Forecasting & Predictive Analytics

ESSEC MSc in Management
ESSEC-CentraleSupélec MSc in Data Sciences & Business Analytics
Guillaume Chevillon

EXAMINATION
March 29, 2017
4.30pm-7.00pm

Calculators are authorized
Please answer the four exercises below.
You can always assume one question answered and proceed to the next

Univariate Time Series

1. Answer the following questions.

   (a) Precisely provide the definition of covariance (or weak) stationarity. Why is stationarity a useful property?

   (b) In which of the following models are the $\{y_t\}$ stationary, assuming $\{\epsilon_t\}$ is a mean-zero white noise process? If the answer depends on extra conditions, explain the conditions required. In all cases, explain your answers.

      i. $\Delta y_t = -0.2y_{t-1} + \epsilon_t$;

      ii. $y_t = \phi_0 + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \epsilon_t + \theta\epsilon_{t-1}$;

      iii. $y_t = \phi_0 + 0.1x_{t-1} + \epsilon_t, \quad x_t = x_{t-1} + \epsilon_t$;

      iv. $y_t = 0.8y_{t-1} + \epsilon_t$;

      v. $y_t = y_{t-1} + \epsilon_t - \epsilon_{t-1}$.

      vi. $y_t - y_{t-2} = \epsilon_t$

2. Consider the $\mathsf{AR}(2)$ process defined as

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \epsilon_t$$

   (a) Rewrite the model with Δy_t on the left-hand side and y_{t-1} and Δy_{t-1} on the right-hand side.

   (b) What restrictions are needed on ϕ_1 and ϕ_2 for this model to collapse to an $\mathsf{AR}(1)$ in the first differences Δy_t?

   (c) When the model collapses to an AR(1) for Δy_t, what does this tell you about the properties of y_t?

   (d) Explain and discuss the Augmented Dickey-Fuller test, including the null and alternative hypotheses, the methodology and properties.

   (e) We consider forecasting y_{t+h} using information up to time t.

      i. Derive an explicit expression for the 1-step and 2-step ahead forecast errors $e_{t+1|t}$ and $e_{t+2|t}$ for y_t.

      ii. What is the autocorrelation function of a time-series of forecast errors, the processes $\{e_{t+1|t}\}$ and $\{e_{t+2|t}\}$ observed for $t \geq 1$? (in words or graphically, no need to prove it mathematically).

Multivariate models

3. Consider the VAR(1) (remember that at all times, you can assume that you have proved previous questions)

$$\begin{bmatrix} x_t \\ y_t \end{bmatrix} = \begin{bmatrix} 0.4 & 0.3 \\ 0.8 & 0.6 \end{bmatrix} \begin{bmatrix} x_{t-1} \\ y_{t-1} \end{bmatrix} + \begin{bmatrix} \epsilon_t \\ \eta_t \end{bmatrix}$$

where

$$\begin{bmatrix} \epsilon_t \\ \eta_t \end{bmatrix} \overset{i.i.d}{\sim} \mathsf{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_\epsilon^2 & \sigma_{\epsilon\eta} \\ \sigma_{\epsilon\eta} & \sigma_\eta^2 \end{bmatrix} \right)$$

(a) Verify that the vector process $(x_t, y_t)'$ is non-stationary (hint: one eigenvalue is trivial)

(b) Find the two values of β such that $x_t - \beta y_t$ follows an AR(1). What are these two values, and comment on the properties of the resulting AR(1) processes (careful, one of them is trivial).

(c) Define the concept of cointegration. Do x_t and y_t cointegrate here?

(d) What assumption do you need to make if you want to compute the impulse response function

$$\frac{\partial x_{t+k}}{\partial \epsilon_t}$$

as a function of k? (do not compute the function).

EMPIRICAL ANALYSIS

4. We study the empirical properties of the sales and advertizing dataset we studied in the course. We denote by a prefix L, the logarithm and by D, the difference, so DLSales is the first difference in the logarithm of Sales. Figures 1, 2 and 3 present the data. Figure 4 reports the autocorrelation and partial autocorrelation functions of the logarithm of Advertizing and Sales and their first differences.

(a) Comment on the stationarity of the time series presented Figures 1 to 4.

(b) We want to model LSales and LAdvert via ARIMA(p, d, q) models. What values do you suggest for either variable based on Figures 1 to 4.

(c) We estimate an AR(2) model for DLAdvert (standard errors in parentheses below the parameter estimates).

$$\text{DLAdvert}_t = \underset{(0.0328)}{0.0015} + \underset{(0.138)}{0.00491} \text{ DLAdvert}_{t-1} - \underset{(0.136)}{0.287} \text{ DLAdvert}_{t-2} + \epsilon_t$$

Comment on the quality of this equation: would you simplify the model by reducing the number of parameters?

(d) We now add a dummy variable for 1936, D1936,

$$\begin{aligned} \text{DLAdvert}_t = \; & \underset{(0.0275)}{0.0229} - \underset{(0.123)}{0.216} \text{ DLAdvert}_{t-1} \\ & - \underset{(0.112)}{0.278} \text{ DLAdvert}_{t-2} - \underset{(0.211)}{1.02} \text{ D1936}_t + \epsilon_t \end{aligned}$$

Comment on the role of the dummy variable. (The model results are drawn on Figure 5).

(e) Imagine you are in 1936 and want to use the models above to forecast LAdvert in 1937 and 1938. Propose a method for doing so making all the assumptions you may require. You may use that the values of LAdvert for 1931-1936 are

Year	LAdvert
1931	6.89
1932	6.95
1933	7.28
1934	7.31
1935	6.69
1936	5.82

(f) We now propose to use the following model for forecasting two years ahead,

$$\begin{aligned}\text{D2LAdvert}_t \;=\;& \underset{(0.119)}{0.597}\ \text{D2LAdvert}_{t-1} - \underset{(0.117)}{0.624}\ \text{D2LAdvert}_{t-2} \\ &+ \underset{(0.114)}{0.375}\ \text{D2LAdvert}_{t-3} + \underset{(0.0313)}{0.0223} - \underset{(0.231)}{1.08}\ \text{dumm1936}_t\end{aligned}$$

where $\text{D2LAdvert}_t = \text{LAdvert}_t - \text{LAdvert}_{t-2}$. (The model results are drawn on Figure 6). What forecasts does this new model generate?

(g) Discuss the difference and relative benefits from Iterated and Direct multistep forecasting methods (both from a theoretical and empirical point of view).

(h) We now estimate a VAR(1) model as follows:

$$\begin{aligned}\text{DLAdvert}_t \;=\;& \underset{(0.0272)}{0.0193} - \underset{(0.152)}{0.466}\ \text{DLAdvert}_{t-1} \\ &+ \underset{(0.274)}{0.746}\ \text{DLSales}_{t-1} - \underset{(0.209)}{1.06}\ \text{D1936}_t + \epsilon_t \\ \text{DLSales}_t \;=\;& \underset{(0.0162)}{0.00975} - \underset{(0.0909)}{0.0523}\ \text{DLAdvert}_{t-1} \\ &+ \underset{(0.164)}{0.364}\ \text{DLSales}_{t-1} - \underset{(0.125)}{0.306}\ \text{D1936}_t + \eta_t\end{aligned}$$

Comment on the Granger Causality properties of the model by carefully defining the concept.

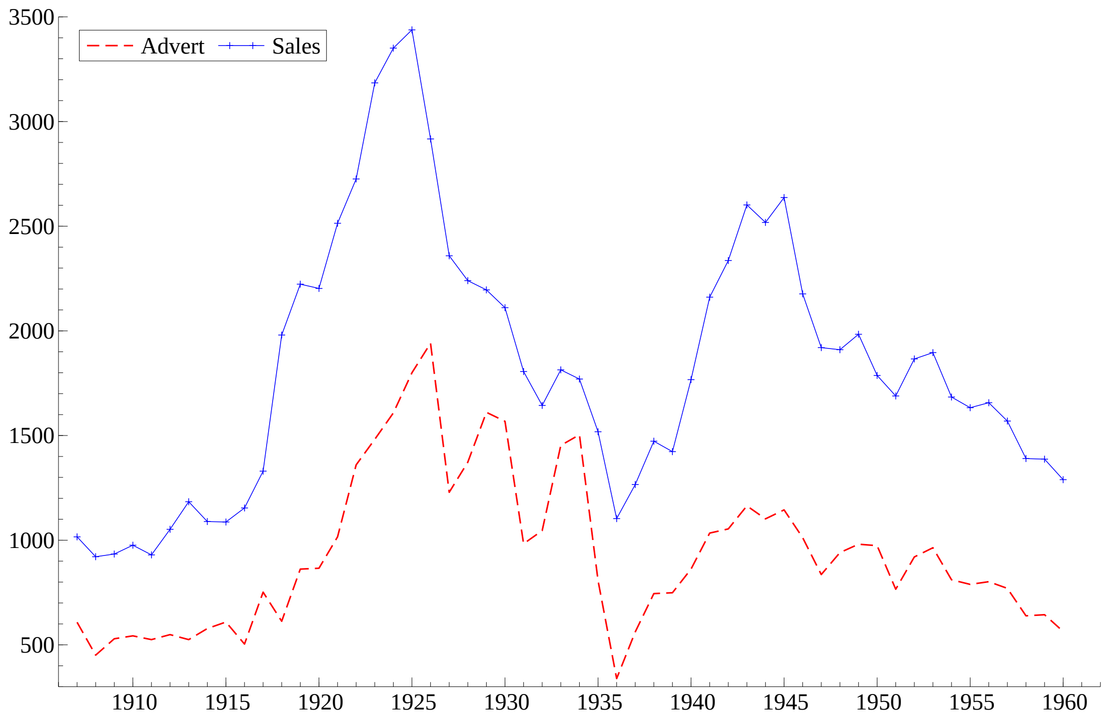

Figure 1:

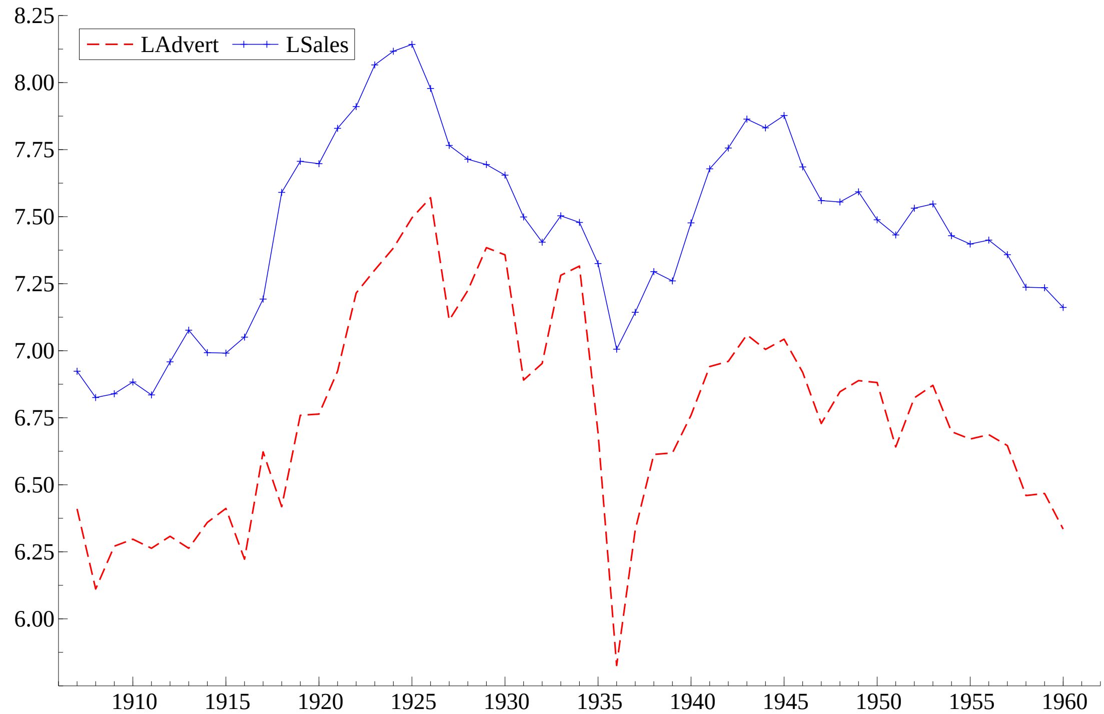

Figure 2:

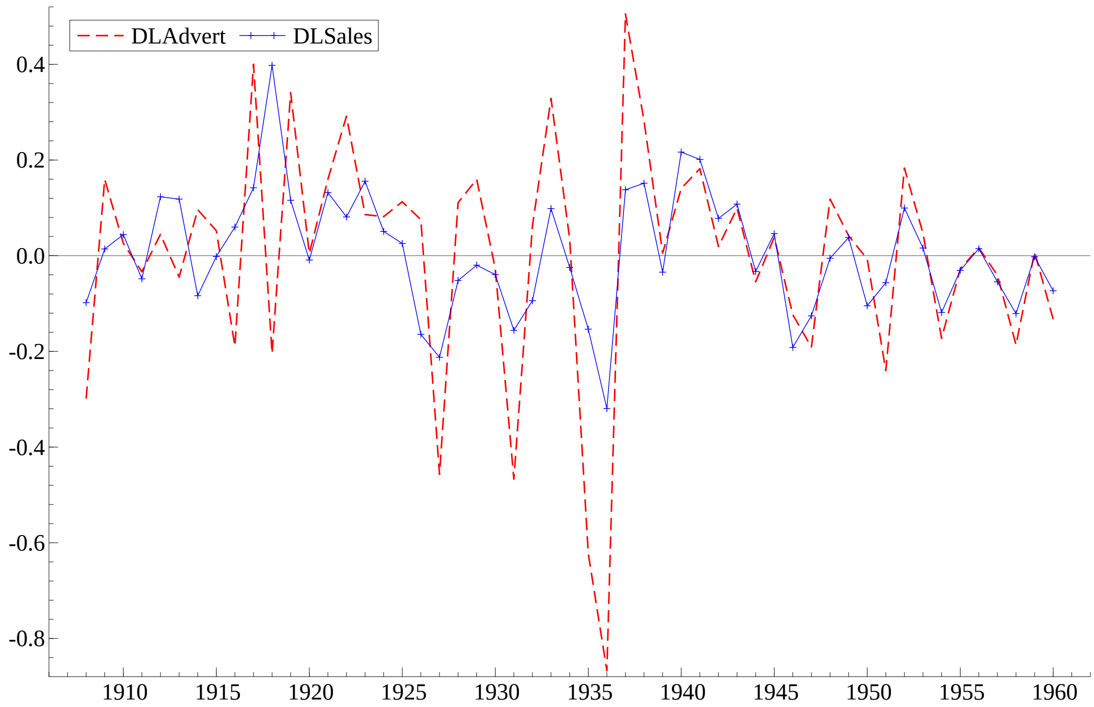

Figure 3:

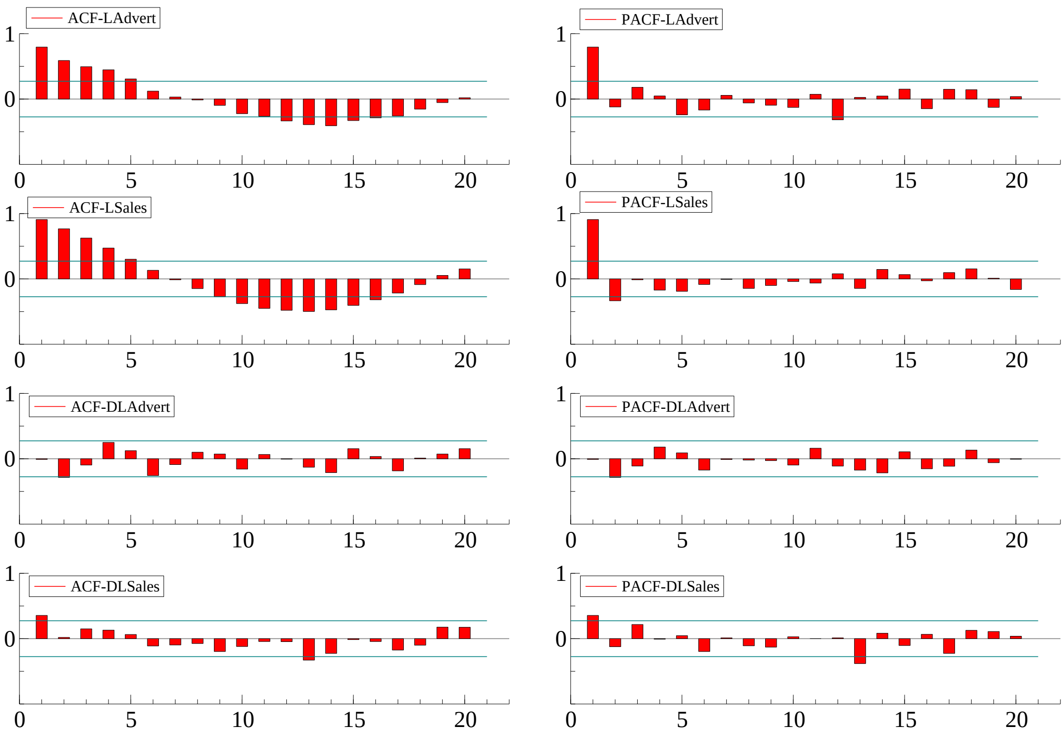

Figure 4:

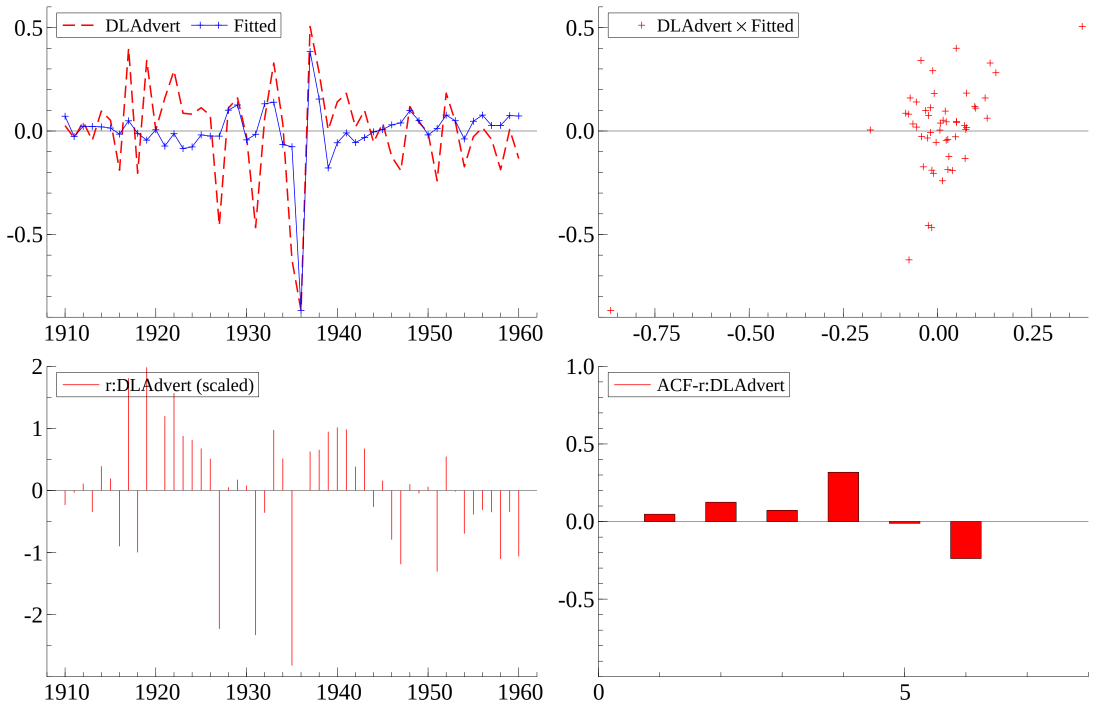

Figure 5:

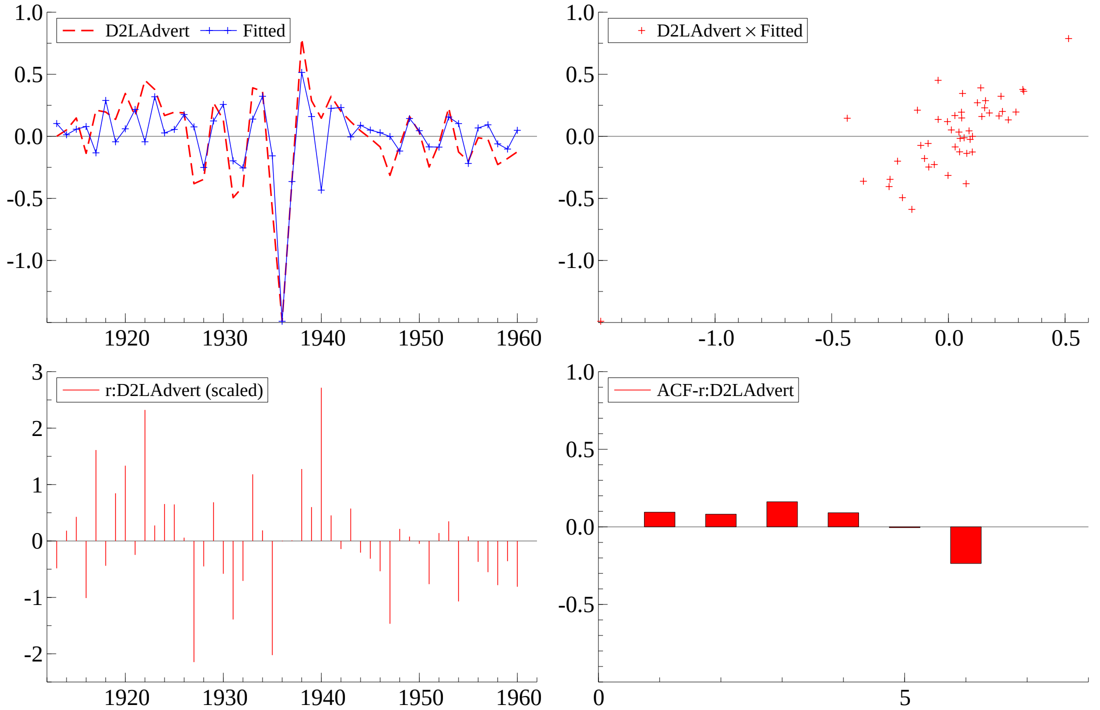

Figure 6: