EXERCISE SHEET 5: AJR (2001) ON THE COLONIAL ORIGINS OF GROWTH

In this and next week's exercises we will re-estimate some of the empirical work done by Acemoglu, Johnson & Robinson (2001) on the role of institutions for long term development.

Main learning goals: 1. Merging data by using loops
2. Once again, descriptive statistics
3. Running and interpreting simple OLS regressions
4. Plotting data

Hand-in: As usual: do-file and the corresponding log-file (saved in .log-format) containing the following steps (numbered in the do-file).

Moreover, upload a .pdf-file containing the graph (exercise 5) saved in under `familyname_exc5.pdf`.

1. The data has been provided in seven different pieces (zipped as maketable.zip), in which some of the variables are recurring, others are unique. Combine all the datasets into one common dataset using the `merge` command. Use a loop for merging. Inspect the resulting dataset and save it under the name `AJRdata.dta`.
2. Check the log settler mortality variable; is it entirely consistently derived from `extmort4`? Adjust it if necessary.
3. Reproduce the descriptive statistics given in Table 1 as close as you can, p. 1377.
4. Estimate the equation

$$\ln y_i = \mu + \alpha R_i + \boldsymbol{X}_i'\boldsymbol{\gamma} + \epsilon_i \tag{1}$$

   (p. 1378) using the command `regress`. Regress with and without heteroskedasticity robust standard errors, using the option `robust`.

   Reproduce columns (1), (2) and (6) and (8) of table 2 (p. 1379). Display all estimates in a simple table.

   Examine your results. 1. How do they compare to the ones by AJR (2001)? 2. Would you use conventional or robust standard errors to determine the significance of your coefficients?[1]

   Interpret your results statistically (in terms of their sign and significance). Comment on the results from column (6) in your do-file in detail.
5. Re-estimate columns 4 and 10 of panel A and columns 6 and 10 of panel B of table 3, p. 1385. Restrict your sample to the sample used in the paper.

   Why is table 3 and the intuition behind it important for the following IV approach? Explain briefly.
6. Reproduce Figure 2 from the paper as close as you can.

   Use the same style to plot the share of European settlers in total population in 1900 against settler mortality. Convert this latter figure into a .pdf-file and submit it along with the other output.

   Comment in your do-file shortly on what you see in this new figure.
7. Discuss shortly the following question in your do-file:

   What are the potential biases affecting the coefficient of interest in equation (1) and in which direction are they going to bias the coefficient? If now instruments are introduced to correct for these biases, is the coefficient estimated by IV expected to turn out smaller or larger?

Further comments: Submit 3 files (.do, .log, and .pdf(graph)) to Ilias. Usual procedures apply.

[1] The original data used in the paper contains 110 countries in the world sample. Our dataset contains 111 countries. Therefore the results will not be exactly the same as in the paper. However, all conclusions from the results will be the same.