

## Problem Set 2

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.
  - (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.
  - (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.
  - (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.
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  $s$ .
- (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.
- (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.)
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.)
- (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.)
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  $[y, \bar{y}]$  is

$$\frac{1}{\bar{y} - y} \int_y^{\bar{y}} y dy$$

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  $\hat{y}$ .
- (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 ( $w_1 - w_0$ )?
- (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.)
- (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  $\hat{\beta}$ . Do we expect  $\hat{\beta}$  to overestimate the individual's return, underestimate it, or get it right?
- (e) (5 points) A government official is considering a program that would increase college attendance by a proportion  $\alpha$  of the population. If he uses the same regression as above and concludes that society will be  $\alpha\hat{\beta}$  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)?