Home assigment # 3

(To be submitted by Tuesday, December 14, 00:00 a.m.)

1. For this exercise, use the data in *fertil.gdt* file. It contains information about a sample of women in Botswana: their education (variable *educ*), their age (variable *age*), and the number of children they have (variable *children*).

   (a) Estimate the model
   $$children = \beta_0 + \beta_1 educ + \varepsilon \ ,$$
   interpret the coefficient β_1.

   (b) Conduct the Breusch-Pagan test for heteroskedasticity. Does its result justify the use of robust standard errors? If yes, reestimate the model using these robust standard errors and comment on the difference with respect to the model estimated by simple OLS.

   (c) Redefine the model as
   $$children = \beta_0 + \beta_1 educ + \beta_2 age + \beta_3 age^2 + \varepsilon$$
   and estimate it.

      i. Does age have a significant impact on number of children of Botswana women? State the hypothesis, and compute the test statistics by hand. Interpret the results and compare it to the results of the test in Gretl.

      ii. Do you find justification for the inclusion of *age* in quadratic form? [Hint: State the four specification criteria and argue if they are satisfied for the quadratic in age.]

      iii. How does the coefficient β_1 change when compared to part (a)? Does this signal any bias in the model from part (a)? Where does it come from? Explain the sign of this bias.

   (d) Do you think the coefficient β_1 from the model in part (c) may suffer from some omitted variable bias? What variable(s) could be missing in the model and how would the coefficient β_1 change if they were included in the regression? (What is the sign of the potential bias in coefficient β_1?)

   (e) Conduct the RESET test of the model from part (c) and interpret the results.

2. The aim of this exercise is to utilize Econometric tools learned in classes and to analyze published peer-reviewed papers. You do not need to read all papers thoroughly, rather you have to understand what the main idea of the paper(s) are, which conclusions these papers lay down, and which type of methodological caveats are present. More detailed instructions are below.

Read AER 2001 paper "The Colonial Origins of Comparative Development: An Empirical Investigation" (Acemoglu, Johnson, Robinson, 2001) and answer the following questions:

- What is the econometric model that the paper investigates? Carry out the corresponding equation.
- That is, what is the outcome/dependent variable and what is the main independent variable?
- Why is it difficult to estimate the model using a simple OLS method?
- How does the paper resolve the endogeneity problem? What are the assumptions that the proposed instrument should satisfy?
- How does the paper argues that the exclusion restriction is satisfied?

Now, read AER 2012 comment "The Colonial Origins of Comparative Development: An Empirical Investigation: Comment." What are the main argument(s) that Albouy (2012) develops to undermine the conclusions of Acemoglu, Johnson, Robinson (2001)?

Finally, read Acemoglu, Johnson, Robinson (2012) "The Colonial Origins of Comparative Development: An Empirical Investigation: Reply." How did the authors claim that their findings were robust to Albouy (2012) critique?

(Bonus) Develop your own argument. Are you convinced by the claims/critique provided by the author(s)? Try to explain your reasoning.