ECON 0150 | Spring 2026 | Homework 5.3

Due: Sunday April 19, at 11:59PM

Homework is designed to both test your knowledge and challenge you to apply familiar concepts in new applications. Answer clearly and completely. You are welcomed and encouraged to work in groups so long as your work is your own. Use the provided datasets to answer the following questions. Then submit your figures and answers to Gradescope.

Q1. Adding Numerical Controls to the Recession and Health Model

Over the past several assignments, you have progressively built up a model of the relationship between county unemployment and BMI using 2011 BRFSS data. You started with a simple regression in HW 4.1, added a categorical control in HW 5.1, and tested for interaction effects in HW 5.2. Zhang et al. (2014) also control for numerical variables like age and income. Let's do the same.

a) A researcher wants to estimate the effect of county unemployment on BMI while controlling for age, gender, college education, and marital status. Write the regression model.

b) Fit a multiple regression with both unemployment and age:

$$\texttt{BMI5}_i = \beta_0 + \beta_1 \times \texttt{unemployment\_rate}_i + \beta_2 \times \texttt{AGE}_i + \varepsilon_i \tag{1}$$

Report the estimated coefficients. Interpret each one using "holding constant" language. How does the coefficient on `unemployment_rate` compare to your simple regression from HW 4.1?

c) Now fit the full model you wrote in part (a):

$$\texttt{BMI5}_i = \beta_0 + \beta_1 \times \texttt{unemployment\_rate}_i + \beta_2 \times \texttt{Female}_i + \beta_3 \times \texttt{AGE}_i + \beta_4 \times \texttt{College}_i + \beta_5 \times \texttt{Married}_i + \varepsilon_i \tag{2}$$

Report the estimated coefficients. How does the coefficient on `unemployment_rate` compare to the simple model from HW 4.1? Has the sign, magnitude, or significance changed?

Q2. Reading Regression Output with Multiple Controls

A labor economist studies the determinants of hourly wages using a sample of $n = 500$ full-time workers. They estimate:

$$\texttt{wage}_i = \beta_0 + \beta_1 \times \texttt{experience}_i + \beta_2 \times \texttt{education}_i + \beta_3 \times \texttt{urban}_i + \varepsilon_i \tag{3}$$

where `experience` is years of work experience, `education` is years of schooling, and `urban` = 1 if the worker lives in an urban area.

	coef	std err	t	P>|t|	[0.025	0.975]
Intercept	3.20	1.450	2.207	0.028	0.35	6.05
experience	0.45	0.060	7.500	0.000	0.33	0.57
education	1.85	0.180	10.278	0.000	1.50	2.20
urban	2.90	0.720	4.028	0.000	1.49	4.31

a) Interpret the coefficient on `experience` (0.45) in context. What does "holding constant" mean here?

b) Interpret the coefficient on `education` (1.85) in context.

c) Interpret the coefficient on `urban` (2.90) in context. How is the interpretation of a binary control variable different from a numerical one?

d) What is the predicted hourly wage for a worker with 10 years of experience, 16 years of education, living in an urban area?

e) What is the predicted hourly wage for a worker with the same experience and education, but living in a rural area?