

Independent Exercise 1

Week 8

March 23, 2020

Your work on this exercise does not need to be submitted and will not be graded. I can provide feedback and help as needed. You may work on this exercise with others (practicing appropriate social distancing).

Consider the following equation, which can be used to study the effects of skipping class on college GPA:

$$\widehat{colGPA} = 1.369 + 0.421 \text{ hsGPA} + 0.012 \text{ ACT} - 0.095 \text{ skipped} + 0.030 \text{ alcohol}$$

(0.331) (0.094) (0.011) (0.027) (0.022)

Standard errors are in parentheses below the corresponding coefficients, $N = 64$, and $R^2 = 0.458$.

- (i) Use the data in GPA1.DTA to replicate these results.
- (ii) Reestimate the marginal effect of skipping class on GPA for male and female subsamples.
- (ii) In terms of the model parameters, state the null hypothesis that *skipped* has no effect on *colGPA*, controlling for high school GPA, ACT score, and weekly alcohol consumption.
- (iii) Identify which regression coefficients are statistically significant and which are not. Explain your reasoning.
- (iv) Evaluate the economic significance of the statistically significant coefficients.
- (v) Compare the 95% confidence intervals for each of the regressors. Explain how to identify statistically significant estimates using only these intervals.
- (vi) Can you say with certainty that skipping college lectures decreases one's GPA? Suggest some unobserved factors that may be correlated with *skipped*.

If you can do everything here, as well as interpret the coefficient of determination (R^2) and have a good sense of what your variables are and what they mean, you're on track. These are the basics of estimating a regression and understanding what you get back from your econometric software. Most of the above will be your first steps every time you estimate any econometric model.