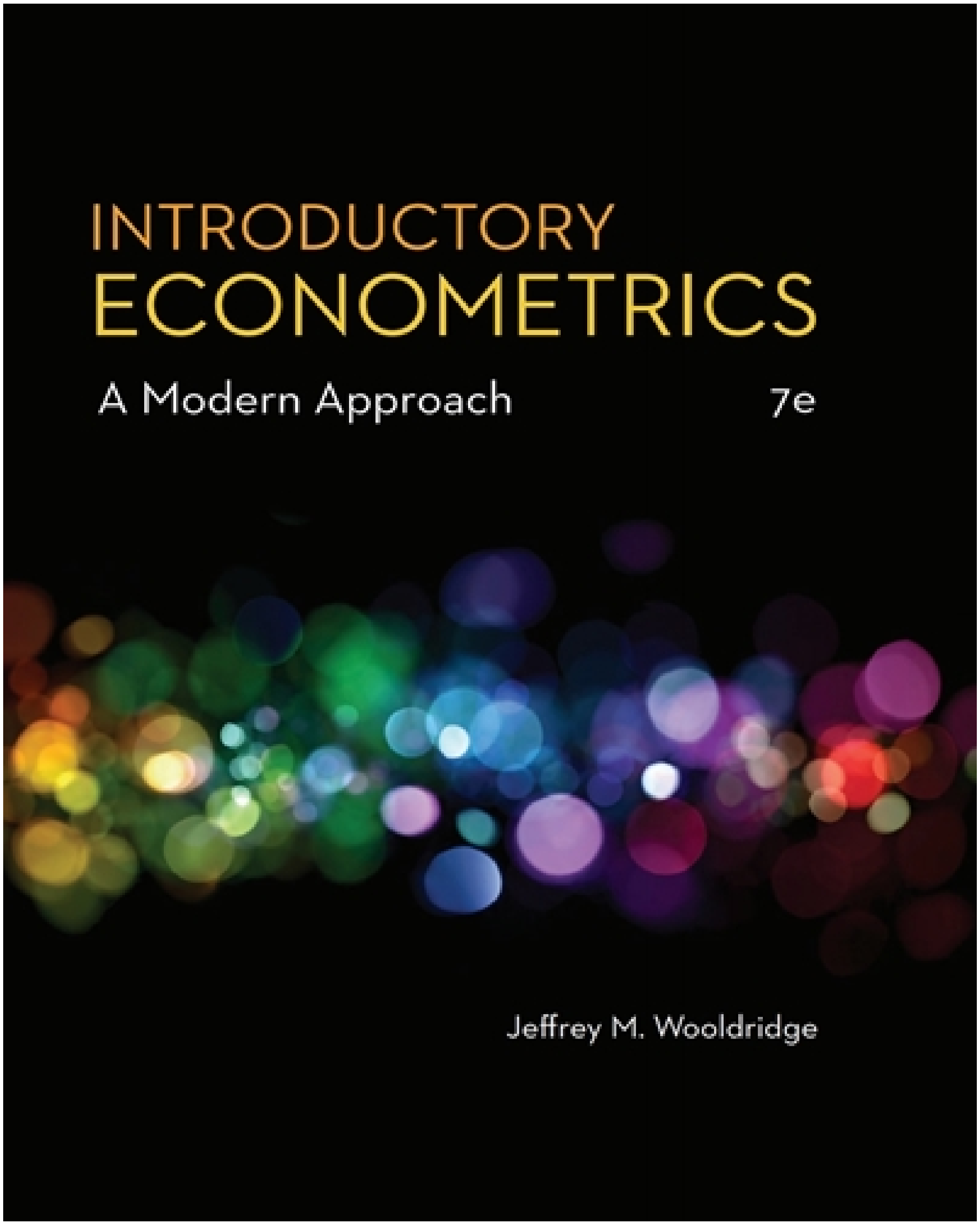
Solutions for Introductory Econometrics A Modern Approach 7th Edition by Wooldridge



Solutions

CHAPTER 1

The Nature of Econometrics and Economic Data

SOLUTIONS TO PROBLEMS

1.1 (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 a larger class size is associated with lower performance. We might find a negative correlation because a 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 might 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 to be placed in 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.

1.3 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, which 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.

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* in the sample is about \$5.90, which seems low.

(iii) Using Table B-60 in the 2004 *Economic Report of the President*, the CPI was 56.9 in 1976 and 233 in 2013.

(iv) To convert 1976 dollars into 2013 dollars, we use the ratio of the CPIs, which is $233 / 56.9 \approx 4.09$. Therefore, the average hourly wage in 2013 dollars is roughly $4.09(\$5.90) \approx \$24.13$, 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) 289 out of 1,823, or about 15.85 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.]

(vii) The percentage by which school A outspends school B is

(vii) The percentage by which school A outspends school B is

$$100 \cdot \frac{(6{,}000 - 5{,}500)}{5{,}500} \approx 9.09\%.$$

When we use the approximation based on the difference in the natural logs we get a somewhat smaller number:

$$100 \cdot [\log(6{,}000) - \log(5{,}500)] \approx 8.71\%.$$

C1.5 (i) The smallest and largest values of *children* are 0 and 13, respectively. The average is about 2.27.

(ii) Out of 4,358 women, only 611 have electricity in the home, or about 14.02 percent.

(iii) The average of *children* for women without electricity is about 2.33, and for those with electricity it is about 1.90. So, on average, women with electricity have .43 fewer children than those who do not.

(iv) We cannot infer causality here. There are many confounding factors that may be related to the number of children and the presence of electricity in the home; household income and level of education are two possibilities. For example, it could be that women with more education have fewer children and are more likely to have electricity in the home (the latter due to an income effect).

C1.7 (i) The percentage of men in the sample report abusing alcohol is 9.9. The employment rate is 24.3.

(ii) The employment rate of men who abuse alcohol is 22.6.

(iii) The employment rate who do not abuse alcohol is 24.5.

(iv) The employment rates of men who abuse alcohol and who do not are 22.6 and 24.5, respectively. The difference in these employment rates is very less, which means that alcohol abuse does not cause unemployment.