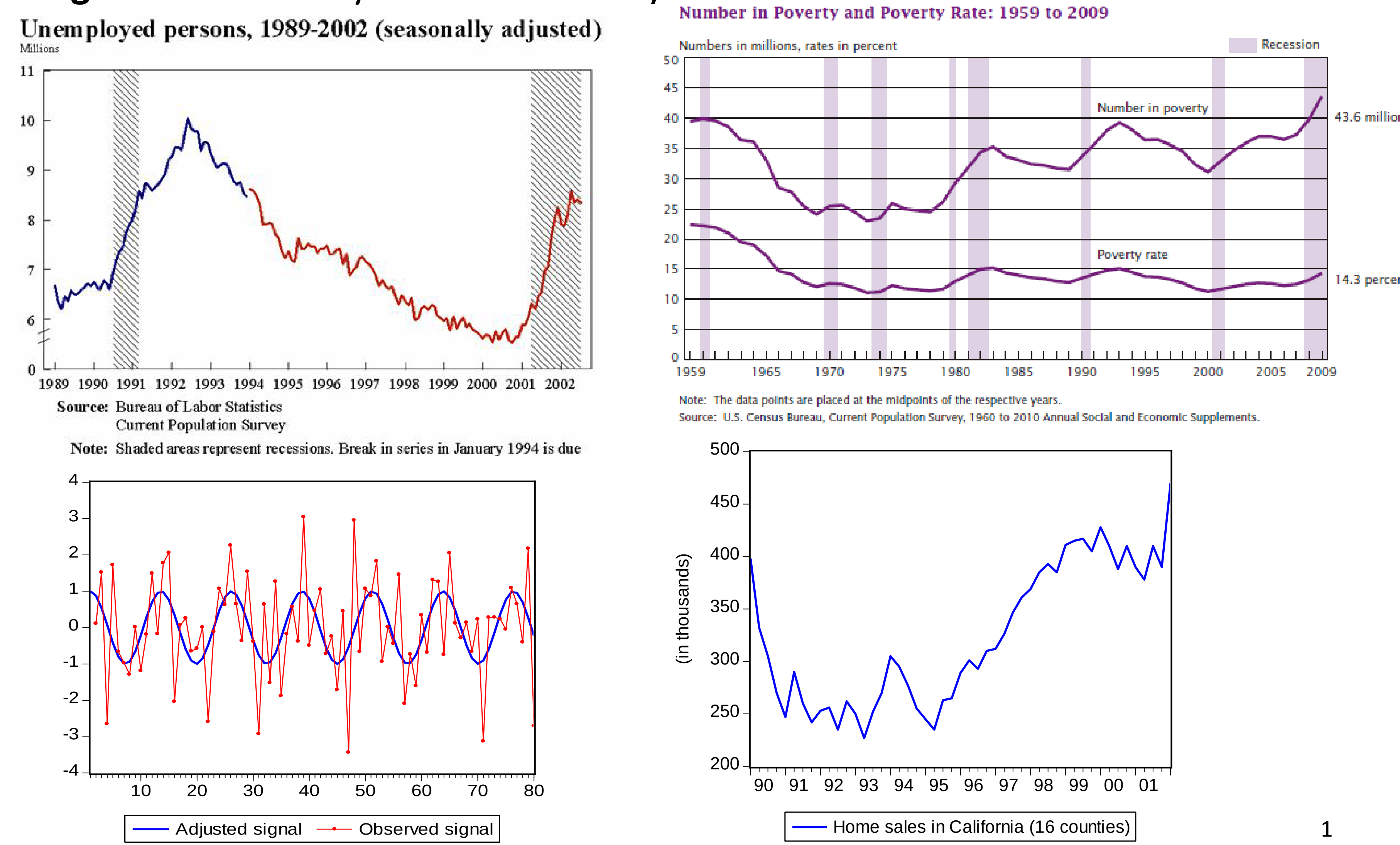
CHAPTER 7
FORECASTING WITH AUTOREGRESSIVE (AR) MODELS

Figure 7.1 A Variety of Time Series Cycles



7.1 Cycles

A cycle is a time series pattern of periodic fluctuations.

cycle → deterministic

cycle → stochastic

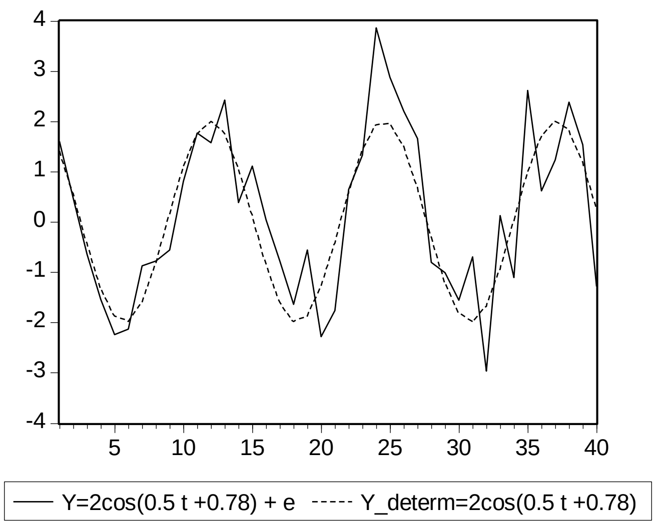


Figure 7.2 Deterministic Cycle

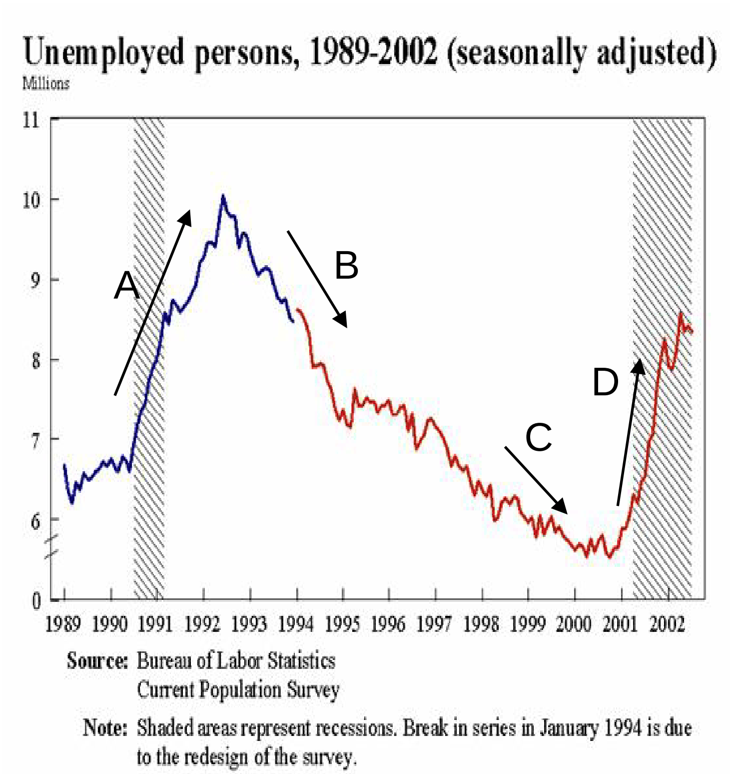


Figure 7.3 Unemployed Persons, 1989-2002 (Seasonally Adjusted)

An autoregressive model of order $p \geq 0$, referred as AR(p), has the following functional form

$$Y_t = c + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + \varepsilon_t$$

where ε_t is a white noise process.

A process is **covariance stationary (causal)** if it can be written as a linear function of past shocks:

$$X_t = \varepsilon_t + \varphi_1 \varepsilon_{t-1} + \varphi_2 \varepsilon_{t-2} + \varphi_3 \varepsilon_{t-3} + \ldots$$

This happens *iif* all the roots ξ_i of the $\varphi(L)$ polynomial are outside the unit circle:

$$|\xi_i| > 1,$$

i.e., *iff a*ll the modules of the inverse roots are smaller than 1:

$$|1/\xi_i| < 1$$

(if $1/\xi = a + b\,i$, where $i = \sqrt{-1}$, $\sqrt{(a^2 + b^2)} < 1$)

NB: An AR(*p*) is always *invertible.* A MA(*q*) is always *stationary.*

7.2.1 The AR(1) Process

$$Y_t = c + \phi Y_{t-1} + \varepsilon_t$$

$$E[Y_t] = \frac{c}{1-\phi}$$

$$\sigma_Y{}^2 = Var(Y_t) = \frac{\sigma_\varepsilon{}^2}{1-\phi^2}$$

$$\rho_k = \phi^k$$

$$r_1 = \rho_k,$$

$$r_k = 0, k \geq 2$$

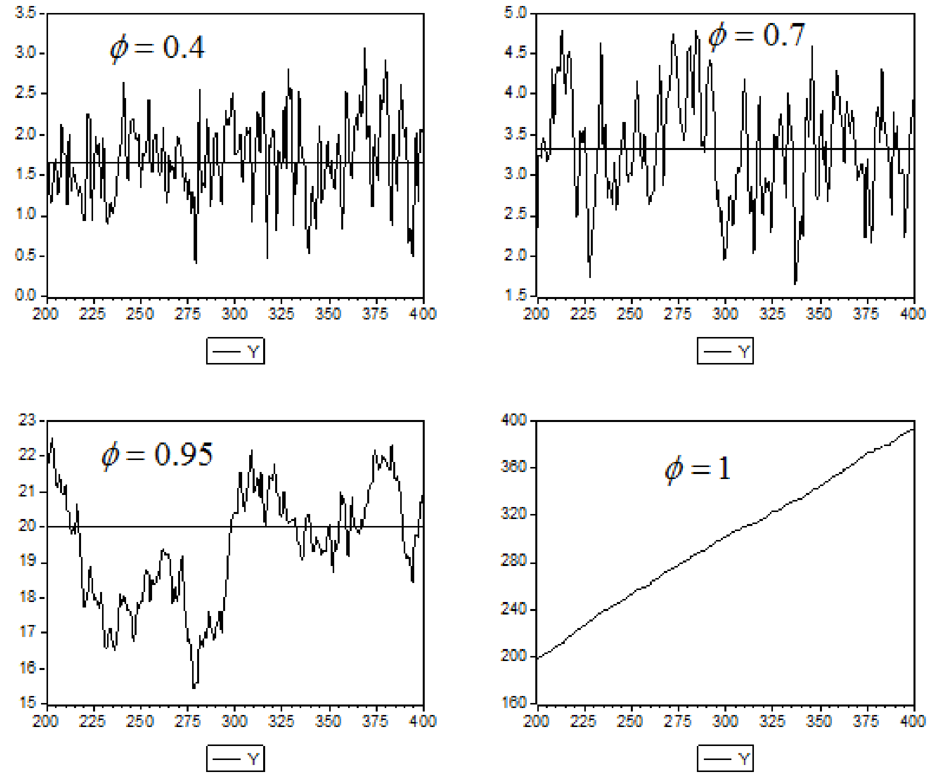
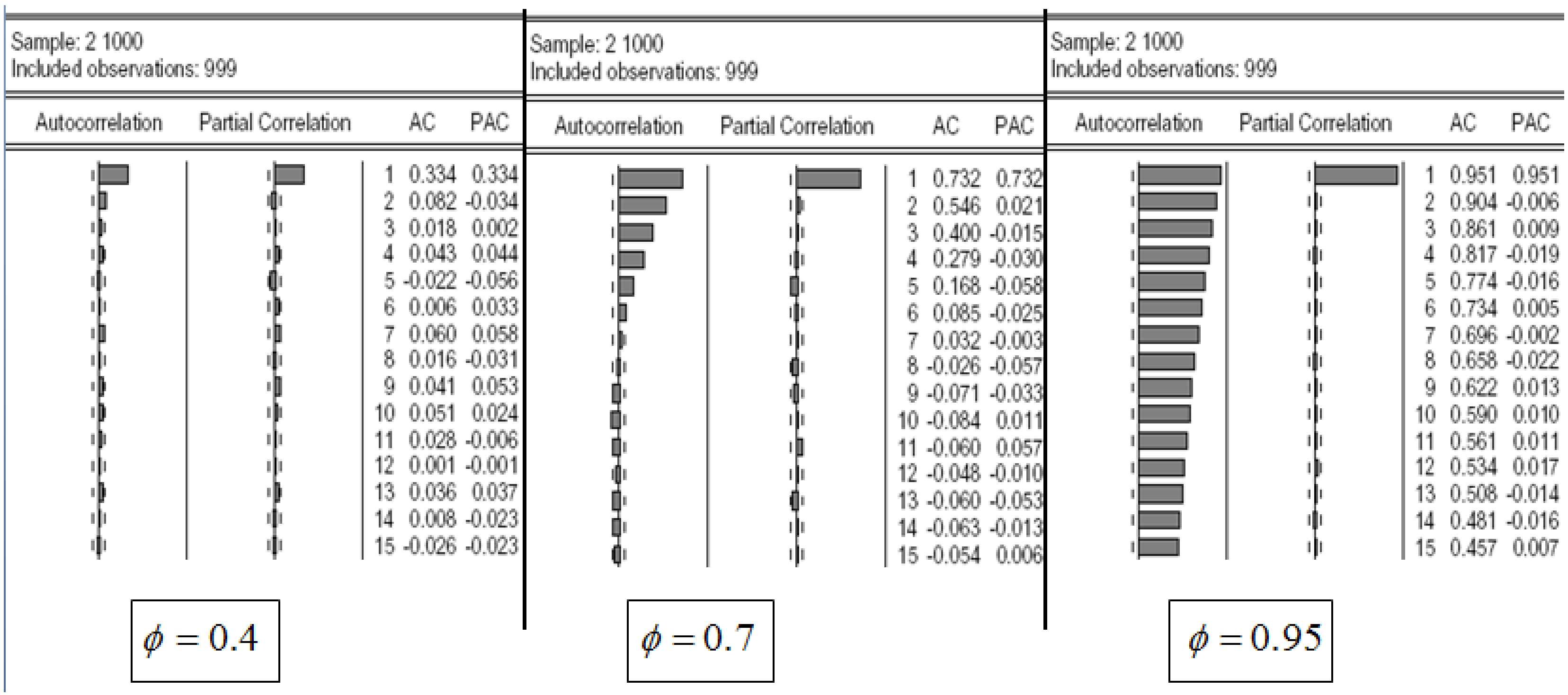


Figure 7.4 Autoregressive Processes AR(1)

Figure 7.5 Autocorrelation Functions of Covariance-Stationary AR(1) Processes



$\phi = 0.4$

	AC	PAC
1	0.334	0.334
2	0.082	-0.034
3	0.018	0.002
4	0.043	0.044
5	-0.022	-0.056
6	0.006	0.033
7	0.060	0.058
8	0.016	-0.031
9	0.041	0.053
10	0.051	0.024
11	0.028	-0.006
12	0.001	-0.001
13	0.036	0.037
14	0.008	-0.023
15	-0.026	-0.023

$\phi = 0.7$

	AC	PAC
1	0.732	0.732
2	0.546	0.021
3	0.400	-0.015
4	0.279	-0.030
5	0.168	-0.058
6	0.085	-0.025
7	0.032	-0.003
8	-0.026	-0.057
9	-0.071	-0.033
10	-0.084	0.011
11	-0.060	0.057
12	-0.048	-0.010
13	-0.060	-0.053
14	-0.063	-0.013
15	-0.054	0.006

$\phi = 0.95$

	AC	PAC
1	0.951	0.951
2	0.904	-0.006
3	0.861	0.009
4	0.817	-0.019
5	0.774	-0.016
6	0.734	0.005
7	0.696	-0.002
8	0.658	-0.022
9	0.622	0.013
10	0.590	0.010
11	0.561	0.011
12	0.534	0.017
13	0.508	-0.014
14	0.481	-0.016
15	0.457	0.007

A necessary and sufficient condition for an AR(1) process $Y_t = c + \phi Y_{t-1} + \varepsilon_t$ to be covariance stationary is that $|\phi| < 1$.

Figure 7.6 Time Series Plot and Autocorrelation Functions of AR(1) with Negative Parameter

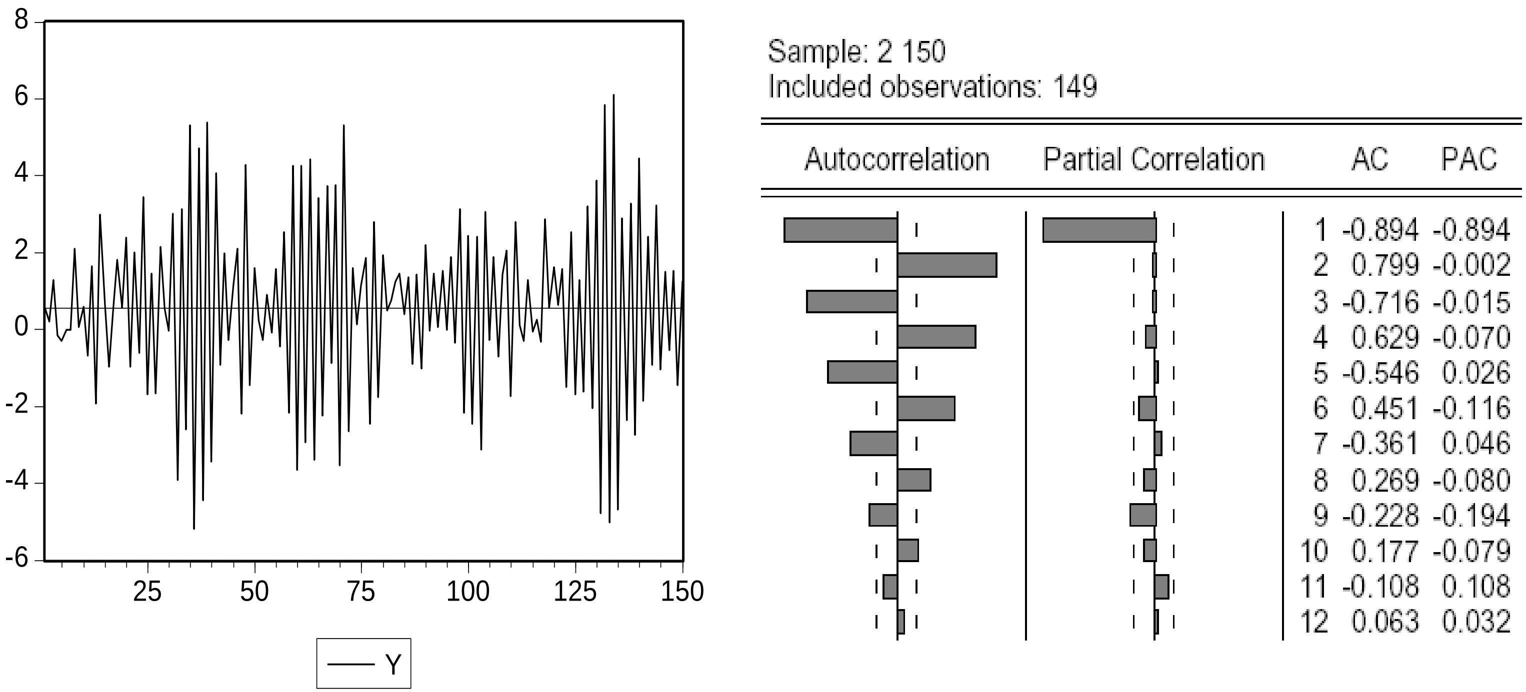
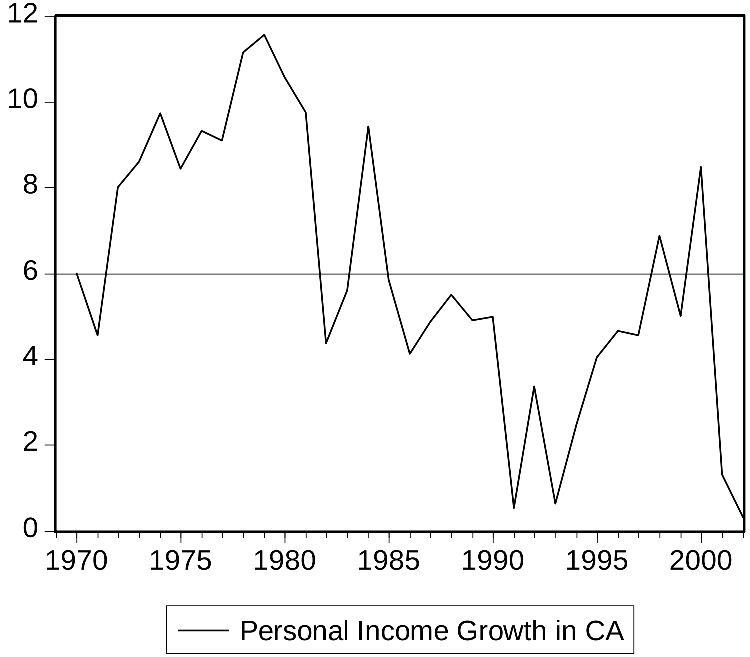
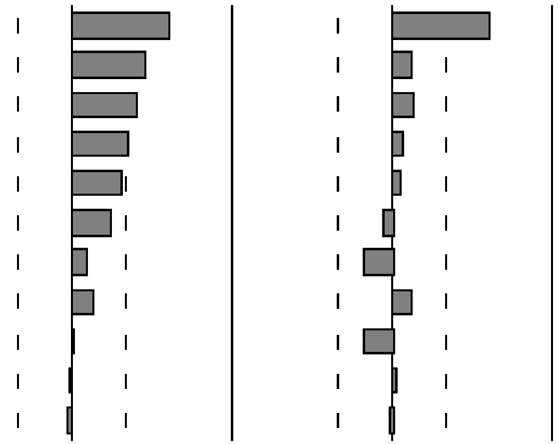


Figure 7.7 Per Capita Income Growth (California, 1969-2002)



Sample: 1969 2002
Included observations: 33



Autocorrelation	Partial Correlation		AC	PAC
		1	0.629	0.629
		2	0.471	0.125
		3	0.417	0.134
		4	0.365	0.059
		5	0.327	0.051
		6	0.247	-0.050
		7	0.098	-0.180
		8	0.135	0.126
		9	0.024	-0.179
		10	-0.009	0.021
		11	-0.021	-0.006

7.2.2 AR(2) Process

$$Y_t = c + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \varepsilon_t$$

$$E[Y_t] = \frac{c}{1 - \phi_1 - \phi_2} \qquad Var(Y_t) = \frac{\sigma_\varepsilon^{\ 2}}{1 - \phi_1\rho_1 - \phi_2\rho_2}$$

$Y_t = 1 + Y_{t-1} - 0.5Y_{t-2} + \varepsilon_t$ $\quad$ $\phi_1 + \phi_2 = 0.5$

$Y_t = 1 - 0.5Y_{t-1} + 0.4Y_{t-2} + \varepsilon_t$ $\quad$ $\phi_1 + \phi_2 = -0.1$

$Y_t = 1 + 0.5Y_{t-1} + 0.3Y_{t-2} + \varepsilon_t$ $\quad$ $\phi_1 + \phi_2 = 0.8$

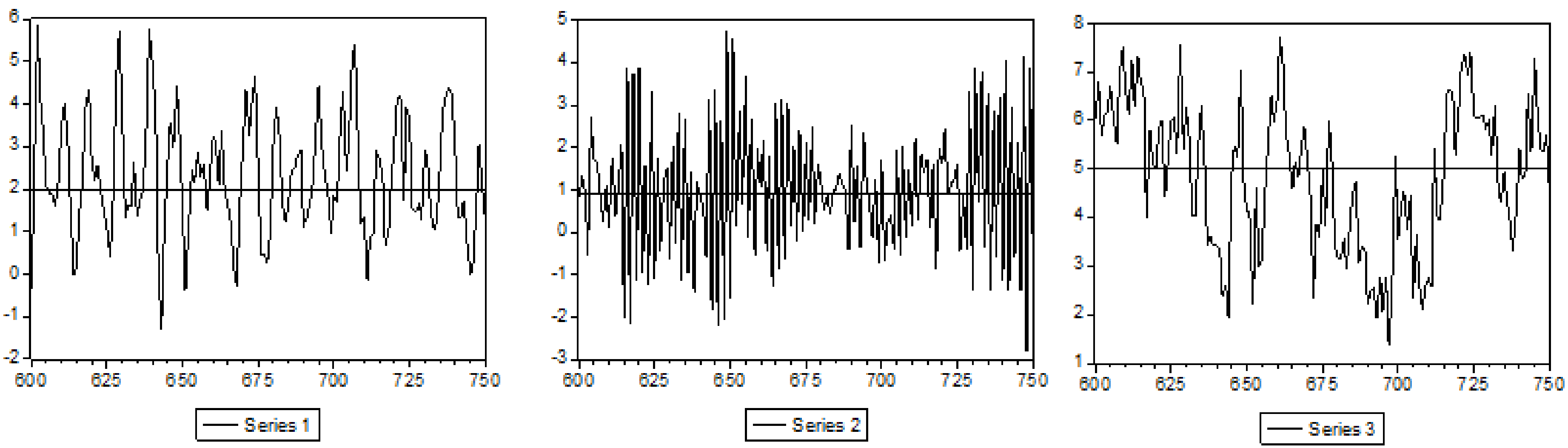


Figure 7.8 Autoregressive Processes AR(2)

The necessary conditions for an AR(2) process to be covariance stationary are

$$-1 < \phi_2 < 1$$

$$-2 < \phi_1 < 2$$

and the sufficient conditions are

$$\phi_1 + \phi_2 < 1$$

$$\phi_2 - \phi_1 < 1$$

$$Y_t = 1 + 0.5Y_{t-1} + 0.5Y_{t-2} + \varepsilon_t$$

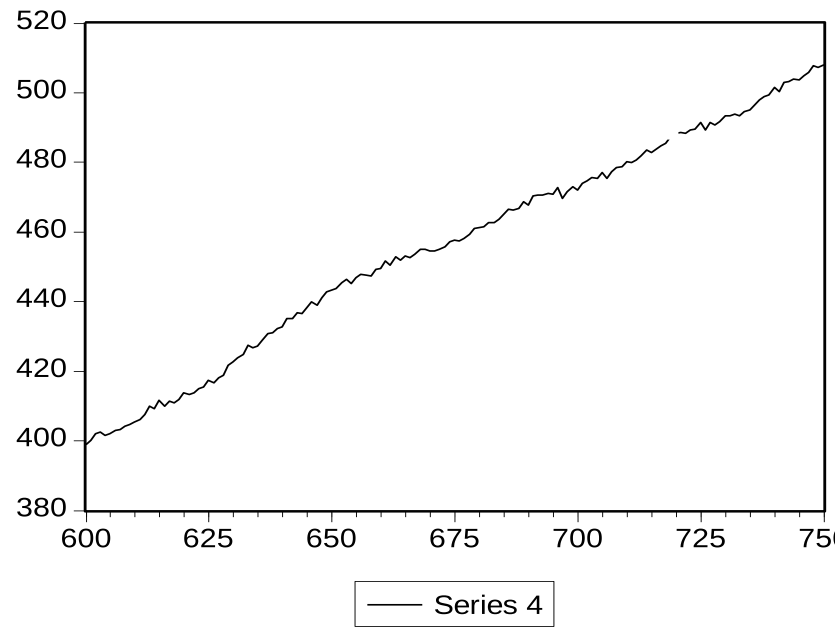
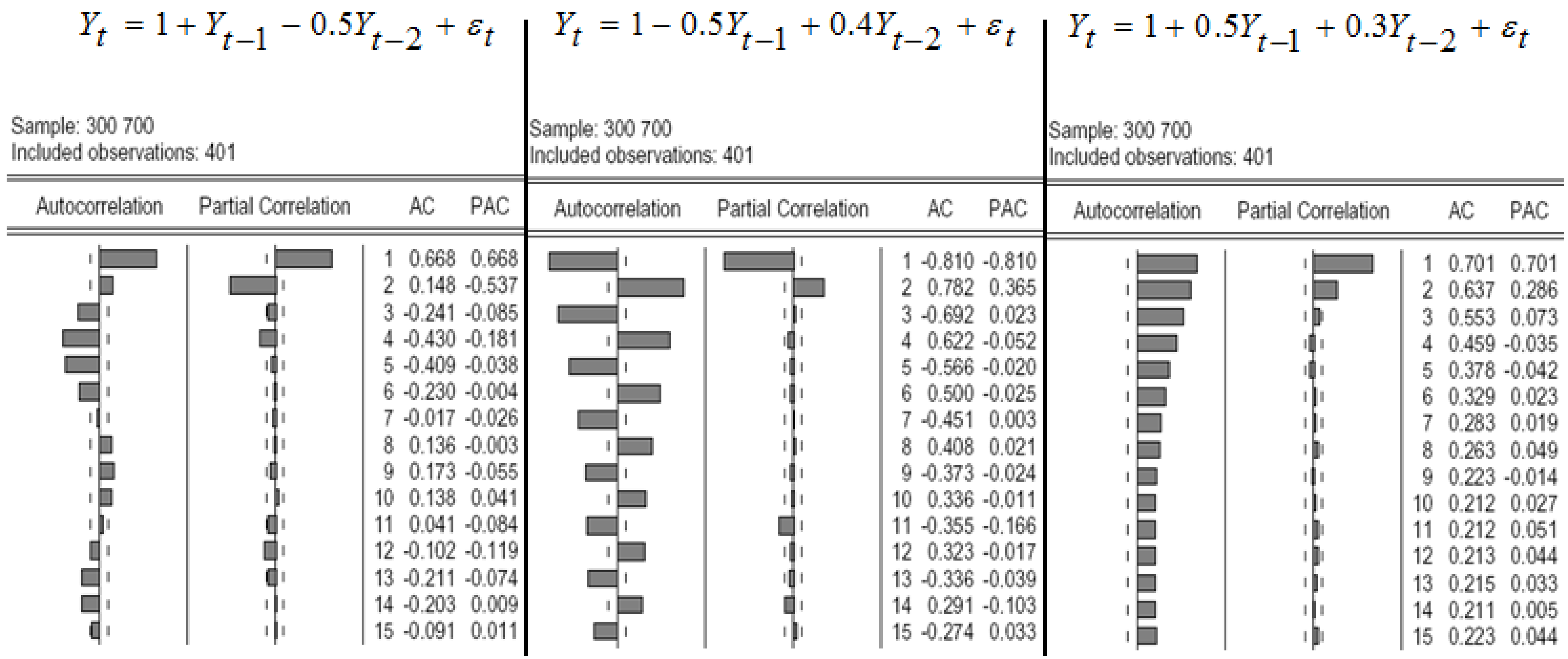


Figure 7.9 Nonstationary AR(2)

Figure 7.10 Autocorrelation Functions of Covariance-Stationary AR(2) Processes



$Y_t = 1 + Y_{t-1} - 0.5Y_{t-2} + \varepsilon_t$

	AC	PAC
1	0.668	0.668
2	0.148	-0.537
3	-0.241	-0.085
4	-0.430	-0.181
5	-0.409	-0.038
6	-0.230	-0.004
7	-0.017	-0.026
8	0.136	-0.003
9	0.173	-0.055
10	0.138	0.041
11	0.041	-0.084
12	-0.102	-0.119
13	-0.211	-0.074
14	-0.203	0.009
15	-0.091	0.011

$Y_t = 1 - 0.5Y_{t-1} + 0.4Y_{t-2} + \varepsilon_t$

	AC	PAC
1	-0.810	-0.810
2	0.782	0.365
3	-0.692	0.023
4	0.622	-0.052
5	-0.566	-0.020
6	0.500	-0.025
7	-0.451	0.003
8	0.408	0.021
9	-0.373	-0.024
10	0.336	-0.011
11	-0.355	-0.166
12	0.323	-0.017
13	-0.336	-0.039
14	0.291	-0.103
15	-0.274	0.033

$Y_t = 1 + 0.5Y_{t-1} + 0.3Y_{t-2} + \varepsilon_t$

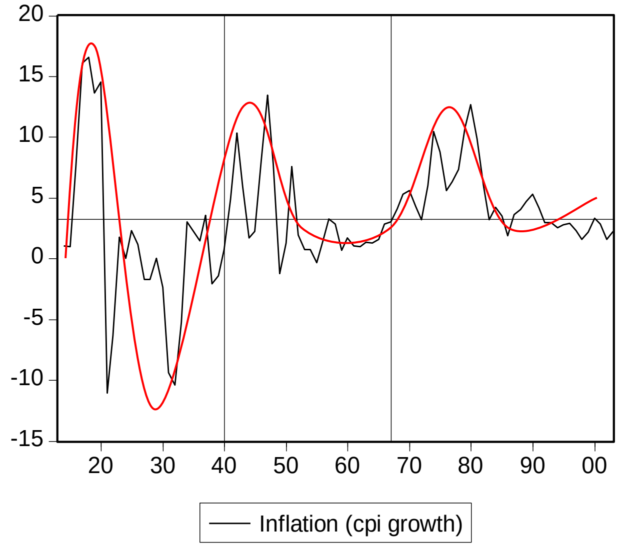
	AC	PAC
1	0.701	0.701
2	0.637	0.286
3	0.553	0.073
4	0.459	-0.035
5	0.378	-0.042
6	0.329	0.023
7	0.283	0.019
8	0.263	0.049
9	0.223	-0.014
10	0.212	0.027
11	0.212	0.051
12	0.213	0.044
13	0.215	0.033
14	0.211	0.005
15	0.223	0.044

PACF: $r_1 \neq 0, r_2 \neq 0, r_k = 0 \; k \geq 2$

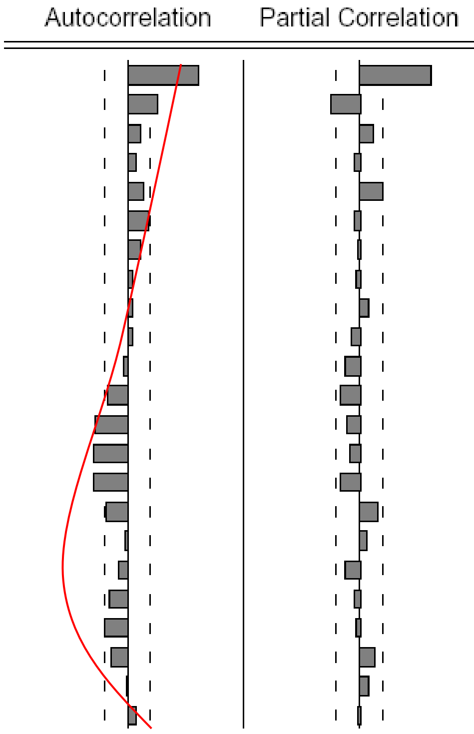
ACF: decays to zero

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2}, \rho_0 = 1, \rho_1 = \frac{\phi_1}{1 - \phi_2}$$

Figure 7.11 U.S. Inflation Rate



Sample: 1913 2003
Included observations: 90



	AC	PAC
1	0.639	0.639
2	0.259	-0.252
3	0.117	0.128
4	0.066	-0.038
5	0.144	0.210
6	0.181	-0.039
7	0.115	-0.016
8	0.039	-0.032
9	0.039	0.089
10	0.035	-0.067
11	-0.047	-0.128
12	-0.174	-0.162
13	-0.280	-0.114
14	-0.303	-0.081
15	-0.306	-0.172
16	-0.184	0.156
17	-0.032	0.073
18	-0.075	-0.125
19	-0.161	-0.034
20	-0.208	-0.031
21	-0.155	0.137
22	-0.008	0.081
23	0.073	-0.005

Table 7.1 Estimation Results, U.S. Inflation Rate, AR(2) Model

Dependent Variable: CPI_GR
Method: Least Squares
Sample (adjusted): 1916 2003
Included observations: 88 after adjustments
Convergence achieved after 3 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	3.311924	0.880693	3.760588	0.0003
AR(1)	0.799420	0.104934	7.618326	0.0000
AR(2)	-0.251858	0.104879	-2.401413	0.0185

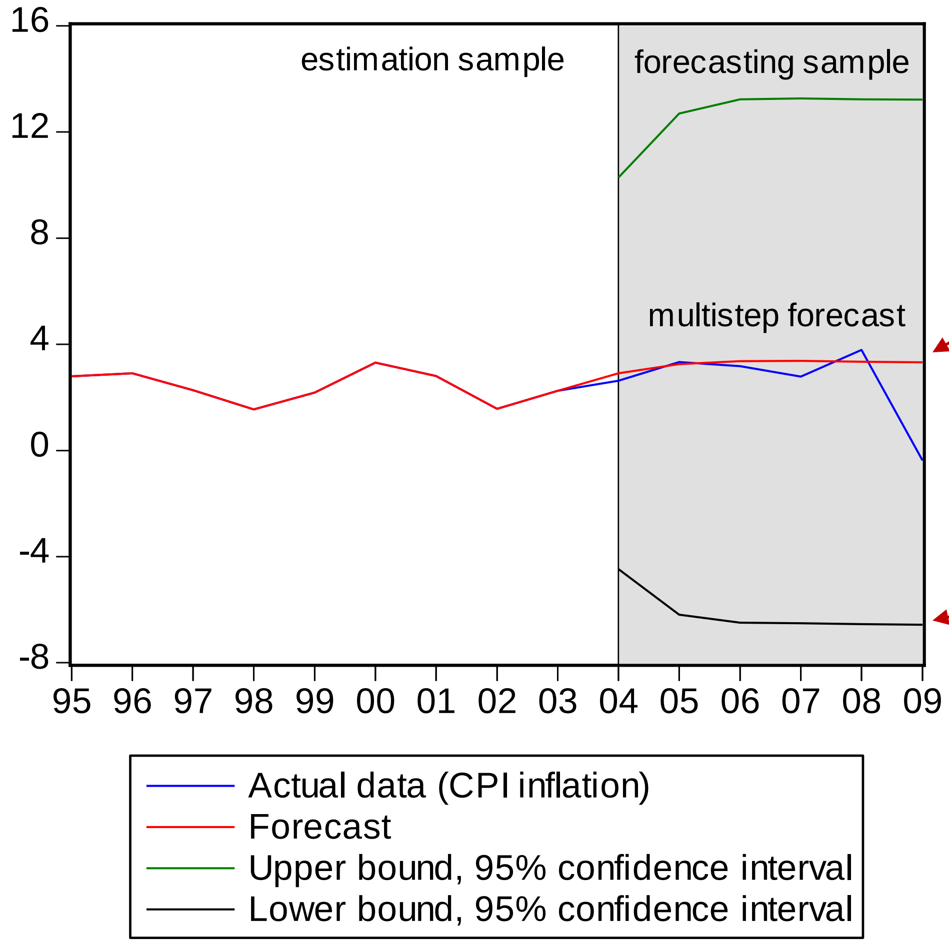
R-squared	0.445779	Mean dependent var	3.298182
Adjusted R-squared	0.432739	S.D. dependent var	4.962744
S.E. of regression	3.737777	Akaike info criterion	5.508356
Sum squared resid	1187.533	Schwarz criterion	5.592810
Log likelihood	-239.3676	F-statistic	34.18425
Durbin-Watson stat	1.904683	Prob(F-statistic)	0.000000

Inverted AR Roots	.40-.30i	.40+.30i

Table 7.2 Multistep Forecast of U.S. Inflation Rate

$h=1$ 2004	$f_{t,1} = \hat{c} + \hat{\phi}_1 Y_t + \hat{\phi}_2 Y_{t-1} =$ $= 1.49 + 0.79 \times 2.25 - 0.25 \times 1.56 =$ ≈ 2.90	$\sigma^2_{t+1\|t} = \hat{\sigma}^2_{\varepsilon} = 3.74^2$	$f(Y_{t+1} \mid I_t) \to N(\mu_{t+1\|t}, \sigma^2_{t+1\|t})$ $= N(2.90, 3.74^2)$
$h=2$ 2005	$f_{t,2} = \hat{c} + \hat{\phi}_1 f_{t,1} + \hat{\phi}_2 Y_t =$ $= 1.49 + 0.79 \times 2.90 - 0.25 \times 2.25 =$ ≈ 3.25	$\sigma^2_{t+2\|t} = \hat{\sigma}^2_{\varepsilon}(1 + \hat{\phi}_1^{\ 2}) =$ $= 3.74^2(1 + 0.79^2) =$ $\approx 4.81^2$	$f(Y_{t+2} \mid I_t) \to N(3.25, 4.81^2)$
$h=3$ 2006	$f_{t,3} = \hat{c} + \hat{\phi}_1 f_{t,2} + \hat{\phi}_2 f_{t,1} =$ $= 1.49 + 0.79 \times 3.25 - 0.25 \times 2.90 =$ ≈ 3.36	$\sigma^2_{t+3\|t} = \hat{\sigma}^2_{\varepsilon}(1 + \hat{\phi}_1^2 +$ $+ (\hat{\phi}_2 + \hat{\phi}_1^2)^2) =$ $\approx 5.03^2$	$f(Y_{t+3} \mid I_t) \to N(3.36, 5.03^2)$
$h=4$ 2007	$f_{t,4} = \hat{c} + \hat{\phi}_1 f_{t,3} + \hat{\phi}_2 f_{t,2} =$ $= 1.49 + 0.79 \times 3.36 - 0.25 \times 3.25 =$ ≈ 3.37	$\sigma^2_{t+4\|t} \approx 5.04^2$	$f(Y_{t+4} \mid I_t) \to N(3.37, 5.04^2)$
$h = 5$ 2008	$f_{t,5} = \hat{c} + \hat{\phi}_1 f_{t,4} + \hat{\phi}_2 f_{t,3} =$ $= 1.49 + 0.79 \times 3.37 - 0.25 \times 3.36 =$ ≈ 3.34	$\sigma^2_{t+5\|t} \approx 5.04^2$	$f(Y_{t+5} \mid I_t) \to N(3.34, 5.04^2)$
$h=6$ 2009	$f_{t,6} = \hat{c} + \hat{\phi}_1 f_{t,5} + \hat{\phi}_2 f_{t,4} =$ $= 1.49 + 0.79 \times 3.34 - 0.25 \times 3.37 =$ ≈ 3.32	$\sigma^2_{t+6\|t} \approx 5.04^2$	$f(Y_{t+6} \mid I_t) \to N(3.32, 5.04^2)$

Figure 7.12 U.S. Inflation Rate, Multistep Forecast

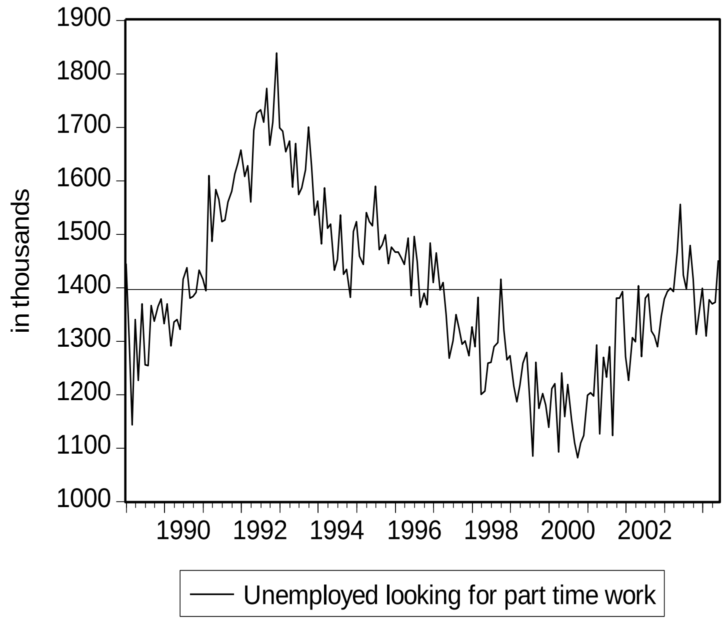


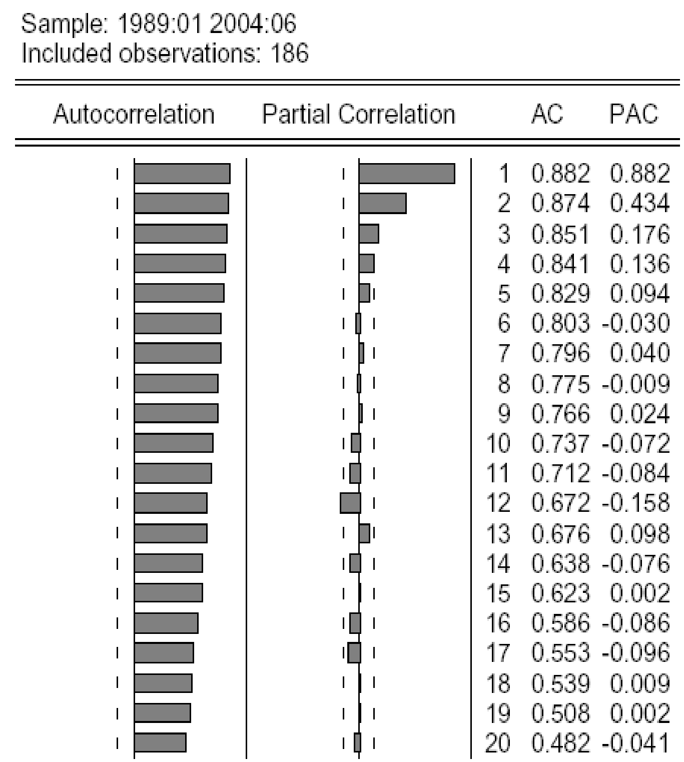
Unlike MA, the AR forecasts keep changing, but also converging to the unconditional mean

Bands converge slowly to unconditional uncertainty bands

7.2.3 AR(p) Process

Figure 7.13 Number of Unemployed People Looking for Part-Time Work

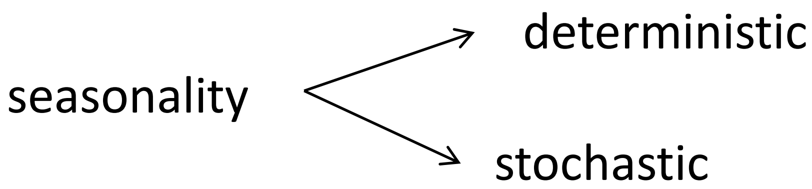




	AC	PAC
1	0.882	0.882
2	0.874	0.434
3	0.851	0.176
4	0.841	0.136
5	0.829	0.094
6	0.803	-0.030
7	0.796	0.040
8	0.775	-0.009
9	0.766	0.024
10	0.737	-0.072
11	0.712	-0.084
12	0.672	-0.158
13	0.676	0.098
14	0.638	-0.076
15	0.623	0.002
16	0.586	-0.086
17	0.553	-0.096
18	0.539	0.009
19	0.508	0.002
20	0.482	-0.041

7.3.1 Deterministic and Stochastic Seasonal Cycles

A seasonal cycle is defined as a periodic fluctuation in the data associated with the calendar.



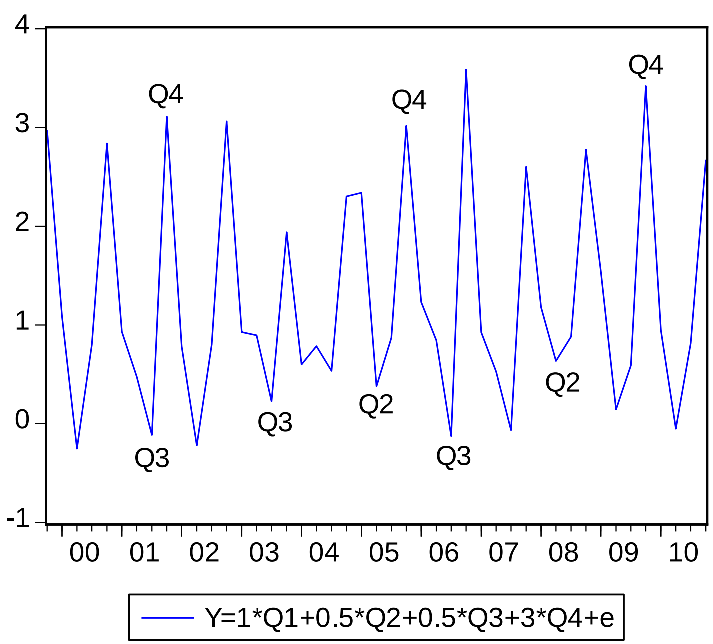


Figure 7.14 Deterministic Seasonality

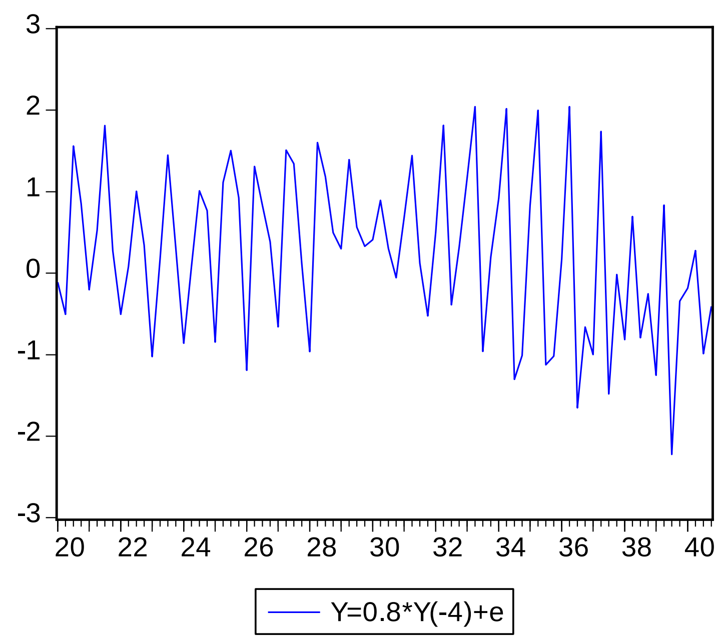


Figure 7.15 Stochastic Seasonality

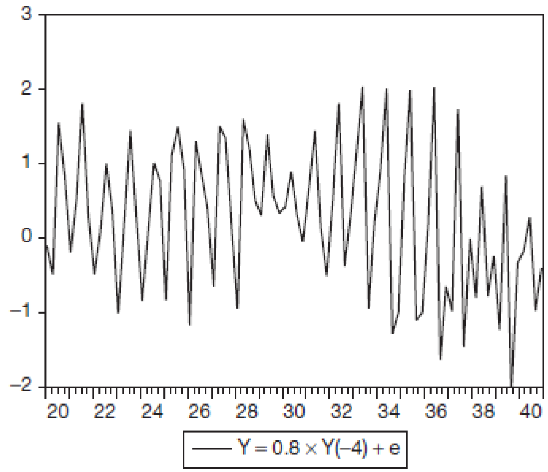
7.3.2 Seasonal ARMA Models

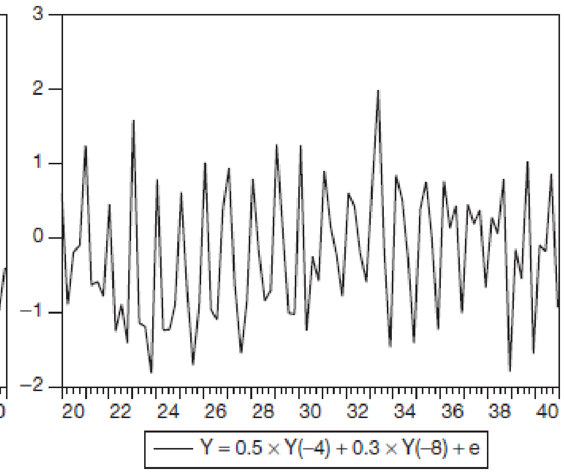
Figure 7.16 Seasonal AR(1) and AR(2), Time Series Plots and Autocorrelograms

$$Y_t = c + \phi_s Y_{t-s} + \phi_{2s} Y_{t-2s} + \phi_{3s} Y_{t-3s} + \cdots + \phi_{ps} Y_{t-ps} + \varepsilon_t$$

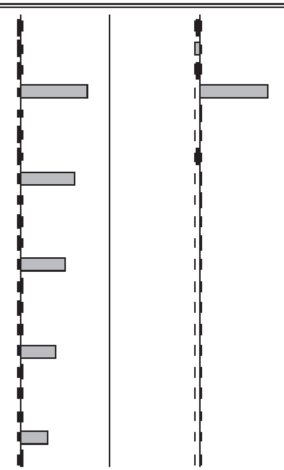
In the seasonal context, the order of the process needs to be understood in light of the data seasonality.

For identifying the order of seasonal AR and MA models, we extrapolate what we know from nonseasonal AR and MA models.



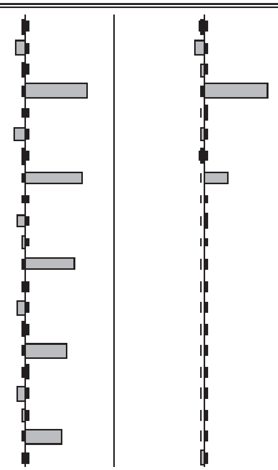


Included observations: 4201



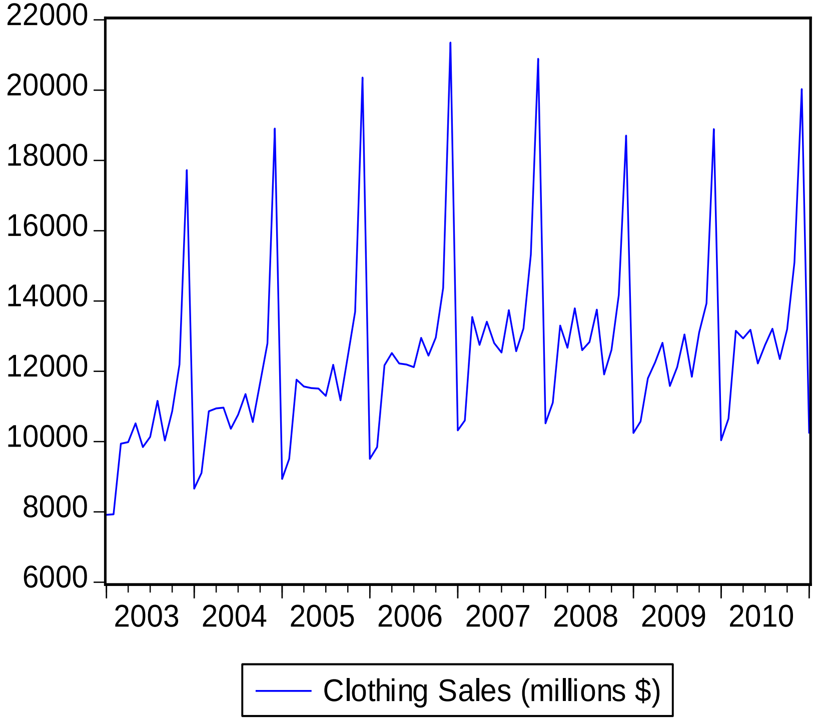
Lag	AC	PAC
1	-0.020	-0.020
2	-0.027	-0.027
3	-0.022	-0.023
4	0.796	0.796
5	-0.007	0.030
6	-0.024	0.004
7	-0.025	-0.009
8	0.641	0.021
9	0.012	0.027
10	-0.017	0.014
11	-0.014	0.036
12	0.515	-0.001
13	0.021	-0.005
14	-0.011	0.004
15	-0.001	0.015
16	0.414	0.002
17	0.025	-0.001
18	-0.004	0.006
19	0.004	-0.007
20	0.337	0.012
21	0.028	0.004

Included observations: 4197

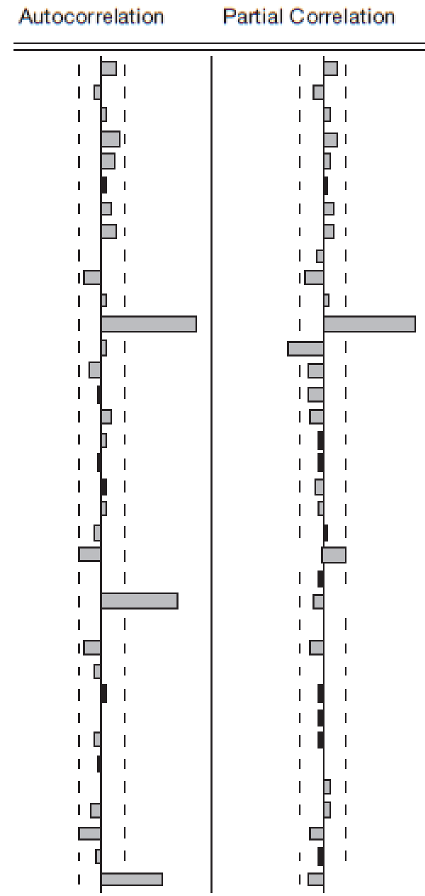


Lag	AC	PAC
1	-0.018	-0.018
2	-0.109	-0.109
3	-0.022	-0.026
4	0.729	0.725
5	-0.003	0.022
6	-0.116	-0.033
7	-0.032	-0.021
8	0.668	0.290
9	0.005	0.010
10	-0.097	0.033
11	-0.035	0.000
12	0.560	0.005
13	0.016	0.008
14	-0.097	-0.005
15	-0.033	0.014
16	0.485	0.002
17	0.021	0.003
18	-0.088	-0.003
19	-0.037	-0.006
20	0.417	0.005
21	0.009	-0.040

Figure 7.17 Monthly Clothing Sales in the United States. Time Series Plot and Autocorrelation Functions



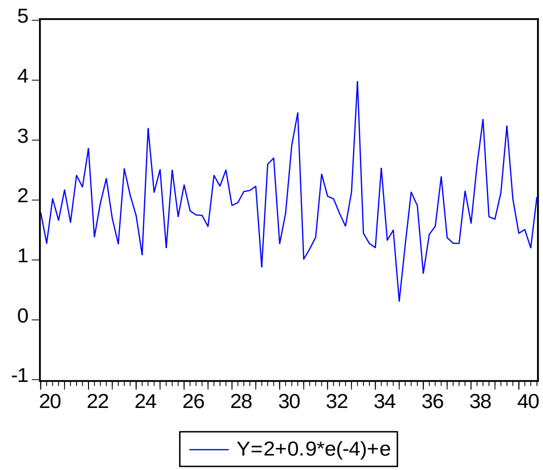
Sample: 2003M01 2011M01
Included observations: 97

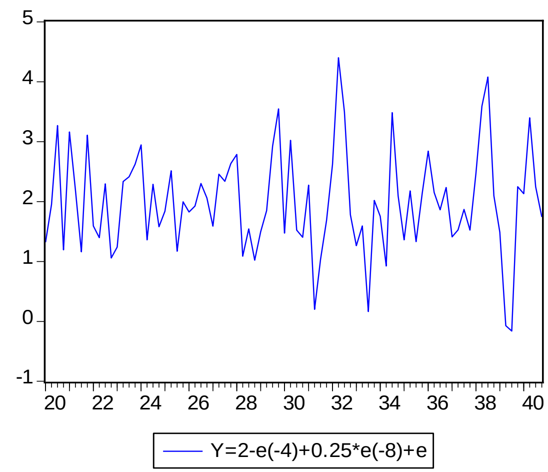


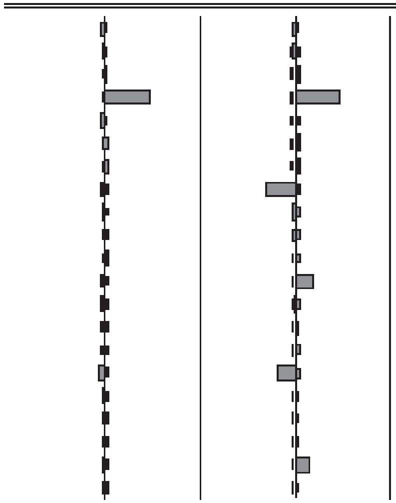
	AC	PAC
1	0.132	0.132
2	-0.069	-0.088
3	0.047	0.070
4	0.157	0.138
5	0.106	0.077
6	0.029	0.026
7	0.095	0.093
8	0.125	0.082
9	-0.001	-0.041
10	-0.150	-0.161
11	0.051	0.051
12	0.826	0.827
13	0.054	-0.375
14	-0.109	-0.132
15	-0.012	-0.119
16	0.088	-0.111
17	0.044	-0.024
18	-0.026	-0.019
19	0.030	-0.066
20	0.055	-0.033
21	-0.059	0.034
22	-0.182	0.197
23	-0.002	-0.024
24	0.672	-0.084
25	0.001	0.006
26	-0.147	-0.112
27	-0.058	-0.004
28	0.031	-0.024
29	0.000	-0.018
30	-0.063	-0.018
31	-0.014	0.004
32	0.011	0.071
33	-0.083	0.059
34	-0.202	-0.096
35	-0.038	-0.011
36	0.536	-0.119

Figure 7.18 Seasonal MA(1) and MA(2), Time Series Plots and Autocorrelograms

$$Y_t = \mu + \theta_s \varepsilon_{t-s} + \theta_{2s} \varepsilon_{t-2s} + \theta_{3s} \varepsilon_{t-3s} + \cdots + \theta_{qs} \varepsilon_{t-qs} + \varepsilon_t$$

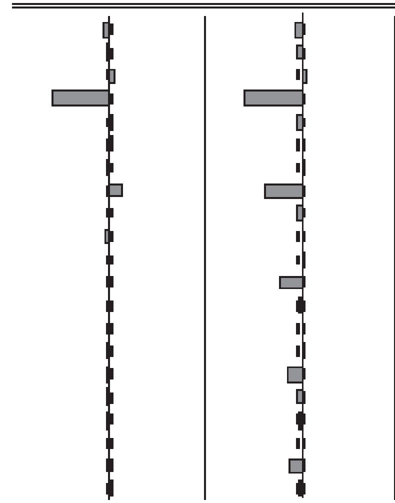






Included observations: 4193

Lag	AC	PAC
1	-0.035	-0.035
2	-0.021	-0.023
3	0.019	0.018
4	0.490	0.491
5	-0.038	0.002
6	0.008	0.028
7	0.044	0.024
8	-0.014	-0.331
9	-0.019	-0.016
10	-0.003	-0.041
11	0.026	-0.003
12	-0.034	0.205
13	-0.022	-0.014
14	-0.007	0.027
15	0.010	0.008
16	-0.051	-0.202
17	-0.013	0.004
18	0.012	-0.006
19	0.007	0.009
20	-0.028	0.157
21	0.008	0.009



Included observations: 4189

Lag	AC	PAC
1	-0.056	-0.056
2	-0.023	-0.027
3	0.047	0.044
4	-0.613	-0.612
5	0.026	-0.040
6	0.047	0.012
7	-0.016	0.034
8	0.134	-0.383
9	-0.005	-0.031
10	-0.034	0.002
11	0.011	0.040
12	-0.000	-0.229
13	0.006	-0.010
14	0.013	0.001
15	-0.012	0.025
16	-0.016	-0.155
17	-0.019	-0.034
18	-0.010	-0.020
19	0.001	0.003
20	0.003	-0.117
21	0.022	-0.020

7.3.3 Combining ARMA and Seasonal ARMA Models

Purely seasonal ARMA: $$Y_t = c + \phi_s Y_{t-s} + \phi_{2s} Y_{t-2s} + \cdots\cdots + \phi_{ps} Y_{t-ps} + \theta_s \varepsilon_{t-s} + \theta_{2s} \varepsilon_{t-2s} + \theta_{3s} \varepsilon_{t-3s} + \cdots + \theta_{qs} \varepsilon_{t-qs} + \varepsilon_t$$

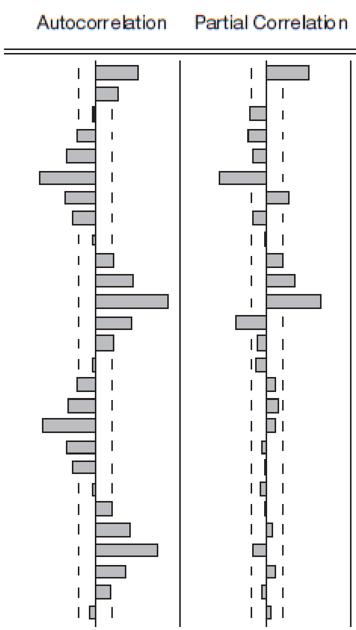
<u>Example:</u>
Combined S-ARMA$(1,2)_4$ + non-seasonal ARMA(2,1):

$$(1 - \phi_1 L - \phi_2 L^2)(1 - \phi_4 L^4) Y_t = c + (1 + \theta_4 L^4 + \theta_8 L^8)(1 + \theta_1 L)\varepsilon_t$$

<u>Data example</u>

Figure 7.19 Monthly Changes of Private Residential Construction in U.S. (Millions of Dollars), Time Series Plot and Autocorrelograms

Sample: 2002M01 2011M01
Included observations: 108



	AC	PAC
1	0.507	0.507
2	0.260	0.004
3	-0.009	-0.191
4	-0.225	-0.209
5	-0.357	-0.179
6	-0.678	-0.571
7	-0.365	0.263
8	-0.258	-0.176
9	-0.041	-0.010
10	0.225	0.203
11	0.453	0.326
12	0.876	0.660
13	0.446	-0.362
14	0.228	-0.114
15	-0.033	-0.115
16	-0.215	0.098
17	-0.327	0.138
18	-0.627	0.100
19	-0.341	-0.033
20	-0.251	-0.015
21	-0.046	-0.088
22	0.198	-0.015
23	0.412	0.066
24	0.746	-0.160
25	0.380	0.085
26	0.187	-0.054
27	-0.051	0.057

It makes sense to propose: ARMA(1,0) + S-ARMA$(1,0)_{12}$

ARMA(1,0) + S-ARMA(1,0)$_{12}$

$$(1-\phi_1 L)(1-\phi_{12}L^{12})Y_t = c + \varepsilon_t$$

Or ARMA(13,0) with parameter restrictions

$$(1-\phi_1 L-\phi_{12}L^{12}+\phi_1\phi_{12}L^{13})Y_t = c + \varepsilon_t$$

TABLE 7.3 Monthly Changes in Residential Construction, Estimation Results of AR(1) and S-AR(1) Model

Dependent Variable: change CONST
Method: Least Squares
Sample (adjusted): 2003M03 2011M01
Included observations: 95 after adjustments
Convergence achieved after 6 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-593.2408	2399.622	-0.247223	0.8053
AR(1)	0.439971	0.093551	4.703012	0.0000
SAR(12)	0.923569	0.038771	23.82102	0.0000

R-squared	0.894790	Mean dependent var	-128.3158
Adjusted R-squared	0.892502	S.D. dependent var	3036.076
S.E. of regression	995.4326	Akaike info criterion	16.67530
Sum squared resid	91161518	Schwarz criterion	16.75595
Log likelihood	-789.0768	F-statistic	391.2194
Durbin–Watson stat	2.115719	Prob(F-statistic)	0.000000

Figure 7.20 Monthly Changes in Residential Construction, Multistep Forecast