STOCKHOLM UNIVERSITY
Department of Statistics
Matias Quiroz

FINAL EXAM, ECONOMETRICS II
2023-05-31

Time for examination: 08:00-13:00

Allowed tools: Pocket calculator, own formula sheet (1 double-sided A4 page), Course textbooks (any edition): Wooldridge, J.M. *Introductory Econometrics - a Modern Approach* and Montgomery, Jennings and Kulachi - *Introduction to Time Series and Forecasting*.

The exam consists of 6 problems. **The problems are not sorted by their degree of difficulty**. Write well motivated worked solutions, preferably on a single side of the paper. State any necessary assumptions or conditions where needed. Answers may be given in English or Swedish.

Passing rate: 50% of the total 100 marks. See the course description for a detailed grading criteria.

`Good luck!`

Problem 1. (15 marks)

Without motivation, indicate which single alternative is correct for each of the following sub-questions. Answering more than one alternative in a sub-question results in 0 marks for the sub-question.

(a.) The time series model

$$y_t = 10 + 0.7y_{t-1} + 0.2y_{t-2} + \epsilon_t,$$

where ϵ_t is iid with $\epsilon_t \sim \mathcal{N}(0, \sigma^2)$:

(i.) has $\mathrm{E}(y_t) = 10$.

(ii.) is covariance stationary.

(iii.) has $\mathrm{V}(y_t) = \sigma^2(1 + 0.7^2 + 0.2^2)$.

(iv.) is non-stationary.

(b.) Suppose that the true model is

$$y_t = 0.4x_t + z_t + u_t,$$

where x_t and z_t are weakly dependent and stationary processes, and u_t is a white noise process with $\mathrm{E}(u_t) = 0$. Assume that

$$z_t = 0.95z_{t-1} + v_t,$$

where v_t is white noise with $\mathrm{E}(v_t) = 0$. Suppose that we regress y_t against x_t and obtain the fitted values

$$\hat{y}_t = \hat{\beta}x_t, \tag{1}$$

where $\hat{\beta}$ is obtained via OLS. Then:

(i.) The proposed model underlying (1), i.e.

$$y_t = \beta x_t + \eta_t,$$

where η_t is a white noise process with $\mathrm{E}(\eta_t) = 0$, is dynamically complete.

(ii.) We can expect that the residuals obtained from (1) are independent.

(iii.) We can expect that the residuals obtained from (1) have a unit root.

(iv.) We can expect that the residuals obtained from (1) are serially correlated.

(c.) Which of the following four is the most suitable model for the time series in Figure 1:

(i.) $y_t = \beta_0 + \beta_1 t + \phi y_{t-11} + \varepsilon_t$, with $|\phi| < 1$ and ε_t iid with $\mathrm{E}(\varepsilon_t) = 0$.

(ii.) $y_t = L_t S_t + \varepsilon_t$, where L_t is the linear trend component and S_t is the seasonal adjustment with

$$S_t = S_{t+6} = S_{t+12} = S_{t+18} = \ldots,$$

and ε_t iid with $\mathrm{E}(\varepsilon_t) = 0$.

(iii.) $y_t = L_t + S_t + \varepsilon_t$, where L_t is the linear trend component and S_t is the seasonal adjustment with

$$S_t = S_{t+12} = S_{t+24} = S_{t+36} = \dots,$$

and ε_t iid with $\mathrm{E}(\varepsilon_t) = 0$.

(iv.) $y_t = L_t S_t + \varepsilon_t$, where L_t is the linear trend component and S_t is the seasonal adjustment with

$$S_t = S_{t+12} = S_{t+24} = S_{t+36} = \dots,$$

and ε_t iid with $\mathrm{E}(\varepsilon_t) = 0$.

(d.) Suppose that we have two periods of data, $t = 1, 2$, and N individuals, $i = 1, \dots, N$. Consider the fixed effects (unobserved effects) model

$$y_{it} = \beta x_{it} + a_i + u_{it}, \tag{2}$$

where u_{it} is the time-varying error and a_i is the fixed effect. What is **not true** about this fixed effects model:

(i.) We can take the first differences to obtain

$$\Delta y_i = \beta \Delta x_i + \Delta u_i, \tag{3}$$

with $\Delta y_i = y_{i2} - y_{i1}$, $\Delta x_i = x_{i2} - x_{i1}$, and $\Delta u_i = u_{i2} - u_{i1}$.

(ii.) To estimate β by OLS using the approach in (i.), i.e. using (3), it is required that x_{it} is constant across time for each individual i.

(iii.) To estimate β by OLS using the approach in (i.), i.e. using (3), strict exogeneity needs to hold for u_{it} in (2), which means that $\mathrm{E}(u_{it}|x_{i1}, x_{i2}) = 0$ for $t = 1, 2$.

(iv.) The parameter a_i in (2) captures unobserved effects for individual i that are constant across time.

(e.) Which of the following is the main problem with serially correlated errors u_t in the time series regression

$$y_t = \alpha + \beta x_t + u_t.$$

(i.) They necessarily violate contemporaneous exogeneity, i.e. $\mathrm{E}(u_t|x_t) \neq 0$.

(ii.) The OLS estimator of β is inconsistent, even if $\mathrm{E}(u_t|x_t) = 0$.

(iii.) Testing the hypothesis $H_0 : \beta = 0$ vs $H_1 : \beta \neq 0$ using the sampling distribution of the OLS estimator may give misleading results.

(iv.) $\mathrm{Cov}(u_t, \alpha) \neq 0$, which causes inconsistent estimation of α.

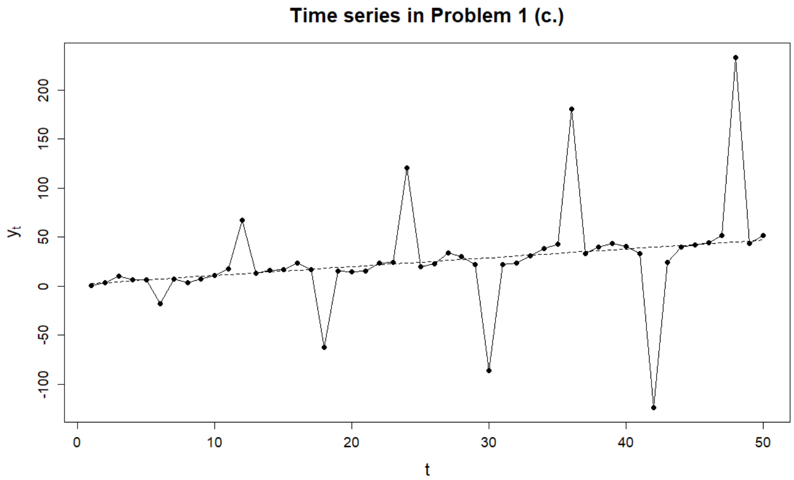

Figure 1: Time series data in Problem 1 (c.).

Problem 2. (15 marks)
Consider the following three time series models:

- Model 1:

$$y_t = 5 + \epsilon_t - 0.7\epsilon_{t-1} + 0.3\epsilon_{t-2}, \tag{4}$$

  with $\epsilon_t \sim \mathcal{N}(0, 1)$, and independent for all t.

- Model 2:

$$y_t = 5 + \epsilon_t + 0.9\epsilon_{t-1} + 0.5\epsilon_{t-2}, \tag{5}$$

  with $\epsilon_t \sim \mathcal{N}(0, 1)$, and independent for all t.

- Model 3:

$$y_t = 0.35 + 0.93y_{t-1} + \epsilon_t, \tag{6}$$

  with $\epsilon_t \sim \mathcal{N}(0, 1)$, and independent for all t.

For each of the models above, T = 10,000 samples are simulated and the autocorrelation function $\rho(k)$, $k = 1, \ldots, 12$, for each model is estimated. Figure 2 shows the estimated autocorrelation functions (B1, B2, B3) and snapshots of the first $t = 1, \ldots, 50$ observations for each time series (A1, A2, A3). Note that the figures in each panel are in random order, i.e. a time series in a given row does not necessarily correspond to the autocorrelation function in the same row.

Pair each of the three models, i.e. (4), (5), and (6), with one realisation (A1, A2, A3) and one autocorrelation function (B1, B2, B3). Clearly motivate your answers. Providing a correct pairing for a model without motivation results in 0 marks for the model.

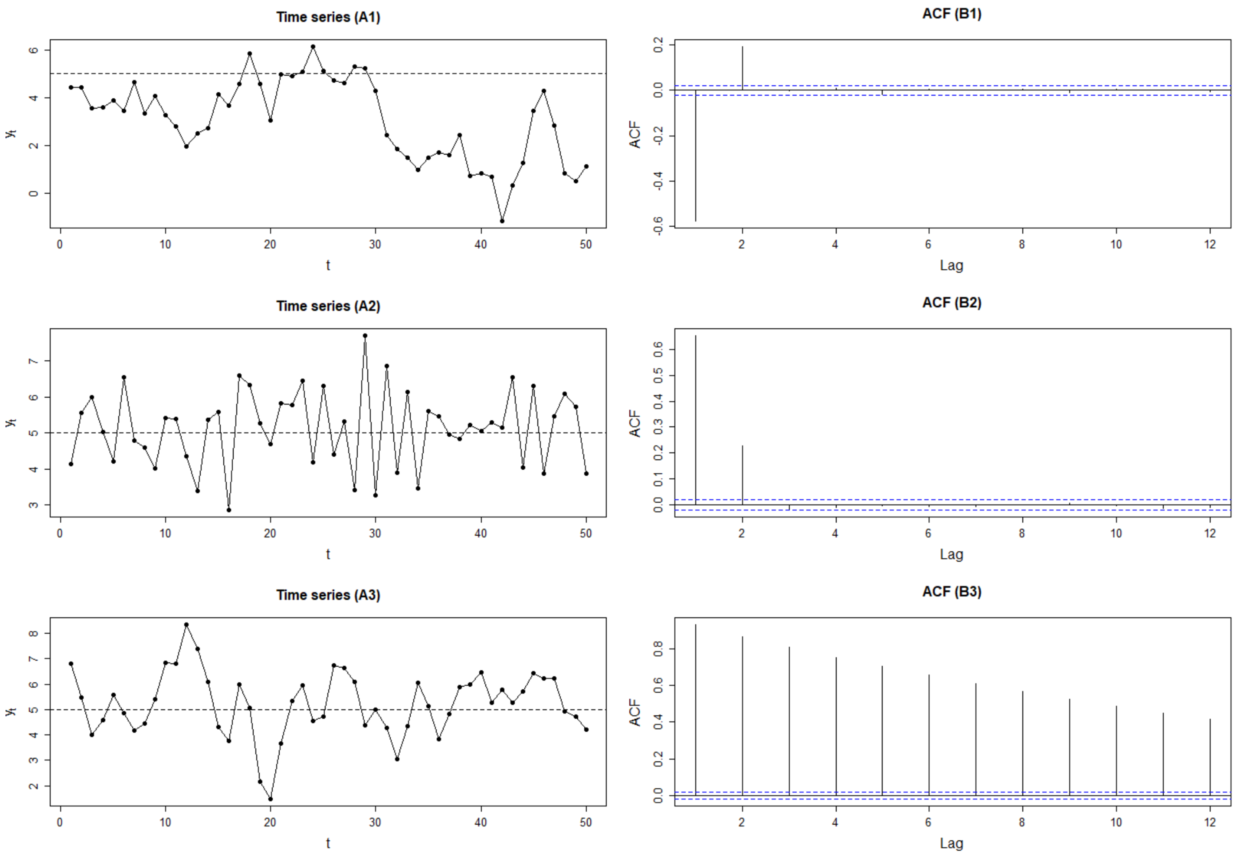

Figure 2: Three time series y_t, $t = 1, \dots, 50$, (left panel) and three estimated autocorrelation functions (right panel) in Problem 2. The dashed horisontal lines in the figures on the left panel correspond to $E(y_t)$. The two dashed horisontal lines in the figures on the right panel correspond to the 95% confidence bands for the sample autocorrelation. Note that the figures in each panel are in random order, i.e. a time series in a given row does not necessarily correspond to the autocorrelation function in the same row.

Problem 3. (10 marks)

Write the following models using backshift operators. Clearly indicate the order of the different lag-polynomials, e.g. for an ARIMA(1,1,3),

$$(1 - \phi_1 B)(1 - B)y_t = \delta + (1 - \theta_1 B - \theta_2 B^2 - \theta_3 B^3)\epsilon_t.$$

(a.) ARIMA(1,2,1).

(b.) ARIMA(0,2,2).

(c.) ARIMA(1,1,1)×(1,1,2) with seasonal period $s = 4$.

Problem 4. (20 marks)

A marketing company wants to model the average monthly amount y_t (in SEK) a car company spends on advertising, i.e marketing spending, as a function of their monthly sales x_t (in SEK) as follows

$$y_t = \alpha_0 + \delta_1 x_{t-1} + \delta_2 x_{t-2} + \delta_3 x_{t-3} + u_t, \tag{7}$$

where u_t denotes the error term. Both y_t and x_t can be assumed to be stationary and weakly dependent.

(a.) What kind of model is (7)? Suppose OLS is used to estimate (7). What is required to obtain consistent estimates of δ_1, δ_2 and δ_3? Suppose further that u_t is serially correlated. Does your stated result for consistency still hold?

(b.) Suppose there is a temporary 10 million SEK increase in monthly sales at time t. What is the immediate change in marketing spending due to this increase?

(c.) Suppose there is a temporary 20 million SEK increase in monthly sales at time t. What is the change in marketing spending two periods after this increase?

(d.) Suppose there is a permanent 5 million SEK increase in monthly sales at time t. What is the cumulative effect on marketing spending two time periods after this permanent change?

(e.) The car company hires a new CEO at time t who promises that she will deliver a permanent 10 million SEK increase in monthly sales from the moment she starts. What is the long-term change in marketing spending with this new CEO?

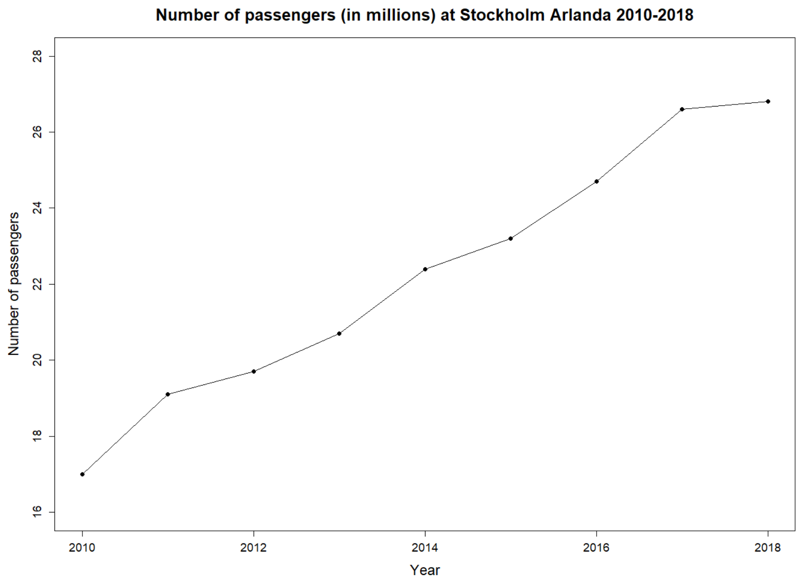

Figure 3: Number of passengers (in millions) that used Stockholm Arlanda airport during 2010-2018 in Problem 5.

Problem 5. (25 marks)

Table 1 shows the number of passengers (in millions) that used Stockholm Arlanda airport during 2010-2018. Figure 3 shows a time series plot of the data.

Table 1: Number of passengers (in millions) that used Stockholm Arlanda airport during 2010-2018 in Problem 5.

Year	2010	2011	2012	2013	2014	2015	2016	2017	2018
y_T	17	19.1	19.7	20.7	22.4	23.2	24.7	26.6	26.8

Table 2 shows the data y_T together with the smoothed values $\tilde{y}_T^{(1)}$ and $\tilde{y}_T^{(2)}$ obtained as

$$\begin{aligned}\tilde{y}_T^{(1)} &= \lambda y_T \quad + (1-\lambda)\tilde{y}_{T-1}^{(1)} \\ \tilde{y}_T^{(2)} &= \lambda \tilde{y}_T^{(1)} + (1-\lambda)\tilde{y}_{T-1}^{(2)},\end{aligned}$$

for $T = 1, \ldots, 5$, ($T = 1$ and $T = 5$ correspond to 2010 and 2014, respectively) with $\tilde{y}_0^{(1)} = \tilde{y}_0^{(2)} = 17$ and $\lambda = 0.6$.

Table 2: Number of passengers that used Stockholm Arlanda airport during 2010-2018 and smoothed values (up to 2014) using $\lambda = 0.6$ in Problem 5.

Year	2010	2011	2012	2013	2014	2015	2016	2017	2018
y_T	17	19.1	19.7	20.7	22.4	23.2	24.7	26.6	26.8
$\tilde{y}_T^{(1)}$	17.00	18.26	19.12	20.07	21.47	?	?	?	?
$\tilde{y}_T^{(2)}$	17.00	17.76	18.58	19.47	20.67	?	?	?	?

(a.) Explain why simple exponential smoothing is not a suitable model for these data.

(b.) Complete the cells with question marks (?) in Table 2 using double exponential smoothing (with the same λ as above, i.e. $\lambda = 0.6$). Moreover, compute the unbiased estimates $\hat{y}_T$, $T = 1, \ldots, 9$, of the expected value of the underlying process

$$y_t = \beta_0 + \beta_1 t + \varepsilon_t,$$

where ε_t are iid for all t with $E(\varepsilon_t) = 0$. You do not need to recompute the smoothed values that are given in Table 2.

(c.) Forecast the number of passengers that use Stockholm Arlanda airport in 2019 using double exponential smoothing. It turned out that the actual number of passengers that used Stockholm Arlanda airport in 2019 was 25.6 (in millions). Compute the relative forecast error.

(d.) Consider only the first four years of data, i.e. 2010, 2011, 2012, 2013. Forecast the number of passengers that will visit Arlanda in 2014 and 2015 using Holt's method. Use the starting values for the level and the trend as $L_0 = 16$ and $T_0 = 1.2$, respectively, and the smoothing constants $\alpha = 0.5$ and $\gamma = 0.6$. Compute the forecast errors (the one-step ahead forecast error for 2014 and the two-step ahead forecast error for 2015).

Problem 6. (15 marks)

Suppose that y_t follows the MA(5) process

$$y_t = \mu - \theta_2 \epsilon_{t-2} - \theta_4 \epsilon_{t-4} - \theta_5 \epsilon_{t-5}, \qquad (8)$$

where the error ϵ_t is white noise with $\mathrm{E}(\epsilon_t) = 0$ and $\mathrm{V}(\epsilon_t) = \sigma^2$. In this problem, you are **not allowed** to use the formulas for the autocovariance and autocorrelation of an MA(5) process (Equations (5.9) and (5.10) in MJK): you have to do the calculations yourself.

(a.) Compute the autocovariance function of y_t, $\gamma_y(k) = \mathrm{Cov}(y_t, y_{t+k})$ for $k = 1, 2, 3, 4, 5$, and the autocorrelation function of y_t, $\rho_y(-k) = \mathrm{Corr}(y_t, y_{t-k})$ for $k = 2, 3$.

(b.) Suppose that $\epsilon_t \sim \mathcal{N}(0, 1)$, $\mu = 2$, $\theta_2 = -0.2$, and $\theta_4 = \theta_5 = 0.4$. Assume that, at time $t = 15$, $\epsilon_{15} = 0.28$, $\epsilon_{14} = -0.5$, $\epsilon_{13} = 0.7$, $\epsilon_{12} = 0.2$, $\epsilon_{11} = 1.05$, $\epsilon_{10} = -1.21$, and $\epsilon_9 = 2.11$. Let $\hat{y}_{t+1}$ be the optimal one-step ahead forecast that minimises $\mathrm{E}\left((y_{t+1} - \hat{y}_{t+1})^2\right)$. Compute $\hat{y}_{t+1}$ for $t = 15$.
Hint: $\hat{y}_{t+1}$ is a special type of expectation (that is often played in the broken record!).