CHAPTER 6.

FORECASTING WITH MOVING AVERAGE PROCESSES

SOLUTIONS

by
Wei Lin and Yingying Sun
(University of California, Riverside)

Exercise 1

a. Most financial returns, e.g. stocks, bonds, exchange rates, interest rates, behave very close to noise processes. This means that future returns cannot be linearly predicted from their past returns as there is no linear dependence to exploit.

b. We generate 100 observations of a white noise process $N(0, 4)$ by writing in the command line of EViews 'series whitenoise=2*nrnd'. The corresponding autocorrelograms are in Figure 1. Observe that all autocorrelation coefficients in ACF and PACF are statistically zero: the spikes are within the 95% confidence bands and the Q-statistics have very large p-values. This is the profile of a white noise process.

Sample: 1 100
Included observations: 100

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.036	0.036	0.1347	0.714
		2	0.024	0.022	0.1928	0.908
		3	-0.008	-0.010	0.1997	0.978
		4	-0.116	-0.116	1.6269	0.804
		5	-0.094	-0.087	2.5838	0.764
		6	-0.162	-0.154	5.4184	0.491
		7	0.066	0.079	5.9024	0.551
		8	0.057	0.048	6.2653	0.618
		9	-0.002	-0.029	6.2658	0.713
		10	0.000	-0.047	6.2658	0.792

Figure 1: ACF, PACF of a white noise process $N(0, 4)$

c. We choose three industries: oil, computer technology, and utility sector. The price indexes are: AMEX Oil Index (ˆXOI), NYSE ARCA Computer Technology Index (ˆXCI), and PHLX Utility Sector Index (ˆUTY). In Figure 2, we plot the daily indexes and their respective daily returns from 2004 up to July 16, 2012 for a total of 2150 observations. Though the industries are very different, the time series profiles are very similar: the indexes exhibit an upward trend until the end of 2007 approximately, followed by a steep drop during the financial crisis of 2008. Though this crisis originated in the real estate and banking sectors, it rattled over all sectors of the economy. It was a systemic event that brought substantial volatility and uncertainty to the markets. The 2008 episode of very high volatility is clearly observed in the series of returns. The technology sector recovered very nicely from the 2008 crisis, but the oil and utilities sectors have not yet reached the pre-crisis peaks. What it seems a permanent feature of the post-crisis times is a substantial increase in volatility compared to the pre-crisis times.

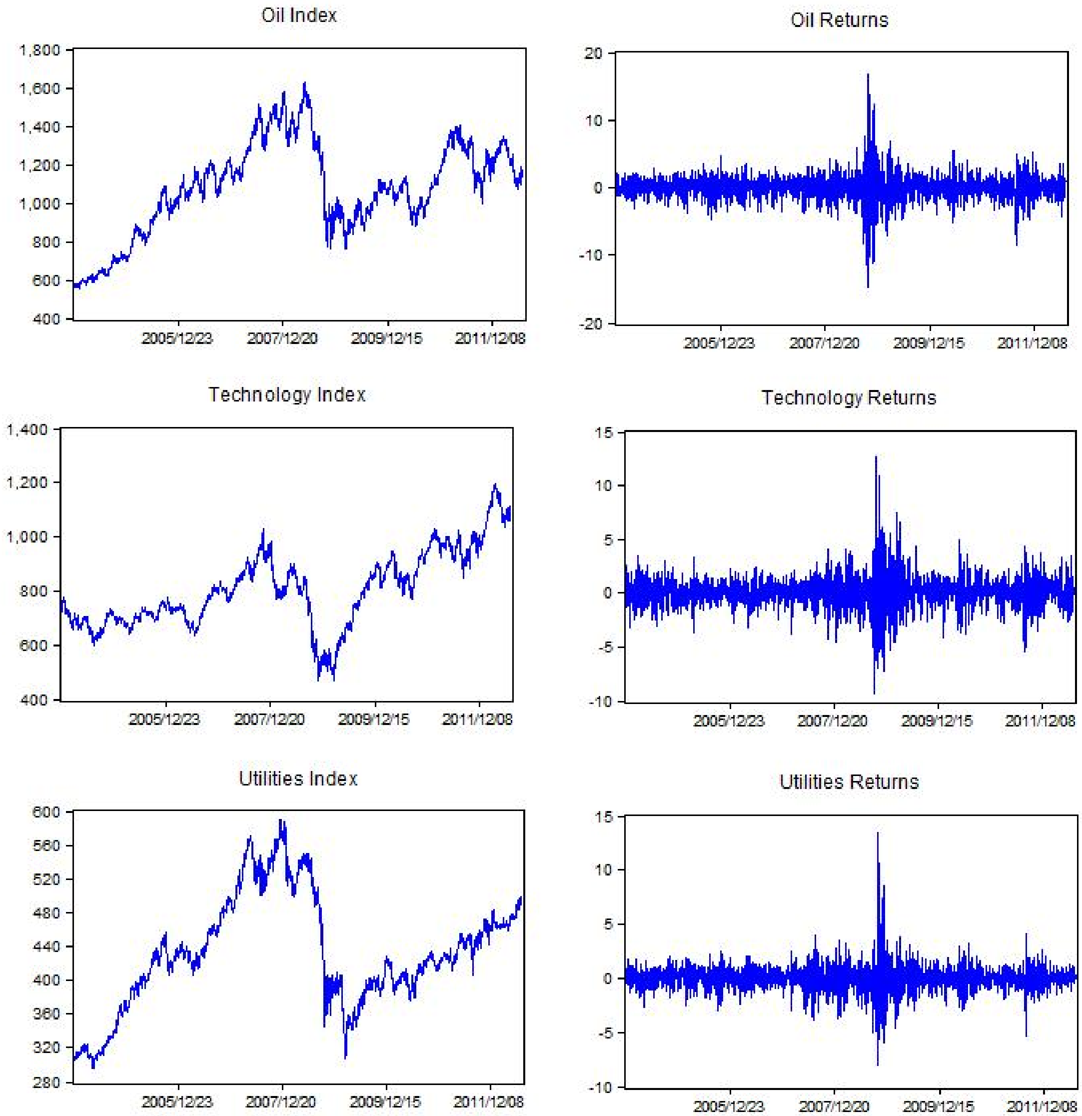
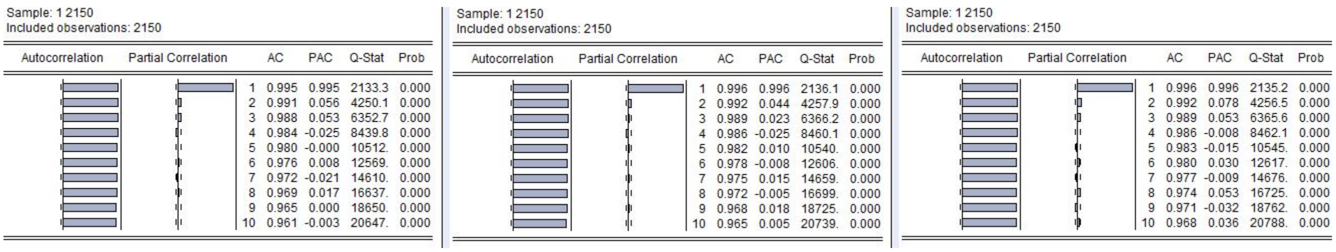


Figure 2: Time Series of Oil, Technology, and Utilities Indexes and Returns

The autocorrelation functions of the level of the indexes are presented in Figure 3 and those of their corresponding returns in Figure 4. The indexes have almost identical ACFs and PACFs. All autocorrelation coefficients are almost one and the partial autocorrelation coefficients of order one are also one and the rest are zero. These profiles are very typical of price indexes and returns in markets where public information is very fluid and it is immediately absorbed by the markets. No participant has an informational advantage to exploit before than anybody else. We call these markets 'efficient' in the sense that the best prediction of tomorrow's price is today's price.

The autocorrelograms of the series of returns are also very similar. The common feature across the three autocorrelograms is that the autocorrelations are very small to be economically significant. Only the utility sector has an autocorrelation of order one larger than 0.10 (in absolute value); the rest of the autocorrelation coefficients are below 0.1 (in absolute value). Strictly speaking, these returns do not follow the profile of white noise processes because we observe that the first two spikes are outside of the confidence bands (the standard error of the autocorrelation coefficient is very small $\sim 0.021 = 1/\sqrt{2149}$), though their values are pretty small. The statistical significance of these spikes translate into Q-statistics that reject the null hypothesis of no autocorrelation. As an exercise, we recommend splitting the sample before and after the crisis and calculate the autocorrelograms. Before the crisis, the white noise profile emerges more clearly.



Sample: 1 2150
Included observations: 2150

Lag	AC	PAC	Q-Stat	Prob
1	0.995	0.995	2133.3	0.000
2	0.991	0.056	4250.1	0.000
3	0.988	0.053	6352.7	0.000
4	0.984	-0.025	8439.8	0.000
5	0.980	-0.000	10512.	0.000
6	0.976	0.008	12569.	0.000
7	0.972	-0.021	14610.	0.000
8	0.969	0.017	16637.	0.000
9	0.965	0.000	18650.	0.000
10	0.961	-0.003	20647.	0.000

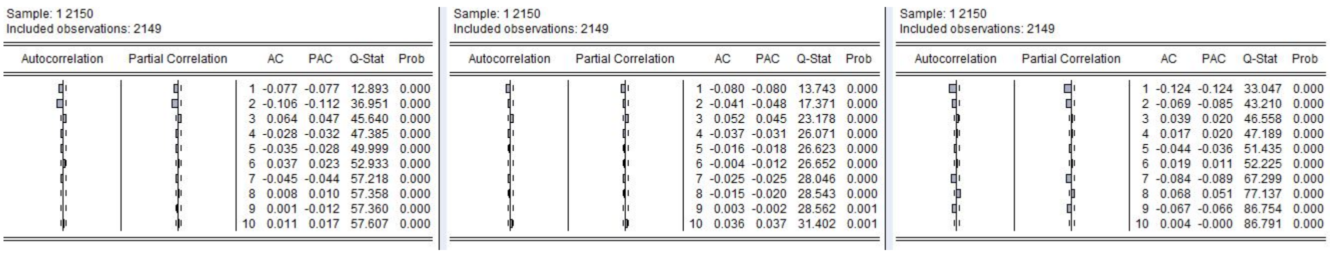
Sample: 1 2150
Included observations: 2150

Lag	AC	PAC	Q-Stat	Prob
1	0.996	0.996	2136.1	0.000
2	0.992	0.044	4257.9	0.000
3	0.989	0.023	6366.2	0.000
4	0.986	-0.025	8460.1	0.000
5	0.982	0.010	10540.	0.000
6	0.978	-0.008	12606.	0.000
7	0.975	0.015	14659.	0.000
8	0.972	-0.005	16699.	0.000
9	0.968	0.018	18725.	0.000
10	0.965	0.005	20739.	0.000

Sample: 1 2150
Included observations: 2150

Lag	AC	PAC	Q-Stat	Prob
1	0.996	0.996	2135.2	0.000
2	0.992	0.078	4256.5	0.000
3	0.989	0.053	6365.6	0.000
4	0.986	-0.008	8462.1	0.000
5	0.983	-0.015	10545.	0.000
6	0.980	0.030	12617.	0.000
7	0.977	-0.009	14676.	0.000
8	0.974	0.053	16725.	0.000
9	0.971	-0.032	18762.	0.000
10	0.968	0.036	20788.	0.000

Figure 3: ACF and PACF of Oil, Technology, and Utilities Indexes



Sample: 1 2150
Included observations: 2149

Lag	AC	PAC	Q-Stat	Prob
1	-0.077	-0.077	12.893	0.000
2	-0.106	-0.112	36.951	0.000
3	0.064	0.047	45.640	0.000
4	-0.028	-0.032	47.385	0.000
5	-0.035	-0.028	49.999	0.000
6	0.037	0.023	52.933	0.000
7	-0.045	-0.044	57.218	0.000
8	0.008	0.010	57.358	0.000
9	0.001	-0.012	57.360	0.000
10	0.011	0.017	57.607	0.000

Sample: 1 2150
Included observations: 2149

Lag	AC	PAC	Q-Stat	Prob
1	-0.080	-0.080	13.743	0.000
2	-0.041	-0.048	17.371	0.000
3	0.052	0.045	23.178	0.000
4	-0.037	-0.031	26.071	0.000
5	-0.016	-0.018	26.623	0.000
6	-0.004	-0.012	26.652	0.000
7	-0.025	-0.025	28.046	0.000
8	-0.015	-0.020	28.543	0.000
9	0.003	-0.002	28.562	0.001
10	0.036	0.037	31.402	0.001

Sample: 1 2150
Included observations: 2149

Lag	AC	PAC	Q-Stat	Prob
1	-0.124	-0.124	33.047	0.000
2	-0.069	-0.085	43.210	0.000
3	0.039	0.020	46.558	0.000
4	0.017	0.020	47.189	0.000
5	-0.044	-0.036	51.435	0.000
6	0.019	0.011	52.225	0.000
7	-0.084	-0.089	67.299	0.000
8	0.068	0.051	77.137	0.000
9	-0.067	-0.066	86.754	0.000
10	0.004	-0.000	86.791	0.000

Figure 4: ACF and PACF of Oil, Technology, and Utilities Returns

We could fit either a MA(1) or MA(2) process to the returns but we should expect small MA parameters. We present the estimation of an MA(2) model for the utilities sector, which shows slightly larger autocorrelations than those from the other two sectors. In Table 1, we present the estimation results. As we were expecting the estimates of the MA parameters are small and θ_2 is not statistically significant from zero at the conventional levels. Observe that the constant, which is the mean of the process, is also not significant. Based on this model we build a forecast up to 4 days ahead. The funcional form of the optimal forecasts for $h = 1, 2, 3, 4$ (using a quadratic loss function) is:

$f_{t,1} = 0.029 - 0.13\hat{\varepsilon}_t - 0.05\hat{\varepsilon}_{t-1}$
$f_{t,2} = 0.029 - 0.05\hat{\varepsilon}_t$
$f_{t,3} = f_{t,4} = 0.029$

Dependent Variable: RET_UTI				
Method: Least Squares				
Sample (adjusted): 2 2150				
Included observations: 2149 after adjustments				
Convergence achieved after 6 iterations				
White heteroskedasticity-consistent standard errors & covariance				
MA Backcast: 0 1				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.029745	0.021045	1.413393	0.1577
MA(1)	-0.13069	0.046831	-2.79064	0.0053
MA(2)	-0.05933	0.051709	-1.14747	0.2513
R-squared	0.021306	Mean dependent var		0.029739
Adjusted R-squared	0.020393	S.D. dependent var		1.216826
S.E. of regression	1.204354	Akaike info criterion		3.211159
Sum squared resid	3112.707	Schwarz criterion		3.219078
Log likelihood	-3447.39	Hannan-Quinn criter.		3.214056
F-statistic	23.35856	Durbin-Watson stat		2.004313
Prob(F-statistic)	0			
Inverted MA Roots	0.32	-0.19		

Table 1: Estimation of MA(2) for Utillities Sector Returns

We will need to back up $\hat{\varepsilon}_t$ and $\hat{\varepsilon}_{t-1}$ from the past histories of the returns (see Section 6.3.1.3 in the textbook). The computer software will do it for us; use the dynamic forecast option (under the Proc tab) and set the forecasting horizon from observation 2151 to 2154 to obtain the values of the forecasts:

Forecast horizon	Forecast	St. error	95% confidence int.
2012.07.17 (h=1)	-0.0195	1.2056	[-2.382, 2.343]
2012.07.18 (h=2)	0.0337	1.2148	[-2.347, 2.414]
2012.07.19 (h=3)	0.0297	1.2168	[-2.324, 2.414]
2012.07.20 (h=4)	0.0297	1.2168	[-2.324, 2.414]

The corresponding forecast errors will be computed once we know the realized values from 2012.07.17 to 2012.07.20 (unknown at the time of this writing) but we can compute the variance of the forecast errors:

$$E(e_{t,1}^2) = \sigma_{\hat{\varepsilon}}^2 \sim 1.204^2$$
$$E(e_{t,2}^2) = \sigma_{\hat{\varepsilon}}^2(1 + \hat{\theta}_1^2) \sim 1.215^2$$
$$E(e_{t,3}^2) = E(e_{t,4}^2) = \hat{\sigma}_{ret}^2 = 1.216^2.$$

The third column of the above table reports the standard error of the forecasts, i.e $\hat{\sigma}_{t+h|t}$, which is the square root of the variances that we have just calculated. Under the assumption of normality, we construct a 95 % confidence interval for the forecast as $f_{t,h} \pm 1.96\hat{\sigma}_{t+h|t}$, which is reported in the last column of the above table. Observe that the intervals are very wide, going from negative to positive values and including zero. This reveals that it is very difficult to forecast returns based on their past history, and that a sensible process to characterize returns is a white noise process, though we may find small autocorrelation of MA type that it may be due to some micro structure features of the markets (see Chapter 5).

Exercise 2
The only difference between the two MA processes is the value of the moving average parameter θ. Observe that they are inversely related, i.e., 0.8 is the inverse of 1.25. The autocorrelation of order 1 for the first process is $\rho_1 = \theta/(1+\theta^2) = 0.8/(1+0.8^2) = 0.4878$, and for the second

$\rho_1 = 1.25/(1+1.25^2) = 0.4878$. Thus, both autocorrelation coefficients are identical; in addition the autocorrelations of higher order than one are all zero because the processes are MA(1). Both processes have identical autocorrelation functions. The invertible process is $y_t = 1.2 + 0.8\varepsilon_{t-1} + \varepsilon_t$ because $0.8 < 1$. In Figures 5 and 6, we plot the autocorrelation functions of both simulated processes for a sample size of 100 observations and assuming a white noise process $\varepsilon_t \rightarrow N(0, 0.25)$. Observe that the autocorrelation functions are very similar but they are also noisy because the sample is very small. The estimated $\hat{\rho}_1 = 0.306$ is not very close to the theoretical value of $\rho_1 = 0.487$. Increase the sample size, say to 1000 observations, and observe that the sample autocorrelograms are almost identical to the theoretical autocorrelograms of a MA(1) process.

Sample: 1 100
Included observations: 99

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.306	0.306	9.5489	0.002
		2	-0.364	-0.505	23.202	0.000
		3	-0.157	0.244	25.783	0.000
		4	0.146	-0.096	28.034	0.000
		5	0.040	-0.027	28.203	0.000
		6	-0.108	-0.017	29.456	0.000
		7	0.050	0.134	29.723	0.000
		8	0.127	-0.054	31.496	0.000
		9	-0.021	0.025	31.545	0.000
		10	-0.022	0.097	31.601	0.000

Figure 5: ACF, PACF of $y_t = 1.2 + 0.8\varepsilon_{t-1} + \varepsilon_t$

Sample: 1 100
Included observations: 99

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.307	0.307	9.6029	0.002
		2	-0.369	-0.511	23.610	0.000
		3	-0.159	0.251	26.252	0.000
		4	0.148	-0.105	28.562	0.000
		5	0.039	-0.025	28.723	0.000
		6	-0.110	-0.017	30.025	0.000
		7	0.046	0.128	30.254	0.000
		8	0.126	-0.052	32.011	0.000
		9	-0.025	0.016	32.079	0.000
		10	-0.027	0.101	32.161	0.000

Figure 6: ACF, PACF of $y_t = 1.2 + 1.25\varepsilon_{t-1} + \varepsilon_t$

Exercise 3

The optimal forecasts for $h = 1, 2, 3$ are:

$f_{t,1} = 2.3 - 0.95\varepsilon_t = 2.3 - 0.95 \times 0.4 = 1.92$
$f_{t,2} = f_{t,3} = 2.3$

with corresponding forecast errors:

$e_{t,1} = \varepsilon_t$
$e_{t,2} = -0.95\varepsilon_{t+1} + \varepsilon_{t+2}$
$e_{t,3} = -0.95\varepsilon_{t+2} + \varepsilon_{t+3}$.

The variances of the forecast errors are:

$E(e_{t,1}^2) = \sigma_\varepsilon^2 = 1$
$E(e_{t,2}^2) = E(e_{t,3}^2) = (1 + 0.95^2)\sigma_\varepsilon^2 = 1.90.$

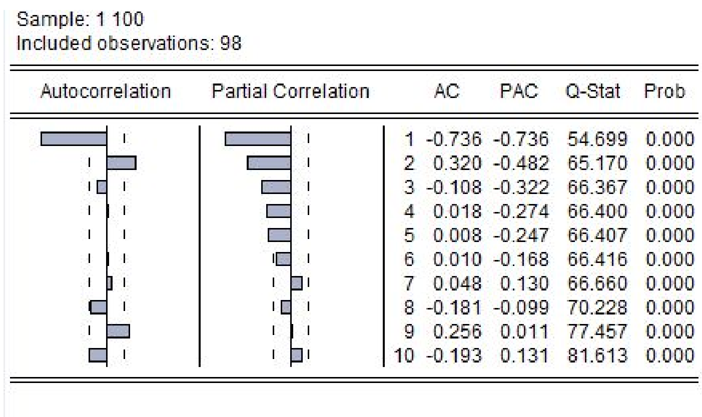
The density forecasts are:

$f(Y_{t+1}|I_t) \to N(1.92, 1)$
$f(Y_{t+2}|I_t) = f(Y_{t+3}|I_t) \to N(2.3, 1.90).$

Exercise 4
a. The theoretical autocorrelations up to order 10 are:

$\rho_1 = (\theta_1\theta_2 + \theta_1)/(1 + \theta_1^2 + \theta_2^2) = ((-2) \times 1.35 - 2)/(1 + (-2)^2 + 1.35^2) = -0.688$
$\rho_2 = \theta_2/(1 + \theta_1^2 + \theta_2^2) = 1.35/(1 + (-2)^2 + 1.35^2) = 0.197$
$\rho_3 = \rho_4 = ... = \rho_{10} = 0.$

b. In Figure 7, we plot the sample autocorrelograms of the simulated MA(2) process for 100 observations and a white noise process ε_t normally distributed with mean zero and variance one.



Sample: 1 100
Included observations: 98

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.736	-0.736	54.699	0.000
		2	0.320	-0.482	65.170	0.000
		3	-0.108	-0.322	66.367	0.000
		4	0.018	-0.274	66.400	0.000
		5	0.008	-0.247	66.407	0.000
		6	0.010	-0.168	66.416	0.000
		7	0.048	0.130	66.660	0.000
		8	-0.181	-0.099	70.228	0.000
		9	0.256	0.011	77.457	0.000
		10	-0.193	0.131	81.613	0.000

Figure 7: ACF, PACF of $y_t = 0.7 - 2\varepsilon_{t-1} + 1.35\varepsilon_{t-2} + \varepsilon_t$

The sample autocorrelogram has two significant spikes: the sample autocorrelation of order 1 is negative $\hat{\rho}_1 = -0.736$, which is close to the theoretical value; and the sample autocorrelation of order 2 $\hat{\rho}_2 = 0.32$, which has the same sign as the theoretical autocorrelation but it is larger. This is due to the small sample size. Increase the sample size to about 1000 observations, and the sample autocorrelations will converge to the theoretical values. The shapes of both autocorrelograms are the distinctive profiles of an MA(2) process: two significant spikes in the ACF and a slow decay towards zero in the PACF.

Exercise 5
a. We present a simulated time series from the MA(2) process $y_t = 0.7 - 2\varepsilon_{t-1} + 1.35\varepsilon_{t-2} + \varepsilon_t$ in Figure 8. The series is very ragged because of the negative autocorrelation of order one. The unconditional mean of the process is 0.7 (black horizontal line); the process fluctuates around this central value, thus it is clearly first-order stationary. Since we have generated the data, we know that it is also covariance stationary.

With this time series, we proceed to estimate a MA(2) model. We present the estimation output in Table 2. It is surprising to find that the estimated parameters are very different from the theoretical values. The estimated model is $y_t = 0.7 - 1.52\varepsilon_{t-1} + 0.74\varepsilon_{t-2} + \varepsilon_t$. We should expect some small

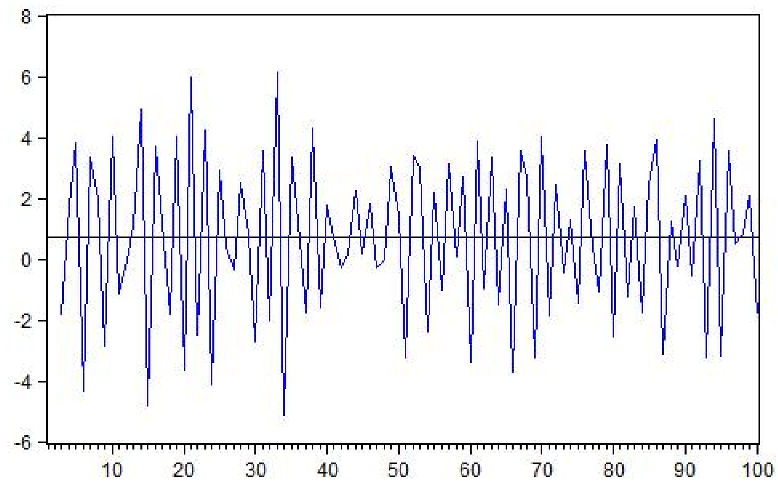


Figure 8: Simulated time series from $y_t = 0.7 - 2\varepsilon_{t-1} + 1.35\varepsilon_{t-2} + \varepsilon_t$

differences with the theoretical values because the sample size is small, but the differences that we observe are too large. However, the estimated autocorrelations are the same as those calculated in Exercise 4: $\hat{\rho}_1 = (-1.524 \times 1.744)/3.876 = -0.685$ and $\hat{\rho}_2 = 0.744/3.876 = 0.191$. What we see here again is the invertibility issue. We have two representations of a MA(2) process, invertible and non-invertible, which are indistinguishable from the autocorrelation functions as these are identical for both representations.

Dependent Variable: YY2				
Method: Least Squares				
Sample (adjusted): 3 100				
Included observations: 98 after adjustments				
Convergence achieved after 14 iterations				
MA Backcast: 1 2				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.71964	0.030125	23.88847	0
MA(1)	-1.524302	0.071434	-21.33859	0
MA(2)	0.744076	0.072296	10.292	0
R-squared	0.757944	Mean dependent var		0.694942
Adjusted R-squared	0.752848	S.D. dependent var		2.710594
S.E. of regression	1.347555	Akaike info criterion		3.464595
Sum squared resid	172.511	Schwarz criterion		3.543727
Log likelihood	-166.7652	Hannan-Quinn criter.		3.496602
F-statistic	148.7354	Durbin-Watson stat		2.026077
Prob(F-statistic)	0			
Inverted MA Roots	.76-.40i	.76+.40i		

Table 2: Estimation of a MA(2) process with simulated time series

We need a further concept to assess invertibility in MA(2) processes. A MA(2) is invertible if the roots of its characteristics equation are larger than one. Write the theoretical process using the lag operator, i.e. $y_t = 0.7 + (1 - 2L + 1.35L^2)\varepsilon_t$. The characteristic equation is $1 - 2x + 1.35x^2 = 0$; see that we have substituted the lag operator by a 'dummy' variable x. Solving for the roots of this quadratic equation, we have that $x_1 = 0.740 + 0.438i$ and $x_2 = 0.740 - 0.438i$ where i is the imaginary number $i = \sqrt{-1}$. With imaginary roots we need to compute the modulus, i.e. $m = 0.740^2 + 0.438^2 = 0.740$. Since the modulus is less than one, this theoretical process is non-invertible. Now, if we implement the same procedure for the estimated model $y_t = 0.7 - 1.52\varepsilon_{t-1} + 0.74\varepsilon_{t-2} + \varepsilon_t$, we will obtain imaginary roots with modulus greater than one, which means that this is the invertible representation. With some additional algebra, and given the non-invertible representation we can always find the invertible representation, and vice versa. The EViews output provides the invertible representation, which we will use to produce the forecasts.

b. From the invertible MA(2) process $y_t = 0.7 - 1.52\varepsilon_{t-1} + 0.74\varepsilon_{t-2} + \varepsilon_t$, we calculate the forecasts for $h = 1, 2, 3$:

$f_{t,1} = 0.7 - 1.52\varepsilon_t + 0.74\varepsilon_{t-1} = 0.7 - 1.52 \times 0.4 + 0.74 \times (-1.2) = -0.796$
$f_{t,2} = 0.7 + 0.74\varepsilon_t = 0.7 + 0.74 \times 0.4 = 0.994$
$f_{t,3} = 0.7$

Exercise 6
The process $y_t = 1.2 + 0.8\varepsilon_{t-1} + \varepsilon_t$ can be written as $y_t - 1.2 = (1 + 0.8L)\varepsilon_t$, which implies that $\frac{y_t - 1.2}{(1-(-0.8L))} = \varepsilon_t$. Since $0.8 < 1$, the ratio $1/(1 - (-0.8L))$ is the limit of the following sequence $\frac{1}{1-(-0.8L)} = 1 - 0.8L + 0.8^2L^2 - 0.8^3L^3 + ...$ Thus, $(y_t - 1.2)(1 - 0.8L + 0.8^2L^2 - 0.8^3L^3 + ...) = \varepsilon_t$. Then, the autoregressive representation is

$$(y_t - 1.2) = 0.8(y_{t-1} - 1.2) - 0.8^2(y_{t-2} - 1.2) + 0.8^3(y_{t-3} - 1.2) - ... + \varepsilon_t$$

The process $y_t = 1.2 + 1.25\varepsilon_{t-1} + \varepsilon_t$ is non-invertible, i.e $\theta = 1.25 > 1$. There is not limit to the ratio $\frac{1}{1-(-1.25L)}$ but we can transform this ratio by using the forward operator $F = 1/L$,

$$\frac{1}{1 - (-1.25L)} = \frac{1}{1 - (-1/0.8F)} = \frac{0.8F}{1 - (-0.8F)} = 0.8F(1 - 0.8F + 0.8^2F^2 - 0.8^3F^3 + ...)$$

Then, the autoregressive representation is

$$(y_t - 1.2) \times 0.8F(1 - 0.8F + 0.8^2F^2 - 0.8^3F^3 + ...) = \varepsilon_t$$

which makes y_t a function of the future values $y_{t+1}, y_{t+2}, ...$ Obviously, this representation is not useful for forecasting purposes. Thus, we choose the invertible process $y_t = 1.2 + 0.8\varepsilon_{t-1} + \varepsilon_t$ because the present is a function of the past.

Exercise 7
In Figure 9, we have updated the monthly time series of the 5-year constant maturity yield on US Treasury securities; the sample runs from April 1953 to June 2012. The upper plot is the level in percentage, the middle plot is the monthly change, and the lower plot is the *percentage* monthly change. The lower plot is comparable to that in Figure 6.5 in the textbook. There are two features in the updated sample: from April 2008 to June 2012, the yield has been going down as the result of the monetary policy of the Fed, which has maintained a low rate environment to stimulate the economy; the second feature in the updated sample is the increase in volatility from 2008 to now, which is most obvious in the lower plot.

Regarding the linear dependence of the updated time series, we do not observe any major change. The autocorrelation functions corresponding to the time series in the middle and lower plots of Figure 9 is in Figure 10. There is a prominent autocorrelation of order one (around 0.3) in the

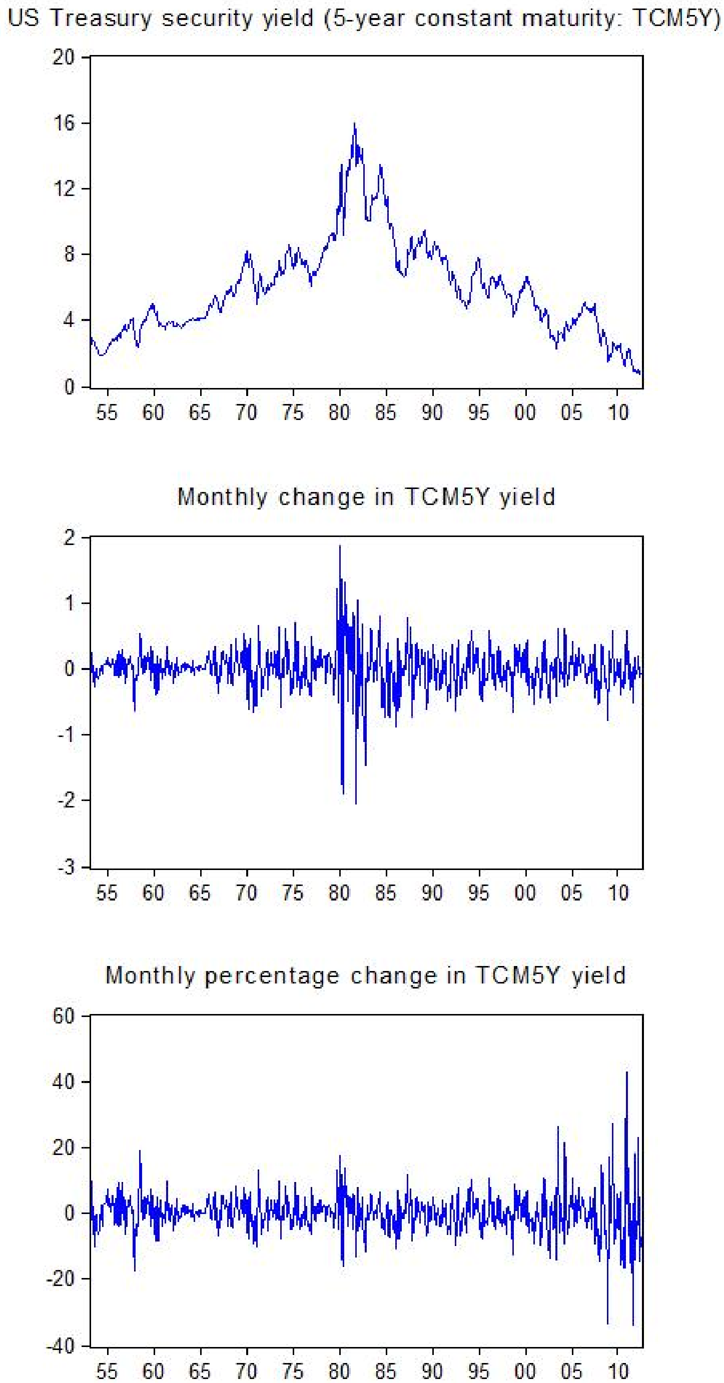
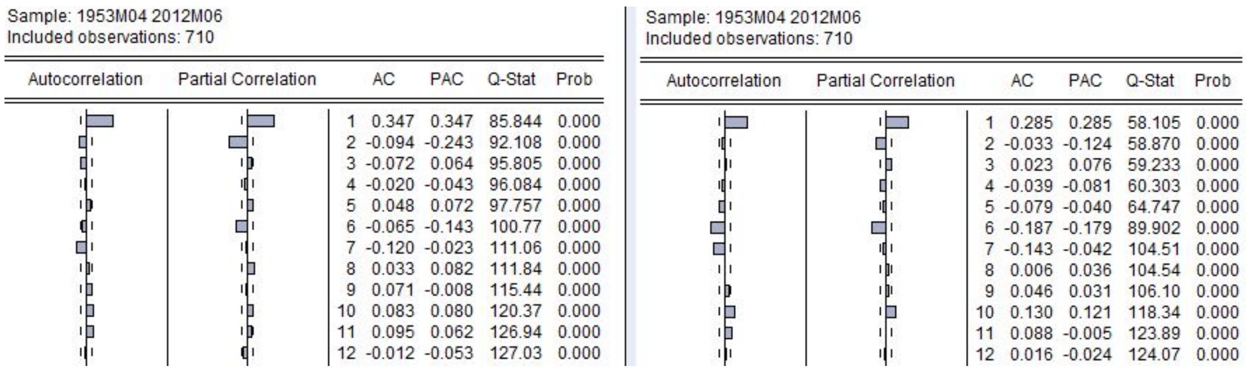


Figure 9: 5-year constant maturity yield on US Treasury securities

ACFs and a decay towards zero in the profile of the PACFs. This indicates that a MA(1) model will be sufficient to capture the linear dependence of the time series. Observe that this is very similar to the information in Figure 6.5 and Table 6.1 in the textbook.



Sample: 1953M04 2012M06
Included observations: 710

	AC	PAC	Q-Stat	Prob
1	0.347	0.347	85.844	0.000
2	-0.094	-0.243	92.108	0.000
3	-0.072	0.064	95.805	0.000
4	-0.020	-0.043	96.084	0.000
5	0.048	0.072	97.757	0.000
6	-0.065	-0.143	100.77	0.000
7	-0.120	-0.023	111.06	0.000
8	0.033	0.082	111.84	0.000
9	0.071	-0.008	115.44	0.000
10	0.083	0.080	120.37	0.000
11	0.095	0.062	126.94	0.000
12	-0.012	-0.053	127.03	0.000

Sample: 1953M04 2012M06
Included observations: 710

	AC	PAC	Q-Stat	Prob
1	0.285	0.285	58.105	0.000
2	-0.033	-0.124	58.870	0.000
3	0.023	0.076	59.233	0.000
4	-0.039	-0.081	60.303	0.000
5	-0.079	-0.040	64.747	0.000
6	-0.187	-0.179	89.902	0.000
7	-0.143	-0.042	104.51	0.000
8	0.006	0.036	104.54	0.000
9	0.046	0.031	106.10	0.000
10	0.130	0.121	118.34	0.000
11	0.088	-0.005	123.89	0.000
12	0.016	-0.024	124.07	0.000

Figure 10: ACF and PACF of changes (left) and percentage changes (right) in TCM5Y yield

We proceed to estimate a MA(1) model for the series of changes in the yield of the TCM5Y securities and the estimation output is shown in Table 3. The estimated MA parameter, i.e. $\hat{\theta}_1 = 0.497$ is statistically significant from zero at any conventional level, and the mean of the process is statistically zero. Consequently, the optimal forecasts are: $f_{t,1} = 0.497\hat{\varepsilon}_t = -0.021$, which implies a prediction for July 2012 of a yield of $0.71 - 0.021 = 0.689\%$ and $f_{t,h} = 0$ for all $h > 1$. The standard error of these forecasts are: $\hat{\sigma}_{t+1|t} = 0.294$ and $\hat{\sigma}_{t+h|t} = 0.328$ for all $h > 1$. With this information we can construct any confidence interval for the forecast. For instance, the 95% confidence interval (under the assumption of normality) for one-month ahead forecast is $f_{t,1} \pm 1.96 \times \hat{\sigma}_{t+1|t} = -0.021 \pm 1.96 \times 0.294 = [-0.597, 0.555]$.

Dependent Variable: D5Y				
Method: Least Squares				
Sample (adjusted): 1953M05 2012M06				
Included observations: 710 after adjustments				
Convergence achieved after 6 iterations				
White heteroskedasticity-consistent standard errors & covariance				
MA Backcast: 1953M04				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.002686	0.016513	-0.162678	0.8708
MA(1)	0.497664	0.051536	9.656565	0
R-squared	0.175469	Mean dependent var		-0.00269
Adjusted R-squared	0.174305	S.D. dependent var		0.323379
S.E. of regression	0.293847	Akaike info criterion		0.3913
Sum squared resid	61.13314	Schwarz criterion		0.40416
Log likelihood	-136.9115	Hannan-Quinn criter.		0.396268
F-statistic	150.6703	Durbin-Watson stat		2.06991
Prob(F-statistic)	0			
Inverted MA Roots	-0.5			

Table 3: Estimation of MA(1) for changes in the TCM5Y yield

Exercise 8

We download the 1-year and 20-year constant maturity Treasury yields, i.e. TCM1Y and TCM20Y respectively, from the Board of Governors of the Federal Reserve System databases:

http://www.federalreserve.gov/releases/h15/data.htm.

In Figure 11 we plot both time series. The series TCM20Y was discontinued from January 1987 to September 1993. Most of the time, the yield on the longer maturity is larger than the yield on the short maturity since investors expect a higher reward for holding long run versus short run assets. However, there are times in which we observe an inversion, i.e. the shorter maturity has larger yield than the longer maturity. This inversion is considered a leading indicator of recessions; for instance observe that in the months prior to the financial crisis of 2008, both securities were offering similar yields.

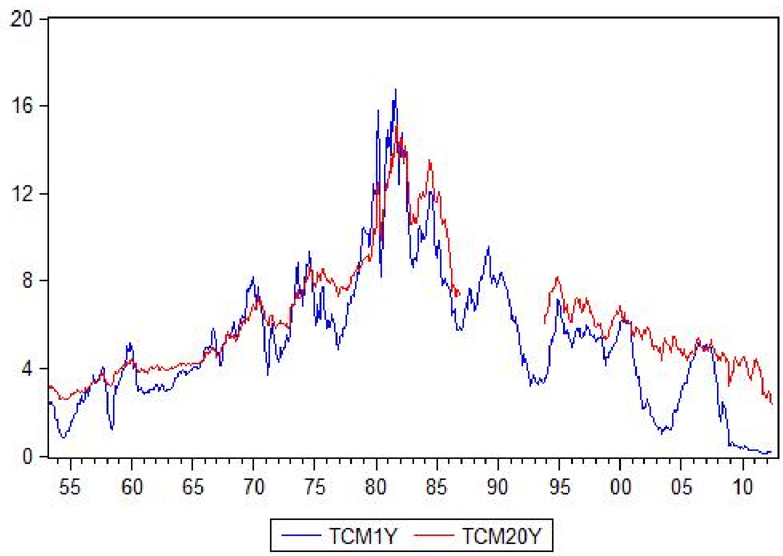


Figure 11: 1-year and 20-year constant maturity yields on US Treasury securities

In Figure 12 we plot the time series of the changes in the yields, which profiles are very similar to the TCM5Y yields; and in Figure 13 we present their corresponding autocorrelation functions. For TCM20Y we only consider the last part of the sample from October 1993 on. In both series, the profiles of the ACFs and PACFs are very similar: the autocorrelation coefficients are small and the best linear model is a MA(1), which is the same model that we propose for TCM5Y yields in Exercise 7.

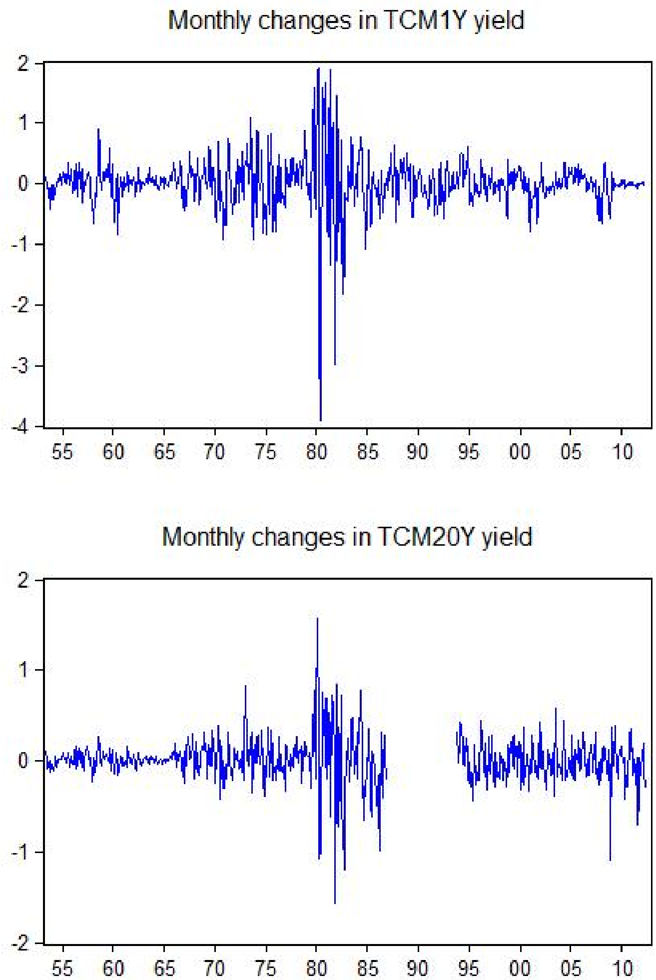


Figure 12: Monthly changes in 1-year and 20-year constant maturity yields on US Treasury securities

Sample: 1953M04 2012M10
Included observations: 710

Lag	AC	PAC	Q-Stat	Prob
1	0.364	0.364	94.486	0.000
2	-0.075	-0.239	98.509	0.000
3	-0.106	0.019	106.50	0.000
4	-0.049	-0.033	108.24	0.000
5	0.080	0.110	112.85	0.000
6	-0.113	-0.251	121.97	0.000
7	-0.162	0.015	140.91	0.000
8	0.066	0.131	144.04	0.000
9	0.113	-0.011	153.24	0.000
10	0.077	0.016	157.56	0.000
11	0.087	0.135	162.97	0.000
12	-0.049	-0.135	164.74	0.000

Sample: 1993M10 2012M10
Included observations: 224

Lag	AC	PAC	Q-Stat	Prob
1	0.205	0.205	9.5599	0.002
2	-0.075	-0.122	10.845	0.004
3	0.069	0.118	11.939	0.008
4	0.031	-0.023	12.156	0.016
5	-0.148	-0.141	17.220	0.004
6	-0.121	-0.063	20.595	0.002
7	-0.052	-0.048	21.236	0.003
8	0.005	0.035	21.242	0.007
9	0.023	0.028	21.369	0.011
10	0.019	0.003	21.452	0.018
11	0.012	-0.013	21.487	0.029
12	-0.083	-0.119	23.147	0.027

Figure 13: ACF and PACF of changes in 1-year (left) and 20-year (right) constant maturity yields on US Treasury securities

Exercise 9

In the upper panel of Figure 14 we plot the daily price of Microsoft and the corresponding return from January 8, 2003 to July 20, 2012 for a total of 2402 observations; in the lower panel, we plot the smoothed price based on a 4-day moving average, i.e. $p_t^{sm} = (p_t + p_{t-1} + p_{t-2} + p_{t-3})/4$, and the corresponding return. The smoothed price removes short term volatility but the long run features persist. Volatility is different when we compare the returns based on the close price with the returns based on the smoothed price. The former returns have a range from -12% to almost 20% while the smoothed returns move between -4% and 4%.

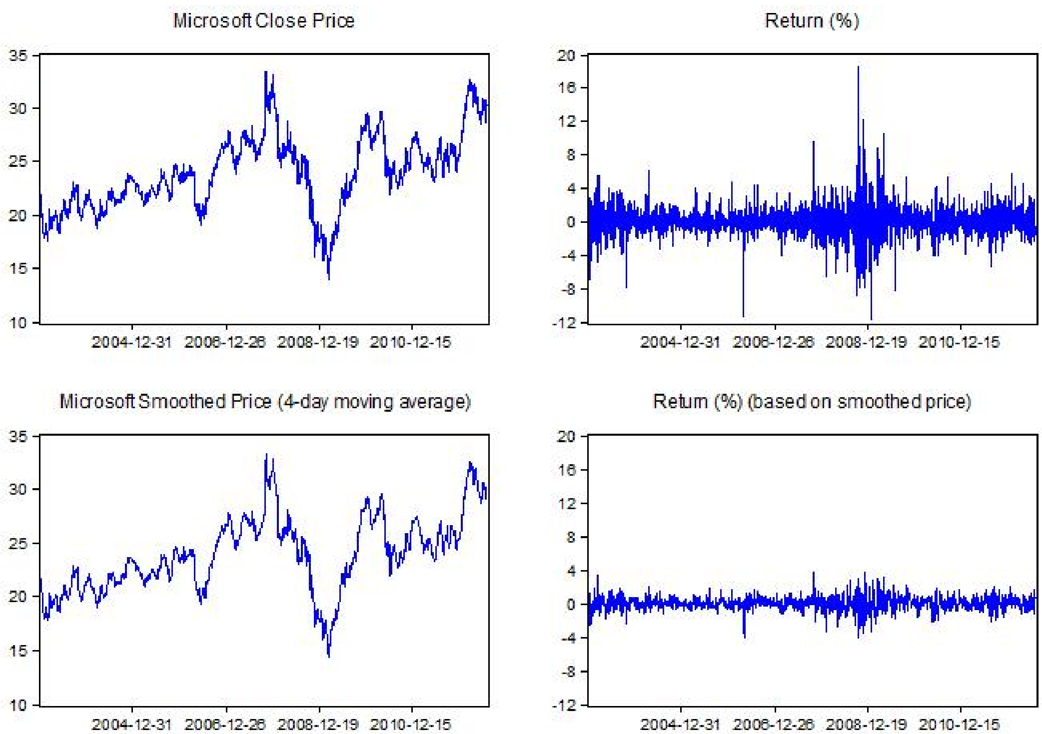


Figure 14: Microsoft prices and returns

In Figure 15 we present the autocorrelation functions of the returns and the smoothed returns. The autocorrelations of the returns are practically zero and their profiles are very close to a white noise process, but the smoothed returns are very autocorrelated. The autocorrelations of the smoothed returns follow a MA(3) process as we see three significant spikes in the ACF and partial autocorrelations decaying towards zero in the PACF. This is the result of smoothing the price by constructing a 4-day moving average price.

Sample: 1 2402
Included observations: 2401

	AC	PAC	Q-Stat	Prob
1	-0.091	-0.091	19.961	0.000
2	-0.045	-0.053	24.774	0.000
3	0.051	0.042	30.946	0.000
4	-0.043	-0.038	35.492	0.000
5	-0.035	-0.038	38.397	0.000
6	0.013	0.000	38.787	0.000
7	-0.056	-0.055	46.333	0.000
8	-0.017	-0.026	47.060	0.000
9	0.000	-0.013	47.060	0.000
10	0.033	0.034	49.669	0.000
11	0.031	0.034	51.926	0.000
12	0.026	0.030	53.506	0.000
13	-0.023	-0.019	54.787	0.000
14	-0.006	-0.011	54.879	0.000
15	0.005	0.002	54.942	0.000

Sample: 1 2402
Included observations: 2398

	AC	PAC	Q-Stat	Prob
1	0.692	0.692	1150.3	0.000
2	0.443	-0.069	1621.6	0.000
3	0.211	-0.134	1728.2	0.000
4	-0.090	-0.322	1747.5	0.000
5	-0.077	0.338	1761.8	0.000
6	-0.074	-0.048	1775.1	0.000
7	-0.082	-0.079	1791.5	0.000
8	-0.037	-0.115	1794.8	0.000
9	0.014	0.273	1795.2	0.000
10	0.052	-0.021	1801.7	0.000
11	0.071	-0.084	1813.9	0.000
12	0.057	-0.114	1821.6	0.000
13	0.024	0.174	1823.0	0.000
14	0.011	0.026	1823.3	0.000
15	-0.005	-0.090	1823.3	0.000

Figure 15: Microsoft returns (left) and smoothed returns (right)

Exercise 10

Here we present our favorite stock: Apple Inc. (AAPL). We download daily prices from October 17, 2007 to July 20, 2012. The time series of prices and corresponding returns is presented in Figure 16. Looking at the time series of prices, it is obvious why this stock could be a favorite of many investors. It has triplicated its value, from about \$200 at the end of 2007 to \$600 in July 2012. However, the stock was not immune to the crisis of 2008, showing high volatility during that time. In Figure 17, we present the autocorrelation functions of the returns, and it is very obvious that the process is white noise and it will not be possible to predict these returns on the basis of past information (see the Q-statistics and their p-values). Consequently, our best forecast would be the unconditional mean of the returns. In Figure 18, we have the histogram and other descriptive statistics of the returns. The daily return average is 0.134 % and the standard deviation is 2.391. We would like to test the null hypothesis $H_0 : \mu = 0$ versus $H_1 : \mu \neq 0$. We construct a t-ratio, i.e. $t = 0.134/\sqrt{2.391^2/1200} = 1.941$, from which we conclude that the unconditional mean is not statistically different from zero at the 5% significance level. Consequently, our best forecast at any horizon for the daily return of Apple is zero, which implies that our best forecast for tomorrow's price is today's price.

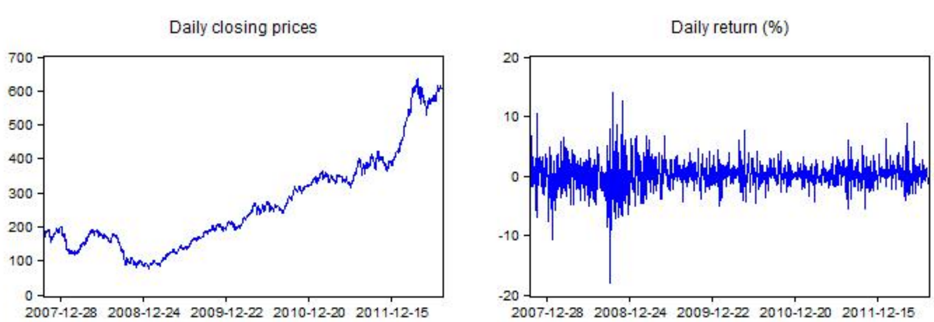


Figure 16: Apple Inc. daily prices (left) and returns (right)

Sample: 200 1399
Included observations: 1200

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.015	-0.015	0.2654	0.606
		2	-0.037	-0.037	1.9010	0.387
		3	0.002	0.001	1.9066	0.592
		4	0.068	0.067	7.5037	0.112
		5	0.010	0.013	7.6326	0.178
		6	-0.060	-0.055	11.997	0.062
		7	-0.013	-0.014	12.197	0.094
		8	0.025	0.017	12.982	0.112
		9	-0.010	-0.012	13.107	0.158
		10	0.017	0.025	13.439	0.200
		11	0.001	0.004	13.441	0.265
		12	0.053	0.049	16.827	0.156
		13	0.021	0.022	17.342	0.184
		14	0.004	0.008	17.365	0.237
		15	-0.039	-0.039	19.174	0.206

Figure 17: ACF and PACF of Apple daily returns

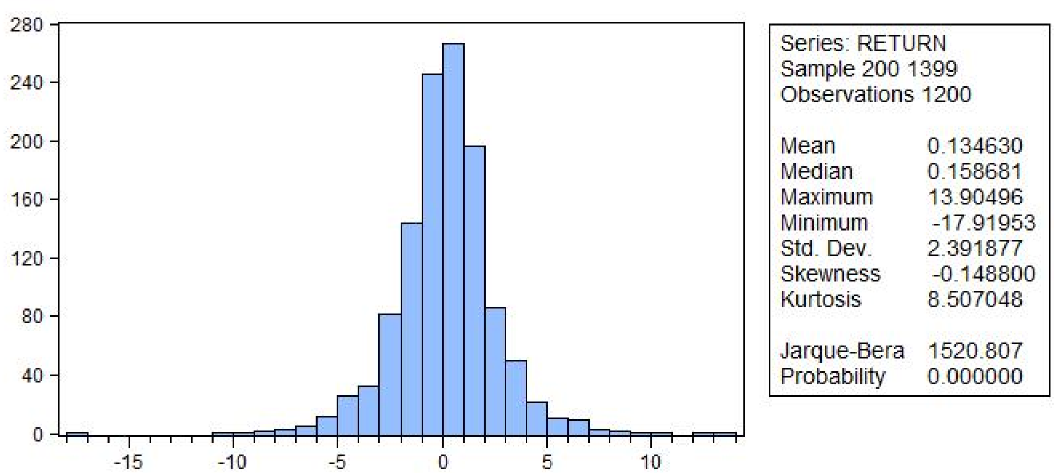


Figure 18: Histogram of Apple daily returns