Economics 423
Econometrics II   Summer Session II 2020

Computer Project #1 (out of 100 points)

Instructions:

-You must submit this to soriano@umd.edu by 11:59 PM by the due date specified on the most recent version of the syllabus.

-A proper submission includes a log file (or whatever the equivalent is in R) in a text or PDF format. I should not have to open Stata or R to see your code and output.

-For each line of code, include a briefcomment that explains what you are intending to do with that line. (For example: sum y x z *Summarize variables I am using). This is your best way to receive partial credit if you make a mistake. This is especially true for people using R, since I do not use R myself.

-Comment your answers to the questions at the beginning or the end of the log. Keep them all in one place so I do not need to search the output for each answer.

-At the beginning of the file, specify the one person that you worked with (if anyone).

-Please note: I reserve the right to ask anyone to stay in office hours to explain their answers if I have reason to believe their work in not their own.

Background and Purpose

The purpose of this project is to appreciate the pitfalls that may occur when estimating models using time series data. The paper by Whittington, Alm and Peters in the June 1990 AER, "Fertility and the Personal Exemption: Implicit Pronatalist Policy in the United States," argues that the US income tax system can affect the birth rate in the US by altering the personal income tax exemption. Using data from 1913 through 1984, the authors argue that the real value of the personal exemption was positively related the generalized fertility rate among women 15 to 44. And, the effect is large: they estimate that a $100 increase in child tax benefits would increase fertility by 3.2 to 6.5 percent. A comment on their paper by Crump, Goda and Mumford, published in the June 2011 AER, challenges their results.

Wooldridge uses the data from the Whittington et al. study to illustrate the pitfalls of simplistic estimation of time series models. You are asked below to estimate various models relating the generalized fertility rate (gfr) to the personal exemption (pe). These models illustrate issues with trending data and data that follow a random walk.

Tasks **(6 points each)**

1. Download the FERTIL3.DTA dataset from ELMS. Plot the gfr_t and pe_t using line graphs.
2. Reproduce equations (10.18) and (10.19) in the text.
3. Estimate the cumulative effect of pe_t, pe_{t-1} and pe_{t-2} and compute its standard error.
4. De-trend the variables in (10.19) using (a) linear; (b) quadratic and (c) cubic trends.
5. Regress grf_t on t and t^2 and save the residuals. Then regress the residuals on the variables on the RHS of equation (10.35)—pe_t, ww2, pill, t and t^2. Compare the R^2 with (10.35)

6. Now estimate the relationship using first differences (equation (11.27) in the text.
7. Estimate the cumulative effect of the pe in equation (11.27) and compute its standard error.
8. Add a time trend to (11.27).
9. Test for the presence of autocorrelation in (11.27).

Questions to Answer

A. What do the results in Task 4 suggest about the need to de-trend the variables? (7 points)
B. Why is it an even better idea to run the equation in first differences? (8 points)
C. Do you need to add a time trend to the equation when using first differences? Explain. (7 points)
D. What happens to R^2 when you run the equation in first differences? Why is this happening? Is it a problem? (8 points)
E. What do the results of Task 9 imply for the standard errors you have calculated? (8 points)
F. What do you conclude about the impact of the personal exemption on the generalized fertility rate based on the equations you have estimated? (8 points)

Some useful formulas:

To plot a line graph: graph twoway line <name of y variable> <name of x variable>

To save the residuals from an equation: predict <name of residual>, resid

To lag a time series variable: uhat_1 = uhat[_n-1]
<lagged variable name> = <name of variable to be lagged>[_n-1]