

## ECO 102: Topics in Economics

Ecole Polytechnique, Spring 2022

*Professors:* Geoffroy Barrows, Benoit Schmutz

*Teaching Assistants:* Arnault Chatelain, Maddalena Conte

---

### TD3: Omitted Variable Bias

In this assignment, you are going to replicate the empirical findings in Mankiew et al (1992) and extend. Download the folder TD3.zip, type your answers directly in the Tex script and your Stata commands in a do file.

#### Exercise

1. Load the data MRW QJE1992.xlsx into Stata. Do all the numerical variables show up as numbers? If not, perhaps you need to reformat some variables.
2. Export descriptive statistics for relevant variables by sample (non-oil, intermediate, OECD).
3. Generate a scatter plot of growth rate in GDP per capita against log GDP per worker in 1960, similar to Figure 1A in the paper. Does your figure look exactly the same Figure 1A? Why not? But does it look roughly similar?
4. Now estimate the “text book” model from Table I of the paper. I.e., estimate by OLS the model

$$\ln GDP_c^{pc,1985} = \alpha + \beta_1 \ln s_c + \beta_2 \ln (n + g + \delta)_c + \epsilon_c$$

where  $\ln GDP_c^{pc,1985}$  is log of GDP per capita in 1985 in country  $c$ ,  $\ln s_c$  is the log of the savings rate ( $I/Y$ ) in country  $c$ , and  $(n + g + \delta)_c$  is the log of population growth rate + productivity growth rate + depreciation. Not in the footnotes to Table I, Mankiew et al say that they set  $n + \delta$  equal to 5% for all countries.

Run the regression by sample and report coefficient estimates and standard errors along with other usual statistics related to the estimation.

Are your answers exactly the same as in Table I? Are they roughly the same?

5. According to the augmented model in Makiew et al, human capital investment is correlated with physical capital investment, and they both positively affect GDP per capita in 1985. If this is the true model, in which way are the point estimates from the “text book” model biased?

Check that the education investment variable is correlated with the savings rate. Present results in a regression table.

Are Makiew et al right to be concerned about this omitted variables?

6. Now estimate the “full” model from Table 2 of the paper. I.e., estimate by OLS the model

$$\ln GDP_c^{pc,1985} = \alpha + \beta_1 \ln s_c + \beta_2 \ln (n + g + \delta)_c + \beta_3 \ln (school)_c + \epsilon_c$$

Run the regression by sample and report coefficient estimates and standard errors along with other usual statistics related to the estimation.

Are your answers exactly the same as in Table II? Are they roughly the same?

Do the point estimates move in the expected direction relative to Table I?

7. Could there be other omitted variables? Think of one and write it down. If you have time, go to the Penn World Tables

<https://www.rug.nl/ggdc/productivity/pwt/?lang=en>

Download the omitted variable you are worried about. Merge it to the dataset. Re-run the regression including your hypothesized omitted variable. Present the regression table. Do the point estimates change relative to the Mankiew et al paper? Should you send your results to a journal for publication?