

FINANCIAL ECONOMETRICS
SHANGHAI JIAO TONG UNIVERSITY
SUMMER SCHOOL 2025
EMPIRICAL EXERCISE 5A

FATIH KANSOY

Reference: Stock, J.H. and Watson, M.W., Introduction to Econometrics (4th Edition)

Data: Birthweight_Smoking.csv

Empirical Exercise 7.1. Use the `Birthweight_Smoking` data set to answer the following questions. To begin, run three regressions:

- (1) *Birthweight* on *Smoker*
 - (2) *Birthweight* on *Smoker*, *Alcohol*, and *Nprevist*
 - (3) *Birthweight* on *Smoker*, *Alcohol*, *Nprevist*, and *Unmarried*
- (a) What is the value of the estimated effect of smoking on birth weight in each of the regressions?
 - (b) Construct a 95% confidence interval for the effect of smoking on birth weight, using each of the regressions.
 - (c) Does the coefficient on *Smoker* in regression (1) suffer from omitted variable bias? Explain.
 - (d) Does the coefficient on *Smoker* in regression (2) suffer from omitted variable bias? Explain.
 - (e) Consider the coefficient on *Unmarried* in regression (3).
 - (i) Construct a 95% confidence interval for the coefficient.
 - (ii) Is the coefficient statistically significant? Explain.

- (iii) Is the magnitude of the coefficient large? Explain.
- (iv) A family advocacy group notes that the large coefficient suggests that public policies that encourage marriage will lead, on average, to healthier babies. Do you agree? (*Hint:* Review the discussion of control variables in Section 6.8.)
- (f) Consider the various other control variables in the data set. Which do you think should be included in the regression? Using a table like Table 7.1, examine the robustness of the confidence interval you constructed in (b). What is a reasonable 95% confidence interval for the effect of smoking on birth weight?