

# Project Task Forecasting, BAN430

Johan Lyhagen

2021

Hand-out date: 2021-03-18, 09.00

Hand-in date: 2021-04-09, 14:00

## 1 Instructions

**Start to read the exam question from the beginning to the end!** First, note that you are allowed to make this exam in groups of maximum of two! Note also that you should hand this in using WISEflow.

You should hand in:

1. A written report explaining what you do, showing your results and interpreting them. Note that the report should include an introduction, conclusion and references as well as the main body of the report. Write with your fellow students in mind. The purpose of this is that when you work you will write reports all the time and/or you will read reports by others. If you are curious about the structure of the report take a look at a scientific paper within your field of interest, or a forecasting journal like Journal of Forecasting or International Journal of Forecasting. *I should understand not only what you do but also why!*
2. An R-program, with comments, such that I can reproduce your results.
3. A data file in Rdata format with the data you used.
4. You may write the report in English or Norwegian but I would appreciate English as my Norwegian is rather basic.
5. As usual, all tables and figures in the report should be relevant and commented upon. Some additional less relevant might be put in an Appendix.
6. The maximum length of the report is 20 pages, excluding a cover page if you intend to use one. To this you may add the appendix. Adding irrelevant things will be negative.

Some additional information:

1. When we say forecasts below, we are interested in forecasts of the level of the variable.
2. For the forecasts below, we also want the relevant forecast accuracy measures and plots!
3. The forecast horizon should be at least 8 periods.
4. When we want you to do a forecast it is for the seasonally unadjusted series we refer to, unless otherwise stated.

The tasks are the following:

- Download a seasonally unadjusted macro economic time series that has seasonality for a country of your choice. Also download the seasonally adjusted version of the same variable. There are many sources like the national statistics bureaus and OECD. Another good source is the Federal Reserve Bank of St. Louis.
  - Make sure that the variable of interest has a clear seasonal component and trend/cycle component.
  - For later use, download two additional economic variables.
- Divide the variables into a training set and a test set.
- Make appropriate summary statistics and plots for an initial analysis of training set of the time series.
- The first major task is to try to replicate the seasonal adjustment. How well can you replicate the seasonal adjusted series?
- Decompose the series into its components, i.e. the season, trend, etc. Answer interesting questions like: Can you identify business cycles? How has the seasonality developed through time?
- Forecast the individual components and use them to form forecasts of the test set.
- Select an appropriate ETS model and use it to forecast the test set.
- Identify an appropriate  $ARIMA(p,d,q)(P,D,Q)$  model.
- Make forecasts using the ARIMA model.
- Here you are going to add the two additional variables to make an appropriate multivariate model.
- Use the multivariate model to forecast.
- Make a dynamic regression model with the two additional variables as explanatory variables. Forecast the two additional variables and then use the dynamic regression model to forecast.

- Instead of stochastic trends one can use deterministic trends, splines and Fourier terms. Model the variable using deterministic terms. How do you treat the error? Make forecast.
- A very simple way to combine forecasts is just to take the average. Compare the average with another more advanced way to combine forecasts. How well do the forecast averaging methods perform compared to the previous methods?
- Here you are going to compare the effect of estimating a univariate model with including an irrelevant variable by performing a Monte Carlo simulation.
  1. Generate a variable  $y_t$  according to the estimated ARIMA model above.
  2. Generate another variable as  $x_t = 0.5x_{t-1} + 0.5y_{t-1} + \varepsilon_t$
  3. Divide into test and train set
  4. Forecast  $y_t$  by estimate an ARIMA on  $y_t$ .
  5. Forecast  $y_t$  by estimate an VAR on  $y_t, x_t$ .
  6. Repeat this a sufficiently many times and for some sample sizes
  7. What conclusion can you draw?

**Good Luck!**