Financial Econometrics Rauli Susmel
FINA 4397/7354-A

Second Midterm

Note: Please, follow the instructions received in the email you received this file.

1. (**30 points. Modeling Strategies**). Download the data Real_Estate_2022.csv from my homepage, which can be reached using the following R command.

RE_da <- read.csv("https://www.bauer.uh.edu/rsusmel/4397/Real_Estate_2022.csv", head=TRUE, sep=",")

The file contains log changes in home prices for San Francisco (SF_c), and Los Angeles (LA_c). The file also contains changes in unemployment for each city (SF_u, and LA_u), log changes in the leading economic indicators Index for California (Cind_c), changes in a Tech Index (Tech_c), log changes in oil prices (Oil), log changes in gold prices (Gold), and the Fama-French factors: Mkt_RF, SMB, HML, RMW and CMA, along the risk-free rate, RF. You have data from Jan 1990 to Aug 2022.

The last column of the file contains simulated data, Simu_1, which you will use to answer question 3.

You want to model the log changes in home prices for **San Francisco**, p_t. Real Estate agents say that there is more activity in the summer, thus, you consider dummy variables for Spring, Summer, and Fall. You also include a 2008 Financial Crisis dummy, which you date as starting in August 2008 (observation **224**).

a. Starting from a General Unrestricted Model (GUM), using all the variables you can think of that make sense to include, select an appropriate (reduced) model for SF, (SF_c).
a. Report GUM regression.
b. Pick a p-value to reduce the GUM. Report reduced regression.
c. Report and interpret the R^2.
d. What are the drivers of the your reduced model?
e. Do you have evidence of seasonality –i.e., are the dummy variables for Spring, Summer or Fall significant?
f. Check if the errors of your reduced model are normal (use a Jarque-Bera test).
g. Check that the model's errors do not show autocorrelation.
h. Check that the model's errors do not show heteroscedasticity.
i. 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.

2. (**25 points. Forecasting**). You want to forecast log changes in home prices, p_t,using an AR(1) model.

a. Estimate an AR(1) model for p_t. using data from Jan 1990 to Dec 2020 (estimation period). That is, you estimate the following AR(1) model:

$$p_t = \mu + \phi_1 p_{t-1} + \varepsilon_t. \qquad \varepsilon_t. \sim \text{WN}.$$

Report the regression.

b. Using your AR(1) estimates from 2.a, forecast p_t from Jan 2021 to August 2022 (validation period). Compute the MSE of your AR(1) forecast.

c. Using a the random walk model (RW) for p_t, forecast p_t from Jan 2021 to August 2022. Compute the MSE of your RW forecast.

d. Test the equality of MSEs using an MGN/HLN test. Interpret the test results.

3. (**15 points. ARMA identification**). For this question you will use the simulated series (Simu_1) for parts 3.a to 3.d. For part 3.e (the last part), you will use log changes in oil prices (oil):

4. True of False (**20 points**). Briefly justify all your answers.

a. GLS estimates are more efficient than OLS estimates.

b. On average, FGLS estimates should be similar to the OLS estimates.

c. Returns are stationary when the correlations decay over time.

d. A Random Walk model without a drift has constant mean, but exploding variance.

e. If one of the roots of an AR(p) is equal to one and the rest are greater than one, the process is stationary.