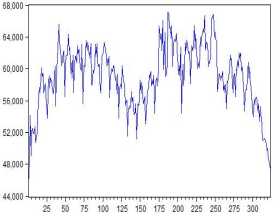
Seasonality and SARIMA models

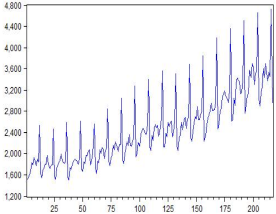
Nuno Sobreira

ISEG - Institute of Economics & Management

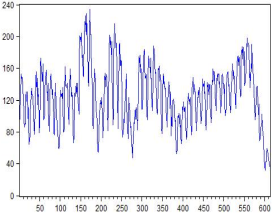
Examples of seasonal data



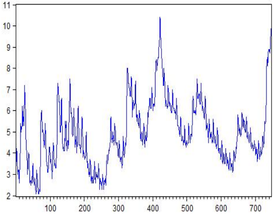
(a) Gasoline sales volume



(b) Licor sales (in millions of $)



(c) New Housing Starts



(d) Unemployment rate

Figure 1: Examples of (not seasonally adjusted) US time series data

Introduction I

- Throughout the sample, the time series illustrated in Figure 1 exhibit specific patterns in a certain period of time, say during a day, week, month or quarter that more a less repeat year after year or, in a broader sense, after a fixed time interval.
- This periodic behaviour is very common in time series and is denominated as seasonality.
- This seasonal component may be present when we have intra-annual data for our time series of interest. For example, our time series sample has quarterly, monthly, weekly or even daily frequency.
- There are many and strong motives to believe that most economic time series have a seasonal component:
  1. The consumption of gasoline rises during Summer due increased travelling by automobile;
  2. The international airline prices increase in Summer due to the holiday season;
  3. The electricity consumption rises during some periods of the Summer and Winter to control the building's temperature;

Introduction II

4. The private consumption increases in November and December due to Christmas season;
5. Construction activity and jobs decrease in Winter due to obstacles posed by cold, wind and rain.
6. The production of the agricultural goods strongly depends on the climate condition. Consequently, it has a strong seasonal component.
7. . . .

- These examples show that seasonality may have many different possible manifestations in a given time series.
- Naturally, time series reacts to these different possible seasonal patterns by proposing different modelling strategies for the seasonal component.
- For macro/financial data, the choice of the appropriate modelling technique depends if we consider the seasonality as:
  1. Deterministic Seasonality
  2. Stochastic Seasonality

Deterministic Seasonality I

- **Deterministic Seasonality** - assumes that the seasonal flutuations are more a less equal/similar year after year. The mean of one season may be different from another season. But year after year the pattern roughly repeats itself.
- A very extreme example is the sales of Chrismas trees which we expect to have a very similar pattern year by year, independently of the economic conditions.
- For example, with monthly data we have:

$$E(X_t) = \begin{cases} \mu_1, \text{ if month=January} \\ \mu_2, \text{ if month=February} \\ \vdots \\ \mu_{12}, \text{ if month=December} \end{cases}$$

Deterministic Seasonality II

- To allow for a different mean in each month we augment the econometric model with a dummy variable for each month:

$$D_{1,t} = \begin{cases} 1, \text{ if month=January} \\ 0, \text{ otherwise} \end{cases} \qquad D_{2,t} = \begin{cases} 1, \text{ if month=February} \\ 0, \text{ otherwise} \end{cases}$$

$$\ldots\ldots \qquad D_{12,t} = \begin{cases} 1, \text{ if month=December} \\ 0, \text{ otherwise} \end{cases}$$

- For a very simple and unrealistic model with no serial correlation we have:

$$\boxed{X_t = \sum_{s=1}^{12} \mu_s D_{s,t} + u_t, u_t \overset{w.n.}{\sim} \left(0, \sigma_u^2\right)}$$

- For an ARMA(p,q) model we have:

$$\boxed{X_t = \sum_{s=1}^{12} \mu_s D_{s,t} + u_t,\ \phi\left(L\right) u_t = \theta\left(L\right) \varepsilon_t,\ \varepsilon_t \overset{w.n.}{\sim} \left(0, \sigma_\varepsilon^2\right)}$$

Stochastic Seasonality I

- The most standard line of thought is to consider **seasonality** as **stochastic**. Here, we observe seasonal time persistence but the seasonal patterns change over time.
- For example, tourism expenditures are seasonal but also shift according to the disposable income that depends on the business cycle.
- In this case, we need a different approach to model the seasonality component of the time series of interest.
- A possible path is to use **automatic procedures** to remove the seasonality component. The most popular are the **TramoSeats** and **Census X12-ARIMA** (implementable in EViews in Proc→ Seasonal adjustment). For more details consult, for example, http://www.census.gov/srd/www/x12a/ or EViews manual.
- Then we apply the standard **Box-Jenkins methodology** to fit an **ARIMA model** to the already seasonally adjusted data.

Stochastic Seasonality II

- However, there are two **main disadvantages** with this approach:

  (a) Many times, not all seasonal effects are removed with these automatic procedures.

  (b) Automatic procedures are not efficient. As argued by Bell and Hilmer(1984), it is more efficient to analyse and model jointly the seasonal and nonseasonal components of the time series of interest.

- The more efficient approach advocated in (b) is the one followed by the **SARIMA** class of **models**. We study in detail this models in this group of slides.
- The most general form of the SARIMA class models the time dependence according to two different dimensions:
  1. **Nonseasonal dependence** - relationship between observations for successive “seasons" (months, quarters,. . .) in a particular year;
  2. **Seasonal dependence** - relationship between the observations for the same “season" (month, quarter,. . .) in successive years;

Stochastic Seasonality III

- Examples of pure seasonal models:
  1. For quarterly data:

$$X_t = \Phi X_{t-4} + \varepsilon_t, \ \varepsilon_t \overset{w.n.}{\sim} (0, \sigma_\varepsilon^2)$$

$$X_t = \varepsilon_t - \Theta \varepsilon_{t-4}, \ \varepsilon_t \overset{w.n.}{\sim} (0, \sigma_\varepsilon^2)$$

  2. For monthly data:

$$X_t = \Phi X_{t-12} + \varepsilon_t, \ \varepsilon_t \overset{w.n.}{\sim} (0, \sigma_\varepsilon^2)$$

$$X_t = \varepsilon_t - \Theta \varepsilon_{t-12}, \ \varepsilon_t \overset{w.n.}{\sim} (0, \sigma_\varepsilon^2)$$

Stochastic Seasonality IV

- Examples of (not pure) seasonal models:
  1. For quarterly data:

$$X_t = \phi X_{t-1} + \Phi X_{t-4} + \varepsilon_t,\ \varepsilon_t \overset{w.n.}{\sim} (0, \sigma_\varepsilon^2)$$

$$X_t = \varepsilon_t - \theta \varepsilon_{t-1} - \Theta \varepsilon_{t-4},\ \varepsilon_t \overset{w.n.}{\sim} (0, \sigma_\varepsilon^2)$$

  2. For monthly data:

$$X_t = \phi X_{t-1} + \Phi X_{t-12} + \varepsilon_t,\ \varepsilon_t \overset{w.n.}{\sim} (0, \sigma_\varepsilon^2)$$

$$X_t = \varepsilon_t - \theta \varepsilon_{t-1} - \Theta \varepsilon_{t-12},\ \varepsilon_t \overset{w.n.}{\sim} (0, \sigma_\varepsilon^2)$$

Stochastic Seasonality V

- The model building procedure for the SARIMA class follows the same steps as the (non seasonal) ARIMA. Recall that the Box-Jenkins methodological principles are:
  1. **Tentative identification** - Start by taking (seasonal and nonseasonal) first-differences to stationarize the time series if the time series is non stationary both at seasonal and non seasonal frequencies. As a practical matter we take, at most, one seasonal and one nonseasonal difference (in very rare occasions we may use two). After this process, examine carefully the SACF and SPACF and select different candidate seasonal ARMA models that are compatible with the SACF/SPACF.
  2. **Estimation of the SARIMA model**
  3. **Diagnostic checking (residuals)**
- However, the use of the SACF and SPACF in the tentative idenfication stage is more complicated with seasonal time series. This is due to the interaction between the seasonal and the nonseasonal ARMA components.

Stochastic Seasonality VI

- Throughout the next slides we present the main theoretical properties for the most important SARIMA models. The knowledge of these properties is very useful for interpretation and for the tentative identification stage.
- We illustrate the applicability of these results with the application of the Box-Jenkins methodology to a real dataset with a clear seasonal component.

Seasonal Moving Average process $\text{SMA}(Q)_S$

- The Seasonal Moving Average process $SMA(Q)_S$ is defined by the following equation:

$$\boxed{X_t = \Theta_0 + \varepsilon_t - \Theta_1 \varepsilon_{t-S} - \Theta_2 \varepsilon_{t-2S} - \ldots - \Theta_Q \varepsilon_{t-QS}, \; \varepsilon_t \overset{w.n.}{\sim} (0, \sigma_\varepsilon^2)}$$

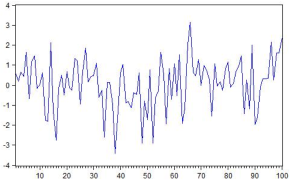
- Using the lag operator, L, this model can rewritten in a more simplified form as:

$$\boxed{X_t = \Theta_0 + \Theta(L^S)\varepsilon_t}$$

where $\Theta(L^S) = 1 - \Theta_1 L^S - \Theta_2 L^{2S} - \ldots - \Theta_Q L^{QS}$.

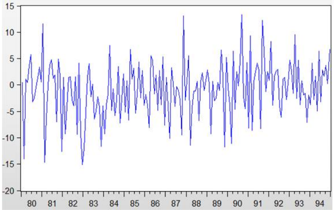
- X_t is **invertible** if the (inverse) roots of the MA polynomial, $\Theta(L^S)$ are outside (inside) the unit circle.

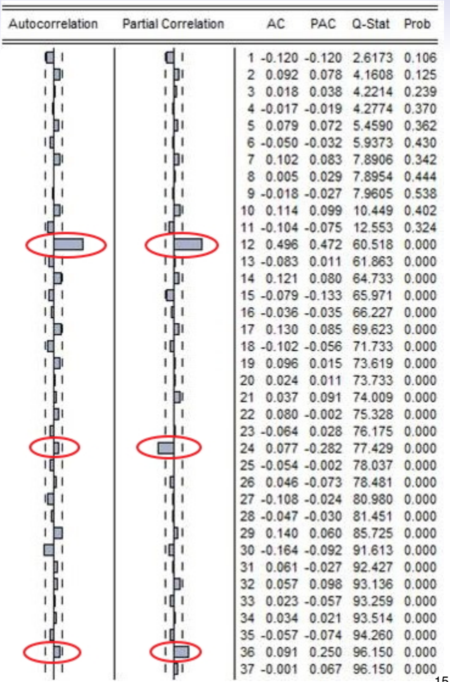
$SMA(1)_4$ process



Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.006	-0.006	0.0124	0.911
		2	0.069	0.069	1.9300	0.381
		3	-0.002	-0.001	1.9314	0.587
		4	0.450	0.447	84.044	0.000
		5	-0.041	-0.045	84.721	0.000
		6	-0.015	-0.078	84.807	0.000
		7	0.000	0.007	84.807	0.000
		8	-0.095	-0.368	88.514	0.000
		9	-0.011	0.048	88.561	0.000
		10	-0.006	0.078	88.574	0.000
		11	0.037	0.034	89.127	0.000
		12	-0.072	0.217	91.301	0.000
		13	0.040	0.001	91.982	0.000
		14	-0.015	-0.099	92.077	0.000
		15	0.057	0.032	93.455	0.000
		16	-0.047	-0.209	94.385	0.000
		17	-0.014	-0.061	94.464	0.000
		18	-0.093	-0.031	98.129	0.000
		19	0.012	-0.030	98.193	0.000
		20	-0.061	0.105	99.773	0.000
		21	-0.028	0.056	100.10	0.000
		22	-0.114	-0.081	105.62	0.000
		23	-0.010	0.021	105.66	0.000
		24	-0.041	-0.111	106.39	0.000
		25	0.030	-0.008	106.76	0.000
		26	-0.061	0.008	108.34	0.000
		27	0.039	0.048	109.01	0.000
		28	0.019	0.117	109.17	0.000
		29	0.031	0.018	109.59	0.000
		30	-0.031	-0.061	110.01	0.000
		31	0.041	-0.008	110.75	0.000
		32	-0.004	-0.141	110.76	0.000
		33	-0.037	-0.060	111.38	0.000
		34	-0.041	-0.008	112.11	0.000
		35	0.027	0.061	112.43	0.000
		36	-0.041	0.080	113.15	0.000

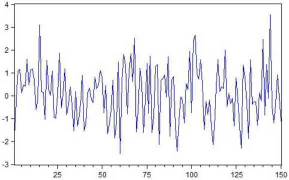
SMA$(1)_{12}$ process

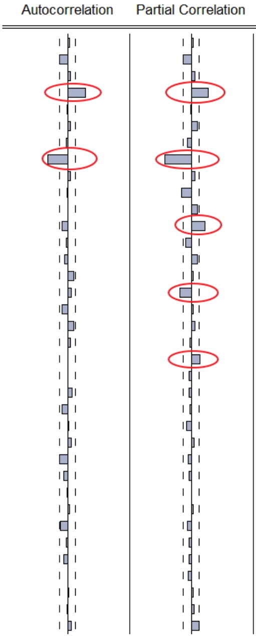




	AC	PAC	Q-Stat	Prob
1	-0.120	-0.120	2.6173	0.106
2	0.092	0.078	4.1608	0.125
3	0.018	0.038	4.2214	0.239
4	-0.017	-0.019	4.2774	0.370
5	0.079	0.072	5.4590	0.362
6	-0.050	-0.032	5.9373	0.430
7	0.102	0.083	7.8906	0.342
8	0.005	0.029	7.8954	0.444
9	-0.018	-0.027	7.9605	0.538
10	0.114	0.099	10.449	0.402
11	-0.104	-0.075	12.553	0.324
12	0.496	0.472	60.518	0.000
13	-0.083	0.011	61.863	0.000
14	0.121	0.080	64.733	0.000
15	-0.079	-0.133	65.971	0.000
16	-0.036	-0.035	66.227	0.000
17	0.130	0.085	69.623	0.000
18	-0.102	-0.056	71.733	0.000
19	0.096	0.015	73.619	0.000
20	0.024	0.011	73.733	0.000
21	0.037	0.091	74.009	0.000
22	0.080	-0.002	75.328	0.000
23	-0.064	0.028	76.175	0.000
24	0.077	-0.282	77.429	0.000
25	-0.054	-0.002	78.037	0.000
26	0.046	-0.073	78.481	0.000
27	-0.108	-0.024	80.980	0.000
28	-0.047	-0.030	81.451	0.000
29	0.140	0.060	85.725	0.000
30	-0.164	-0.092	91.613	0.000
31	0.061	-0.027	92.427	0.000
32	0.057	0.098	93.136	0.000
33	0.023	-0.057	93.259	0.000
34	0.034	0.021	93.514	0.000
35	-0.057	-0.074	94.260	0.000
36	0.091	0.250	96.150	0.000
37	-0.001	0.067	96.150	0.000

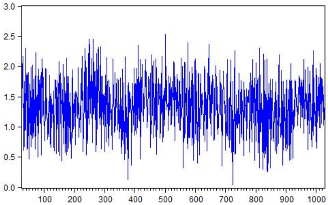
$SMA(2)_4$ process



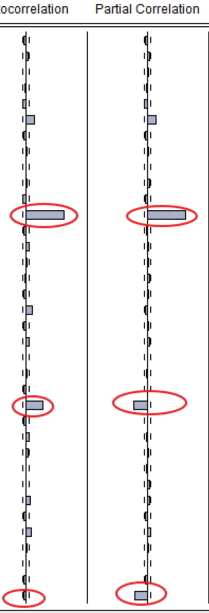


	AC	PAC	Q-Stat	Prob
1	0.026	0.026	0.1711	0.679
2	-0.125	-0.125	4.1130	0.128
3	0.051	0.059	4.7803	0.189
4	0.298	0.285	27.565	0.000
5	-0.009	-0.013	27.585	0.000
6	0.041	0.113	28.021	0.000
7	-0.020	-0.062	28.129	0.000
8	-0.329	-0.445	56.281	0.000
9	0.042	0.064	56.752	0.000
10	-0.008	-0.171	56.768	0.000
11	0.001	0.114	56.768	0.000
12	-0.091	0.223	58.952	0.000
13	-0.022	-0.084	59.082	0.000
14	-0.053	0.105	59.823	0.000
15	0.107	0.025	62.912	0.000
16	0.065	-0.196	64.046	0.000
17	-0.089	0.029	66.189	0.000
18	0.112	0.058	69.587	0.000
19	0.041	-0.019	70.049	0.000
20	0.002	0.153	70.050	0.000
21	0.001	-0.033	70.051	0.000
22	0.077	-0.039	71.677	0.000
23	-0.095	-0.024	74.171	0.000
24	0.020	-0.081	74.281	0.000
25	0.063	0.046	75.401	0.000
26	-0.125	-0.055	79.782	0.000
27	-0.064	-0.029	80.954	0.000
28	-0.004	0.016	80.960	0.000
29	0.040	0.026	81.405	0.000
30	-0.117	-0.073	85.319	0.000
31	-0.022	-0.041	85.453	0.000
32	-0.062	-0.030	86.556	0.000
33	0.000	-0.053	86.556	0.000
34	0.020	0.033	86.677	0.000
35	-0.001	0.044	86.677	0.000
36	0.073	0.133	88.245	0.000

SMA(2)$_{12}$ process



	AC	PAC	Q-Stat	Prob
1	-0.035	-0.035	1.2490	0.264
2	0.038	0.037	2.7121	0.258
3	0.007	0.010	2.7640	0.429
4	-0.019	-0.020	3.1346	0.536
5	-0.056	-0.058	6.3151	0.277
6	0.145	0.143	27.525	0.000
7	-0.044	-0.031	29.498	0.000
8	0.024	0.012	30.089	0.000
9	-0.001	-0.002	30.091	0.000
10	0.023	0.027	30.649	0.001
11	-0.053	-0.039	33.465	0.000
12	0.648	0.640	459.42	0.000
13	-0.039	-0.021	460.92	0.000
14	0.052	0.018	463.72	0.000
15	0.018	0.012	464.04	0.000
16	-0.014	0.031	464.23	0.000
17	-0.044	0.020	466.17	0.000
18	0.102	-0.042	476.87	0.000
19	-0.030	0.016	477.80	0.000
20	0.061	0.036	481.59	0.000
21	-0.000	0.008	481.59	0.000
22	0.008	-0.020	481.67	0.000
23	-0.049	0.006	484.09	0.000
24	0.284	-0.232	566.70	0.000
25	-0.043	-0.001	568.63	0.000
26	0.048	-0.009	570.99	0.000
27	0.025	0.013	571.65	0.000
28	-0.002	-0.013	571.65	0.000
29	-0.000	0.041	571.65	0.000
30	0.068	0.025	576.38	0.000
31	-0.041	-0.038	578.16	0.000
32	0.098	0.064	588.06	0.000
33	0.014	0.023	588.27	0.000
34	0.000	0.004	588.27	0.000
35	-0.042	-0.035	590.12	0.000
36	-0.041	-0.215	591.84	0.000



Seasonal Autoregressive process $\text{SAR}(P)_S$

- The Seasonal Autoregressive model $SAR(P)_S$ is defined by the following equation:

$$\boxed{X_t = \Phi_0 + \Phi_1 X_{t-S} + \Phi_2 X_{t-2S} + \ldots + \Phi_P X_{t-PS} + \varepsilon_t,\ \varepsilon_t \overset{w.n.}{\sim} (0, \sigma_\varepsilon^2)}$$

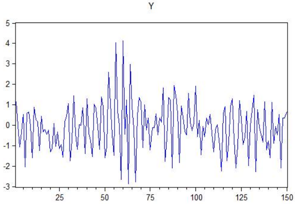
- Using the lag operator, L, this model can be written in a more compact form as:

$$\boxed{\Phi(L^S) X_t = \Phi_0 + \varepsilon_t}$$

  where $\Phi(L^S) = 1 - \Phi_1 L^S - \Phi_2 L^{2S} - \ldots - \Phi_P L^{PS}$.

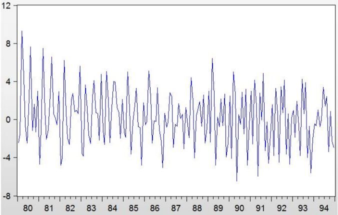
- X_t is **stationary** if the (inverse) roots of the AR polynomial, $\Phi(L^S)$, are outside (inside) the unit circle.

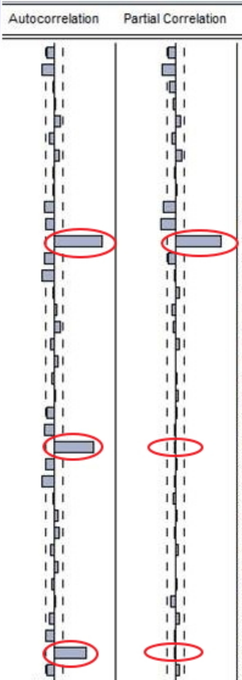
SAR$(1)_4$ process



Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.015	0.015	0.1146	0.735
		2	0.038	0.038	0.8457	0.655
		3	0.052	0.051	2.1920	0.534
		4	0.792	0.793	319.76	0.000
		5	-0.031	-0.062	320.24	0.000
		6	0.018	-0.056	320.40	0.000
		7	0.080	0.026	323.67	0.000
		8	0.639	0.045	532.24	0.000
		9	-0.091	-0.090	536.43	0.000
		10	0.002	-0.018	536.43	0.000
		11	0.107	0.040	542.28	0.000
		12	0.551	0.113	698.57	0.000
		13	-0.126	-0.015	706.79	0.000
		14	0.018	0.044	706.96	0.000
		15	0.109	-0.036	713.14	0.000
		16	0.467	0.005	826.44	0.000
		17	-0.133	0.032	835.61	0.000
		18	0.039	0.020	836.41	0.000
		19	0.121	0.018	844.01	0.000
		20	0.377	-0.031	918.30	0.000
		21	-0.125	0.034	926.54	0.000
		22	0.061	0.024	928.47	0.000
		23	0.120	-0.023	935.99	0.000
		24	0.297	-0.022	982.47	0.000
		25	-0.115	0.017	989.50	0.000
		26	0.090	0.050	993.77	0.000
		27	0.107	-0.013	999.87	0.000
		28	0.212	-0.072	1023.7	0.000
		29	-0.084	0.063	1027.5	0.000
		30	0.095	-0.026	1032.3	0.000
		31	0.101	0.021	1037.8	0.000
		32	0.145	-0.016	1049.0	0.000
		33	-0.065	-0.018	1051.3	0.000
		34	0.100	0.013	1056.7	0.000
		35	0.084	-0.006	1060.5	0.000
		36	0.080	-0.051	1064.0	0.000

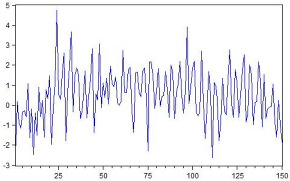
SAR$(1)_{12}$ process

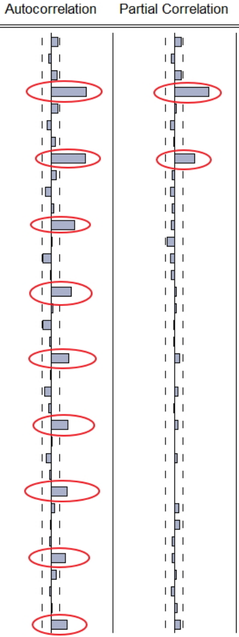




	AC	PAC	Q-Stat	Prob
1	-0.135	-0.135	3.3393	0.068
2	-0.203	-0.226	10.953	0.004
3	-0.031	-0.102	11.127	0.011
4	0.029	-0.043	11.284	0.024
5	0.101	0.077	13.204	0.022
6	-0.083	-0.060	14.509	0.024
7	0.078	0.103	15.656	0.028
8	-0.020	-0.010	15.735	0.046
9	-0.017	0.011	15.793	0.071
10	-0.179	-0.208	21.959	0.015
11	-0.152	-0.243	26.447	0.006
12	0.816	0.778	156.27	0.000
13	-0.159	-0.134	161.21	0.000
14	-0.218	-0.019	170.60	0.000
15	-0.020	0.053	170.68	0.000
16	0.046	-0.068	171.10	0.000
17	0.093	-0.044	172.82	0.000
18	-0.070	0.057	173.81	0.000
19	0.071	0.005	174.83	0.000
20	-0.038	0.030	175.12	0.000
21	0.013	0.049	175.16	0.000
22	-0.139	0.019	179.17	0.000
23	-0.157	-0.015	184.29	0.000
24	0.664	-0.028	276.75	0.000
25	-0.151	0.026	281.60	0.000
26	-0.218	-0.022	291.66	0.000
27	-0.023	-0.042	291.77	0.000
28	0.061	0.032	292.57	0.000
29	0.075	-0.020	293.81	0.000
30	-0.051	0.033	294.38	0.000
31	0.058	-0.004	295.14	0.000
32	-0.042	0.031	295.53	0.000
33	0.009	-0.086	295.54	0.000
34	-0.086	0.074	297.19	0.000
35	-0.152	-0.016	302.43	0.000
36	0.539	-0.020	368.64	0.000
37	-0.121	0.058	372.02	0.000

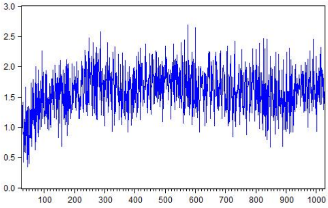
SAR$(2)_4$ process

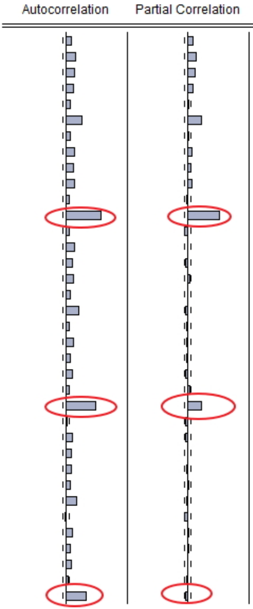




	AC	PAC	Q-Stat	Prob
1	0.116	0.116	2.7130	0.100
2	-0.032	-0.046	2.9237	0.232
3	0.109	0.120	5.3665	0.147
4	0.591	0.578	77.297	0.000
5	0.119	0.035	80.231	0.000
6	-0.059	-0.060	80.965	0.000
7	0.078	-0.008	82.239	0.000
8	0.579	0.350	152.74	0.000
9	0.093	-0.035	154.57	0.000
10	-0.083	-0.062	156.02	0.000
11	0.045	-0.039	156.46	0.000
12	0.399	-0.047	190.68	0.000
13	0.017	-0.123	190.74	0.000
14	-0.127	-0.059	194.26	0.000
15	-0.006	-0.045	194.26	0.000
16	0.351	0.032	221.26	0.000
17	0.028	0.034	221.44	0.000
18	-0.139	-0.010	225.75	0.000
19	-0.020	-0.001	225.84	0.000
20	0.302	0.084	246.33	0.000
21	-0.002	0.013	246.33	0.000
22	-0.108	0.073	248.96	0.000
23	-0.040	-0.008	249.32	0.000
24	0.293	0.073	269.06	0.000
25	0.022	0.003	269.17	0.000
26	-0.077	0.045	270.55	0.000
27	-0.016	0.006	270.61	0.000
28	0.267	0.006	287.34	0.000
29	0.070	0.074	288.50	0.000
30	-0.007	0.089	288.51	0.000
31	-0.022	-0.035	288.63	0.000
32	0.241	-0.029	302.59	0.000
33	0.093	0.030	304.67	0.000
34	-0.025	-0.052	304.81	0.000
35	0.022	0.048	304.93	0.000
36	0.276	0.109	323.71	0.000

$SAR(2)_{12}$ process





	AC	PAC	Q-Stat	Prob
1	0.090	0.090	8.0448	0.005
2	0.158	0.151	33.081	0.000
3	0.139	0.117	52.383	0.000
4	0.124	0.087	67.809	0.000
5	0.073	0.025	73.121	0.000
6	0.270	0.232	146.54	0.000
7	0.069	0.010	151.33	0.000
8	0.146	0.070	172.95	0.000
9	0.118	0.046	187.09	0.000
10	0.148	0.078	209.39	0.000
11	0.052	-0.016	212.14	0.000
12	0.585	0.534	558.71	0.000
13	0.050	-0.065	561.21	0.000
14	0.143	0.006	581.93	0.000
15	0.103	-0.034	592.69	0.000
16	0.122	0.031	607.85	0.000
17	0.070	0.007	612.81	0.000
18	0.207	-0.030	656.39	0.000
19	0.041	-0.021	658.12	0.000
20	0.117	-0.024	672.05	0.000
21	0.077	-0.012	678.14	0.000
22	0.105	-0.030	689.42	0.000
23	0.054	0.036	692.37	0.000
24	0.498	0.233	947.49	0.000
25	0.021	-0.048	947.95	0.000
26	0.106	-0.039	959.53	0.000
27	0.088	0.003	967.57	0.000
28	0.085	-0.013	974.99	0.000
29	0.066	0.014	979.47	0.000
30	0.179	-0.013	1012.6	0.000
31	-0.009	-0.061	1012.7	0.000
32	0.112	0.018	1025.7	0.000
33	0.071	0.015	1030.9	0.000
34	0.088	0.006	1038.9	0.000
35	0.030	-0.018	1039.9	0.000
36	0.342	-0.028	1161.4	0.000

Seasonal Autoregressive and Moving Average process SARMA$(P, Q)_S$

- The Seasonal Autoregressive and Moving Average process *SARMA*$(P, Q)_S$ is defined by the following equation:

$$X_t = \Phi_0 + \Phi_1 X_{t-S} + \ldots + \Phi_P X_{t-PS} + \varepsilon_t - \Theta_1 \varepsilon_{t-S} - \ldots - \Theta_Q \varepsilon_{t-QS}$$

  where $\varepsilon_t \overset{w.n.}{\sim} (0, \sigma_\varepsilon^2)$.

- With the lag operator, L, this model can be rewritten in a more compact form:

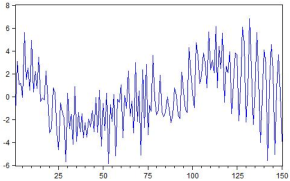
$$\Phi\left(L^S\right) X_t = \Phi_0 + \Theta\left(L^S\right) \varepsilon_t$$

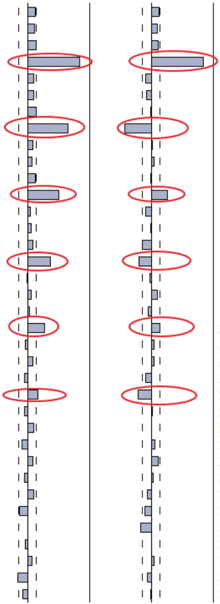
  where:

$$\Phi(L^S) = 1 - \Phi_1 L^S - \Phi_2 L^{2S} - \ldots - \Phi_P L^{PS}$$
$$\Theta(L^S) = 1 - \Theta_1 L^S - \Theta_2 L^{2S} - \ldots - \Theta_Q L^{QS}$$

- X_t is **stationary** if the (inverse) roots of the AR polynomial, $\Phi(L^S)$, are outside (inside) the unit circle.
- X_t is **invertible** if the (inverse) roots of the MA polynomial, $\Theta(L^S)$, are outside (inside) the unit circle.

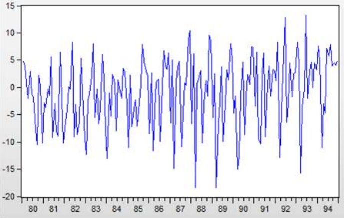
SARMA$(1,1)_4$ process

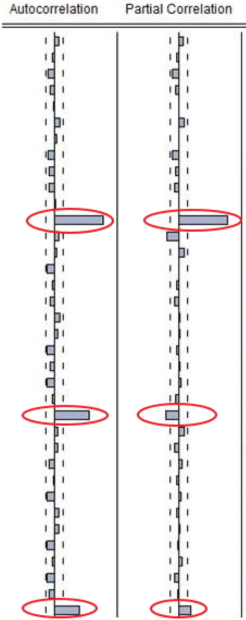




	AC	PAC	Q-Stat	Prob
1	0.138	0.138	3.8585	0.049
2	0.120	0.103	6.8194	0.033
3	0.143	0.117	11.004	0.012
4	0.882	0.877	171.40	0.000
5	0.114	-0.082	174.11	0.000
6	0.101	-0.075	176.23	0.000
7	0.140	-0.008	180.34	0.000
8	0.680	-0.439	277.58	0.000
9	0.084	0.024	279.07	0.000
10	0.082	0.056	280.50	0.000
11	0.132	-0.006	284.20	0.000
12	0.522	0.264	342.73	0.000
13	0.046	-0.096	343.19	0.000
14	0.069	-0.016	344.22	0.000
15	0.094	-0.141	346.13	0.000
16	0.390	-0.212	379.52	0.000
17	-0.004	-0.018	379.52	0.000
18	0.069	0.102	380.59	0.000
19	0.027	-0.050	380.75	0.000
20	0.280	0.144	398.39	0.000
21	-0.041	0.046	398.77	0.000
22	0.090	0.042	400.58	0.000
23	-0.044	-0.091	401.04	0.000
24	0.178	-0.216	408.30	0.000
25	-0.052	0.024	408.94	0.000
26	0.109	0.013	411.68	0.000
27	-0.095	0.070	413.81	0.000
28	0.083	0.117	415.41	0.000
29	-0.045	0.037	415.90	0.000
30	0.103	-0.070	418.40	0.000
31	-0.127	-0.102	422.25	0.000
32	0.002	-0.173	422.26	0.000
33	-0.025	0.015	422.41	0.000
34	0.080	0.043	423.97	0.000
35	-0.169	-0.066	431.00	0.000
36	-0.075	-0.013	432.40	0.000

SARMA$(1,1)_{12}$ process





	AC	PAC	Q-Stat	Prob
1	0.080	0.080	1.1856	0.276
2	-0.028	-0.035	1.3279	0.515
3	-0.113	-0.108	3.6739	0.299
4	-0.069	-0.053	4.5580	0.336
5	-0.009	-0.005	4.5717	0.470
6	0.095	0.083	6.2681	0.394
7	0.028	0.001	6.4119	0.493
8	-0.105	-0.112	8.5220	0.384
9	-0.093	-0.061	10.165	0.337
10	-0.092	-0.075	11.809	0.298
11	0.010	0.001	11.827	0.377
12	0.810	0.812	139.80	0.000
13	0.066	-0.197	140.65	0.000
14	0.031	0.095	140.84	0.000
15	-0.115	-0.003	143.45	0.000
16	-0.043	-0.031	143.82	0.000
17	-0.073	-0.082	144.88	0.000
18	0.093	0.021	146.65	0.000
19	0.056	0.012	147.29	0.000
20	-0.122	-0.032	150.35	0.000
21	-0.067	0.022	151.27	0.000
22	-0.122	0.038	154.33	0.000
23	-0.045	-0.062	154.76	0.000
24	0.571	-0.216	223.27	0.000
25	0.053	0.093	223.87	0.000
26	0.054	-0.077	224.49	0.000
27	-0.097	0.043	226.51	0.000
28	-0.024	-0.034	226.64	0.000
29	-0.122	0.031	229.86	0.000
30	0.081	-0.047	231.30	0.000
31	0.071	0.022	232.42	0.000
32	-0.117	-0.010	235.44	0.000
33	-0.028	0.044	235.62	0.000
34	-0.136	-0.077	239.77	0.000
35	-0.090	-0.023	241.59	0.000
36	0.424	0.200	282.45	0.000

Seasonal Autoregressive, Integrated and Moving Average process SARIMA$(P, D, Q)_S$

- The seasonal autoregressive, integrated and moving Average process $SARIMA(P, D, Q)_S$ is defined by the following equation:

$$\boxed{\Delta_S^D X_t = \Phi_0 + \Phi_1 \Delta_S^D X_{t-S} + \ldots + \Phi_p \Delta_{12}^S X_{t-PS} + \varepsilon_t - \Theta_1 \varepsilon_{t-S} - \ldots - \Theta_Q \varepsilon_{t-QS}}$$

where $\Delta_S^D = \left(1 - L^S\right)^D$ and $\varepsilon_t \overset{w.n.}{\sim} \left(0, \sigma_\varepsilon^2\right)$.

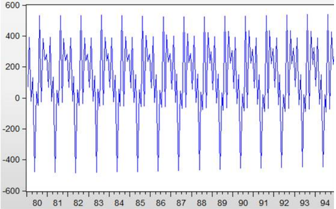
- Using the lag operator, L, this model can be rewriten in a more compact form as:

$$\boxed{\Phi\left(L^S\right) \Delta_S^D X_t = \Phi_0 + \Theta\left(L^S\right) \varepsilon_t}$$

where:

$$\Phi(L^S) = 1 - \Phi_1 L^S - \Phi_2 L^{2S} - \ldots - \Phi_P L^{PS}$$
$$\Theta(L^S) = 1 - \Theta_1 L^S - \Theta_2 L^{2S} - \ldots - \Theta_Q L^{QS}$$

SARIMA$(1,1,0)_{12}$ process



	AC	PAC	Q-Stat	Prob
1	-0.210	-0.210	8.0647	0.005
2	0.495	0.471	53.109	0.000
3	-0.531	-0.508	105.36	0.000
4	0.309	0.118	123.14	0.000
5	-0.646	-0.415	201.20	0.000
6	0.182	-0.204	207.47	0.000
7	-0.637	-0.338	284.34	0.000
8	0.300	-0.203	301.46	0.000
9	-0.499	-0.503	349.08	0.000
10	0.467	-0.418	391.06	0.000
11	-0.199	-0.439	398.77	0.000
12	0.934	0.677	568.73	0.000
13	-0.196	-0.011	576.28	0.000
14	0.464	-0.485	618.73	0.000
15	-0.496	0.316	667.67	0.000
16	0.291	-0.137	684.53	0.000
17	-0.601	-0.060	757.15	0.000
18	0.171	0.201	763.03	0.000
19	-0.594	-0.078	834.79	0.000
20	0.279	0.029	850.68	0.000
21	-0.463	-0.090	894.91	0.000
22	0.432	0.048	933.55	0.000
23	-0.187	0.064	940.86	0.000
24	0.867	-0.135	1098.7	0.000
25	-0.182	0.112	1105.7	0.000
26	0.433	-0.030	1145.5	0.000
27	-0.461	-0.051	1191.1	0.000
28	0.271	0.046	1207.0	0.000
29	-0.557	-0.035	1274.2	0.000
30	0.159	0.050	1279.7	0.000
31	-0.550	-0.029	1346.3	0.000
32	0.257	-0.001	1360.9	0.000
33	-0.428	0.062	1401.7	0.000
34	0.396	-0.066	1437.0	0.000
35	-0.174	0.031	1443.8	0.000
36	0.800	-0.024	1589.4	0.000
37	-0.167	-0.010	1595.8	0.000

General multiplicative model $\mathbf{SARIMA(p, d, q) \times (P, D, Q)_S}$ I

- The general multiplicative model $\mathbf{SARIMA(p, d, q) \times (P, D, Q)_S}$ is defined by the following equation using the Lag operator:

$$\boxed{\phi(L)\,\Phi(L^S)\Delta^d\Delta_S^D X_t = \Phi_0 + \theta(L)\,\Theta(L^S)\varepsilon_t,\ \varepsilon_t \overset{w.n.}{\sim} (0, \sigma_\varepsilon^2)}$$

where:

$$\phi(L) = 1 - \phi_1 L - \phi_2 L^2 - \ldots - \phi_p L^p$$
$$\theta(L) = 1 - \theta_1 L - \theta_2 L^2 - \ldots - \theta_q L^q$$
$$\Phi(L^S) = 1 - \Phi_1 L^S - \Phi_2 L^{2S} - \ldots - \Phi_P L^{PS}$$
$$\Theta(L^S) = 1 - \Theta_1 L^S - \Theta_2 L^{2S} - \ldots - \Theta_Q L^{QS}$$

and the (inverse) roots of the polynomials $\phi(L)$ and $\Phi(L^S)$ are outside (inside) the unit circle or, in other words, $\Delta^d\Delta_S^D X_t$ is stationary.

General multiplicative model $\text{SARIMA}(p, d, q) \times (P, D, Q)_S$ II

- To ease the understanding of the general multiplicative specification the following examples (without constant) might be useful:
    1. **$\text{SARIMA}(1, 0, 0) \times (1, 0, 0)_{12}$**

$$\boxed{(1 - \phi L)\left(1 - \Phi L^{12}\right) X_t = \varepsilon_t}$$

$$\Leftrightarrow \boxed{X_t = \phi X_{t-1} + \Phi X_{t-12} - \phi \Phi X_{t-13} + \varepsilon_t}$$

    2. **$\text{SARIMA}(0, 0, 1) \times (0, 0, 1)_{12}$**

$$\boxed{X_t = (1 - \theta L)\left(1 - \Theta L^{12}\right) \varepsilon_t}$$

$$\Leftrightarrow \boxed{X_t = \varepsilon_t - \theta \varepsilon_{t-1} - \Theta \varepsilon_{t-12} + \theta \Theta \varepsilon_{t-13}}$$

General multiplicative model $\mathbf{SARIMA(p,d,q)\times(P,D,Q)_S}$ III

3. **SARIMA**$(\mathbf{1,0,0})\times(\mathbf{0,0,1})_{\mathbf{12}}$

$$\boxed{(1-\phi L)\,X_t=\left(1-\Theta L^{12}\right)\varepsilon_t}$$

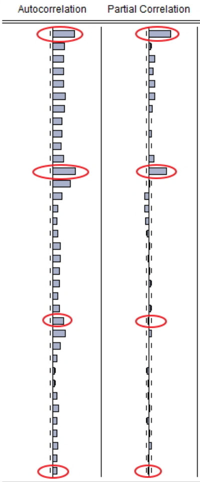
$$\Leftrightarrow \boxed{X_t=\phi X_{t-1}+\varepsilon_t-\Theta\varepsilon_{t-12}}$$

4. **SARIMA**$(\mathbf{0,0,1})\times(\mathbf{1,0,0})_{\mathbf{12}}$

$$\boxed{\left(1-\Phi L^{12}\right)X_t=(1-\theta L)\,\varepsilon_t}$$

$$\Leftrightarrow \boxed{X_t=\Phi X_{t-12}+\varepsilon_t-\theta\varepsilon_{t-1}}$$

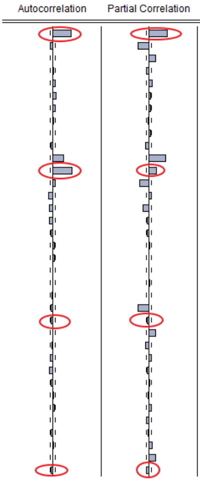
SACF/PACF of a SARIMA$(1,0,0) \times (1,0,0)_{12}$



	AC	PAC	Q-Stat	Prob
1	0.466	0.466	218.26	0.000
2	0.248	0.039	280.07	0.000
3	0.227	0.125	331.96	0.000
4	0.222	0.087	381.60	0.000
5	0.243	0.115	440.91	0.000
6	0.271	0.121	514.90	0.000
7	0.253	0.069	579.52	0.000
8	0.193	0.007	617.27	0.000
9	0.197	0.061	656.70	0.000
10	0.173	0.001	686.83	0.000
11	0.231	0.113	740.89	0.000
12	0.483	0.383	977.85	0.000
13	0.378	0.009	1123.1	0.000
14	0.192	-0.087	1160.4	0.000
15	0.114	-0.083	1173.6	0.000
16	0.093	-0.075	1182.5	0.000
17	0.111	-0.035	1195.1	0.000
18	0.169	0.016	1224.3	0.000
19	0.168	-0.015	1253.1	0.000
20	0.134	0.014	1271.4	0.000
21	0.141	0.031	1291.7	0.000
22	0.111	-0.007	1304.2	0.000
23	0.132	0.035	1322.1	0.000
24	0.223	-0.023	1373.2	0.000
25	0.269	0.057	1447.5	0.000
26	0.170	-0.003	1477.1	0.000
27	0.091	0.002	1485.7	0.000
28	0.034	-0.038	1486.9	0.000
29	0.036	-0.011	1488.2	0.000
30	0.084	-0.013	1495.5	0.000
31	0.121	0.024	1510.7	0.000
32	0.084	-0.037	1518.0	0.000
33	0.090	0.014	1526.3	0.000
34	0.107	0.054	1538.3	0.000
35	0.087	-0.011	1546.1	0.000
36	0.097	-0.019	1556.0	0.000

The SACF/SPACF is compatible with an AR model with both the seasonal and nonseasonal part: we find the most relevant spikes in the SPACF at lags 1 (nonseasonal) and 12 (seasonal) and it seems to cut off after the lag 12. The SACF decays but it seems to be infinite in extent with no explicit cut off.

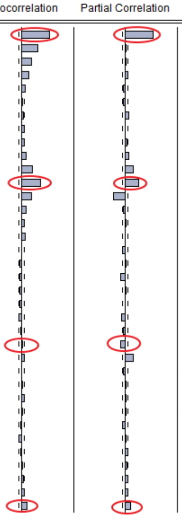
SACF/PACF of a SARIMA$(0,0,1) \times (0,0,1)_{12}$



	AC	PAC	Q-Stat	Prob
1	0.387	0.387	149.90	0.000
2	-0.050	-0.235	152.42	0.000
3	-0.006	0.135	152.46	0.000
4	0.022	-0.053	152.94	0.000
5	0.044	0.076	154.93	0.000
6	0.072	0.029	160.08	0.000
7	0.064	0.037	164.26	0.000
8	0.012	-0.022	164.42	0.000
9	0.005	0.026	164.45	0.000
10	-0.034	-0.067	165.60	0.000
11	0.229	0.353	218.93	0.000
12	0.405	0.170	385.16	0.000
13	0.049	-0.192	387.61	0.000
14	-0.093	0.062	396.32	0.000
15	-0.078	-0.138	402.45	0.000
16	-0.081	-0.032	409.17	0.000
17	-0.045	-0.038	411.22	0.000
18	0.039	0.024	412.78	0.000
19	0.036	-0.005	414.13	0.000
20	-0.004	0.005	414.15	0.000
21	0.001	0.022	414.15	0.000
22	-0.016	0.001	414.41	0.000
23	-0.027	-0.241	415.19	0.000
24	-0.042	-0.043	416.96	0.000
25	-0.004	0.147	416.98	0.000
26	-0.014	-0.071	417.19	0.000
27	-0.058	0.065	420.61	0.000
28	-0.082	-0.032	427.45	0.000
29	-0.037	0.017	428.86	0.000
30	0.019	-0.024	429.21	0.000
31	0.038	0.041	430.74	0.000
32	-0.001	-0.051	430.74	0.000
33	-0.029	-0.024	431.64	0.000
34	0.012	0.079	431.81	0.000
35	0.001	0.140	431.81	0.000
36	-0.026	-0.054	432.49	0.000

The SACF/SPACF is compatible with an MA model with both the seasonal and nonseasonal part: we find the most relevant spikes in the SACF at lags 1 (nonseasonal) and 12 (seasonal) and it seems to cut off after the lag 12. The SPACF displays relatively high spikes even for high lags.

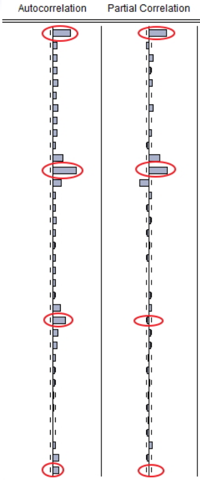
SACF/PACF of a SARIMA$(1,0,0) \times (0,0,1)_{12}$



	AC	PAC	Q-Stat	Prob
1	0.553	0.553	306.74	0.000
2	0.316	0.014	406.74	0.000
3	0.196	0.023	445.23	0.000
4	0.150	0.044	467.98	0.000
5	0.075	-0.045	473.61	0.000
6	0.008	-0.042	473.68	0.000
7	0.041	0.076	475.33	0.000
8	0.042	-0.002	477.14	0.000
9	0.055	0.033	480.15	0.000
10	0.095	0.076	489.25	0.000
11	0.209	0.168	533.58	0.000
12	0.385	0.277	683.64	0.000
13	0.192	-0.245	721.11	0.000
14	0.089	-0.031	729.14	0.000
15	0.061	0.019	732.97	0.000
16	0.067	0.007	737.51	0.000
17	-0.003	-0.049	737.52	0.000
18	-0.030	0.021	738.41	0.000
19	-0.049	-0.084	740.82	0.000
20	-0.045	0.013	742.92	0.000
21	-0.039	-0.006	744.49	0.000
22	-0.053	-0.070	747.32	0.000
23	-0.010	-0.031	747.42	0.000
24	0.012	-0.098	747.56	0.000
25	0.042	0.169	749.37	0.000
26	0.004	-0.039	749.39	0.000
27	0.019	0.009	749.75	0.000
28	0.042	0.020	751.55	0.000
29	0.014	0.012	751.75	0.000
30	-0.020	-0.050	752.15	0.000
31	-0.052	0.015	754.94	0.000
32	-0.010	0.046	755.04	0.000
33	0.017	0.041	755.32	0.000
34	0.026	0.059	756.03	0.000
35	0.060	0.058	759.75	0.000
36	0.108	0.112	771.81	0.000

Both the SACF and SPACF are apparently infinite in extent with no explicit cutoff. Thus, we may suspect of a model with both the MA and AR parts. The explicit form should be done by trying more parsimonious models and the best model should be selected according to the Box-Jenkins principles.

SACF/SPACF of SARIMA$(0,0,1) \times (1,0,0)_{12}$



	AC	PAC	Q-Stat	Prob
1	0.380	0.380	145.16	0.000
2	0.092	-0.062	153.63	0.000
3	0.090	0.090	161.86	0.000
4	0.087	0.029	169.40	0.000
5	0.114	0.081	182.37	0.000
6	0.071	-0.006	187.40	0.000
7	0.110	0.096	199.70	0.000
8	0.085	-0.002	207.04	0.000
9	0.091	0.062	215.39	0.000
10	0.049	-0.028	217.82	0.000
11	0.209	0.231	262.22	0.000
12	0.497	0.397	512.69	0.000
13	0.172	-0.195	542.86	0.000
14	0.065	0.064	547.10	0.000
15	0.057	-0.035	550.38	0.000
16	0.073	0.025	555.83	0.000
17	0.064	-0.047	560.01	0.000
18	0.039	0.011	561.56	0.000
19	0.064	-0.025	565.78	0.000
20	0.044	-0.008	567.76	0.000
21	0.058	-0.002	571.25	0.000
22	0.039	0.031	572.84	0.000
23	0.167	0.057	601.54	0.000
24	0.264	-0.026	672.95	0.000
25	0.103	0.041	683.84	0.000
26	0.086	0.038	691.48	0.000
27	0.056	-0.010	694.71	0.000
28	0.037	-0.041	696.10	0.000
29	0.029	0.016	696.94	0.000
30	0.023	-0.017	697.49	0.000
31	0.019	-0.030	697.86	0.000
32	0.000	-0.024	697.86	0.000
33	0.000	-0.030	697.86	0.000
34	0.053	0.072	700.76	0.000
35	0.126	-0.029	717.29	0.000
36	0.125	0.000	733.55	0.000

According to the SPACF/SACF it seems that we have a pure AR model such as *SARIMA*$(1,0,0) \times (1,0,0)_{12}$. However, the time series was simulated according to a *SARIMA*$(0,0,1) \times (1,0,0)_{12}$. This case illustrates the difficulties of tentative identification and how important is to try different models and select the best fit according to Box-Jenkins methodology.

EXERCISE

- **In this exercise we use monthly data about the total number of international airline passengers (in thousands of passengers) in the U.S. during the period 01/1949-12/1960, AIRLINE. Wf1. Answer the following questions:**

EXERCISE I

(a) **Sketch the plot of the series. What do you conclude regarding stationarity and seasonality? Do you think it is necessary to apply any transformation?**

(b) **Sketch the plot and the correlogram of the transformed series. What do you conclude?**

(c) **Would you apply another transformation to the series? Why? If your answer is affirmative, comment the results obtained for the new transformed series.**

(d) **Sketch the plot and the correlogram of the series $\Delta^2 \log(X_t)$. What do you conclude?**

EXERCISE II

(e) **Sketch the correlogram of the series $\Delta_{12}\Delta\log(X_t)$. What do you conclude?**

(f) **Remove the seasonal component from the series $\Delta log(X_{t)}$ using TramoSeats and Census X12. Sketch the plot and the correlogram of the series of the seasonally adjusted series. What do you conclude?**

(g) **Apply the Box-Jenkins methodology to select the model(s) of the class SARIMA$(p, d, q) \times (P, D, Q)_S$ that best fit the data. Recall that the multiplicative model SARIMA$(p, d, q) \times (P, D, Q)_S$ is defined by the equation:**

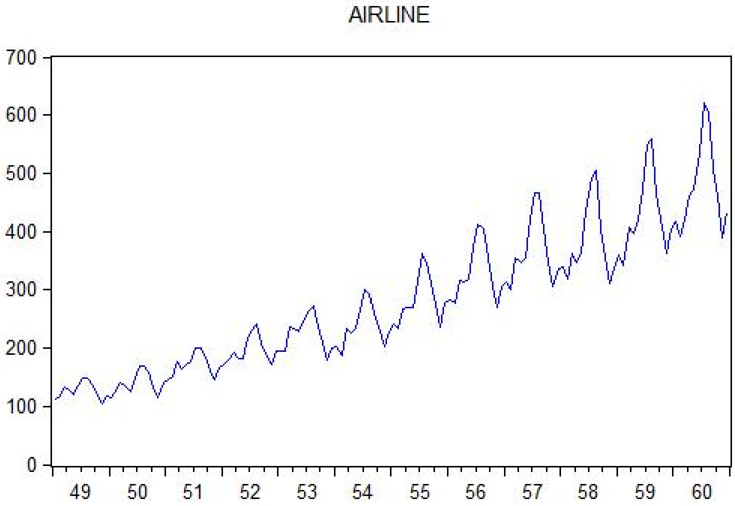
$$\phi(L)\,\Phi(L^S)\Delta^d\Delta_S^D X_t = \theta(L)\,\Theta(L^S)\varepsilon_t,\ \varepsilon_t \overset{w.n.}{\sim} (0, \sigma_\varepsilon^2)$$

EXERCISE III

Examine carefully the residual of the proposed model(s).

(h) **Estimate the model without the last 12 observations and make dynamic and static forecasts for 1960:01 until 1960:12. Compare your forecasts with the realized values. Draw a time series plot with the realized values, point and interval forecasts.**

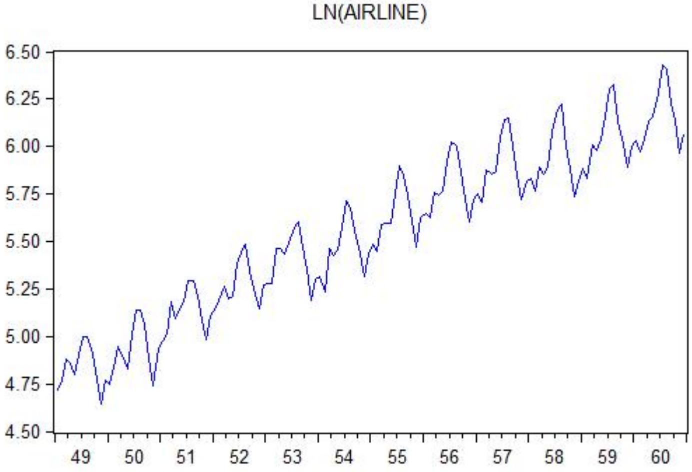
(a) **Sketch the plot of the series. What do you conclude regarding stationarity and seasonality? Do you think it is necessary to apply any transformation?**



Transformation: $\log(X_t)$ I

- From the plot it is clear that the time series possesses a trend and a seasonal pattern with a big spike occurring during Summer and a smaller one during the Spring Break.
- We also see that the variance is not constant, in particular, the series is more volatile in the second half of the sample. We try to stabilize the variance using the log transformation.

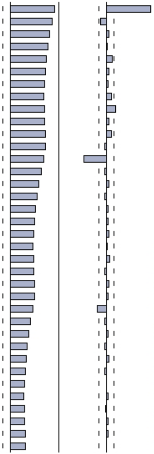
(b) **Sketch the plot and the correlogram of the transformed series. What do you conclude?**



The transformed series continues to display a trending and seasonal pattern.

SACF/SPACF of $\log(X_t)$

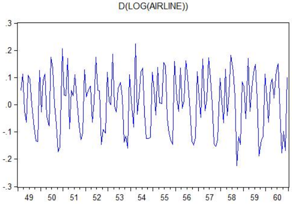
Sample: 1949M01 1960M12
Included observations: 144



Lag	AC	PAC	Q-Stat	Prob
1	0.954	0.954	133.72	0.000
2	0.899	-0.118	253.36	0.000
3	0.851	0.054	361.29	0.000
4	0.808	0.024	459.44	0.000
5	0.779	0.116	551.20	0.000
6	0.756	0.044	638.37	0.000
7	0.738	0.038	721.86	0.000
8	0.727	0.100	803.60	0.000
9	0.734	0.204	887.42	0.000
10	0.744	0.064	974.33	0.000
11	0.758	0.106	1065.2	0.000
12	0.762	-0.042	1157.6	0.000
13	0.717	-0.485	1240.0	0.000
14	0.663	-0.034	1311.1	0.000
15	0.618	0.042	1373.4	0.000
16	0.576	-0.044	1428.0	0.000
17	0.544	0.028	1476.9	0.000
18	0.519	0.037	1521.9	0.000
19	0.501	0.042	1564.1	0.000
20	0.490	0.014	1604.9	0.000
21	0.498	0.073	1647.3	0.000
22	0.506	-0.033	1691.5	0.000
23	0.517	0.061	1737.9	0.000
24	0.520	0.031	1785.3	0.000
25	0.484	-0.194	1826.6	0.000
26	0.437	-0.035	1860.7	0.000
27	0.400	0.036	1889.5	0.000
28	0.364	-0.035	1913.5	0.000
29	0.337	0.044	1934.3	0.000
30	0.315	-0.045	1952.6	0.000
31	0.297	-0.003	1969.0	0.000
32	0.289	0.034	1984.6	0.000
33	0.295	-0.020	2001.1	0.000
34	0.305	0.028	2018.8	0.000
35	0.315	0.029	2038.0	0.000
36	0.319	-0.004	2057.8	0.000

The SACF decays very slowly confirming the nonstationarity of the series. The presence of the nonseasonal unit root makes it impossible to obtain any information from the SACF regarding the seasonal pattern.

(c) **Would you apply another transformation to the series? Why? If your answer is affirmative, comment the results obtained for the new transformed series.**



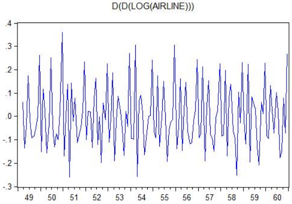
Sample: 1949M01 1960M12
Included observations: 143

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.200	0.200	5.8263	0.016
		2	-0.120	-0.167	7.9476	0.019
		3	-0.151	-0.096	11.314	0.010
		4	-0.322	-0.311	26.788	0.000
		5	-0.084	0.008	27.848	0.000
		6	0.026	-0.075	27.949	0.000
		7	-0.111	-0.210	29.826	0.000
		8	-0.337	-0.495	47.240	0.000
		9	-0.116	-0.192	49.308	0.000
		10	-0.109	-0.532	51.169	0.000
		11	0.206	-0.302	57.825	0.000
		12	0.841	0.586	169.89	0.000
		13	0.215	0.026	177.27	0.000
		14	-0.140	-0.181	180.40	0.000
		15	-0.116	0.120	182.58	0.000
		16	-0.279	0.000	195.28	0.000
		17	-0.052	0.025	195.72	0.000
		18	0.012	-0.125	195.75	0.000
		19	-0.114	0.087	197.94	0.000
		20	-0.337	-0.054	217.10	0.000
		21	-0.107	-0.062	219.06	0.000
		22	-0.075	-0.025	220.03	0.000
		23	0.199	0.033	226.90	0.000
		24	0.737	-0.010	321.53	0.000
		25	0.197	-0.048	328.37	0.000
		26	-0.124	0.018	331.09	0.000
		27	-0.103	0.028	332.97	0.000
		28	-0.211	0.016	341.00	0.000
		29	-0.065	-0.093	341.77	0.000
		30	0.016	0.008	341.82	0.000
		31	-0.115	0.071	344.28	0.000
		32	-0.289	0.110	359.91	0.000
		33	-0.127	-0.093	362.95	0.000
		34	-0.041	0.063	363.26	0.000
		35	0.147	-0.092	367.44	0.000
		36	0.657	0.058	451.19	0.000

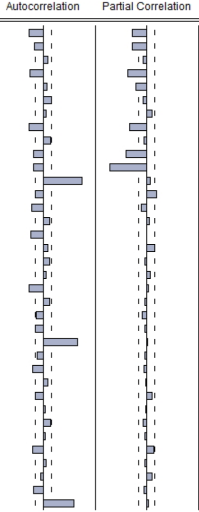
Transformation: $\Delta \log(X_t)$

- We take first differences to eliminate the non seasonal unit root from the $log(airline)$ series.
- The SACF of $\Delta log(airline)$ produces a very clear seasonal autocorrelation pattern with very large positive autocorrelations at the seasonal frequencies (lag 12, 24, 36,...) with lower but still relevant autocorrelations at the "neighbour" lags.
- Moreover we observe a slow decline of the seasonal autocorrelations. This implies that the first difference was not sufficient to stationarize the series.

(d) **Sketch the plot and the correlogram of the series $\Delta^2 \log(X_t)$. What do you conclude?**



Sample: 1949M01 1960M12
Included observations: 142



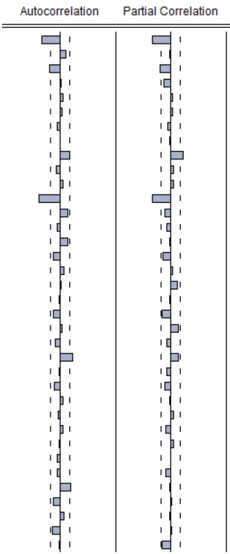
	AC	PAC	Q-Stat	Prob
1	-0.293	-0.293	12.416	0.000
2	-0.182	-0.292	17.238	0.000
3	0.092	-0.075	18.477	0.000
4	-0.265	-0.371	28.893	0.000
5	0.074	-0.206	29.720	0.000
6	0.157	-0.071	33.429	0.000
7	0.061	0.106	33.991	0.000
8	-0.282	-0.339	46.141	0.000
9	0.136	-0.063	48.973	0.000
10	-0.193	-0.425	54.756	0.000
11	-0.202	-0.744	61.119	0.000
12	0.787	0.069	158.47	0.000
13	-0.162	0.203	162.66	0.000
14	-0.234	-0.113	171.42	0.000
15	0.120	0.048	173.73	0.000
16	-0.256	0.000	184.40	0.000
17	0.098	0.160	185.99	0.000
18	0.123	-0.043	188.49	0.000
19	0.064	0.082	189.16	0.000
20	-0.279	0.035	202.23	0.000
21	0.117	-0.026	204.53	0.000
22	-0.150	-0.090	208.36	0.000
23	-0.166	-0.029	213.11	0.000
24	0.675	0.014	292.09	0.000
25	-0.131	-0.042	295.09	0.000
26	-0.211	-0.051	302.97	0.000
27	0.083	-0.021	304.19	0.000
28	-0.168	0.106	309.26	0.000
29	0.036	-0.012	309.49	0.000
30	0.133	-0.068	312.71	0.000
31	0.033	-0.045	312.91	0.000
32	-0.207	0.139	320.89	0.000
33	0.041	-0.039	321.21	0.000
34	-0.061	0.108	321.91	0.000
35	-0.203	-0.063	329.79	0.000
36	0.608	0.027	401.09	0.000

Transformation: $\Delta^2 \log(X_t)$?

- We try to stationarize the time series by taking second differences to the series. large positive autocorrelations at the seasonal frequencies (lag 12, 24, 36,. . .) with lower but still relevant autocorrelations at the "neighbour" lags.
- This filter was clearly unsuccessful as we continue to have large positive seasonal autocorrelations that decay very slowly. Any alternative?
- The slow decline of the SACF at the seasonal frequencies indicates seasonal nonstationarity in the data: $s = 12$ in this case since we are using monthly data.

(e) Sketch the correlogram of the series $\Delta_{12}\Delta\log(X_t)$. What do you conclude?

Sample: 1949M01 1960M12
Included observations: 131

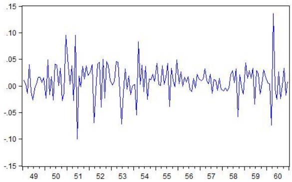


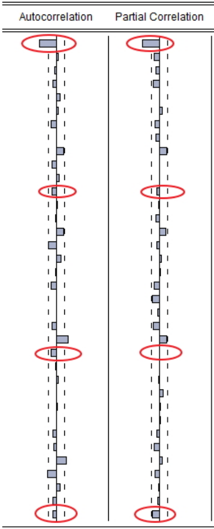
	AC	PAC	Q-Stat	Prob
1	-0.341	-0.341	15.596	0.000
2	0.105	-0.013	17.086	0.000
3	-0.202	-0.193	22.648	0.000
4	0.021	-0.125	22.710	0.000
5	0.056	0.033	23.139	0.000
6	0.031	0.035	23.271	0.001
7	-0.056	-0.060	23.705	0.001
8	-0.001	-0.020	23.705	0.003
9	0.176	0.226	28.147	0.001
10	-0.076	0.043	28.987	0.001
11	0.064	0.047	29.589	0.002
12	-0.387	-0.339	51.473	0.000
13	0.152	-0.109	54.866	0.000
14	-0.058	-0.077	55.361	0.000
15	0.150	-0.022	58.720	0.000
16	-0.139	-0.140	61.645	0.000
17	0.070	0.026	62.404	0.000
18	0.016	0.115	62.442	0.000
19	-0.011	-0.013	62.460	0.000
20	-0.117	-0.167	64.598	0.000
21	0.039	0.132	64.834	0.000
22	-0.091	-0.072	66.168	0.000
23	0.223	0.143	74.210	0.000
24	-0.018	-0.067	74.265	0.000
25	-0.100	-0.103	75.918	0.000
26	0.049	-0.010	76.310	0.000
27	-0.030	0.044	76.463	0.000
28	0.047	-0.090	76.839	0.000
29	-0.018	0.047	76.894	0.000
30	-0.051	-0.005	77.344	0.000
31	-0.054	-0.096	77.848	0.000
32	0.196	-0.015	84.590	0.000
33	-0.122	0.012	87.254	0.000
34	0.078	-0.019	88.340	0.000
35	-0.152	0.023	92.558	0.000
36	-0.010	-0.165	92.577	0.000

Transformation: $\Delta_{12}\Delta\mathbf{log}(\mathbf{X_t})$

- Given what we exposed in the answer to the last question we apply seasonal differencing, Δ_{12}.
- Now we are able to analyse the SACF/PACF of the transformed series, $\Delta_{12}\Delta log(X_t)$, and choose the most adequate SARIMA model.
- It is important to realize that the pattern of the SACF/SPACF of a seasonal series such as is much harder to interpret than a nonseasonal series.

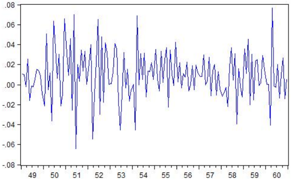
(f) **Remove the seasonal component from the series $\Delta log(X_{t)}$ using TramoSeats and Census X12. Sketch the plot and the correlogram of the series of the seasonally adjusted series. What do you conclude?**

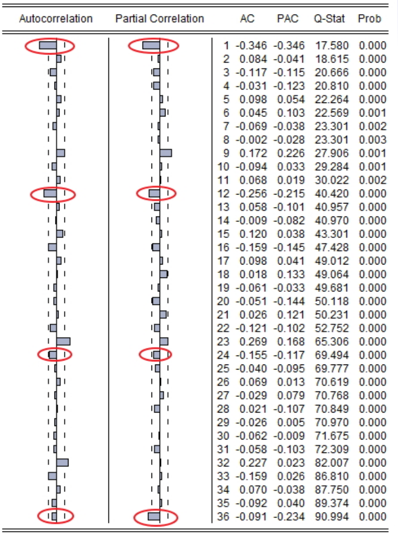




	AC	PAC	Q-Stat	Prob
1	-0.351	-0.351	18.141	0.000
2	0.030	-0.107	18.274	0.000
3	-0.044	-0.080	18.556	0.000
4	-0.067	-0.125	19.221	0.001
5	0.073	-0.001	20.023	0.001
6	0.036	0.062	20.216	0.003
7	-0.115	-0.096	22.230	0.002
8	0.006	-0.081	22.235	0.004
9	0.148	0.150	25.663	0.002
10	-0.097	0.002	27.139	0.002
11	0.048	-0.003	27.499	0.004
12	-0.088	-0.059	28.743	0.004
13	0.011	-0.013	28.762	0.007
14	-0.015	-0.070	28.798	0.011
15	0.141	0.117	32.015	0.006
16	-0.158	-0.055	36.102	0.003
17	0.097	0.030	37.656	0.003
18	-0.019	0.018	37.713	0.004
19	-0.103	-0.108	39.507	0.004
20	-0.006	-0.142	39.514	0.006
21	-0.003	-0.037	39.515	0.009
22	-0.089	-0.131	40.891	0.008
23	0.227	0.138	49.853	0.001
24	-0.099	-0.005	51.569	0.001
25	-0.023	-0.006	51.658	0.001
26	0.040	-0.017	51.946	0.002
27	-0.002	0.072	51.947	0.003
28	0.009	0.014	51.963	0.004
29	-0.004	0.011	51.966	0.005
30	-0.072	-0.064	52.913	0.006
31	-0.062	-0.108	53.627	0.007
32	0.198	0.048	60.960	0.002
33	-0.176	-0.089	66.825	0.000
34	0.080	-0.010	68.050	0.000
35	-0.072	-0.032	69.041	0.001
36	-0.071	-0.144	70.035	0.001

Figure 2: Plot and SACF/SPACF of the series $\Delta log(X_{t)}$ with the Census X12 procedure





	AC	PAC	Q-Stat	Prob
1	-0.346	-0.346	17.580	0.000
2	0.084	-0.041	18.615	0.000
3	-0.117	-0.115	20.666	0.000
4	-0.031	-0.123	20.810	0.000
5	0.098	0.054	22.264	0.000
6	0.045	0.103	22.569	0.001
7	-0.069	-0.038	23.301	0.002
8	-0.002	-0.028	23.301	0.003
9	0.172	0.226	27.906	0.001
10	-0.094	0.033	29.284	0.001
11	0.068	0.019	30.022	0.002
12	-0.256	-0.215	40.420	0.000
13	0.058	-0.101	40.957	0.000
14	-0.009	-0.082	40.970	0.000
15	0.120	0.038	43.301	0.000
16	-0.159	-0.145	47.428	0.000
17	0.098	0.041	49.012	0.000
18	0.018	0.133	49.064	0.000
19	-0.061	-0.033	49.681	0.000
20	-0.051	-0.144	50.118	0.000
21	0.026	0.121	50.231	0.000
22	-0.121	-0.102	52.752	0.000
23	0.269	0.168	65.306	0.000
24	-0.155	-0.117	69.494	0.000
25	-0.040	-0.095	69.777	0.000
26	0.069	0.013	70.619	0.000
27	-0.029	0.079	70.768	0.000
28	0.021	-0.107	70.849	0.000
29	-0.026	0.005	70.970	0.000
30	-0.062	-0.009	71.675	0.000
31	-0.058	-0.103	72.309	0.000
32	0.227	0.023	82.007	0.000
33	-0.159	0.026	86.810	0.000
34	0.070	-0.038	87.750	0.000
35	-0.092	0.040	89.374	0.000
36	-0.091	-0.234	90.994	0.000

Figure 3: Plot and SACF/SPACF of the series $\Delta log(X_{t)}$ with the TramoSeats procedure

(g) **Apply the Box-Jenkins methodology to select the model(s) of the class SARIMA$(\mathbf{p}, \mathbf{d}, \mathbf{q}) \times (\mathbf{P}, \mathbf{D}, \mathbf{Q})_{\mathbf{S}}$ that best fit the data. Recall that the multiplicative model SARIMA$(\mathbf{p}, \mathbf{d}, \mathbf{q}) \times (\mathbf{P}, \mathbf{D}, \mathbf{Q})_{\mathbf{S}}$ is defined by the equation:**

$$\boxed{\phi(L)\,\Phi(L^S)\Delta^d\Delta_S^D X_t = \theta(L)\,\Theta(L^S)\varepsilon_t,\ \varepsilon_t \overset{w.n.}{\sim} (0, \sigma_\varepsilon^2)}$$

Examine carefully the residual of the proposed model(s).

Model estimation and selection of (P,Q) and (p,q)

Dependent Variable: D(LOG(AIRLINE),1,12)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.000401	0.001723	-0.232626	0.8165
AR(1)	-0.413851	0.084830	-4.878584	0.0000
SAR(12)	-0.453451	0.082965	-5.465539	0.0000

R-squared	0.320812	Mean dependent var	-0.000931
Adjusted R-squared	0.309000	S.D. dependent var	0.046208
S.E. of regression	0.038411	Akaike info criterion	-3.655850
Sum squared resid	0.169672	Schwarz criterion	-3.585409
Log likelihood	218.6951	Hannan-Quinn criter.	-3.627249
F-statistic	27.15987	Durbin-Watson stat	2.029454
Prob(F-statistic)	0.000000		

Inverted AR Roots	.90+.24i	.90-.24i	.66-.66i	.66-.66i
	.24-.90i	.24+.90i	-.24-.90i	-.24+.90i
	-.41	-.66-.66i	-.66+.66i	-.90+.24i
	-.90-.24i			

(a) SARIMA$(1,1,0) \times (1,1,0)_{12}$

Dependent Variable: D(LOG(AIRLINE),1,12)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.000247	0.001183	-0.209255	0.8346
AR(1)	-0.342205	0.083423	-4.102040	0.0001
MA(12)	-0.591455	0.072679	-8.137934	0.0000

R-squared	0.360422	Mean dependent var	-8.14E-06
Adjusted R-squared	0.350350	S.D. dependent var	0.045897
S.E. of regression	0.036994	Akaike info criterion	-3.733338
Sum squared resid	0.173803	Schwarz criterion	-3.667164
Log likelihood	245.6670	Hannan-Quinn criter.	-3.706449
F-statistic	35.78419	Durbin-Watson stat	2.035123
Prob(F-statistic)	0.000000		

Inverted AR Roots	-.34			
Inverted MA Roots	.96	.83-.48i	.83+.48i	.48+.83i
	.48-.83i	.00+.96i	-.00-.96i	-.48+.83i
	-.48-.83i	-.83-.48i	-.83+.48i	-.96

(b) SARIMA$(1,1,0) \times (0,1,1)_{12}$

Dependent Variable: D(LOG(AIRLINE),1,12)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.000469	0.001230	-0.381527	0.7035
AR(12)	-0.470959	0.081654	-5.767755	0.0000
MA(1)	-0.486677	0.081465	-5.974096	0.0000

R-squared	0.337529	Mean dependent var	-0.001406
Adjusted R-squared	0.326107	S.D. dependent var	0.046303
S.E. of regression	0.038010	Akaike info criterion	-3.677039
Sum squared resid	0.167594	Schwarz criterion	-3.606977
Log likelihood	221.7838	Hannan-Quinn criter.	-3.648589
F-statistic	29.55101	Durbin-Watson stat	1.983132
Prob(F-statistic)	0.000000		

Inverted AR Roots	.91-.24i	.91+.24i	.66+.66i	.66-.66i
	.24-.91i	.24+.91i	-.24-.91i	-.24+.91i
	-.66-.66i	-.66-.66i	-.91+.24i	-.91-.24i
Inverted MA Roots	.49			

(c) SARIMA$(0,1,1) \times (1,1,0)_{12}$

Dependent Variable: D(LOG(AIRLINE),1,12)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.000322	0.000947	-0.340320	0.7342
MA(1)	-0.383664	0.079128	-4.848656	0.0000
SMA(12)	-0.611143	0.071432	-8.555572	0.0000

R-squared	0.371373	Mean dependent var	0.000291
Adjusted R-squared	0.361551	S.D. dependent var	0.045848
S.E. of regression	0.036634	Akaike info criterion	-3.753036
Sum squared resid	0.171784	Schwarz criterion	-3.687192
Log likelihood	248.8239	Hannan-Quinn criter.	-3.726281
F-statistic	37.80918	Durbin-Watson stat	1.986416
Prob(F-statistic)	0.000000		

Inverted MA Roots	.96	.83+.48i	.83-.48i	.48-.83i
	.48+.83i	.38	.00+.96i	-.00-.96i
	-.48+.83i	-.48-.83i	-.83-.48i	-.83+.48i
	-.96			

(d) SARIMA$(0,1,1) \times (0,1,1)_{12}$

Figure 4: Candidate SARIMA$(p,d,q) \times (P,D,Q)_{12}$ models

Plot and SACF/SPACF of the residuals from the candidate models

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.021	-0.021	0.0514	
		2	-0.117	-0.117	1.7203	
		3	-0.124	-0.131	3.6090	0.057
		4	-0.131	-0.157	5.7416	0.057
		5	0.104	0.063	7.1016	0.069
		6	0.104	0.063	8.4643	0.076
		7	-0.086	-0.099	9.4103	0.094
		8	-0.017	-0.005	9.4472	0.150
		9	0.148	0.184	12.306	0.091
		10	-0.126	-0.133	14.381	0.072
		11	-0.039	-0.063	14.581	0.103
		12	-0.022	0.001	14.648	0.145
		13	-0.022	-0.003	14.714	0.196
		14	0.067	-0.029	15.330	0.224
		15	0.070	0.054	16.011	0.249
		16	-0.145	-0.095	18.948	0.167
		17	0.016	0.016	18.984	0.214
		18	0.039	0.012	19.203	0.258
		19	-0.100	-0.088	20.622	0.244
		20	-0.043	-0.109	20.889	0.285
		21	0.036	0.050	21.076	0.333
		22	-0.044	-0.060	21.361	0.376
		23	0.169	0.100	25.618	0.221
		24	-0.198	-0.247	31.523	0.086
		25	-0.065	0.052	32.166	0.097
		26	0.078	0.019	33.100	0.102
		27	-0.062	-0.131	33.697	0.115
		28	-0.106	-0.203	35.470	0.102
		29	0.029	0.065	35.599	0.124
		30	-0.029	-0.076	35.731	0.150
		31	-0.033	-0.137	35.904	0.176
		32	0.160	0.052	40.144	0.102
		33	-0.187	-0.080	46.000	0.041
		34	-0.027	-0.107	46.118	0.051
		35	0.013	-0.085	46.145	0.064
		36	-0.026	-0.053	46.265	0.078

(a) SARIMA$(1,1,0) \times (1,1,0)_{12}$

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.019	-0.019	0.0484	
		2	-0.090	-0.090	1.1281	
		3	-0.126	-0.131	3.2713	0.071
		4	-0.152	-0.171	6.4229	0.040
		5	0.077	0.042	7.2338	0.065
		6	0.086	0.045	8.2549	0.083
		7	-0.097	-0.128	9.5687	0.088
		8	-0.029	-0.040	9.6873	0.138
		9	0.126	0.150	11.937	0.103
		10	-0.085	-0.099	12.968	0.113
		11	-0.003	-0.046	12.970	0.164
		12	-0.023	0.001	13.047	0.221
		13	0.008	0.049	13.057	0.290
		14	0.069	0.004	13.751	0.317
		15	0.034	0.017	13.922	0.379
		16	-0.152	-0.108	17.404	0.235
		17	0.048	0.066	17.752	0.276
		18	0.037	0.014	17.966	0.326
		19	-0.102	-0.128	19.584	0.296
		20	-0.094	-0.153	20.970	0.281
		21	-0.047	-0.015	21.324	0.319
		22	-0.024	-0.066	21.417	0.373
		23	0.228	0.117	29.718	0.098
		24	0.002	-0.039	29.719	0.125
		25	-0.041	0.045	29.993	0.150
		26	0.076	0.095	30.955	0.155
		27	-0.029	0.013	31.100	0.186
		28	-0.067	-0.097	31.852	0.198
		29	-0.024	-0.016	31.954	0.234
		30	-0.092	-0.062	33.403	0.221
		31	-0.022	-0.068	33.484	0.259
		32	0.142	0.013	37.035	0.176
		33	-0.129	-0.100	39.979	0.130
		34	-0.002	0.002	39.979	0.157
		35	-0.060	-0.110	40.639	0.169
		36	-0.026	-0.067	40.763	0.197

(b) SARIMA$(1,1,0) \times (0,1,1)_{12}$

Figure 5: SACF/SPACF of the residuals from the candidate models

SACF/SPACF of the residuals from the candidate models

	AC	PAC	Q-Stat	Prob
1	0.005	0.005	0.0027	
2	0.060	0.060	0.4517	
3	-0.123	-0.124	2.3447	0.126
4	-0.091	-0.095	3.3846	0.184
5	0.077	0.096	4.1368	0.247
6	0.080	0.079	4.9554	0.292
7	-0.046	-0.086	5.2324	0.388
8	-0.029	-0.030	5.3436	0.501
9	0.111	0.168	6.9692	0.432
10	-0.126	-0.142	9.0584	0.337
11	-0.022	-0.088	9.1246	0.426
12	-0.029	0.042	9.2379	0.510
13	0.001	0.027	9.2380	0.600
14	0.040	-0.040	9.4595	0.663
15	0.059	0.050	9.9406	0.699
16	-0.151	-0.105	13.141	0.515
17	0.010	0.003	13.156	0.590
18	-0.013	-0.004	13.181	0.660
19	-0.085	-0.093	14.221	0.651
20	-0.065	-0.117	14.835	0.673
21	0.031	0.081	14.972	0.724
22	-0.078	-0.074	15.877	0.724
23	0.146	0.077	19.087	0.580
24	-0.201	-0.223	25.184	0.288
25	-0.071	-0.002	25.967	0.302
26	-0.010	0.004	25.982	0.354
27	-0.075	-0.140	26.853	0.363
28	-0.103	-0.196	28.544	0.332
29	-0.002	0.059	28.545	0.383
30	-0.024	-0.028	28.638	0.431
31	-0.056	-0.142	29.145	0.458
32	0.142	0.059	32.495	0.345
33	-0.175	-0.063	37.637	0.191
34	0.001	-0.099	37.637	0.227
35	0.005	-0.031	37.642	0.265
36	0.018	-0.006	37.696	0.304

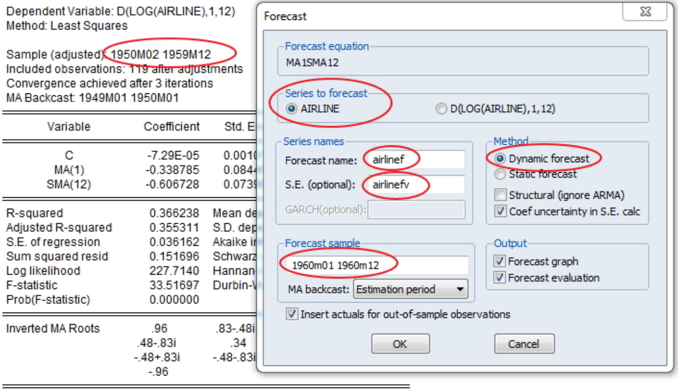
(a) SARIMA$(0, 1, 1) \times (1, 1, 0)_{12}$

	AC	PAC	Q-Stat	Prob
1	0.003	0.003	0.0010	
2	0.014	0.014	0.0285	
3	-0.129	-0.129	2.3059	0.129
4	-0.145	-0.146	5.1705	0.075
5	0.050	0.055	5.5221	0.137
6	0.061	0.052	6.0465	0.196
7	-0.075	-0.119	6.8348	0.233
8	-0.039	-0.054	7.0557	0.316
9	0.103	0.148	8.5823	0.284
10	-0.082	-0.096	9.5459	0.298
11	0.023	-0.040	9.6239	0.382
12	-0.014	0.026	9.6516	0.472
13	0.030	0.069	9.7837	0.550
14	0.042	-0.012	10.048	0.612
15	0.049	0.032	10.409	0.660
16	-0.157	-0.120	14.142	0.439
17	0.027	0.044	14.251	0.507
18	0.000	0.002	14.251	0.580
19	-0.106	-0.139	16.004	0.524
20	-0.101	-0.162	17.606	0.482
21	-0.031	0.027	17.763	0.538
22	-0.029	-0.042	17.897	0.594
23	0.220	0.129	25.715	0.218
24	0.028	-0.017	25.847	0.258
25	-0.019	0.036	25.906	0.305
26	0.063	0.077	26.558	0.325
27	-0.041	-0.004	26.846	0.364
28	-0.064	-0.109	27.543	0.381
29	-0.049	-0.035	27.946	0.414
30	-0.079	-0.045	29.017	0.412
31	-0.048	-0.060	29.418	0.443
32	0.125	0.024	32.169	0.360
33	-0.133	-0.106	35.323	0.271
34	0.002	-0.005	35.324	0.314
35	-0.063	-0.083	36.036	0.328
36	-0.018	-0.071	36.094	0.371

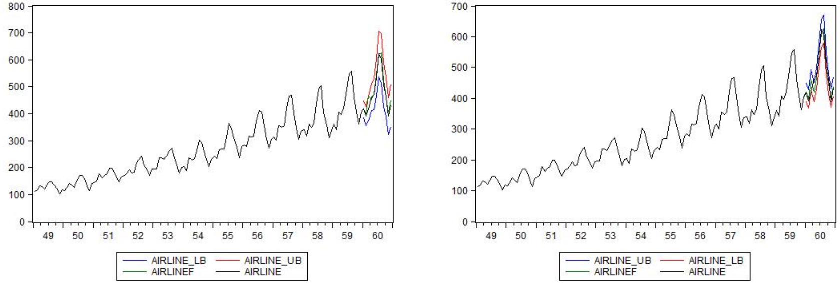
(b) SARIMA$(0, 1, 1) \times (0, 1, 1)_{12}$

Figure 6: SACF/SPACF of the residuals from the candidate models

(h) **Estimate the model without the last 12 observations and make dynamic and static forecasts for 1960:01 until 1960:12.**
Compare your forecasts with the realized values.
Draw a time series plot with the realized values, point and interval forecasts.



Forecasting with SARIMA models



(a) Dynamic forecast **(b) Static forecast**

Figure 7: Forecasting with SARIMA models

EXERCISE I

Choose a macroeconomic time series, either with monthly or quarterly frequency, with a clear <u>stochastic seasonality</u> pattern, and with a sample period of at least 10 years (if monthly) or 30 years (if quarterly). Now, disregard the last 12 observations of the sample and solve the following questions:

(a) **Construct the models from the SARIMA class that better caracterize the dynamic properties of the data (select at least 2 models). Justify your choices in detail.**

(b) **Make a complete diagnostic checking to the residuals.**

EXERCISE II

(c) **Acrescente dummies aos 2 modelos apresentados, interpret your estimates of the <u>constant</u> term and <u>one</u> of the dummies. Evaluate the statistical significance of the coefficients. Hint: The EViews command @seas may be useful.**

(d) **Use the two best models to construct dynamic and static forecasts for the disregarded forecasts with origin on the last observation used for model estimation.**

(e) **For the 2 best models, compare your predictions with the observed values. Represent the series plot of the realized values and the forecasts. Comment on your results.**

Bibliography

- Enders, W. (2009). Applied Econometric Time Series, chapter 2.11.
- Mills, T. C. (1991). Time series techniques for economists, chapter 10.
- Wei, W. W. S. (2005). Time series analysis: Univariate and Multivariate Methods, chapter 8.