Financial Econometrics Rauli Susmel
FINA 4397

Practice Final Exam

Download the monthly Fin.csv dataset (Fin.csv) from my homepage. You have CAT and CAT stock prices (CAT and CAT respectively), EUR exchange rates against the USD (USD_EUR), the U.S. 1-month risk-free rate (RF), Market Dividends (Dividends), Market Earnings (Earnings), and the Fama-French 5 factors, market excess returns (Mkt_RF), SMB, HML, RMW and CMA. The Fama-French factors and the risk free rate are in percentage terms. The data covers the period 1980:December – 2020:July.

You can read the data set with the following R line:

```
Fin_da <- read.csv("https://www.bauer.uh.edu/rsusmel/4397/Fin.csv", head=TRUE, sep=",")
```

1. Compute CAT log returns (lr_cat). Then, you model Caterpillar excess return (cat_x) as a function of 4 Fama-French factors: Excess market returns (Mkt_RF), SMB, HML and RMW.That is,

$$cat\_x_i = \beta_0 + \beta_1 \text{ Mkt\_RF}_i + \beta_2 \text{ SMB}_i + \beta_3 \text{ HML}_i + \beta_4 \text{ RMW}_i + \varepsilon_i \quad (*)$$

a. Report the regression.
b. Interpret the coefficients β_1 and β_3.
c. What are the drivers of CAT excess returns?
d. Interpret the R^2 and the F-goodness of fit test.
e. Using a t-test, test if the risk of CAT is similar to the market risk –i.e., H_0: $\beta_1 = 1$. State your conclusion from the result of the test.
f. Test H_0: $\beta_2 = \beta_4 = 0.5$ vs H_1: at least one β_2, and/or β_4, different from 0.5. State your conclusion from the test.
g. Using a QLR test, check it the model (*) suffers from a structural break.

2. Continuation from Question 1.
a. Using an LM test, test if Mkt_RF squared and SMB squared cause heteroscedasticity. State if you reject the null hypothesis of no heteroscedasticity.
b. Test for heteroscedasticity using the GQ tests and the studentized LM-BP. State if you reject the null hypothesis of no heteroscedasticity.
c. Test for autocorrelation using the Durbin-Watson test. State if you reject the null hypothesis of no autocorrelation.
d. Test for autocorrelation using a Box-Pierce test (a Q test), with 4 lags and 12 lags. State if you reject the null hypothesis of no autocorrelation.
e. Test for autocorrelation, using a Q test, allowing for heteroscedasticity and automatic selection of lag structure.
f. If you find autocorrelation and/or heteroscedasticity, use the appropriate HAC SE to calculate t-values. Report the new t-values. Does any coefficient lose significance? If so, re-specify model.

3. Continuation from Questions 2.
a. Estimate your model from 2.e using the estimation period January 1981 to December 2016. Report the regression.
b. Assume the explanatory variables follow a Random Walk, that is, the previous month observation is the best forecast for next month. With this assumption, generate forecasts for January 2017 till the end of your sample. Calculate the MSE.

c. Assume that CAT log returns follow a Random Walk. Compute the MSE of the Random Walk Model (RWM) for CAT.
d. Test if the MSEs from model (*) and the RWM are significantly different. State your conclusion from the test.

4. You have estimated the following ARMA(1,3) model for some time series y_t:

$$y_t = 0.56 + .46\, y_{t-1} + 0.5\, \varepsilon_{t-1} - 0.4\, \varepsilon_{t-2} - 0.2\, \varepsilon_{t-3} + \varepsilon_t$$

Suppose that you have data for time to $t-1$, i.e. you know that $y_{t-1} = 1.34$, $\hat{\varepsilon}_{t-1} = 0.17$, $\hat{\varepsilon}_{t-2} = -.95$, and $\hat{\varepsilon}_{t-3} = 0.77$
a. Obtain forecasts for the series y_t for times t, t+1, and t+2 using the estimated ARMA model.
b. Suppose the actual values for the series turned out to be −0.21, -0.42, 3.01 for t, t +1, t +2, calculate the (out-of-sample) mean squared error (MSE).
c. Now, you try to use a SES model to forecast the values for the series y_t for times t, t+1, and t+2 The estimated value of the smoothing constant, α, is 0.26, with the most recently available smoothed value, $S_{t-1} = 0.607$. Obtain forecasts for the series y_t for times t, t+1, and t+2 using this model. Compute the (out-of-sample) MSE.
d. Based on your answers in b and c, in this case, which model is a better forecasting model.

5. Now, you will analyze log changes in Earnings (lr_earn):
a. Report the ACF/PACF.
b. Based on your ACF/PACF, suggest an ARIMA model for lr_earn. Justify your answer.
c. Is your selected model stationary? Justify your answer.
d. Check the residuals. Do you find evidence of additional autocorrelation in the residuals?
e. Do you have monthly seasonality? Regress the monthly stock returns, lr_earn, against monthly dummies and state if you find joint significance for the seasonal dummies.

6. True of False (Provide a very brief statement justifying your answer. No justification, no points.)
a. Outliers can bias a regression and cause heteroscedasticity.
b. White SE are only appropriate in the presence of heteroscedasticity.
c. A normal distribution is needed to estimate a model using Maximum Likelihood.
d. A bootstrap Confidence Interval does not rely on the normality assumption.
e. If a series is 2nd-order stationary the third and fourth moments may not be constant.
f. An ARMA(1,2) model is always stationary.
g. A Random Walk has an undefined (infinite) variance.