Financial Econometrics Rauli Susmel
FINA 4397

Practice Questions for Midterm 2

2.1 (**Heteroscedasticiy and Autocorrelation**) For this question, use the FX_USA_MX data set from my homepage (FX_USA_MX.csv).
FMX_da <- read.csv("http://www.bauer.uh.edu/rsusmel/4397/FX_USA_MX.csv", head=TRUE, sep=",")

You fit the following regression:

$$MX\_int_i = \beta_0 + \beta_1 \, US\_int_i + \beta_2 \, e_i + \beta_3 \, MX\_I + \beta_4 \, MX\_y_i + \varepsilon_i$$

a. Report the regression and interpret β_1.
b. Test for heteroscedasticity using the GQ tests and the studentized LM-BP (be precise and describe the variables you use to run the test).
c. If you find heteroscedasticity correct the SE using the appropriate HC SE, report the adjusted t-values. Does any coefficient loses significance?
d. DW test for autocorrelation.
e. Test for autocorrelation using the BG LM test, with 4 lags.
f. If you find autocorrelation, use the appropriate HAC SE and report the adjusted t-values. Does any coefficient loses significance?
g. Estimate the model with data from 1978.2 to 2018.4. Then, assuming that all your explanatory variables follow a Random Walk –i.e., the best predictor of next quarter's value is today's value-, forecast Mexican interest rates for the period 2019.1 to 2021.2. Report the MSE.
h. Using the same Random Walk assumption for the driving variables, forecast Mexican interest rates for 2021.3 (out-of-sample forecast).

2.2 (**Modeling Strategies**). Download the data Real_Estate_2020.csv from my homepage.
RE_da <- read.csv("http://www.bauer.uh.edu/rsusmel/4397/Real_Estate_2020.csv", head=TRUE, sep=",")

The file contains log changes in home prices for Los Angeles (LA), San Francisco (SF), and Las Vegas (LV), the notation for prices changes is XX_c, where XX is the city. The file also contains changes in unemployment for each city (notation: ZZ_u, where ZZ is the city), an index of economic conditions for each city (WW_EC, where WW is the city), changes in the leading economic indicators for different states: California, and Nevada (notation, Xind_c, where X is the first initial of the state), changes in the Federal Reserve Tech Indicator and the Fama-French 5 factors: Mkt_RF, SMB, HML, RMW and CMA. You have data from Feb 1990 to Sep 2019. You want to model the log changes in home prices for Las Vegas (LV). Real Estate agents say that there is more activity in the summer, thus, you consider dummy variables for Spring, Summer, and Fall. Since Las Vegas was seriously affected by the 2008 Financial Crisis you add a dummy variable for the financial crisis.

a. Starting from a General Unrestricted Model, using all the variables you can think of that make sense to include, select an appropriate model for Las Vegas (LV).
b. What are the driver of LAS home prices in your reduced (specific) model?

c. Did the 2008 Financial Crisis affect LAS prices? Do you have evidence of seasonality –i.e., are the dummy variables for Spring, Summer or Fall significant?
d. Check if the errors are normal (use a Jarque-Bera test).
e. Check that the model's errors do not show autocorrelation
f. e. Check that the model's errors do not show heteroscedasticity.
g. If they do show autocorrelation and/or heteroscedasticity, use proper SE to conduct tests of significance for the coefficients for the driver variables in the reduced model.
h. Check if there is monthly seasonality in the residuals. If there is re-specify the model.

2.3 (**ARMA Process**). You obtain the following estimates for an AR(1) model of some returns data

$$y_t = 0.1 + 0.73\ y_{t-1} + 0.30\ \ \varepsilon_{t-1}\ +\ \varepsilon_t$$

where ε_t is a white noise error process.
a. Check the estimated model for stationarity.
b. What is the expected patterns for the ACF and PACF? (You don't need to calculate them, just describe the pattern).
c. Calculate the first 3 ACF.

2.4 (**Significance of ACF/PACF and Identification**). Consider the following autocorrelation and partial autocorrelation coefficients estimated using 356 observations for a weakly stationary series, y_t :

Autocorrelations of series 'x_sea', by lag

0	1	2	3	4	5	6	7	8	9	10	11	12
1.000	-0.268	0.158	-0.044	-0.020	0.004	0.023	-0.069	0.053	-0.002	-0.017	-0.052	0.092

Partial autocorrelations of series 'sim_ar1', by lag

1	2	3	4	5	6	7	8	9	10	11	12
-0.268	0.093	0.022	-0.044	-0.011	0.034	-0.061	0.015	0.035	-0.020	-0.077	0.079

a. Use the Bartlett SE –i.e., 1/sqrt(*T*)-, to determine which, if any, of the ACF coefficients are significant at the 5% level.
b. Use the Ljung–Box statistics to test the joint null hypothesis that the first three autocorrelation coefficients are jointly zero.

2.5 (Theory Review) – True or False
a. If the data is heteroscedastic, we cannot use OLS.
b.White Standard errors can be used when the errors show autocorrelation and heteroscedasticity.
c. OLS is still unbiased if we use the wrong variance structure –i.e., wrong (**A3'**) assumption.
d. OLS and GLS produce the same estimates of β.
e. We can apply the Law of Large Numbers to time series data
f. Describe the use of the ACF and PACF to identify ARIMA models.
g. An AR(1) model is always stationary.
h. An MA(4) model is always stationary