All Exercises (W1-W6)

Time Series Analysis (Erasmus Universiteit Rotterdam)

(International) Bachelor Econometrics and Operations Research
FEB23001(X)-19 Tijdreeksanalyse / Time Series Analysis

Exercises - Key features of economic time series (Chapter 2)

Exercise 2.1

The Excel file `USUNINSCE.xlsx` contains quarterly observations of US initial claims for unemployment insurance[1] over the period 1974Q1–2013Q4 (160 observations). The series is referred to as 'initial claims' below. A dummy variable DREC is also included, which takes the value 1 during recession periods and 0 during expansions.[2]

Examine whether the time series of initial claims displays any of the 'key features' [(i) trend; (ii) seasonality; (iii) aberrant observations; (iv) heteroskedasticity; (v) nonlinearity] discussed in Chapter 2. You may also want to consider transformations of the time series such as quarterly changes or quarterly growth rates (obtained by taking first differences of the logarithm of the series), and use the graphical and auxiliary regression tools discussed in Chapter 2.

Solution

In terms of the 'key features', we observe the following for the time series of quarterly US initial claims for unemployment insurance (referred to as 'initial claims' in the following):

(i) **Trend**: Figure 1 shows the original time series of initial claims, denoted as y_t. We do not observe a pronounced trend in the original time series. This is confirmed by the regression

$$y_t = \alpha + \beta t + u_t, \qquad t = 1, 2, \ldots, T,$$

which gives an estimate (obtained by least squares using the complete sample period 1974Q1–2013Q4 ($T = 160$), with the standard error in parentheses) of $\hat{\beta} = -0.200\,(0.123)$, which is not significantly different from zero at the 10% significance level (p-value: 0.106).

[1]This variable, collected by the US Department of Labor is one of the ten components of The Conference Board's Leading Economic Index, which is one of the most important indicators of the (future) state of the US economy; see `http://www.conference-board.org/data/bcicountry.cfm?cid=1` for more information. Also see `https://www.dol.gov/ui/data.pdf`.

[2]Based on the turning points provided at `http://www.nber.org/cycles/`

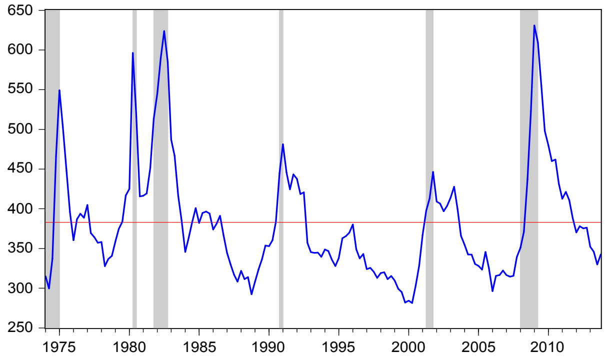


Figure 1: Initial claims. The shaded areas are NBER recession periods.

(ii) **Seasonality**: From Figure 1 it also seems that the time series does not exhibit a pronounced seasonal pattern, in the sense that we do not observe a regular 'saw-tooth' pattern for the different quarters within each year. This is confirmed by the 'vector-of-quarters' plot of the first differences in Figure 2, which shows that it is not the case that the change in initial claims in certain quarters of the year is consistently higher/lower than in other quarters. We check for the presence of seasonality by means of the regression model

$$y_t - y_{t-1} = \mu_1 D_{1,t} + \mu_2 D_{2,t} + \cdots + \mu_4 D_{4,t} + u_t \qquad t = 2, 3, \ldots, T,$$

where $D_{s,t}$ is a seasonal dummy variable which is equal to 1 if the observation at time t is for quarter s and 0 otherwise. We find estimates $\hat{\mu}_1 = -0.218$, $\hat{\mu}_2 = 2.620$, $\hat{\mu}_3 = -1.235$, and $\hat{\mu}_4 = -0.480$, with a standard error of 5.36. A formal F-test of the null hypothesis that $\mu_1 = \mu_2 = \mu_3 = \mu_4$ takes the value 0.099 with a corresponding p-value of 0.96, such that the null hypothesis of equal mean changes in different quarters cannot be rejected. The R^2 of this regression is equal to 0.002, which also shows that seasonal variation can explain only a tiny fraction of the variation in (the change of) initial claims.

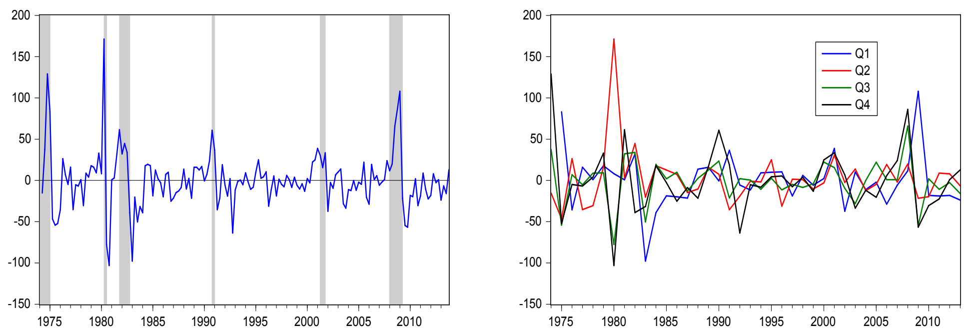


Figure 2: First differences of initial claims.

(iii) **Aberrant observations**: From the graph of the initial claims series in Figure 1 we observe that it tends to decline during periods of economic expansions, while it increases during recessions. This pattern makes sense, as (relatively) more people will be hired (fired) during expansions (recessions). The peaks in initial claims mostly coincide with troughs in the business cycle (i.e. the end of recessions). In terms of the changes in initial claims (see the left panel of Figure 2), we observe some very large positive and negative values exceeding 100 (in magnitude). Given that the first differences have a mean of approximately zero and a standard deviation of 33.6, such values are rather 'unlikely' (in the sense that the probability that such large values occur is small, for example under a normal distribution). It would require a further detailed analysis to establish whether these observations indeed should be classified as 'outliers'.

(iv) **Heteroskedasticity**: From the graphs of initial claims we do not observe that the fluctuations in the time series tend to increase or decrease over time. The plot of the first differences in the left panel of Figure 2, however, suggests that the variance might be larger during recessions than during expansions. This is confirmed by the sample standard deviations, which is equal to 55.33 during recessions (as determined by the NBER dummy) compared to only 22.87 during expansions.

(v) **Nonlinearity**: From Figure 1, it seems that the increases in initial claims during economic recessions occur more rapidly than the declines during expansions. We examine the relevance of this possible 'nonlinearity' by means of the regression

$$y_t - y_{t-1} = \mu_1 D_{t,\mathrm{R}} + \mu_2(1 - D_{t,\mathrm{R}}) + u_t, \qquad t = 1, 2, \ldots, T,$$

where $D_{t,\mathrm{R}}$ is the dummy variable for recessions. We find estimates $\hat{\mu}_1 = 41.26\,(6.26)$ and $\hat{\mu}_2 = -6.42\,(2.51)$, with standard errors in parentheses. The average change during recessions thus is significantly different from the average change during expansions, suggesting that this type of 'nonlinearity' may be a relevant characteristic of the time series.

Exercise 2.2

The Excel file `USCC.xlsx` contains quarterly observations on US total consumer credit outstanding[3] over the period 1959Q1–2017Q4 (236 observations). Column B contains the original series (converted to an index such that the value in 2005Q1 is equal to 100, and adjusted for inflation), while column C contains the seasonally adjusted series. A dummy variable DREC is also included, which takes the value 1 during recession periods and 0 during expansions.[4]

Examine whether the time series of consumer credit displays any of the 'key features' [(i) trend; (ii) seasonality; (iii) aberrant observations; (iv) heteroskedasticity; (v) nonlinearity] discussed in Chapter 2. You may involve both the original and the seasonally adjusted series in your analysis. You may also want to consider transformations of the time series such as quarterly growth rates (obtained by taking first differences of the (natural) logarithm of the series), and use the graphical and auxiliary regression tools discussed in Chapter 2.

Solution

In terms of 'key features', we observe the following for the time series of quarterly consumer credit:

(i) **Trend**: The left panel of Figure 1 shows the original time series of consumer credit, while the right panel shows its (natural) logarithm. We observe a pronounced upward trend, which is interrupted by short periods of declining credit levels around economic recessions. The slope of the trend seems to increase after 1993, although for the logarithmic series this does not seem so apparent (suggesting that in terms of growth rates the trend may be stable). Denoting the log consumer credit by x_t, the presence of a trend is confirmed by the regression

$$x_t = \alpha + \beta t + u_t, \qquad t = 1, 2, \ldots, T,$$

which gives an estimate (obtained by least squares, using the complete sample period 1959Q1–2017Q4 ($T = 236$)) of $\hat{\beta} = 0.0086\,(8.96 \times 10^{-5})$ with the standard error in parentheses.

[3] This variable is considered to be a good indicator of potential future consumption spending levels. It is watched closely by firms, investors, and policymakers, see `https://www.investopedia.com/terms/c/consumercredit.asp`. It also is (part of) one of the components of the Conference Board's Lagging Economic Index, which is an important indicator of the state of the US economy; see `http://www.conference-board.org/data/bcicountry.cfm?cid=1` for more information.

[4] Based on the turning points provided at `http://www.nber.org/cycles/`

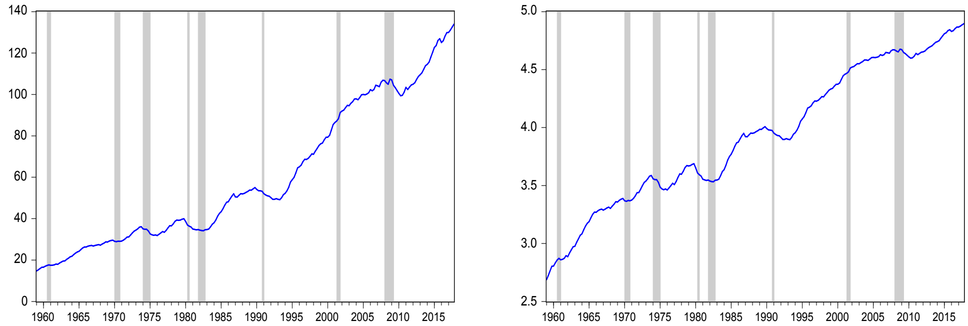


Figure 1: Levels of consumer credit (left panel) and its logarithm (right panel). The shaded areas are NBER recession periods.

(ii) **Seasonality**: The left panel of Figure 2 shows 100 times the first differences of x_t (i.e. quarterly growth rates of consumer credit). We observe a fairly regular 'saw-tooth' pattern for the different quarters within each year. This is confirmed by the 'vector-of-quarters' plot in the right panel of Figure 2, which shows that for most years the growth in consumer credit is higher in Q3 and Q4 than in Q1. For Q2, the growth rate is quite close to the values of Q3 and Q4 before 1980, but declines after this date to a level that is comparable with Q1. We check for the presence of seasonality by means of the regression model

$$100 * (x_t - x_{t-1}) = \mu_1 D_{1,t} + \mu_2 D_{2,t} + \cdots + \mu_4 D_{4,t} + u_t \qquad t = 2, 3, \ldots, T,$$

where $D_{s,t}$ is a seasonal dummy variable which is equal to 1 if the observation at time t is for quarter s and 0 otherwise. We find estimates $\hat{\mu}_1 = -0.18$, $\hat{\mu}_2 = 0.83$, $\hat{\mu}_3 = 1.51$, and $\hat{\mu}_4 = 1.58$, with a standard error of 0.19. A formal F-test of the null hypothesis that $\mu_1 = \mu_2 = \mu_3 = \mu_4$ takes the value 18.75 with a corresponding p-value of 0.000, such that the null can be convincingly rejected. The R^2 of the regression is equal to 0.196, suggesting that seasonality is moderately important in explaining the variation of consumer credit.

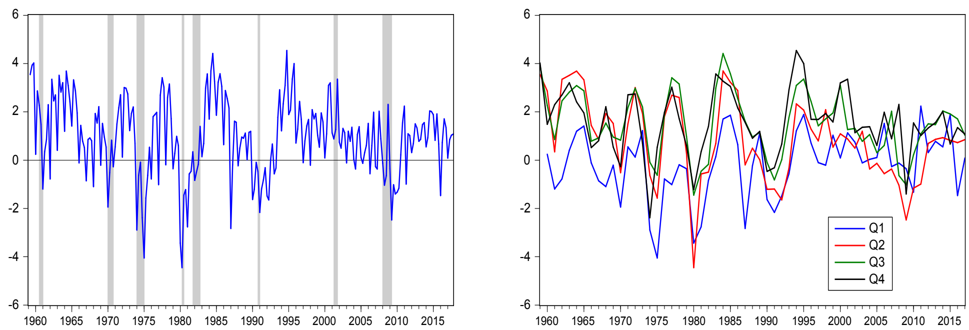


Figure 2: Quarterly growth rate of consumer credit.

(iii) **Aberrant observations**: The time series does not seem to contain observations that obviously would be classified as 'outliers'. The plot of the first differences in Figure 2 suggests that perhaps the observations in 1980Q2 and 1987Q1 may be considered as 'aberrant' (given that a sudden spike downwards in the growth rate occurs in those quarters). This is confirmed by Figure 3, showing the quarterly growth of the seasonally adjusted time series. In fact, in addition to the observations mentioned above, this graph brings forward more observations for which the value is quite different from what might be expected given the values of neighboring observations: 1989Q1, 1998Q2, 2006Q1, 2008Q3, 2008Q4, 2011Q1, and 2016Q1. Whether or not these observations should be classified as 'outliers' requires further investigation though.

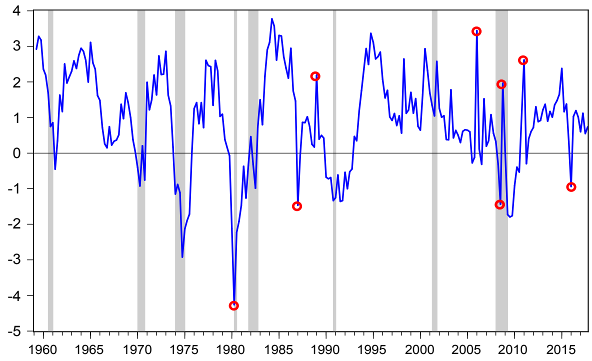


Figure 3: Quarterly growth rate of seasonally adjusted consumer credit.

(iv) **Heteroskedasticity**: The graphs of the growth rates in Figures 2 and 3 do not show obvious signs of (conditional) heteroskedasticity. Structural breaks in volatility also are not apparent. For example, while many US macroeconomic time series experienced a large and sudden reduction in volatility around 1984, this does not seem to be the case for consumer credit: The standard deviation of the quarterly growth rates of the seasonally adjusted time series for the period 1959Q2-1983Q4 is equal to 1.54, while it is equal to 1.25 for the period 1984Q1-2017Q4. At the same time, Figure 4 shows the standard deviation of these growth rates computed using a 10-year moving window, which suggests that volatility has fluctuated substantially over time: starting from around 1.0 in the 1960s it increased to levels around 1.9 for windows ending around 1985, and then declined again to 0.73 for the window 1996Q1-2005Q4, and finally increased again to a level around 1.1 for the final windows.

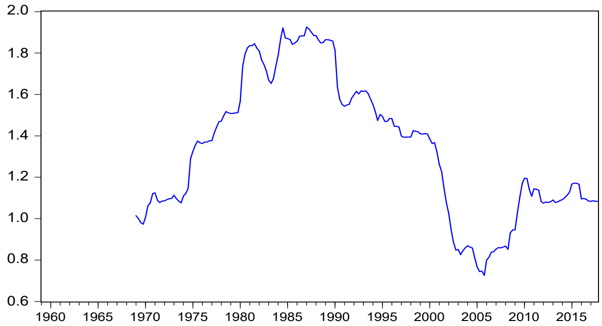


Figure 4: Standard deviation of quarterly growth rate of seasonally adjusted consumer credit based on 10-year moving window.

(v) **Nonlinearity**: As noted above, the positive trend in consumer credit is interrupted by short periods of declining levels around economic recessions. We examine the relevance of this possible 'nonlinearity' by means of the regression

$$100 * (x_t - x_{t-1}) = \mu_1 D_{t,\mathrm{R}} + \mu_2(1 - D_{t,\mathrm{R}}) + u_t, \qquad t = 1, 2, \ldots, T,$$

where $D_{t,\mathrm{R}}$ is the dummy variable for recessions. We find estimates $\hat{\mu}_1 = -0.48\,(0.27)$ and $\hat{\mu}_2 = 1.15\,(0.11)$, with standard errors in parentheses. An F-test of the null hypothesis that $\mu_1 = \mu_2$ takes the value 30.8 with a corresponding p-value of 0.000. The average growth rate during recessions thus is significantly different from the average growth rate during expansions. The R^2 of the regression is 0.117, suggesting that the difference in means across expansions and recessions explains a limited (but still noteworthy) fraction of the variation of the consumer credit growth rates.[5]

[5]This finding is not affected by the seasonality in consumer credit, in the sense that for the seasonally adjusted growth rates we find similar results with estimates $\hat{\mu}_1 = -0.42\,(0.23)$ and $\hat{\mu}_2 = 1.14\,(0.09)$, with an $R^2 = 0.143$. Similarly, augmenting the regression with three quarterly dummy variables, we still find a significant difference between $\hat{\mu}_1$ and $\hat{\mu}_2$. Interestingly, the R^2 of this augmented regression is equal to 0.31, showing that jointly the seasonality and nonlinearity explain quite a substantial part of the variation in consumer credit growth.

Exercise 2.3

The Excel file `LSAGDP.xlsx` contains quarterly observations on the natural logarithm of South African Gross Domestic Product (GDP) over the period 1960Q1–2016Q2 (226 observations), as obtained from OECD Statistics.[6] A dummy variable DREC is also included, which takes the value 1 during recession periods and 0 during expansions, as determined with the BBQ algorithm of Harding and Pagan (2002).[7]

Examine whether the time series of log South African GDP displays any of the 'key features' [(i) trend; (ii) seasonality; (iii) aberrant observations; (iv) heteroskedasticity; (v) nonlinearity] discussed in Chapter 2. You may also want to consider transformations of the series such as first differences (corresponding with quarterly growth rates), and use the auxiliary graphs and regressions discussed in Chapter 2.

Solution

We observe the following 'key features' for the time series of quarterly log GDP:

(i) **Trend**: The time series shows a clear upward trend, which is interrupted by short periods of declining GDP levels during economic recessions - see Figure 1. The presence of a trend is confirmed by the regression

$$y_t = \alpha + \beta t + u_t, \qquad t = 1, 2, \ldots, T,$$

which gives an estimate (obtained by least squares) of $\hat{\beta} = 0.0066\,(8.7 \times 10^{-5})$ with the standard error in parentheses. Note that the left panel of Figure 1 suggests that the assumption of a constant trend slope may not be appropriate, in the sense that substantial changes in the 'strength' of the trend seem to occur at various points during the sample period.[8]

[6] `https://stats.oecd.org/`

[7] Harding, D. and A.R. Pagan (2002), Dissecting the cycle: a methodological investigation, *Journal of Monetary Economics* **49**, 365–381.

[8] This is confirmed by (a) the residuals from the linear trend regression, which still display trending behavior, and (b) a regression that allows for changes in the trend slope (and intercept) in 1970Q1, 1980Q1, 1993Q1 and 2006Q2, which provides a much better fit (and statistically significant changes in the slope).

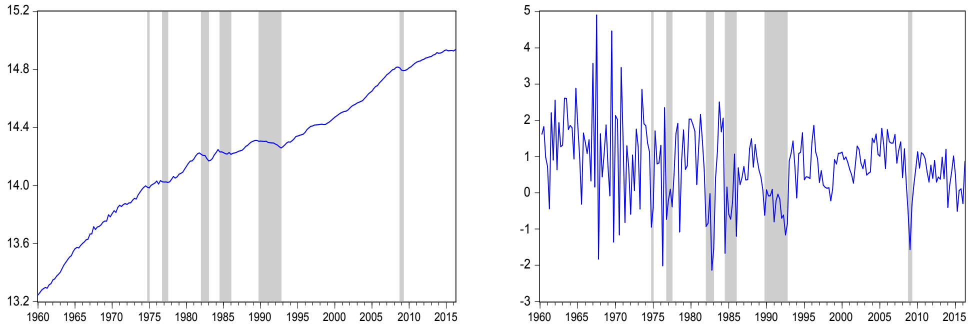


Figure 1: Levels and first differences of the log of quarterly South African GDP. The shaded areas are recession periods.

(ii) **Seasonality**: The time series does not seem to exhibit a pronounced seasonal pattern. This is confirmed by the regression model

$$y_t - y_{t-1} = \mu_1 D_{1,t} + \mu_2 D_{2,t} + \cdots + \mu_4 D_{4,t} + u_t \qquad t = 2, 3, \ldots, T,$$

where $D_{s,t}$ is a seasonal dummy variable which is equal to 1 if the observation at time t is for quarter s and 0 otherwise. We find estimates $\hat{\mu}_1 = 0.77$, $\hat{\mu}_2 = 0.69$, $\hat{\mu}_3 = 0.84$, and $\hat{\mu}_4 = 0.70$, which are not significantly different from each other given their standard error of 0.14. Also, the R^2 of the regression is equal to only 0.003, indicating that seasonality is not an important aspect of the variation in GDP growth.

(iii) **Aberrant observations**: The time series does not seem to contain observations that obviously would be classified as 'outliers'. The first differences in the right panel of Figure 4 suggest that perhaps the observations in 1967Q3, 1969Q3 and 1976Q2 may be considered as 'aberrant' (given that a sudden spike upwards in the growth rate occurs in 1967Q3 and 1969Q3, and a sudden spike downwards in 1976Q2), but it would require a further detailed analysis to confirm this.

(iv) **Heteroskedasticity**: The time series does not show strong signs of heteroskedasticity. There might be a permanent change in volatility, in that the quarterly growth rates seem to have become smaller since 1985. Indeed, the standard deviation of the quarterly growth rates for the period 1960Q2-1984Q4 is equal to 1.34, while it is equal to only 0.68 for the period 1985Q1-2016Q2.

(v) **Nonlinearity**: As noted above, the positive trend in GDP is interrupted by short periods of declining levels during economic recessions. We examine the relevance of this possible 'nonlinearity' by means of the regression

$$y_t - y_{t-1} = \mu_1 D_{t,\mathrm{R}} + \mu_1(1 - D_{t,\mathrm{R}}) + u_t, \qquad t = 1, 2, \ldots, T,$$

where $D_{t,\mathrm{R}}$ is the dummy variable for recessions. We find estimates $\hat{\mu}_1 = -0.56\,(0.15)$ and $\hat{\mu}_2 = 0.99\,(0.06)$, with standard errors in parentheses. The average growth rate during recessions thus is significantly different from the average growth rate during expansions.

(International) Bachelor Econometrics and Operations Research
FEB23001(X)-19 Tijdreeksanalyse / Time Series Analysis

Exercises - Basic concepts (Chapter 3)

Exercise 3.1

Derive the first four values π_1, π_2, π_3, and π_4 in the polynomial

$$\pi(L) = 1 - \pi_1 L - \pi_2 L^2 - \pi_3 L^3 - \pi_4 L^4 - \dots \tag{1}$$

when it holds true that

$$\pi(L) = \frac{1 - \phi_1 L - \phi_2 L^2}{1 + \theta_1 L}. \tag{2}$$

Note that the ratio on the right-hand side of equation (2) corresponds with an ARMA(2,1) model.

Solution

⇒ The purpose of this exercise is to write ARMA models as AR models, or as MA models, and vice versa. This can be useful for various reasons.

⇒ You can think of the above as writing an ARMA model $\phi_p(L)y_t = \theta_q(L)\varepsilon_t$ as an AR model $\pi(L)y_t = \varepsilon_t$, where $\pi(L) = \phi_p(L)/\theta_q(L)$, where in this case $p = 2$ and $q = 1$ with $\phi_p(L) = 1 - \phi_1 L - \phi_2 L^2$ and $\theta_q(L) = 1 + \theta_1 L$. Hence, it follows that

$$\pi(L)\theta_q(L) = \phi_p(L)$$

that is

$$(1 - \pi_1 L - \pi_2 L^2 - \pi_3 L^3 - \pi_4 L^4 - \dots)(1 + \theta_1 L) = (1 - \phi_1 L - \phi_2 L^2).$$

The left-hand side is equal to

$$(1 - (\pi_1 - \theta_1)L - (\pi_2 + \pi_1\theta_1)L^2 - (\pi_3 + \pi_2\theta_1)L^3 - (\pi_4 + \pi_3\theta_1)L^4 - \dots)$$

From this, you can solve for the π_i coefficients, by equating the coefficients for the different lag orders on the left- and hand-right. This gives $\pi_1 = \phi_1 + \theta_1$, $\pi_2 = \phi_2 - \pi_1\theta_1 = \phi_2 - (\phi_1 + \theta_1)\theta_1$, $\pi_3 = -\pi_2\theta_1 = -(\phi_2 - (\phi_1 + \theta_1)\theta_1)\theta_1$, etc.

Exercise 3.2

Consider the variables y_t and x_t, such that y_t can be described by an AR(1) model

$$y_t = \phi_1 y_{t-1} + \varepsilon_t,$$

and x_t by a (restricted) ARMA(4,1) model

$$x_t = \beta x_{t-4} + u_t + \theta u_{t-1},$$

where both ε_t and u_t are white noise series, and $|\phi_1| < 1$, $|\beta| < 1$, and $|\theta| < 1$, to guarantee stationarity and invertibility of y_t and x_t.

Show that the variable z_t defined by $z_t = y_t + x_t$ can be described by an ARMA(5,4) model.

Solution

Write $y_t = \phi_1 y_{t-1} + \varepsilon_t$ as

$$(1 - \phi_1 L) y_t = \varepsilon_t \quad \text{or} \quad y_t = \frac{\varepsilon_t}{1 - \phi_1 L}.$$

Similarly, write $x_t = \beta x_{t-4} + u_t + \theta u_{t-1}$ as

$$(1 - \beta L^4) x_t = (1 + \theta L) u_t \quad \text{or} \quad x_t = \frac{(1 + \theta L) u_t}{1 - \beta L^4}.$$

Then, $z_t = y_t + x_t$ can be written as

$$z_t = \frac{\varepsilon_t}{1 - \phi_1 L} + \frac{(1 + \theta L) u_t}{1 - \beta L^4},$$

or

$$(1 - \phi_1 L)(1 - \beta L^4) z_t = (1 - \beta L^4) \varepsilon_t + (1 - \phi_1 L)(1 + \theta L) u_t.$$

The model for z_t

$$(1 - \phi_1 L)(1 - \beta L^4) z_t = (1 - \beta L^4) \varepsilon_t + (1 - \phi_1 L)(1 + \theta L) u_t$$

can be written as

$$\underbrace{(1 - \phi_1 L - \beta L^4 + \phi_1 \beta L^5)}_{\text{AR(5) polynomial}} z_t = \underbrace{\varepsilon_t - \beta \varepsilon_{t-4} + u_t + (\theta - \phi_1) u_{t-1} - \phi_1 \theta u_{t-2}}_{\text{no dependence beyond lag 4}}.$$

Hence, indeed z_t can be described by an ARMA(5,4) model.

Exercise 3.3
Consider a time series y_t, which can be described by

$$y_t = \phi_1 y_{t-1} + \varepsilon_t + \theta_1 \varepsilon_{t-1}, \tag{3}$$

where ε_t is a standard white noise variable with variance σ_ε^2. Unfortunately, it turns out that y_t is only observed with measurement error, that is, one observes z_t, given by

$$z_t = y_t + u_t, \tag{4}$$

instead of y_t, where u_t is also a standard white noise variable with variance σ_u^2. It is known that ε_t and u_t are mutually uncorrelated at all lags.

Derive the ARMA model for z_t.

Solution
This can be done as follows.

$$z_t = \frac{1+\theta_1 L}{1-\phi_1 L}\varepsilon_t + u_t, \tag{5}$$

or

$$(1-\phi_1 L)z_t = (1+\theta_1 L)\varepsilon_t + (1-\phi_1 L)u_t. \tag{6}$$

The right-hand side of this model is an MA(1) variable, as it can be written as

$$= \underbrace{\varepsilon_t + \theta_1\varepsilon_{t-1} + u_t - \phi_1 u_{t-1}}_{\text{no dependence beyond lag 1}}.$$

We can also understand this by considering the structure of its autocovariance function. The unconditional variance γ_0 of $(1+\theta_1 L)\varepsilon_t + (1-\phi_1 L)u_t$ is equal to

$$\gamma_0 = (1+\theta_1^2)\sigma_\varepsilon^2 + (1+\phi_1^2)\sigma_u^2, \tag{7}$$

and its first-order autocovariance γ_1 is found to be

$$\gamma_1 = \theta_1\sigma_\varepsilon^2 - \phi_1\sigma_u^2, \tag{8}$$

while higher-order autocovariances γ_k are equal to 0, for $k = 2, 3, 4, ...$

Hence, combined with the AR(1) component $(1-\phi_1 L)z_t$, the resulting model for z_t is an ARMA(1,1) model.

Exercise 3.4
Derive expressions for the autocorrelations ρ_k for the AR(2) model given by

$$y_t - \phi_1 y_{t-1} - \phi_2 y_{t-2} = \varepsilon_t. \tag{9}$$

Solution
Multiplying both sides by y_{t-1}, taking expectations, and dividing by γ_0 results in

$$\rho_1 - \phi_1 \rho_0 - \phi_2 \rho_1 = 0, \tag{10}$$

and since $\rho_0 = 1$, we obtain

$$\rho_1 = \phi_1/(1 - \phi_2). \tag{11}$$

To obtain an expression for ρ_2, analogous operations are carried out on (9) except for using y_{t-2} instead of y_{t-1}, yielding

$$\rho_2 - \phi_1 \rho_1 - \phi_2 \rho_0 = 0. \tag{12}$$

Substituting (11) into (12) gives

$$\rho_2 = \phi_1^2/(1 - \phi_2) + \phi_2. \tag{13}$$

Analogous to (11) and (12) we can derive that in general

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2}. \qquad \text{for } k = 2, 3, 4, \ldots \tag{14}$$

Exercise 3.5

Consider the general MA(∞) representation for a stationary time series y_t, that is,

$$y_t = \varepsilon_t + \theta_1\varepsilon_{t-1} + \theta_2\varepsilon_{t-2} + \theta_3\varepsilon_{t-3} + \theta_4\varepsilon_{t-4} + \dots$$

and suppose the parameters θ_1, $\theta_2, \dots$ are given.

a. What are the forecast errors for 3 and for 4 steps ahead?
b. What is the covariance between these errors?

Solution

What is known at time T about the observation for time $T+h$?

$$y_{T+h} = \underbrace{\varepsilon_{T+h} + \theta_1\varepsilon_{T+h-1} + \dots + \theta_{h-1}\varepsilon_{T+1}}_{\text{unknown at time } T} + \underbrace{\theta_h\varepsilon_T + \theta_{h+1}\varepsilon_{T-1} + \dots}_{\text{known at time } T}$$

The optimal h-step ahead point forecast is equal to

$$y_{T+h|T} = \mathsf{E}[y_{T+h}|\mathcal{Y}_T] = \theta_h\varepsilon_T + \theta_{h+1}\varepsilon_{T-1} + \theta_{h+2}\varepsilon_{T-2} + \dots$$

with forecast error

$$e_{T+h|T} = \varepsilon_{T+h} + \theta_1\varepsilon_{T+h-1} + \dots + \theta_{h-1}\varepsilon_{T+1}.$$

For 3- and 4-steps ahead, this gives

$$\begin{aligned} e_{T+3|T} &= \varepsilon_{T+3} + \theta_1\varepsilon_{T+2} + \theta_2\varepsilon_{T+1}, \\ e_{T+4|T} &= \varepsilon_{T+4} + \theta_1\varepsilon_{T+3} + \theta_2\varepsilon_{T+2} + \theta_3\varepsilon_{T+1} \end{aligned}$$

For the covariance between these errors it follows that

$$\begin{aligned} \mathsf{E}[e_{T+3|T}e_{T+4|T}] &= \mathsf{E}[\theta_1\varepsilon_{T+3}^2 + \theta_1\theta_2\varepsilon_{T+2}^2 + \theta_2\theta_3\varepsilon_{T+1}^2] \\ &= (\theta_1 + \theta_1\theta_2 + \theta_2\theta_3)\sigma^2 \end{aligned}$$

Exercise 3.6

Consider the ARMA(1,1) model for a time series variable y_t,

$$y_t = \phi_1 y_{t-1} + \varepsilon_t + \theta_1 \varepsilon_{t-1}, \qquad t = 1, \ldots, T, \tag{15}$$

where the parameters are such that the model is stationary and invertible, that is $|\phi_1| < 1$ and $|\theta_1| < 1$. Furthermore, the shocks ε_t are independent and normally distributed with constant variance σ^2, that is $\varepsilon_t \sim N(0, \sigma^2)$.

a. Derive the optimal 1-step, 2-steps and 3-steps ahead point forecasts of y_t in case of the standard quadratic loss function (that is, derive expressions for $\hat{y}_{T+k|T} = \mathsf{E}[y_{T+k}|\mathcal{Y}_T]$ for $k = 1$, 2 and 3, where $\mathcal{Y}_T$ denotes the available history of the time series up to time T).

b. Express the corresponding forecast errors

$$e_{T+k|T} \equiv y_{T+k} - \hat{y}_{T+k|T},$$

for $k = 1$, 2 and 3 in terms of the shocks ε_t in (15).

c. Suppose that the parameters in (15) have been estimated using observations y_t for $t = 1, \ldots, T$, with the following results:

$$\hat{\phi}_1 = 0.78, \quad \hat{\theta}_1 = 0.32, \quad \text{and } \hat{\sigma}^2 = 0.51.$$

Furthermore, assume that $y_{T-1} = 9.79$, $y_T = 7.24$ and the corresponding residuals $\hat{\varepsilon}_{T-1} = 0.51$ and $\hat{\varepsilon}_T = -0.56$.

c.1 What is the implied 2-step ahead point forecast $\hat{y}_{T+2|T}$?

c.2 What is the variance of the corresponding 2-step ahead forecast error $e_{T+2|T}$?

c.3 Give a 90% "equal-tail probability" interval forecast for y_{T+2} at time T (i.e. 2-steps ahead). That is, derive the lower bound $\widehat{L}_{T+2|T}$ and upper bound $\widehat{U}_{T+2|T}$ such that

$$P[\widehat{L}_{T+2|T} < y_{T+2} < \widehat{U}_{T+2|T}|\mathcal{Y}_T] = 0.90, \quad \text{and}$$
$$P[y_{T+2} < \widehat{L}_{T+2|T}|\mathcal{Y}_T] = P[y_{T+2} > \widehat{U}_{T+2|T}|\mathcal{Y}_T] = 0.05.$$

Solution

a. The 1-step ahead forecast at time T is given by

$$\begin{aligned}\hat{y}_{T+1|T} &= \mathsf{E}[y_{T+1}|\mathcal{Y}_T] \\ &= \mathsf{E}[\phi_1 y_T + \varepsilon_{T+1} + \theta_1 \varepsilon_T|\mathcal{Y}_T] \\ &= \phi_1 y_T + \theta_1 \varepsilon_T,\end{aligned}$$

because $\mathsf{E}[\varepsilon_{T+1}|\mathcal{Y}_T] = 0$.

The 2-step ahead forecast at time T is given by

$$\begin{aligned}
\hat{y}_{T+2|T} &= \mathsf{E}[y_{T+2}|\mathcal{Y}_T] \\
&= \mathsf{E}[\phi_1 y_{T+1} + \varepsilon_{T+2} + \theta_1 \varepsilon_{T+1}|\mathcal{Y}_T] \\
&= \phi_1 \mathsf{E}[y_{T+1}|\mathcal{Y}_T] \\
&= \phi_1 \hat{y}_{T+1|T} \\
&= \phi_1^2 y_T + \phi_1 \theta_1 \varepsilon_T.
\end{aligned}$$

The 3-step ahead forecast at time T is given by

$$\begin{aligned}
\hat{y}_{T+3|T} &= \mathsf{E}[y_{T+3}|\mathcal{Y}_T] \\
&= \mathsf{E}[\phi_1 y_{T+2} + \varepsilon_{T+3} + \theta_1 \varepsilon_{T+2}|\mathcal{Y}_T] \\
&= \phi_1 \mathsf{E}[y_{T+2}|\mathcal{Y}_T] \\
&= \phi_1 \hat{y}_{T+2|T} \\
&= \phi_1^3 y_T + \phi_1^2 \theta_1 \varepsilon_T.
\end{aligned}$$

b. For the 1-step ahead forecast error, it holds that

$$\begin{aligned}
e_{T+1|T} &= y_{T+1} - \mathsf{E}[y_{T+1}|\mathcal{Y}_T] \\
&= y_{T+1} - \phi_1 y_T - \theta_1 \varepsilon_T \\
&= \varepsilon_{T+1}.
\end{aligned}$$

The 2-step ahead forecast error is equal to

$$\begin{aligned}
e_{T+2|T} &= y_{T+2} - \mathsf{E}[y_{T+2}|\mathcal{Y}_T] \\
&= y_{T+2} - \phi_1^2 y_T - \phi_1 \theta_1 \varepsilon_T \\
&= \phi_1 y_{T+1} + \varepsilon_{T+2} + \theta_1 \varepsilon_{T+1} - \phi_1^2 y_T - \phi_1 \theta_1 \varepsilon_T \\
&= \phi_1^2 y_T + \phi_1 \varepsilon_{T+1} + \phi_1 \theta_1 \varepsilon_T + \varepsilon_{T+2} + \theta_1 \varepsilon_{T+1} - \phi_1^2 y_T - \phi_1 \theta_1 \varepsilon_T \\
&= (\phi_1 + \theta_1)\varepsilon_{T+1} + \varepsilon_{T+2},
\end{aligned}$$

while the 3-step ahead forecast error is equal to

$$\begin{aligned}
e_{T+3|T} &= y_{T+3} - \mathsf{E}[y_{T+3}|\mathcal{Y}_T] \\
&= y_{T+3} - \phi_1^3 y_T - \phi_1^2 \theta_1 \varepsilon_T \\
&= \phi_1 y_{T+2} + \varepsilon_{T+3} + \theta_1 \varepsilon_{T+2} - \phi_1^3 y_T - \phi_1^2 \theta_1 \varepsilon_T \\
&= \phi_1^2 y_{T+1} + \phi_1 \varepsilon_{T+2} + \phi_1 \theta_1 \varepsilon_{T+1} + \varepsilon_{T+3} + \theta_1 \varepsilon_{T+2} - \phi_1^3 y_T - \phi_1^2 \theta_1 \varepsilon_T \\
&= \phi_1^3 y_T + \phi_1^2 \varepsilon_{T+1} + \phi_1^2 \theta_1 \varepsilon_T + \phi_1 \varepsilon_{T+2} + \phi_1 \theta_1 \varepsilon_{T+1} + \varepsilon_{T+3} + \theta_1 \varepsilon_{T+2} - \phi_1^3 y_T - \phi_1^2 \theta_1 \varepsilon_T \\
&= \phi_1(\phi_1 + \theta_1)\varepsilon_{T+1} + (\phi_1 + \theta_1)\varepsilon_{T+2} + \varepsilon_{T+3}
\end{aligned}$$

c.1 The two-step ahead forecast is given by $\hat{y}_{T+2|T} = \phi_1^2 y_T + \phi_1\theta_1\varepsilon_T$. Substituting the observed value for y_T and the estimates of ϕ_1, θ_1 and ε_t gives $\hat{y}_{T+2|T} = 4.265$.

c.2 Note that $\mathsf{E}[e_{T+2|T}|\mathcal{Y}_T] = 0$. The variance of the 2-step ahead forecast error is given by

$$\begin{aligned}\mathsf{V}[e_{T+2|T}] &= \mathsf{E}[e_{T+2|T}^2|\mathcal{Y}_T] \\ &= \mathsf{E}[((\phi_1 + \theta_1)\varepsilon_{T+1} + \varepsilon_{T+2})^2] \\ &= (\phi_1 + \theta_1)^2\mathsf{E}[\varepsilon_{T+1}^2] + \mathsf{E}[\varepsilon_{T+2}^2] \\ &= [(\phi_1 + \theta_1)^2 + 1]\sigma^2.\end{aligned}$$

Again using the estimates of ϕ_1, θ_1, and σ^2, it follows that $\mathsf{V}[e_{T+2|T}] = 1.127$.

c.3 Because $\varepsilon_t \sim N(0, \sigma^2)$, it follows that $e_{T+2|T}|\mathcal{Y}_T \sim N(0, \mathsf{V}[e_{T+2|T}])$, and because $y_{T+2} = \hat{y}_{T+2|T} + e_{T+2|T}$, it follows that $y_{T+2}|\mathcal{Y}_T \sim N(\hat{y}_{T+2|T}, \mathsf{V}[e_{T+2|T}])$. Hence, the required 90% interval forecast is given by

$$(L_T, U_T) = \left(\hat{y}_{T+2|T} - z_{0.95}\sqrt{\mathsf{V}[e_{T+2|T}]}, \hat{y}_{T+2|T} + z_{0.95}\sqrt{\mathsf{V}[e_{T+2|T}]}\right),$$

where $z_{0.95}$ is the 95^{th} percentile of the standard normal distribution, that is, $z_{0.95} = 1.645$. Thus, $L_T = \hat{y}_{T+2|T} - z_{0.95}\sqrt{\mathsf{V}[e_{T+2|T}]} = 2.519$ and $U_T = 6.011$.

Exercise 3.7
Consider the AR(2) model for a time series variable y_t,

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \varepsilon_t, \qquad t = 1, \ldots, T, \tag{16}$$

where the parameters ϕ_1 and ϕ_2 are such that both roots of the AR(2)-polynomial $\phi_2(L) = 1 - \phi_1 L - \phi_2 L^2$ are outside the unit circle, where L is the usual lag operator (defined as $L^k y_t \equiv y_{t-k}$ for all integers k). Furthermore, the shocks ε_t are independent and normally distributed with constant variance σ_ε^2, that is $\varepsilon_t \sim N(0, \sigma_\varepsilon^2)$.

a. Derive the optimal 1-step, 2-steps and 3-steps ahead point forecasts of y_t in case of the standard quadratic loss function (that is, derive expressions for $\hat{y}_{T+k|T} = \mathsf{E}[y_{T+k}|\mathcal{Y}_T]$ for $k = 1$, 2 and 3, where $\mathcal{Y}_T$ denotes the available history of the time series up to time T). [You may assume that the parameters ϕ_1 and ϕ_2 are known.]
b. Express the corresponding forecast errors

$$e_{T+k|T} \equiv y_{T+k} - \hat{y}_{T+k|T},$$

for $k = 1$, 2 and 3 in terms of the shocks ε_t in (16).
c. What is the covariance between the errors for subsequent 1-step, 2-steps and 3-steps ahead forecasts made at time T and $T+1$? That is, what is $\mathsf{E}[e_{T+j|T}e_{T+j+1|T+1}]$ for $j = 1$, 2, and 3?

Solution
a. The 1-step ahead forecast at time T given by

$$\begin{aligned}\hat{y}_{T+1|T} &= \mathsf{E}[y_{T+1}|\mathcal{Y}_T] \\ &= \mathsf{E}[\phi_1 y_T + \phi_2 y_{T-1} + \varepsilon_{T+1}|\mathcal{Y}_T] \\ &= \phi_1 y_T + \phi_2 y_{T-1}.\end{aligned} \tag{17}$$

because $\mathsf{E}[\varepsilon_{T+1}|\mathcal{Y}_T] = 0$.

The 2-step ahead forecast at time T is given by

$$\begin{aligned}\hat{y}_{T+2|T} &= \mathsf{E}[\phi_1 y_{T+1} + \phi_2 y_T + \varepsilon_{T+2}|\mathcal{Y}_T] \\ &= \phi_1 \hat{y}_{T+1|T} + \phi_2 y_T \\ &= \phi_1(\phi_1 y_T + \phi_2 y_{T-1}) + \phi_2 y_T \\ &= (\phi_1^2 + \phi_2) y_T + \phi_1\phi_2 y_{T-1}.\end{aligned} \tag{18}$$

The 3-step ahead forecast at time T is given by

$$\begin{aligned}\hat{y}_{T+3|T} &= \mathsf{E}[\phi_1 y_{T+2} + \phi_2 y_{T+1} + \varepsilon_{T+3}|\mathcal{Y}_T] \\ &= \phi_1 \hat{y}_{T+2|T} + \phi_2 \hat{y}_{T+1|T} \\ &= \phi_1(\phi_1(\phi_1 y_T + \phi_2 y_{T-1}) + \phi_2 y_T) + \phi_2(\phi_1 y_T + \phi_2 y_{T-1}) \\ &= (\phi_1^3 + 2\phi_1\phi_2) y_T + (\phi_1^2\phi_2 + \phi_2^2) y_{T-1}. \end{aligned} \tag{19}$$

b. For the 1-step ahead forecast error, it holds that

$$\begin{aligned} e_{T+1|T} &= y_{T+1} - \mathsf{E}[y_{T+1}|\mathcal{Y}_T] \\ &= y_{T+1} - \phi_1 y_T - \phi_2 y_{T-1} \\ &= \varepsilon_{T+1}. \end{aligned} \tag{20}$$

The 2-step ahead forecast error $e_{T+2|T} = y_{T+2} - \mathsf{E}[y_{T+2}|\mathcal{Y}_T]$ is equal to

$$e_{T+2|T} = \varepsilon_{T+2} + \phi_1 \varepsilon_{T+1}, \tag{21}$$

as

$$\begin{aligned} y_{T+2} &= \phi_1 y_{T+1} + \phi_2 y_T + \varepsilon_{T+2} \\ &= \phi_1(\phi_1 y_T + \phi_2 y_{T-1} + \varepsilon_{T+1}) + \phi_2 y_T + \varepsilon_{T+2}. \end{aligned}$$

As

$$\begin{aligned} y_{T+3} &= \phi_1 y_{T+2} + \phi_2 y_{T+1} + \varepsilon_{T+3} \\ &= \phi_1(\phi_1(\phi_1 y_T + \phi_2 y_{T-1} + \varepsilon_{T+1}) + \phi_2 y_T + \varepsilon_{T+2}) \\ &\quad + \phi_2(\phi_1 y_T + \phi_2 y_{T-1} + \varepsilon_{T+1}) + \varepsilon_{T+3}, \end{aligned}$$

the 3-step ahead forecast error

$$e_{T+3|T} = \varepsilon_{T+3} + \phi_1 \varepsilon_{T+2} + (\phi_1^2 + \phi_2)\varepsilon_{T+1}. \tag{22}$$

c. For the 1-step ahead forecast errors, we have $e_{T+1|T} = \varepsilon_{T+1}$ and $e_{T+2|T+1} = \varepsilon_{T+2}$. Their covariance is equal to zero, as ε_t is a white noise process.

For the 2-step ahead forecast errors, we have $e_{T+2|T} = \phi_1\varepsilon_{T+1} + \varepsilon_{T+2}$ and $e_{T+3|T+1} = \phi_1\varepsilon_{T+2} + \varepsilon_{T+3}$. Their covariance is equal to $\phi_1\sigma^2$.

For the 3-step ahead forecast errors, we have $e_{T+3|T} = \varepsilon_{T+3} + \phi_1\varepsilon_{T+2} + (\phi_1^2 + \phi_2)\varepsilon_{T+1}$ and $e_{T+4|T+1} = \varepsilon_{T+4} + \phi_1\varepsilon_{T+3} + (\phi_1^2 + \phi_2)\varepsilon_{T+2}$. Their covariance is equal to $\phi_1(1 + (\phi_1^2 + \phi_2))\sigma^2$.

Exercise 3.8
Suppose that the time series y_t is generated according to an AR(1) process

$$y_t = \phi_1 y_{t-1} + \varepsilon_t, \tag{23}$$

with $|\phi_1| < 1$ and ε_t is a white noise series with $\mathsf{E}[\varepsilon_t] = 0$ and $\mathsf{E}[\varepsilon_t^2] = \sigma^2$ for all t.

Consider the series of first differences $\Delta y_t = y_t - y_{t-1}$. Answer the following questions:

a. Show that the unconditional mean $\mathsf{E}[\Delta y_t]$ is equal to 0.
b. Show that the unconditional variance $\mathsf{E}[(\Delta y_t - \mathsf{E}[\Delta y_t])^2]$ is equal to $\frac{2}{1+\phi_1}\sigma^2$.
c. Derive an expression for the k-th order autocorrelation of Δy_t. What happens with these autocorrelations if $\phi_1 \to 0$ in the AR(1) model (23)? And what happens if $\phi_1 \to 1$?

Solution
a. Assuming that $|\phi_1| < 1$, it follows from the properties of a stationary AR(1) process that y_t has unconditional mean equal to zero. Hence,

$$\mathsf{E}[\Delta y_t] = \mathsf{E}[y_t - y_{t-1}] = \mathsf{E}[y_t] - \mathsf{E}[y_{t-1}] = 0 - 0 = 0.$$

b. There are (at least) two ways to solve this.

First, again assuming that $|\phi_1| < 1$, it follows from the properties of a stationary AR(1) process that y_t has unconditional variance equal to $\mathsf{E}[(y_t - \mathsf{E}[y_t])^2] = \mathsf{E}[y_t^2] = \sigma^2/(1-\phi_1^2)$. Also note that we can write Δy_t as $\Delta y_t = y_t - y_{t-1} = (\phi_1 - 1)y_{t-1} + \varepsilon_t$. Hence, it follows that

$$\begin{aligned}
\mathsf{E}[(\Delta y_t - \mathsf{E}[\Delta y_t])^2] &= \mathsf{E}[\Delta y_t^2] \\
&= \mathsf{E}[((\phi_1 - 1)y_{t-1} + \varepsilon_t)^2] \\
&= \mathsf{E}[(\phi_1 - 1)^2 y_{t-1}^2 + \varepsilon_t^2 + 2(\phi_1 - 1)y_{t-1}\varepsilon_t] \\
&= (\phi_1 - 1)^2 \mathsf{E}[y_{t-1}^2] + \mathsf{E}[\varepsilon_t^2] \\
&= (\phi_1 - 1)^2 \sigma^2/(1-\phi_1^2) + \sigma^2 \\
&= \frac{\phi_1^2 - 2\phi_1 + 1 + 1 - \phi_1^2}{1-\phi_1^2}\sigma^2 \\
&= \frac{2(1-\phi_1)}{1-\phi_1^2}\sigma^2 = \frac{2}{1+\phi_1}\sigma^2,
\end{aligned}$$

where we have used the fact that $\mathsf{E}[y_{t-1}\varepsilon_t] = 0$.

Second, from the AR(1) specification for y_t, it follows that

$$\begin{aligned}\Delta y_t &= y_t - y_{t-1} \\ &= \phi_1 y_{t-1} + \varepsilon_t - (\phi_1 y_{t-2} + \varepsilon_{t-1}) \\ &= \phi_1 (y_{t-1} - y_{t-2}) + \varepsilon_t - \varepsilon_{t-1} \\ &= \phi_1 \Delta y_{t-1} + \Delta\varepsilon_t.\end{aligned}$$

Hence,

$$\begin{aligned}\mathsf{E}[\Delta y_t^2] &= \mathsf{E}[(\phi_1 \Delta y_{t-1} + \Delta\varepsilon_t)^2] \\ &= \phi_1^2 \mathsf{E}[\Delta y_{t-1}^2] + \mathsf{E}[\Delta\varepsilon_t^2] + 2\phi_1 \mathsf{E}[\Delta y_{t-1}\Delta\varepsilon_t] \\ &= \phi_1^2 \mathsf{E}[\Delta y_{t-1}^2] + 2\sigma^2 - 2\phi_1\sigma^2, \qquad (24)\end{aligned}$$

Here we have used that

$$\begin{aligned}\mathsf{E}[\Delta\varepsilon_t^2] &= \mathsf{E}[(\varepsilon_t - \varepsilon_{t-1})^2] \\ &= \mathsf{E}[\varepsilon_t^2] + \mathsf{E}[\varepsilon_{t-1}^2] - 2\mathsf{E}[\varepsilon_t\varepsilon_{t-1}] \\ &= \sigma^2 + \sigma^2 + 0 = 2\sigma^2,\end{aligned}$$

and

$$\begin{aligned}\mathsf{E}[\Delta y_{t-1}\Delta\varepsilon_t^2] &= \mathsf{E}[(y_{t-1} - y_{t-2})(\varepsilon_t - \varepsilon_{t-1})] \\ &= \mathsf{E}[y_{t-1}\varepsilon_t] - \mathsf{E}[y_{t-2}\varepsilon_t] - \mathsf{E}[y_{t-1}\varepsilon_{t-1}] + \mathsf{E}[y_{t-2}\varepsilon_{t-1}] \\ &= 0 - 0 - \sigma^2 + 0 = -\sigma^2.\end{aligned}$$

Setting $\mathsf{E}[\Delta y_t^2] = \mathsf{E}[\Delta y_{t-1}^2]$ in (24) and solving for $\mathsf{E}[\Delta y_t^2]$ gives

$$\mathsf{E}[\Delta y_t^2] = \frac{2(1-\phi_1)}{1-\phi_1^2}\sigma^2 = \frac{2}{1+\phi_1}\sigma^2.$$

Note that the unconditional variance of Δy_t becomes equal to $2\sigma^2$ when $\phi_1 = 0$, which makes sense given that $\Delta y_t = \varepsilon_t - \varepsilon_{t-1}$ in that case. The unconditional variance becomes equal to σ^2 when $\phi_1 = 1$, which also makes sense given that $\Delta y_t = \varepsilon_t$ in that case.

c. As shown above, we can write $\Delta y_t = \phi_1 \Delta y_{t-1} + \Delta\varepsilon_t$. Hence (and using the fact that $\mathsf{E}[\Delta y_t] = 0$, the first-order autocovariance of Δy_t is given by

$$\begin{aligned}
\gamma_{1,\Delta y} &= \mathsf{E}[\Delta y_t \Delta y_{t-1}] \\
&= \mathsf{E}[(\phi_1 \Delta y_{t-1} + \Delta\varepsilon_t)\Delta y_{t-1}] \\
&= \phi_1 \mathsf{E}[\Delta y_{t-1}^2] + \mathsf{E}[\varepsilon_t \Delta y_{t-1}] - \mathsf{E}[\varepsilon_{t-1}\Delta y_{t-1}] \\
&= \frac{2\phi_1}{1+\phi_1}\sigma^2 - \sigma^2 \\
&= \frac{2\phi_1 - 1 - \phi_1}{1+\phi_1}\sigma^2 \\
&= \frac{\phi_1 - 1}{1+\phi_1}\sigma^2,
\end{aligned}$$

where we have used the results that $\mathsf{E}[\Delta y_{t-1}^2] = 2\sigma^2/(1+\phi_1)$, $\mathsf{E}[\varepsilon_t \Delta y_{t-1}] = 0$ and $\mathsf{E}[\varepsilon_{t-1}\Delta y_{t-1}] = \sigma^2$.

For the k-th order autocovariance of Δy_t, for $k > 1$, it follows that

$$\begin{aligned}
\gamma_{k,\Delta y} &= \mathsf{E}[\Delta y_t \Delta y_{t-k}] \\
&= \mathsf{E}[(\phi_1 \Delta y_{t-1} + \Delta\varepsilon_t)\Delta y_{t-k}] \\
&= \phi_1 \mathsf{E}[\Delta y_{t-1}\Delta y_{t-k}] + \mathsf{E}[\varepsilon_t \Delta y_{t-k}] - \mathsf{E}[\varepsilon_{t-1}\Delta y_{t-k}] \\
&= \phi_1 \gamma_{k-1,\Delta y} \\
&= \phi_1^{k-1}\frac{\phi_1 - 1}{1+\phi_1}\sigma^2,
\end{aligned}$$

where we have used the fact that $\mathsf{E}[\varepsilon_t \Delta y_{t-k}] = 0$ and $\mathsf{E}[\varepsilon_{t-1}\Delta y_{t-k}] = 0$ for all $k > 1$.

Hence, for the autocorrelations of Δy_t it follows that

$$\rho_{k,\Delta y} \equiv \frac{\gamma_{k,\Delta y}}{\gamma_{0,\Delta y}} = \frac{\phi_1^{k-1}\frac{\phi_1-1}{1+\phi_1}\sigma^2}{\frac{2}{1+\phi_1}\sigma^2} = \frac{\phi_1^{k-1}(\phi_1 - 1)}{2}.$$

(i) When $\phi_1 \to 0$, we find that the first order autocorrelation $\rho_{1,\Delta y}$ becomes equal to -0.5, while all higher-order autocorrelations become equal to zero. Again this makes sense, as $\Delta y_t = \varepsilon_t - \varepsilon_{t-1}$ in this case [which indeed has this autocorrelation properties].

(ii) When $\phi_1 \to 1$, we find that all autocorrelations $\rho_{k,\Delta y}$ become equal to zero. This also makes sense, as $\Delta y_t = \varepsilon_t$ in this case.

[(iii) In general, when $|\phi_1| < 1$, the sum of all autocorrelations $\rho_{k,\Delta y}$, $k = 1, 2, \ldots$, is equal to -0.5 (because in that case $\sum_{k=1}^{\infty}\phi_1^{k-1} = 1/(1-\phi_1)$). Notice that taking first differences of the time series y_t is not necessary if $|\phi_1| < 1$, as it is already covariance-stationary. Put differently, finding that the sum of autocorrelations is equal to -0.5 generally is a sign of *overdifferencing*.]

Exercise 3.9

Suppose that the time series y_t has an AR(1) data generating process (DGP) given by

$$y_t = y_{t-1} + \varepsilon_t, \tag{25}$$

where ε_t is a white noise series with $\mathsf{E}[\varepsilon_t] = 0$ and $\mathsf{E}[\varepsilon_t^2] = \sigma^2$ for all t.

Suppose that the variable y is observed every six months, but that it is aggregated to an annually observed time series x_T by taking the sum of the two observations of y in year T. Show that x_T can be described by

$$x_T = x_{T-1} + u_T,$$

where u_T is a MA(1) process with first order autocorrelation equal to $\frac{1}{6}$.

Solution

From the AR(1) specification for the observed series y, it follows that y_t can be expressed as

$$y_t = y_{t-2} + \varepsilon_t + \varepsilon_{t-1} = y_{t-2} + (1+L)\varepsilon_t.$$

Multiplying both sides with $1+L$ results in

$$(1+L)y_t = (1+L)y_{t-2} + (1+L)(1+L)\varepsilon_t.$$

If the observations at times t and $t-1$ are in year T, this is equivalent to

$$x_T = x_{T-1} + u_T,$$

where u_T corresponds with $(1+L)(1+L)\varepsilon_t = \varepsilon_t + 2\varepsilon_{t-1} + \varepsilon_{t-2}$. It then follows that

$$\mathsf{E}[u_T] = \mathsf{E}[\varepsilon_t + 2\varepsilon_{t-1} + \varepsilon_{t-2}] = 0, \tag{26}$$

$$\mathsf{E}[u_T^2] = \mathsf{E}[(\varepsilon_t + 2\varepsilon_{t-1} + \varepsilon_{t-2})^2] = 6\sigma^2, \tag{27}$$

$$\mathsf{E}[u_T u_{T-1}] = \mathsf{E}[(\varepsilon_t + 2\varepsilon_{t-1} + \varepsilon_{t-2})(\varepsilon_{t-2} + 2\varepsilon_{t-3} + \varepsilon_{t-4})] = \sigma^2, \tag{28}$$

$$\mathsf{E}[u_T u_{T-k}] = \mathsf{E}[(\varepsilon_t + 2\varepsilon_{t-1} + \varepsilon_{t-2})(\varepsilon_{t-2k} + 2\varepsilon_{t-2k-1} + \varepsilon_{t-2k-2})] = 0, \quad \text{for all } k > 1. \tag{29}$$

Thus, indeed u_T has an MA(1) structure (i.e. a non-zero first-order autocorrelation and all higher-order autocorrelations equal to zero), with first-order autocorrelation equal to $1/6$.

Exercise 3.10

Suppose we are interested in a time series x_t that is generated according to an MA(1) process

$$x_t = \delta + \theta_1 \varepsilon_{t-1} + \varepsilon_t, \tag{30}$$

with $|\theta_1| < 1$ and ε_t is a white noise series with $\mathsf{E}[\varepsilon_t] = 0$ and $\mathsf{E}[\varepsilon_t^2] = \sigma_\varepsilon^2$ for all t.

The variable x_t is observed with measurement error, that is, in practice we observe the variable y_t, which is defined as

$$y_t = x_t + \eta_t, \tag{31}$$

where the measurement error η_t is independent and normally distributed with constant variance σ_η^2, that is $\eta_t \sim N(0, \sigma_\eta^2)$. Furthermore, ε_t and η_t are independent.

Derive an expression (in terms of the parameters δ, θ_1, σ_ε^2 and σ_η^2) for the following characteristics of the observed time series y_t:

a. the unconditional mean $\mu_y = \mathsf{E}[y_t]$,

b. the unconditional variance $\gamma_0(y) = \mathsf{E}[(y_t - \mathsf{E}[y_t])^2]$, and

c. the first-order autocorrelation $\rho_1(y) = \gamma_1(y)/\gamma_0(y)$, where $\gamma_1(y)$ is the first-order autocovariance of y_t, that is, $\gamma_1(y) = \mathsf{E}[(y_t - \mathsf{E}[y_t])(y_{t-1} - \mathsf{E}[y_{t-1}])]$.

Compare these to the corresponding characteristics of the underlying series x_t and interpret the differences (in terms of the effects of the measurement error η_t.

Solution

a. From the specification for y_t and x_t it follows straightforwardly that

$$\mathsf{E}[y_t] = \mathsf{E}[x_t + \eta_t] = \mathsf{E}[\delta + \theta_1 \varepsilon_{t-1} + \varepsilon_t + \eta_t] = \delta,$$

given that $\mathsf{E}[\varepsilon_t] = 0$ and $\mathsf{E}[\eta_t] = 0$.

b. For the unconditional variance $\gamma_0(y)$, we find

$$\begin{aligned}
\gamma_0(y) &= \mathsf{E}[(y_t - \mathsf{E}[y_t])^2] = \mathsf{E}[(y_t - \delta)^2] \\
&= \mathsf{E}[(\theta_1 \varepsilon_{t-1} + \varepsilon_t + \eta_t)^2] \\
&= \mathsf{E}[\theta_1^2 \varepsilon_{t-1}^2 + \varepsilon_t^2 + \eta_t^2] \\
&= (1 + \theta_1^2)\sigma_\varepsilon^2 + \sigma_\eta^2,
\end{aligned}$$

making use of the properties that (i) ε_t is a white noise series with $\mathsf{E}[\varepsilon_t^2] = \sigma_\varepsilon^2$; (ii) η_t is a white noise series with variance σ_η^2; and (iii) ε_t and η_t are independent.

c. For the first-order autovariance $\gamma_1(y)$, the same three properties imply that

$$\begin{aligned}
\gamma_1(y) &= \mathsf{E}[(y_t - \mathsf{E}[y_t])(y_{t-1} - \mathsf{E}[y_{t-1}])] \\
&= \mathsf{E}[(\theta_1\varepsilon_{t-1} + \varepsilon_t + \eta_t)(\theta_1\varepsilon_{t-2} + \varepsilon_{t-1} + \eta_{t-1})] \\
&= \mathsf{E}[\theta_1\varepsilon_{t-1}^2] \\
&= \theta_1\sigma_\varepsilon^2,
\end{aligned}$$

such that the first-order autocorrelation

$$\rho_1(y) = \gamma_1(y)/\gamma_0(y) = \frac{\theta_1\sigma_\varepsilon^2}{(1+\theta_1^2)\sigma_\varepsilon^2 + \sigma_\eta^2} = \frac{\theta_1}{(1+\theta_1^2) + \sigma_\eta^2/\sigma_\varepsilon^2}.$$

For the 'clean' MA(1) series x_t, we have

$$\begin{aligned}
\mathsf{E}[x_t] &= \delta, \\
\gamma_0(x) &= (1+\theta_1^2)\sigma_\varepsilon^2, \\
\gamma_1(x) &= \theta_1\sigma_\varepsilon^2, \\
\rho_1(x) &= \frac{\theta_1}{(1+\theta_1^2)}.
\end{aligned}$$

Hence, the observed series y_t has (i) the same unconditional mean as x_t, (ii) an unconditional variance that is larger by σ_η^2 (due to the measurement error), (iii) the same first-order autocovariance, and (iv) a smaller first-order autocorrelation (which follows from (ii) and (iii)).

Exercise 3.11

The Excel file `USUNINSCE.xlsx` contains quarterly observations of US initial claims for unemployment insurance[1] over the period 1974Q1–2013Q4 (160 observations). The series is referred to as 'initial claims' below. A dummy variable DREC is also included, which takes the value 1 during recession periods and 0 during expansions.[2]

a. Is the series of initial claims normally distributed?

b. Compute the first 40 empirical autocorrelations and partial autocorrelations for the initial claims series. Examine the significance of the autocorrelations. What kind of AR, MA or ARMA model do the (partial) autocorrelations suggest?

c. Estimate AR(p) and MA(q) models for $p = 0, \ldots, 4$ and $q = 1, \ldots, 4$ for the quarterly initial claims series, and record the values of the Akaike Information Criterion and the Schwarz Information Criterion. Use the sample period 1975Q1-2013Q4 to make sure that you use the same number of observations for all models. (Include a constant in all models.) Which model do these criteria indicate as the preferred choice?

d. Estimate an AR(2) model for the quarterly initial claims series using the complete sample period 1974Q1-2013Q4. What are the roots of the AR(2) polynomial? Compute the values of the first 20 'implied' autocorrelations (that is, the values of the autocorrelations of an AR(2) process for the estimates $\hat{\phi}_1$ and $\hat{\phi}_2$ obtained here) and compare these with the empirical autocorrelations. How closely do they match?

e. Inspect the properties of the time series of residuals from the estimated AR(2) model. In particular, address the following questions:

   i) Are the residuals normally distributed?

   ii) What are the autocorrelation properties of the residuals, and what does this imply for the adequacy of the estimated AR(2) model?

[1]This variable, collected by the US Department of Labor is one of the ten components of The Conference Board's Leading Economic Index, which is one of the most important indicators of the (future) state of the US economy; see `http://www.conference-board.org/data/bcicountry.cfm?cid=1` for more information. Also see `https://www.dol.gov/ui/data.pdf`.

[2]Based on the turning points provided at `http://www.nber.org/cycles/`

iii) What are the autocorrelation properties of the squared residuals? What does this imply for the assumption of homoskedasticity of the shocks in the model?

f. Re-estimate the AR(2) model for the quarterly initial claims series using the sample period 1974Q1–1993Q4 (T =80 observations). Construct one-step a-head point forecasts for the period 1994Q1–2013Q4 (80 observations). Evaluate the accuracy of these point forecasts in 'absolute' terms, by considering their a) unbiasedness, b) accuracy, and c) efficiency. Also evaluate the 'relative' accuracy of the point forecasts obtained from the AR(2) model by comparing them with so-called 'random walk' forecasts, where $\hat{y}_{T+1|T} = y_T$, that is, the forecast is simply the observed value in the previous quarter.

Solution

a. The initial claims show pronounced deviations from normality: Kurtosis is equal to 4.79 and thus substantially larger than the normal value of 3, while skewness is equal to 1.36 instead of 0. These two features lead to rejection of normality (at the 1% significance level) when we use the Jarque-Bera statistic, for which we find a value of 70.48 with p-value of 0.000.

The histogram in Figure 1 (left panel) clearly shows the skewness in the time series, with an almost left tail in the empirical distribution and more observations than expected under normality in the extreme right tail. The positive skewness might be a reason to consider the log of initial claims, but this does not resolve the issue: skewness declines but remains substantially positive at a value of 0.89.

The first differences also do not seem normality distributed, with skewness equal to 1.10 and kurtosis equal to 8.80. Note that the histogram in Figure 1 (right panel) shows that the large kurtosis is not only due to 'fat tails' (i.e. more observations in the tails than expected) but also due to a 'high peak' (i.e. more observations clustered around the mean than expected under normality).

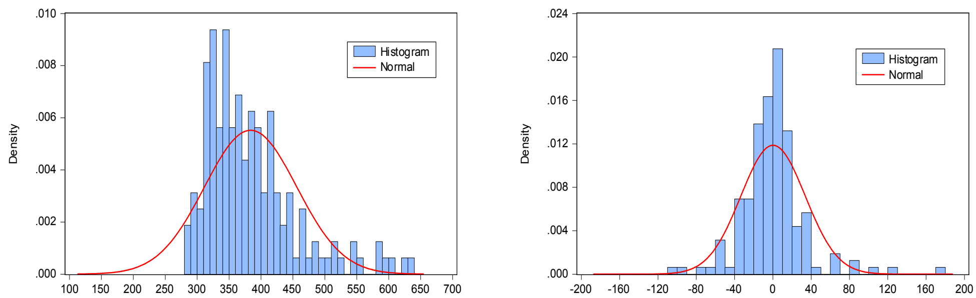


Figure 1: Histogram of level (left panel) and first differences (right panel) of initial claims

b. The empirical (partial) autocorrelation functions shown in Figure 2 have patterns that suggest a fairly straightforward interpretation: the first- and second-order partial autocorrelations are significantly different from zero, while higher order partial autocorrelations are not. The autocorrelations decline towards zero in an oscillatory manner – positive autocorrelations for orders 1-9 are followed by a sequence of negative autocorrelations until order 20, again followed by positive autocorrelations. These patterns correspond quite nicely with the theoretical (partial) autocorrelations of an AR(2) model with complex roots.

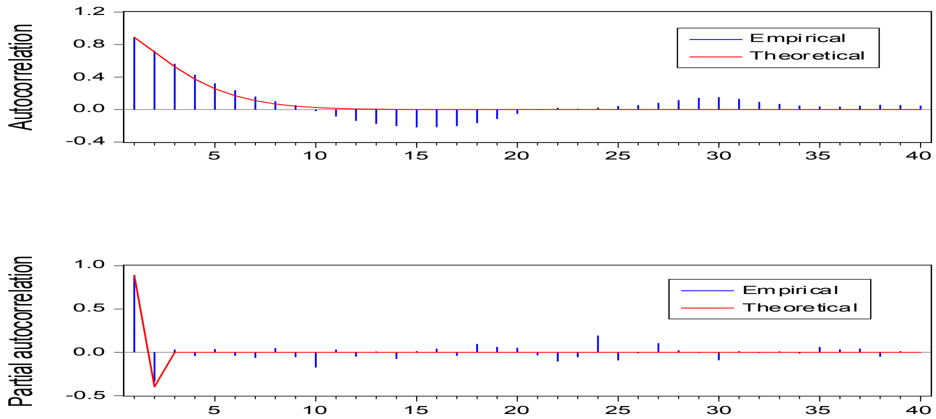


Figure 2: Empirical and theoretical [AR(2)] autocorrelations

c. The values of the AIC and SIC for AR(p) and MA(q) models for $p = 0, \ldots, 4$ and $q = 1, \ldots, 4$ are as follows:

	AIC	SIC		AIC	SIC
AR(0)	11.405	11.424			
AR(1)	9.744	9.783	MA(1)	10.483	10.542
AR(2)	9.592	9.651	MA(2)	10.036	10.114
AR(3)	9.605	9.683	MA(3)	9.829	9.927
AR(4)	9.618	9.715	MA(4)	9.726	9.843

Both AIC and SIC achieve their minimum value for the AR(2) model.

d. Estimating an AR(2) model

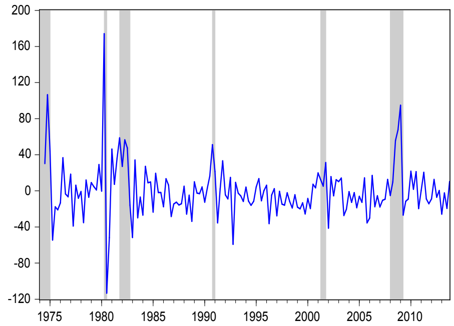
$$y_t = \alpha + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \varepsilon_t$$

for the time series of initial claims y_t by least squares gives estimates $\hat{\phi}_1 = 1.243\,(0.074)$ and $\hat{\phi}_2 = -0.398\,(0.073)$, with standard errors in parentheses. Note that these estimates are perfectly in line with the empirical (partial) autocorrelations computed above.

The AR(2) polynomial $\phi_2(L)$ has a pair of complex roots equal to $1.564 \pm 0.264i$. The modulus of these roots is 1.586 – hence the roots are outside the unit circle and the AR(2) model is stationary.

Figure 2 shows the theoretical autocorrelations ρ_k for $k = 1, 2, \ldots, 40$ for the AR(2) model implied by the estimates of ϕ_1 and ϕ_2 obtained for the initial claims series. The partial autocorrelations are shown as. Comparing these 'implied' values with the empirical (partial) autocorrelations (also shown in Figure 2), we observe that the non-zero first- and second-order partial autocorrelations are matched very well. In terms of the autocorrelations, the theoretical values are close to the empirical ones up to order 10. The correspondence is much less close for higher-order autocorrelations. In particular, we observe that the theoretical autocorrelations are all positive (and are essentially zero beyond order 10), not matching the (cyclical) pattern in the empirical autocorrelations (with negative values for lags 9-20.

e. The residuals from the AR(2) model have skewness equal to 1.47 and kurtosis equal to 11.34. The Jarque-Bera test statistic has a value 514.9, with p-value 0.000. Hence, we reject the null hypothesis that the residuals are normally distributed at the 5% significance level. The residuals are shown in Figure 3. We observe a small number of extremely large residuals, with values of 106.6 in 1974Q4, 174.1 in 1980Q2, -113.3 in 1980Q3, and 94.9 in 2009Q1. To a large extent these observations are responsible for the non-normality: if we leave these four residuals out of consideration, skewness and kurtosis are substantially lower at 0.42 and 3.73, respectively.



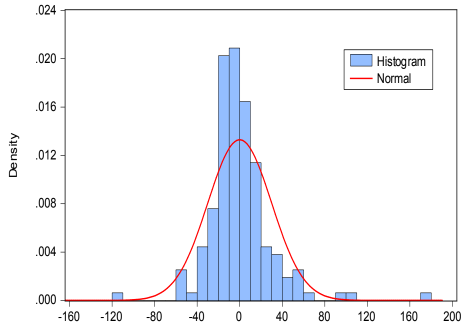


Figure 3: Residuals from AR(2) model

None of the first 20 autocorrelations of the residuals, $\hat{\rho}_k$ are (individually) significantly different from zero, suggesting that the AR(2) model adequately captures the autocorrelation properties of the initial claims series.

We find a highly significant first-order autocorrelation of the squared residuals,

$\hat{\rho}_1 = 0.335$ with a p-value of 0.000. Hence, it seems that the assumption of homoskedasticity is not appropriate. Obviously, it might be that this first-order autocorrelation is inflated by the large residuals in 1980Q2 and 1980Q3. However, leaving those two observations out of consideration, we find $\hat{\rho}_1 = 0.244$, which still is highly significant.

f. The one-step forecasts obtained from the AR(2) model for the period 1994Q1–2013Q4 are shown below. The forecasts seem to track the actual values quite well. The main difference obviously is that the forecasts are 'shifted' forward compared to the actual values. This makes sense given that the point forecasts are formed as $\hat{y}_{T+1|T} = \hat{\phi}_1 y_T + \hat{\phi}_2 y_{T-1}$, with the estimates based on the sample 1974Q1-1993Q4 being equal to $\hat{\phi}_1 = 1.14$ and $\hat{\phi}_2 = -0.36$. In addition, the only actual value outside the 95% interval forecast occurs during the recent crisis in 2009Q1.

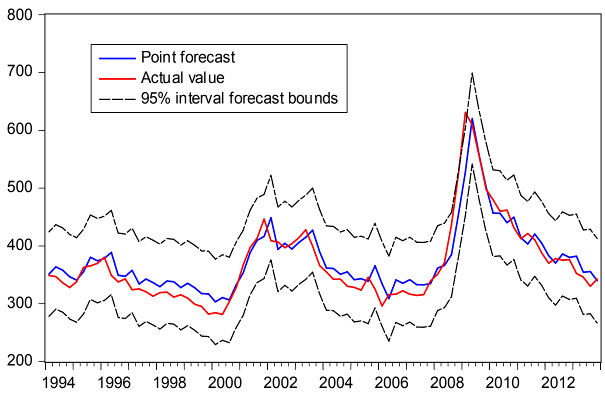


Figure 4: Point forecasts from AR(2) model with actual values
and lower and upper bounds of 95% interval forecast

The mean of the forecast errors $e_{\text{AR2},t+1|t}$ is equal to -7.02, with standard error 2.56. Hence, the mean forecast error is significantly different from zero even at the 1% level, such that we may reject the null hypothesis that the AR(2)-forecasts are unbiased.

The mean squared prediction error (MSPE) is equal to 566.3. This is substantially lower than the variance of the AR(2) residuals over the period 1974Q1-1993Q4, which is equal to 1359.4. This variance is heavily affected by the large residuals in 1974Q4, 1980Q2 and 1980Q3 though: leaving these three observations out of consideration reduces the variance of the residuals to 699. The MSPE is also much smaller than the unconditional variance of the initial claims over the forecast period (which is equal to the staggering value of 4690). Hence, the AR(2) forecasts are accurate, at least in the sense that they perform much better than a constant forecast equal to the mean of y over the forecast period.

When we estimate the Mincer-Zarnowitz regression

$$y_{t+1} = \beta_0 + \beta_1 \hat{y}_{t+1|t} + \eta_{t+1},$$

we find estimates $\hat{\beta}_0 = -68.55\,(16.11)$ and $\hat{\beta}_1 = 1.164\,(0.042)$. The intercept is significantly different from 0, while the estimate of β_1 is significantly different from 1. This suggests that the forecasts are not 'efficient' and do not make the best possible use of information from the past initial claims. At the same time, the R^2 of the MZ-regression is equal to 0.906, showing that the forecasts 'explain' a substantial part of the variation in the actual series, which is in agreement with Figure 4.

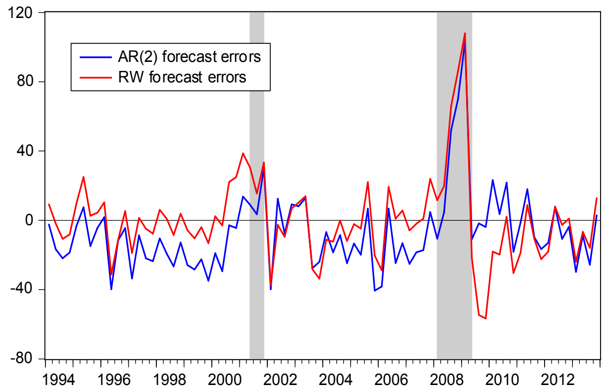


Figure 5: Forecast errors from AR(2) model and random walk (RW)

The MSPE of 'random walk' forecasts, where $\hat{y}_{T+1|T} = y_T$, is equal to 622.2, which seems considerably larger than the MSPE of the AR(2) forecasts. In order to test whether the difference is statistically significant, we may use the Diebold-Mariano test statistic based on the 'loss differential' $d_{t+1} = e^2_{\mathrm{AR2},t+1|t} - e^2_{\mathrm{RW},t+1|t}$, where $e_{\mathrm{RW},t+1|t}$ is the forecast error of the random walk forecast. The value of the DM statistic is equal to -0.67, with a (two-sided) p-value equal to 0.50, so that we do not reject the null hypothesis of equal forecast accuracy at the usual significance levels.

At first sight this conclusion might be surprising, given the substantial difference in MSPE values for the AR(2) and RW forecasts. The forecast errors in Figure 5 and the loss differential d_{t+1} in Figure 6 help to understand this result though. We observe that both forecast errors show substantial variation (also leading to a large variance in d_{t+1}), while in addition the RW forecasts often are more accurate than the AR(2) forecasts. It seems that the AR(2) model is mostly useful (relative to the RW) during recessions (and immediately before and/or after these periods) given that the loss differential takes the largest negative values during the years 2000-2001 and 2008-2009.

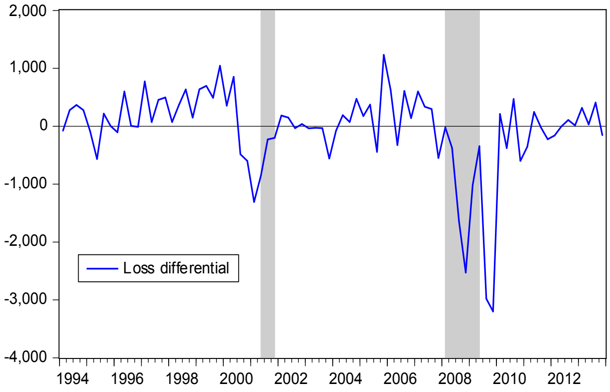


Figure 6: Loss differential: $d_{t+1} = e^2_{\text{AR2},t+1|t} - e^2_{\text{RW},t+1|t}$.

Exercise 3.12
The Excel file `USMSSALES.xlsx` contains quarterly observations on the natural logarithm of US manufacturing sales[3] over the period 1959Q1–2015Q4. A dummy variable DREC is also included, which takes the value 1 during recession periods and 0 during expansions.[4]

Create a time series of quarterly growth rates of sales by taking the first differences of the log sales series (and multiplying by 100 to put the growth rates in terms of percentage points). Estimate an AR(3) model for these quarterly growth rates using the sample period 1960Q1–1979Q4 (T =80 observations). Construct one-step ahead point forecasts for the sample period 1980Q1–2015Q4 (144 observations). Evaluate the accuracy of these point forecasts in 'absolute' terms, by considering their a) unbiasedness, b) accuracy, and c) efficiency. Also evaluate the 'relative' accuracy of the point forecasts obtained from the AR(3) model by comparing them with so-called 'random walk' forecasts, where the forecast is simply the observed value in the previous quarter, that is, $\hat{y}_{t+1|t} = y_t$.

Solution
The one-step forecasts obtained from the AR(3) model for the period 1980Q1–2015Q4 are shown in Figure 7 below. The forecasts seem quite close to the actual values in general, except during recession periods (1991, 2001, and 2008-9) when the actual growth rates are negative (and often large in magnitude), while the forecasts remain positive (or only become moderately negative as in 2009).

The mean of the forecast errors $e_{\mathrm{AR3},t+1|t}$ is equal to -0.183, with standard error 0.097. Hence, the mean forecast error is significantly different from zero at the 10% level but not at the 5% level. Hence, depending on the significance level use, we reject / do not reject the null hypothesis that the AR(3)-forecasts are unbiased. The forecasts are too high on average.

[3]This variable is one of the four components of The Conference Board's Coincident Economic Index, which is one of the most important indicators of the state of the US economy; see `http://www.conference-board.org/data/bcicountry.cfm?cid=1` for more information.
[4]Based on the turning points provided at `http://www.nber.org/cycles/`

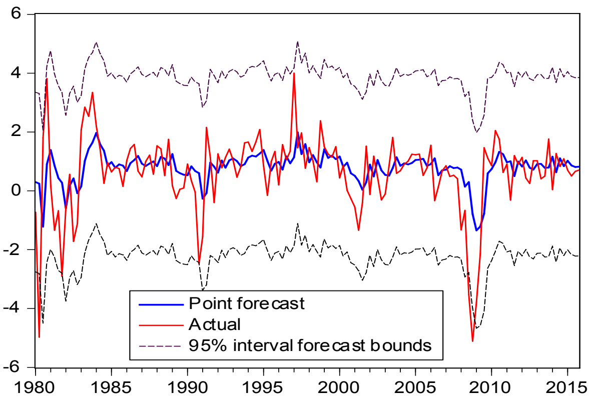


Figure 7: Point forecasts obtained from AR(3) model, together with actual values and lower and upper bounds of 95% interval forecast

The mean squared prediction error (MSPE) is equal to 1.39, which is more than 20% smaller than the variance of the quarterly growth rates over the forecast period (at 1.80). Hence, the AR(3) forecasts are reasonably accurate (in the sense that at least they perform better than a constant forecast equal to the mean of y over the forecast period).

When we estimate the Mincer-Zarnowitz regression

$$y_{t+1} = \beta_0 + \beta_1 \hat{y}_{t+1|t} + \eta_{t+1},$$

we find estimates $\hat{\beta}_0 = -0.434\,(0.178)$ and $\hat{\beta}_1 = 1.318\,(0.188)$. The intercept is significantly different from 0, confirming that the AR(3) forecasts are biased. The estimate of β_1 is not significantly different from 1. The R^2 of the MZ-regression is equal to 0.256, suggesting that the forecasts 'explain' only a relatively small part of the variation in the actual growth rates.

The MSPE of 'random walk' forecasts, where $\hat{y}_{T+1|T} = y_T$, is equal to 1.757, which obviously is substantially larger than the MSPE of the AR(3) forecasts. In order to test whether the difference is statistically significant, we may use the Diebold-Mariano test statistic based on the 'loss differential' $d_{t+1} = e^2_{\text{AR3},t+1|t} - e^2_{\text{RW},t+1|t}$, where $e_{\text{RW},t+1|t}$ is the forecast error of the random walk forecast. The value of the DM statistic is equal to -1.102, with a (one-sided) p-value equal to 0.135, so that we do not reject the null hypothesis of equal forecast accuracy at conventional significance levels.

Exercise 3.13

The time series y_t is generated according to an AR(1) process:

$$y_t = \phi_1 y_{t-1} + \varepsilon_t, \tag{32}$$

where $|\phi_1| < 1$, and ε_t is a white noise series with $\mathsf{E}[\varepsilon_t] = 0$, $\mathsf{E}[\varepsilon_t^2] = \sigma_\varepsilon^2$ for all t, and $\mathsf{E}[\varepsilon_t \varepsilon_s] = 0$ for all $s \neq t$.

Suppose that a researcher uses a 'subset' AR(2) model[5]

$$y_t = \phi_2^* y_{t-2} + \eta_t, \tag{33}$$

where she sets the parameter ϕ_2^* equal to the square of ϕ_1 in equation (32), that is, $\phi_2^* = \phi_1^2$. Furthermore, the researcher assumes that η_t is a white noise sequence with (among others) mean equal to zero. (Note that this assumption may be incorrect!)

The researcher uses the subset AR(2) model in equation (33) to compute one-step ahead point forecasts for y_t at $t = T+1, T+2, T+3, \dots$ using a quadratic loss function (that is, she chooses the point forecast in order to minimize the expected mean squared prediction error). The resulting point forecast for $T+j+1$ ($j = 0, 1, 2, 3, \dots$) is denoted as $\hat{y}_{T+j+1|T+j}$, with corresponding forecast error $e_{T+j+1|T+j} = y_{T+j+1} - \hat{y}_{T+j-1|T+j}$.

Answer the following questions:

a. Show that the one-step ahead point forecasts obtained from the subset AR(2) model in equation (33) are (unconditionally) unbiased, that is $\mathsf{E}[e_{T+j+1|T+j}] = 0$ for all $j > 0$.

b. Derive an expression for the unconditional variance of the forecast errors $e_{T+j+1|T+j}$ in terms of the parameters σ_ε^2 and ϕ_1.

c. Derive an expression for the k-th order autocorrelation $\rho_{k,e}$ of the one-step ahead forecast errors (that is, $\rho_{k,e}$ is the correlation between $e_{T+j+1|T+j}$ and $e_{T+j+1-k|T+j-k}$) for $k = 1, 2, 3, \dots$. Interpret your results.

[5] A subset AR(p) model includes only a subset of the lags $1, 2, \dots, p$. Put differently, the coefficients of some lags are equal to zero.

Solution

a. Given that the researcher uses a quadratic loss function, the one-step ahead point forecast of y_{T+j+1}, $j > 0$, is equal to the conditional expectation in the subset AR(2) model as given in equation (33), that is,

$$\hat{y}_{T+j+1|T+j} = \mathsf{E}[y_{T+j+1}|\mathcal{Y}_{T+j}] = \phi_1^2 y_{T+j-1},$$

where $\mathcal{Y}_{T+j}$ denotes the history of the time series until $t = T + j$, that is, $\mathcal{Y}_{T+j} = \{y_{T+j}, y_{T+j-1}, \dots\}$, and where we have used the assumption of the researcher that $\mathsf{E}[\eta_{T+j+1}|\mathcal{Y}_{T+j}] = 0$.

Note that the AR(1) data-generating process (DGP) given in equation (32) implies that we can write y_t as

$$y_t = \phi_1 y_{t-1} + \varepsilon_t = \phi_1(\phi_1 y_{t-2} + \varepsilon_{t-1}) + \varepsilon_t = \phi_1^2 y_{t-2} + \varepsilon_t + \phi_1 \varepsilon_{t-1}.$$

Hence, the one-step forecast error for the subset AR(2) model is equal to

$$\begin{aligned} e_{T+j+1|T+j} &= y_{T+j+1} - \hat{y}_{T+j+1|T+j} \\ &= \phi_1^2 y_{T+j-1} + \varepsilon_{T+j+1} + \phi_1 \varepsilon_{T+j} - \phi_1^2 y_{T+j-1} \\ &= \varepsilon_{T+j+1} + \phi_1 \varepsilon_{T+j}. \end{aligned} \tag{34}$$

Given that $\mathsf{E}[\varepsilon_t] = 0$ for all t, it follows immediately that $\mathsf{E}[e_{T+j+1|T+j}] = 0$, such that the forecasts indeed are unconditionally unbiased.

b. Observing that the one-step ahead forecast error for the subset AR(2) model is equal to $e_{T+j+1|T+j} = \varepsilon_{T+j+1} + \phi_1 \varepsilon_{T+j}$, it follows that its unconditional variance is equal to

$$\begin{aligned} \mathsf{E}[e_{T+j+1|T+j}^2] &= \mathsf{E}[(\varepsilon_{T+j+1} + \phi_1 \varepsilon_{T+j})^2] \\ &= \mathsf{E}[\varepsilon_{T+j+1}^2 + 2\phi_1 \varepsilon_{T+j+1}\varepsilon_{T+j} + \phi_1^2 \varepsilon_{T+j}^2] \\ &= (1 + \phi_1^2)\sigma_\varepsilon^2, \end{aligned}$$

where we have used the fact that ε_t is white noise with variance σ_ε^2.

c. Note that the one-step ahead forecast error for the subset AR(2) model $e_{T+j+1|T+j} = \varepsilon_{T+j+1} + \phi_1 \varepsilon_{T+j}$ essentially has a first-order moving average [MA(1)] structure. Hence, it follows that it only has a non-zero first-order autocorrelation, because $e_{T+j+1|T+j}$ and $e_{T+j+1-k|T+j-k}$ have no terms in common whenever $k > 1$.

For the first-order autocovariance of $e_{T+j+1|T+j}$ we find

$$\begin{aligned}
\mathsf{E}[e_{T+j+1|T+j}e_{T+j|T+j-1}] &= \mathsf{E}[(\varepsilon_{T+j+1} + \phi_1\varepsilon_{T+j})(\varepsilon_{T+j} + \phi_1\varepsilon_{T+j-1})] \\
&= \mathsf{E}[\varepsilon_{T+j+1}\varepsilon_{T+j} + \phi_1\varepsilon_{T+j+1}\varepsilon_{T+j-1} + \phi_1\varepsilon_{T+j}^2 + \phi_1^2\varepsilon_{T+j}\varepsilon_{T+j-1}] \\
&= 0 + 0 + \phi_1\sigma_\varepsilon^2 + 0 \\
&= \phi_1\sigma_\varepsilon^2,
\end{aligned}$$

again using the fact that ε_t is a white noise process.

Dividing this by the variance of the one-step ahead forecast error gives the first-order autocorrelation

$$\rho_{e,1} = \frac{\phi_1\sigma_\varepsilon^2}{(1+\phi_1^2)\sigma_\varepsilon^2} = \frac{\phi_1}{(1+\phi_1^2)}.$$

(International) Bachelor Econometrics and Operations Research
FEB23001(X)-19 Tijdreeksanalyse / Time Series Analysis

Exercises - Trends (Chapter 4)

Exercise 4.1
Consider the AR(3) model

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \phi_3 y_{t-3} + \varepsilon_t. \tag{1}$$

a. Show that when $\phi_1 + \phi_2 + \phi_3 = 1$, this AR(3) model can be written as an AR(2) model for $(1-L)y_t$. Rewrite the AR(3) model such that this parameter restriction can be tested using a $t-$test.
b. Show that when $\phi_1 - \phi_2 + \phi_3 = -1$, this AR(3) model can be written as an AR(2) model for $(1+L)y_t$. Rewrite the AR(3) model such that this parameter restriction can be tested using a $t-$test.

Solution
a. Substituting $\phi_1 = 1 - \phi_2 - \phi_3$, it follows that

$$\begin{aligned} y_t &= (1 - \phi_2 - \phi_3)y_{t-1} + \phi_2 y_{t-2} + \phi_3 y_{t-3} + \varepsilon_t \\ &= y_{t-1} - (\phi_2 + \phi_3)(y_{t-1} - y_{t-2}) - \phi_3(y_{t-2} - y_{t-3}) + \varepsilon_t, \end{aligned}$$

or

$$(1-L)y_t = -(\phi_2 + \phi_3)(1-L)y_{t-1} - \phi_3(1-L)y_{t-2} + \varepsilon_t. \tag{2}$$

An alternative way to see this is by substituting the restriction $\phi_1 = 1 - \phi_2 - \phi_3$ in the AR(3) polynomial

$$\begin{aligned} (1 - \phi_1 L - \phi_2 L^2 - \phi_3 L^3) &= (1 - (1 - \phi_2 - \phi_3)L - \phi_2 L^2 - \phi_3 L^3) \\ &= (1 - L) + (\phi_2 + \phi_3)(L - L^2) + \phi_3(L^2 - L^3) \\ &= (1 - L)(1 + (\phi_2 + \phi_3)L + \phi_3 L^2). \end{aligned}$$

To test the restriction $\phi_1 + \phi_2 + \phi_3 = 1$, it is convenient to rewrite the model as

$$\Delta_1 y_t = (\phi_1 + \phi_2 + \phi_3 - 1)y_{t-1} - (\phi_2 + \phi_3)\Delta_1 y_{t-1} - \phi_3 \Delta_1 y_{t-2} + \varepsilon_t. \tag{3}$$

b. Notice that the restriction $\phi_1 - \phi_2 + \phi_3 = -1$ implies that the AR(3)-polynomial $\phi_3(L)$ has a root equal to -1, in the sense that $z = -1$ is a solution to

$$\phi_3(z) = 1 - \phi_1 z - \phi_2 z^2 - \phi_3 z^3 = 0.$$

[To see that a root equal to -1 leads to a cycle of two periods in y_t, consider the simple case of an AR(1) model

$$y_t = \phi_1 y_{t-1} + \varepsilon_t.$$

When $\phi_1 = -1$, this gives $y_t = y_{t-2} + \varepsilon_t - \varepsilon_{t-1}$.]

Substituting $\phi_1 = \phi_2 - \phi_3 - 1$, it follows that

$$\begin{aligned} y_t &= (\phi_2 - \phi_3 - 1)y_{t-1} + \phi_2 y_{t-2} + \phi_3 y_{t-3} + \varepsilon_t \\ &= -y_{t-1} + (\phi_2 - \phi_3)(y_{t-1} + y_{t-2}) + \phi_3(y_{t-2} + y_{t-3}) + \varepsilon_t. \end{aligned}$$

or

$$(1+L)y_t = (\phi_2 - \phi_3)(1+L)y_{t-1} + \phi_3(1+L)y_{t-2} + \varepsilon_t, \tag{4}$$

To test the restriction $\phi_1 - \phi_2 + \phi_3 + 1 = 0$, it is convenient to rewrite the model as

$$\begin{aligned} (1+L)y_t = {} & (\phi_1 - \phi_2 + \phi_3 + 1)y_{t-1} \\ & + (\phi_2 - \phi_3)(1+L)y_{t-1} + \phi_3(1+L)y_{t-2} + \varepsilon_t. \end{aligned} \tag{5}$$

Exercise 4.2

Consider a time series y_t which experiences a permanent change in mean, as described by the model

$$y_t = \delta D_{t,[T/2]} + \varepsilon_t, \qquad t = 1, 2, \ldots, T, \tag{6}$$

where $D_{t,[T/2]} = 0$ for $t = 1, \ldots, T/2$ and $D_{t,[T/2]} = 1$ for $t = T/2+1, \ldots, T$, and ε_t is a white noise series with variance σ^2.

a. Show that asymptotically ($T \to \infty$) the following holds for the sample mean $\bar{y} = \frac{1}{T}\sum_{t=1}^{T} y_t$, the sample variance $\hat{\gamma}_0 = \frac{1}{T}\sum_{t=1}^{T}(y_t - \bar{y})^2$, and the k-th order autocovariance $\hat{\gamma}_k = \frac{1}{T}\sum_{t=1}^{T}(y_t - \bar{y})(y_{t-k} - \bar{y})$ for any $k > 0$:

$$\bar{y} \to \delta/2, \quad \hat{\gamma}_0 \to \sigma^2 + \delta^2/4, \quad \text{and} \quad \hat{\gamma}_k \to \delta^2/4.$$

b. What does this imply for the behavior of the Dickey-Fuller test for a unit root, when applied to the series y_t, when $|\delta|/\sigma^2$ becomes very large?

Solution

a. For the sample mean $\bar{y}$, we find that as $T \to \infty$

$$\begin{aligned}
\bar{y} &= \frac{1}{T}\sum_{t=1}^{T} y_t \\
&= \frac{1}{T}\sum_{t=1}^{T}(\delta D_{t,[T/2]} + \varepsilon_t) \\
&= \delta \underbrace{\frac{1}{T}\sum_{t=1}^{T} D_{t,[T/2]}}_{\to 1/2} + \underbrace{\frac{1}{T}\sum_{t=1}^{T} \varepsilon_t}_{\to \mathsf{E}[\varepsilon_t] = 0} \to \delta/2,
\end{aligned} \tag{7}$$

For the sample variance $\hat{\gamma}_0$, the following holds as $T \to \infty$

$$\begin{aligned}
\hat{\gamma}_0 &= \frac{1}{T}\sum_{t=1}^{T}(y_t - \bar{y})^2 = \frac{1}{T}\sum_{t=1}^{T}(y_t^2 - 2y_t\bar{y} + \bar{y}^2) \\
&= \frac{1}{T}\sum_{t=1}^{T} y_t^2 \; - \; \bar{y}^2
\end{aligned} \tag{8}$$

and because

$$\begin{aligned}
\frac{1}{T}\sum_{t=1}^{T} y_t^2 &= \frac{1}{T}\sum_{t=1}^{T}(\delta D_{t,[T/2]} + \varepsilon_t)^2 \\
&= \underbrace{\frac{1}{T}\sum_{t=1}^{T} \delta^2 D_{t,[T/2]}^2}_{\to \delta^2/2} + \underbrace{\frac{1}{T}\sum_{t=1}^{T} 2\delta D_{t,[T/2]}\varepsilon_t}_{\to 0} + \underbrace{\frac{1}{T}\sum_{t=1}^{T} \varepsilon_t^2}_{\to \sigma^2}
\end{aligned}$$

we have that $\hat{\gamma}_0 \to \delta^2/2 + \sigma^2 - (\delta/2)^2 = \delta^2/4 + \sigma^2$.

For the 1st-order autocovariance $\hat{\gamma}_1$, the following holds as $T \to \infty$

$$\begin{aligned}\hat{\gamma}_1 &= \frac{1}{T}\sum_{t=1}^{T}(y_t - \bar{y})(y_{t-1} - \bar{y}) = \frac{1}{T}\sum_{t=1}^{T}(y_t y_{t-1} - y_t\bar{y} - y_{t-1}\bar{y} + \bar{y}^2) \\ &= \frac{1}{T}\sum_{t=1}^{T} y_t y_{t-1} \ - \ \bar{y}^2 \qquad (9)\end{aligned}$$

and because

$$\begin{aligned}\frac{1}{T}\sum_{t=1}^{T} y_t y_{t-1} &= \frac{1}{T}\sum_{t=1}^{T}(\delta D_{t,[T/2]} + \varepsilon_t)(\delta D_{t-1,[T/2]} + \varepsilon_{t-1}) \\ &= \underbrace{\frac{1}{T}\sum_{t=1}^{T}\delta^2 D_{t,[T/2]} D_{t-1,[T/2]}}_{\to \delta^2/2} \ + \ \underbrace{\frac{1}{T}\sum_{t=1}^{T}(\delta D_{t,[T/2]}\varepsilon_{t-1} + \delta D_{t-1,[T/2]}\varepsilon_t)}_{\to 0} \ + \ \underbrace{\frac{1}{T}\sum_{t=1}^{T}\varepsilon_t\varepsilon_{t-1}}_{\to 0}\end{aligned}$$

we have that $\hat{\gamma}_1 \to \delta^2/2 - (\delta/2)^2 = \delta^2/4$.

For the k-th order autocovariance $\hat{\gamma}_k$ with $k > 1$, the same derivations as above can be applied with y_{t-1} replaced with y_{t-k}, with the same outcome, namely that $\hat{\gamma}_k \to \delta^2/4$ as $T \to \infty$.

b. Because $\hat{\gamma}_0 \to \sigma^2 + \delta^2/4$ and $\hat{\gamma}_k \to \delta^2/4$, we have that

$$\hat{\rho}_k = \frac{\delta^2/4}{\sigma^2 + \delta^2/4} \to 1,$$

when $|\delta|/\sigma^2$ becomes very large. All empirical autocorrelations approach 1 as the sample size T and the magnitude of the mean shift δ become large. Hence, the EACF of y_t looks just like what we would find for a time series with a unit root. For that reason, the DF test will typically **not** reject the null hypothesis of a unit root.

Exercise 4.3

Suppose that y_t is a trending time series with data generating process (DGP)

$$y_t = \alpha + \beta t + \eta_t, \tag{10}$$

with $\beta \neq 0$ for $t = 1, \ldots, T-1$, and $\eta_t \sim N(0, \sigma^2)$. At time T there is a sudden level shift in the time series and the DGP becomes

$$y_t = \alpha + \alpha^* + \beta t + \eta_t,$$

for $t = T, T+1, \ldots$, with $\alpha^* \neq 0$.

Suppose that we are unaware of this level shift and want to forecast future values of y_t at time T. We use the following two models to describe the trending time series

$$\text{Model A: (Deterministic trend)} \qquad y_t = \alpha + \beta t + e_t, \tag{11}$$

$$\text{Model B: (Stochastic trend)} \qquad \Delta y_t = \beta + u_t, \tag{12}$$

where Δ is the first-difference operator, that is, $\Delta y_t = y_t - y_{t-1}$. Assume that we know the true values of α and β from the DGP that holds for y_t for $t = 1, \ldots, T-1$ as given in (10), so that we do not have to estimate the parameters in these models.

a. Derive the 1-step and 2-step ahead point forecasts from models A and B made at time T (that is, $\hat{y}_{T+1|T}$ and $\hat{y}_{T+2|T}$). Compute the corresponding 1-step and 2-step ahead forecast errors. Which model provides the best forecasts? Explain this result!

b. Suppose now that the level shift is temporary and only lasts for 1 time period. In other words, the true DGP is

$$\begin{aligned} y_t &= \alpha + \beta t + \eta_t && \text{for } t = 1, \ldots, T-1, \\ y_t &= \alpha + \alpha^* + \beta t + \eta_t && \text{for } t = T, \\ y_t &= \alpha + \beta t + \eta_t && \text{for } t = T+1, T+2, \ldots. \end{aligned}$$

Again we use models A and B to construct 1-step and 2-step ahead forecasts at period T. Which of the two models do you prefer in this case? Motivate your answer!

Solution

a. Model A would give the following forecasts:

$$\begin{aligned} \hat{y}_{T+1|T,A} &= \mathsf{E}[y_{T+1}|\mathcal{Y}_T] = \alpha + \beta(T+1), \\ \hat{y}_{T+2|T,A} &= \mathsf{E}[y_{T+2}|\mathcal{Y}_T] = \alpha + \beta(T+2), \end{aligned}$$

while Model B (which says $y_t = \beta + y_{t-1} + u_t$) would give forecasts

$$\begin{aligned}
\hat{y}_{T+1|T,B} &= \mathsf{E}[y_{T+1}|\mathcal{Y}_T] = \beta + y_T, \\
\hat{y}_{T+2|T,B} &= \mathsf{E}[y_{T+2}|\mathcal{Y}_T] = \beta + \mathsf{E}[y_{T+1}|\mathcal{Y}_T] = 2\beta + y_T.
\end{aligned}$$

As the realizations are equal to

$$\begin{aligned}
y_{T+1} &= \alpha + \alpha^* + \beta(T+1) + \eta_{T+1}, \\
y_{T+2} &= \alpha + \alpha^* + \beta(T+2) + \eta_{T+2},
\end{aligned}$$

the forecast errors for Model A are given by

$$\begin{aligned}
e_{T+1|T,A} &= \alpha^* + \eta_{T+1}, \\
e_{T+2|T,A} &= \alpha^* + \eta_{T+2}.
\end{aligned}$$

Notice that the realizations can also be written as

$$\begin{aligned}
y_{T+1} &= (\alpha + \alpha^* + \beta T) + \beta + \eta_{T+1} \\
&= (y_T - \eta_T) + \beta + \eta_{T+1}, \\
y_{T+2} &= (\alpha + \alpha^* + \beta(T+1)) + \beta + \eta_{T+2} \\
&= (y_{T+1} - \eta_{T+1}) + \beta + \eta_{T+2} = y_T - \eta_T + 2\beta + \eta_{T+2},
\end{aligned}$$

such that for Model B we have

$$\begin{aligned}
e_{T+1|T,B} &= \eta_{T+1} - \eta_T, \\
e_{T+2|T,B} &= \eta_{T+2} - \eta_T.
\end{aligned}$$

Hence, the variances of the one- and two-step ahead forecast error for Model A equal σ^2, while for Model B they are $2\sigma^2$. Hence, in terms of variance Model A is preferred. However, Model A has the disadvantage that it delivers biased forecasts, as $\mathsf{E}[e_{T+1|T,A}] = \mathsf{E}[e_{T+2|T,A}] = \alpha^* \neq 0$, while Model B's one-step ahead forecasts are unbiased.

b. The forecasts obviously remain the same. The realizations now are given by

$$\begin{aligned}
y_{T+1} &= \alpha + \beta(T+1) + \eta_{T+1} \\
&= (\alpha + \beta T) + \beta + \eta_{T+1} \\
&= (y_T - \alpha^* - \eta_T) + \beta + \eta_{T+1}, \\
y_{T+2} &= \alpha + \beta(T+2) + \eta_{T+2} \\
&= (\alpha + \beta(T+1)) + \beta + \eta_{T+2} \\
&= (y_{T+1} - \eta_{T+1}) + \beta + \eta_{T+2} \\
&= (y_T - \alpha^* - \eta_T) + 2\beta + \eta_{T+2}.
\end{aligned}$$

The forecast errors for Model A now are given by

$$e_{T+1|T,A} = \eta_{T+1},$$
$$e_{T+2|T,A} = \eta_{T+2},$$

while for Model B we have

$$e_{T+1|T,B} = -\alpha^* - \eta_T + \eta_{T+1},$$
$$e_{T+2|T,B} = -\alpha^* - \eta_T + \eta_{T+2}.$$

Hence, in this case the forecast errors from Model A not only have the smaller variance but the forecasts also are unbiased, whereas the forecasts from Model B now are biased, with $\mathsf{E}[e_{T+1|T,B}] = \mathsf{E}[e_{T+2|T,B}] = -\alpha^*$.

Exercise 4.4

The Excel file `USUNINSCE.xlsx` contains quarterly observations of US initial claims for unemployment insurance[1] over the period 1974Q1–2013Q4 (160 observations). The series is referred to as 'initial claims' below. A dummy variable DREC is also included, which takes the value 1 during recession periods and 0 during expansions.[2]

Apply the Augmented Dickey-Fuller (ADF) test to examine the presence of a s-tochastic trend in the initial claims series. Use the Schwarz Information Criterion to select the 'lag length' in the test regression. Apply the test two times: first without any deterministic terms (that is, no intercept and no trend) included in the test regression, and then with an intercept included. Can you explain the different outcomes of the test?

Solution

In both cases, the Schwarz Information Criterion indicates that only the first lagged first difference should be included in the test regression. As a result, the ADF test is based on the regression

$$\Delta y_t = \alpha + \rho y_{t-1} + \phi_1 \Delta y_{t-1} + \varepsilon_t,$$

where y_t denotes the initial claims in quarter t, and $\Delta y_t = y_t - y_{t-1}$ indicates the first difference. In the first implementation of the ADF test, we omit the intercept, that is, we set $\alpha = 0$.

In case the intercept is omitted from the test regression, we find an estimate of the coefficient ρ on the lagged level y_{t-1} equal to -0.0044, with a Dickey-Fuller t-statistic of -0.67. The corresponding p-value is equal to 0.425, such that we do not reject the null hypothesis of a stochastic trend in this series. In case an intercept is included in the test regression, we find an estimate of the coefficient ρ on the lagged level y_{t-1} equal to -0.154, with a Dickey-Fuller t-statistic of -4.51. The corresponding p-value is equal to 0.0003, such that we then strongly reject the null hypothesis of a stochastic trend in this series.

[1] This variable, collected by the US Department of Labor is one of the ten components of The Conference Board's Leading Economic Index, which is one of the most important indicators of the (future) state of the US economy; see `http://www.conference-board.org/data/bcicountry.cfm?cid=1` for more information. Also see `https://www.dol.gov/ui/data.pdf`.

[2] Based on the turning points provided at `http://www.nber.org/cycles/`

The rather different outcomes of the ADF test depending on whether or not we include an intercept in the test regression can be understood by observing the initial claims series as shown in Figure 1 below. First, note that the series displays quite prolonged deviations from its (unconditional) mean level, hence the question whether this series contains a stochastic trend is a legitimate one to ask. However, if we examine this hypothesis with an ADF test without an intercept in the test regression, we impose that in the model under the alternative hypothesis of stationarity ($\rho < 0$), the time series has an unconditional mean equal to zero. Clearly, this assumption is inappropriate, given (for example) the sample mean of 383. As such, the model under the alternative hypothesis is not realistic for this time series, and hence the ADF test becomes 'biased' towards the null hypothesis. Accounting for the nonzero unconditional mean by including an intercept in the test regression, we then find that while the deviations from the mean are quite long-lasting, they probably should not be characterized as 'permanent'.

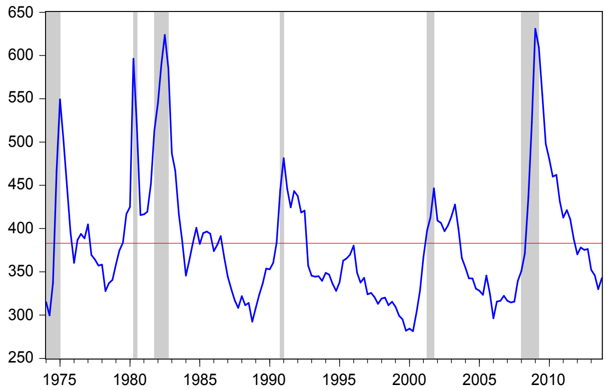


Figure 1: Initial claims

The shaded areas are NBER recession periods.

Exercise 4.5

The Excel file `LSAGDP.xlsx` contains quarterly observations on the natural logarithm of South African Gross Domestic Product (GDP) over the period 1960Q1–2016Q2 (226 observations), as obtained from OECD Statistics.[3] A dummy variable DREC is also included, which takes the value 1 during recession periods and 0 during expansions, as determined with the BBQ algorithm of Harding and Pagan (2002).[4]

Apply the Augmented Dickey-Fuller (ADF) test to examine the presence of a stochastic trend in the log GDP series. Given the properties of the time series, which deterministic components do you include in the test regression? Does the outcome of the test depend on whether you use the Schwarz Information Criterion or the Akaike Information Criterion to select the 'lag length' in the test regression?

Solution

The time series of quarterly log GDP shows a clear upward trend, interrupted by periods of declining levels during economic recessions. Given this property of the time series, we include an intercept and a linear trend in the regression model that is used for the Dickey-Fuller test. Applying the Schwarz Information Criterion for automatic selection of the lag length (which produces the result that the first two lagged first differences are included), gives an estimate of the coefficient ρ on the lagged level y_{t-1} equal to -0.022, with a Dickey-Fuller t-statistic of -2.74. The corresponding p-value is equal to 0.221, such that we do not reject the null hypothesis of a unit root in this series. Applying the Akaike Information Criterion for automatic lag length selection suggests to include seven lagged first differences. This gives an estimate of the coefficient ρ on the lagged level y_{t-1} equal to -0.028, with a Dickey-Fuller t-statistic of -3.18. The corresponding p-value is equal to 0.092, so in this case we would reject the null hypothesis of a unit root at the 10% but not at the 5% significance level.

[3] `https://stats.oecd.org/`

[4] Harding, D. and A.R. Pagan (2002), Dissecting the cycle: a methodological investigation, *Journal of Monetary Economics* **49**, 365–381.

Exercise 4.6

The Excel file `USMSSALES.xlsx` contains quarterly observations on the natural logarithm of US manufacturing sales[5] over the period 1959Q1–2015Q4. A dummy variable DREC is also included, which takes the value 1 during recession periods and 0 during expansions.[6]

a. Apply the Augmented Dickey-Fuller (ADF) test to examine the presence of a stochastic trend in the log sales series. Given the properties of the time series, which deterministic components do you include in the test regression? Does the outcome of the test depend on whether you use the Schwarz Information Criterion or the Akaike Information Criterion to select the 'lag length' in the test regression?

b. Estimate an AR(3) model with intercept and deterministic trend for the log sales series y_t, that is,

$$y_t = \alpha + \beta t + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \phi_3 y_{t-3} + \varepsilon_t,$$

using the sample period 1960Q1–1979Q4 (T =80 observations). Construct one-step ahead point forecasts for the sample period 1980Q1–2015Q4 (144 observations). Convert these to forecasts for the quarterly growth rate in percent, that is, construct the series $\hat{g}_{t+1|t} = 100 \times (\hat{y}_{t+1|t} - y_t)$, where $\hat{y}_{t+1|t}$ denotes the one-step ahead forecast for log sales in quarter $t+1$ and y_t is the actual value of log sales in quarter t.

Evaluate the accuracy of these point forecasts for the quarterly growth rates in 'absolute' terms, by considering their a) unbiasedness, b) accuracy, and c) efficiency. Also evaluate their 'relative' accuracy by comparing the point forecasts with those obtained from an AR(3) model (with an intercept but without a linear trend) for quarterly growth rates $g_t = 100 \times (y_t - y_{t-1})$. Are the results of this forecast comparison in agreement with the conclusions from the ADF test applied in part a.?

[5]This variable is one of the four components of The Conference Board's Coincident Economic Index, which is one of the most important indicators of the state of the US economy; see `http://www.conference-board.org/data/bcicountry.cfm?cid=1` for more information.

[6]Based on the turning points provided at `http://www.nber.org/cycles/`

c. Repeat the forecasting exercise of part b., but now updating the estimates of the coefficient in the AR(3) model with intercept and trend using a moving window of 80 observations as you move through the out-of-sample period. That is, first estimate the model using the sample period 1960Q1–1979Q4 and compute a one-step ahead forecast for 1980Q1. Then re-estimate the model using the sample period 1960Q2–1980Q1 and compute a forecast for 1980Q2. Continue in this way until you compute the forecast for 2015Q4 using the model coefficients as estimated on the sample period 1995Q4–2015Q3.

Again evaluate the quality of the resulting forecasts and compare this with the forecasts obtained in part b. Can you explain the differences in accuracy?

Also compare the forecasts obtained here with the forecasts obtained from an AR(3) model for quarterly growth rates , also updating the coefficient estimates using a moving window of 80 observations. Are the results of this forecast comparison in agreement with the conclusions from the ADF test applied in part a.?

Solution

a. The time series of quarterly (log) manufacturing sales shows a clear upward trend (interrupted by short periods of decline during economic recessions). Given this property of the time series, we include an intercept and a linear trend in the regression model that is used for the Dicky-Fuller test. Applying the Schwarz Information Criterion for automatic selection of the lag length (which produces the result that one lagged first differences is included), gives an estimate of the coefficient ρ on the lagged level y_{t-1} equal to -0.025, with a Dickey-Fuller t-statistic of -2.16. The corresponding p-value is equal to 0.512, such that we do not reject the null hypothesis of a unit root in this series. Applying the Akaike Information Criterion for automatic lag length selection suggests to include 9 lagged first differences. This gives an estimate of the coefficient ρ on the lagged level y_{t-1} equal to -0.032, with a Dickey-Fuller t-statistic of -2.46. The corresponding p-value is equal to 0.350, so also in this case we do not reject the null hypothesis of a unit root.

b. The one-step forecasts obtained from the AR(3) model with deterministic trend for the period 1980Q1–2015Q4 are shown below.

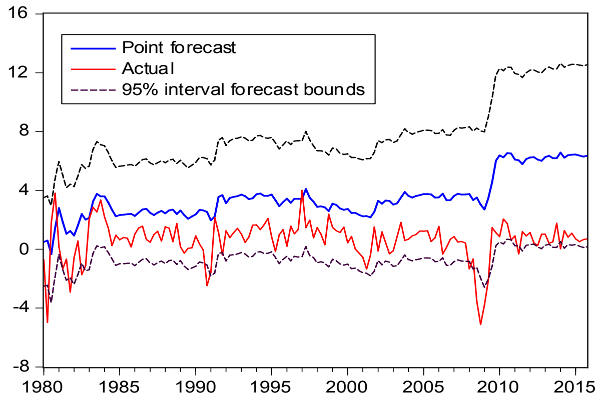


Figure 2: Point forecasts from AR(3) model with deterministic trend, together with actual values and lower and upper bounds of 95% interval forecast

The mean of the forecast errors $e_{\text{AR3DT},t+1|t}$ is equal to -2.88, with standard error 0.148. Hence, the mean forecast error is significantly different from zero, which means that we reject the null hypothesis that the forecasts are unbiased. Figure 3 also shows quite clearly that the forecasts are too high, on average.

The mean squared prediction error (MSPE) is equal to 11.40, which is more than six times as large as the variance of the quarterly growth rates over the forecast period (at 1.80). Hence, the AR(3)-with-trend forecasts seem very inaccurate. Note, however, that the high value of the MSPE is, to a large extent, due to the bias: recall that the MSPE is the sum of the variance of the forecast errors and the squared bias. The latter component is 8.29 ($= -2.88^2$), such that the variance of the forecast errors is 3.11. Still, this is about 70% larger than the variance of the quarterly growth rates over the forecast period (at 1.80).

Due to the large bias, the AR(3) forecasts also seem to be inefficient: when we estimate the Mincer-Zarnowitz regression

$$g_{t+1} = \beta_0 + \beta_1 \hat{g}_{t+1|t} + \eta_{t+1},$$

we find $\hat{\beta}_0 = -0.048\,(0.28)$ and $\hat{\beta}_1 = 0.188\,(0.075)$. Obviously, the estimate $\hat{\beta}_1$ is significantly different from 1. The R^2 of the MZ-regression is equal to 0.042, showing that the forecasts have almost no explanatory power for the actual growth rates.

The MSPE of the forecasts obtained with an AR(3) model for the growth rates is equal to 1.39, which obviously is substantially smaller than the MSPE of the AR(3)-with-trend forecasts. In order to test whether the difference is statistically significant, we may use the Diebold-Mariano test statistic based on the 'loss differential' $d_{t+1} = e^2_{\mathrm{AR3},t+1|t} - e^2_{\mathrm{AR3DT},t+1|t}$, where $e_{\mathrm{AR3},t+1|t}$ is the forecast error of the AR(3) model for the growth rates. The value of the DM statistic is equal to -10.55, with a p-value equal to 0.0000, so that we may reject the null hypothesis of equal forecast accuracy.

The results of this forecast comparison are in agreement with the conclusions from the Dickey-Fuller test applied in part a., as the DF-test provides quite strong evidence that the null hypothesis of a unit root should not be rejected. Here we find that the AR(3) model for growth rates (hence, with the unit root imposed) also produces much more accurate forecasts than the AR(3) model with deterministic trend.

c. The one-step forecasts obtained when the coefficients in the AR(3) model with deterministic trend are updated using a moving window of 80 observations are shown below.

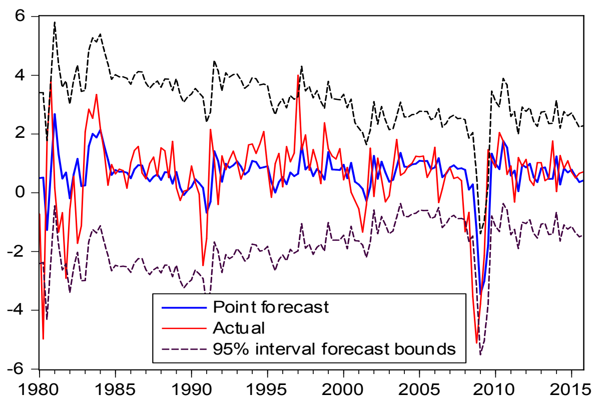


Figure 3: Point forecasts from AR(3) model with deterministic trend, together with actual values and lower and upper bounds of 95% interval forecast

The mean of the forecast errors $e_{\mathrm{AR3DTR},t+1|t}$ is equal to -0.029, with standard error 0.103. Hence, the mean forecast error is not significantly different from zero even at the 10% significance level, which means that we do not reject the null hypothesis that the forecasts are unbiased.

The mean squared prediction error (MSPE) is equal to 1.527, which is about 15% smaller than the variance of the quarterly growth rates over the forecast period (at 1.80). Given that the bias is very small in this case, the MSPE is almost equal to the variance of the forecast errors.

When we estimate the Mincer-Zarnowitz regression

$$g_{t+1} = \beta_0 + \beta_1 \hat{g}_{t+1|t} + \eta_{t+1},$$

we find estimates $\hat{\beta}_0 = 0.130\,(0.137)$ and $\hat{\beta}_1 = 0.751\,(0.141)$. The intercept is not significantly different from 0, confirming that the forecasts are unbiased. The estimate of β_1 is not significantly different from 1 at the 5% level (although it is at the 10% level), such that we may conclude that the forecasts are 'efficient'. Still the R^2 of the MZ-regression is equal to 0.165, suggesting that the forecasts 'explain' only a relatively small part of the variation in the actual growth rates.

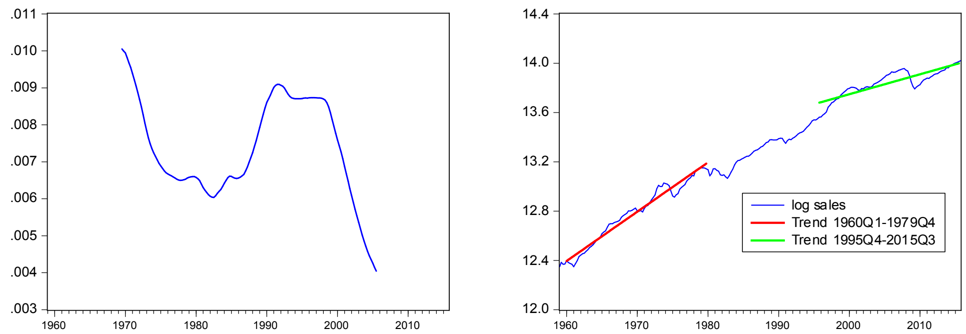


Figure 4: Moving window estimates of trend coefficient (left) and estimated trends on first and last moving window (right).

The reason that the forecasts obtained with recursively estimated/updated coefficients are much more accurate than those obtained in part b. with coefficients that are 'fixed' (to those values obtained using the sample period 1960Q1-1979Q4) is that the trend properties of the log sales series change over time. We examine this by estimating the coefficients α and β in

$$y_t = \alpha + \beta t + \varepsilon_t,$$

using a moving window of 80 observations, starting with 1960Q1-1979Q4 and ending with 1995Q4-2015Q3. The right panel in Figure 4 shows the estimated trends $\hat{\alpha} + \hat{\beta} t$ using the first and last windows. The left panel in Figure 4 shows the least squares estimates of β for all moving windows, where the date on the horizontal axis corresponds with the mid-point of the moving window. It is seen that the estimates vary considerably over time and, in particular, that the estimate obtained using the sample period 1960Q1-1979Q4 is substantially larger than that obtained using later windows. This eventually leads to out-of-sample forecasts that are too high, as seen in Figure 2.

When we also update the coefficient estimates in the AR(3) model for the growth rates before making forecasts, we find an MSPE equal to 1.449, which still is smaller than the MSPE of the AR(3)-with-trend forecasts. In order to test whether the difference is statistically significant, we may use the Diebold-Mariano test statistic based on the 'loss differential' $d_{t+1} = e^2_{\mathrm{AR3R},t+1|t} - e^2_{\mathrm{AR3DTR},t+1|t}$, where $e_{\mathrm{AR3R},t+1|t}$ is the forecast error of the AR(3) model for the growth rates. The value of the DM statistic is equal to -1.07, with a p-value equal to 0.286, so that we do not reject the null hypothesis of equal forecast accuracy even at the 10% significance level.

The results of this forecast comparison are not really in agreement with the conclusions from the Dickey-Fuller test applied in part a., in the sense that the DF-test provides quite strong evidence that the null hypothesis of a unit root cannot be rejected. In contrast, here we find that the AR(3) model for growth rates (hence, with the unit root imposed) does not produce significantly more accurate forecasts than the AR(3) model with deterministic trend, at least when the coefficient estimates are updated.

(International) Bachelor Econometrics and Operations Research
FEB23001(X)-19 Tijdreeksanalyse / Time Series Analysis

Exercises - Seasonality (Chapter 5)

Exercise 5.1

Consider the following AR(3) model for a quarterly series y_t:

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \phi_3 y_{t-3} + \varepsilon_t, \tag{1}$$

where ε_t is a white noise series with variance σ^2.

Derive the parameter restrictions for $\{\phi_1, \phi_2, \phi_3\}$ that should hold in case y_t has

a. A single nonseasonal unit root
b. Two nonseasonal unit roots
c. Three seasonal unit roots: -1, and $\pm i$

Solution

The AR(3) polynomial is given by

$$\phi_3(L) = 1 - \phi_1 L - \phi_2 L^2 - \phi_3 L^3. \tag{2}$$

a. In case of a single nonseasonal unit root, this can be written as

$$\begin{aligned}\phi_3(L) &= (1-L)(1-\phi_1^* L - \phi_2^* L^2) \\ &= 1 - (1+\phi_1^*)L - (\phi_2^* - \phi_1^*)L^2 + \phi_2^* L^3 \end{aligned} \tag{3}$$

Equating the coefficients in (2) and (3) gives

$$\begin{aligned}\phi_1 &= 1 + \phi_1^*, \\ \phi_2 &= \phi_2^* - \phi_1^*, \\ \phi_3 &= -\phi_2^*.\end{aligned}$$

Combining these three expressions gives the restriction

$$\phi_1 + \phi_2 + \phi_3 = 1.$$

b. In case of another nonseasonal unit root, the polynomial $\phi_2^*(L) = (1 - \phi_1^* L - \phi_2^* L^2)$ in $\phi_3(L) = (1 - L)\phi_2^*(L)$, see (3) above, can be written as

$$\begin{aligned}(1 - \phi_1^* L - \phi_2^* L^2) &= (1 - L)(1 - \phi_1^{**} L) \\ &= 1 - (1 + \phi_1^{**})L + \phi_1^{**} L^2. \end{aligned} \tag{4}$$

Equating coefficients in (4), and using the relationships between ϕ_1^*, ϕ_2^* and ϕ_1, ϕ_2, ϕ_3 gives

$$\begin{aligned}\phi_1^* &= 1 + \phi_1^{**} = -\phi_3 - \phi_2 \\ \phi_2^* &= -\phi_1^{**} = -\phi_3.\end{aligned}$$

Combining these two expressions for ϕ_1^{**} gives the additional restriction

$$1 + \phi_2 + 2\phi_3 = 0.$$

c. In case of three seasonal unit roots -1 and $\pm i$, the AR(3) polynomial can be written as

$$(1 + L)(1 + L^2) = 1 + L + L^2 + L^3, \tag{5}$$

from which the restrictions on ϕ_1, ϕ_2 and ϕ_3 are obvious.

Exercise 5.2
Consider the following restricted AR(4) model for a quarterly series y_t:

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \phi_4 y_{t-4} + \varepsilon_t, \tag{6}$$

where ε_t is a white noise series with variance σ^2.

If possible, derive the parameter restrictions for $\{\phi_1, \phi_2, \phi_4\}$ that should hold in case y_t has

a. A single nonseasonal unit root

b. Two nonseasonal unit roots

c. Three seasonal unit roots: -1, and $\pm i$

Solution
The restricted AR(4) polynomial is given by

$$\phi_4(L) = 1 - \phi_1 L - \phi_2 L^2 - \phi_4 L^4. \tag{7}$$

a. In case of a single nonseasonal unit root, this can be written as

$$\begin{aligned}\phi_4(L) &= (1 - L)(1 - \phi_1^* L - \phi_2^* L^2 - \phi_3^* L^3) \\ &= 1 - (1 + \phi_1^*)L - (\phi_2^* - \phi_1^*)L^2 - (\phi_3^* - \phi_2^*)L^3 + \phi_3^* L^4 \end{aligned} \tag{8}$$

Equating the coefficients in (7) and (8) gives

$$\begin{aligned}\phi_1 &= 1 + \phi_1^*, \\ \phi_2 &= \phi_2^* - \phi_1^*, \\ \phi_3(= 0) &= \phi_3^* - \phi_2^*, \\ \phi_4 &= -\phi_3^*.\end{aligned}$$

Combining these three expressions (by simply adding them up) gives the restriction

$$\phi_1 + \phi_2 + \phi_4 = 1.$$

b. In case of another nonseasonal unit root, the polynomial $\phi_3^*(L) = (1 - \phi_1^* L - \phi_2^* L^2 - \phi_3^* L^3)$ in $\phi_4(L) = (1-L)\phi_3^*(L)$, see (8) above, can be written as

$$\begin{aligned}(1 - \phi_1^* L - \phi_2^* L^2 - \phi_3^* L^3) &= (1-L)(1 - \phi_1^{**} L - \phi_2^{**} L^2) \\ &= 1 - (1 + \phi_1^{**})L - (\phi_2^{**} - \phi_1^{**})L^2 + \phi_2^{**} L^3. \qquad (9)\end{aligned}$$

Equating coefficients in (9), we find

$$\begin{aligned}\phi_1^* &= 1 + \phi_1^{**} \\ \phi_2^* &= \phi_2^{**} - \phi_1^{**} \\ \phi_3^* &= -\phi_2^{**}.\end{aligned}$$

Combining these expressions (by adding them up) gives the restriction $\phi_1^* + \phi_2^* + \phi_3^* = 1$. Using the relationships between ϕ_1^*, ϕ_2^* and ϕ_1, ϕ_2, ϕ_3 gives the following restriction in terms of the original AR-coefficients:

$$1 + \phi_2 + 3\phi_4 = 0.$$

(Note that this restriction can be written in various alternative ways. For example, because it also holds that $\phi_1 + \phi_2 + \phi_4 = 1$ (see a.), we can also write the additional restriction as $\phi_1 = 2(1 + \phi_4)$.)

c. In case of three seasonal unit roots -1 and $\pm i$, the restricted AR(4) polynomial can be written as

$$\begin{aligned}\phi_4(L) &= (1+L)(1+L^2)(1 - \psi_1 L) && (10)\\ &= (1 + L + L^2 + L^3)(1 - \psi_1 L) && (11)\\ &= 1 - (\psi_1 - 1)L - (\psi_1 - 1)L^2 - (\psi_1 - 1)L^3 - \psi_1 L^4. && (12)\end{aligned}$$

Because the third-order coefficient in the original AR(4)-polynomial $\phi_4(L)$ is equal to zero it follows that $\psi_1 = 1$, which in turn implies the restrictions $\phi_1 = 0$, $\phi_2 = 0$, and $\phi_4 = 1$.
(Note that the resulting model is a seasonal random walk, that is, y_t also has a nonseasonal unit root at 1.)

Exercise 5.3

Suppose we want to consider the possibility that a quarterly time series y_t has deterministic seasonality by allowing the mean to vary across the different seasons in a first-order autoregressive [AR(1)] model for y_t. For that purpose, let $D_{s,t}$ for $s = 1, \ldots, 4$ denote quarterly dummy variables, defined as

$$D_{s,t} = 1 \text{ if time } t \text{ corresponds with season } s\text{, and 0 otherwise.}$$

Consider the following three model representations:

$$\phi_1(L)(y_t - \mu_1 D_{1,t} - \mu_2 D_{2,t} - \mu_3 D_{3,t} - \mu_4 D_{4,t}) = \varepsilon_t, \tag{13}$$

$$\phi_1(L)(y_t - \mu^* - \mu_1^* D_{1,t} - \mu_2^* D_{2,t} - \mu_3^* D_{3,t}) = \varepsilon_t, \tag{14}$$

$$\phi_1(L)(y_t - \mu^{**} - \mu_1^{**} D_{1,t}^* - \mu_2^{**} D_{2,t}^* - \mu_3^{**} D_{3,t}^*) = \varepsilon_t, \tag{15}$$

where $\phi_1(L) = 1 - \phi_1 L$ and ε_t is a white noise process with mean equal to zero and variance equal to σ^2, and $D_{s,t}^* = D_{s,t} - D_{4,t}$ for $s = 1, 2, 3$.

Discuss the interpretation of the "mean coefficients" in each of these representations (that is, the coefficients μ_1, μ_2, μ_3, and μ_4 in (13), the coefficients μ^*, μ_1^*, μ_2^*, and μ_3^* in (14), and the coefficients μ^{**}, μ_1^{**}, μ_2^{**}, and μ_3^{**} in (15)). Also point out how the coefficients in the different representations are related. How would you test for the presence of deterministic seasonality in each of the three representations?

Solution

For simplicity, consider the case $\phi_1 = 0$ (but all results below continue to hold for the general AR(1) case). In (13), the coefficients μ_s correspond to the unconditional mean of y_t during seasons s, as the model reads

$$y_t = \mu_s + \varepsilon_t, \qquad \text{when time } t \text{ corresponds to season } s.$$

The model in (14) reads

$$y_t = \begin{cases} \mu^* + \mu_s^* + \varepsilon_t, & \text{when time } t \text{ corresponds to season } s = 1, 2 \text{ or } 3 \\ \mu^* + \varepsilon_t, & \text{when time } t \text{ corresponds to season } s = 4. \end{cases}$$

Hence, in this model μ^* is the unconditional mean of y_t during the fourth season, while μ_s^* is the difference between the unconditional mean of y_t during season s and the unconditional mean of y_t during the fourth season, for $s = 1, 2, 3$.

The model in (15) reads

$$y_t = \begin{cases} \mu^{**} + \mu_s^{**} + \varepsilon_t, & \text{when time } t \text{ corr. to season } s = 1, 2 \text{ or } 3 \\ \mu^{**} - \mu_1^{**} - \mu_2^{**} - \mu_3^{**} + \varepsilon_t, & \text{when time } t \text{ corresponds to season } s = 4. \end{cases}$$

To understand the interpretation of the coefficients in this specification, it is useful to take the sum of the four observations within a year, which gives:

$$\sum_{i=0}^{3} y_{t-i} = 4\mu^{**} + \varepsilon_t + \varepsilon_{t-1} + \varepsilon_{t-2} + \varepsilon_{t-3},$$

which makes clear that μ^{**} is the average unconditional mean of y_t across the four seasons, which follows from the fact that (13) with $\phi_1 = 0$ gives $\sum_{i=0}^{3} y_{t-i} = \mu_1 + \mu_2 + \mu_3 + \mu_4 + \varepsilon_t + \varepsilon_{t-1} + \varepsilon_{t-2} + \varepsilon_{t-3}$. Consequently, μ_s^{**} is the difference between the unconditional mean of y_t during season s and the overall unconditional mean of y_t for $s = 1, 2, 3$. The sum $-\mu_1^{**} - \mu_2^{**} - \mu_3^{**}$ is the difference between the unconditional mean of y_t during season 4 and the overall unconditional mean of y_t.

Comparing the representations in (13) and (14), we have that $\mu^* = \mu_4$, while $\mu^* + \mu_s^* = \mu_s$ or $\mu_s^* = \mu_s - \mu_4$ for $s = 1, 2, 3$. Comparing the representations in (13) and (15), we have that $\mu^{**} = \frac{1}{4} \sum_{s=1}^{4} \mu_s$. In addition, $\mu^{**} + \mu_s^{**} = \mu_s$, or $\mu_s^{**} = \mu_s - \frac{1}{4} \sum_{s=1}^{4} \mu_s$ for $s = 1, 2, 3$.

We can test for the presence of deterministic seasonality by testing the null hypothesis composed of the restrictions that imply that the unconditional means in all seasons are equal. These restrictions are $\mu_1 = \mu_2 = \mu_3 = \mu_4$ in (13), $\mu_1^* = \mu_2^* = \mu_3^* = 0$ in (14), and $\mu_1^{**} = \mu_2^{**} = \mu_3^{**} = 0$ in (15).

Exercise 5.4

Suppose y_t is a quarterly observed time series, and ε_t is a white noise process with mean equal to zero and variance equal to σ^2. Questions a., b., and c. should be considered independently from each other.

a. Suppose the data generating process for a y_t is $(1 - L^2)y_t = \varepsilon_t$. What are the autocorrelation properties of the seasonally differenced series $u_t = (1 - L^4)y_t$?

b. Suppose the data generating process for y_t is $(1 - L^4)y_t = \varepsilon_t$. What are the autocorrelation properties of the series $u_t = (1 - L^2)y_t$?

c. Suppose the data generating process for y_t is $(1 + L^2)y_t = \varepsilon_t$. What are the autocorrelation properties of the seasonally differenced series $u_t = (1 - L^4)y_t$?

Solution

a. Notice that $(1 - L^2)y_t = (1 - L)(1 + L)y_t$. Thus, y_t has unit roots equal to 1 and -1. Applying the filter $(1 - L^4)$ amounts to *overdifferencing* as this also implies the complex pair of unit roots $\pm i$.

Note that $y_t = y_{t-2} + \varepsilon_t = y_{t-4} + \varepsilon_t + \varepsilon_{t-2}$, such that

$$u_t = (1 - L^4)y_t = \varepsilon_t + \varepsilon_{t-2}.$$

Hence, it follows that (i) $\mathsf{E}[u_t] = 0$; and (ii) $\mathsf{V}[u_t] = \mathsf{E}[u_t^2] = \mathsf{E}[(\varepsilon_t + \varepsilon_{t-2})^2] = 2\sigma^2$.

Furthermore,

$$\begin{aligned}
\gamma_1 &= \mathsf{E}[u_t u_{t-1}] = \mathsf{E}[(\varepsilon_t + \varepsilon_{t-2})(\varepsilon_{t-1} + \varepsilon_{t-3})] = 0,\\
\gamma_2 &= \mathsf{E}[u_t u_{t-2}] = \mathsf{E}[(\varepsilon_t + \varepsilon_{t-2})(\varepsilon_{t-2} + \varepsilon_{t-4})] = \sigma^2,\\
\gamma_k &= \mathsf{E}[u_t u_{t-k}] = \mathsf{E}[(\varepsilon_t + \varepsilon_{t-2})(\varepsilon_{t-k} + \varepsilon_{t-k-2})] = 0 \quad \text{for all } k \geq 3.
\end{aligned}$$

Thus, all autocorrelations of $u_t = (1 - L^4)y_t$ are equal to zero, except $\rho_2 = \gamma_2/\gamma_0 = 0.5$.

b. Notice that $(1-L^4)y_t = (1-L)(1+L)(1+L^2)y_t$. Thus, y_t has unit roots equal to 1, -1 and $\pm i$. Applying the filter $(1-L^2)$ amounts to *underdifferencing* as it only implies the roots 1 and -1 but neglects the complex pair $\pm i$.

Note that

$$u_t = y_t - y_{t-2} = -y_{t-2} + y_{t-4} + y_t - y_{t-4} = -u_{t-2} + \varepsilon_t.$$

Hence, asymptotically, u_t has the following autocorrelation properties:

$$\rho_k = \begin{cases} -1 & \text{for } k = 2, 6, 10, \ldots, \\ 1 & \text{for } k = 4, 8, 12, \ldots, \\ 0 & \text{for all odd values of } k. \end{cases}$$

c. Notice that $(1+L^2)y_t = (1-iL)(1+iL)y_t$, with $i^2 = -1$. Thus, y_t has the complex pair of seasonal unit roots $\pm i$. Applying the filter $(1-L^4)$ amounts to *overdifferencing* as this also implies the unit roots 1 and -1.

Note that $y_t = -y_{t-2} + \varepsilon_t = y_{t-4} + \varepsilon_t - \varepsilon_{t-2}$, such that

$$u_t = (1-L^4)y_t = \varepsilon_t - \varepsilon_{t-2}.$$

Hence, it follows that (i) $\mathsf{E}[u_t] = 0$; and (ii) $\mathsf{V}[u_t] = \mathsf{E}[u_t^2] = \mathsf{E}[(\varepsilon_t - \varepsilon_{t-2})^2] = 2\sigma^2$. Furthermore,

$$\begin{aligned} \gamma_1 &= \mathsf{E}[u_t u_{t-1}] = \mathsf{E}[(\varepsilon_t - \varepsilon_{t-2})(\varepsilon_{t-1} - \varepsilon_{t-3})] = 0, \\ \gamma_2 &= \mathsf{E}[u_t u_{t-2}] = \mathsf{E}[(\varepsilon_t - \varepsilon_{t-2})(\varepsilon_{t-2} - \varepsilon_{t-4})] = -\sigma^2, \\ \gamma_k &= \mathsf{E}[u_t u_{t-k}] = \mathsf{E}[(\varepsilon_t - \varepsilon_{t-2})(\varepsilon_{t-k} - \varepsilon_{t-k-2})] = 0 \quad \text{for all } k \geq 3. \end{aligned}$$

Thus, all autocorrelations of $u_t = (1-L^4)y_t$ are equal to zero, except $\rho_2 = \gamma_2/\gamma_0 = -0.5$.

(International) Bachelor Econometrics and Operations Research
FEB23001(X)-19 Tijdreeksanalyse / Time Series Analysis

Exercises - Aberrant observations (Chapter 6)

Exercise 6.1

Suppose we are interested in a time series x_t, which follows a stationary AR(1) process,

$$x_t = \phi_1 x_{t-1} + \varepsilon_t, \tag{1}$$

where $|\phi_1| < 1$, and $\varepsilon_t \sim$ i.i.d. $N(0, \sigma^2)$. Instead of x_t we actually observe the time series y_t, where

$$y_t = x_t + \zeta d_t, \qquad t = 0, 1, 2, \ldots, T, \tag{2}$$

where d_t can take the values -1, 0, and 1, with probabilities

$$P(d_t = 1) = P(d_t = -1) = p_o/2 \qquad \text{and} \qquad P(d_t = 0) = 1 - p_o, \tag{3}$$

for certain $0 < p_o < 1$.

a. Derive the expressions for the variance $\gamma_{0,x}$ and for the first-order autocovariance $\gamma_{1,x}$ of x_t in terms of the AR(1) parameter ϕ_1 and the variance of shocks ε_t.

b. What type of outlier is present in the observed time series y_t, according to the specification as given in (1)-(3)?

c. Show that asymptotically ($T \to \infty$) the following holds for the sample mean $\bar{y} = \frac{1}{T}\sum_{t=1}^{T} y_t$, the sample variance $\hat{\gamma}_{0,y} = \frac{1}{T}\sum_{t=1}^{T}(y_t - \bar{y})^2$, and the first-order autocovariance $\hat{\gamma}_{1,y} = \frac{1}{T}\sum_{t=1}^{T}(y_t - \bar{y})(y_{t-1} - \bar{y})$ of the observed series y_t:

$$\bar{y} \to 0, \tag{4}$$

$$\hat{\gamma}_{0,y} \to \gamma_{0,x} + \zeta^2 p_o, \tag{5}$$

$$\hat{\gamma}_{1,y} \to \gamma_{1,x}. \tag{6}$$

d. Suppose we consider an AR(1) model for the observed time series y_t, that is,

$$y_t = \tilde{\phi}_1 y_{t-1} + \eta_t, \qquad t = 1, \ldots, T. \tag{7}$$

What do the properties (4), (5), and (6) imply for the OLS estimator of the parameter $\tilde{\phi}_1$ in this model?

Solution

a. From (1), it follows that x_t has unconditional mean equal to zero, that is $\mathsf{E}[x_t] = 0$. To obtain the unconditional variance, note that

$$\begin{aligned}\mathsf{E}[x_t^2] &= \mathsf{E}[(\phi_1 x_{t-1} + \varepsilon_t)^2] \\ &= \mathsf{E}[\phi_1^2 x_{t-1}^2 + 2\phi_1 x_{t-1}\varepsilon_t + \varepsilon_t^2] \\ &= \phi_1^2 \mathsf{E}[x_{t-1}^2] + \mathsf{E}[\varepsilon_t^2].\end{aligned}$$

The unconditional variance of x_t should be the same as that of x_{t-1} given that the AR(1) model is assumed to be stationary. Hence, we set $\mathsf{E}[x_{t-1}^2] = \mathsf{E}[x_t^2]$, such that $\gamma_{0,x} = \mathsf{E}[x_t^2] = \sigma^2/(1-\phi_1^2)$.

(Alternatively, recursively substitute for lagged x_t in (1) until you obtain

$$x_t = \phi_1^t x_0 + \sum_{i=0}^{t-1} \phi_1^i \varepsilon_{t-i},$$

then take expectations of the squares on the left- and right-hand side while letting $t \to \infty$)

For the first-order autocovariance, multiply (1) with x_{t-1} on both the left- and right-hand side, and take expectations:

$$\mathsf{E}[x_t x_{t-1}] = \mathsf{E}[\phi_1 x_{t-1}^2 + x_{t-1}\varepsilon_t],$$

which gives $\gamma_{1,x} = \phi_1 \gamma_{0,x}$.

b. The process d_t controls the occurrence of outliers. ζ is the absolute value of the outlier (the outlier can be both positive and negative, as d_t can be equal to $+1$ or -1). As shown by (2), only the observation at time t is affected by an outlier occurring at time t. Hence, this is the typical specification of an additive outlier.

c. From the definition of d_t, it follows that $\mathsf{E}[d_t] = 0$. Hence, for the sample mean $\bar{y}$, we find that

$$\begin{aligned}\bar{y} &= \frac{1}{T}\sum_{t=1}^{T} y_t \\ &= \frac{1}{T}\sum_{t=1}^{T}(x_t + \zeta d_t) \\ &= \underbrace{\frac{1}{T}\sum_{t=1}^{T} x_t}_{\to\, \mathsf{E}[x_t] = 0} + \zeta \underbrace{\frac{1}{T}\sum_{t=1}^{T} d_t}_{\to\, \mathsf{E}[d_t] = 0} \to 0\end{aligned}$$

For the sample variance $\hat{\gamma}_0$, the following holds

$$\begin{aligned}\hat{\gamma}_0 &= \frac{1}{T}\sum_{t=1}^{T}(y_t-\bar{y})^2 = \frac{1}{T}\sum_{t=1}^{T}(y_t^2 - 2y_t\bar{y} + \bar{y}^2)\\ &= \frac{1}{T}\sum_{t=1}^{T}y_t^2 \ - \ \bar{y}^2 \end{aligned} \tag{8}$$

and

$$\frac{1}{T}\sum_{t=1}^{T}y_t^2 = \frac{1}{T}\sum_{t=1}^{T}(x_t+\zeta d_t)^2 = \frac{1}{T}\sum_{t=1}^{T}x_t^2 + 2\zeta\frac{1}{T}\sum_{t=1}^{T}x_t d_t + \zeta^2\frac{1}{T}\sum_{t=1}^{T}d_t^2.$$

The probability limit of the first term is $\gamma_{0,x}$. The second term converges to the covariance between x_t and d_t, which is equal to 0 by assumption. The limit of the final term is equal to $\zeta^2 p_o$, which can be understood by noting that

$$\mathsf{E}[d_t^2] = (1\times P(d_t=-1)) + (0\times P(d_t=0)) + (1\times P(d_t=1)) = p_o.$$

For the 1st-order autocovariance $\hat{\gamma}_1$, the following holds

$$\begin{aligned}\hat{\gamma}_1 &= \frac{1}{T}\sum_{t=1}^{T}(y_t-\bar{y})(y_{t-1}-\bar{y}) = \frac{1}{T}\sum_{t=1}^{T}(y_t y_{t-1} - y_t\bar{y} - y_{t-1}\bar{y} + \bar{y}^2)\\ &= \frac{1}{T}\sum_{t=1}^{T}y_t y_{t-1} \ - \ \bar{y}^2, \end{aligned} \tag{9}$$

and

$$\begin{aligned}\frac{1}{T}\sum_{t=1}^{T}y_t y_{t-1} &= \frac{1}{T}\sum_{t=1}^{T}(x_t+\zeta d_t)(x_{t-1}+\zeta d_{t-1})\\ &= \frac{1}{T}\sum_{t=1}^{T}x_t x_{t-1} + \zeta\frac{1}{T}\sum_{t=1}^{T}(x_t d_{t-1} + x_{t-1}d_t) + \zeta^2\frac{1}{T}\sum_{t=1}^{T}d_t d_{t-1},\end{aligned}$$

and because the probability limits of the second and third term in the last line are equal to zero (due to the assumptions of independence between x_t and d_t and the lack of autocorrelation in d_t), the result follows.

d. The OLS estimate of $\tilde{\phi}_1$ is equal to

$$\hat{\tilde{\phi}}_1 = \frac{\hat{\gamma}_{1,y}}{\hat{\gamma}_{0,y}},$$

which asymptotically converges to $\gamma_{1,x}/(\gamma_{0,x}+\zeta^2 p_o)$ according to the results derived above. This will be smaller (in absolute value) than the true value ϕ_1. Hence, the coefficient estimate is biased towards zero.

Exercise 6.2

Suppose we are interested in a time series x_t, which follows an MA(1) process,

$$x_t = \varepsilon_t + \theta_1 \varepsilon_{t-1}, \tag{10}$$

where $|\theta_1| < 1$, and the shocks $\varepsilon_t \sim$ i.i.d. $N(0, \sigma^2)$.

a. Derive the expressions for the variance $\gamma_{0,x}$ and for the first-order autocovariance $\gamma_{1,x}$ of x_t in terms of the MA(1) parameter θ_1 and the variance of ε_t.

In addition to the time series x_t, we also observe another time series y_t, where

$$y_t = \eta_t + \theta_1 \eta_{t-1}, \tag{11}$$

where the shocks η_t are given by $\varepsilon_t + \zeta d_t$ with ε_t as above, and d_t can take the values -1, 0, and 1, with probabilities

$$P(d_t = 1) = P(d_t = -1) = p_o/2 \qquad \text{and} \qquad P(d_t = 0) = 1 - p_o, \tag{12}$$

for certain $0 < p_o < 1$.

b. What type of outlier would you say is present in the time series y_t, according to the specification as given in (10)-(12)?

c. Show that asymptotically ($T \to \infty$) the following holds for the sample mean $\bar{y} = \frac{1}{T}\sum_{t=1}^{T} y_t$, the sample variance $\hat{\gamma}_{0,y} = \frac{1}{T}\sum_{t=1}^{T}(y_t - \bar{y})^2$, and the first-order autocovariance $\hat{\gamma}_{1,y} = \frac{1}{T}\sum_{t=1}^{T}(y_t - \bar{y})(y_{t-1} - \bar{y})$ of the observed series y_t:

$$\bar{y} \to 0, \tag{13}$$

$$\hat{\gamma}_{0,y} \to \gamma_{0,x} + (1 + \theta_1^2)\zeta^2 p_o, \tag{14}$$

$$\hat{\gamma}_{1,y} \to \gamma_{1,x} + \theta_1 \zeta^2 p_o. \tag{15}$$

What does this imply for the properties of the autocorrelations of y_t, compared to those of x_t?

Solution

a. From (10) and because $\mathsf{E}[\varepsilon_t] = 0$, it follows directly that x_t has unconditional mean equal to zero, that is $\mathsf{E}[x_t] = 0$. To obtain the unconditional variance, note that

$$\begin{aligned}\mathsf{E}[x_t^2] &= \mathsf{E}[(\varepsilon_t + \theta_1 \varepsilon_{t-1})^2] \\ &= \mathsf{E}[\varepsilon_t^2 + 2\theta_1 \varepsilon_{t-1}\varepsilon_t + \theta_1^2 \varepsilon_{t-1}^2] \\ &= (1+\theta_1^2)\sigma^2,\end{aligned}$$

as $\varepsilon_t \sim$ i.i.d. $(0, \sigma^2)$ such that $\mathsf{E}[\varepsilon_{t-1}\varepsilon_t] = 0$.

For the first-order autocovariance, multiply (10) with x_{t-1} on both the left- and right-hand side, and take expectations:

$$\begin{aligned}\mathsf{E}[x_t x_{t-1}] &= \mathsf{E}[(\varepsilon_t + \theta_1 \varepsilon_{t-1})(\varepsilon_{t-1} + \theta_1 \varepsilon_{t-2})] \\ &= \mathsf{E}[\theta_1 \varepsilon_{t-1}^2] = \theta_1 \sigma^2,\end{aligned}$$

again because $\varepsilon_t \sim$ i.i.d. $(0, \sigma^2)$.

b. The process d_t controls the occurrence of outliers. ζ is the absolute value of the outlier (the outlier can be both positive and negative, as d_t can be equal to $+1$ or -1). As shown by (11), the outlier ζd_t takes the form a large shock at time t, which affects the time series y_t and y_{t+1} in the same way as 'regular' shocks ε_t. This is the typical specification of an innovation outlier.

c. From the definition of d_t, it follows that $\mathsf{E}[d_t] = 0$. Hence, for the sample mean $\bar{y}$ we find that

$$\begin{aligned}\bar{y} &= \frac{1}{T}\sum_{t=1}^{T} y_t \\ &= \frac{1}{T}\sum_{t=1}^{T}(\eta_t + \theta_1 \eta_{t-1}) \\ &= \frac{1}{T}\sum_{t=1}^{T}(\varepsilon_t + \zeta d_t + \theta_1(\varepsilon_{t-1} + \zeta d_{t-1})) \\ &= \underbrace{\frac{1}{T}\sum_{t=1}^{T}\varepsilon_t}_{\to 0} + \zeta \underbrace{\frac{1}{T}\sum_{t=1}^{T} d_t}_{\to 0} + \theta_1 \underbrace{\frac{1}{T}\sum_{t=1}^{T}\varepsilon_{t-1}}_{\to 0} + \theta_1 \zeta \underbrace{\frac{1}{T}\sum_{t=1}^{T} d_{t-1}}_{\to 0} \to 0\end{aligned}$$

For the sample variance $\hat{\gamma}_0$, the following holds

$$\begin{aligned}\hat{\gamma}_0 &= \frac{1}{T}\sum_{t=1}^{T}(y_t-\bar{y})^2 = \frac{1}{T}\sum_{t=1}^{T}(y_t^2 - 2y_t\bar{y} + \bar{y}^2)\\ &= \frac{1}{T}\sum_{t=1}^{T} y_t^2 \;-\; \bar{y}^2 \end{aligned} \tag{16}$$

and

$$\begin{aligned}\frac{1}{T}\sum_{t=1}^{T} y_t^2 &= \frac{1}{T}\sum_{t=1}^{T}(\eta_t + \theta_1\eta_{t-1})^2\\ &= \frac{1}{T}\sum_{t=1}^{T}(\varepsilon_t + \zeta d_t + \theta_1(\varepsilon_{t-1} + \zeta d_{t-1}))^2\\ &= \frac{1}{T}\sum_{t=1}^{T}(\varepsilon_t + \theta_1\varepsilon_{t-1})^2 + 2\zeta\frac{1}{T}\sum_{t=1}^{T}(\varepsilon_t + \theta_1\varepsilon_{t-1})(d_t + \theta_1 d_{t-1})\\ &\quad + \zeta^2\frac{1}{T}\sum_{t=1}^{T}(d_t + \theta_1 d_{t-1})^2.\end{aligned}$$

The probability limit of the first term is $\gamma_{0,x}$. The second term converges to 0 due to the assumption that ε_t and d_t are independent. The limit of the final term is equal to $(1+\theta_1^2)\zeta^2 p_o$, which can be understood by noting that

$$\mathsf{E}[d_t^2] = (1 \times P(d_t = -1)) + (0 \times P(d_t = 0)) + (1 \times P(d_t = 1)) = p_o.$$

For the 1st-order autocovariance $\hat{\gamma}_1$, the following holds

$$\begin{aligned}\hat{\gamma}_1 &= \frac{1}{T}\sum_{t=1}^{T}(y_t-\bar{y})(y_{t-1}-\bar{y}) = \frac{1}{T}\sum_{t=1}^{T}(y_t y_{t-1} - y_t\bar{y} - y_{t-1}\bar{y} + \bar{y}^2)\\ &= \frac{1}{T}\sum_{t=1}^{T} y_t y_{t-1} \;-\; \bar{y}^2, \end{aligned} \tag{17}$$

and

$$\begin{aligned}\frac{1}{T}\sum_{t=1}^{T} y_t y_{t-1} &= \frac{1}{T}\sum_{t=1}^{T}(\eta_t + \theta_1\eta_{t-1})(\eta_{t-1} + \theta_1\eta_{t-2})\\ &= \frac{1}{T}\sum_{t=1}^{T}(\varepsilon_t + \zeta d_t + \theta_1(\varepsilon_{t-1} + \zeta d_{t-1}))(\varepsilon_{t-1} + \zeta d_{t-1} + \theta_1(\varepsilon_{t-2} + \zeta d_{t-2}))\\ &= \frac{1}{T}\sum_{t=1}^{T} x_t x_{t-1} + \frac{1}{T}\sum_{t=1}^{T}\zeta(x_t(d_{t-1} + \theta_1 d_{t-2}) + x_{t-1}(d_t + \theta_1 d_{t-1})\\ &\quad + \frac{1}{T}\sum_{t=1}^{T}\zeta^2(d_t + \theta_1 d_{t-1})(d_{t-1} + \theta_1 d_{t-2}),\end{aligned}$$

and because the probability limit of the second term in the last line is equal to zero (due to the assumptions of independence between x_t and d_t) and that of the third term is equal to $\theta_1\zeta^2 p_o$, the result follows.

To see the implications for the autocorrelation properties of y_t, note that (14) and (15) imply that

$$\begin{aligned}\rho_{1,y} = \frac{\gamma_{1,y}}{\gamma_{0,y}} &\to \frac{\gamma_{1,x} + \theta_1\zeta^2 p_o}{\gamma_{0,x} + (1+\theta_1^2)\zeta^2 p_o} \\ &= \frac{\theta_1\sigma^2 + \theta_1\zeta^2 p_o}{(1+\theta_1^2)\sigma^2 + (1+\theta_1^2)\zeta^2 p_o} \\ &= \frac{\theta_1}{(1+\theta_1^2)} = \rho_{1,x}.\end{aligned}$$

Hence, y_t has the same autocorrelation properties as x_t.

Exercise 6.3

Consider a time series y_t which experiences a temporary change in mean, as described by the model

$$y_t = \delta D_{t,[T/4,3T/4]} + \varepsilon_t, \qquad t = 1, 2, \ldots, T, \tag{18}$$

where $D_{t,[T/4,3T/4]} = 1$ for $t = T/4 + 1, \ldots, 3T/4$ and $D_{t,[T/4,3T/4]} = 0$ otherwise, and ε_t is a white noise series with variance σ^2.

a. Show that asymptotically ($T \to \infty$) the following holds for the sample mean $\bar{y} = \frac{1}{T}\sum_{t=1}^{T} y_t$, the sample variance $\hat{\gamma}_{0,y} = \frac{1}{T}\sum_{t=1}^{T}(y_t - \bar{y})^2$, and the first-order autocovariance $\hat{\gamma}_{1,y} = \frac{1}{T}\sum_{t=1}^{T}(y_t - \bar{y})(y_{t-1} - \bar{y})$:

$$\bar{y} \to \delta/2, \tag{19}$$

$$\hat{\gamma}_{0,y} \to \sigma^2 + \delta^2/4, \tag{20}$$

$$\hat{\gamma}_{1,y} \to \delta^2/4. \tag{21}$$

b. Suppose we consider an AR(1) model for the observed time series y_t, that is,

$$y_t = \alpha + \phi_1 y_{t-1} + \eta_t, \qquad t = 1, \ldots, T. \tag{22}$$

What do the properties (19), (20), and (21) imply for the OLS estimator of the parameter ϕ_1 in this model? [Hint: can you express the OLS estimate $\hat{\phi}_1$ in terms of $\hat{\gamma}_{0,y}$ and $\hat{\gamma}_{1,y}$?]

c. What does this imply for the behavior of the Dickey-Fuller test for a unit root, when applied to the series y_t,when $|\delta|/\sigma^2$ becomes very large?

Solution

a. For the sample mean $\bar{y}$, we find that

$$\begin{aligned}
\bar{y} &= \mathsf{E}[\frac{1}{T}\sum_{t=1}^{T} y_t] \\
&= \mathsf{E}[\frac{1}{T}\sum_{t=1}^{T}(\delta D_{t,[T/4,3T/4]} + \varepsilon_t)] \\
&= \delta \underbrace{\frac{1}{T}\sum_{t=1}^{T} D_{t,[T/4,3T/4]}}_{\to 1/2} + \underbrace{\frac{1}{T}\sum_{t=1}^{T}\varepsilon_t}_{\to \mathsf{E}[\varepsilon_t] = 0} \to \delta/2,
\end{aligned} \tag{23}$$

For the sample variance $\hat{\gamma}_0$, the following holds

$$\begin{aligned}\hat{\gamma}_0 &= \frac{1}{T}\sum_{t=1}^{T}(y_t-\bar{y})^2 = \frac{1}{T}\sum_{t=1}^{T}(y_t^2 - 2y_t\bar{y} + \bar{y}^2)\\ &= \frac{1}{T}\sum_{t=1}^{T} y_t^2 \;-\; \bar{y}^2\end{aligned}$$

and

$$\begin{aligned}\frac{1}{T}\sum_{t=1}^{T} y_t^2 &= \frac{1}{T}\sum_{t=1}^{T}(\delta D_{t,[T/4,3T/4]} + \varepsilon_t)^2\\ &= \frac{1}{T}\sum_{t=1}^{T}\delta^2 D^2_{t,[T/4,3T/4]} + \frac{1}{T}\sum_{t=1}^{T} 2\delta D_{t,[T/4,3T/4]}\varepsilon_t + \frac{1}{T}\sum_{t=1}^{T}\varepsilon_t^2.\end{aligned}$$

The probability limits of the three terms in the last line are $\delta^2/2$, 0 and σ^2, respectively. Hence, we have that $\hat{\gamma}_0 \to \delta^2/2 + \sigma^2 - (\delta/2)^2 = \delta^2/4 + \sigma^2$.

For the 1st-order autocovariance $\hat{\gamma}_1$, the following holds

$$\begin{aligned}\hat{\gamma}_1 &= \frac{1}{T}\sum_{t=1}^{T}(y_t-\bar{y})(y_{t-1}-\bar{y}) = \frac{1}{T}\sum_{t=1}^{T}(y_t y_{t-1} - y_t\bar{y} - y_{t-1}\bar{y} + \bar{y}^2)\\ &= \frac{1}{T}\sum_{t=1}^{T} y_t y_{t-1} \;-\; \bar{y}^2\end{aligned}$$

and

$$\begin{aligned}\frac{1}{T}\sum_{t=1}^{T} y_t y_{t-1} &= \frac{1}{T}\sum_{t=1}^{T}(\delta D_{t,[T/4,3T/4]} + \varepsilon_t)(\delta D_{t-1,[T/4,3T/4]} + \varepsilon_{t-1})\\ &= \frac{1}{T}\sum_{t=1}^{T}\delta^2 D_{t,[T/4,3T/4]} D_{t-1,[T/4,3T/4]}\\ &\quad + \frac{1}{T}\sum_{t=1}^{T}(\delta D_{t,[T/4,3T/4]}\varepsilon_{t-1} + \delta D_{t-1,[T/4,3T/4]}\varepsilon_t) + \frac{1}{T}\sum_{t=1}^{T}\varepsilon_t\varepsilon_{t-1}\end{aligned}$$

The first term has limiting value $\delta^2/2$, while the second and third terms converge to zero. Hence, we have that $\mathsf{E}[\hat{\gamma}_1] = \delta^2/2 - (\delta/2)^2 = \delta^2/4$.

b. The OLS estimate of ϕ_1 is equal to

$$\hat{\phi}_1 = \frac{\hat{\gamma}_{1,y}}{\hat{\gamma}_{0,y}},$$

which has probability limt $(\delta^2/4)/(\sigma^2 + \delta^2/4)$ according to the results derived above. This will be larger than 0, the true value of ϕ_1. In fact, when $|\delta|/\sigma^2$ becomes larger, it approaches 1.

c. The Dickey-Fuller statistic tests the null hypothesis $\mathsf{H}_0 : \phi_1 = 1$ against the alternative $\mathsf{H}_1 : \phi_1 < 1$ in (22). As $|\delta|/\sigma^2$ becomes very large, $\hat{\phi}_1$ approaches 1, such that the null hypothesis is not likely to be rejected.

Exercise 6.4

Suppose we are interested in a time series x_t, which follows an AR(2) process,

$$x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + \varepsilon_t, \tag{24}$$

or

$$(1 - \phi_1 L - \phi_2 L^2)x_t = \varepsilon_t,$$

where ϕ_1 and ϕ_2 are such that the AR(2) process is stationary, L is the usual lag operator defined as $L^k x_t \equiv x_{t-k}$, and $\varepsilon_t \sim$ i.i.d. $N(0, \sigma^2)$.

Instead of x_t we actually observe the 'contaminated' series y_t, containing a single outlier at time $t = \tau$, that is

$$y_t = x_t + \zeta_t d_t, \tag{25}$$

where $d_t = 1$ for $t = \tau$ and 0 otherwise.

Question: How would you test whether the outlier at time $t = \tau$ is an additive outlier (AO) or an innovation outlier (IO)?

Solution

In case of an AO we have $\zeta_t \equiv \zeta \neq 0$, such that

$$y_t = x_t + \zeta d_t,$$

which implies that

$$(1 - \phi_1 L - \phi_2 L^2)y_t = (1 - \phi_1 L - \phi_2 L^2)x_t + (1 - \phi_1 L - \phi_2 L^2)\zeta d_t,$$

or

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \varepsilon_t + \zeta d_t - \phi_1 \zeta d_{t-1} - \phi_2 \zeta d_{t-2}.$$

In case of an IO we have $\zeta_t \equiv \zeta/(1 - \phi_1 L - \phi_2 L^2)$, or, given that an IO is simply an unusual shock at $t = \tau$,

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \varepsilon_t + \zeta d_t.$$

Define $e_t = y_t - \phi_1 y_{t-1} - \phi_2 y_{t-2}$, such that

$$\text{AO:} \quad e_t = \varepsilon_t + \zeta(d_t - \phi_1 d_{t-1} - \phi_2 d_{t-2}) \tag{26}$$

$$\text{IO:} \quad e_t = \varepsilon_t + \zeta d_t, \tag{27}$$

We can interpret (26)-(27) as a regression model for e_t:

$$e_t = \zeta z_{\mathrm{x},t} + \varepsilon_t, \qquad t = 1, \ldots, T, \;\; \mathrm{x} = \mathrm{AO,IO}, \tag{28}$$

with

$$\text{AO:} \quad z_{\mathrm{AO},t} = \begin{cases} 1 & \text{for } t = \tau \\ -\phi_1 & \text{for } t = \tau + 1 \\ -\phi_2 & \text{for } t = \tau + 2 \\ 0 & \text{for } t > \tau + 2 \end{cases}$$

$$\text{IO:} \quad z_{\mathrm{IO},t} = \begin{cases} 1 & \text{for } t = \tau \\ 0 & \text{for } t > \tau \end{cases}$$

and $z_{\mathrm{x},t} = 0$, x=AO,IO, for all $t < \tau$.

For each outlier type, we can obtain estimate of ζ by regressing e_t on $z_{\mathrm{x},t}$ as

$$\hat{\zeta}_{\mathrm{x}}(\tau) = \frac{\sum_{t=1}^{T} e_t z_{\mathrm{x},t}}{\sum_{t=1}^{T} z_{\mathrm{x},t}^2} = \frac{\sum_{t=\tau}^{T} e_t z_{\mathrm{x},t}}{\sum_{t=\tau}^{T} z_{\mathrm{x},t}^2}.$$

Using the definitions of $z_{\mathrm{AO},t}$ and $z_{\mathrm{IO},t}$, it is straightforward to show that $\hat{\zeta}_{\mathrm{AO}}(\tau) = (e_\tau - \phi_1 e_{\tau+1} - \phi_2 e_{\tau+2})/(1 + \phi_1^2 + \phi_2^2)$ and $\hat{\zeta}_{\mathrm{IO}}(\tau) = e_\tau$.

The significance of $\hat{\zeta}_{\mathrm{x}}$ can be assessed by considering the associated t-statistic:

$$\hat{\lambda}_{\mathrm{x}}(\tau) = \frac{\hat{\zeta}_{\mathrm{x}}(\tau)}{\hat{\sigma}_\varepsilon \left(\sum_{t=\tau}^{T} z_{\mathrm{x},t}^2 \right)^{-1/2}},$$

where $\hat{\sigma}_\varepsilon$ is the OLS estimate of the standard deviation of ε_t.

Exercise 6.5

Suppose we are interested in a time series x_t, which follows an ARMA(1,1) process:

$$x_t = \phi_1 x_{t-1} + \varepsilon_t + \theta_1 \varepsilon_{t-1}, \tag{29}$$

where $|\phi_1| < 1$, $|\theta_1| < 1$, and the shocks $\varepsilon_t \sim$ i.i.d. $N(0, \sigma_\varepsilon^2)$. However, instead of x_t we observe the time series y_t, where

$$y_t = x_t + \zeta_t d_t, \tag{30}$$

where d_t can take the values -1, 0, and 1, with probabilities

$$P(d_t = 1) = P(d_t = -1) = p_o/2 \qquad \text{and} \qquad P(d_t = 0) = 1 - p_o, \tag{31}$$

for certain $0 < p_o < 1$.

a. Argue that, according to the specification given in equations (29)-(31), the observed time series y_t possibly contains additive outliers [AOs] in case $\zeta_t = \zeta$.

b. How would you need to specify ζ_t such that equations (29)-(31) accommodate innovation outliers [IOs]?

c. In general, which of these two types of outliers [AOs and IOs] is most harmful if they are neglected/overlooked? Motivate your answer (for example by discussing the effects that neglected AOs and IOs may have)!

Suppose that we have observations y_t for $t = 1, \ldots, T$, which are generated according to equations (29)-(31), with $\zeta_t = \zeta$, $\varepsilon_0 = 0$ and the realizations of d_t are such that $d_\tau = 1$ for some $1 < \tau < T$ and $d_t = 0$ for all $t \neq \tau$ (that is, only a single AO occurs in the time series, at time $t = \tau$).

Assume that we know the true value of the parameters ϕ_1 and θ_1 in equation (29). Let e_t denote the ARMA(1,1) residuals for the observed time series y_t, that is $e_t = y_t - \phi_1 y_{t-1} - \theta_1 e_{t-1}$, where e_0 is set equal to 0 for initialization.

d. Derive the relation between the residuals e_t and the shocks ε_t in equation (29) for $t = 1, \ldots, T$.

Solution

a. According to the specification as given in equations (29)-(31) in the exercise, with $\zeta_t = \zeta$ the outlier only affects the observation at the time it occurs, and it does not affect subsequent observations. This corresponds exactly with the definition of an additive outlier [AO].

b. An innovation outlier occurring at time $t = \tau$ also affects future observations at $t = \tau + 1, \tau + 2, \ldots$, and its effect disappears in the same way as "regular" shocks ε_t. In other words, an IO in y_t can be interpreted as an AO in the error process ε_t, or as an "unusual shock". Note that the ARMA(1,1) process for x_t as given in equation (29) can be written as $(1 - \phi_1 L)x_t = (1 + \theta_1 L)\varepsilon_t$ or $x_t = \frac{(1+\theta_1 L)}{(1-\phi_1 L)}\varepsilon_t$, where L is the usual lag operator defined as $L^k w_t \equiv w_{t-k}$ for all integers k. Inserting this in the general relation $y_t = x_t + \zeta_t d_t$ gives

$$y_t = \frac{(1 + \theta_1 L)}{(1 - \phi_1 L)}\varepsilon_t + \zeta_t d_t, \tag{32}$$

which shows that the dynamic effects of regular shocks ε_t on y_t is determined by the ratio of the MA(1)- and AR(1)-polynomials $\frac{(1+\theta_1 L)}{(1-\phi_1 L)}$. Hence, in order to achieve the same for the effects of an outlier occurring at a particular point in time, we should specify ζ_t as $\zeta_t = \zeta(1 + \theta_1 L)/(1 - \phi_1 L)$. Inserting this in (32) and multiplying with the AR(1)-polynomial $(1 - \phi_1 L)$ also gives the apprpriate model for the observed time series y_t as

$$y_t = \phi_1 y_{t-1} + \varepsilon_t + \zeta d_t + \theta_1(\varepsilon_{t-1} + \zeta d_{t-1}), \tag{33}$$

which again shows the interpretation of an IO as an unusual shock.

c. In general, AO's are more harmful than IO's. For example, AO's are known to affect the estimates of parameters in ARMA-type models, whereas IO's do not. Also, while AO's only affect a single observation of the observed time series, they do affect multiple residuals, which might for example lead to spurious finding of non-zero residual autocorrelation. Both AO's and IO's affect the accuracy of out-of-sample forecasts, especially when the outlier occurs near the forecast origin – although again the effects of an AO are likely to be larger given that the observations following the outlier do not satisfy the underlying ARMA(1,1) relation while they do in case of an IO.

d. Given that $y_t = x_t + \zeta d_t$, it follows that

$$\frac{(1-\phi_1 L)y_t}{1+\theta_1 L} = \frac{(1-\phi_1 L)x_t}{1+\theta_1 L} + \zeta\frac{(1-\phi_1 L)d_t}{1+\theta_1 L}. \tag{34}$$

Now we use that

(i) the definition of $e_t = y_t - \phi_1 y_{t-1} - \theta_1 e_{t-1}$ implies that $y_t = \phi_1 y_{t-1} + e_t + \theta_1 e_{t-1}$ or $e_t = (1-\phi_1 L)y_t/(1+\theta_1 L)$, and

(ii) the fact that x_t follows an ARMA(1,1) process $x_t = \phi_1 x_{t-1} + \varepsilon_t + \theta_1 \varepsilon_{t-1}$ such that $\varepsilon_t = (1-\phi_1 L)x_t/(1+\theta_1 L)$

Together, these two properties imply that (34) can be written as

$$e_t = \varepsilon_t + \zeta\frac{(1-\phi_1 L)d_t}{1+\theta_1 L}.$$

The ratio $\frac{(1-\phi_1 L)}{1+\theta_1 L}$ corresponds with an infinite order polynomial, say, $\pi(L) = 1 - \pi_1 L - \pi_2 L^2 - \pi_3 L^3 - \pi_4 L^4 - \dots$. The coefficients π_i, $i = 1, 2\dots$, can be found by noting that $\pi(L) = \frac{(1-\phi_1 L)}{1+\theta_1 L}$ can be written as

$$(1 - \pi_1 L - \pi_2 L^2 - \pi_3 L^3 - \pi_4 L^4 - \dots)(1+\theta_1 L) = (1-\phi_1 L).$$

The left-hand side of this equation is equal to

$$(1 - (\pi_1 - \theta_1)L - (\pi_2 + \theta_1\pi_1)L^2 - (\pi_3 + \theta_1\pi_2)L^3 - (\pi_4 + \theta_1\pi_3)L^4 - \dots)$$

From this, we can solve for the π_i coefficients recursively, by equating the coefficients for the different lag orders on the left- and hand-right. This gives $\pi_1 = \phi_1 + \theta_1$, $\pi_2 = -\theta_1\pi_1 = -\theta_1(\phi_1 + \theta_1)$, $\pi_3 = -\theta_1\pi_2 = \theta_1^2(\phi_1 + \theta_1)$, and in general $\pi_k = -\theta_1\pi_{k-1} = (-\theta_1)^{k-1}(\phi_1 + \theta_1)$.

Hence, we have

$$e_t = \varepsilon_t + \zeta(d_t - (\theta_1 + \phi_1)d_{t-1} + \theta_1(\theta_1 + \phi_1)d_{t-2} - \theta_1^2(\theta_1 + \phi_1)d_{t-3} + \dots).$$

As it is given that only a single outlier occurs at time τ, such that $d_\tau = 1$ and $d_t = 0$ for all $t \neq \tau$, the above implies that the relation between e_t and ε_t is given by

$$\begin{aligned}
e_t &= \varepsilon_t, && \text{for all } 1 \leq t < \tau, \\
e_t &= \varepsilon_t + \zeta, && \text{for } t = \tau, \\
e_t &= \varepsilon_t - (-\theta_1)^{t-\tau-1}(\theta_1 + \phi_1)\zeta, && \text{for all } \tau > t \leq T.
\end{aligned}$$

Alternatively, the relation between e_t and ε_t can be derived ('recursively') as follows. Given that (i) $\varepsilon_0 = 0$, (ii) the realizations of d_t are such that $d_\tau = 1$ for some $1 < \tau < T$ and $d_t = 0$ for all $t \neq \tau$ (that is, only a single outlier occurs in the time series, at time $t = \tau$), and (iii) e_0 is set equal to 0, it follows immediately that

$$\begin{aligned} e_t &= y_t - \phi_1 y_{t-1} - \theta_1 e_{t-1} \\ &= x_t - \phi_1 x_{t-1} - \theta_1 \varepsilon_{t-1} \\ &= \varepsilon_t \qquad \text{for all } t < \tau. \end{aligned}$$

The outlier occurs at $t = \tau$, such that

$$\begin{aligned} e_\tau &= y_\tau - \phi_1 y_{\tau-1} - \theta_1 e_{\tau-1} \\ &= x_\tau + \zeta - \phi_1 x_{\tau-1} - \theta_1 \varepsilon_{\tau-1} \\ &= \varepsilon_\tau + \zeta. \end{aligned}$$

For $t = \tau + 1$, it then follows that

$$\begin{aligned} e_{\tau+1} &= y_{\tau+1} - \phi_1 y_\tau - \theta_1 e_\tau \\ &= x_{\tau+1} - \phi_1 (x_\tau + \zeta) - \theta_1 (\varepsilon_\tau + \zeta) \\ &= \varepsilon_{\tau+1} - (\theta_1 + \phi_1)\zeta. \end{aligned}$$

Similarly, for $t = \tau + 2$,

$$\begin{aligned} e_{\tau+2} &= y_{\tau+2} - \phi_1 y_{\tau+1} - \theta_1 e_{\tau+1} \\ &= x_{\tau+2} - \phi_1 x_{\tau+1} - \theta_1 (\varepsilon_{\tau+1} - (\theta_1 + \phi_1)\zeta) \\ &= \varepsilon_{\tau+2} + \theta_1(\theta_1 + \phi_1)\zeta, \end{aligned}$$

and for $t = \tau + 3$,

$$\begin{aligned} e_{\tau+3} &= y_{\tau+3} - \phi_1 y_{\tau+2} - \theta_1 e_{\tau+2} \\ &= x_{\tau+3} - \phi_1 x_{\tau+2} - \theta_1 (\varepsilon_{\tau+2} + \theta_1(\theta_1 + \phi_1)\zeta) \\ &= \varepsilon_{\tau+3} - \theta_1^2(\theta_1 + \phi_1)\zeta. \end{aligned}$$

In general, for $t = \tau + j$, $j = 2, 3, \ldots, T - \tau$, we have that

$$\begin{aligned} e_{\tau+j} &= y_{\tau+j} - \phi_1 y_{\tau+j-1} - \theta_1 e_{\tau+j-1} \\ &= x_{\tau+j} - \phi_1 x_{\tau+j-1} - \theta_1 (\varepsilon_{\tau+j-1} - (-\theta_1)^{j-2}(\theta_1 + \phi_1)\zeta) \\ &= \varepsilon_{\tau+j} - (-\theta_1)^{j-1}(\theta_1 + \phi_1)\zeta. \end{aligned}$$

Exercise 6.6

Consider the time series of quarterly growth rates of US investment, for the period 1969Q3-1999Q4 in the Excel file `USINV.xlsx`.

a. Estimate an AR(2) model (including an intercept) for this time series, using the complete sample period. Evaluate the model by inspecting the properties of the residuals (check (i) normality; (ii) [partial] autocorrelations; (iii) homoskedasticity). What kind of patterns do you observe in the residuals around 1978Q2 and 1980Q2?

b. Test for the presence of additive outliers (AOs) and innovation outliers (IO) at an unspecified point in time.

c. Estimate a model where the observation in 1978Q2 is treated as an IO and the observation in 1980Q2 is treated as an AO. Evaluate the model by inspecting the properties of the residuals. Compare the results with those obtained in part a. of this exercise.

Solution

a. Estimating the AR(2) model $y_t - \mu = \phi_1(y_{t-1} - \mu) + \phi_2(y_{t-2} - \mu) + \varepsilon_t$ by conditional least squares gives coefficient estimates (with standard errors in parentheses) of $\hat{\mu} = 1.24(0.46)$, $\hat{\phi}_1 = 0.39(0.09)$, and $\hat{\phi}_2 = 0.21(0.09)$, with an $R^2 = 0.28$.

With a skewness equal to -0.20 and kurtosis equal to 4.05, the Jarque-Bera test for normality of the residuals gives a p-value of 0.041. A significant (partial) autocorrelation appears at lag 8, with $\hat{\rho}_8(\hat{\varepsilon}) = -0.226$ ($\hat{\psi}_8(\hat{\varepsilon}) = -0.218$). The same applies to the squared residuals, for which $\hat{\rho}_8(\hat{\varepsilon}^2) = 0.314$, for example.

b. The residual patterns are 'suspect' around 1978Q2 and 19802, in the sense that (i) the residual for 1978Q2 is unusually large positive, and (ii) the residual for 1980Q2 is unusually large negative while residuals for the two following observations are somewhat large positive, see Figure 1. This raises the suspicion that 1978Q2 might be an innovation outlier (IO) and 1980Q2 might be an additive outlier (AO).

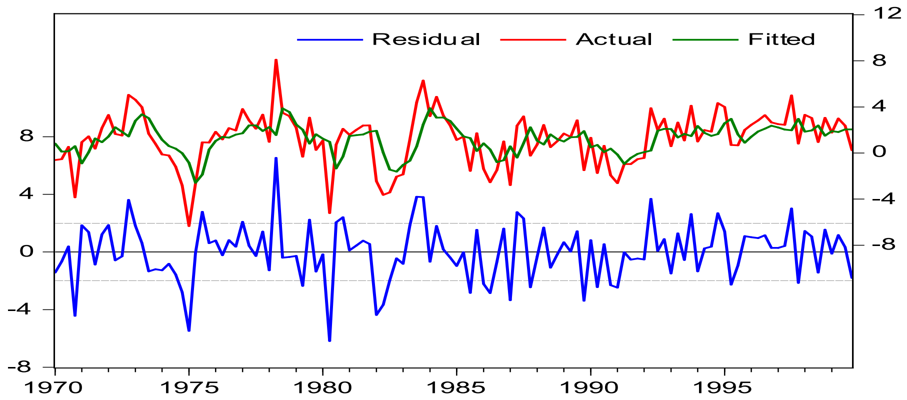


Figure 1: Residuals in AR(2) model for quarterly investment growth rates

This suspicion is confirmed by the test statistics $\hat{\lambda}_{\mathrm{x}}(\tau)$ for x=AO and IO, as shown in Figure 2. The (absolute) value of $\hat{\lambda}_{\mathrm{AO}}(\tau)$ is largest for 1980Q2 and equal to -3.68, while the (absolute) value of $\hat{\lambda}_{\mathrm{IO}}(\tau)$ is largest for 1978Q2 and equal to 3.50.

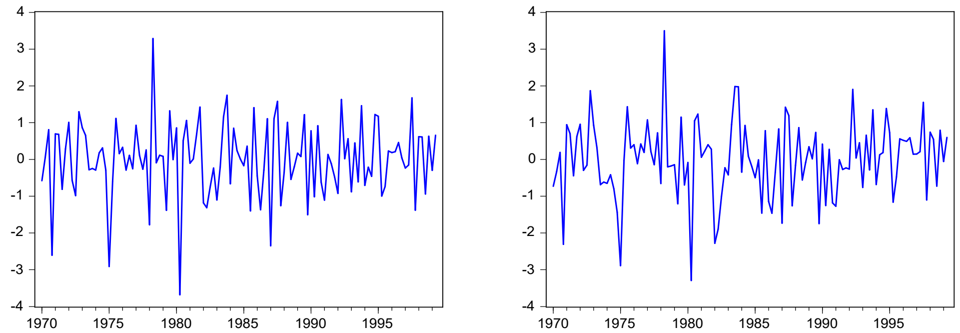


Figure 2: Time series of $\hat{\lambda}_{\mathrm{AO}}(\tau)$ (left) and $\hat{\lambda}_{\mathrm{IO}}(\tau)$ (right)

We use simulation to obtain the (finite-sample) distribution of the test statistics $\hat{\lambda}_{\mathrm{AO}} = \max_{1\leq\tau\leq T} |\hat{\lambda}_{\mathrm{AO}}(\tau)|$ and $\hat{\lambda}_{\mathrm{IO}} = \max_{1\leq\tau\leq T} |\hat{\lambda}_{\mathrm{IO}}(\tau)|$ under the null hypothesis of no outliers, using 50,000 replications. The results of this exercise are shown in Figure 3. For $\hat{\lambda}_{\mathrm{AO}}$ we find a p-value of 0.042, such that we can reject the null hypothesis of no additive outliers at the 5% significance level. For $\hat{\lambda}_{\mathrm{IO}}$ find a p-value of 0.084, such that we can reject the null hypothesis of no additive outliers at the 10% but not the 5% significance level.

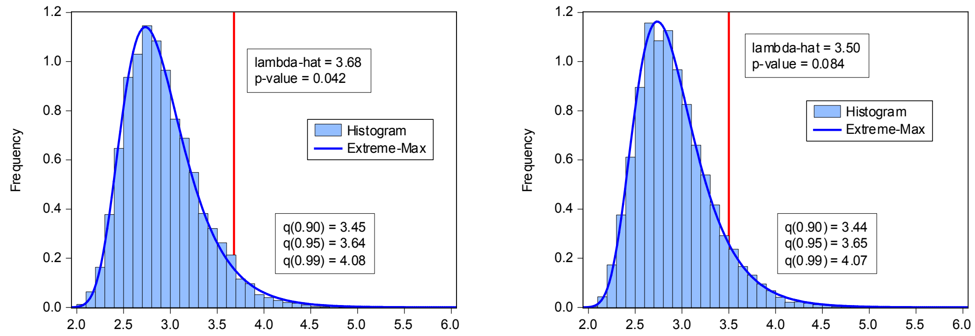


Figure 3: Simulated distribution under the null hypothesis of $\hat{\lambda}_i(\tau)$ statistics for testing for the presence of an AO or IO are shown on the left and right, respectively.

c. Estimating the AR(2) model

$$y_t - \mu - \delta_2 d_{1980Q2,t} = \phi_1(y_{t-1} - \mu - \delta_2 d_{1980Q2,t-1}) \\ + \phi_2(y_{t-2} - \mu - \delta_2 d_{1980Q2,t-2}) + \delta_1 d_{1978Q2,t} + \varepsilon_t$$

by conditional least squares gives coefficient estimates (with standard errors in parentheses) of $\hat{\mu} = 1.15(0.43)$, $\hat{\delta}_1 = 6.63(1.81)$, $\hat{\delta}_2 = -6.22(1.63)$, $\hat{\phi}_1 = 0.43(0.09)$, and $\hat{\phi}_2 = 0.18(0.09)$, with an $R^2 = 0.42$. Hence, the fit improves substantially compared to the model without any outlier treatment, see part a.

With a skewness equal to -0.31 and kurtosis equal to 3.26, the Jarque-Bera test for normality of the residuals gives a p-value of 0.31. The (partial) autocorrelation at lag 8 is reduced to $\hat{\rho}_8(\hat{\varepsilon}) = -0.204$ ($\hat{\psi}_8(\hat{\varepsilon}) = -0.181$). The same applies to the squared residuals, for which $\hat{\rho}_8(\hat{\varepsilon}^2) = 0.062$, for example. Hence, the properties of the residuals improve compared to the model without any outlier treatment, see part a.

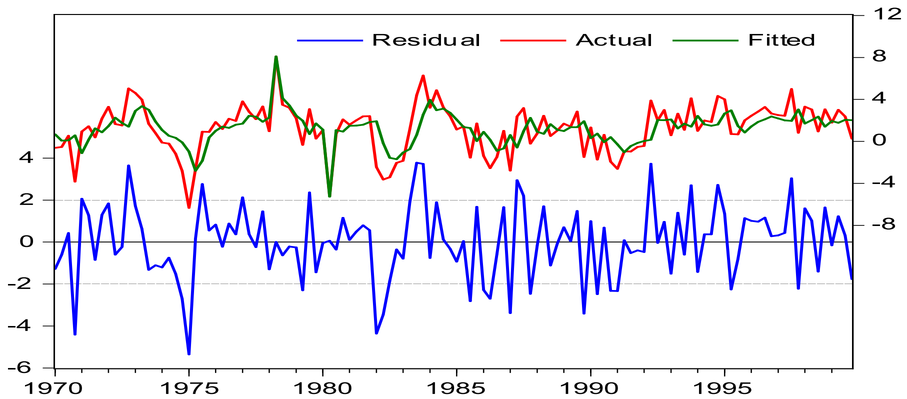


Figure 4: Residuals in AR(2)-AO model for quarterly investment growth rates

Exercise 6.7

The Excel file `TOTALSA.xlsx` contains monthly observations on (seasonaly adjusted) total vehicle sales (in millions of units) in the US over the period January 1988 - April 2019 (376 observations).

a. Examine the basic statistical properties of this time series for the sample period January 1988 - December 2007 (240 observations). Address at least the following three questions, and for each of these questions also examine the influence that the observations in October 2001 and July 2005 have on your answers:

   1. What do the (partial) autocorrelations look like?
   2. Does the time series contain a stochastic trend?
   3. Is it reasonable to assume a normal distribution for this time series?

b. Estimate an AR(4) model (allowing for a non-zero unconditional mean) for the time series of total vehicle sales for the sample period January 1988 - December 2007. Evaluate the model by inspecting the properties of the residuals (check (i) normality; (ii) [partial] autocorrelations; (iii) homoskedasticity). What kind of patterns do you observe in the residuals around October 2001 and July 2005?

c. Using the AR(4) model as estimated in part b., compute the outlier detection statistics $\hat{\lambda}_{AO}$ and $\hat{\lambda}_{IO}$, as discussed in Chapter 6. Use simulation to obtain the distribution of these test statistics under the null hypothesis of no outliers. Use this simulated distribution to evaluate the significance of the values of $\hat{\lambda}_{AO}$ and $\hat{\lambda}_{IO}$ that you found for the empirical time series. Are the results in line with the patterns observed in the AR(4)-residuals, in particular around October 2001 and July 2005?

d. Re-estimate the AR(4) model for the sample period January 1988 - December 2007, but now possibly treating the observations for October 2001 and July 2005 as outliers. For each of these observations, make a motivated choice for treating it either as an additive outlier (AO) or as an innovation outlier (IO) (or for not treating it as an outlier at all). Evaluate the model in the usual way, in particular by inspecting the properties of the residuals. Also, address the question whether it is sufficient to treat only those two observations as outliers, or whether there are more 'aberrant observations' that might need treatment?

e. Use the two estimated AR(4) models (from parts b. and d.) to obtain one-step ahead point forecasts for total vehicle sales for the period January 2008 - April 2019. Evaluate the quality of these forecasts, by examining their (i) bias, (ii) accuracy, and (iii) efficiency. Also compare the relative accuracy of the forecasts by means of the Diebold-Mariano statistic. Is it useful to treat the observations in October 2001 and July 2005 as outliers from a forecasting perspective? How does the period January 2008 - December 2009 affect your conclusions (that is, how do the different forecasts perform for the period January 2010 - April 2019)?

Solution

a. The time series of monthly total vehicle sales is shown in Figure 1 (for the complete sample period). The time series shows several interesting features. Very prominent is of course the huge decline before and during the most recent recession, from 17 million in January 2007 to slightly more than 9 million in the first months of 2009. We also observe several 'spikes' in the time series. Specifically, in October 2001, July 2005 and August 2009, the amount of vehicle sales was much higher than expected given the preceding observations. Obviously, these observations are 'candidate' outliers. Note that these spikes occur for good reasons, in the sense that in each of these three months, vehicle sales were stimulated either through actions taken by large car producers or a government policy/program. Specifically, in October 2001 the three large US car producers (General Motors, Ford, DaimlerChrysler) offered interest-free financing of new vehicle purchases. In July 2005, the big three gave all customers large discounts on new cars, which normally were available only to employees. And finally, in August 2009, the US government activated the so-called CARS program, providing a subsidy of $4000 to consumers buying a new more fuel-efficient car.

Figure 1 also shows that the time series rarely crosses its (sample) mean. Hence, mean-reversion (if any) is rather slow, and in fact it seems relevant to test for the presence of a stochastic trend. For this purpose we apply the (Augmented) Dickey-Fuller statistic, using the sample period January 1988 - December 2007. Because the time series does not contain a pronounced trend, we only include an intercept in the test regression. Using the Schwarz Information Criterion (SIC) to select the appropriate lag order p (which gives $p = 4$, that is, three lagged first differences are included), the ADF test statistic takes the value -1.67 with a p-value of 0.45,[1]

[1] Using the Akaike Information Criterion (AIC) gives $p = 11$ and an ADF test statistic of -1.22 with a p-value of 0.67.

such that we do not reject the null hypothesis of a stochastic trend. This result is influenced quite substantially by the two potential outliers in October 2001 and July 2005 though: when we include dummy variables for these two observations in the ADF test regression, we find a test statistic of -2.66 with a p-value of 0.082, such that we may reject the null hypothesis at the 10% significance level.

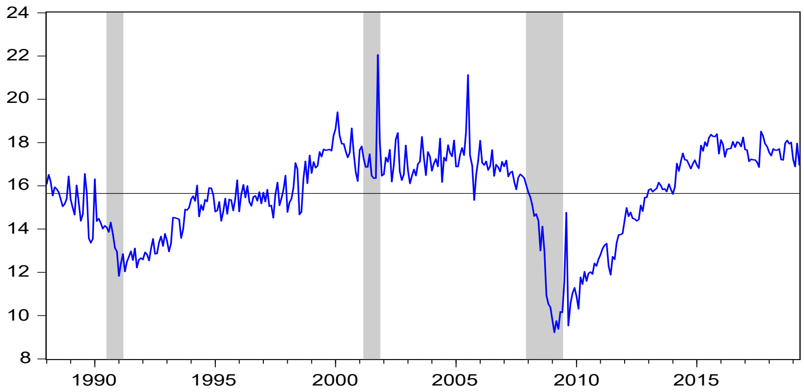


Figure 1: Monthly US total vehicle sales, January 1988-April 2019.
Shaded areas are NBER-defined recessions; the horizontal lines indicates the sample mean

Figure 2 shows the histogram of the time series (plus a normal density with the same mean and variance) for the sample period January 1988 - December 2007. Skewness over this period is equal to -0.12, while kurtosis is equal to 3.35. The Jarque-Bera test statistic takes a value of 1.74, with p-value 0.42, such that normality cannot be rejected. Again, the observations in October 2001 and July 2005 are quite important for this conclusion: excluding these from the sample, skewness turns negative at -0.46 while kurtosis declines to 2.52. This results in a Jarque-Bera statistic of 10.48, such that normality can be rejected even at the 1% significance level.

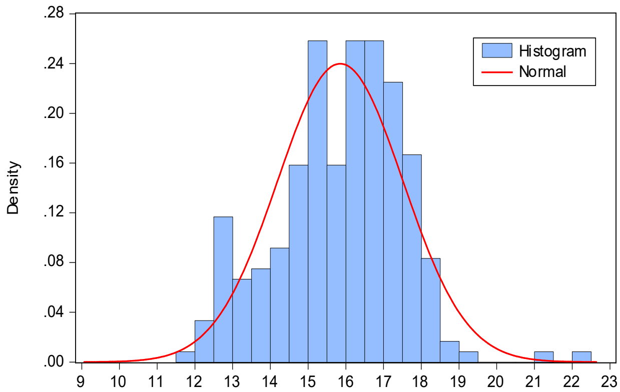


Figure 2: Histogram of monthly vehicle sales, January 1988-December 2007.

The first 20 (partial) autocorrelations are shown in the second and third column of the table below, with an asterisk indicating statistical significance at the 5% level.[2] We find significant partial autocorrelations at lags 1-6, 10, and 15, and an autocorrelation function with all positive values and with magnitudes that gradually decline towards zero. Again, the two observations in October 2001 and July 2005 are quite influential. Omitting these observations from the sample period changes the (partial) autocorrelations, although the overall pattern remains the same. The effects also go in different directions: For example, the first three partial autocorrelations increase (to 0.883, 0.279, 0.293) while the fourth declines (to 0.291). The partial autocorrelations at orders six and 15 become very close to zero (0.007 and 0.006), such that they are no longer significant.

	Time series		AR(4) residuals		AR(4)-AO residuals	
k	AC	PAC	AC	PAC	AC	PAC
1	0.853*	0.853*	−0.048	−0.048	−0.051	−0.051
2	0.788*	0.220*	−0.076	−0.078	−0.052	−0.054
3	0.780*	0.261*	−0.061	−0.070	−0.042	−0.048
4	0.818*	0.351*	−0.144*	−0.159*	−0.139*	−0.148*
5	0.812*	0.135*	−0.045	−0.077	0.044	0.023
6	0.798*	0.132*	0.122	0.087	0.091	0.079
7	0.777*	0.048	0.054	0.039	0.049	0.052
8	0.788*	0.118	0.067	0.065	0.105	0.108
9	0.794*	0.091	0.178*	0.206*	0.148*	0.195*
10	0.744*	−0.174*	−0.069	0.009	−0.096	−0.032
11	0.710*	−0.098	−0.062	0.002	−0.034	−0.006
12	0.711*	−0.043	−0.016	0.013	−0.084	−0.069
13	0.721*	−0.032	0.031	0.072	0.139*	0.153*
14	0.716*	0.015	0.155*	0.161*	0.096	0.054
15	0.670*	−0.141*	−0.040	−0.074	−0.007	−0.029
16	0.655*	0.008	0.016	0.021	0.089	0.081
17	0.641*	−0.077	−0.007	0.006	−0.060	−0.024
18	0.622*	−0.102	−0.111	−0.118	−0.079	−0.083
19	0.614*	0.075	0.013	−0.005	−0.001	−0.017
20	0.609*	0.036	0.098	0.057	−0.002	−0.003

b. We estimate an AR(4) model $\phi_4(L)(y_t - \mu) = \varepsilon_t$, where y_t denotes the monthly total vehicle sales, and $\phi_4(L) = 1 - \phi_1 L - \phi_2 L^2 - \phi_3 L^3 - \phi_4 L^4$. Using the observations for $t =$May 1988-December 2007 (because the observations in January-April 1988 are used as 'pre-sample' observations), we find the following estimates, with standard errors in parentheses: $\hat{\mu} = 15.90(0.93)$, $\hat{\phi}_1 = 0.514(0.062)$, $\hat{\phi}_2 = 0.032(0.070)$, $\hat{\phi}_3 =$

[2]The sample period January 1988-December 2007 contains 240 observations, such that the standard error of the partial autocorrelations is equal to 0.0645.

0.048(0.070), and $\hat{\phi}_4 = 0.352(0.062)$.[3] Note that the estimated coefficients for the second and third lags are not significantly different from zero even at the 10% level. The R^2 of the model is equal to 0.79.

The residuals from the AR(4) model have skewness equal to 1.56 and kurtosis equal to 13.13. Normality is clearly rejected as the Jarque-Bera statistic is equal to 1105, with a p-value of 0.000. The (partial) autocorrelations of the residuals are shown in the fourth and fifth column of the table above; for both, we find significant value at $k = 4$, 9 and 14. For the squared residuals, we find very small and insignificant (partial) autocorrelations (the largest value is 0.070 for $k = 1$), such that heteroskedasticity does not seem to be a relevant issue.

The observations in October 2001 and July 2005 have a rather large influence on these properties of the AR(4) residuals. Leaving the residuals for these two months out of consideration, we find that skewness and kurtosis are equal to -0.13 and 3.67. This results in a Jarque-Bera statistic of 5.01 with p-value 0.08, such that it becomes doubtful whether normality actually should be rejected or not. For the (partial) autocorrelations, we find smaller values at virtually all orders. For example, for $k = 4$, 9 and 14 we find autocorrelations equal to -0.132, 0.138 and 0.116, and partial autocorrelations equal to -0.136, 0.128 and 0.085. Given the standard error of 0.065, it is not so obvious whether the residuals actually show significant autocorrelation. Interestingly, for the autocorrelations of the squared residuals we find an opposite effect: Dropping the residuals in October 2001 and July 2005 results in a second-order autocorrelation equal to 0.279, which is significantly different from zero. Whether heteroskedasticity really is a relevant issue requires further investigation.

[3] These are estimates obtained with conditional least squares. In case the parameters are estimated with Maximum Likelihood, we find slightly different results: $\hat{\mu} = 15.93(0.91)$, $\hat{\phi}_1 = 0.513(0.048)$, $\hat{\phi}_2 = 0.030(0.067)$, $\hat{\phi}_3 = 0.048(0.060)$, and $\hat{\phi}_4 = 0.347(0.056)$.

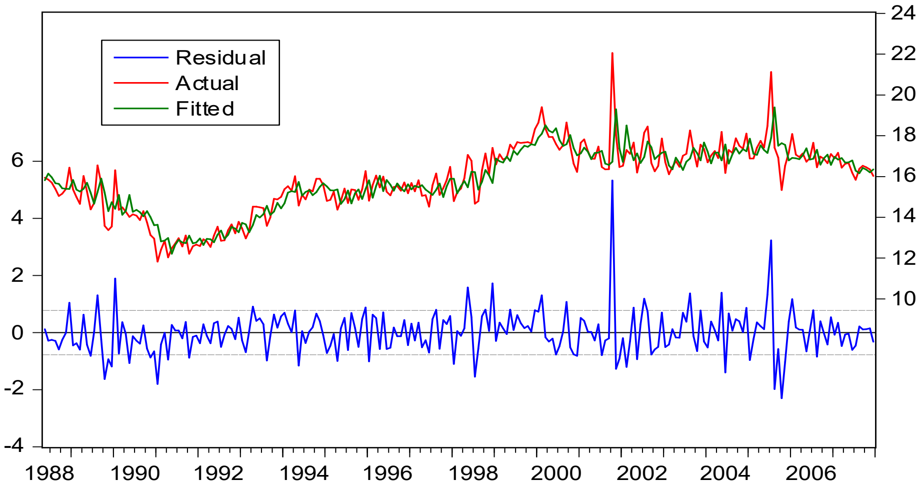


Figure 3: Residuals in AR(4) model for monthly vehicle sales

The residuals are shown in Figure 3. For October 2001, we observe a large positive residual, followed by (relatively) large negative residuals in November 2001 and February 2002. This matches quite well with the patterns we may expect from an additive outlier (AO) in an AR(4) model: An AO of magnitude ζ occurring at $t = \tau$ has an effect equal to $-\phi_j\zeta$ on the residuals in periods $t = \tau + j$ for $j = 1, 2, 3$ and 4. Given that $\hat{\phi}_1 = 0.514$ and $\hat{\phi}_4 = 0.352$, we would indeed expect a substantial negative impact on the residuals in November 2001 and February 2002 following the positive AO in October 2001.

For July 2005 we also find a very large positive residual, followed by a number of rather large negative residuals for the months August-November 2005. While this suggests that also July 2005 may be an AO, this conclusion is in fact not so clear-cut: The sequence of large negative residuals in subsequent months does not match very well with the pattern that might be expected from an AO given the estimates of the AR(4) coefficients. In particular, given the small estimates of ϕ_2 and ϕ_3, we would expect to find large negative residuals for August and November, but not for September and October 2005.

Note that the possible different character of the outliers in October 2001 and July 2005 is also apparent in the time series plot in Figure 1. October 2001 seems like a classic example of an AO, with a single isolated spike in the time series. Things seem to be different for July 2005, in the sense that the large upward spike in that month is followed by several months with sales levels below normal. Hence, it might be that a substantial number of people simply advanced ther car purchase to benefit from the employee discount program.

c. Based on the AR(4) model estimated in part b., we compute the t-statistics $\hat{\lambda}_i(\tau)$ for testing whether an additive outlier (AO) or innovation outlier (IO) occurred at time τ, for $\tau = 1, \ldots, T$. These sequences of test statistics are shown in Figure 4.

The largest absolute value of the AO test statistic is obtained for October 2001, with $\hat{\lambda}_{\mathrm{AO}}(\tau) = 8.02$. This is not surprising, given the patterns observed in the AR(4) residuals for this and subsequent observations, see 3.b. The largest absolute value of the IO test statistic is also obtained for October 2001, with $\hat{\lambda}_{\mathrm{IO}}(\tau) = 7.78$.

For July 2005, we find $\hat{\lambda}_{\mathrm{AO}}(\tau) = 5.56$ and $\hat{\lambda}_{\mathrm{IO}}(\tau) = 4.37$. For both test statistics, these are the second-largest values following the ones obtained for October 2001. Strictly speaking, we should first accommodate the outlier in October 2001 in the model, and then redo the testing procedure to see whether there is any evidence for the presence of additional outliers. Given the values for $\hat{\lambda}_{\mathrm{AO}}(\tau)$ and $\hat{\lambda}_{\mathrm{IO}}(\tau)$ found here for July 2005, it is quite likely that this observation would come out as the next 'candidate outlier', and given the magnitude of the test statistics, we would probably find that it is significant. Given that $\hat{\lambda}_{\mathrm{AO}}(\tau)$ is substantially larger than $\hat{\lambda}_{\mathrm{IO}}(\tau)$, the corresponding p-value for $\hat{\lambda}_{\mathrm{AO}}(\tau)$ can be expected to be the smaller of the two. This could be interpreted as suggesting that it is more appropriate to characterize the outlier in July 2005 as an AO rather than as an IO, although the pattern in the residuals in subsequent months does not really match the expected pattern for an AO, as discussed before.

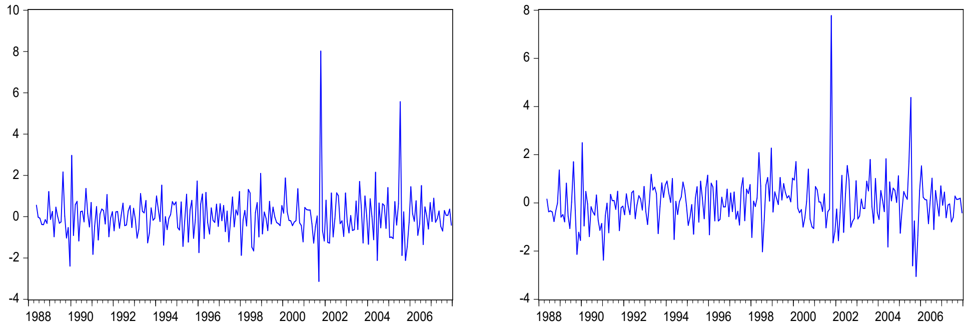


Figure 4: Values of the $\hat{\lambda}_i(\tau)$ statistics for testing whether an AO or IO occurred at time τ are shown on the left and right, respectively.

We use simulation to obtain the (finite-sample) distribution of the test statistics $\hat{\lambda}_{\mathrm{AO}} = \max_{1 \le \tau \le T} |\hat{\lambda}_{\mathrm{AO}}(\tau)|$ and $\hat{\lambda}_{\mathrm{IO}} = \max_{1 \le \tau \le T} |\hat{\lambda}_{\mathrm{IO}}(\tau)|$ under the null hypothesis of no outliers, using 50,000 replications. The results of this exercise are shown in Figure 5. For both test statistics we find very small p-values. In fact, none (only one) of the simulated test statistics exceeds the empirical values for the AO (IO) test statistic. We therefore reject the null hypothesis of no outliers.

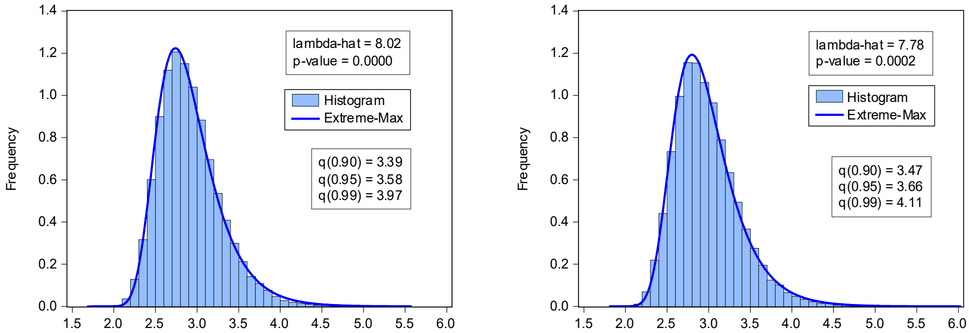


Figure 5: Simulated distribution under the null hypothesis of $\hat{\lambda}_i(\tau)$ statistics for testing for the presence of an AO or IO are shown on the left and right, respectively.

d. The patterns in the AR(4) residuals as discussed in part b. together with the results of the outlier test statistics in part c. indicate that October 2001 corresponds with an AO. Indeed, treating this observation as such gives a better fit of the observations following this month compared to treating it as an IO. This is visualized in the left panel of Figure 6, showing the residuals for the original AR(4) model (without any outlier treatment) and AR(4) models where October 2001 is treated as either AO or IO for the period January 2001 - December 2002. It can be seen that treating October 2001 as IO only eliminates the large positive residual occurring in that month. Treating October 2001 as AO also reduces the magnitude of the residuals in subsequent months (especially November 2001 and February 2002).

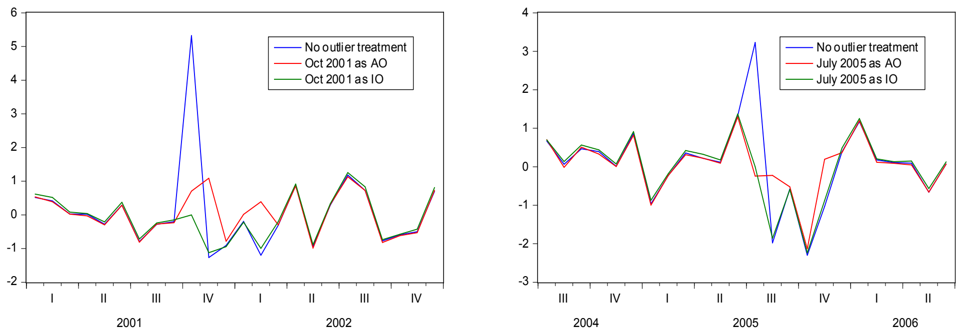


Figure 6: Residuals in AR(4) models for monthly vehicle sales in selected periods.

For July 2005 similar considerations apply. The patterns in the AR(4) residuals and the outlier test statistics indicate that it most likely corresponds with an AO, although as discussed before, the residual pattern does not align perfectly with the expected effects based on the AR(4) coefficient estimates. The right panel of Figure 6 shows the residuals for the original AR(4) model (without any outlier treatment) and AR(4) models where July 2005 is treated as either AO or IO for the period July 2004 - June 2006. This shows quite convincingly that treating this observation as an AO gives a better fit of the observations in following months compared to treating it as an IO. Specifically, treatment as an IO only eliminates the large positive residual occurring in July 2005 itself. Treatment as AO also reduces the magnitude of the residuals in subsequent months, especially August and November 2005.[4]

The preferred AR(4)-AO model is thus given by

$$\phi_4(L)(y_t - \mu - \delta_1 d_{2001M10,t} - \delta_2 d_{2005M7,t}) = \varepsilon_t,$$

where $\phi_4(L) = 1 - \phi_1 L - \phi_2 L^2 - \phi_3 L^3 - \phi_4 L^4$, and where $d_{2001M10,t}$ and $d_{2005M7,t}$ are dummy variables for October 2001 and July 2005, respectively. Using the observations for t =May 1988-December 2007, we find the following nonlinear least squares estimates for the coefficients in this model, with standard errors in parentheses: $\hat{\mu} = 15.89(1.07)$, $\hat{\phi}_1 = 0.538(0.063)$, $\hat{\phi}_2 = 0.013(0.072)$, $\hat{\phi}_3 = 0.082(0.072)$, $\hat{\phi}_4 = 0.328(0.063)$ $\hat{\delta}_1 = 4.637(0.540)$, and $\hat{\delta}_2 = 3.411(0.545)$. We find positive values for $\hat{\delta}_1$ and $\hat{\delta}_2$ as expected, given that vehicle sales experienced large positive 'spikes' in both quarters October 2001 and July 2005.

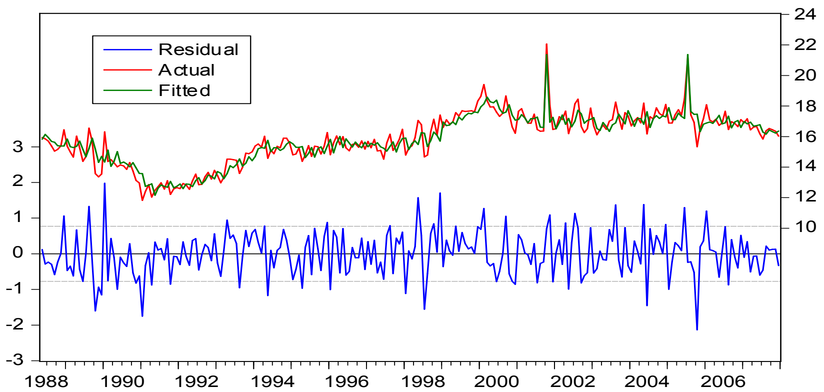


Figure 7: Residuals in AR(4)-AO model for monthly vehicle sales.

The R^2 of the AR(4)-AO model is equal to 0.86, which is quite a bit larger than the R^2 for the AR(4) model estimated in 3.b, also given that only two observations are treated as outliers. The residuals from the AR(4) model are shown in Figure 7. Comparing this with the 'standard' AR(4) model, the main differences (obviously)

[4]Note that treatment as AO leaves large residuals in September and October 2005 though. This corresponds with the rather small estimates of ϕ_2 and ϕ_3 in the AR(4) model. In order to accommodate these effects, it would be necessary to treat July 2005 as a 'transient change' type outlier.

occur in the residuals for October 2001 and July 2005 and subsequent months, see also Figure 6. In the AR(4)-AO model, the residuals have skewness equal to -0.019 and kurtosis equal to 3.59. Normality is not rejected, as the Jarque-Bera statistic is equal to 3.43, with a p-value of 0.18. The (partial) autocorrelations of the residuals are shown in the table above; Significant values still occur at lags 4 and 9, and now also 13. For the (partial) autocorrelations of the squared residuals we find no significant values.

Finally, comparing the residuals in Figure 7 with those in Figure 3 shows that the largest residuals in October 2001 and July 2005 obviously have disappeared die to the outlier treatment. However, several other large residuals remain; for example in January 1990, December 1998 and (still) October 2005. It may be worthwhile to examine whether some of these observations also should be considered as outliers.

e. We use the 'standard' AR(4) model from part b. and the AR(4)-AO model from part d. to obtain one-step ahead forecasts for the monthly vehicle sales for the period January 2008 - April 2019.

The two series of forecasts are shown below, together with the actual values and the resulting forecast errors. The two forecasts seem almost identical (which is not very surprising because the estimates of the autoregressive parameters are quite similar in the two models). The MSPE's are equal to 0.586 for the AR(4) model and 0.571 for the AR(4)-AO model. Hence, incorporating outliers in the model actually leads to somewhat more accurate forecasts, reducing the MSPE by 2.6 percentage points. Applying the Diebold-Mariano test of equal predictive ability, we find a (one-sided) p-value of 0.033, indicating that this difference in MSPE is significant.

Both models produce unbiased forecasts, in the sense that the mean forecast errors of -0.019 and -0.009 (for the AR(4) and AR(4)-AO models, respectively) are not significantly different from 0 at the 5% significance level. The forecasts from both models seem to be efficient: for the Mincer-Zarnowitz regression $y_{t+1} = \alpha + \beta \hat{y}_{t+1|t} + \eta_{t+1}$, we find $\hat{\alpha} = -0.528(0.418)$ and $\hat{\beta} = 1.033(0.027)$ for the AR(4) model and $\hat{\alpha} = -0.273(0.408)$ and $\hat{\beta} = 1.017(0.026)$ for the AR(4)-AO model (with standard errors in parentheses). In both case, $\hat{\alpha}$ and $\hat{\beta}$ are not significantly different from 0 and 1, respectively.

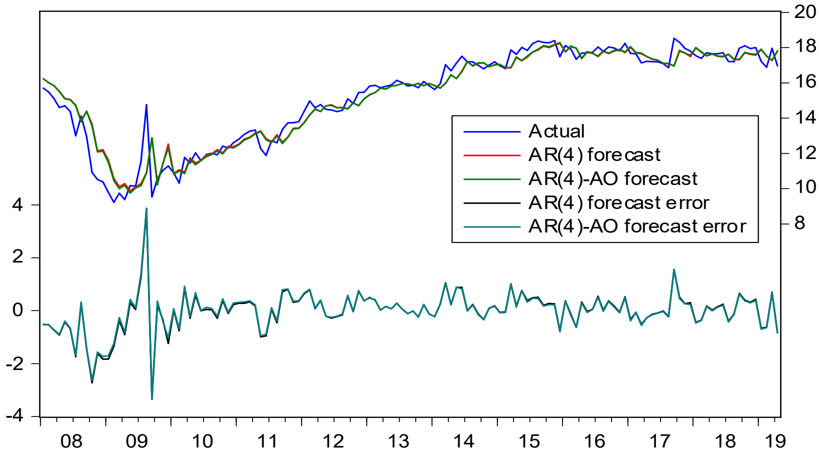


Figure 8: Forecasts for monthly vehicle sales, Jan 2008-Apr 2019.

The first two years of the out-of-sample period contain the large decline in the level of vehicle sales and a possible outlier in August 2009. Not surprisingly then, this period has a strong impact on the forecasting performance. For the period January 2010 - April 2019, both sets of forecasts are much more accurate, with MSPE values equal to 0.2065 and 0.2056 for the AR(4) and AR(4)-AO models, respectively. The difference in MSPE is no longer significant, in the sense that the (one-sided) p-value of the DM statistic now is equal to 0.40. Both sets of forecasts also seem somewhat biased, with mean forecast errors equal to 0.125 and 0.124 (with standard error of 0.041 in both cases).[5]

In sum, treating the observations in October 2001 and July 2005 as outliers is somewhat useful from a forecasting perspective as it gives a significant reduction in MSPE; but this conclusion is quite sensitive to which period is used for forecast evaluation.

[5]Given that over the full out-of-sample period 2008-2019 we find that the forecasts are biased, this also suggests that the forecasts are particularly strongly biased during the years 2008-2009. This is indeed the case: both forecasts are too high, failing to predict the large decline in vehicle sales. Mean forecast errors are equal to -0.692 and -0.628 with standard error 0.286.

(International) Bachelor Econometrics and Operations Research
FEB23001(X)-19 Tijdreeksanalyse / Time Series Analysis

Exercises - Nonlinearity (Chapter 8)

Exercise 8.1

Consider the following model, where y_t is the quarterly growth rate of US GDP:

$$y_t = \phi_0 + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \theta_1 CDR_{t-1} + \varepsilon_t, \tag{1}$$

where $\varepsilon_t \sim \text{iid}N(0, \sigma^2)$, and where CDR_t is defined as follows:

$$CDR_t = Y_t - \max_{j \geq 0} Y_{t-j}, \tag{2}$$

where Y_t denotes the logarithm of the level of US GDP.

Estimating the parameters in the above model using observations for the period 1954Q1-2003Q4 gives estimates with the following properties: $\hat{\phi}_0 > 0$, $0 < \hat{\phi}_1 < 1$, $0 < \hat{\phi}_2 < 1$, $\hat{\phi}_1 + \hat{\phi}_2 < 1$, and $\hat{\theta}_1 < 0$.

Describe in as much detail as possible which features of the US GDP growth are captured by this model. In particular, what is the function of the variable CDR_{t-1} in the model? How would you label the regimes consisting of observations for which $CDR_t = 0$ and consisting of observations for which $CDR_t < 0$?

Solution

$\Rightarrow$ Note that CDR_t is the difference between the current level of GDP and the historic maximum up to time t. Hence, CDR_t is negative during recessions and perhaps during the first few quarters of expansions, and equal to zero during the remainder of the expansion periods.

The model in (1) can be interpreted as an AR(2) model with a time-varying intercept equal to $\phi_0 + \theta_1 CDR_{t-1}$. Hence, the model allows for different intercepts, or different average growth rates, in the regimes corresponding to $CDR_{t-1} = 0$ and $CDR_{t-1} < 0$. Given that $\theta_1 < 0$ and that $CDR_{t-1} < 0$ during recessions and early expansions, this leads to a higher value of the intercept during those periods and, hence, higher average growth. This may be seen as an attempt to capture the "bounce-back" -effect of output.

Exercise 8.2
Consider the Switching AR(1) model

$$y_t = (\mu_0 + \phi_0 y_{t-1}) + (\mu_1 + \phi_1 y_{t-1})s_t + \varepsilon_t$$

for $t = 1, \ldots, T$ with $\varepsilon_t \sim N(0, \sigma^2)$, where s_t is a latent binary random variable with

$$\Pr[s_t = 1] = F(\delta + \gamma y_{t-1}) = \frac{\exp(\delta + \gamma y_{t-1})}{1 + \exp(\delta + \gamma y_{t-1})}.$$

a. What is/are the key difference(s) between the Switching AR(1) model and the LSTAR(1) model given by

$$y_t = (\mu_0 + \phi_0 y_{t-1}) + (\mu_1 + \phi_1 y_{t-1})F(\delta + \gamma y_{t-1}) + \varepsilon_t$$

for $t = 1, \ldots, T$ with $\varepsilon_t \sim N(0, \sigma^2)$ or are both models the same?

b. Derive the unbiased 1-step ahead forecasts of y_{T+1} made at time T for the Switching AR(1) model and for the LSTAR(1) model.

c. Consider the Switching AR(1) model with

$$\Pr[s_t = 1] = \frac{\exp(\delta + \gamma s_{t-1})}{1 + \exp(\delta + \gamma s_{t-1})}.$$

Show that this model is the same as a Markov Switching AR(1) model where

$$\Pr[s_t = 1 | s_{t-1} = 1] = p \text{ and } \Pr[s_t = 0 | s_{t-1} = 0] = q.$$

Express p and q in terms of δ and γ.

d. Suppose that you have observed y_t for $t = 1, \ldots, T$. A recession is defined as a period where for at least 2 consecutive periods $s_t = 1$. Suppose that $S_T = 0$. Derive the probability in terms of p and q that the period $[T+1, T+3]$ contains a recession.

Solution:
a. In the LSTAR model, the regime is "deterministic" in the sense that the regime at time t is known with certainty when y_{t-1} is observed. In the Switching AR(1) model, the regime at time t is still uncertain / stochastic even when y_{t-1} is known.

b. For the LSTAR model, we have

$$\hat{y}_{T+1|T} = \mathsf{E}_T[y_{T+1}] = (\mu_0 + \phi_0 y_T) + (\mu_1 + \phi_1 y_T)F(\delta + \gamma y_T).$$

For the Switching AR(1) model, we have

$$\begin{aligned}\hat{y}_{T+1|T} &= \mathsf{E}_T[(\mu_0 + \phi_0 y_T) + (\mu_1 + \phi_1 y_T)s_{T+1} + \varepsilon_{T+1}] \\ &= (\mu_0 + \phi_0 y_T) + (\mu_1 + \phi_1 y_T)F(\delta + \gamma y_T),\end{aligned}$$

as $\mathsf{E}_T[s_{T+1}] = \Pr[s_{T+1} = 1]$. Hence, the forecasts from the two models are identical.

c. We simply have

$$\Pr[s_t = 1|s_{t-1} = 1] = \frac{\exp(\delta + \gamma)}{1 + \exp(\delta + \gamma)} = p,$$
$$\Pr[s_t = 0|s_{t-1} = 0] = 1 - \Pr[s_t = 1|s_{t-1} = 0] = 1 - \frac{\exp(\delta)}{1 + \exp(\delta)} = \frac{1}{1 + \exp(\delta)} = q.$$

d. A recession can occur in three different ways:

(i) $s_{T+1} = 1$, $s_{T+2} = 1$, $s_{T+3} = 0$,

(ii) $s_{T+1} = 1$, $s_{T+2} = 1$, $s_{T+3} = 1$,

(iii) $s_{T+1} = 0$, $s_{T+2} = 1$, $s_{T+3} = 1$.

As $s_T = 0$, the probabilities of each of those sequences of s_{T+1}, s_{T+2} and s_{T+3} is given by

(i) $(1 - q)p(1 - p)$,

(ii) $(1 - q)p^2$,

(iii) $q(1 - q)p$.

Exercise 8.3

Consider the following model, where y_t is the quarterly growth rate of US GDP:

$$y_t - \mu_t = \phi_1(y_{t-1} - \mu_{t-1}) + \varepsilon_t, \tag{3}$$

where $\varepsilon_t \sim \text{iid}N(0, \sigma^2)$, and μ_t is given by

$$\mu_t = \mu_0 + \mu_1 \mathtt{I}(S_t = 1) + \lambda \mathtt{I}(S_t = 0) \sum_{j=1}^{m} \mathtt{I}(S_{t-j} = 1), \tag{4}$$

where $\mathtt{I}(A)$ is the indicator function for the event A (that is, $\mathtt{I}(A) = 1$ if A is true, and $\mathtt{I}(A) = 0$ otherwise), and $S_t \in \{0, 1\}$ is a first-order Markov process with constant transition probabilities given by

$$P(S_t = 0|S_{t-1} = 0) = p \qquad \text{and} \qquad P(S_t = 1|S_{t-1} = 1) = q.$$

Estimating the parameters in the above model with $m = 6$ using observations for the period 1954Q1-2003Q4 gives $0 < \widehat{\phi}_1 < 1$, $\widehat{\mu}_0 > 0$, $\widehat{\mu}_0 + \widehat{\mu}_1 < 0$, $\widehat{\lambda} > 0$, $\widehat{p} = 0.94$, and $\widehat{q} = 0.78$.

Describe which features of the US GDP growth are captured by this model. In particular, how would you label the regimes $S_t = 0$ and $S_t = 1$? What is the function of the last component of μ_t in (4), that is $\lambda \mathtt{I}(S_t = 0) \sum_{j=1}^{m} \mathtt{I}(S_{t-j} = 1)$?

Solution

This is an AR(1)-model with time-varying mean μ_t. The positive estimate of the AR-parameter ϕ_1 reflects the positive autocorrelation in GDP growth rates.

From the specification of μ_t it follows that the unconditional mean for quarter t is equal to

- $\mu_0 + \mu_1$ if $S_t = 1$;
- μ_0 if $S_t = 0$ and $\sum_{j=1}^{m} \mathtt{I}(S_{t-j} = 1) = 0$
- $\mu_0 + \lambda \sum_{j=1}^{m} \mathtt{I}(S_{t-j} = 1)$ if $S_t = 0$ and $\sum_{j=1}^{m} \mathtt{I}(S_{t-j} = 1) > 0$

From the estimation results that $\widehat{\mu}_0 > 0$, $\widehat{\mu}_0 + \widehat{\mu}_1 < 0$, $\widehat{\lambda} > 0$, we conclude that $S_t = 0$ and $S_t = 1$ correspond with regimes of expansions and recessions, respectively. The last component of μ_t in (4), that is $\lambda \mathtt{I}(S_t = 0) \sum_{j=1}^{m} \mathtt{I}(S_{t-j} = 1)$, captures the 'bounce-back' effect, as it implies that average growth is higher immediately following a recession (given that the estimate of λ is positive). Suppose that a recession has lasted for four quarters; then average growth in the first three quarters of the subsequent expansion is equal to $\mu_0 + 4\lambda$, which gradually declines to μ_0 during the next four quarters.

Exercise 8.4
The Excel file `USGDP.xlsx` contains quarterly observations for US real GDP over the period 1958Q4-2018Q4 (241 observations).

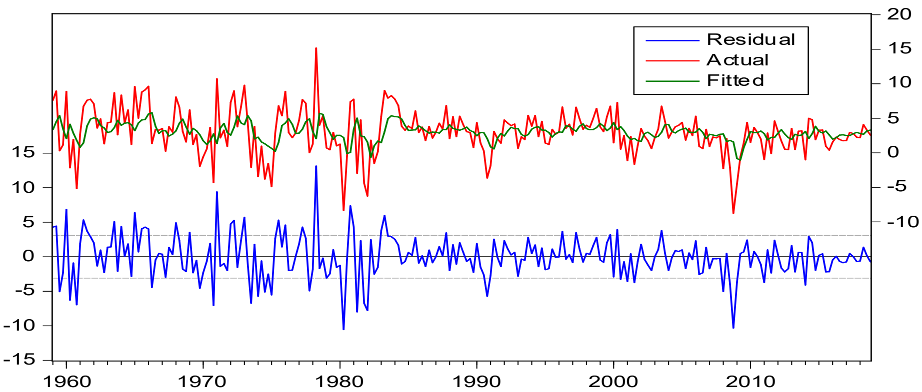
a. Estimate an AR(2) model for annualized quarterly real GDP growth rates and obtain the residuals. Test for a break in volatility of these residuals in 1984Q1, that is, test the null hypothesis $\mathsf{H}_0 : \delta_1 = \delta_2$ in the regression

$$\sqrt{\pi/2}|\hat{\varepsilon}_t| = \delta_1 \mathtt{I}[t \leq \tau] + \delta_2 \mathtt{I}[t > \tau] + \varepsilon_t, \quad t = 1, \dots, n, \tag{5}$$

where $\hat{\varepsilon}_t$ denotes the AR(2) residual for quarter t and τ corresponds to 1984Q1.

b. Apply Quandt's Sup$F$ test to test for a break in volatility of these residuals at an unknown date. Use a trimming fraction $\lambda = 0.05$.

Solution
a. Denote the quarterly growth rate in annualized percentage points by y_t, that is, $y_t = 400 \times \ln(Y_t/Y_{t-1})$, where Y_t is the level of real GDP. Estimating the AR(2) model $y_t - \mu = \phi_1(y_{t-1} - \mu) + \phi_2(y_{t-2} - \mu) + \varepsilon_t$ gives coefficient estimates (with standard errors in parentheses) of $\hat{\mu} = 3.04(0.36)$, $\hat{\phi}_1 = 0.25(0.057)$, and $\hat{\phi}_2 = 0.19(0.055)$, with an $R^2 = 0.12$. The growth rates, as well as the AR(2)-residuals, appear to exhibit a substantial reduction in volatility around the mid-1980s.



Estimating (5), yields estimates $\hat{\delta}_1 = 4.13(0.25)$ and $\hat{\delta}_2 = 1.89(0.20)$ with an $R^2 = 0.18$. The Wald-test for the restriction $\delta_1 = \delta_2$ equals 51.37 with p-value 0.000, such that we convincingly reject the null hypothesis of constant volatility.

b. The pointwise Wald test for a structural break in volatility based on (5) is shown below. The global maximum is reached when $\tau = 1984Q1$, where the test statistic is equal to 51.37. This is much larger than the critical values of the SupF test as derived by Andrews (1993), which equal 8.19, 9.84, and 13.01 for significance levels of 10, 5 and 1%, respectively. Hence, we reject the null hypothesis of constant volatility.