London School of Economics and Political Science
Department of Economics

Tatiana Komarova
Summer 2023

EC312: Econometrics for MRes Students
Part 2. Computer Exercise 1 on 2SLS

The data *FERTIL2.dta* in 'fertil2' include, for women in Botswana during 1988, information on number of children, years of education, age, and religious and economic status variables.

(a) Estimate the model

$$children = \beta_0 + \beta_1 educ + \beta_2 age + \beta_3 age^2 + \beta_4 urban + \varepsilon$$

by OLS and interpret the estimates. In particular, holding age fixed, what is the estimated effect of another year of education on fertility? If 100 women receive another year of education, how many fewer children are they expected to have?

Discuss why it is very likely that regressor *educ* is endogenous.

(b) The variable $frsthalf$ is a dummy variable equal to one if the woman was born during the first six months of the year. Assuming that $frsthalf$ is uncorrelated with the error term from part (a), show that $frsthalf$ is a reasonable excluded IV candidate for *educ*.

(c) Estimate the model from part (a) by using $frsthalf$ as an IV for *educ* (calculate heteroskedasticity-robust standard errors). Compare the estimated effect of education with the OLS estimate from part (a).

(d) Add the binary variables *electric*, *tv*, and *bicycle* to the model and assume these are exogenous. Estimate the equation by OLS and 2SLS and compare the estimated coefficients on *educ*. Interpret the coefficient on *tv* and explain why television ownership has a negative effect on fertility in the OLS regression.