Third and final test in Introductory Econometrics

Robert M. Kunst

January 24, 2018

1. For a static and simple linear regression in a time-series context $y_t = \beta_0 + \beta_1 x_t + u_t$, five assumptions are often invoked in order to guarantee nice properties of the OLS estimator.[12 points]

   (a) Write down the five assumptions, and indicate which of them are not needed for the unbiasedness property of OLS.

   (b) If the additional regressor y_{t-1} is inserted, such that the regression reads $y_t = \beta_0 + \beta_1 x_t + \beta_2 y_{t-1} + u_t$, the regression becomes dynamic. Which of the assumptions listed in (a) cannot hold any more in this context, and which assumption is used instead?

   (c) Why is it more important in a dynamic regression than in a static regression that errors are uncorrelated over time?

2. Can we do it all by hand on a primitive computer using non-academic software? Imagine we only can do OLS (showing coefficient estimates, standard errors, and R^2) and some basic algebraic manipulation of variables. You consider and estimate a simple static linear regression in a time-series context $y_t = \beta_0 + \beta_1 x_t + u_t$.[12 points]

   (a) You would like to run a Breusch-Godfrey test for one lag only. You remember that you have to run an auxiliary regression. What variable is regressed on what explanatory variables in this auxiliary regression? How do you calculate the test statistic from this auxiliary regression?

   (b) The Breusch-Godfrey test rejects (assume you have a table of chi-square significance points), so you consider a GLS variant according to Cochrane-Orcutt, in which you apply OLS to transformed variables for $t = 2, \ldots, n$. You use the ρ estimate $\hat{\rho}$ from the Breusch-Godfrey test as an estimate of the error process $u_t = \rho u_{t-1} + \varepsilon_t$ with serially uncorrelated ε_t. How do you transform the variables?

3. Stop crime, get rid of the police? A linear regression of crime rates C_i on police-per-population P_i across 200 U.S. districts indeed yields a positive and significant coefficient on P. The empirical effect persists even in the presence of further regressors, such as an index of ethnic fractionalization F_i and an index of income inequality Q_i. [12 points]

   (a) The errors u are difficult to interpret. Maybe there is some tendency of some regions to create a fertile ground for crime. Can you name tentatively some of such unobserved effects in u? Do you think that these are correlated with P?

   (b) Why would you instinctively argue that reducing the police force will not decrease crime, even when the regression shows the reported effects?

   (c) Someone has suggested the number of public servants per population S as an instrument for P. Which two conditions should this instrument fulfill? Taking the idea into account that a region that tends to rely on public services will hire more teachers, more bus drivers, and also more police, do you think that they hold here? If they hold, will the IV estimator with S as an instrument for P, and F, Q assumed exogenous, be consistent?

   (d) It turns out that F and Q are positively correlated, as ethnically heterogeneous areas also tend to be socially unequal. Does this correlation make your IV estimator inconsistent?

4. For each of the following tests, write down the null hypotheses, the alternative hypotheses, and the (maybe asymptotic, if an exact distribution does not exist) distribution of the test statistic under the null. You can use words or symbols, but unusual notation must be explained. [9 points]:

   (a) The Breusch-Godfrey test;

   (b) The Breusch-Pagan test;

   (c) The F–test for the last two regression coefficients in a regression with 4 regressors.