

FINANCIAL ECONOMETRICS
SHANGHAI JIAO TONG UNIVERSITY
SUMMER SCHOOL 2025
EMPIRICAL EXERCISE 4A

FATIH KANSOY

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

Data: Birthweight_Smoking.csv

Empirical Exercise 6.1. Use the `Birthweight_Smoking` data set introduced in Empirical Exercise E5.3 to answer the following questions.

- (a) Regress *Birthweight* on *Smoker*. What is the estimated effect of smoking on birth weight?
- (b) Regress *Birthweight* on *Smoker*, *Alcohol*, and *Nprevist*.
 - (i) Using the two conditions in Key Concept 6.1, explain why the exclusion of *Alcohol* and *Nprevist* could lead to omitted variable bias in the regression estimated in (a).
 - (ii) Is the estimated effect of smoking on birth weight substantially different from the regression that excludes *Alcohol* and *Nprevist*? Does the regression in (a) seem to suffer from omitted variable bias?
 - (iii) Jane smoked during her pregnancy, did not drink alcohol, and had 8 prenatal care visits. Use the regression to predict the birth weight of Jane's child.
 - (iv) Compute R^2 and $\bar{R}^2$. Why are they so similar?
 - (v) How should you interpret the coefficient on *Nprevist*? Does the coefficient measure a causal effect of prenatal visits on birth weight? If not, what does it measure?

- (c) Estimate the coefficient on *Smoking* for the multiple regression model in (b), using the three-step process in Appendix 6.3 (the Frisch–Waugh theorem). Verify that the three-step process yields the same estimated coefficient for *Smoking* as that obtained in (b).
- (d) An alternative way to control for prenatal visits is to use the binary variables *Trip0* through *Trip3*. Regress *Birthweight* on *Smoker*, *Alcohol*, *Trip0*, *Trip2*, and *Trip3*.
- (i) Why is *Trip1* excluded from the regression? What would happen if you included it in the regression?
 - (ii) The estimated coefficient on *Trip0* is large and negative. What does this coefficient measure? Interpret its value.
 - (iii) Interpret the value of the estimated coefficients on *Trip2* and *Trip3*.
 - (iv) Does the regression in (d) explain a larger fraction of the variance in birth weight than the regression in (b)?

Calculations performed using Stata. See accompanying EE_4A_solution.do file.