CHAPTER 6
FORECASTING WITH MOVING AVERAGE (MA) MODELS

6.1 A Model with No Dependence: White Noise

Definition: (White noise)

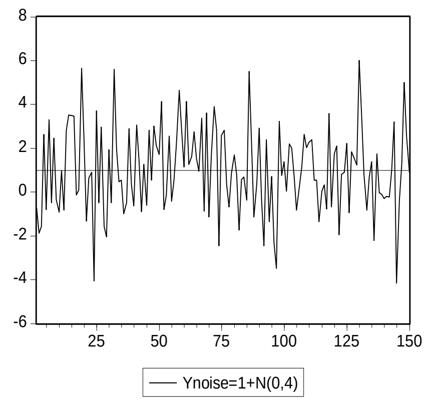
Let ε_t be a stochastic process.

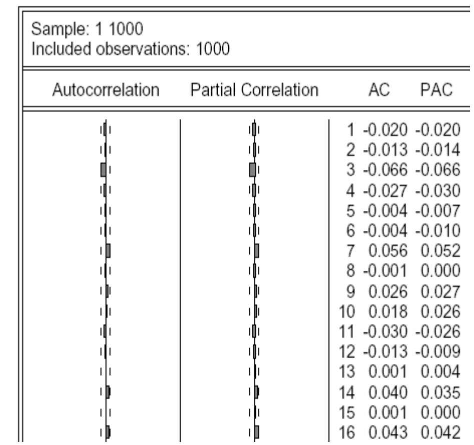
If $E[\varepsilon_t] = 0$, $Var(\varepsilon_t) = {\sigma_\varepsilon}^2$(constant)

and $Cov(\varepsilon_t, \varepsilon_{t-k}) = 0, k \neq 0$

Then we say that the process ε_t is white noise.

The white noise process is stationary and it does not exhibit any linear dependence.





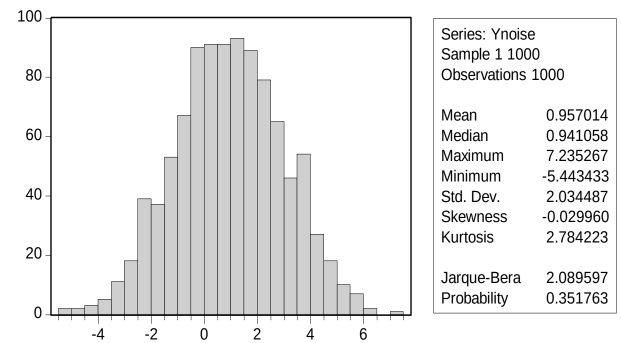
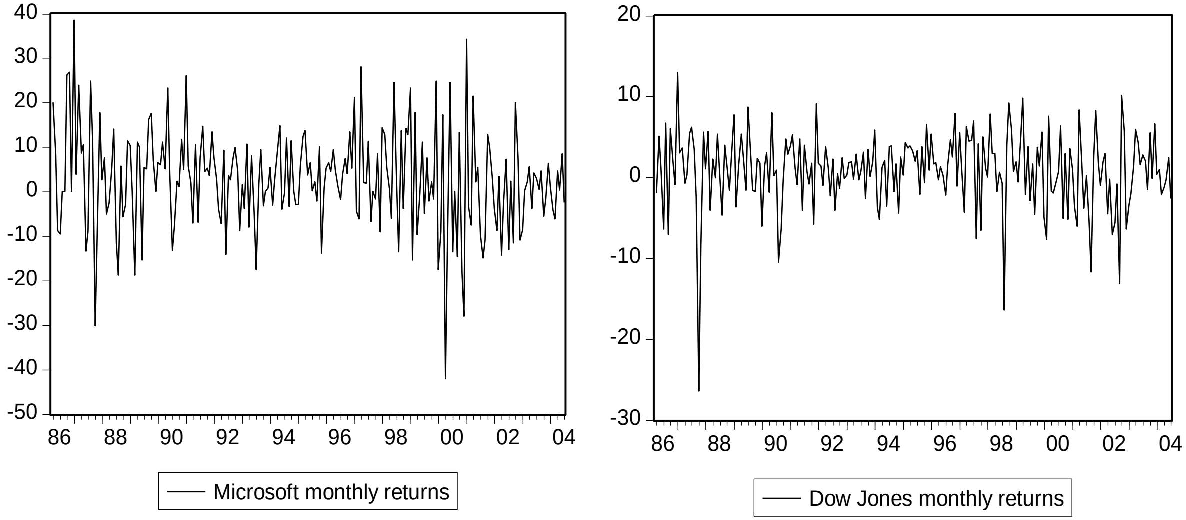


Figure 6.1 Time Series and Autocorrelation Functions of a Simulated White Noise Process

Figure 6.2 Autocorrelation Functions of Monthly Returns to Microsoft and the Dow Jones Index

Difficult to predict



Sample: 1986:03 2004:07
Included observations: 220

Lag	AC	PAC
1	-0.081	-0.081
2	-0.094	-0.101
3	0.132	0.117
4	-0.017	-0.006
5	0.008	0.030
6	-0.013	-0.029
7	0.106	0.113
8	0.015	0.023
9	-0.006	0.024
10	0.131	0.112
11	0.013	0.035
12	-0.016	0.005
13	-0.020	-0.045
14	0.030	0.013
15	-0.075	-0.091
16	0.064	0.068
17	0.085	0.051
18	-0.094	-0.064
19	-0.049	-0.079
20	0.046	0.005

Sample: 1986:03 2004:07
Included observations: 220

Lag	AC	PAC
1	-0.021	-0.021
2	-0.044	-0.044
3	-0.056	-0.058
4	-0.126	-0.131
5	0.048	0.037
6	-0.031	-0.045
7	0.092	0.081
8	-0.044	-0.057
9	-0.043	-0.030
10	0.043	0.035
11	-0.012	0.006
12	0.015	-0.006
13	-0.003	0.003
14	-0.034	-0.034
15	-0.059	-0.060
16	0.039	0.042
17	0.031	0.012
18	0.079	0.076
19	-0.026	-0.028
20	-0.016	0.005

The Wold Decomposition Theorem

If $\{Y_t\}$ is a covariance stationary process and $\{\varepsilon_t\}$ is a white noise process then there exists a unique linear representation as:

$$Y_t = V_t + \varepsilon_t + \psi_1 \varepsilon_{t-1} + \psi_2 \varepsilon_{t-2} + \cdots = V_t + \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}$$

where V_t is a deterministic component and $\sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}$ is the stochastic component with $\psi_0 = 1, \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j} < \infty$.

The Wold decomposition theorem guarantees that any (purely nondeterministic) covariance stationary stochastic process can be expressed as a linear combination of past shocks

Finite Representation of the Wold Decomposition Theorem

$$Y_t = \varepsilon_t + \psi_1\varepsilon_{t-1} + \psi_2\varepsilon_{t-2} + \cdots = (1 + \psi_1 L + \psi_2 L^2 + \cdots)\varepsilon_t = \boldsymbol{\Psi}(L)\varepsilon_t$$

The infinite polynomial can be appproximated by the ratio of two finite polynominals :

$$\boldsymbol{\Psi}(L) \approx \frac{\boldsymbol{\Theta}_q(L)}{\boldsymbol{\Phi}_p(L)}$$

where $\boldsymbol{\Theta}_q(L) = 1 + \theta_1 L + \theta_2 L^2 + \cdots + \theta_q L^q$

and $\boldsymbol{\Phi}_p(L) = 1 - \phi_1 L - \phi_2 L^2 - \cdots - \phi_p L^p$

and the Wold decomposition can be approximated (or written) as:

$$Y_t = \boldsymbol{\Psi}(L)\varepsilon_t \approx \frac{\boldsymbol{\Theta}_q(L)}{\boldsymbol{\Phi}_p(L)}\varepsilon_t \Leftrightarrow \boldsymbol{\Phi}_p(L)Y_t = \boldsymbol{\Theta}_q(L)\varepsilon_t \Leftrightarrow$$

$$\Leftrightarrow Y_t - \underbrace{\phi_1 Y_{t-1} - \phi_2 Y_{t-2} - \cdots - \phi_p Y_{t-p}}_{AR(p)} = \underbrace{\varepsilon_t + \theta_1\varepsilon_t + \theta_2\varepsilon_t + \cdots + \theta_q\varepsilon_{t-q}}_{MA(q)}$$

A moving average process of order $q \geq 0$, referred as MA(q), has the following functional form:

$$Y_t = \mu + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} \cdots + \theta_q \varepsilon_{t-q} + \varepsilon_t$$

where ε_t is a zero-mean white noise process.

A process is **invertible** if it can be written as a linear function of past observations (up to an unpredictable shock):

$$X_t = \varepsilon_t + \pi_1 X_{t-1} + \pi_2 X_{t-2} + \pi_3 X_{t-3} + \ldots$$

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

$$|\xi_i| > 1,$$

I.e., *iff* the modules of the inverse roots are smaller than 1:

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

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

6.3.1 MA(1) Process

$$Y_t = \mu + \theta\varepsilon_{t-1} + \varepsilon_t$$

$$E[Y_t] = \mu$$

$$\sigma_Y{}^2 = Var(Y_t) = (1 + \theta^2)\sigma_\varepsilon{}^2$$

$$\rho_1 = \frac{\theta}{1 + \theta^2}$$

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

r_k decays to zero

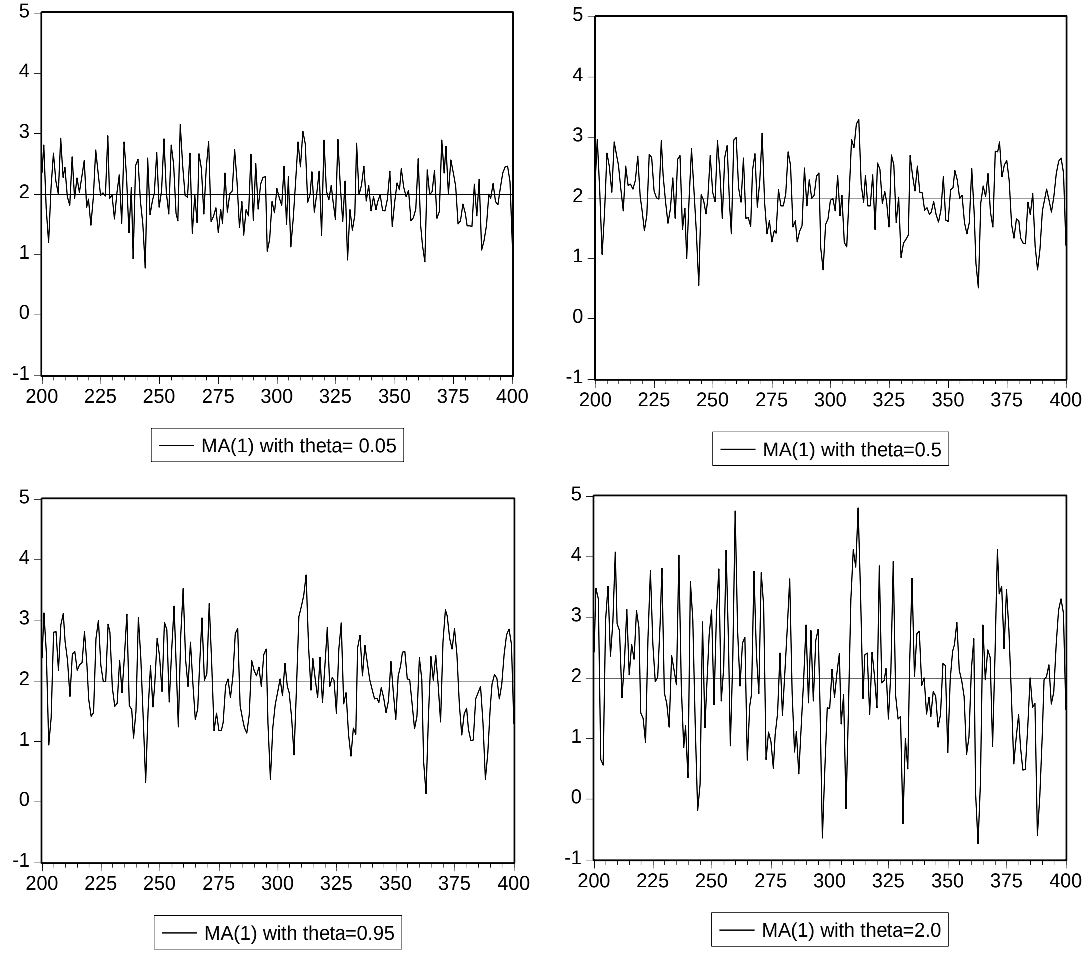


Figure 6.3 Simulated MA(1) Processes

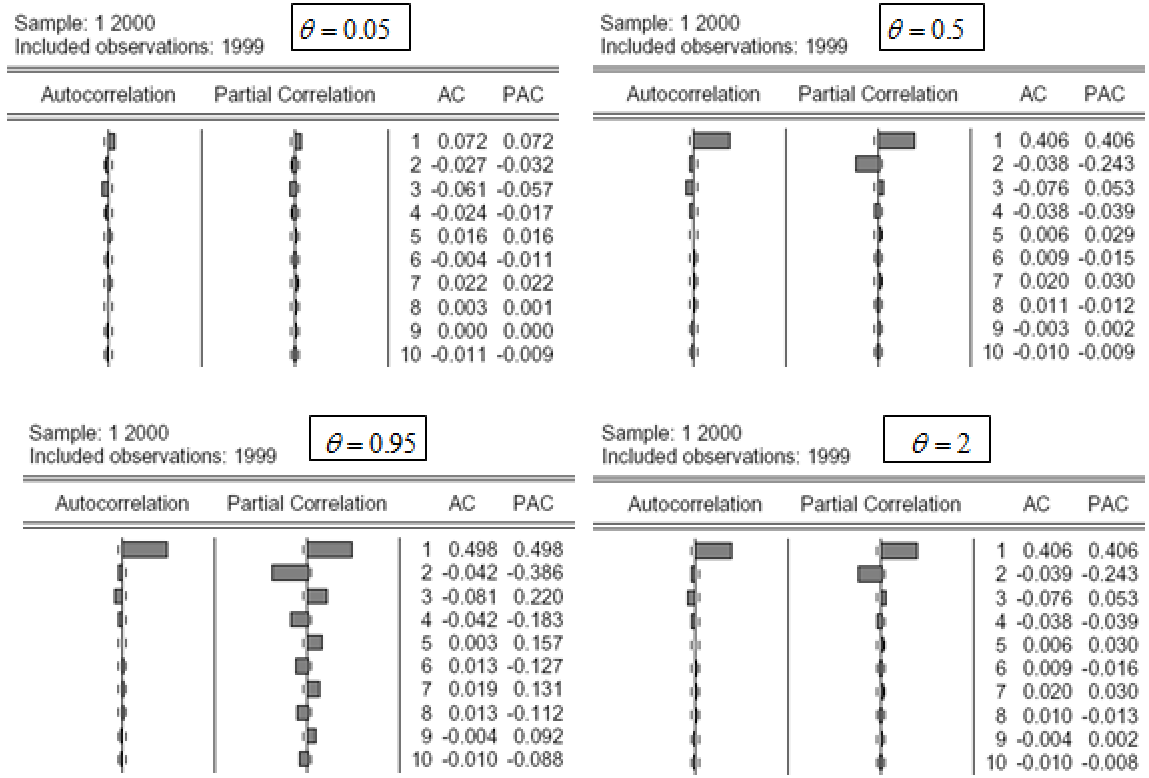
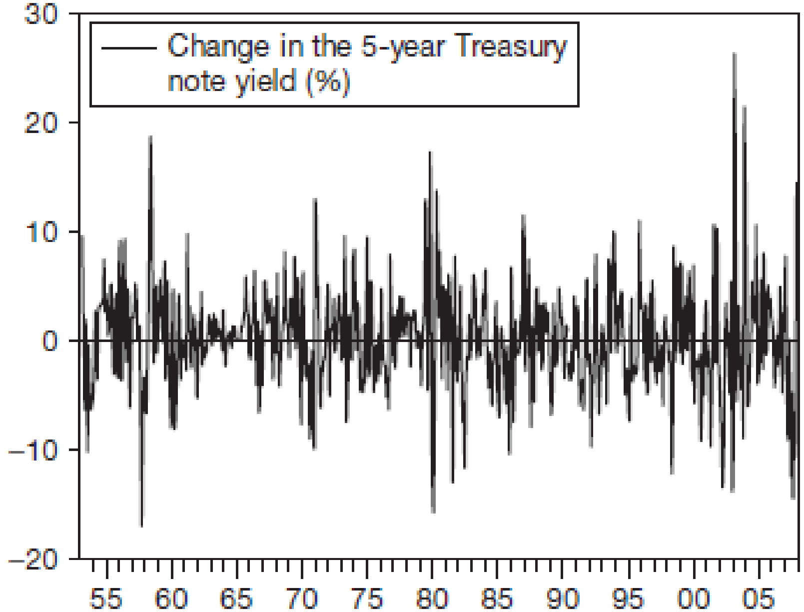


Figure 6.4 Autocorrelation Functions of Simulated MA(1) Processes

We say that an MA(1) process is **invertible** if $|\theta| < 1$

Figure 6.5 Percentage Changes in the 5-Year Treasury Note Yield



Sample: 1953M04 2008M04
Included observations: 660

Autocorrelation	Partial Correlation		AC	PAC
		1	0.339	0.339
		2	-0.073	-0.213
		3	0.007	0.129
		4	0.014	-0.063
		5	-0.043	-0.017
		6	-0.073	-0.060
		7	-0.069	-0.035

Table 6.1 Estimation Output: 5-Year Treasury Yield (Monthly Percentage Changes)

Dependent Variable: DY
Method: Least Squares
Sample (adjusted): 1953M05 2007M11
Included observations: 655 after adjustments
Convergence achieved after 7 iterations
Backcast: 1953M04

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.160159	0.258095	0.620544	0.5351
MA(1)	0.485011	0.034468	14.07130	0.0000

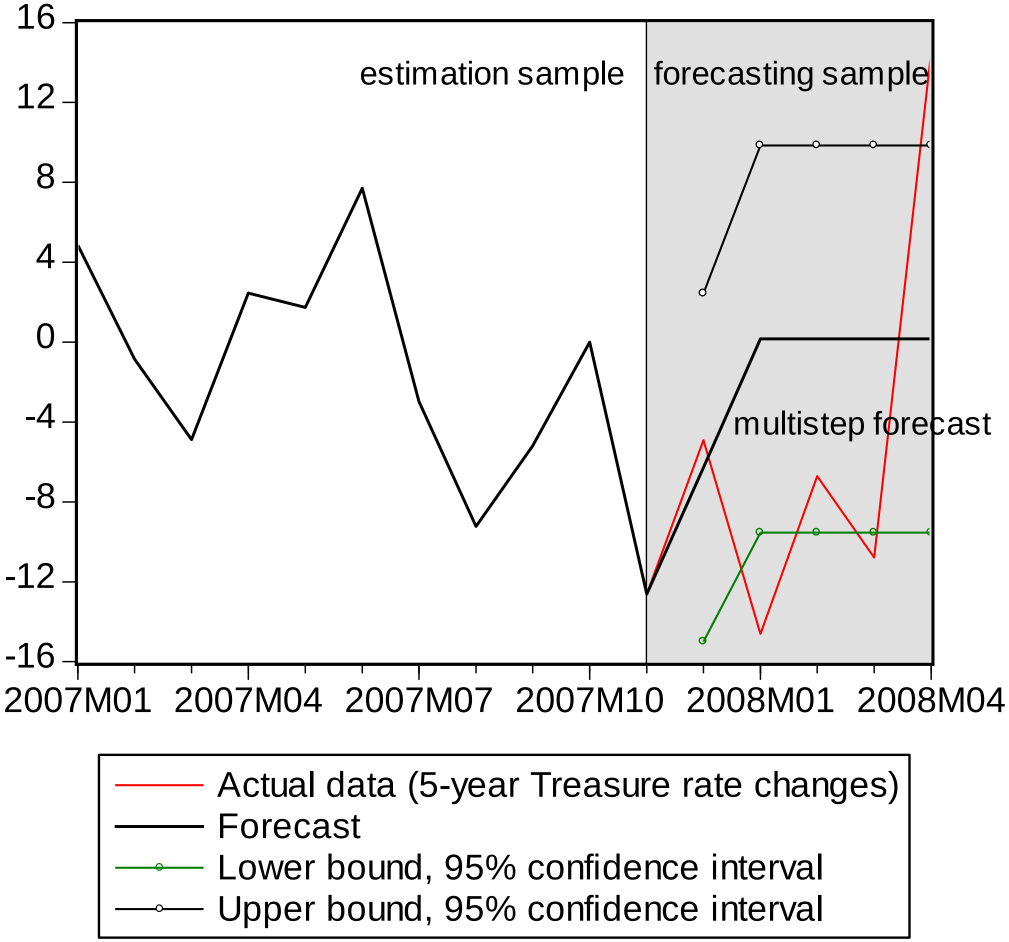
R-squared	0.165370	Mean dependent var	0.168613
Adjusted R-squared	0.164092	S.D. dependent var	4.866609
S.E. of regression	4.449443	Akaike info criterion	5.826484
Sum squared resid	12927.79	Schwarz criterion	5.840177
Log likelihood	-1906.173	F-statistic	129.3829
Durbin-Watson stat	2.055799	Prob(F-statistic)	0.000000

Inverted MA Roots	-.49

$h=1$ 12/2007	$f_{t,1}=\hat{\mu}+\hat{\theta}\varepsilon_t=$ $=0.160+0.485\hat{\varepsilon}_t$ $=-6.276\%$	$\sigma^2_{t+1\|t}=\hat{\sigma}^2_\varepsilon=4.449^2$	$f(Y_{t+1}\|I_t)\rightarrow N(\mu_{t+1\|t},\sigma^2_{t+1\|t})$ $=N(-6.276, 4.449^2)$
$h=2$ 1/2008	$f_{t,2}=\hat{\mu}=0.160\%$	$\sigma^2_{t+2\|t}=\hat{\sigma}^2_\varepsilon(1+\hat{\theta}^2)$ $=4.449^2(1+0.485^2)$ $=23.683=\hat{\sigma}^2_Y$	$f(Y_{t+2}\|I_t)\rightarrow N(0.16, 23.683)$
$h=3$ 2/2008	$f_{t,3}=\hat{\mu}=0.160\%$	$\sigma^2_{t+3\|t}=23.683=\hat{\sigma}^2_Y$	$f(Y_{t+3}\|I_t)\rightarrow N(0.16, 23.683)$
$h=4$ 3/2008	$f_{t,4}=\hat{\mu}=0.160\%$	$\sigma^2_{t+4\|t}=23.683=\hat{\sigma}^2_Y$	$f(Y_{t+4}\|I_t)\rightarrow N(0.16, 23.683)$
$h=5$ 4/2008	$f_{t,5}=\hat{\mu}=0.160\%$	$\sigma^2_{t+5\|t}=23.683=\hat{\sigma}^2_Y$	$f(Y_{t+5}\|I_t)\rightarrow N(0.16, 23.683)$

Table 6.2 December 2007-April 2008 Forecasts of 5-year Treasure Yield Changes

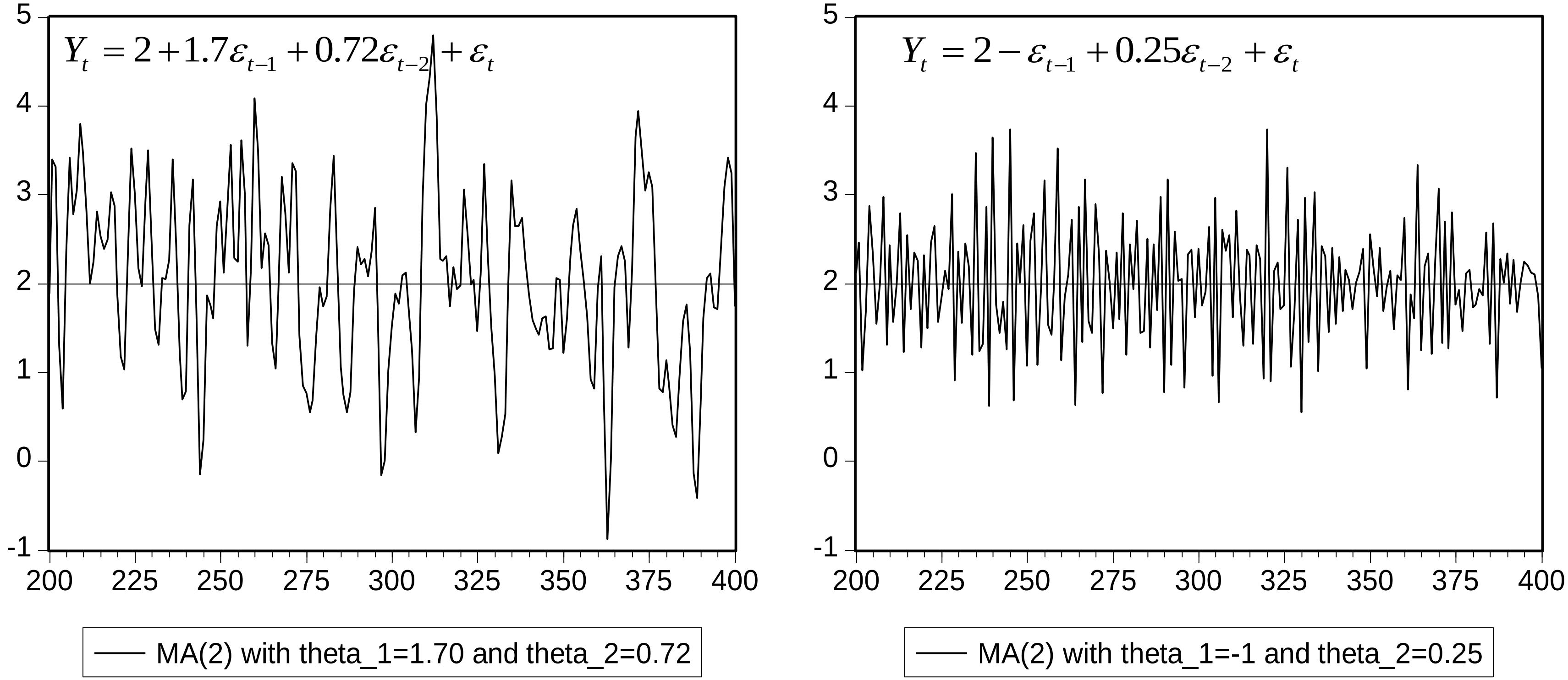
Figure 6.6 Multistep Forecast of Monthly Changes of 5-year Treasury Yield



MA(2) Process $Y_t = \mu + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \varepsilon_t$

$$E[Y_t] = \mu$$

$$\sigma_Y{}^2 = Var(Y_t) = (1 + \theta_1{}^2 + \theta_2{}^2)\sigma_\varepsilon{}^2$$



(a) (b)

Figure 6.7 Simulated MA(2) Processes

Figure 6.8 Autocorrelation Functions of Simulated MA(2) Processes

$$\rho_1 = \frac{\theta_1 + \theta_1\theta_2}{1 + \theta_1{}^2 + \theta_2^2}$$

$$\rho_2 = \frac{\theta_2}{1 + \theta_1{}^2 + \theta_2^2}$$

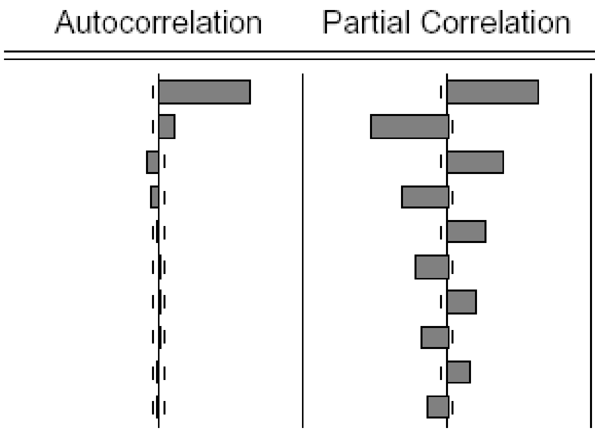
$$\rho_k = 0, k \geq 3$$

r_k decays to zero

The order q of an MA(q) process can be identified as the last k such that $\rho_k \neq 0$

Sample: 1 2000
Included observations: 1998

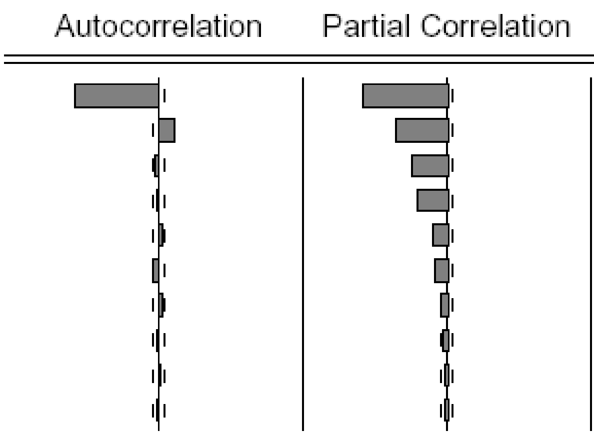
$$Y_t = 2 + 1.7\varepsilon_{t-1} + 0.72\varepsilon_{t-2} + \varepsilon_t$$



	AC	PAC
1	0.649	0.649
2	0.109	-0.542
3	-0.082	0.394
4	-0.054	-0.315
5	-0.007	0.271
6	0.016	-0.215
7	0.022	0.199
8	0.014	-0.184
9	-0.002	0.157
10	-0.010	-0.143

Sample: 1 2000
Included observations: 1998

$$Y_t = 2 - \varepsilon_{t-1} + 0.25\varepsilon_{t-2} + \varepsilon_t$$

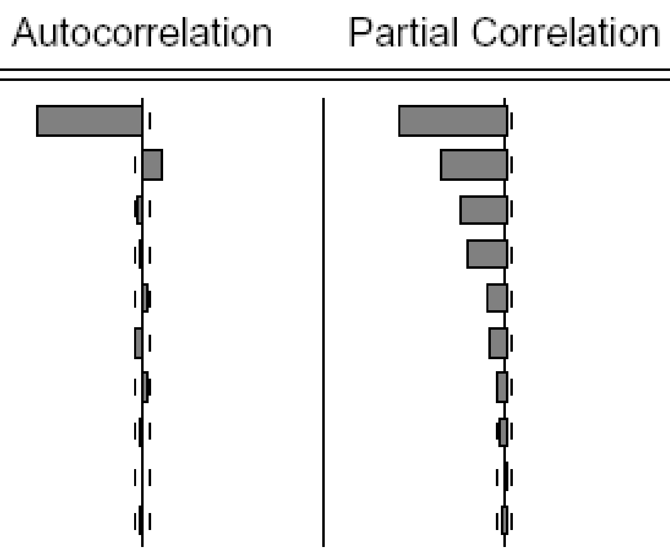


	AC	PAC
1	-0.592	-0.592
2	0.119	-0.356
3	-0.026	-0.250
4	-0.001	-0.199
5	0.031	-0.106
6	-0.034	-0.088
7	0.028	-0.043
8	-0.014	-0.022
9	0.008	-0.001
10	-0.012	-0.008

The MA(2) process is invertible if the roots of the characteristic equation are, in absolute value, greater than one.

$$1 - \theta_1 x + \theta_2 x^2 = 0$$

Sample: 1 2000
Included observations: 1998



	AC	PAC
1	-0.592	-0.592
2	0.119	-0.356
3	-0.026	-0.250
4	-0.001	-0.199
5	0.031	-0.106
6	-0.033	-0.087
7	0.027	-0.041
8	-0.014	-0.020
9	0.008	0.000
10	-0.011	-0.007

Figure 6.9 Autocorrelation Functions of MA Process $Y_t = 2 - 4\varepsilon_{t-1} + 4\varepsilon_{t-2} + \varepsilon_t$