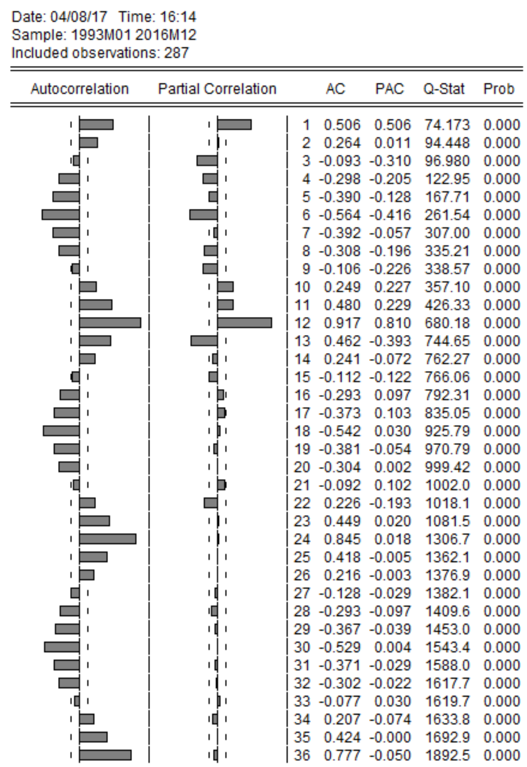
Homework 5

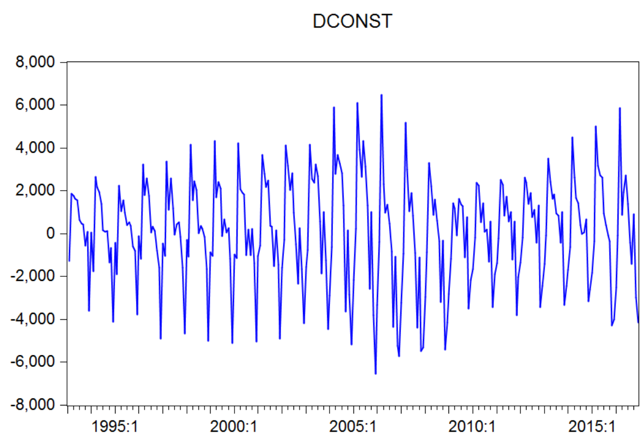
Eco 4306 Economic and Business Forecasting
Spring 2018
Due: Thursday, March 8, before the class

Problem 1

(a) The time series plot and correlogram of the total private residential construction spending for the period 1993M01-2013M12 are below.


Date: 04/08/17 Time: 16:14
Sample: 1993M01 2016M12
Included observations: 287

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.506	0.506	74.173	0.000
		2	0.264	0.011	94.448	0.000
		3	-0.093	-0.310	96.980	0.000
		4	-0.298	-0.205	122.95	0.000
		5	-0.390	-0.128	167.71	0.000
		6	-0.564	-0.416	261.54	0.000
		7	-0.392	-0.057	307.00	0.000
		8	-0.308	-0.196	335.21	0.000
		9	-0.106	-0.226	338.57	0.000
		10	0.249	0.227	357.10	0.000
		11	0.480	0.229	426.33	0.000
		12	0.917	0.810	680.18	0.000
		13	0.462	-0.393	744.65	0.000
		14	0.241	-0.072	762.27	0.000
		15	-0.112	-0.122	766.06	0.000
		16	-0.293	0.097	792.31	0.000
		17	-0.373	0.103	835.05	0.000
		18	-0.542	0.030	925.79	0.000
		19	-0.381	-0.054	970.79	0.000
		20	-0.304	0.002	999.42	0.000
		21	-0.092	0.102	1002.0	0.000
		22	0.226	-0.193	1018.1	0.000
		23	0.449	0.020	1081.5	0.000
		24	0.845	0.018	1306.7	0.000
		25	0.418	-0.005	1362.1	0.000
		26	0.216	-0.003	1376.9	0.000
		27	-0.128	-0.029	1382.1	0.000
		28	-0.293	-0.097	1409.6	0.000
		29	-0.367	-0.039	1453.0	0.000
		30	-0.529	0.004	1543.4	0.000
		31	-0.371	-0.029	1588.0	0.000
		32	-0.302	-0.022	1617.7	0.000
		33	-0.077	0.030	1619.7	0.000
		34	0.207	-0.074	1633.8	0.000
		35	0.424	-0.000	1692.9	0.000
		36	0.777	-0.050	1892.5	0.000



(b) The shape of the correlogram in (a) is similar to the one in Figure 7.19 suggesting that the U.S. residential construction in 1993M01-2013M12 follows a similar process as in 2002M01-2011M01

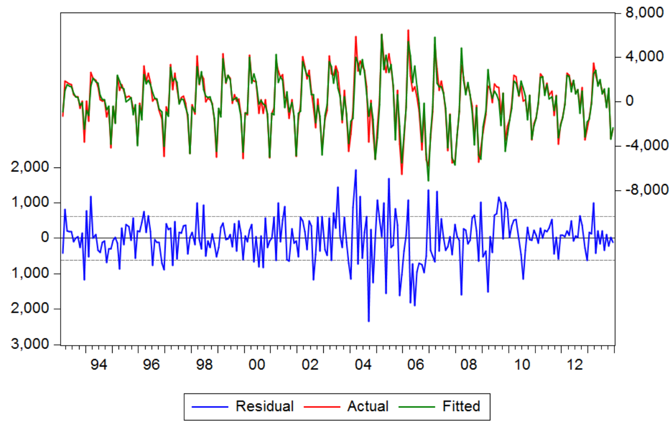
(c) Since in the correlogram in (a) ACF decays in a wave like pattern and PACF has dominant lags 1, 6, 12, 13, we first consider a multiplicative AR(1)+SAR(1) model

Dependent Variable: DCONST
Method: ARMA Maximum Likelihood (BFGS)
Date: 04/08/17 Time: 16:14
Sample: 1993M02 2013M12
Included observations: 251
Convergence achieved after 5 iterations
Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	218.5690	1235.621	0.176890	0.8597
AR(1)	0.563509	0.044817	12.57349	0.0000
SAR(12)	0.944410	0.013547	69.71517	0.0000
SIGMASQ	379202.0	27048.41	14.01938	0.0000

R-squared	0.933686	Mean dependent var	49.44223
Adjusted R-squared	0.932881	S.D. dependent var	2396.078
S.E. of regression	620.7600	Akaike info criterion	15.82347
Sum squared resid	95179704	Schwarz criterion	15.87965
Log likelihood	-1981.845	Hannan-Quinn criter.	15.84608
F-statistic	1159.242	Durbin-Watson stat	2.130377
Prob(F-statistic)	0.000000		

Inverted AR Roots	1.00	.86-.50i	.86+.50i	.56
	.50+.86i	.50-.86i	.00+1.00i	-.00-1.00i
	-.50+.86i	-.50-.86i	-.86+.50i	-.86-.50i
	-1.00			



Date: 04/08/17 Time: 16:14
Sample: 1993M01 2013M12
Included observations: 251
Q-statistic probabilities adjusted for 2 ARMA terms

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.066	-0.066	1.1109	
		2	0.136	0.132	5.8246	
		3	-0.022	-0.005	5.9460	0.015
		4	-0.010	-0.030	5.9704	0.051
		5	0.058	0.061	6.8484	0.077
		6	-0.075	-0.065	8.3212	0.080
		7	-0.081	-0.109	10.030	0.074
		8	0.116	0.132	13.549	0.035
		9	-0.088	-0.054	15.602	0.029
		10	-0.017	-0.073	15.675	0.047
		11	-0.044	-0.013	16.196	0.063
		12	-0.004	0.010	16.201	0.094
		13	-0.016	-0.046	16.267	0.132
		14	-0.008	0.010	16.284	0.179
		15	0.004	0.030	16.288	0.234
		16	0.009	-0.027	16.308	0.295
		17	0.050	0.054	16.997	0.319
		18	-0.093	-0.085	19.357	0.251
		19	0.047	0.020	19.956	0.276
		20	0.068	0.098	21.224	0.268
		21	-0.041	-0.048	21.684	0.300
		22	0.029	-0.011	21.918	0.345
		23	-0.020	0.022	22.033	0.398
		24	0.152	0.147	28.514	0.159
		25	-0.000	-0.029	28.514	0.197
		26	-0.059	-0.051	29.501	0.202
		27	-0.028	-0.033	29.724	0.235
		28	0.025	0.020	29.903	0.272
		29	-0.047	-0.041	30.532	0.291
		30	-0.033	-0.026	30.842	0.324
		31	-0.043	0.000	31.369	0.348
		32	0.058	0.025	32.353	0.351
		33	0.002	0.019	32.354	0.400
		34	0.033	0.043	32.675	0.434
		35	-0.022	-0.003	32.818	0.476
		36	-0.072	-0.129	34.329	0.452

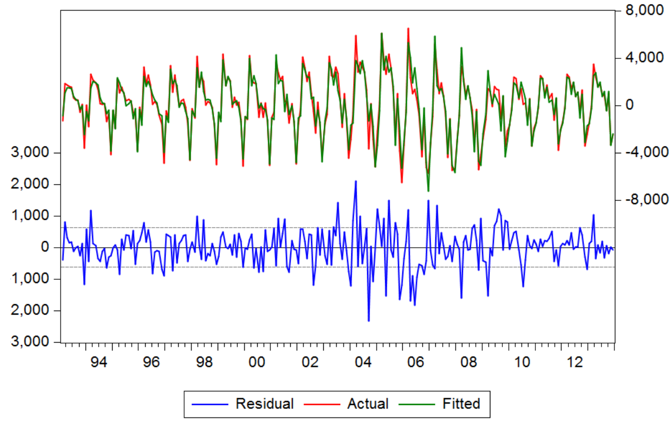
The low p-value of the Ljung-Box test for residuals at lag 3 and the marginally significant lag 2 in the PACF of the residuals from AR(1)-SAR(1) model suggest that a modified AR(2)+SAR(1) model will likely yield better results. As shown below, adding the second lag really improves the results of ACF and PACF for residuals, and the AIC also favors AR(2)+SAR(1) model (but note that the SIC prefers the AR(1)+SAR(1) model).

Dependent Variable: DCONST
Method: ARMA Maximum Likelihood (BFGS)
Date: 04/08/17 Time: 16:14
Sample: 1993M02 2013M12
Included observations: 251
Convergence achieved after 7 iterations
Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	214.4219	1374.109	0.156044	0.8761
AR(1)	0.497140	0.054595	9.105947	0.0000
AR(2)	0.116143	0.052147	2.227211	0.0268
SAR(12)	0.944592	0.013249	71.29646	0.0000
SIGMASQ	373960.6	26485.07	14.11968	0.0000

R-squared	0.934603	Mean dependent var	49.44223
Adjusted R-squared	0.933540	S.D. dependent var	2396.078
S.E. of regression	617.7066	Akaike info criterion	15.81783
Sum squared resid	93864109	Schwarz criterion	15.88805
Log likelihood	-1980.137	Hannan-Quinn criter.	15.84609
F-statistic	878.9103	Durbin-Watson stat	1.975678
Prob(F-statistic)	0.000000		

Inverted AR Roots	1.00	.86+.50i	.86-.50i	.67
	.50+.86i	.50-.86i	.00+1.00i	-.00-1.00i
	-.17	-.50+.86i	-.50-.86i	-.86+.50i
	-.86-.50i	-1.00		



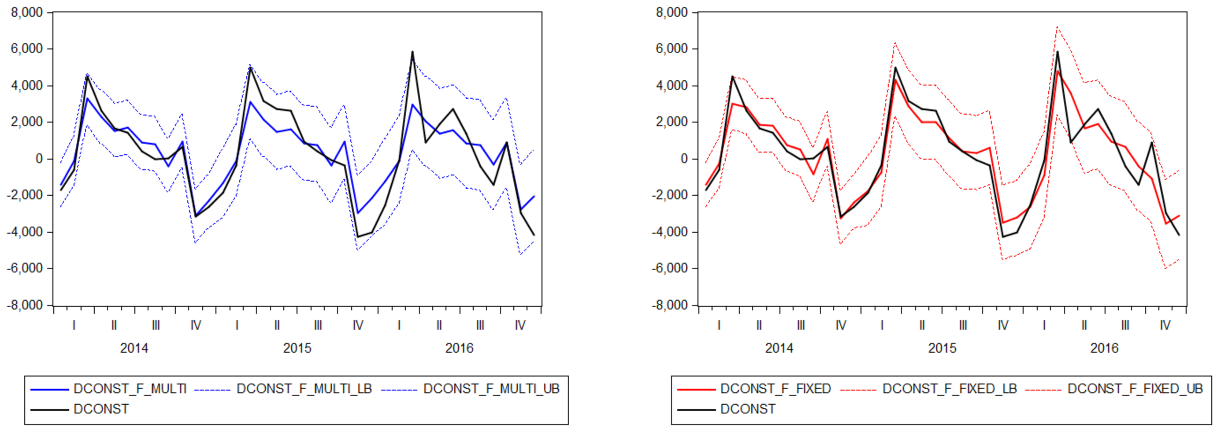
Date: 04/08/17 Time: 16:14
Sample: 1993M01 2013M12
Included observations: 251
Q-statistic probabilities adjusted for 3 ARMA terms

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.011	0.011	0.0322	
		2	0.054	0.054	0.7743	
		3	-0.054	-0.056	1.5282	
		4	-0.032	-0.034	1.7972	0.180
		5	0.042	0.049	2.2539	0.324
		6	-0.091	-0.092	4.3953	0.222
		7	-0.081	-0.089	6.1091	0.191
		8	0.112	0.132	9.3965	0.094
		9	-0.073	-0.078	10.799	0.095
		10	-0.031	-0.065	11.048	0.137
		11	-0.041	-0.009	11.501	0.175
		12	-0.005	-0.001	11.507	0.243
		13	-0.013	-0.053	11.553	0.316
		14	-0.007	0.014	11.565	0.397
		15	0.006	0.021	11.575	0.480
		16	0.019	-0.024	11.670	0.555
		17	0.040	0.045	12.109	0.598
		18	-0.095	-0.096	14.580	0.482
		19	0.042	0.037	15.068	0.520
		20	0.070	0.082	16.416	0.495
		21	-0.041	-0.061	16.875	0.532
		22	0.010	-0.006	16.904	0.596
		23	-0.005	0.036	16.912	0.659
		24	0.158	0.146	23.933	0.296
		25	0.008	-0.038	23.952	0.350
		26	-0.075	-0.038	25.527	0.324
		27	-0.033	-0.021	25.837	0.362
		28	0.024	0.018	25.999	0.408
		29	-0.044	-0.047	26.564	0.432
		30	-0.043	-0.024	27.097	0.459
		31	-0.037	-0.000	27.498	0.491
		32	0.064	0.024	28.673	0.482
		33	0.022	0.028	28.816	0.527
		34	0.044	0.058	29.388	0.549
		35	-0.016	-0.004	29.459	0.596
		36	-0.071	-0.119	30.955	0.569

(d) The left panel in the figure below shows the multistep forecast for period 2014M01-2016M12. The RMSE for this forecast is 1031.828.

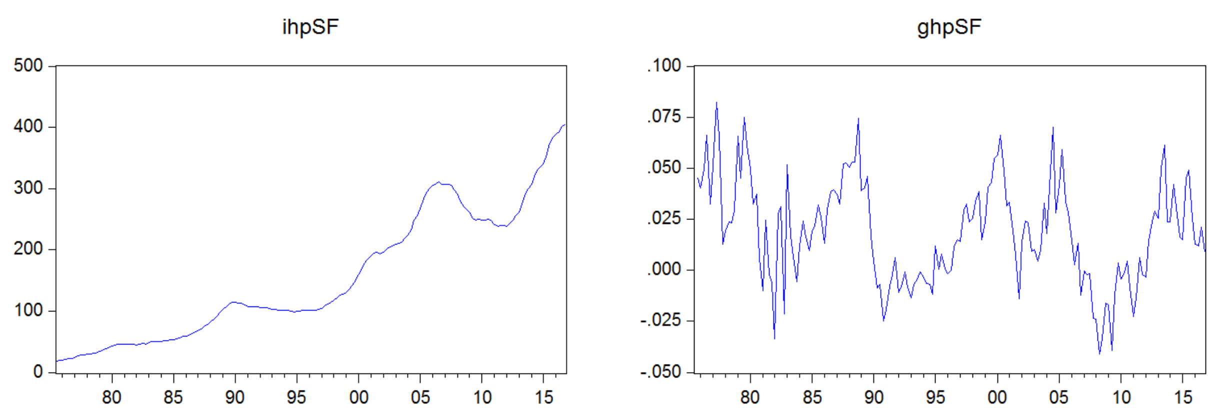
(e) The right panel in the figure below shows the fixed scheme sequence of one step ahead forecasts for period 2014M01-2016M12. The RMSE for this fixed scheme forecast is 846.067.

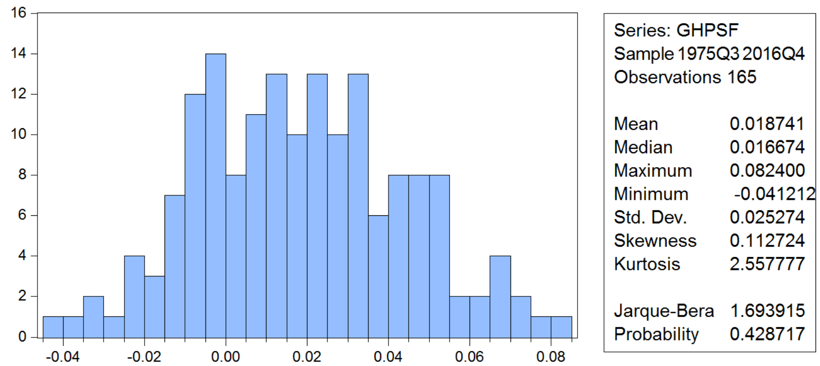
(f) The fixed scheme yields a lower RMSE than the multistep forecast, which is also visible in the two plots - the confidence interval of the fixed scheme forecast is narrower, and the forecast is closer to the



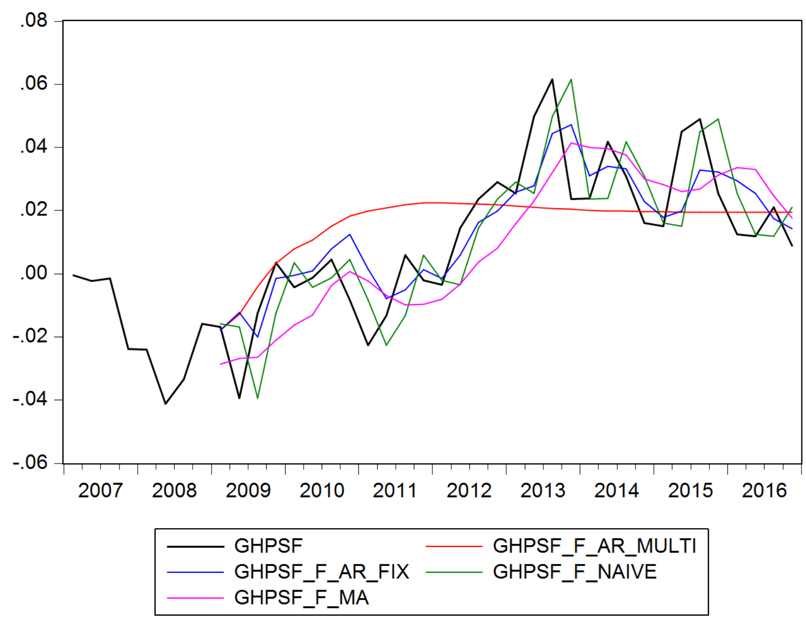
Problem 2

(a) Figure below shows the time series plots for the house price index `ihpsf`, and its growth rate `ghpsf`, and also the histogram for the growth rate, for San Francisco in 1975Q3-2016Q4. The growth rate shows a lot of persistence, but fluctuates around its mean, and the variance is constant over time - the time series thus appears weakly stationary. Jarque-Bera statistic is relatively small, since skewness is close to 0 and kurtosis close to 3. The p-value associated with Jarque-Bera statistic is 0.428 so we do not reject the hypothesis that the growth of house price index in San Francisco is normally distributed.





(b) Figure below shows a time series plot with actual growth in the house price index together with four forecasts: multistep forecast using an AR(p) `ghpsf_f_ar_multi`, sequence of one step ahead forecasts based on fixed scheme, using an AR(p) model `ghpsf_f_ar_fix`, sequence of one step ahead forecasts using naive forecast method `ghpsf_f_naive`, sequence of one step ahead forecasts using simple four observation moving average forecast `ghpsf_f_ma`.

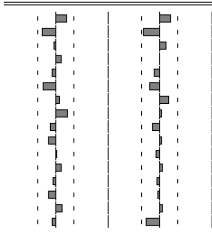


(c) The MSE for the four forecasts are $\text{MSE}_{multi} = 0.000357$, $\text{MSE}_{fix} = 0.000162$, $\text{MSE}_{naive} = 0.000238$, $\text{MSE}_{ma} = 0.000250$. The multistep forecast is the least precise, followed by the moving average forecast and the naive forecast. The fixed scheme forecast is the most precise out of these four.

(d) The correlograms for the one step ahead forecast error from the three forecasts reveal that only fixed scheme forecast produces forecast errors that do not contain any linear time dependence and can be considered white noise. Both the naive forecast and the simple moving average forecast fail this test.

fixed scheme

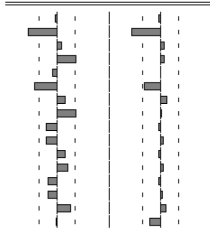
Date: 04/08/17 Time: 19:21
Sample: 2009Q1 2016Q4
Included observations: 32



	AC	PAC	Q-Stat	Prob
1	0.209	0.209	1.5396	0.215
2	-0.269	-0.327	4.1648	0.125
3	-0.042	0.118	4.2323	0.237
4	0.103	-0.006	4.6473	0.325
5	-0.076	-0.107	4.8788	0.431
6	-0.261	-0.201	7.7360	0.258
7	0.074	0.180	7.9750	0.335
8	0.224	0.026	10.245	0.248
9	-0.114	-0.144	10.860	0.285
10	-0.147	0.025	11.924	0.290
11	0.018	-0.066	11.941	0.368
12	0.112	0.050	12.623	0.397
13	-0.064	-0.059	12.855	0.459
14	-0.144	-0.038	14.104	0.442
15	0.116	0.062	14.967	0.454
16	-0.081	-0.271	15.413	0.495

naive method

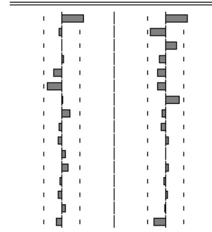
Date: 04/08/17 Time: 19:21
Sample: 2009Q1 2016Q4
Included observations: 32



	AC	PAC	Q-Stat	Prob
1	-0.028	-0.028	0.0278	0.868
2	-0.580	-0.581	12.220	0.002
3	0.092	0.075	12.535	0.006
4	0.385	0.079	18.281	0.001
5	-0.095	0.003	18.647	0.002
6	-0.454	-0.323	27.277	0.000
7	0.162	0.122	28.419	0.000
8	0.381	0.013	34.995	0.000
9	-0.210	-0.041	37.085	0.000
10	-0.226	0.054	39.607	0.000
11	0.170	-0.031	41.098	0.000
12	0.219	0.060	43.717	0.000
13	-0.182	-0.042	45.614	0.000
14	-0.187	0.033	47.731	0.000
15	0.274	0.109	52.545	0.000
16	-0.011	-0.211	52.553	0.000

simple moving average

Date: 04/08/17 Time: 19:21
Sample: 2009Q1 2016Q4
Included observations: 32



	AC	PAC	Q-Stat	Prob
1	0.433	0.433	6.5955	0.010
2	-0.056	-0.300	6.7094	0.035
3	-0.005	0.212	6.7103	0.082
4	0.025	-0.124	6.7347	0.151
5	-0.163	-0.161	7.8087	0.167
6	-0.292	-0.160	11.377	0.077
7	0.008	0.266	11.380	0.123
8	0.164	-0.076	12.605	0.126
9	-0.050	-0.092	12.722	0.176
10	-0.075	0.058	13.004	0.223
11	0.081	0.002	13.347	0.271
12	0.126	0.043	14.210	0.288
13	-0.042	-0.043	14.313	0.352
14	-0.078	0.037	14.678	0.401
15	0.067	-0.022	14.962	0.454
16	-0.101	-0.236	15.658	0.477

(e) Given regression

$$\Delta L_{t+j,1} = \beta_0 + u_{t+j} \quad \text{with } j = 0, 1, 2, \ldots, T-t-1$$

where $\Delta L_{t+j,1} = (e^{naive}_{t+j,1})^2 - (e^{fixed}_{t+j,1})^2$, the test of equal predictive ability for fixed scheme forecast vs naive forecast is equivalent to $H_0 : \beta_0 = 0$. The results below show that β_0 is not statistically significant at 1% or 5% levels, we thus can not reject that fixed scheme forecast and naive forecast are equally precise at these two levels.

Dependent Variable: DL_NAIVE
Method: Least Squares
Date: 04/08/17 Time: 19:21
Sample: 2009Q1 2016Q4
Included observations: 32
HAC standard errors & covariance (Bartlett kernel, Newey-West fixed bandwidth = 4.0000)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	7.63E-05	3.81E-05	2.004553	0.0538

R-squared	0.000000	Mean dependent var	7.63E-05
Adjusted R-squared	0.000000	S.D. dependent var	0.000259
S.E. of regression	0.000259	Akaike info criterion	-13.65162
Sum squared resid	2.07E-06	Schwarz criterion	-13.60582
Log likelihood	219.4260	Hannan-Quinn criter.	-13.63644
Durbin-Watson stat	2.442208		

(f) The results of a regression for the test of equal predictive ability for fixed scheme forecast vs simple four quarter moving average forecast

$$\Delta L_{t+j,1} = \beta_0 + u_{t+j} \quad \text{with } j = 0, 1, 2, \ldots, T-t-1$$

where $\Delta L_{t+j,1} = (e^{ma}_{t+j,1})^2 - (e^{fixed}_{t+j,1})^2$ are below. Since β_0 is statistically significant at 5% level, we can reject that fixed scheme forecast and simple moving average forecast are equally precise.

Dependent Variable: DL_MA
Method: Least Squares
Date: 04/08/17 Time: 19:21
Sample: 2009Q1 2016Q4
Included observations: 32
HAC standard errors & covariance (Bartlett kernel, Newey-West fixed bandwidth = 4.0000)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	8.81E-05	3.96E-05	2.224905	0.0335

R-squared	0.000000	Mean dependent var	8.81E-05
Adjusted R-squared	0.000000	S.D. dependent var	0.000237
S.E. of regression	0.000237	Akaike info criterion	-13.82388
Sum squared resid	1.75E-06	Schwarz criterion	-13.77808
Log likelihood	222.1821	Hannan-Quinn criter.	-13.80870
Durbin-Watson stat	2.025517		