Empirical Exercises - 2
Analysis of US Private Consumption
Applied Statistics and Econometrics II
Spring 2018, NYU
Ercan Karadas

(Due April 12)

Motivation: As private consumption is the largest demand component it plays a crucial role in determining economic activity and employment. Consequently, what determines private consumption is an essential element in macroeconomic models. Macroeconomic theory suggests that private consumption is driven by disposable income and wealth. Using quarterly data covering the period 1973:1 - 2016:4, the goal of this assignment is to investigate empirically to what extent these factors have influenced private consumption.

The Assignment: Conduct an empirical analysis to determine which factors influence private consumption. First, carry out a graphical analysis to explore the joint evolution of private consumption, disposable income, and private wealth. Then, estimate a linear regression model that explains private consumption. Estimate your model with OLS and explain, interpret, and discuss your results.

You are only required to assess stationarity of the variables based on the graphs of the data and to use the linear regression model with contemporaneous explanatory variables. However, if you include the variable ECM_t it should be included with a lag.

Data: Download the dataset `consumptionUSA.csv` from the course webpage on NYU Home. The dataset contains observations for the variables:

CPRIV	Private sector aggregate consumption, constant prices.
CPRIVP	Deflator for private consumption, 1995=100.
DPRIV	Private disposable income, constant prices.
REALW	Private wealth including owner occupied housing, constant prices.
ELOSS	Expected income loss from changes in unemployment.
BNDR	Average bond rate, fractions, p.a.

All the variables, except the interest rate, are seasonally adjusted and are taken from FRED.

[1] Construct time series graphs of the variables and characterize the time series behavior of the variables.

[2] For the empirical modeling, construct the derived variables

$$c_t = \ln(\text{CPRIV}_t)$$
$$y_t = \ln(\text{DPRIV}_t)$$
$$w_t = \ln(\text{REALW}_t)$$
$$\pi_t = \Delta p_t = p_t - p_{t-1}$$
$$r_t = (1/4)\,\text{BNDR}_t$$

[3] Draw time series graphs of the transformed variables, c_t, y_t, w_t, π_t, and ELOS_t, and discuss the economic development in the US over the period.

From a graphical inspection, do any of the variables appear stationary? In the construction of time series regression models, why is it important that the included variables are stationary?

[4] To obtain stationary variables consider the following transformations:

$$\Delta c_t = c_t - c_{t-1}$$
$$\Delta y_t = y_t - y_{t-1}$$
$$\Delta w_t = w_t - w_{t-1}$$
$$\Delta \pi_t = \pi_t - \pi_{t-1}$$
$$\text{ECM}_t = c_t - 0.53 - 0.26 y_t - 0.48 w_t + 2.12 \pi_t.$$

Draw time series graphs of the transformed variables, $\Delta c_t, \Delta y_t, \Delta w_t, \Delta \pi_t$, and ECM_t. From a graphical inspection, do the transformed variables appear stationary?

*H*int: The variable ECM_t can be thought of as the deviation of consumption, c_t, from the equilibrium value in period t and it is formally constructed using cointegration. We will return to this issue later in the course, for now just think of ECM_t as an additional explanatory variable.

[5] When you construct an empirical model for the consumption data, do you think it is preferable to start with a simple model and then successively include more explanatory variables; or do you think it is preferable to begin with a general model and delete insignificant variables? Motivate your answer.

[6] Consider the following regression model

$$\Delta c_t = \beta_1 + \beta_2 \Delta y_t + \beta_3 \Delta w_t + \beta_4 \Delta \pi_t$$
$$+ \beta_5 \text{ECM}_t + \beta_6 \text{ECM}_{t-1} + \epsilon_t$$

Run the regression using OLS. Interpret the signs and magnitudes of the coefficients.

[7] Do the residuals look well behaved?

[8] Use R to calculate the LM test for no first order autocorrelation in the residuals. What is the distribution of the test statistic, and what do you conclude? Save the estimated residuals $\hat{\epsilon}_t$ from the original regression and construct manually the LM test.

*H*int: Remember to insert zero values for the residuals in the beginning of the sample and use the same estimation sample for the auxiliary regression that you used for the original regression. You should get exactly the same test statistic as the one reported by R.

[9] Calculate more misspecification tests using **Car** package. Discuss how the tests are constructed and how they should be interpreted. What do you conclude regarding the specification of the model?

[10] Now simplify the model by removing regressors with insignificant coefficients. Begin by deleting the variable with smallest t-ratio and continue until all coefficients are significant. Always retain the constant term. Recalculate the misspecification tests.

[11] How do you think that the interest rate should affect consumption?

[12] Analyze whether there is a role for the bond rate, r_t, or the real bond rate, $r_t - \pi_t$, in your empirical model.

Hints:

- For the graphical analysis, use any transformations of the variables you find are relevant (among the transformations proposed above, or any other ones), and explain your choices.

- Carefully discuss which assumptions are required to derive the estimator you use and explain if these assumptions are fulfilled or not empirically. If not, what are the consequences on your estimation results? Can you improve on this?

- If you come up with several model specifications, select the one that seems the most relevant to you, and justify your decision. Remember that in most cases, the simpler the better.

- Be precise about the statistical tests you use for testing various hypotheses. Explain which null hypotheses you test, how you test them, and what your conclusions are.

Formalities:

[1] You must hand in a report that (i) presents your graphical analysis, (ii) describes the econometric model, (iii) outlines the modeling progress (e.g., the approach you have taken, the alternative models you have tried, etc.), (iv) presents your preferred model including interpretation and statements on economic and statistical significance, and (v) discusses the potential weaknesses of the model.

[2] The report must be a maximum of 10 pages of pdf produced from an R markdown file. You must hand in your R markdown file and pdf files together. (If you don't have Latex installed you can't render a pdf. In that case, either first install Latex and try again or submit R markdown together with its HTML output).

[3] You can use the R markdown template that I have posted on course webpage. Please name your files as follows: `empEx2-surname(s).Rmd`

[4] If you prefer, you are allowed to work in groups of up to three persons (not necessarily in the same exercise class as yours). The requirements and assessment criteria are the same for assignments written by one, two, or three persons.