CHAPTER 7.

FORECASTING WITH AUTOREGRESSIVE PROCESSES

SOLUTIONS

by

Wei Lin and Yingying Sun

(University of California, Riverside)

Exercise 1

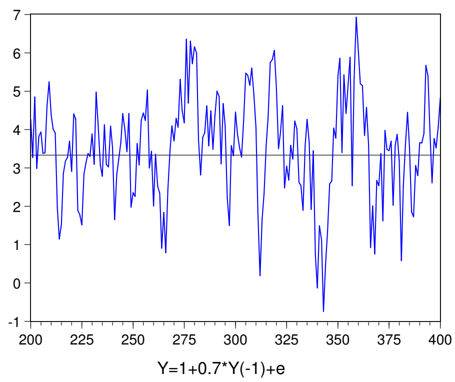
We simulate 2000 observations from the following two AR(1) processes and plot 200 observations.

$$\begin{aligned} Y_t &= 1 + 0.7Y_{t-1} + \epsilon_t \\ Y_t &= 1 - 0.7Y_{t-1} + \epsilon_t \end{aligned}$$

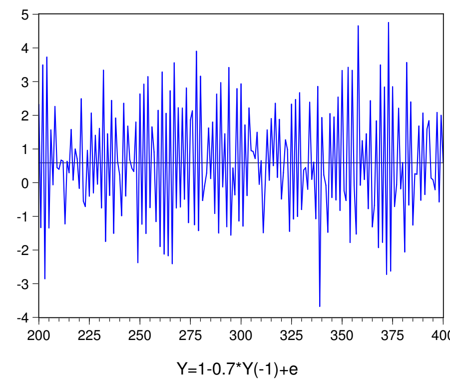
In Figure 1, we plot the time series Y1 (with persistence parameter 0.7) and Y2 (with persistence parameter -0.7). The series come from covariance stationary processes since the magnitude of their persistence parameter is less than one. They move around their respective unconditional means (the horizontal line), $\mu_1 = 1/0.3$ and $\mu_2 = 1/1.7$. The time series Y1 exhibits smoother dynamics than the time series Y2, which shows a zig-zag behavior. This is due to the different sign of the autoregressive parameters; a positive parameter means that positive (negative) observations tend to be followed by positive (negative) observations, while a negative parameter means that positive (negative) observations tend to be followed by negative (positive) observations. The autocorrelation functions in Figure 2 deliver the same message. For Y1, the autocorrelation coefficients are all positive, and for Y2, they have alternating signs. The profile of ACFs and PACFs of these two series are similar in that there is a smooth decay of the autocorrelations towards zero in ACFs and only a significant partial autocorrelation of order one in the PACFs. Since the autoregressive parameter of Y2 is negative, observe the alternating signs of the autocorrelations of this process and the negative value of the partial autocorrelation of order one.

The following EViews code provides the simulated time series Y1 and Y2:

```
series e=nrnd
smpl @first @first
series y1=1/0.3
series y2=1/1.7
smpl @first+1 @last
series y1=1+0.7*y1(-1)+e
series y2=1-0.7*y2(-1)+e
```

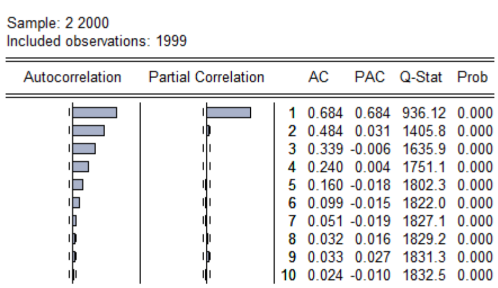


(a) Time Series Y1



(b) Time Series Y2

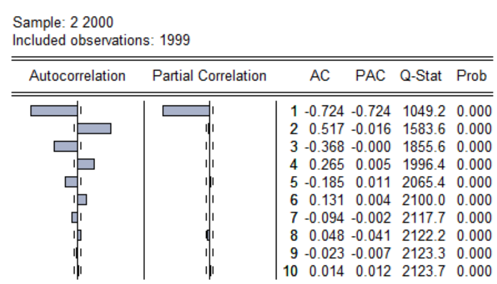
Figure 1: Time Series Plots



Sample: 2 2000
Included observations: 1999

	AC	PAC	Q-Stat	Prob
1	0.684	0.684	936.12	0.000
2	0.484	0.031	1405.8	0.000
3	0.339	-0.006	1635.9	0.000
4	0.240	0.004	1751.1	0.000
5	0.160	-0.018	1802.3	0.000
6	0.099	-0.015	1822.0	0.000
7	0.051	-0.019	1827.1	0.000
8	0.032	0.016	1829.2	0.000
9	0.033	0.027	1831.3	0.000
10	0.024	-0.010	1832.5	0.000

(a) ACF and PACF of Series Y1



Sample: 2 2000
Included observations: 1999

	AC	PAC	Q-Stat	Prob
1	-0.724	-0.724	1049.2	0.000
2	0.517	-0.016	1583.6	0.000
3	-0.368	-0.000	1855.6	0.000
4	0.265	0.005	1996.4	0.000
5	-0.185	0.011	2065.4	0.000
6	0.131	0.004	2100.0	0.000
7	-0.094	-0.002	2117.7	0.000
8	0.048	-0.041	2122.2	0.000
9	-0.023	-0.007	2123.3	0.000
10	0.014	0.012	2123.7	0.000

(b) ACF and PACF of Series Y2

Figure 2: ACF and PACF

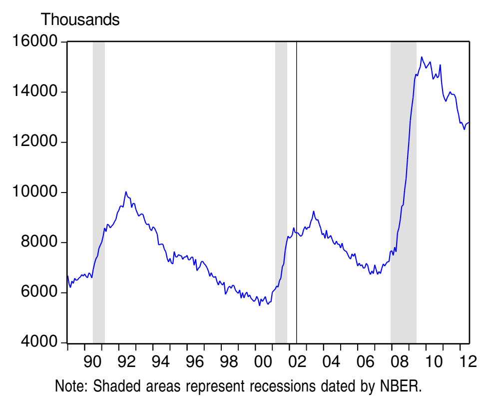
Exercise 2

We download the "unemployed persons" series from the Bureau of Labor Statistics (BLS) website:

`http://www.bls.gov/data/#unemployment`

In Figure 3a we plot the time series of monthly unemployed persons from January 1989 to July 2012.The shaded areas represent recessions recorded by NBER. The vertical line on June 2002 separates the sample in the textbook from the updated sample. One of the major consequences of the financial crisis of 2008 is the tremendous and fast uprise of unemployment in USA compared with the previous recessions of 1990-91 and 2001. The time series shows that there is a lot of persistence in unemployment with strong positive autocorrelation in the short run (within one to two years, $\rho_{12} = 0.76$), almost zero autocorrelation in the medium run (three to five years, $\rho_{36} = 0.05$), and low negative autocorrelation in the long run (ten years, $\rho_{120} = -0.18$).

In Figure 3b we report the autocorrelation functions.The profile of the autocorrelograms indicate that the process is autoregressive of at least order one. The ACF decays very slowly towards zero since the process is highly persistent. The partial autocorrelation coefficient of order one is 0.99, which is very close to the boundary for the process to be covariance-stationary. We can ask whether the process is non-stationary by conducting a test with a null hypothesis such as $H_0 : \phi = 1$. This is the topic that we will explain in Chapter 10. At this stage, we only raise the possibility that this process may not be covariance-stationary.



(a) Time Series Plot

Sample: 1989M01 2012M07
Included observations: 283

Autocorrelation	Partial Correlation		AC	PAC
		1	0.991	0.991
		2	0.979	-0.104
		3	0.966	-0.126
		4	0.951	-0.066
		5	0.933	-0.092
		6	0.914	-0.095
		7	0.893	-0.070
		8	0.869	-0.108
		9	0.844	-0.049
		10	0.817	-0.094
		11	0.789	-0.021
		12	0.759	-0.034

(b) ACF and PACF for 12 Months

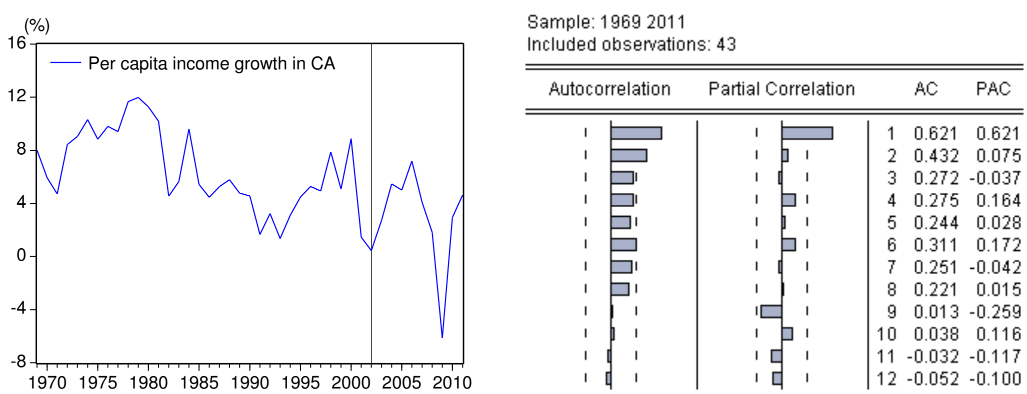
Figure 3: Unemployed Persons

Exercise 3

We download annual per capita income growth data from the Bureau of Economic Analysis (BEA) website:

`http://www.bea.gov/iTable/iTable.cfm?reqid=70&step=1&isuri=1&acrdn=3`

The updated time series runs from 1969 to 2011. In Figure 4a, we plot the data, and the vertical line on year 2002 separates the sample in the textbook from the updated sample. Observe that in the aftermath of the financial crisis of 2008, the growth was negative; in 2009 there was a severe decline of -6.14% (unprecedented in California given its past history) but it bounced back to positive territory in the years after. The ACF and PACF plots in Figure 4b are not very different from those in the textbook. They show the profile of an autoregressive process AR(1) because the only significant partial autocorrelation coefficient is of order one and the autocorrelations in the ACF decay slowly to zero. Table 1 reports the estimation results. When estimating the model, we have excluded the last three observations (from 2009 to 2011) in order to assess the multistep forecast.



(a) Time Series Plot

(b) ACF and PACF

Figure 4: Per Capita Income Growth (California, 1969 – 2011)

Dependent Variable: GPCI_CA				
Method: Least Squares				
Sample (adjusted): 1970 2008				
Included observations: 39 after adjustments				
Convergence achieved after 3 iterations				
Newey-West HAC Standard Errors & Covariance (lag truncation=3)				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	5.712405	1.042373	5.480192	0.0000
AR(1)	0.677893	0.133723	5.069379	0.0000
R-squared	0.440726	Mean dependent var		6.045897
Adjusted R-squared	0.42561	S.D. dependent var		3.094258
S.E. of regression	2.345092	Akaike info criterion		4.592447
Sum squared resid	203.48	Schwarz criterion		4.677758
Log likelihood	-87.55272	Hannan-Quinn criter.		4.623056
F-statistic	29.15716	Durbin-Watson stat		2.072829
Prob(F-statistic)	0.000004			
Inverted AR Roots	0.68			

Table 1: Estimation Results of AR(1) Model

Based on the estimated AR(1) model and assuming a quadratic loss function, we obtain the optimal forecast as follows,

Forecasting Horizon $h = 1$

- optimal point forecast $f_{t,1}$

$$f_{t,1} = \mu_{t+1|t} = E(Y_{t+1}|I_t) = E(c + \phi Y_t + \varepsilon_{t+1}|I_t) = c + \phi Y_t.$$

  Therefore, the optimal forecast of per capita income growth of California in 2009 is,

$$f_{2008,1} = \hat{\mu}_{2009|2008} = \hat{c} + \hat{\phi} Y_{2008} \approx 1.84 + 0.68 \times 1.83 \approx 3.08.$$

- 1-period-ahead forecast error $e_{2008,1}$

$$e_{2008,1} = Y_{2009} - f_{2008,1} = \varepsilon_{2009} \approx -6.14 - 3.08 = -9.22$$

- The uncertainty associated with the forecast

$$\sigma^2_{2009|2008} = \mathrm{var}(e_{2008,1}) = E(\varepsilon^2_{2009}) = \sigma^2_\varepsilon \approx 5.50$$

Forecasting Horizon $h = 2$

- optimal point forecast $f_{t,2}$

$$f_{t,2} = \mu_{t+2|t} = E(Y_{t+2}|I_t) = E(c + \phi Y_{t+1} + \varepsilon_{t+2}|I_t) = c + \phi f_{t,1}.$$

  Therefore, the optimal forecast of per capita income growth of California in 2010 is,

$$f_{2008,2} = \hat{\mu}_{2010|2008} = \hat{c} + \hat{\phi} f_{2008,1} \approx 1.84 + 0.68 \times 3.08 \approx 3.93.$$

- 2-period-ahead forecast error $e_{2008,2}$

$$e_{2008,2} = Y_{2010} - f_{2008,2} = \phi\varepsilon_{2009} + \varepsilon_{2010} \approx 2.94 - 3.93 = -0.99$$

- The uncertainty associated with the forecast

$$\begin{aligned} \sigma^2_{2010|2008} &= \mathrm{var}(e_{2008,2}) = E(\phi\varepsilon_{2009} + \varepsilon_{2010})^2 \\ &= \sigma^2_\varepsilon(1 + \phi^2) \approx 8.04 \end{aligned}$$

Forecasting Horizon $h = 3$

- optimal point forecast $f_{t,3}$

$$f_{t,3} = \mu_{t+3|t} = E(Y_{t+3}|I_t) = E(c + \phi Y_{t+2} + \varepsilon_{t+3}|I_t) = c + \phi f_{t,2}.$$

  Therefore, the optimal forecast of per capita income growth of California in 2011 is,

$$f_{2008,3} = \hat{\mu}_{2011|2008} = \hat{c} + \hat{\phi} f_{2008,2} \approx 1.84 + 0.68 \times 3.93 \approx 4.51.$$

- 3-period-ahead forecast error $e_{2008,3}$

$$e_{2008,3} = Y_{2011} - f_{2008,3} = \phi^2 \varepsilon_{2009} + \phi\varepsilon_{2010} + \varepsilon_{2011} \approx 4.63 - 4.51 = 0.12$$

- The uncertainty associated with the forecast

$$\begin{aligned} \sigma^2_{2011|2008} &= \text{var}(e_{2008,3}) = E(\phi^2\varepsilon_{2009} + \phi\varepsilon_{2010} + \varepsilon_{2011})^2 \\ &= \sigma^2_\varepsilon(1 + \phi^2 + \phi^4) \approx 9.22 \end{aligned}$$

Observe that this model is unable to predict the big drop in 2009; this year was too anomalous given the history of this time series. However, the forecast for 2010 and 2011 is very accurate as the time series in the latest years conform to past behavior.

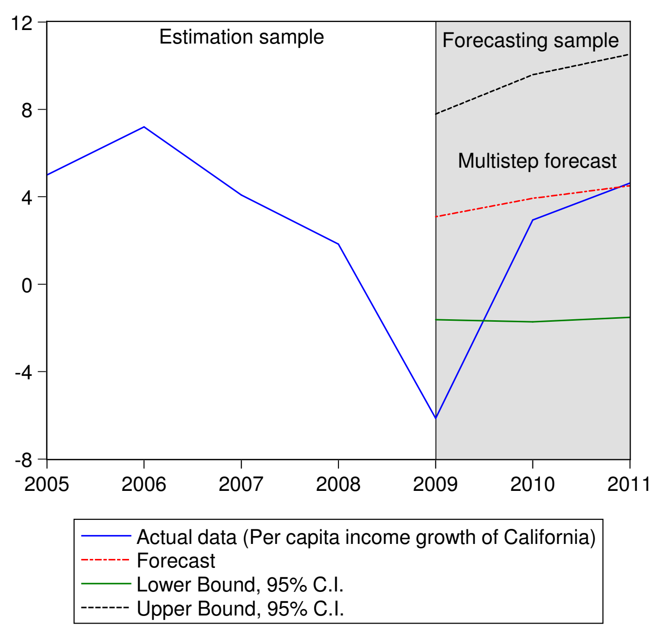
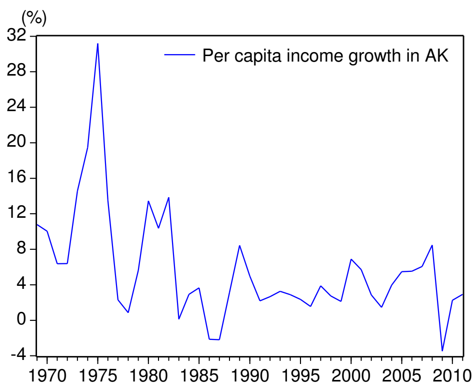


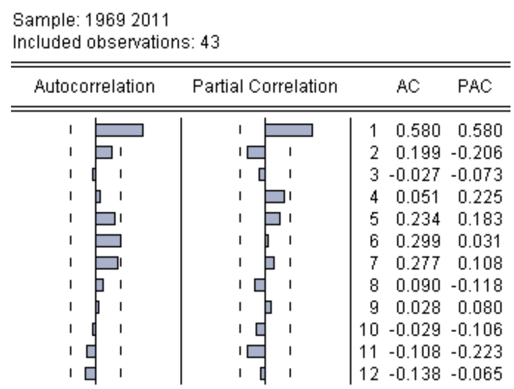
Figure 5: Per Capita Income Growth in California, Multistep Forecast

Exercise 4

We download per capita income growth in Alaska from the same website as in Exercise 3 to investigate whether it follows a similar process to that of California. The time series dates from 1969 to 2011. We plot the data in Figure 6a. The years 1973 to 1975 are very different (extremely high growth rates) from the most recent years and they contribute substantially to the high variance of the series. The cycle is not as evident as in the California time series. The financial crisis of 2008 also affected Alaska, in 2009 the growth rate was negative -3.47%.The ACF and PACF plots in Figure 4b show much weaker autocorrelations than those of California, but nevertheless the process seems to be autoregressive AR(1) or potentially AR(2). In Table 2 we report the estimation results of an AR(2) process. When estimating the model, we exclude the last three observations (from 2009 to 2011) to assess the multistep forecast.



(a) Time Series Plot



(b) ACF and PACF

Figure 6: Per Capita Income Growth (Alaska, 1969 – 2011)

Dependent Variable: GPCI_AK				
Method: Least Squares				
Sample (adjusted): 1971 2008				
Included observations: 38 after adjustments				
Convergence achieved after 3 iterations				
Newey-West HAC Standard Errors & Covariance (lag truncation=3)				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	6.001294	1.645752	3.646536	0.0009
AR(1)	0.773986	0.183084	4.227498	0.0002
AR(2)	-0.269122	0.115800	-2.324021	0.0261
R-squared	0.417162	Mean dependent var		5.976053
Adjusted R-squared	0.383857	S.D. dependent var		6.300417
S.E. of regression	4.9455	Akaike info criterion		6.11049
Sum squared resid	856.0288	Schwarz criterion		6.239773
Log likelihood	-113.0993	Hannan-Quinn criter.		6.156488
F-statistic	12.52548	Durbin-Watson stat		2.009107
Prob(F-statistic)	0.000079			
Inverted AR Roots	.39+.35i	.39-.35i		

Table 2: Estimation Results of AR(2) Model

Based on the estimated AR(2) model and assuming a quadratic loss function, we calculate the optimal forecast as follows,

Forecasting Horizon $h = 1$

- optimal point forecast $f_{t,1}$

$$f_{t,1} = \mu_{t+1|t} = E(Y_{t+1}|I_t) = E(c + \phi_1 Y_t + \phi_2 Y_{t-1} + \varepsilon_{t+1}|I_t) = c + \phi_1 Y_t + \phi_2 Y_{t-1}.$$

  Therefore, the optimal forecast of per capita income growth of Alaska in 2009 is,

$$f_{2008,1} = \hat{\mu}_{2009|2008} = \hat{c} + \hat{\phi}_1 Y_{2008} + \hat{\phi}_2 Y_{2007} \approx 2.97 + 0.774 \times 8.47 - 0.269 \times 6.07 \approx 7.89.$$

- 1-period-ahead forecast error $e_{2008,1}$

$$e_{2008,1} = Y_{2009} - f_{2008,1} = \varepsilon_{2009} \approx -3.47 - 7.89 = -11.36$$

- The uncertainty associated with the forecast

$$\sigma^2_{2009|2008} = \text{var}(e_{2008,1}) = E(\varepsilon^2_{2009}) = \sigma^2_\varepsilon \approx 24.46$$

Forecasting Horizon $h = 2$

- optimal point forecast $f_{t,2}$

$$f_{t,2} = \mu_{t+2|t} = E(Y_{t+2}|I_t) = E(c + \phi_1 Y_{t+1} + \phi_2 Y_t + \varepsilon_{t+2}|I_t) = c + \phi_1 f_{t,1} + \phi_2 Y_t.$$

  Therefore, the optimal forecast of per capita income growth of Alaska in 2010 is,

$$\begin{aligned} f_{2008,2} &= \hat{\mu}_{2010|2008} = \hat{c} + \hat{\phi}_1 f_{2008,1} + \hat{\phi}_2 Y_{2008} \\ &\approx 2.97 + 0.774 \times 7.89 - 0.269 \times 8.47 \approx 6.80. \end{aligned}$$

- 2-period-ahead forecast error $e_{2008,2}$

$$e_{2008,2} = Y_{2010} - f_{2008,2} = \phi_1 \varepsilon_{2009} + \varepsilon_{2010} \approx 2.25 - 6.80 = -4.55$$

- The uncertainty associated with the forecast

$$\begin{aligned} \sigma^2_{2010|2008} &= \text{var}(e_{2008,2}) = E(\phi_1 \varepsilon_{2009} + \varepsilon_{2010})^2 \\ &= \sigma^2_\varepsilon (1 + \phi_1) \approx 24.46(1 + 0.774^2) \approx 39.11. \end{aligned}$$

Forecasting Horizon $h = 3$

- optimal point forecast $f_{t,3}$

$$f_{t,3} = \mu_{t+3|t} = E(Y_{t+3}|I_t) = E(c + \phi_1 Y_{t+2} + \phi_2 Y_{t+1} + \varepsilon_{t+3}|I_t) = c + \phi_1 f_{t,2} + \phi_2 f_{t,1}.$$

Therefore, the optimal forecast of per capita income growth of Alaska in 2011 is,

$$\begin{aligned} f_{2008,3} &= \hat{\mu}_{2011|2008} = \hat{c} + \hat{\phi}_1 f_{2008,2} + \hat{\phi}_2 f_{2008,1} \\ &\approx 2.97 + 0.774 \times 6.80 - 0.269 \times 7.89 \approx 6.11. \end{aligned}$$

- 3-period-ahead forecast error $e_{2008,3}$

$$e_{2008,3} = Y_{2011} - f_{2008,3} = (\phi_1^2 + \phi_2)\varepsilon_{2009} + \phi_1\varepsilon_{2010} + \varepsilon_{2011} \approx 2.93 - 6.11 = -3.18$$

- The uncertainty associated with the forecast

$$\begin{aligned} \sigma^2_{2010|2008} &= \text{var}(e_{2008,2}) = E((\phi_1^2 + \phi_2)\varepsilon_{2009} + \phi_1\varepsilon_{2010} + \varepsilon_{2011})^2 \\ &= \sigma^2_\varepsilon((\phi_1^2 + \phi_2)^2 + \phi_1^2 + 1) \approx 46.06 \end{aligned}$$

Note that the forecast uncertainty in the Alaska series is much higher than in the California series. The primary reason is that the variance of the innovation in the AR(2) model is about twice the innovation variance in the AR(1) model for California. Again, in this series the severe drop in income in 2009 is difficult to predict but nevertheless is almost contained within the 95% confidence bands.

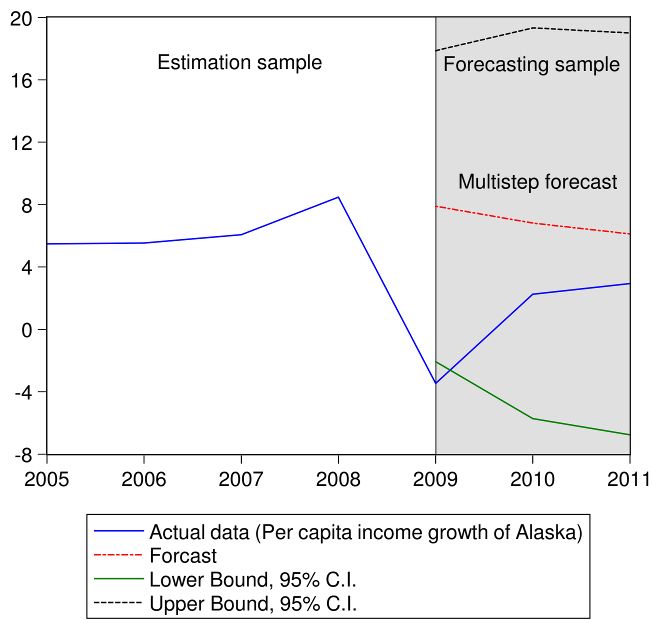


Figure 7: Per Capita Income Growth in Alaska, Multistep Forecast

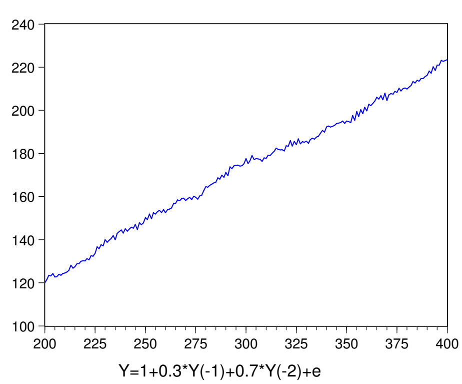
Exercise 5
We simulate 2000 observations from the following two AR(2) processes and we plot 200 observations in Figure 8.

$$
\begin{aligned}
Y_t &= 1 + 0.3Y_{t-1} + 0.7Y_{t-2} + \epsilon_t \\
Y_t &= 1 - 0.3Y_{t-1} - 0.7Y_{t-2} + \epsilon_t
\end{aligned}
$$

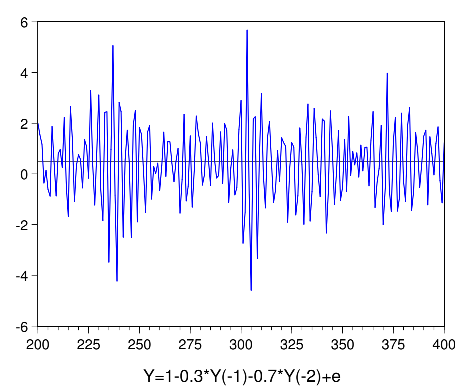
The first time series (Y1) with $\phi_1 = 0.3$ and $\phi_2 = 0.7$ satisfies the necessary conditions for covariance stationarity but it does not satisfy the sufficient condition $\phi_1 + \phi_2 < 1$. Thus, the process is non-stationary. This is obvious in the time series plot of Y1, which exhibits an upward trend. In contrast, the second time series Y2 satisfies both the necessary and the sufficient conditions for covariance-stationarity, i.e. $2 > \phi_1 = -0.3 > -2, 1 > \phi_2 = -0.7 > -1, \phi_1 + \phi_2 = -1 < 1, \phi_2 - \phi_1 = -0.4 < 1$). The time series plot of Y2 shows that the series moves around a central value that is the unconditional mean of the process. We report the corresponding autocorrelation functions in Figure 9. The non-stationary series Y1 has autocorrelation coefficients equal to one for any displacement, and only one significant partial autocorrelation coefficient equal to one. This is the profile of a non-stationary process, known as unit root process, which we will study in Chapter 10. The stationary series Y2 has autocorrelation coefficients decaying towards zero in an alternating fashion, and two significant partial autocorrelation coefficients much smaller than one.

The two time series can be simulated with the following EViews code:

```
series e=nrnd
smpl @first @first+1
series y1=0
series y2=0
smpl @first+2 @last
series y1=1+0.3*y1(-1)+0.7*y1(-2)+e
series y2=1-0.3*y2(-1)-0.7*y2(-2)+e
```

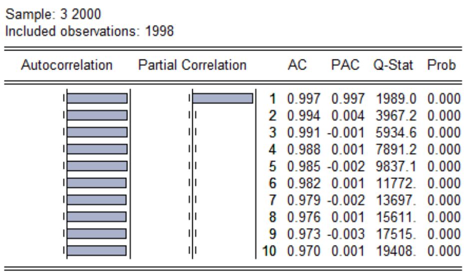


(a) Time Series Y1



(b) Time Series Y2

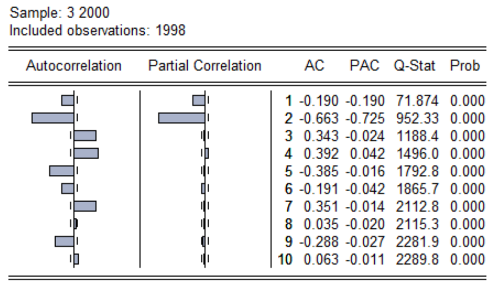
Figure 8: Time Series Plots


Sample: 3 2000
Included observations: 1998

Autocorrelation Partial Correlation

	AC	PAC	Q-Stat	Prob
1	0.997	0.997	1989.0	0.000
2	0.994	0.004	3967.2	0.000
3	0.991	-0.001	5934.6	0.000
4	0.988	0.001	7891.2	0.000
5	0.985	-0.002	9837.1	0.000
6	0.982	0.001	11772.	0.000
7	0.979	-0.002	13697.	0.000
8	0.976	0.001	15611.	0.000
9	0.973	-0.003	17515.	0.000
10	0.970	0.001	19408.	0.000

(a) ACF and PACF of Series Y1


Sample: 3 2000
Included observations: 1998

Autocorrelation Partial Correlation

	AC	PAC	Q-Stat	Prob
1	-0.190	-0.190	71.874	0.000
2	-0.663	-0.725	952.33	0.000
3	0.343	-0.024	1188.4	0.000
4	0.392	0.042	1496.0	0.000
5	-0.385	-0.016	1792.8	0.000
6	-0.191	-0.042	1865.7	0.000
7	0.351	-0.014	2112.8	0.000
8	0.035	-0.020	2115.3	0.000
9	-0.288	-0.027	2281.9	0.000
10	0.063	-0.011	2289.8	0.000

(b) ACF and PACF of Series Y2

Figure 9: ACF and PACF

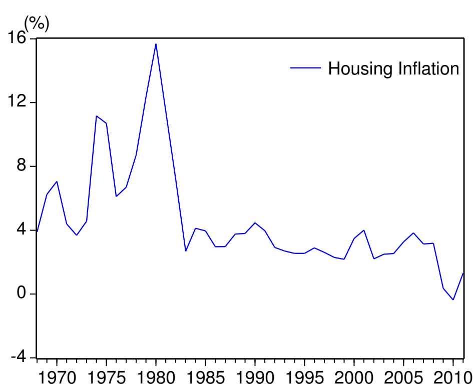
Exercise 6

We download the inflation rate for housing and transportation from the Bureau of Labor Statistics (BLS) website:

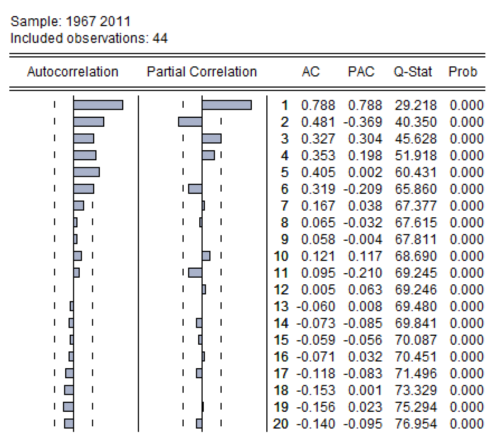
`http://www.bls.gov/data/#prices`

In Figure 10, we plot the time series and the autocorrelation functions for housing inflation. The ACF exhibits a smooth decay towards zero and the PACF has two, potentially three spikes different from zero. We would estimate both models AR(2) and AR(3) and we would analyze whether the autoregressive parameters are statistically different from zero in order to choose the order of the process.

In Figure 11, we plot the time series and the autocorrelation functions for transportation inflation. The time series of housing and transportation inflation have in common the high levels of inflation in the late 70s and early 80s but the autocorrelation functions are very different. Transportation inflation exhibits less dependence than housing inflation. The only autocorrelation coefficient statistically different from zero is ρ_1 though in the overall profile we see some decay towards zero. We propose an AR(1) with a small autoregressive parameter of about $\hat{\phi} = 0.40$.



(a) Time Series Plot

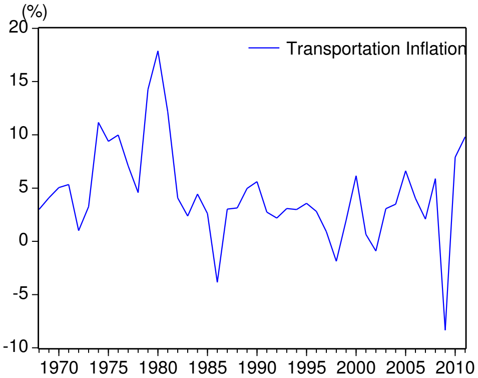

Sample: 1967 2011
Included observations: 44

Autocorrelation Partial Correlation

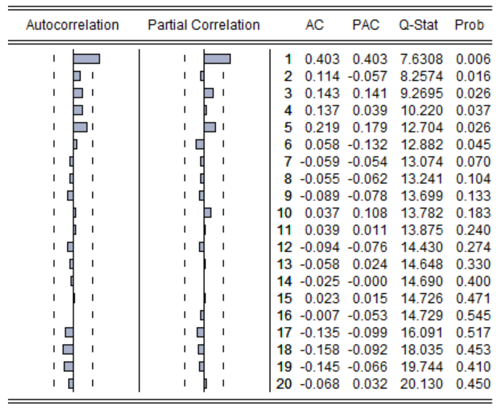
	AC	PAC	Q-Stat	Prob
1	0.788	0.788	29.218	0.000
2	0.481	-0.369	40.350	0.000
3	0.327	0.304	45.628	0.000
4	0.353	0.198	51.918	0.000
5	0.405	0.002	60.431	0.000
6	0.319	-0.209	65.860	0.000
7	0.167	0.038	67.377	0.000
8	0.065	-0.032	67.615	0.000
9	0.058	-0.004	67.811	0.000
10	0.121	0.117	68.690	0.000
11	0.095	-0.210	69.245	0.000
12	0.005	0.063	69.246	0.000
13	-0.060	0.008	69.480	0.000
14	-0.073	-0.085	69.841	0.000
15	-0.059	-0.056	70.087	0.000
16	-0.071	0.032	70.451	0.000
17	-0.118	-0.083	71.496	0.000
18	-0.153	0.001	73.329	0.000
19	-0.156	0.023	75.294	0.000
20	-0.140	-0.095	76.954	0.000

(b) ACF and PACF

Figure 10: Housing Inflation



(a) Time Series Plot



Sample: 1967 2011
Included observations: 44

	AC	PAC	Q-Stat	Prob
1	0.403	0.403	7.6308	0.006
2	0.114	-0.057	8.2574	0.016
3	0.143	0.141	9.2695	0.026
4	0.137	0.039	10.220	0.037
5	0.219	0.179	12.704	0.026
6	0.058	-0.132	12.882	0.045
7	-0.059	-0.054	13.074	0.070
8	-0.055	-0.062	13.241	0.104
9	-0.089	-0.078	13.699	0.133
10	0.037	0.108	13.782	0.183
11	0.039	0.011	13.875	0.240
12	-0.094	-0.076	14.430	0.274
13	-0.058	0.024	14.648	0.330
14	-0.025	-0.000	14.690	0.400
15	0.023	0.015	14.726	0.471
16	-0.007	-0.053	14.729	0.545
17	-0.135	-0.099	16.091	0.517
18	-0.158	-0.092	18.035	0.453
19	-0.145	-0.066	19.744	0.410
20	-0.068	0.032	20.130	0.450

(b) ACF and PACF

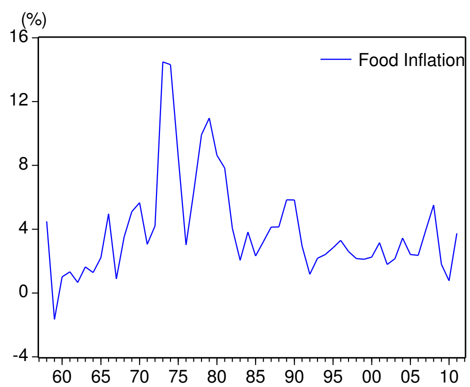
Figure 11: Transportation Inflation

Exercise 7

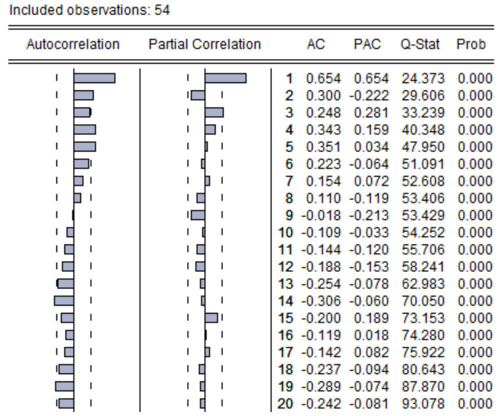
We download the inflation rate for food and gas (not seasonally adjusted) from the Bureau of Labor Statistics (BLS) website:

`http://www.bls.gov/data/#prices`

In Figure 12, we report the annual time series and the autocorrelation functions for food inflation. The ACF and PACF point towards an AR(3) process, which is estimated in Table 3. When estimating the model, we have excluded the last three observations (from 2009 to 2011) to assess the multistep forecast.The estimation results support an AR(3) model with all three autoregressive parameters statistically different from zero at the 5% level.



(a) Time Series Plot



Sample: 1958 2011
Included observations: 54

	AC	PAC	Q-Stat	Prob
1	0.654	0.654	24.373	0.000
2	0.300	-0.222	29.606	0.000
3	0.248	0.281	33.239	0.000
4	0.343	0.159	40.348	0.000
5	0.351	0.034	47.950	0.000
6	0.223	-0.064	51.091	0.000
7	0.154	0.072	52.608	0.000
8	0.110	-0.119	53.406	0.000
9	-0.018	-0.213	53.429	0.000
10	-0.109	-0.033	54.252	0.000
11	-0.144	-0.120	55.706	0.000
12	-0.188	-0.153	58.241	0.000
13	-0.254	-0.078	62.983	0.000
14	-0.306	-0.060	70.050	0.000
15	-0.200	0.189	73.153	0.000
16	-0.119	0.018	74.280	0.000
17	-0.142	0.082	75.922	0.000
18	-0.237	-0.094	80.643	0.000
19	-0.289	-0.074	87.870	0.000
20	-0.242	-0.081	93.078	0.000

(b) ACF and PACF

Figure 12: Food Inflation

Dependent Variable: FOOD_INFL				
Method: Least Squares				
Sample (adjusted): 1961 2008				
Included observations: 48 after adjustments				
Convergence achieved after 3 iterations				
Newey-West HAC Standard Errors & Covariance (lag truncation=3)				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	4.330453	1.071782	4.040422	0.0002
AR(1)	0.946371	0.142761	6.629081	0.0000
AR(2)	-0.50267	0.223109	-2.25301	0.0293
AR(3)	0.279156	0.140173	1.991506	0.0527
R-squared	0.555485	Mean dependent var		4.224793
Adjusted R-squared	0.525177	S.D. dependent var		3.163886
S.E. of regression	2.180151	Akaike info criterion		4.47632
Sum squared resid	209.1345	Schwarz criterion		4.632254
Log likelihood	-103.432	F-statistic		18.32811
Durbin-Watson stat	1.983789	Prob(F-statistic)		0.000000
Inverted AR Roots	0.77	.09+.60i	.09-.60i	

Table 3: Estimation Results of AR(3) Model for Food Inflation

Based on the estimated AR(3) model and assuming a quadratic loss function, we calculate the optimal forecast as follows,

Forecasting Horizon $h = 1$

- optimal point forecast $f_{t,1}$

$$\begin{aligned} f_{t,1} &= \mu_{t+1|t} = E(Y_{t+1}|I_t) = E(c + \phi_1 Y_t + \phi_2 Y_{t-1} + \phi_3 Y_{t-2} + \varepsilon_{t+1}|I_t) \\ &= c + \phi_1 Y_t + \phi_2 Y_{t-1} + \phi_3 Y_{t-2} \end{aligned}$$

Therefore, the optimal forecast of food inflation in 2009 is,

$$\begin{aligned} f_{2008,1} &= \hat{\mu}_{2009|2008} = \hat{c} + \hat{\phi}_1 Y_{2008} + \hat{\phi}_2 Y_{2007} + \hat{\phi}_3 Y_{2006} \\ &\approx 1.17 + 0.95 \times 5.51 - 0.50 \times 3.59 + 0.28 \times 2.36 \approx 5.09 \end{aligned}$$

- 1-period-ahead forecast error $e_{2008,1}$

$$e_{2008,1} = Y_{2009} - f_{2008,1} = \varepsilon_{2009} \approx 1.80 - 5.09 = -3.29$$

- The uncertainty associated with the forecast

$$\sigma^2_{2009|2008} = \text{var}(e_{2008,1}) = E(\varepsilon^2_{2009}) = \sigma^2_\varepsilon \approx 4.84$$

Forecasting Horizon $h = 2$

- optimal point forecast $f_{t,2}$

$$\begin{aligned} f_{t,2} &= \mu_{t+2|t} = E(Y_{t+2}|I_t) = E(c + \phi_1 Y_{t+1} + \phi_2 Y_t + \phi_3 Y_{t-1} + \varepsilon_{t+2}|I_t) \\ &= c + \phi_1 f_{t,1} + \phi_2 Y_t + \phi_3 Y_{t-1} \end{aligned}$$

Therefore, the optimal forecast of food inflation in 2010 is,

$$\begin{aligned} f_{2008,2} &= \hat{\mu}_{2010|2008} = \hat{c} + \hat{\phi}_1 f_{2008,1} + \hat{\phi}_2 Y_{2008} + \hat{\phi}_3 Y_{2007} \\ &\approx 1.17 + 0.95 \times 5.09 - 0.50 \times 5.51 + 0.28 \times 3.59 \approx 4.35 \end{aligned}$$

- 2-period-ahead forecast error $e_{2008,2}$

$$e_{2008,2} = Y_{2010} - f_{2008,2} = \phi_1 \varepsilon_{2009} + \varepsilon_{2010} \approx 0.77 - 4.35 = -3.58$$

- The uncertainty associated with the forecast

$$\begin{aligned} \sigma^2_{2010|2008} &= \operatorname{var}(e_{2008,2}) = E(\phi_1 \varepsilon_{2009} + \varepsilon_{2010})^2 \\ &= \sigma^2_\varepsilon (1 + \phi_1^2) \approx 9.12 \end{aligned}$$

Forecasting Horizon $h = 3$

- optimal point forecast $f_{t,3}$

$$\begin{aligned} f_{t,3} &= \mu_{t+3|t} = E(Y_{t+3}|I_t) \\ &= E(c + \phi_1 Y_{t+2} + \phi_2 Y_{t+1} + \phi_3 Y_t + \varepsilon_{t+3}|I_t) \\ &= c + \phi_1 f_{t,2} + \phi_2 f_{t,1} + \phi_3 Y_t \end{aligned}$$

Therefore, the optimal forecast of food inflation in 2011 is,

$$\begin{aligned} f_{2008,3} &= \hat{\mu}_{2011|2008} = \hat{c} + \hat{\phi}_1 f_{2008,2} + \hat{\phi}_2 f_{2008,1} + \hat{\phi}_2 Y_{2008} \\ &\approx 1.17 + 0.95 \times 4.35 - 0.50 \times 5.09 + 0.28 \times 5.51 \approx 4.29 \end{aligned}$$

- 3-period-ahead forecast error $e_{2008,3}$

$$e_{2008,3} = Y_{2011} - f_{2008,3} = (\phi_1^2 + \phi_2)\varepsilon_{2009} + \phi_1 \varepsilon_{2010} + \varepsilon_{2011} \approx 3.74 - 4.29 = -0.55$$

- The uncertainty associated with the forecast

$$\begin{aligned} \sigma^2_{2011|2008} &= \operatorname{var}(e_{2008,3}) = E((\phi_1^2 + \phi_2)\varepsilon_{2009} + \phi_1 \varepsilon_{2010} + \varepsilon_{2011})^2 \\ &= \sigma^2_\varepsilon ((\phi_1^2 + \phi_2)^2 + \phi_1^2 + 1) \approx 9.85 \end{aligned}$$

In Figure 13, we plot the multistep forecast and the 95% confidence bands; though the point forecast is above the actual values, these are well contained within the bands.

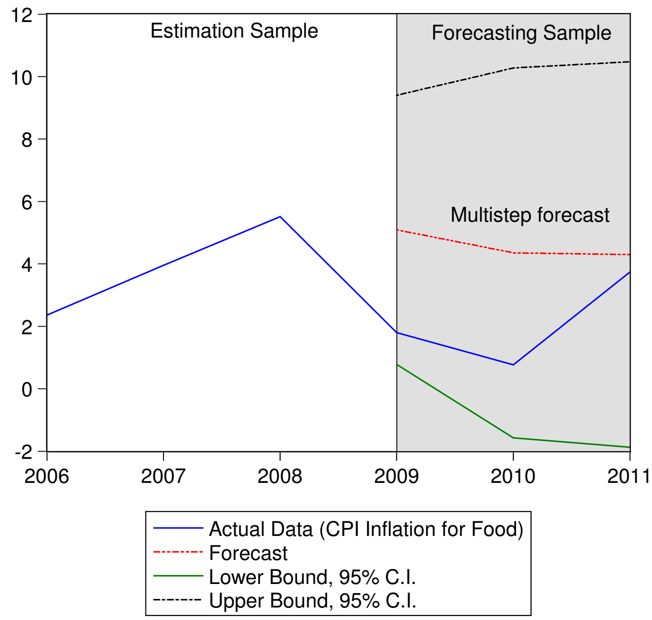
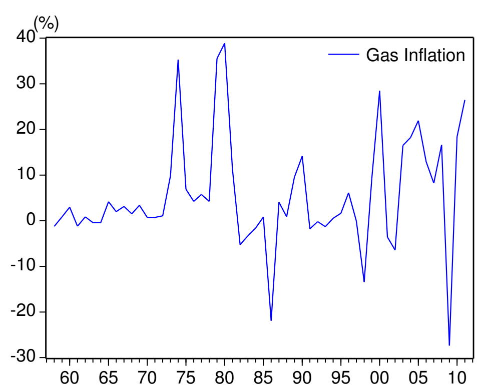
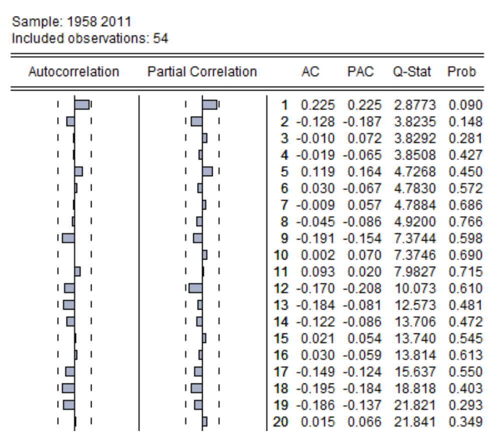


Figure 13: Food inflation, Multistep Forecast

In Figure 14, we plot the annual time series and the autocorrelation functions of gas inflation. The profile of the ACF and PACF shows that the series is very much a white noise process. The Q-statistics fail to reject the null hypothesis of autocorrelations being equal to zero. This means that the best linear forecast is the unconditional mean of the process; there is no dependence that we can exploit. We could explore the possibility of an MA(1) given that the autocorrelation coefficient of order one seems to be marginally significant. The estimation results are presented in Table 4. When estimating this model, we have excluded the last three observations (from 2009 to 2011) to assess the multistep forecast. Observe that the series is also very volatile; the unconditional standard deviation is twice as much as the unconditional mean.



(a) Time Series Plot



Sample: 1958 2011
Included observations: 54

	AC	PAC	Q-Stat	Prob
1	0.225	0.225	2.8773	0.090
2	-0.128	-0.187	3.8235	0.148
3	-0.010	0.072	3.8292	0.281
4	-0.019	-0.065	3.8508	0.427
5	0.119	0.164	4.7268	0.450
6	0.030	-0.067	4.7830	0.572
7	-0.009	0.057	4.7884	0.686
8	-0.045	-0.086	4.9200	0.766
9	-0.191	-0.154	7.3744	0.598
10	0.002	0.070	7.3746	0.690
11	0.093	0.020	7.9827	0.715
12	-0.170	-0.208	10.073	0.610
13	-0.184	-0.081	12.573	0.481
14	-0.122	-0.086	13.706	0.472
15	0.021	0.054	13.740	0.545
16	0.030	-0.059	13.814	0.613
17	-0.149	-0.124	15.637	0.550
18	-0.195	-0.184	18.818	0.403
19	-0.186	-0.137	21.821	0.293
20	0.015	0.066	21.841	0.349

(b) ACF and PACF

Figure 14: Gas Inflation

Dependent Variable: GAS_INFL				
Method: Least Squares				
Sample: 1958 2008				
Included observations: 51				
Convergence achieved after 7 iterations				
Newey-West HAC Standard Errors & Covariance (lag truncation=3)				
MA Backcast: 1957				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	5.604229	2.113868	2.651173	0.0108
MA(1)	0.571961	0.156631	3.651634	0.0006
R-squared	0.202564	Mean dependent var		5.50256
Adjusted R-squared	0.18629	S.D. dependent var		11.3776
S.E. of regression	10.26327	Akaike info criterion		7.533446
Sum squared resid	5161.399	Schwarz criterion		7.609204
Log likelihood	-190.103	F-statistic		12.44693
Durbin-Watson stat	2.100008	Prob(F-statistic)		0.000921
Inverted MA Roots	-0.57			

Table 4: Estimation Results of MA(1) Model for Gas

Based on the estimated MA(1) model and assuming a quadratic loss function, we calculate the optimal forecasting as follows,

FORECASTING HORIZON $h = 1$

- optimal point forecast $f_{t,1}$

$$f_{t,1} = \mu_{t+1|t} = E(Y_{t+1}|I_t) = E(\mu + \theta\varepsilon_t + \varepsilon_{t+1}|I_t) = \mu + \theta\varepsilon_t$$

  Therefore, the optimal forecast of gas inflation in 2009 is,

$$f_{2008,1} = \hat{\mu}_{2009|2008} = \hat{\mu} + \hat{\theta}\hat{\varepsilon}_t = 5.60 + 0.57\hat{\varepsilon}_t \approx 11.02$$

- 1-period-ahead forecast error $e_{2008,1}$

$$e_{2008,1} = Y_{2009} - f_{2008,1} = \varepsilon_{2009} \approx -27.36 - 11.02 = -38.38$$

- The uncertainty associated with the forecast

$$\sigma^2_{2009|2008} = \text{var}(e_{2008,1}) = E(\varepsilon^2_{2009}) = \hat{\sigma}^2_\varepsilon \approx 10.26^2 = 105.27$$

FORECASTING HORIZON $h = 2$

- optimal point forecast $f_{t,2}$

$$f_{t,2} = \mu_{t+2|t} = E(Y_{t+2}|I_t) = E(\mu + \theta\varepsilon_{t+1} + \varepsilon_{t+2}|I_t) = \mu$$

  Therefore, the optimal forecast of gas inflation in 2010 is $\hat{\mu} \approx 5.60$.

- 2-period-ahead forecast error $e_{2008,2}$

$$e_{2008,2} = Y_{2010} - f_{2008,2} = \theta\varepsilon_{2009} + \varepsilon_{2010} = Y_{2010} - f_{2008,2} \approx 18.38 - 5.60 = 12.78$$

- The uncertainty associated with the forecast

$$\sigma^2_{2010|2008} = \text{var}(e_{2008,2}) = (1 + \hat{\theta}^2)\hat{\sigma}^2_{\varepsilon} \approx 10.26^2(1 + 0.57^2) = 139.47$$

FORECASTING HORIZON $h = 3$

- optimal point forecast $f_{t,3}$

$$f_{t,3} = \mu_{t+3|t} = E(Y_{t+3}|I_t) = E(\mu + \theta\varepsilon_{t+2} + \varepsilon_{t+3}|I_t) = \mu$$

Therefore, the optimal forecast of gas inflation in 2011 is $\mu \approx 5.60$

- 3-period-ahead forecast error $e_{2008,3}$

$$e_{2008,3} = Y_{2011} - f_{2008,3} = \theta\varepsilon_{2010} + \varepsilon_{2009} = Y_{2011} - f_{2008,3} \approx 26.45 - 5.60 = 20.58$$

- The uncertainty associated with the forecast

$$\sigma^2_{2011|2008} = \text{var}(e_{2008,3}) = (1 + \hat{\theta}^2)\hat{\sigma}^2_{\varepsilon} \approx 10.26^2(1 + 0.57^2) = 139.47$$

As we expected, from $h = 2$ onwards the forecast reverts to the unconditional mean of the process. The 95% confidence bands are very wide reflecting the high variance of the series. series. In Figure 15, we plot the multistep forecast.

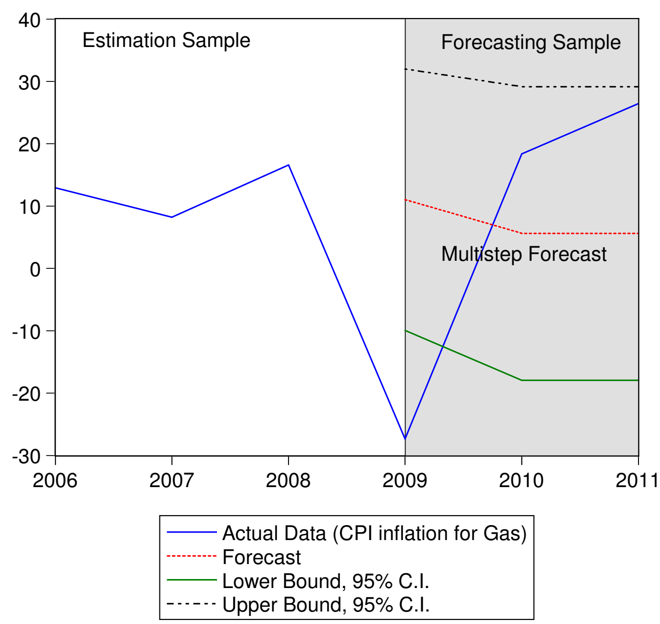
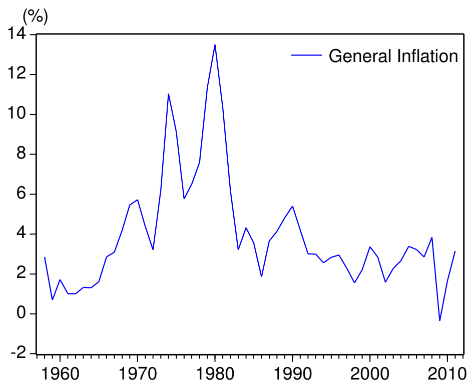


Figure 15: Gas inflation, Multistep Forecast

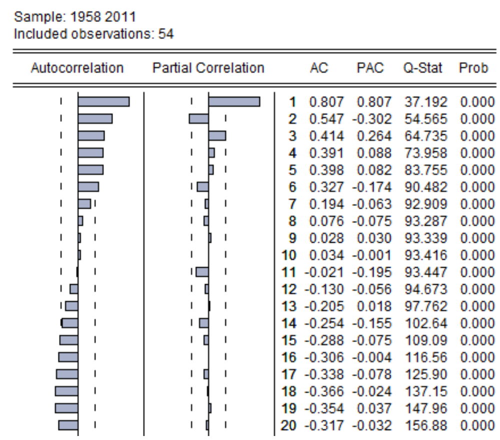
Exercise 8
We download the time series of general inflation (not seasonally adjusted) and the general inflation rates excluding food and gas from the Bureau of Labor Statistics (BLS) website:

`http://www.bls.gov/data/#prices`

In Figure 16, we present the annual time series and autocorrelation functions for the general inflation, and in Figure 17, we report similar plots for the general inflation excluding food and gas share. In both cases, the profiles of the ACF and PACF correspond to an autoregressive process AR(3). The estimation results are shown in Tables 5 and 6. When estimating the models, we have excluded the last two observations (2010 and 2011) to assess the multistep forecast. Observe that the unconditional mean of both series in this time period is about the same (4.13%) but the general inflation series is more volatile than the inflation excluding food and gas. The estimated AR(3) processes are also very similar; the estimated autoregressive parameters are of about the same magnitude and sign and all are statistically significant from zero; and the overall persistence of the process is about 0.86 (the sum of the three autoregressive parameters).



(a) Time Series Plot

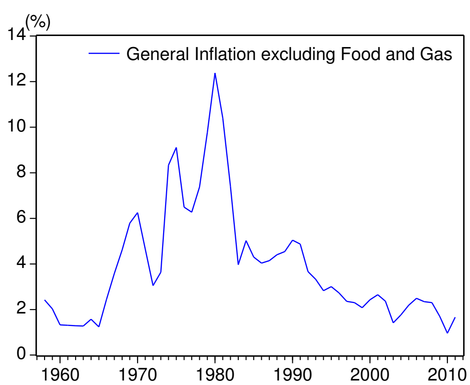


Sample: 1958 2011
Included observations: 54

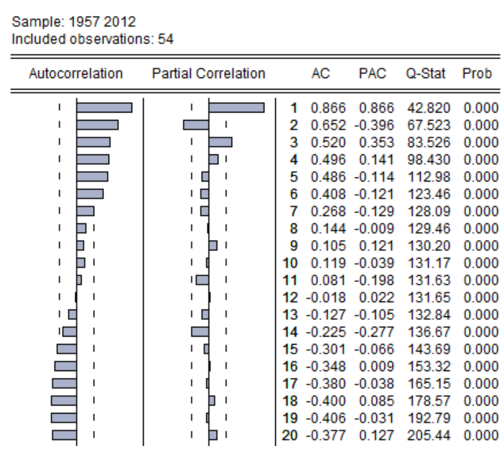
	AC	PAC	Q-Stat	Prob
1	0.807	0.807	37.192	0.000
2	0.547	-0.302	54.565	0.000
3	0.414	0.264	64.735	0.000
4	0.391	0.088	73.958	0.000
5	0.398	0.082	83.755	0.000
6	0.327	-0.174	90.482	0.000
7	0.194	-0.063	92.909	0.000
8	0.076	-0.075	93.287	0.000
9	0.028	0.030	93.339	0.000
10	0.034	-0.001	93.416	0.000
11	-0.021	-0.195	93.447	0.000
12	-0.130	-0.056	94.673	0.000
13	-0.205	0.018	97.762	0.000
14	-0.254	-0.155	102.64	0.000
15	-0.288	-0.075	109.09	0.000
16	-0.306	-0.004	116.56	0.000
17	-0.338	-0.078	125.90	0.000
18	-0.366	-0.024	137.15	0.000
19	-0.354	0.037	147.96	0.000
20	-0.317	-0.032	156.88	0.000

(b) ACF and PACF

Figure 16: General Inflation



(a) Time Series Plot



Sample: 1957 2012
Included observations: 54

	AC	PAC	Q-Stat	Prob
1	0.866	0.866	42.820	0.000
2	0.652	-0.396	67.523	0.000
3	0.520	0.353	83.526	0.000
4	0.496	0.141	98.430	0.000
5	0.486	-0.114	112.98	0.000
6	0.408	-0.121	123.46	0.000
7	0.268	-0.129	128.09	0.000
8	0.144	-0.009	129.46	0.000
9	0.105	0.121	130.20	0.000
10	0.119	-0.039	131.17	0.000
11	0.081	-0.198	131.63	0.000
12	-0.018	0.022	131.65	0.000
13	-0.127	-0.105	132.84	0.000
14	-0.225	-0.277	136.67	0.000
15	-0.301	-0.066	143.69	0.000
16	-0.348	0.009	153.32	0.000
17	-0.380	-0.038	165.15	0.000
18	-0.400	0.085	178.57	0.000
19	-0.406	-0.031	192.79	0.000
20	-0.377	0.127	205.44	0.000

(b) ACF and PACF

Figure 17: General Inflation excluding Food and Gas

Dependent Variable: CPIINFL				
Method: Least Squares				
Sample (adjusted): 1961 2009				
Included observations: 49 after adjustments				
Convergence achieved after 4 iterations				
Newey-West HAC Standard Errors & Covariance (lag truncation=3)				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	3.718094	1.379958	2.694352	0.0099
AR(1)	1.268483	0.161848	7.83748	0.0000
AR(2)	-0.77731	0.24537	-3.16792	0.0028
AR(3)	0.368021	0.160573	2.29192	0.0266
R-squared	0.729847	Mean dependent var		4.163009
Adjusted R-squared	0.711837	S.D. dependent var		2.876272
S.E. of regression	1.544006	Akaike info criterion		3.784746
Sum squared resid	107.278	Schwarz criterion		3.93918
Log likelihood	-88.7263	F-statistic		40.52413
Durbin-Watson stat	1.705549	Prob(F-statistic)		0.000000
Inverted AR Roots	0.86	.20+.62i	.20-.62i	

Table 5: Estimation Results of AR(3) Model for General Inflation

Dependent Variable: CPIINFL_LESS				
Method: Least Squares				
Sample (adjusted): 1961 2009				
Included observations: 49 after adjustments				
Convergence achieved after 3 iterations				
Newey-West HAC Standard Errors & Covariance (lag truncation=3)				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	4.162125	1.376175	3.024416	0.0041
AR(1)	1.365794	0.163088	8.374595	0.0000
AR(2)	-0.86294	0.276598	-3.11984	0.0032
AR(3)	0.382308	0.202927	1.883972	0.066
R-squared	0.823883	Mean dependent var		4.131566
Adjusted R-squared	0.812142	S.D. dependent var		2.595923
S.E. of regression	1.125141	Akaike info criterion		3.151801
Sum squared resid	56.96738	Schwarz criterion		3.306235
Log likelihood	-73.2191	F-statistic		70.17065
Durbin-Watson stat	2.110947	Prob(F-statistic)		0.000000
Inverted AR Roots	0.88	.24+.61i	.24-.61i	

Table 6: Estimation Results of AR(3) Model for General Inflation excluding Food and Gas

Based on the estimated AR(3) models, we calculate the multistep density forecast for both series under the assumption that $\{\varepsilon_t\}$ is normally distributed.

General Inflation

Forecasting Horizon $h = 1$

- optimal point forecast $f_{t,1}$

$$\begin{aligned} f_{t,1} &= \mu_{t+1|t} = E(Y_{t+1}|I_t) = E(c + \phi_1 Y_t + \phi_2 Y_{t-1} + \phi_3 Y_{t-2} + \varepsilon_{t+1}|I_t) \\ &= c + \phi_1 Y_t + \phi_2 Y_{t-1} + \phi_3 Y_{t-2} \end{aligned}$$

  Therefore, the optimal forecast of general inflation in 2010 is,

$$\begin{aligned} f_{2009,1} &= \hat{\mu}_{2010|2009} = \hat{c} + \hat{\phi}_1 Y_{2009} + \hat{\phi}_2 Y_{2008} + \hat{\phi}_3 Y_{2007} \\ &\approx 0.52 + 1.27 \times (-0.36) - 0.78 \times 3.83 + 0.37 \times 2.85 \approx -1.87 \end{aligned}$$

  Observe that the inflation in 2009 was negative (-0.36%, this is the only period with a small but negative rate), and given the high persistence of the process, the forecast for 2010 is forced into negative territory.

- 1-period-ahead forecast error $e_{2008,1}$

$$e_{2009,1} = Y_{2010} - f_{2009,1} = \varepsilon_{2010} \approx 1.64 - (-1.87) = 3.51$$

- The uncertainty associated with the forecast

$$\sigma^2_{2010|2009} = \text{var}(e_{2009,1}) = E(\varepsilon^2_{2010}) = \sigma^2_\varepsilon \approx 1.54^2 = 2.37$$

- The one step ahead density forecast

$$f(Y_{t+1}|I_t) \to N(\mu_{t+1|t}, \sigma^2_{t+1|t}) = N(-1.87, 1.54^2)$$

Forecasting Horizon $h = 2$

- optimal point forecast $f_{t,2}$

$$\begin{aligned} f_{t,2} &= \mu_{t+2|t} = E(Y_{t+2}|I_t) = E(c + \phi_1 Y_{t+1} + \phi_2 Y_t + \phi_3 Y_{t-1} + \varepsilon_{t+2}|I_t) \\ &= c + \phi_1 f_{t,1} + \phi_2 Y_t + \phi_3 Y_{t-1} \end{aligned}$$

  Therefore, the optimal forecast of general inflation in 2011 is,

$$\begin{aligned} f_{2009,2} &= \hat{\mu}_{2011|2009} = \hat{c} + \hat{\phi}_1 f_{2009,1} + \hat{\phi}_2 Y_{2009} + \hat{\phi}_3 Y_{2008} \\ &\approx 0.52 + 1.27 \times (-1.87) - 0.78 \times (-0.36) + 0.37 \times 3.83 \approx -0.15 \end{aligned}$$

- 2-period-ahead forecast error $e_{2009,2}$

$$e_{2009,2} = Y_{2011} - f_{2009,2} = \phi_1 \varepsilon_{2010} + \varepsilon_{2011} \approx 3.16 - (-0.15) = 3.31$$

- The uncertainty associated with the forecast

$$\begin{aligned} \sigma^2_{2011|2009} &= \text{var}(e_{2009,2}) = E(\phi_1 \varepsilon_{2010} + \varepsilon_{2011})^2 \\ &= \sigma^2_\varepsilon (1 + \phi_1^2) \approx 2.71^2 = 7.34 \end{aligned}$$

- The two step ahead density forecast

$$f(Y_{t+2}|I_t) \to N(\mu_{t+2|t}, \sigma^2_{t+2|t}) = N(-0.15, 2.71^2)$$

General Inflation excluding food and gas

Forecasting Horizon $h = 1$

- optimal point forecast $f_{t,1}$

$$\begin{aligned} f_{t,1} &= \mu_{t+1|t} = E(Y_{t+1}|I_t) = E(c + \phi_1 Y_t + \phi_2 Y_{t-1} + \phi_3 Y_{t-2} + \varepsilon_{t+1}|I_t) \\ &= c + \phi_1 Y_t + \phi_2 Y_{t-1} + \phi_3 Y_{t-2} \end{aligned}$$

Therefore, the optimal forecast of general inflation excluding food and gas in 2010 is,

$$\begin{aligned} f_{2009,1} &= \hat{\mu}_{2010|2009} = \hat{c} + \hat{\phi}_1 Y_{2009} + \hat{\phi}_2 Y_{2008} + \hat{\phi}_3 Y_{2007} \\ &\approx 0.46 + 1.37 \times (1.70) - 0.86 \times 2.30 + 0.38 \times 2.35 \approx 1.71 \end{aligned}$$

- 1-period-ahead forecast error $e_{2008,1}$

$$e_{2009,1} = Y_{2010} - f_{2009,1} = \varepsilon_{2010} \approx 0.96 - (1.71) = -0.75$$

- The uncertainty associated with the forecast

$$\sigma^2_{2010|2009} = \text{var}(e_{2009,1}) = E(\varepsilon^2_{2010}) = \sigma^2_\varepsilon \approx 1.13^2 = 1.28$$

- The one step ahead density forecast

$$f(Y_{t+1}|I_t) \to N(\mu_{t+1|t}, \sigma^2_{t+1|t}) = N(1.71, 1.13^2)$$

Forecasting Horizon $h = 2$

- optimal point forecast $f_{t,2}$

$$\begin{aligned} f_{t,2} &= \mu_{t+2|t} = E(Y_{t+2}|I_t) = E(c + \phi_1 Y_{t+1} + \phi_2 Y_t + \phi_3 Y_{t-1} + \varepsilon_{t+2}|I_t) \\ &= c + \phi_1 f_{t,1} + \phi_2 Y_t + \phi_3 Y_{t-1} \end{aligned}$$

Therefore, the optimal forecast of general inflation excluding food and gas in 2011 is,

$$\begin{aligned} f_{2009,2} = \hat{\mu}_{2011|2009} &= \hat{c} + \hat{\phi}_1 f_{2009,1} + \hat{\phi}_2 Y_{2009} + \hat{\phi}_3 Y_{2008} \\ &\approx 0.46 + 1.37 \times (1.71) - 0.86 \times (1.70) + 0.38 \times 2.30 \approx 2.23 \end{aligned}$$

- 2-period-ahead forecast error $e_{2009,2}$

$$e_{2009,2} = Y_{2011} - f_{2009,2} = \phi_1 \varepsilon_{2010} + \varepsilon_{2011} \approx 1.66 - 2.23 = -0.57$$

- The uncertainty associated with the forecast

$$\begin{aligned} \sigma^2_{2011|2009} &= \text{var}(e_{2009,2}) = E(\phi_1 \varepsilon_{2010} + \varepsilon_{2011})^2 \\ &= \sigma^2_\varepsilon (1 + \phi_1^2) \approx 1.93^2 = 3.67 \end{aligned}$$

- The two step ahead density forecast

$$f(Y_{t+2}|I_t) \to N(\mu_{t+2|t}, \sigma^2_{t+2|t}) = N(2.23, 1.93^2)$$

In Figures 18 and 19, we compare the multistep forecasts with the actual values for the general inflation and for inflation excluding food and gas. The uncertainty of the forecast for the general inflation is larger than that of inflation excluding food and gas. The forecast errors calculated for general inflation (2.37 and 7.34) are larger than those for the general inflation excluding food and gas (1.28 and 3.67). In this sense, the forecast of general inflation is less accurate and it is more difficult to predict. Note that the forecast is negative, due to the persistence of the process and the initial negative value in the information set, but is moving into the right direction. Predicting general inflation is a more difficult task because gas inflation behaves as a white noise process (see the analysis in Exercise 7) and it is very volatile. The inclusion of gas inflation introduces additional unpredictability into the time series of general inflation, so that by removing this component we obtain a more predictable series.

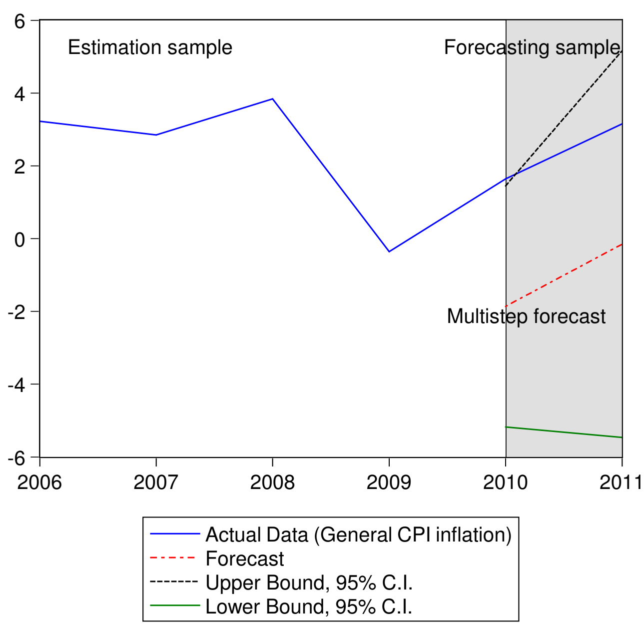


Figure 18: General Inflation, Multistep Forecast

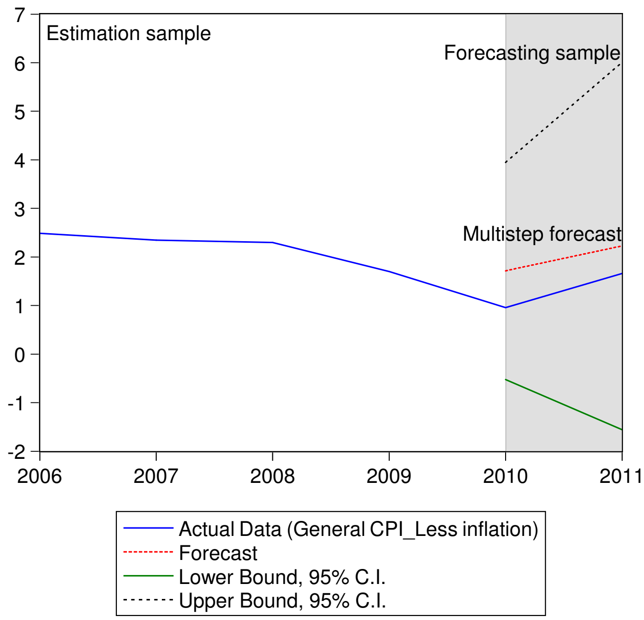
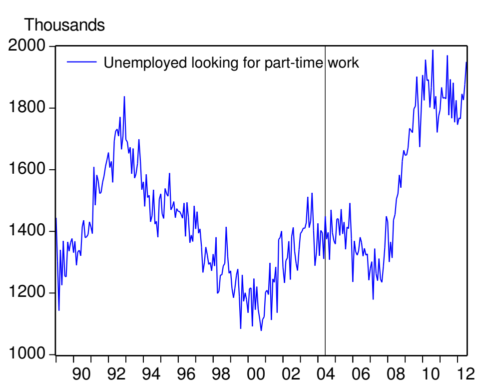


Figure 19: General Inflation excluding Food and Gas, Multistep Forecast

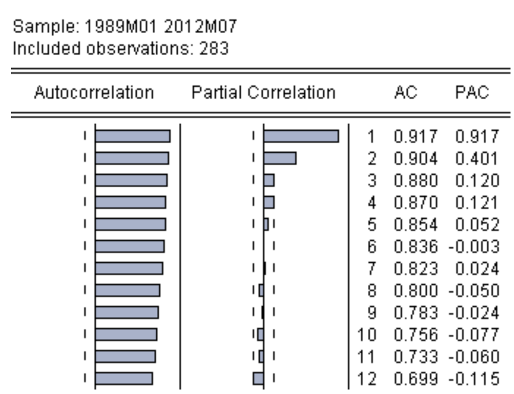
Exercise 9
We download the data of U.S. unemployed looking for part-time work from the Bureau of Labor Statistics (BLS) website:

`http://www.bls.gov/data/#unemployment`

The updated time series runs from January 1989 to July 2012. In Figure 20a, we plot the data, and the vertical line on June 2004 separates the sample in the textbook from the updated sample. The financial crisis of 2008 produced a steep uprise in the number of people looking for part-time work, surpassing the numbers in the recession of the early 90s.



(a) Time Series Plot



(b) ACF and PACF

Figure 20: Number of Unemployed People Looking for Part-Time Work

The ACF and PACF plots in Figure 20b show that the underlying process is an AR(4), quite similar to that in Figure 7.13 of the textbook. The ACF shows strong dependence slowly decaying towards zero and the PACF shows four significant spikes. In Table 7 we report the estimation results. Observe that one of the inverted roots is very close to one, which raises some doubts about the stationarity of the process. Nevertheless, we will proceed and consider that the estimated AR(4) is stationary but with very high persistence. When estimating the model, we exclude the last three observations (from May 2012 to July 2012) to assess the multistep forecast.

Dependent Variable: UNEMPL_PART				
Method: Least Squares				
Sample (adjusted): 1989M05 2012M04				
Included observations: 276 after adjustments				
Convergence achieved after 4 iterations				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1651.207	307.0044	5.378448	0.0000
AR(1)	0.401936	0.059915	6.708476	0.0000
AR(2)	0.327055	0.062948	5.195642	0.0000
AR(3)	0.085621	0.063558	1.347128	0.1791
AR(4)	0.166613	0.058573	2.844525	0.0048
R-squared	0.89442	Mean dependent var		1453.569
Adjusted R-squared	0.892862	S.D. dependent var		201.8053
S.E. of regression	66.05493	Akaike info criterion		11.2368
Sum squared resid	1182442	Schwarz criterion		11.30239
Log likelihood	-1545.678	F-statistic		573.9431
Durbin-Watson stat	2.008814	Prob(F-statistic)		0.000000
Inverted AR Roots	0.99	.01+.52i	.01-.52i	-0.62

Table 7: Estimation Results of AR(4) Model

Based on the estimated AR(4) model and assuming a quadratic loss function, we calculate the optimal forecast as follows,

FORECASTING HORIZON $h = 1$, MAY 2012

- optimal point forecast $f_{t,1}$

$$\begin{aligned} f_{t,1} &= \mu_{t+1|t} = E(Y_{t+1}|I_t) = E(c + \phi_1 Y_t + \phi_2 Y_{t-1} + \phi_3 Y_{t-2} + \phi_4 Y_{t-3} + \varepsilon_{t+1}|I_t) \\ &= c + \phi_1 Y_t + \phi_2 Y_{t-1} + \phi_3 Y_{t-2} + \phi_4 Y_{t-3}. \end{aligned}$$

Therefore, the optimal forecast of the unemployed looking for part-time work in May 2012 is,

$$\begin{aligned} f_{t,1} &= \hat{\mu}_{t+1|t} = \hat{c} + \hat{\phi}_1 Y_t + \hat{\phi}_2 Y_{t-1} + \hat{\phi}_3 Y_{t-2} + \hat{\phi}_4 Y_{t-3} \\ &\approx 31.00 + 0.402 \times 1846 + 0.327 \times 1765 + 0.086 \times 1768 + 0.167 \times 1746 \\ &\approx 1793.88 \end{aligned}$$

- 1-period-ahead forecast error $e_{t,1}$

$$e_{t,1} = Y_{t+1} - f_{t,1} = \varepsilon_{t+1} \approx 1826 - 1793.88 = 32.12$$

- The uncertainty associated with the forecast

$$\sigma^2_{t+1|t} = \mathrm{var}(e_{t,1}) = E(\varepsilon^2_{t+1}) = \sigma^2_\varepsilon \approx 4363.26$$

FORECASTING HORIZON $h = 2$, JUNE 2012

- optimal point forecast $f_{t,2}$

$$\begin{aligned} f_{t,2} &= \mu_{t+2|t} = E(Y_{t+2}|I_t) = E(c + \phi_1 Y_{t+1} + \phi_2 Y_t + \phi_3 Y_{t-1} + \phi_4 Y_{t-2} + \varepsilon_{t+2}|I_t) \\ &= c + \phi_1 f_{t,1} + \phi_2 Y_t + \phi_3 Y_{t-1} + \phi_4 Y_{t-2}. \end{aligned}$$

  Therefore, the optimal forecast of the unemployed looking for part-time work in June 2012 is,

$$\begin{aligned} f_{t,2} &= \hat{\mu}_{t+2|t} = \hat{c} + \hat{\phi}_1 f_{t,1} + \hat{\phi}_2 Y_t + \hat{\phi}_3 Y_{t-1} + \hat{\phi}_4 Y_{t-2} \\ &\approx 31.00 + 0.402 \times 1793.88 + 0.327 \times 1846 + 0.086 \times 1765 + 0.167 \times 1768 \\ &\approx 1802.83 \end{aligned}$$

- 1-period-ahead forecast error $e_{t,2}$

$$e_{t,2} = Y_{t+2} - f_{t,2} = \phi\varepsilon_{t+1} + \varepsilon_{t+2} \approx 1877 - 1802.83 = 74.17$$

- The uncertainty associated with the forecast

$$\begin{aligned} \sigma^2_{t+2|t} &= \operatorname{var}(e_{t,2}) = E(\phi_1 \varepsilon_{t+1} + \varepsilon_{t+2})^2 \\ &= \sigma^2_\varepsilon(1 + \phi_1^2) \approx 5068.38 \end{aligned}$$

FORECASTING HORIZON $h = 3$, JULY 2012

- optimal point forecast $f_{t,3}$

$$\begin{aligned} f_{t,3} &= \mu_{t+3|t} = E(Y_{t+3}|I_t) = E(c + \phi_1 Y_{t+2} + \phi_2 Y_{t+1} + \phi_3 Y_t + \phi_4 Y_{t-1} + \varepsilon_{t+3}|I_t) \\ &= c + \phi_1 f_{t,2} + \phi_2 f_{t,1} + \phi_3 Y_t + \phi_4 Y_{t-1}. \end{aligned}$$

  Therefore, the optimal forecast of the unemployed looking for part-time work in July 2012 is,

$$\begin{aligned} f_{t,3} &= \hat{\mu}_{t+3|t} = \hat{c} + \hat{\phi}_1 f_{t,2} + \hat{\phi}_2 f_{t,1} + \hat{\phi}_3 Y_t + \hat{\phi}_4 Y_{t-1} \\ &\approx 31.00 + 0.402 \times 1802.83 + 0.327 \times 1793.88 + 0.086 \times 1846 + 0.167 \times 1765 \\ &\approx 1795.85 \end{aligned}$$

- 1-period-ahead forecast error $e_{t,3}$

$$e_{t,3} = Y_{t+3} - f_{t,3} = (\phi_1^2 + \phi_2)\varepsilon_{t+1} + \phi_1\varepsilon_{t+2} + \varepsilon_{t+3} \approx 1950 - 1795.85 = 154.15$$

- The uncertainty associated with the forecast

$$\begin{aligned} \sigma^2_{t+3|t} &= \operatorname{var}(e_{t,3}) = E((\phi_1^2 + \phi_2)\varepsilon_{t+1} + \phi_1\varepsilon_{t+2} + \varepsilon_{t+3})^2 \\ &= \sigma^2_\varepsilon((\phi_1^2 + \phi_2)^2 + \phi_1^2 + 1) \approx 6110.04 \end{aligned}$$

We plot the multistep forecast in Figure 21. The forecast is very smooth, which is a reflection of the high persistence of the process. It is well contained within the 95% bands.

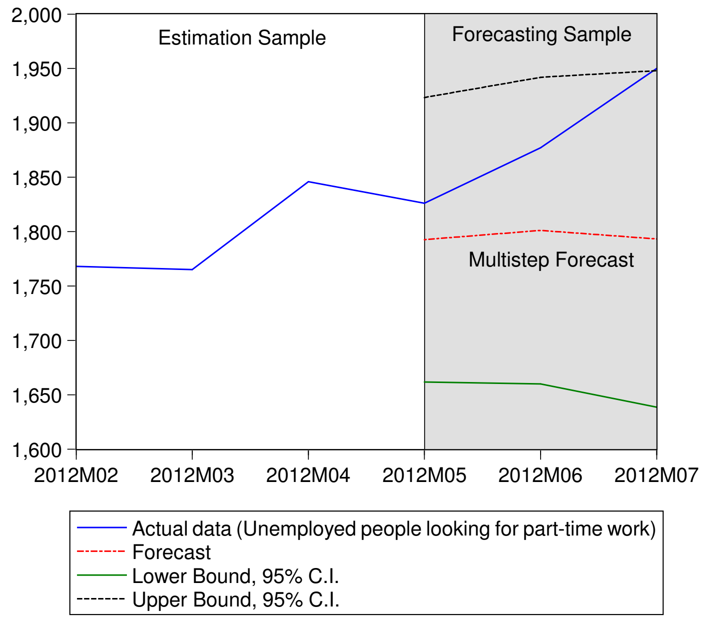


Figure 21: Multistep Forecast

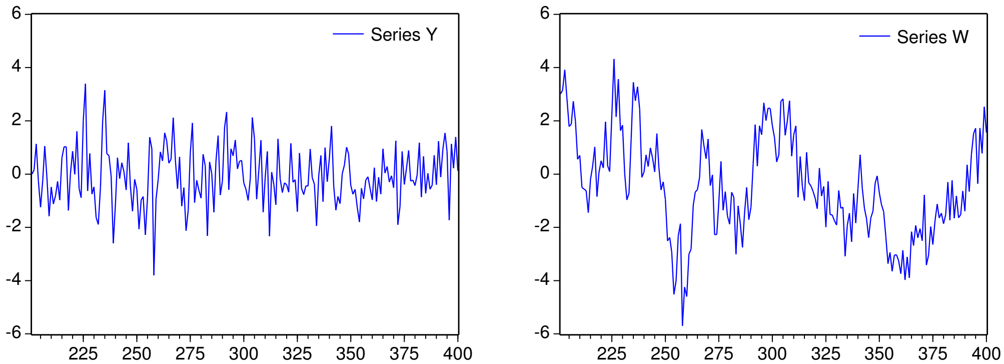
Exercise 10
We simulate 2000 observations of the following two AR(3) processes

$$\begin{aligned} y_t &= 0 + 0.1y_{t-1} - 0.1y_{t-2} + 0.1y_{t-3} + \varepsilon_t \\ w_t &= 0 + 0.5w_{t-1} + 0.3w_{t-2} + 0.1w_{t-3} + \varepsilon_t \end{aligned}$$

and we plot 200 observations in Figure 22. Both series y_t and w_t move around zero, which is their unconditional mean by construction. The series w_t has stronger persistence than the series y_t. Observe that the sum of the three autoregressive parameters is closer to one for w_t, i.e., $\phi_1 + \phi_2 + \phi_3 = 0.9$, than that for y_t, i.e. $\phi_1 + \phi_2 + \phi_3 = 0.1$. The autocorrelation functions in Figure 23 deliver the same message, i.e., w_t has stronger dependence than y_t. Both series have three significant partial autocorrelations, as expected, because both are AR(3) processes by construction but the values for the series y_t are very small compared to those corresponding to w_t.

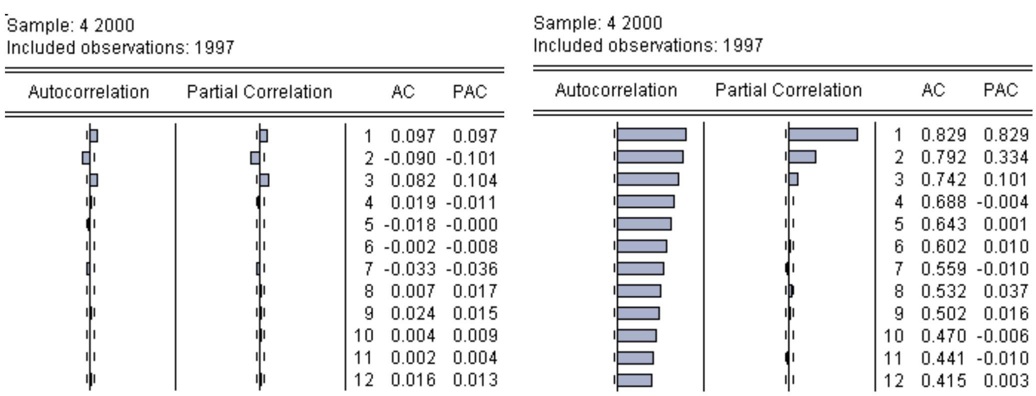
The code to simulate these processes in Eviews is

```
series e=nrnd
smpl @first @first+2
series y=0
series w=0
smpl @first+3 @last
series y=0+0.1*y(-1)-0.1*y(-2)+0.1*y(-3)+e
series w=0+0.5*w(-1)+0.3*w(-2)+0.1*w(-3)+e
```

(a) Time Series Y (b) Time Series W

Figure 22: Time Series Plots



Sample: 4 2000
Included observations: 1997

Autocorrelation	Partial Correlation		AC	PAC
		1	0.097	0.097
		2	-0.090	-0.101
		3	0.082	0.104
		4	0.019	-0.011
		5	-0.018	-0.000
		6	-0.002	-0.008
		7	-0.033	-0.036
		8	0.007	0.017
		9	0.024	0.015
		10	0.004	0.009
		11	0.002	0.004
		12	0.016	0.013

Sample: 4 2000
Included observations: 1997

Autocorrelation	Partial Correlation		AC	PAC
		1	0.829	0.829
		2	0.792	0.334
		3	0.742	0.101
		4	0.688	-0.004
		5	0.643	0.001
		6	0.602	0.010
		7	0.559	-0.010
		8	0.532	0.037
		9	0.502	0.016
		10	0.470	-0.006
		11	0.441	-0.010
		12	0.415	0.003

(a) ACF and PACF of Series Y (b) ACF and PACF of Series W

Figure 23: ACF and PACF