Empirical Exercise: Chapter 8

In the folder, you will find a data file **CPS2015**, which contains data for full-time, full-year workers, ages 25–34, with a high school diploma or B.A./B.S. as their highest degree. A detailed description is given in **CPS2015\_Description**, also available in the folder. (These are the same data as in **CPS96\_15**, iwill use in Empirical Exercise 9, but are limited to the year 2015.) In this exercise, you will investigate the relationship between a worker's age and earnings. (Generally, older workers have more job experience, leading to higher productivity and higher earnings.)

- a. Run a regression of average hourly earnings (*AHE*) on age (*Age*), sex (*Female*), and education (*Bachelor*). If *Age* increases from 25 to 26, how are earnings expected to change? If *Age* increases from 33 to 34, how are earnings expected to change?
- b. Run a regression of the logarithm of average hourly earnings, ln(*AHE*), on *Age*, *Female*, and *Bachelor*. If *Age* increases from 25 to 26, how are earnings expected to change? If *Age* increases from 33 to 34, how are earnings expected to change?
- c. Run a regression of the logarithm of average hourly earnings, ln(AHE), on ln(Age), Female, and Bachelor. If Age increases from 25 to 26, how are earnings expected to change? If Age increases from 33 to 34, how are earnings expected to change?
- d. Run a regression of the logarithm of average hourly earnings, ln(*AHE*), on *Age*, *Age*<sup>2</sup>, *Female*, and *Bachelor*. If *Age* increases from 25 to 26, how are earnings expected to change? If *Age* increases from 33 to 34, how are earnings expected to change?
- e. Do you prefer the regression in (c) to the regression in (b)? Explain.
- f. Do you prefer the regression in (d) to the regression in (b)? Explain.
- g. Do you prefer the regression in (d) to the regression in (c)? Explain.
- h. Plot the regression relation between *Age* and ln(*AHE*) from (b), (c), and (d) for males with a high school diploma. Describe the similarities and differences between the estimated regression functions. Would your answer change if you plotted the regression function for females with college degrees?
- i. Run a regression of ln(*AHE*) on *Age*, *Age*<sup>2</sup>, *Female*, *Bachelor*, and the interaction term . What does the coefficient on the interaction term measure? Alexis is a 30-year-old female with a bachelor's degree. What does the regression predict for her value of ln(*AHE*)? Jane is a 30-year-old female with a high school diploma. What does the regression predict for her value of ln(*AHE*)? What is the predicted difference between Alexis's and Jane's earnings? Bob is a 30-year-old male with a bachelor's degree. What does the regression predict for his value of ln(*AHE*)? Jim is a 30-year-old male with a high school diploma. What does the regression predict for his value of ln(*AHE*)? What is the predicted difference between Bob's and Jim's earnings?
- j. Is the effect of *Age* on earnings different for men than for women? Specify and estimate a regression that you can use to answer this question.
- k. Is the effect of *Age* on earnings different for high school graduates than for college graduates? Specify and estimate a regression that you can use to answer this question.
- 1. After running all these regressions (and any others that you want to run), summarize the effect of age on earnings for young workers.