workers instead. What would be the wage premium for going to college in that case? Explain why. (You shouldn't need to do any calculations.)

Answer: This is slightly confusing because if $c(y) = \frac{1}{2}$ then in equilibrium no one would go to college. However, the idea is that if the cost of college is constant, then it can't serve as a signal; if you got that much, full marks. The "hypothetical" wage premium that a person believes they could get by going to college must certainly be less than $\frac{1}{2}$. We would expect it to be 0 because employers have no particular reason to believe that a renegade who went to college is more likely to be high ability than low ability.

d. (5 points) Go back to the original model with $c(y) = 1 - y$. We run an OLS regression to figure out the return to college from an individual's perspective (i.e., how much more a given person could earn if she decided to go to college); call our estimate $\widehat{\beta}$ Do we expect $\widehat{\beta}$ to overestimate the individual's return, underestimate it, or get it right?

Answer: It gets the individual return right. Even a person with $y = 0$ can get a wage premium $w_1 - w_0$ in this model; she just doesn't want to because going to school is too costly.

e. (5 points) A government official is considering a program that would increase college attendance by a proportion α of the population. If he uses the same regression as above and concludes that society will be $\alpha\widehat{\beta}$ richer as a result of the program, is he overestimating the social benefit, underestimating it, or getting it right (from the perspective of this model)?

Answer: He is overestimating the benefit, because the true social benefit is 0. Remember, no one's productivity increases by going to college–in this model, it would be efficient to prevent anyone from getting a degree!