CHAPTER 4.

TOOLS OF THE FORECASTER

SOLUTIONS

by

Wei Lin and Yingying Sun

(University of California, Riverside)

Note: The house price index for Exercises 1 to 4 is different from the house price index presented in the textbook (Section 4.1.1 and Table 4.1). Both are downloaded from Freddie Mac's website. In the textbook, the time series is the Conventional Mortgage Home Price Index (CMHPI), which is a weighted average of nine census region indexes. In February 2011, Freddie Mac discontinued the publication of CMHPI and replaced it with the Freddie Mac House Price Index (FMHPI) http://www.freddiemac.com/finance/fmhpi. FMHPI is a weighted average of state indexes. The base month of the index is December 2000, i.e. FMHPI = 100. Consequently, the regression results in Table 4.1 of the textbook are not comparable to the regression results presented in Exercises 1 to 4.

Exercise 1

We present the ACFs and PACFs of house prices (P), interest rates (R) and their changes (DP and DR) in Figures 1 to 4. The time series in levels (P and R) have similar profiles: strong autocorrelations that decay slowly towards zero in the ACFs, and a first order prominent spike, larger than 0.90 and statistically significant from zero, in the PACFs (observe the confidence bands for the null hypothesis of zero autocorrelation). The time series of changes in prices and rates exhibit much weaker autocorrelation. However, the two series are very different. While changes in interest rates are not autocorrelated (a barely significant autocorrelation of order one), changes in house prices exhibit short term (about 2 years) autocorrelation. This means that we will be able to exploit time dependence in yearly price growth for forecasting purposes but it will be difficult to do so in forecasting interest rate changes.

Sample: 1975 2011
Included observations: 37

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.944	0.944	35.750	0.000
		2	0.865	-0.252	66.573	0.000
		3	0.772	-0.118	91.894	0.000
		4	0.670	-0.119	111.51	0.000
		5	0.560	-0.099	125.64	0.000
		6	0.452	-0.019	135.15	0.000
		7	0.354	0.021	141.19	0.000
		8	0.269	0.029	144.80	0.000
		9	0.196	-0.000	146.78	0.000
		10	0.133	-0.023	147.72	0.000

Figure 1: ACF, PACF and Q-Statistic of P

Copyright ©2013 Pearson Education, Inc.

Sample: 1975 2011
Included observations: 37

Lag	AC	PAC	Q-Stat	Prob
1	0.906	0.906	32.918	0.000
2	0.772	-0.276	57.488	0.000
3	0.645	0.015	75.149	0.000
4	0.527	-0.052	87.302	0.000
5	0.434	0.058	95.786	0.000
6	0.366	0.033	102.02	0.000
7	0.294	-0.122	106.17	0.000
8	0.222	-0.017	108.63	0.000
9	0.149	-0.069	109.78	0.000
10	0.082	-0.002	110.14	0.000

Figure 2: ACF, PACF and Q-Statistic of R

Sample: 1975 2011
Included observations: 36

Lag	AC	PAC	Q-Stat	Prob
1	0.751	0.751	22.071	0.000
2	0.390	-0.400	28.201	0.000
3	0.038	-0.196	28.262	0.000
4	-0.222	-0.104	30.371	0.000
5	-0.308	0.079	34.545	0.000
6	-0.326	-0.181	39.388	0.000
7	-0.291	-0.064	43.386	0.000
8	-0.257	-0.138	46.623	0.000
9	-0.222	-0.046	49.116	0.000
10	-0.150	-0.007	50.300	0.000

Figure 3: ACF, PACF and Q-Statistic of DP

Sample: 1975 2011
Included observations: 36

Lag	AC	PAC	Q-Stat	Prob
1	0.331	0.331	4.2859	0.038
2	-0.033	-0.161	4.3304	0.115
3	0.018	0.097	4.3444	0.227
4	-0.183	-0.267	5.7816	0.216
5	-0.231	-0.067	8.1359	0.149
6	-0.068	-0.008	8.3475	0.214
7	-0.040	-0.034	8.4211	0.297
8	-0.007	0.006	8.4234	0.393
9	0.002	-0.079	8.4237	0.492
10	0.079	0.099	8.7495	0.556

Figure 4: ACF, PACF and Q-Statistic of DR

Exercise 2
We estimate the following regression models

$$
\begin{aligned}
&\text{i.} && \Delta p_t = \alpha_0 + \alpha_1 \Delta p_{t-1} + \alpha_2 \Delta p_{t-2} + u_t \\
&\text{ii.} && \Delta p_t = \alpha_0 + \alpha_1 \Delta p_{t-1} + \alpha_2 \Delta p_{t-2} + \beta_1 \Delta r_{t-1} + \beta_2 \Delta r_{t-2} + u_t \\
&\text{iii.} && \Delta p_t = \alpha_0 + \alpha_1 \Delta p_{t-1} + \alpha_2 \Delta p_{t-2} + \alpha_3 \Delta p_{t-3} + \alpha_4 \Delta p_{t-4} + u_t \\
&\text{iv.} && \Delta p_t = \alpha_0 + \alpha_1 \Delta p_{t-1} + \alpha_2 \Delta p_{t-2} + \alpha_3 \Delta p_{t-3} + \alpha_4 \Delta p_{t-4} \\
& && \qquad + \beta_1 \Delta r_{t-1} + \beta_2 \Delta r_{t-2} + \beta_3 \Delta r_{t-3} + \beta_4 \Delta r_{t-4} + u_t
\end{aligned}
$$

Tables 1 to 4 report the estimation results.

Models (i) and (ii) are virtually identical. The goodness of fit (adjusted R-squared) is about 70% in both models. It seems that changes in interest rates are not informative to explain changes in house prices in this sample. This claim is also supported by an F-test for a null hypothesis of no effect of interest rates, which is $F = \frac{(322.47-308.83)/2}{308.83/(34-5)} = 0.64$. We compare it with the 5% critical value of an $F_{2,29} = 3.33$ and conclude that the null cannot be rejected, so expanding our information set to include interest rates will not help with the forecasting of house price changes.

In models (iii) and (iv) we include further dynamics of prices and rates. In model (iii), the improvement is marginal. The new regressors are not (or barely) statistically significant at the 5% level. The goodness of fit has a modest increase to 73%. An F-test of model (iii) versus model (i) $F = \frac{(322.47-269.21)/2}{269.21/(32-5)} = 4.65$ reveals that the null hypothesis of no further lags can be rejected at the 5% level ($F_{2,27} = 3.35$) but it cannot at the 1% level ($F_{2,27} = 5.49$). Similar comments apply to model (iv). Comparing model (iv) with model (i), we set the following null hypothesis: lagged interest rates and lagged (lags 3 and 4) price changes do not have any effect on current price changes (6 coefficients are claimed to be zero). The corresponding F-test is $F = \frac{(322.47-244.45)/6}{244.45/(32-9)} = 1.22$ and the 5% critical value is $F_{6,23} = 2.53$. Thus, we fail to reject the null and we settle on model (i). In summary, a multivariate information set, which includes not only past information on prices but also on interest rates, is not any more valuable than the univariate information set to explain house price growth.

Dependent Variable: DP
Method: Least Squares
Sample (adjusted): 1978 2011
Included observations: 34 after adjustments
Newey-West HAC Standard Errors & Covariance (lag truncation=3)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.660277	0.589296	1.12045	0.2711
DP(-1)	1.239116	0.250548	4.945626	0.0000
DP(-2)	-0.538585	0.354404	-1.51969	0.1387

R-squared	0.72172	Mean dependent var	2.655237
Adjusted R-squared	0.703767	S.D. dependent var	5.925817
S.E. of regression	3.225263	Akaike info criterion	5.264003
Sum squared resid	322.4719	Schwarz criterion	5.398682
Log likelihood	-86.48805	F-statistic	40.19936
Durbin-Watson stat	1.468675	Prob(F-statistic)	0.000000

Table 1: Regression Results of Model (i)

Copyright ©2013 Pearson Education, Inc.

Dependent Variable: DP				
Method: Least Squares				
Sample (adjusted): 1978 2011				
Included observations: 34 after adjustments				
Newey-West HAC Standard Errors & Covariance (lag truncation=3)				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.545574	0.614589	0.887706	0.3820
DP(-1)	1.253992	0.257147	4.876565	0.0000
DP(-2)	-0.53199	0.365234	-1.45657	0.1560
DR(-1)	-0.61435	0.300961	-2.04129	0.0504
DR(-2)	0.215435	0.151986	1.417471	0.1670
R-squared	0.733493	Mean dependent var		2.655237
Adjusted R-squared	0.696733	S.D. dependent var		5.925817
S.E. of regression	3.263328	Akaike info criterion		5.338425
Sum squared resid	308.8301	Schwarz criterion		5.56289
Log likelihood	-85.7532	F-statistic		19.95375
Durbin-Watson stat	1.471291	Prob(F-statistic)		0.000000

Table 2: Regression Results of Model (ii)

Dependent Variable: DP				
Method: Least Squares				
Sample (adjusted): 1980 2011				
Included observations: 32 after adjustments				
Newey-West HAC Standard Errors & Covariance (lag truncation=3)				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.748974	0.608225	1.231409	0.2288
DP(-1)	1.576519	0.202442	7.787503	0.0000
DP(-2)	-1.452615	0.370593	-3.919707	0.0005
DP(-3)	1.045887	0.564639	1.852311	0.0749
DP(-4)	-0.485468	0.351456	-1.381305	0.1785
R-squared	0.766471	Mean dependent var		2.550312
Adjusted R-squared	0.731874	S.D. dependent var		6.098095
S.E. of regression	3.157646	Akaike info criterion		5.280132
Sum squared resid	269.2097	Schwarz criterion		5.509153
Log likelihood	-79.48211	F-statistic		22.15435
Durbin-Watson stat	2.0406	Prob(F-statistic)		0.000000

Table 3: Regression Results of Model (iii)

Copyright ©2013 Pearson Education, Inc.

Dependent Variable: DP				
Method: Least Squares				
Sample (adjusted): 1980 2011				
Included observations: 32 after adjustments				
Newey-West HAC Standard Errors & Covariance (lag truncation=3)				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.769484	0.59297	1.297679	0.2073
DP(-1)	1.59055	0.230934	6.88747	0.0000
DP(-2)	-1.49379	0.345609	-4.32219	0.0003
DP(-3)	1.152694	0.539543	2.136428	0.0435
DP(-4)	-0.57769	0.328759	-1.75719	0.0922
DR(-1)	-0.70712	0.382326	-1.84951	0.0773
DR(-2)	0.398556	0.296101	1.346013	0.1914
DR(-3)	-0.51496	0.351942	-1.46319	0.1569
DR(-4)	0.522899	0.289926	1.803558	0.0844
R-squared	0.787946	Mean dependent var		2.550312
Adjusted R-squared	0.714188	S.D. dependent var		6.098095
S.E. of regression	3.260125	Akaike info criterion		5.433667
Sum squared resid	244.4536	Schwarz criterion		5.845905
Log likelihood	-77.9387	F-statistic		10.68287
Durbin-Watson stat	2.143111	Prob(F-statistic)		0.000004

Table 4: Regression Results of Model (iv)

Exercise 3

We plot the ACFs and PACFs of the same data as in Exercise 1 at the *quarterly* frequency in Figures 5 to 8. Observe that for the time series in levels (P and R), the profiles of the ACFs and PACFs are identical to those of the time series at the annual frequency in Exercise 1: strong autocorrelations decaying smoothly towards zero and a large first-order partial autocorrelation coefficient close to one.

Note: Given the student's knowledge at this stage, it is too early to identify stochastic trends on the data but these are examples that illustrate that if there is a stochastic trend, it will show up regardless of the frequency of the data. Consider plotting the quarterly and yearly data in the same plot to explain the trend.

For the time series in differences (changes in prices and interest rates), the autocorrelation is weaker than in levels. For interest rates changes, the message in Exercise 1 remains: there is no meaningful autocorrelation that we could exploit to forecast interest rates changes. For price changes, we find substantial autocorrelation at the quarterly frequency. Now, the ACF shows significant dynamics (read the Q-statistics and the shape of the ACF) and in the PACF there is autocorrelation up to order five, which will be the basis to produce a forecast for price growth. Note that the autocorrelation should be interpreted according to the frequency of the data. For instance, in this exercise, autocorrelation of order four means correlation of observations four quarters apart (one year). In Exercise 1, autocorrelation of order one means correlation of observations one year apart. In general, for stationary data, we should expect different ACFs and PACFs at different frequencies.

Copyright ©2013 Pearson Education, Inc.

Sample: 1975Q1 2011Q4
Included observations: 148

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.988	0.988	147.51	0.000
		2	0.975	-0.094	291.92	0.000
		3	0.960	-0.042	432.95	0.000
		4	0.944	-0.050	570.31	0.000
		5	0.926	-0.101	703.35	0.000
		6	0.906	-0.066	831.62	0.000
		7	0.885	-0.027	954.97	0.000
		8	0.864	-0.021	1073.3	0.000
		9	0.841	-0.063	1186.3	0.000
		10	0.818	-0.020	1293.8	0.000
		11	0.794	0.019	1396.1	0.000
		12	0.771	-0.009	1493.2	0.000
		13	0.746	-0.077	1584.8	0.000
		14	0.720	-0.049	1670.8	0.000
		15	0.695	0.002	1751.3	0.000
		16	0.669	-0.017	1826.5	0.000
		17	0.641	-0.060	1896.2	0.000
		18	0.613	-0.028	1960.5	0.000
		19	0.586	0.007	2019.6	0.000
		20	0.559	0.008	2073.7	0.000

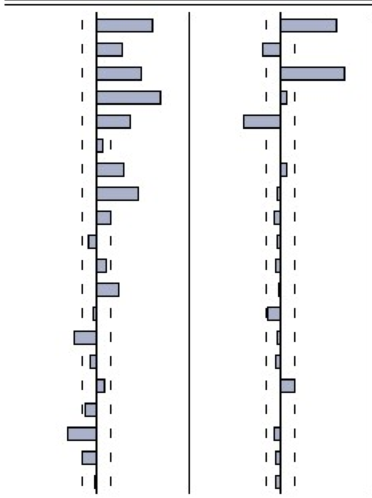
Figure 5: ACF, PACF and Q-Statistic of P

Sample: 1975Q1 2011Q4
Included observations: 148

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.976	0.976	143.99	0.000
		2	0.946	-0.157	280.10	0.000
		3	0.915	-0.018	408.23	0.000
		4	0.882	-0.043	528.19	0.000
		5	0.847	-0.066	639.52	0.000
		6	0.814	0.056	743.20	0.000
		7	0.785	0.036	840.23	0.000
		8	0.757	0.000	931.15	0.000
		9	0.728	-0.065	1015.7	0.000
		10	0.696	-0.051	1093.7	0.000
		11	0.665	-0.015	1165.3	0.000
		12	0.631	-0.059	1230.4	0.000
		13	0.602	0.093	1289.9	0.000
		14	0.574	0.005	1344.6	0.000
		15	0.545	-0.076	1394.2	0.000
		16	0.515	-0.023	1438.8	0.000
		17	0.489	0.034	1479.3	0.000
		18	0.463	-0.003	1516.0	0.000
		19	0.441	0.076	1549.5	0.000
		20	0.422	0.039	1580.4	0.000

Figure 6: ACF, PACF and Q-Statistic of R

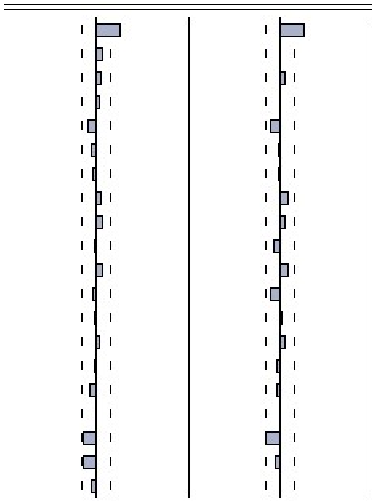
Copyright ©2013 Pearson Education, Inc.



Sample: 1975Q1 2011Q4
Included observations: 147

	AC	PAC	Q-Stat	Prob
1	0.634	0.634	60.317	0.000
2	0.285	-0.196	72.562	0.000
3	0.508	0.725	111.85	0.000
4	0.724	0.068	192.07	0.000
5	0.383	-0.413	214.68	0.000
6	0.080	-0.003	215.69	0.000
7	0.295	0.081	229.32	0.000
8	0.473	-0.031	264.53	0.000
9	0.167	-0.078	268.95	0.000
10	-0.084	-0.027	270.08	0.000
11	0.112	-0.054	272.11	0.000
12	0.257	-0.020	282.84	0.000
13	-0.037	-0.148	283.06	0.000
14	-0.260	-0.041	294.22	0.000
15	-0.075	-0.050	295.16	0.000
16	0.091	0.159	296.55	0.000
17	-0.128	0.001	299.29	0.000
18	-0.321	-0.078	316.77	0.000
19	-0.158	-0.064	321.03	0.000
20	-0.009	-0.053	321.04	0.000

Figure 7: ACF, PACF and Q-Statistic of DP



Sample: 1975Q1 2011Q4
Included observations: 147

	AC	PAC	Q-Stat	Prob
1	0.264	0.264	10.435	0.001
2	0.074	0.004	11.255	0.004
3	0.064	0.047	11.881	0.008
4	0.031	0.003	12.029	0.017
5	-0.092	-0.112	13.336	0.020
6	-0.058	-0.010	13.853	0.031
7	-0.042	-0.022	14.129	0.049
8	0.051	0.085	14.543	0.069
9	0.078	0.060	15.510	0.078
10	-0.016	-0.067	15.552	0.113
11	0.072	0.084	16.390	0.127
12	-0.039	-0.107	16.642	0.164
13	-0.023	0.022	16.726	0.212
14	0.027	0.052	16.844	0.265
15	-0.016	-0.041	16.889	0.326
16	-0.069	-0.038	17.683	0.343
17	-0.004	-0.001	17.686	0.409
18	-0.148	-0.165	21.391	0.260
19	-0.140	-0.058	24.755	0.169
20	-0.050	0.007	25.182	0.195

Figure 8: ACF, PACF and Q-Statistic of DR

Copyright ©2013 Pearson Education, Inc.

Exercise 4

With the same time series of house price changes as in Exercise 3, we run the following four regression models

$$\begin{aligned} \text{i'.} \quad & \Delta p_t = \alpha_0 + \alpha_1 \Delta p_{t-1} + u_t \\ \text{ii'.} \quad & \Delta p_t = \alpha_0 + \alpha_1 \Delta p_{t-1} + \alpha_2 \Delta p_{t-2} + u_t \\ \text{iii'.} \quad & \Delta p_t = \alpha_0 + \alpha_1 \Delta p_{t-1} + \alpha_2 \Delta p_{t-2} + \alpha_3 \Delta p_{t-3} + u_t \\ \text{iv'.} \quad & \Delta p_t = \alpha_0 + \alpha_1 \Delta p_{t-1} + \alpha_2 \Delta p_{t-2} + \alpha_3 \Delta p_{t-3} + \alpha_4 \Delta p_{t-4} + u_t \end{aligned}$$

We report the estimation results in Tables 5 to 8. Strictly speaking we cannot compare these models, where the dependent variable is *quarterly* growth rate, to those in Exercise 2, where the dependent variable is *annual* growth rate. The time variation of the growth rate is different at different frequencies and, as we have seen in Exercise 3, the ACFs and PACFs are also different. Comparing the proposed four models, we observe that we need more than two lags. The adjusted R-squared jumps from 0.43 in model (ii') to 0.75 in model (iii'). This is not a surprise because we have already seen how strong is the partial autocorrelation coefficient of order three. All the three lags are statistically significant (see p-values are zero). Model (iv') does not contribute further to explain the current quarterly price growth. Thus, out of the four models, we prefer model (iii').

Note: the student should estimate an additional model with five lags, which will improve upon model (iii') given the information in the PACF.

For model (iii') we implement a recursive and a rolling estimation schemes. The total number of observations is 144. At the start, we use the first 124 observations. We keep on estimating model (iii'), either recursively or with a rolling window of 124 observations, for the last 20 quarters of the sample until we exhaust all observations. The last 20 quarters correspond to 2007-2011, which is a period of turmoil in the US housing market. In Figures 9 to 12, we plot the estimates in each of the last 20 quarters. The estimates are almost identical for both schemes but they fluctuate greatly, which brings into question the stability of the model over this sample.

Dependent Variable: DP				
Method: Least Squares				
Sample (adjusted): 1975Q3 2011Q4				
Included observations: 146 after adjustments				
Newey-West HAC Standard Errors & Covariance (lag truncation=4)				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.207603	0.12744	1.629034	0.1055
DP(-1)	0.650499	0.078284	8.309495	0.0000
R-squared	0.412479	Mean dependent var		0.637481
Adjusted R-squared	0.408399	S.D. dependent var		1.853774
S.E. of regression	1.425842	Akaike info criterion		3.561006
Sum squared resid	292.7557	Schwarz criterion		3.601877
Log likelihood	-257.953	F-statistic		101.0974
Durbin-Watson stat	1.708299	Prob(F-statistic)		0.000000

Table 5: Regression Results of Model (i')

Copyright ©2013 Pearson Education, Inc.

Dependent Variable: DP				
Method: Least Squares				
Sample (adjusted): 1975Q4 2011Q4				
Included observations: 145 after adjustments				
Newey-West HAC Standard Errors & Covariance (lag truncation=4)				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.261865	0.132498	1.976367	0.0501
DP(-1)	0.78616	0.082303	9.552007	0.0000
DP(-2)	-0.21206	0.074512	-2.84594	0.0051
R-squared	0.438163	Mean dependent var		0.639986
Adjusted R-squared	0.43025	S.D. dependent var		1.859952
S.E. of regression	1.403925	Akaike info criterion		3.536894
Sum squared resid	279.8829	Schwarz criterion		3.598482
Log likelihood	-253.425	F-statistic		55.37111
Durbin-Watson stat	1.668695	Prob(F-statistic)		0.000000

Table 6: Regression Results of Model (ii')

Dependent Variable: DP				
Method: Least Squares				
Sample (adjusted): 1976Q1 2011Q4				
Included observations: 144 after adjustments				
Newey-West HAC Standard Errors & Covariance (lag truncation=4)				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.049461	0.126103	0.39223	0.6955
DP(-1)	0.936222	0.097712	9.581452	0.0000
DP(-2)	-0.79841	0.127894	-6.2427	0.0000
DP(-3)	0.760084	0.099912	7.60757	0.0000
R-squared	0.752741	Mean dependent var		0.641989
Adjusted R-squared	0.747443	S.D. dependent var		1.866287
S.E. of regression	0.937904	Akaike info criterion		2.737046
Sum squared resid	123.1529	Schwarz criterion		2.819541
Log likelihood	-193.067	F-statistic		142.0694
Durbin-Watson stat	2.089587	Prob(F-statistic)		0.000000

Table 7: Regression Results of Model (iii')

Dependent Variable: DP				
Method: Least Squares				
Sample (adjusted): 1976Q2 2011Q4				
Included observations: 143 after adjustments				
Newey-West HAC Standard Errors & Covariance (lag truncation=4)				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.041563	0.137001	0.30338	0.7621
DP(-1)	0.883517	0.191307	4.618329	0.0000
DP(-2)	-0.74106	0.233628	-3.17196	0.0019
DP(-3)	0.690686	0.219458	3.147234	0.0020
DP(-4)	0.074004	0.201848	0.36663	0.7145
R-squared	0.753958	Mean dependent var		0.643274
Adjusted R-squared	0.746826	S.D. dependent var		1.872783
S.E. of regression	0.942317	Akaike info criterion		2.753388
Sum squared resid	122.5386	Schwarz criterion		2.856984
Log likelihood	-191.867	F-statistic		105.7199
Durbin-Watson stat	1.916062	Prob(F-statistic)		0.000000

Table 8: Regression Results of Model (iv')

Copyright ©2013 Pearson Education, Inc.

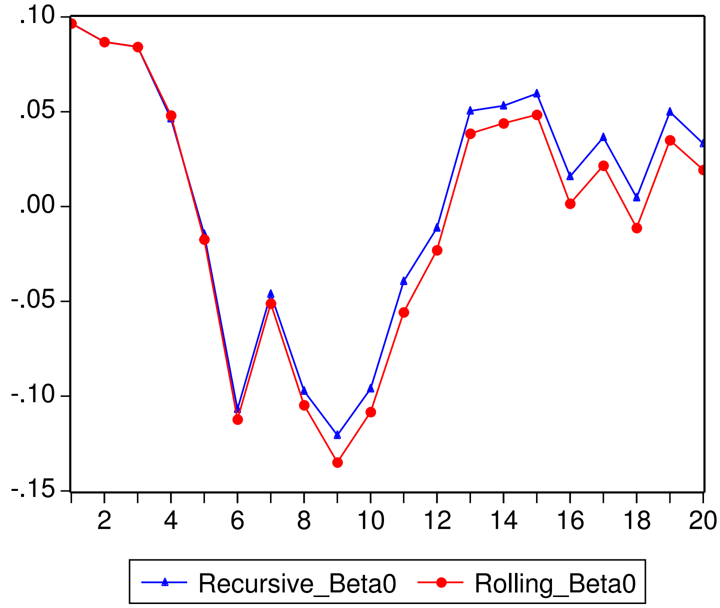


Figure 9: Estimates of the Regression Coefficient (α_0)

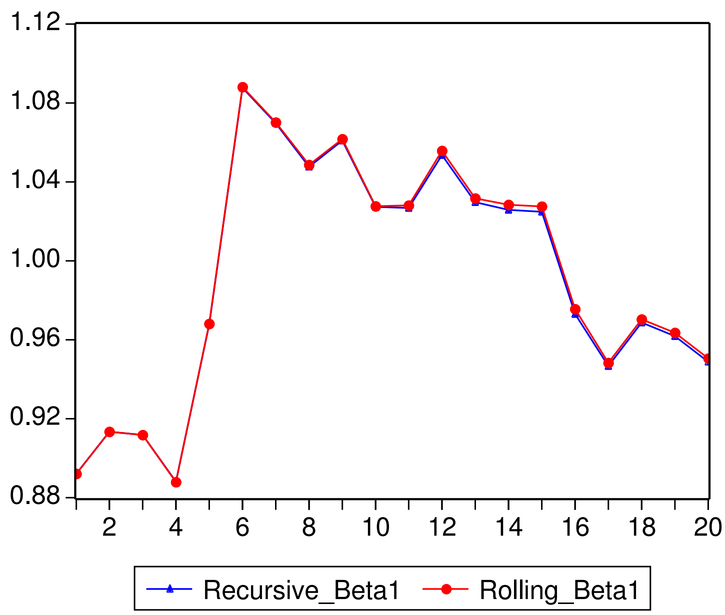


Figure 10: Estimates of the Regression Coefficient (α_1)

Copyright ©2013 Pearson Education, Inc.

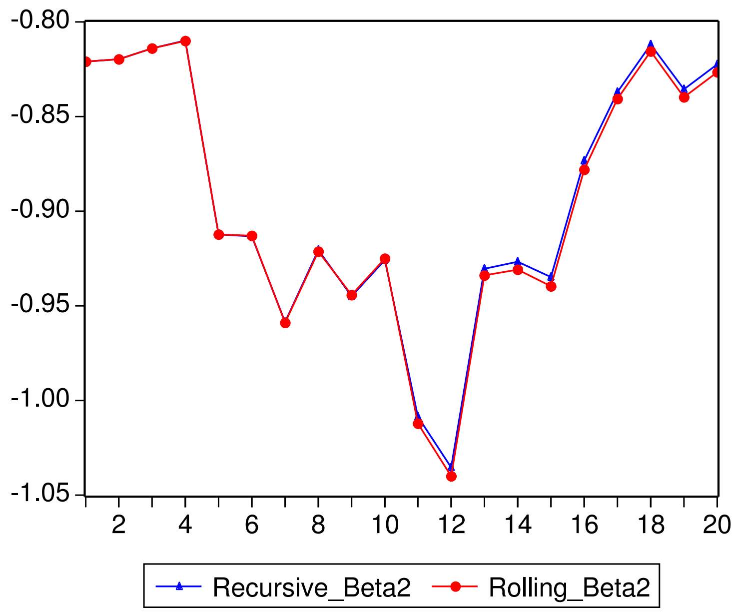


Figure 11: Estimates of the Regression Coefficient (α_2)

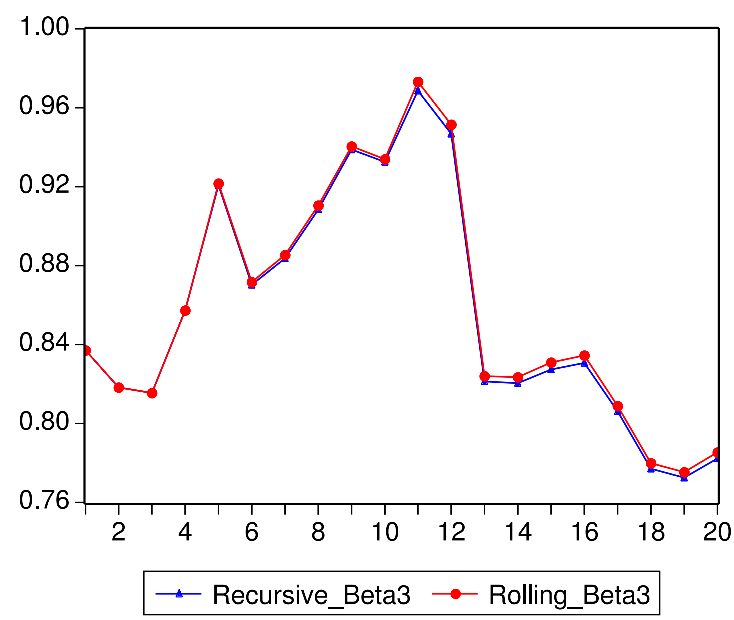


Figure 12: Estimates of the Regression Coefficient (α_3)

Copyright ©2013 Pearson Education, Inc.

Exercise 5
Suppose that deliveries come to the store once a week and they are not perishable. Ideally, a sales manager would like to sell these deliveries within a week so that there is not need for storage or to face unhappy customers whose demand is not met. However, sales fluctuate from week to week and sometimes there will be unsold (overstocking) or oversold (understocking) goods. Overstocking will require additional warehouses and employees to manage the extra goods and potentially it would push costs up. Furthermore, high inventory usually means less cash (or cash equivalents) available for other functions of the store, adding some liquidity constraints. In general, unsold goods trigger store discounts, which may affect the profit margin. On the other hand, understocking translates into lost sales affecting the revenue function of the store, and potentially reducing the customer base. Though the store could directly ship to customers' addresses or expedite the delivery from other warehouses, these actions will also increase costs.

Only the sales manager will know what it is more costly to the store, to overstock or understock. She may need to consider commercial rents in the area, access to credit, employee salaries, market competition, transportation costs, client loyalty, etc. Let us assume that after her analysis, she concludes that overstocking is more costly. Then, her loss function is asymmetric and it should be taken into account to predict department sales. Let y_{t+1} denote the actual department sales at time $t+1$, and let $f_{t,1}$ denote the one-period-ahead forecast of department sales at time t. Then, the forecast error is $e_{t,1} = y_{t+1} - f_{t,1}$. When $e_{t,1} > (<) 0$, actual sales are larger (smaller) than predicted sales, and she will face understocking (overstocking). A good starting point is the following asymmetric loss function

$$L(e) = \begin{cases} a|e|, & e > 0 \\ b|e|, & e \leq 0 \end{cases}$$

Since overstocking is more costly, the loss function will penalize overstocking more severely than understocking. Therefore, b should be larger than a ($a < b$).

Exercise 6
For Exercises 6-10, the data are *quarterly* growth rate of Real GDP and the one-quarter-ahead Greenbook forecast from 1969.Q1 to 2006.Q4 for a total of 152 observations. The Greenbook forecasts are reported as annualized rates but we have converted this rate to quarter-to-quarter growth rate so that we can compare with the actual data. The forecasts are released with a 5-year delay according to Fed policy. The forecast errors are calculated as the difference between the actual value and the forecast, i.e. $e_{t,1} = y_{t+1} - f_{t,1}$. In Figures 13, 14 and 15 , we plot the actual data, the forecast, and the forecast error respectively. Both GDP growth series are first order stationary as they fluctuate around a central value and we do not observe persistence deviations from this value. We observe large volatility at the beginning of the sample (1968.Q4 to mid-1980s) followed by the calmer years of the great moderation.

Copyright ©2013 Pearson Education, Inc.

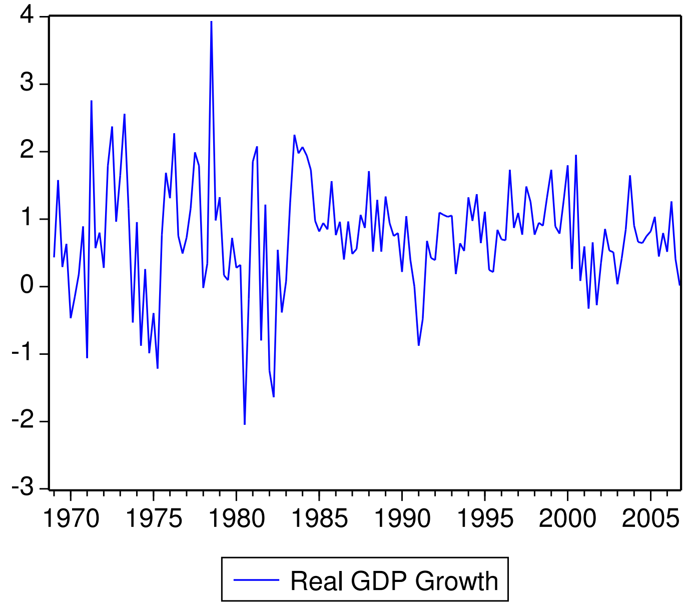


Figure 13: Real GDP Growth

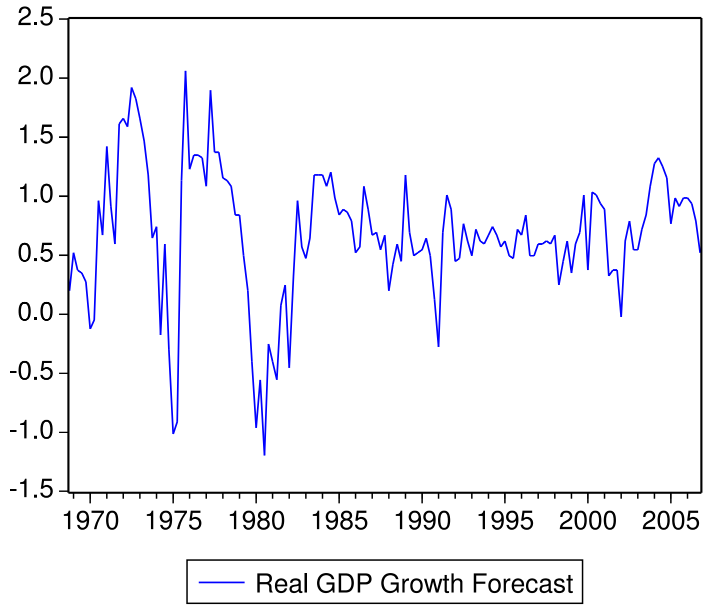


Figure 14: Real GDP Growth Forecast

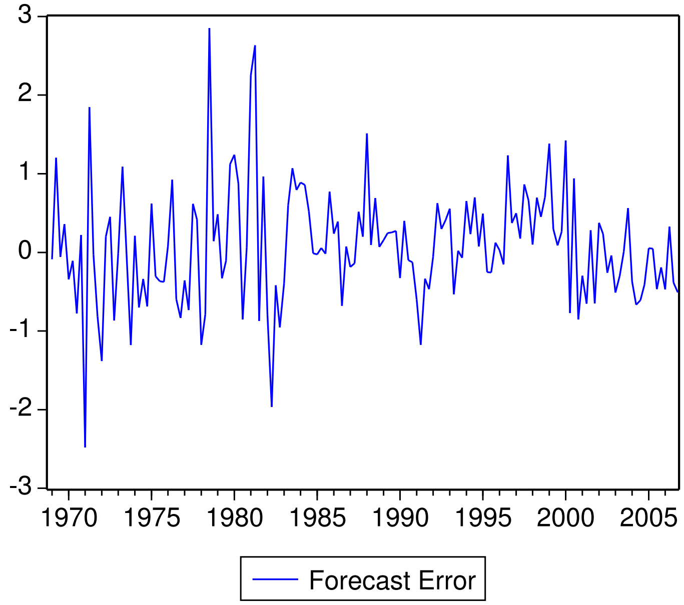


Figure 15: Forecast Error

The descriptive statistics of these series are reported in Figures 16 and 17. The average quarterly growth rate over the full period was 0.75 %, slightly above the one-quarter-ahead average forecast of 0.67% . Observe that the forecast series is less volatile ($\hat{\sigma} = 0.55$) than the actual series ($\hat{\sigma} = 0.84$), which is expected because the actual series has embedded surprises or innovations that, by the nature, are not predictable and add to the uncertainty of the series. The forecast error exhibit similar profile; it is more volatile at the beginning of the sample than at the end, reflecting the difficulty of forecasting in uncertain times. The forecast series seem to be centered around zero, the sample average is 0.077 (descriptive statistics in Figure 18) . We could test the null hypothesis of a zero mean by constructing a t-ratio. Assuming that the observations are uncorrelated, $t = (\bar{e} - 0)/(\sqrt{\hat{\sigma}^2/T}) = 0.077/\sqrt{0.74^2/152} = 1.27$, which means that we fail to reject the null at any reasonable significance level. Thus, the average of the forecast error is zero.

In Figures 19, 20 and 21, we report the ACF and PACF of these three series. The actual GDP growth exhibit some time dependence, it looks like the present growth rate will depend mainly on the rates of the last two quarters. Observe that the autocorrelation functions of the forecast series have similar profile to those of the actual data but the shapes are more pronounced. There is a tendency for the forecast series to be smoother and more dependent over time than the actual series, which is expected as forecasts cannot include unpredictable innovations by definition. The autocorrelation functions of the one-quarter-ahead forecast errors indicate that this series does not have any autocorrelation (check the Q-statistics). This is good news because indicates that the forecast is exploiting as much as possible whatever information set the Fed has to construct its forecast. The one-step-ahead forecast error represents the' surprise' or shock that is not present in the information set, so that we should not expect any time dependence on it.

Copyright ©2013 Pearson Education, Inc.

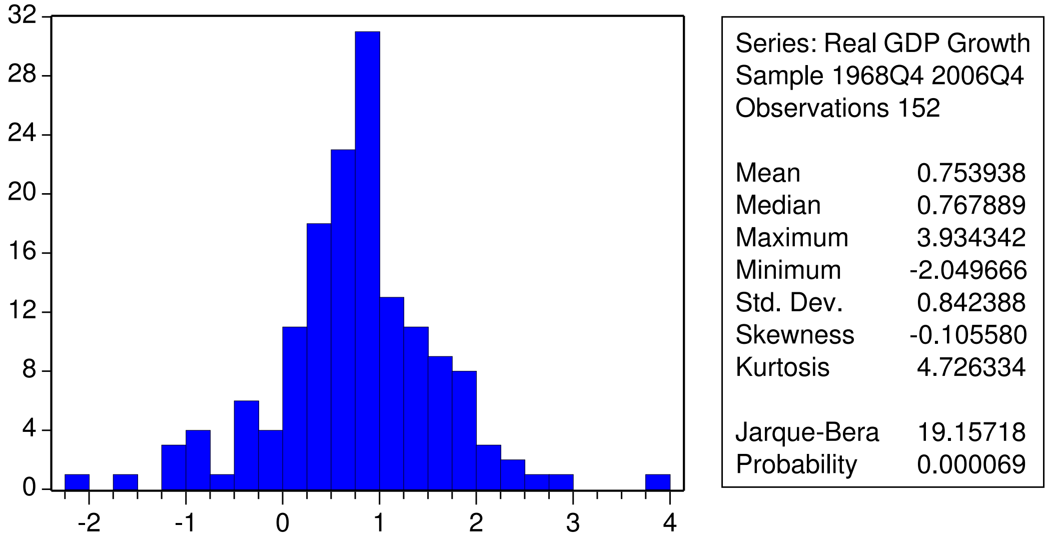


Figure 16: Descriptive Statistics for Real GDP Growth

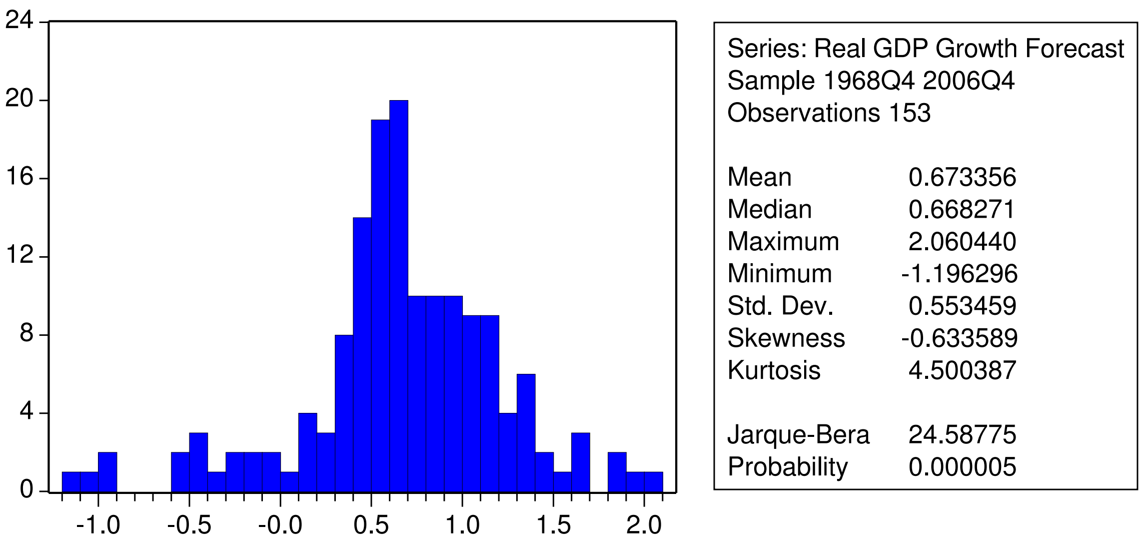


Figure 17: Descriptive Statistics for Real GDP Growth Forecast

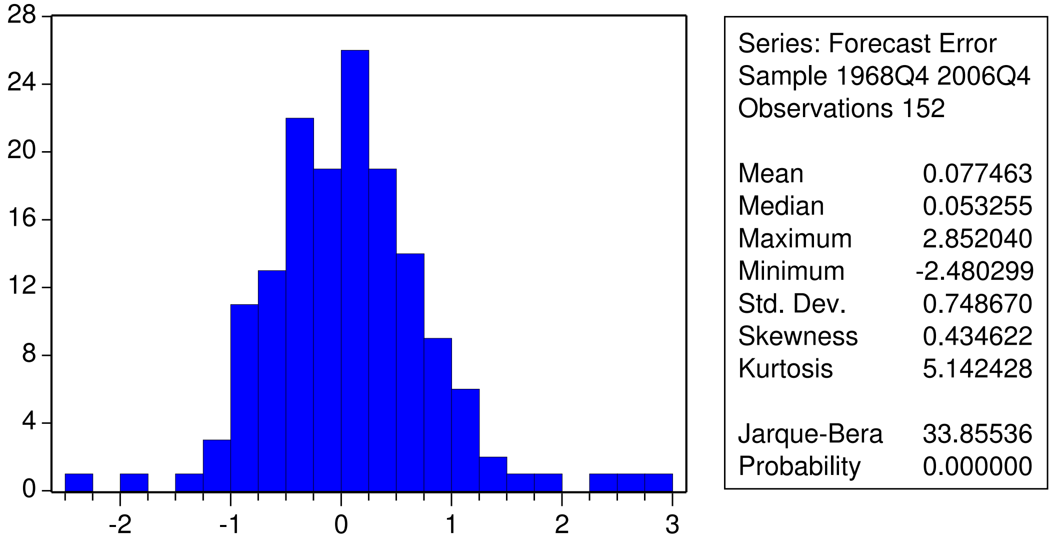


Figure 18: Descriptive Statistics for Forecast Error

Sample: 1968Q4 2006Q4
Included observations: 152

	AC	PAC	Q-Stat	Prob
1	0.245	0.245	9.3202	0.002
2	0.201	0.150	15.637	0.000
3	0.062	-0.018	16.242	0.001
4	0.061	0.023	16.834	0.002
5	-0.070	-0.101	17.608	0.003
6	-0.062	-0.043	18.217	0.006
7	-0.136	-0.095	21.219	0.003
8	-0.216	-0.166	28.788	0.000
9	-0.055	0.075	29.279	0.001
10	-0.051	0.010	29.712	0.001
11	-0.001	0.019	29.713	0.002
12	-0.251	-0.282	40.235	0.000
13	-0.100	-0.049	41.925	0.000
14	-0.153	-0.078	45.917	0.000
15	-0.145	-0.134	49.508	0.000
16	-0.020	0.079	49.574	0.000
17	-0.059	-0.071	50.187	0.000
18	-0.000	0.016	50.187	0.000
19	0.023	-0.022	50.284	0.000
20	0.139	-0.005	53.725	0.000

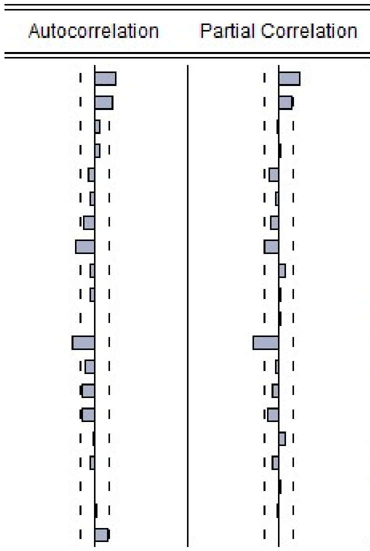
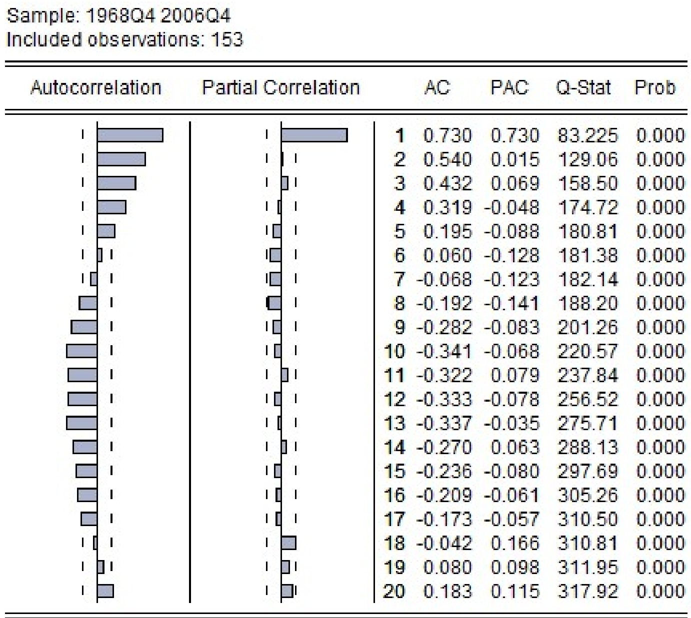


Figure 19: ACF and PACF for Real GDP Growth

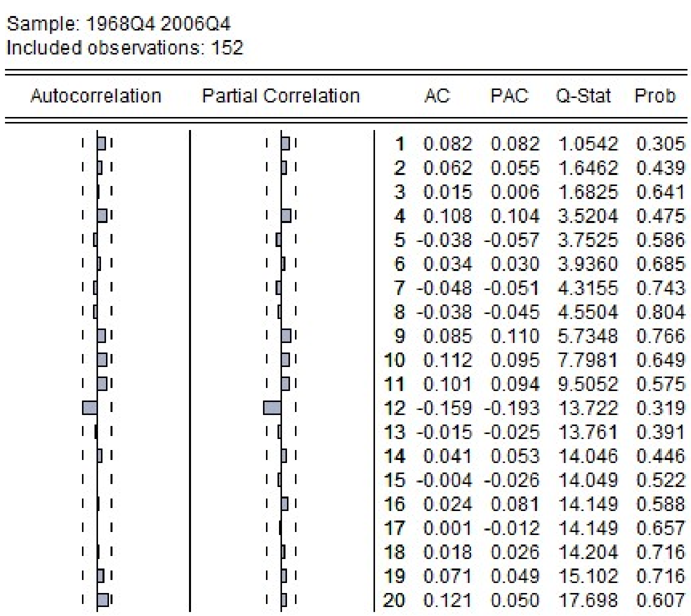
Copyright ©2013 Pearson Education, Inc.



Sample: 1968Q4 2006Q4
Included observations: 153

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.730	0.730	83.225	0.000
		2	0.540	0.015	129.06	0.000
		3	0.432	0.069	158.50	0.000
		4	0.319	-0.048	174.72	0.000
		5	0.195	-0.088	180.81	0.000
		6	0.060	-0.128	181.38	0.000
		7	-0.068	-0.123	182.14	0.000
		8	-0.192	-0.141	188.20	0.000
		9	-0.282	-0.083	201.26	0.000
		10	-0.341	-0.068	220.57	0.000
		11	-0.322	0.079	237.84	0.000
		12	-0.333	-0.078	256.52	0.000
		13	-0.337	-0.035	275.71	0.000
		14	-0.270	0.063	288.13	0.000
		15	-0.236	-0.080	297.69	0.000
		16	-0.209	-0.061	305.26	0.000
		17	-0.173	-0.057	310.50	0.000
		18	-0.042	0.166	310.81	0.000
		19	0.080	0.098	311.95	0.000
		20	0.183	0.115	317.92	0.000

Figure 20: ACF and PACF for Real GDP Growth Forecast



Sample: 1968Q4 2006Q4
Included observations: 152

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.082	0.082	1.0542	0.305
		2	0.062	0.055	1.6462	0.439
		3	0.015	0.006	1.6825	0.641
		4	0.108	0.104	3.5204	0.475
		5	-0.038	-0.057	3.7525	0.586
		6	0.034	0.030	3.9360	0.685
		7	-0.048	-0.051	4.3155	0.743
		8	-0.038	-0.045	4.5504	0.804
		9	0.085	0.110	5.7348	0.766
		10	0.112	0.095	7.7981	0.649
		11	0.101	0.094	9.5052	0.575
		12	-0.159	-0.193	13.722	0.319
		13	-0.015	-0.025	13.761	0.391
		14	0.041	0.053	14.046	0.446
		15	-0.004	-0.026	14.049	0.522
		16	0.024	0.081	14.149	0.588
		17	0.001	-0.012	14.149	0.657
		18	0.018	0.026	14.204	0.716
		19	0.071	0.049	15.102	0.716
		20	0.121	0.050	17.698	0.607

Figure 21: ACF and PACF for Forecast Error

Exercise 7
In Figure 22, we plot the actual data versus the forecast of GDP growth. The green line is the 45-degree line (slope equal to 1). The data seem to scatter around the line though there are some large deviations. We run a regression model of actual values on forecast values. The regression output is provided in Table 9. The regression line (in red) shows the estimates of the intercept $\hat{\beta}_0 = 0.25$ and of the slope $\hat{\beta}_1 = 0.74$. If the forecast is unbiased, and under a quadratic loss function, we should expect $\beta_0 = 0$ and $\beta_1 = 1$; this is to say that the data points should be scattered around the 45-degree line. We can set up an individual hypothesis for each coefficient (t-ratios) and a join hypothesis for both coefficients (F-test). The t-ratio for $H_0 : \beta_0 = 0$ is 2.15 with a p-value of 3.24%, thus we reject the null at the 5 % level but we fail to reject at the 1 % level. The t-ratio for $H_0 : \beta_1 = 1$ is $t = (\hat{\beta}_1 - 1)/\hat{\sigma}_{\hat{\beta}_1} = (0.742971 - 1)/0.122430 = -2.099$, and again we reject the null at the 5% level but fail to reject at the 1 % significance level. The F statistic for the joint hypothesis $H_0 : \beta_0 = 0, \beta_1 = 1$ is calculated by using $F_{m,n-k-1} = \frac{(SSR_0 - SSR_1)/m}{SSR_1/(n-k-1)} = \frac{(85.54869 - 81.57559)/2}{81.57559/150} = 3.6528$. See that under the null hypothesis the model becomes $y_{t+1} = f_{t,1} + u_{t+1}$, then to obtain $SSR_0 = \sum_t (y_{t+1} - f_{t,1})^2 = 85.54869$. For an $F_{2,150}$, the critical value at the 5% significance level is 3.06 and at the 1 % level is 4.75. Once again, we reject the joint null hypothesis at the 5 % level but we fail to reject at the 1 %.

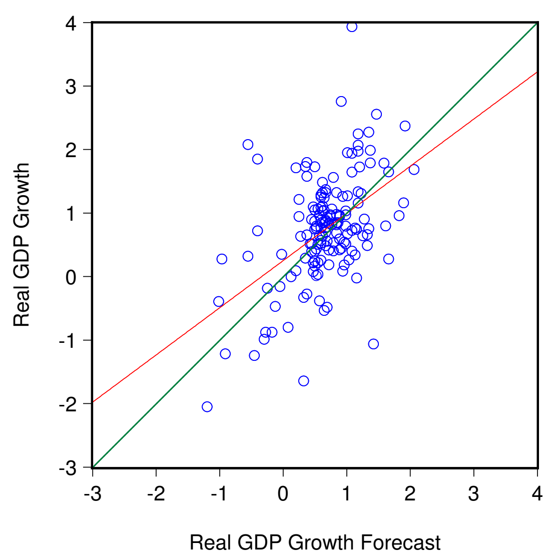


Figure 22: Real GDP Growth Forecast and Real GDP Growth

Copyright ©2013 Pearson Education, Inc.

Dependent Variable: GRGDP				
Method: Least Squares				
Sample (adjusted): 1969Q1 2006Q4				
Included observations: 152 after adjustments				
Newey-West HAC Standard Errors & Covariance (lag truncation=4)				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.251337	0.11638	2.159623	0.0324
GRGDPF2	0.742971	0.12243	6.068519	0.0000
R-squared	0.238695	Mean dependent var		0.753938
Adjusted R-squared	0.233619	S.D. dependent var		0.842388
S.E. of regression	0.737453	Akaike info criterion		2.241842
Sum squared resid	81.57559	Schwarz criterion		2.281630
Log likelihood	-168.38	F-statistic		47.03006
Durbin-Watson stat	1.860853	Prob(F-statistic)		0.000000

Table 9: Regression of Real GDP Growth on Real GDP Growth Forecast

Exercise 8

In Exercise 6, we have already tested that the expected value of the forecast errors is zero, i.e. $t = (\bar{e} - 0)/(\sqrt{\hat{\sigma}^2/T}) = 0.077/\sqrt{0.74^2/152} = 1.27$. In Table 10, we show the estimation results from running a regression of the forecast errors on a constant and three lags. Observe that these regressors do not explain at all the current forecast error. The adjusted R-squared is practically zero; the F-test for overall significance of the regression has a p-value of 66 %, which means that none of the regressors is informative to explain the current error; and the individual t-ratios all have p-values much larger than 5%. In summary, we cannot forecast the error itself based on its own past, which is in agreement with the autocorrelation functions that we have seen in Exercise 6 and with the fact that the one-step-ahead forecast error must be a surprise.

Dependent Variable: ERROR				
Method: Least Squares				
Sample (adjusted): 1969Q4 2006Q4				
Included observations: 149 after adjustments				
Newey-West HAC Standard Errors & Covariance (lag truncation=4)				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.060552	0.06407	0.945089	0.3462
ERROR2(-1)	0.08325	0.09452	0.880770	0.3799
ERROR2(-2)	0.05455	0.093214	0.585218	0.5593
ERROR2(-3)	0.005634	0.053759	0.104810	0.9167
R-squared	0.0108	Mean dependent var		0.071927
Adjusted R-squared	-0.00967	S.D. dependent var		0.750310
S.E. of regression	0.753928	Akaike info criterion		2.299439
Sum squared resid	82.41906	Schwarz criterion		2.380082
Log likelihood	-167.308	F-statistic		0.527706
Durbin-Watson stat	1.995581	Prob(F-statistic)		0.663945

Table 10: Regression of Forecast Error on Its Own Past

Exercise 9

In Table 11 we show the estimation results from regressing the forecast error on two lags of the errors, the current forecast, and one lag of the forecast itself. Overall, this regression confirms that the current forecast error is not predictable based on its past history; the corresponding t-ratios have large p-values. However, the current and one-lag forecasts are marginally significant at the 5 % level, which means that there is common information in the forecast error and the forecast itself that could be exploited further. The F-test of overall significance of the regression also supports

Copyright ©2013 Pearson Education, Inc.

marginally this finding (p-value 2.76 %).

The F statistic for the joint hypothesis $H_0 : \beta_3 = \beta_4 = 0$ is $F_{2,145} = \frac{(82.467-77.339)/2}{77.339/145} = 4.8075$. The estimation of the restricted model is provided in Table 12. For an $F_{2,145}$, the critical value at the 5% significance level is 3.065 and at the 1 % level is 4.76. Thus, we reject the null at the 5 % and 1% levels though at the 1 % is a borderline rejection.

Dependent Variable: ERROR2				
Method: Least Squares				
Sample (adjusted): 1969Q3 2006Q4				
Included observations: 150 after adjustments				
Newey-West HAC Standard Errors & Covariance (lag truncation=4)				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.147305	0.130739	1.126709	0.2617
ERROR2(-1)	0.054618	0.087603	0.623473	0.534
ERROR2(-2)	0.086757	0.108384	0.800454	0.4248
GRGDPF2	-0.49541	0.212711	-2.32901	0.0212
GRGDPF2(-1)	0.366326	0.196735	1.862031	0.0646
R-squared	0.071963	Mean dependent var		0.071060
Adjusted R-squared	0.046362	S.D. dependent var		0.747864
S.E. of regression	0.730322	Akaike info criterion		2.242102
Sum squared resid	77.33865	Schwarz criterion		2.342457
Log likelihood	-163.158	F-statistic		2.810931
Durbin-Watson stat	1.951896	Prob(F-statistic)		0.027675

Table 11: Regression of Forecast Error on Its Own Lags and Real GDP Growth Forecast

Dependent Variable: ERROR2				
Method: Least Squares				
Sample (adjusted): 1969Q3 2006Q4				
Included observations: 150 after adjustments				
Newey-West HAC Standard Errors & Covariance (lag truncation=4)				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.059708	0.063425	0.941399	0.3480
ERROR2(-1)	0.080668	0.092540	0.871707	0.3848
ERROR2(-2)	0.055653	0.091611	0.607494	0.5445
R-squared	0.010427	Mean dependent var		0.071060
Adjusted R-squared	-0.00304	S.D. dependent var		0.747864
S.E. of regression	0.748998	Akaike info criterion		2.279638
Sum squared resid	82.46681	Schwarz criterion		2.339850
Log likelihood	-167.973	F-statistic		0.774427
Durbin-Watson stat	1.990547	Prob(F-statistic)		0.462839

Table 12: Regression of Forecast Error on Its Own Lags

Exercise 10

The results of Exercises 7-9 raise questions about the optimality of the Fed forecasts but the evidence against optimality is statistically weak. Overall, the test statistics that we have implemented reject optimality at the 5% level but fail to reject it at the 1% level . This is a borderline scenario that may prompt some skepticism in the forecaster and for which there is only one response: in the absence of any other inside information, wait for more data to come.

Under a quadratic loss function, the optimal forecast is the conditional mean of y_{t+1}, i.e., $f^*_{t,1} = E(y_{t+1}|I_t)$. The regression in Exercise 7, $y_{t+1} = \beta_0 + \beta_1 f_{t,1} + u_{t+1}$, implies that $E(y_{t+1}|I_t) = \beta_0 + \beta_1 f_{t,1}$, hence the relevance of the null hypothesis $H_0 : \beta_0 = 0, \beta_1 = 1$. We do not find strong evidence against this null given the results of Exercise 7. A note of caution: if we wish to entertain

Copyright ©2013 Pearson Education, Inc.

the rejection of the null (because we work only with a 5% significance level), we would not conclude categorically that the Fed forecast is suboptimal but that the Fed may also use an asymmetric loss function on computing its forecasts. The results of Exercise 8 provide strong support for an unbiased and uncorrelated one-step-ahead forecast error so that this error is unpredictable, as we expected. Most evidence against optimality is coming from the results of Exercise 9. If the forecast error is correlated with the forecast, the Fed is not exploiting optimally the information set because either there is some information in the forecast error that is not properly modeled or there is some information that is left out. But once again, the statistical tests in Exercise 9 are borderline in their significance and we should entertain these results with some reservation.

Copyright ©2013 Pearson Education, Inc.