

CHAPTER 1

SOLUTIONS TO PROBLEMS

1.1 It does not make sense to pose the question in terms of causality. Economists would assume that students choose a mix of studying and working (and other activities, such as attending class, leisure, and sleeping) based on rational behavior, such as maximizing utility subject to the constraint that there are only 168 hours in a week. We can then use statistical methods to measure the association between studying and working, including regression analysis that we cover starting in Chapter 2. But we would not be claiming that one variable “causes” the other. They are both choice variables of the student.

1.2 (i) Ideally, we could randomly assign students to classes of different sizes. That is, each student is assigned a different class size without regard to any student characteristics such as ability and family background. For reasons we will see in Chapter 2, we would like substantial variation in class sizes (subject, of course, to ethical considerations and resource constraints).

(ii) A negative correlation means that larger class size is associated with lower performance. We might find a negative correlation because larger class size actually hurts performance. However, with observational data, there are other reasons we might find a negative relationship. For example, children from more affluent families might be more likely to attend schools with smaller class sizes, and affluent children generally score better on standardized tests. Another possibility is that, within a school, a principal might assign the better students to smaller classes. Or, some parents might insist their children are in the smaller classes, and these same parents tend to be more involved in their children’s education.

(iii) Given the potential for confounding factors – some of which are listed in (ii) – finding a negative correlation would not be strong evidence that smaller class sizes actually lead to better performance. Some way of controlling for the confounding factors is needed, and this is the subject of multiple regression analysis.

SOLUTIONS TO COMPUTER EXERCISES

C1.1 (i) The average of *educ* is about 12.6 years. There are two people reporting zero years of education, and 19 people reporting 18 years of education.

(ii) The average of *wage* is about \$5.90, which seems low in the year 2008.

(iii) Using Table B-60 in the 2004 *Economic Report of the President*, the CPI was 56.9 in 1976 and 184.0 in 2003.

(iv) To convert 1976 dollars into 2003 dollars, we use the ratio of the CPIs, which is $184/56.9 \approx 3.23$. Therefore, the average hourly wage in 2003 dollars is roughly $3.23(\$5.90) \approx \19.06 , which is a reasonable figure.

(v) The sample contains 252 women (the number of observations with *female* = 1) and 274 men.

C1.3 (i) The largest is 100, the smallest is 0.

(ii) 38 out of 1,823, or about 2.1 percent of the sample.

(iii) 17

(iv) The average of *math4* is about 71.9 and the average of *read4* is about 60.1. So, at least in 2001, the reading test was harder to pass.

(v) The sample correlation between *math4* and *read4* is about .843, which is a very high degree of (linear) association. Not surprisingly, schools that have high pass rates on one test have a strong tendency to have high pass rates on the other test.

(vi) The average of *exppp* is about \$5,194.87. The standard deviation is \$1,091.89, which shows rather wide variation in spending per pupil. [The minimum is \$1,206.88 and the maximum is \$11,957.64.]