Exercise for Monday March 6 2023

- By group of 2 people.
- Read Chapter 5 in Diebold. It is the presence of deterministic trends and seasonality in time series.

1 Part 1

The website (in one line)

`https://www.usgs.gov/centers/nmic/historical-statistics-mineral-and-material-commodities-united-states#lead`

contains annual commodity prices historical production series.

Consider one variable. Take production or world production series probably.

1) Look at the series and estimate the 4 following models

$$\begin{aligned}
(a) &: y_t = \alpha + \beta trend + \epsilon_t \\
(b) &: y_t = \alpha + \beta trend + \delta trend^2 + \epsilon_t \\
(c) &: \ln(y_t) = \alpha + \beta trend + \epsilon_t \\
(d) &: \ln(y_t) = \alpha + \beta trend + \delta trend^2 + \epsilon_t
\end{aligned}$$

hint: the command for trend = 1, 2, 3...is simply to add

`@trend`

as a usual regressor. $trend^2$ is obtained with `@trend^2`

2) Based on the graphs (residuals, fitted and realized) as well as on goodness of fit criteria (i.e. R^2, different information criteria), misspecification tests, which model would you overall favor?

3) Reestimate each of the 4 models without the the last 10 observations that you use for forecasting. Compare RMSEs on the last 10 forecasts. Is the best in-sample model (fit on the whole sample) you choosed in 1) also the best out-of-sample model (RMSE).

4) Take your favorite specification and reestimate it until the last available observation (come back to 1) say) . Forecast your series until 2050. In the workfile window, go to **Proc/structure and resize the current page** until 2050. You do not need to enter values for the explanatory variables, EViews knows what the values of the intercept and the trend are in the future.

2 Part 2: On your own data

Find data you are interested in. Download at least three time series that are probably related given your a priory knowledge of economic theory/business relationships. They can be variables from the consumption function, the money demand, indicators of the economic activity for different countries, GDPs, unemployment, number of cars, prices, business data or a company...Take the country you wish, you can also take the same indicator for comparing different countries (e.g. GDP in EU, US, Japan).

You can look for instance at

- Eurostat database http://ec.europa.eu/eurostat/data/database
- OECD database at http://stats.oecd.org/
- IMF at http://www.imf.org/en/Data
- For the US, a very good source is the Federal reserve bank of St. Louis at https://fred.stlouisfed.org/ . It is also directly accessible in EViews (open database).
- Series from any statistical office (BEA, Statistics Canada,...) or central bank (Bundesbank,...).
- others.

Hints and tips:

- Take monthly (or weekly) or quarterly time series such that you have at least 120 say observations. More points is better, hence do not delete series if you have more information.

- The series should be available for the same period.
- One series should be the series you investigate in the univariate model and for which you will introduce explanatory variables in multivariate models next. You can change during the course of course.
- For this course, don't take annual data like the commodities above (too few points) or intra daily data (too volatile).
- Do not take financial time series such as stock prices for your main dependent variable. There will be no dynamics, probably. It can be one of the explanatory variables.
- Take the levels of the data and not the transformed series (e.g. %). Interest rates (%) or unemployment rates (%) are fine of course. But do not take the economic growth or inflation: take instead the level of the real gross domestic product and the price index.
- Take data in real/constant terms (deflated by prices) and not nominal series for investment, consumption, money stock, output,...
- You may choose seasonally adjusted data if you want to avoid the mess of seasonal peaks...You can also do your own seasonal adjustment in EViews.

Your tasks:

1) Test for the presence of a deterministic or a quadratic trend in the level or the log level (choose among models 1 to 4). Put graphs in your output, it will be easier to comment it.

2) Introduce seasonal deterministic dummies and test whether they are significantly different from zero.

3) Test for the null of no autocorrelation using LM and BP tests.

4) Add lags of the dependent variable if you reject the null (and until you reject the null of no autocorrelation).

5) When you are happy with your favorite model, test for other signs of mispecification (heteroskedasticity, non normality, non linearity, etc).

6) Reestimate your model without the last 12 observations and report the RMSE.