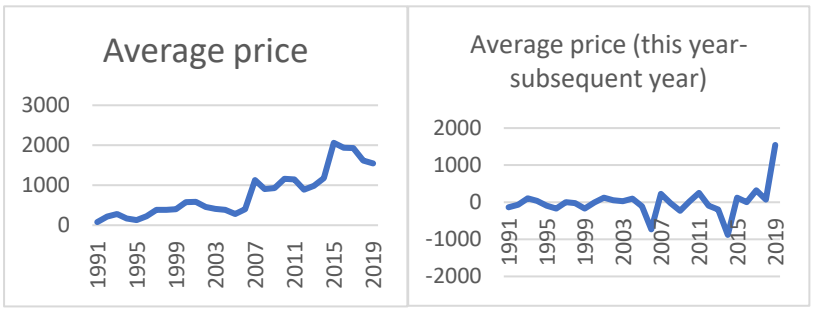
Lixia

Part 1



1. a) The average number of prices per year shows an upward trend. The average price for 2015 is the largest. Some prices in my data are the same in the next 5-15 days that can be considered wrong.

b) Price generally shows an upward trend and it is not stationary. The price reached the maximum in 2015.

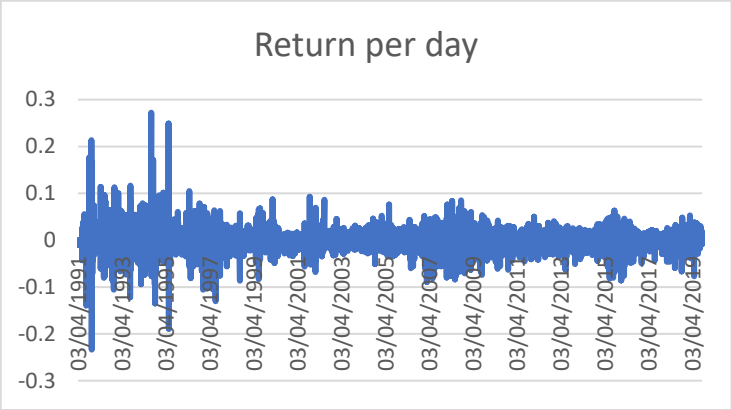
2.a) Rt=ln(Pt) − ln (Pt-1)

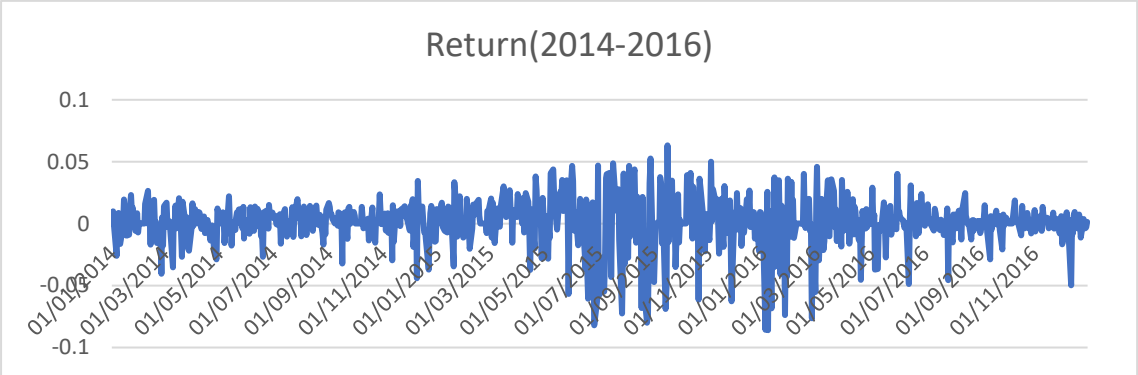
Series: RETURN_PER_DAY	
Sample 1 7424	
Observations 7424	
Mean	0.000380
Median	0.000000
Maximum	0.272152
Minimum	-0.233607
Std. Dev.	0.021175
Skewness	0.446091
Kurtosis	19.35532

b) The largest return numbers is 0.272152. There are no pairs of consecutive days that both have extreme returns, one positie and one negative.

1999/2/10	324.215	0	1995/1/30	125.857	0	2000/1/31	469.649	0
1999/2/11	324.215	0	1995/1/31	125.857	0	2000/2/1	469.649	0
1999/2/12	324.215	0	1995/2/1	125.857	0	2000/2/2	469.649	0
1999/2/15	324.215	0	1995/2/2	125.857	0	2000/2/3	469.649	0
1999/2/16	324.215	0	1995/2/3	125.857	0	2000/2/4	469.649	0
1999/2/17	324.215	0	1995/2/6	125.857	0	2000/2/7	469.649	0
1999/2/18	324.215	0	1995/2/7	125.857	0	2000/2/8	469.649	0
1999/2/19	324.215	0	1995/2/8	125.857	0	2000/2/9	469.649	0
1999/2/22	324.215	0	1995/2/9	125.857	0	2000/2/10	469.649	0
1999/2/23	324.215	0	1995/2/10	125.857	0	2000/2/11	469.649	0
1999/2/24	324.215	0						
1999/2/25	324.215	0						
1999/2/26	324.215	0						

c) There are many sequences of zeros in the series of returns. Every year has sequences of zeros in my return data. Most of these erroneous data appear around holidays, such as the New Year and National Day.

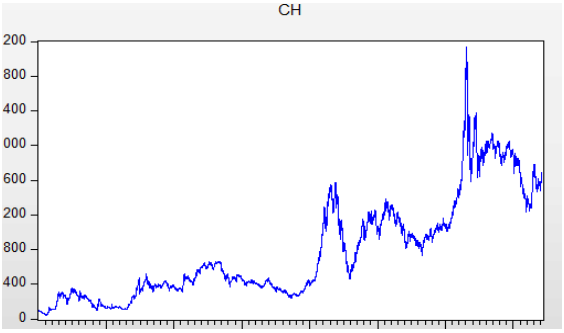




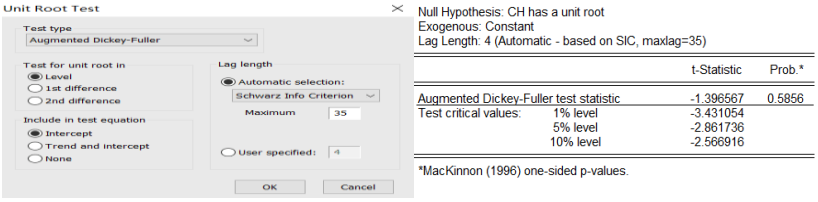
3. The first picture is the return chart. I can see some evidence of volatility clustering. We look at the second return (2014-2016) graph. Some returns come from high volatility periods (2015/7/1-2016/3/1). Other returns come from low volatility periods (2014/1/1-2015/1/1). (Large changes tend to be followed by large changes and small changes tend to be followed by small changes.)

Part 2

1. a. Plot the data. (CH=Price)



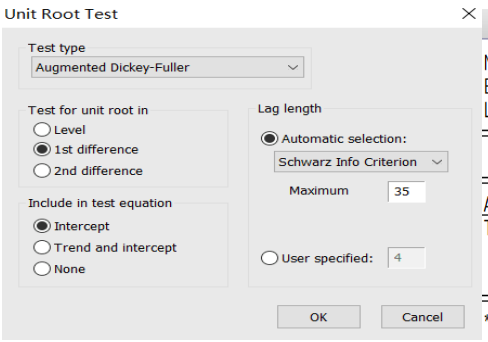
b. Check the data is stationary or not. (Unit root test)



The price is not stationary. (t-Statistic -1.396567 > critical values, can not reject null hypothesis.

So the price has a unit root.)

c. Take first difference.



Augmented Dickey-Fuller Unit Root Test on D(CH)

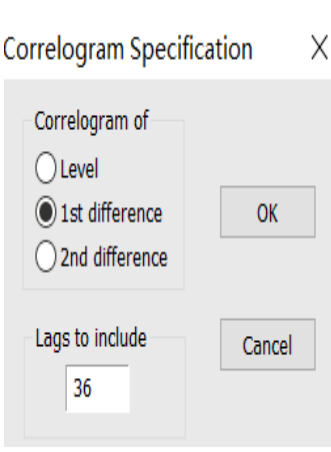
Null Hypothesis: D(CH) has a unit root
Exogenous: Constant
Lag Length: 3 (Automatic - based on SIC, maxlag=35)

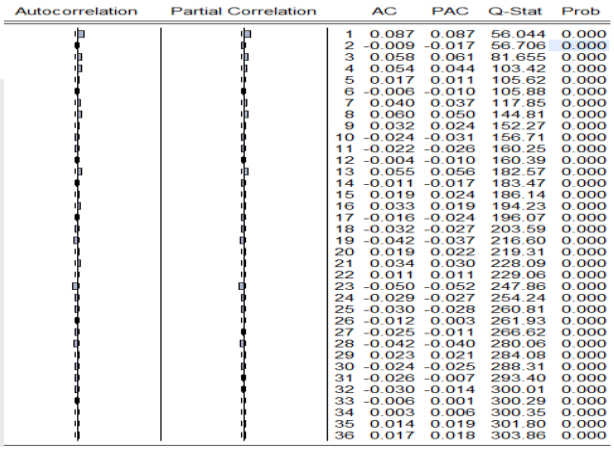
		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-38.75791	0.0000
Test critical values:	1% level	-3.431054	
	5% level	-2.861736	
	10% level	-2.566916	

*MacKinnon (1996) one-sided p-values.

The price is stationary now. (t-Statistic -38.75791 < critical values, reject null hypothesis, so D(CH) does not have a unit root.)

d. Examine the ACF/PACF.





	AC	PAC	Q-Stat	Prob
1	0.087	0.087	56.044	0.000
2	-0.009	-0.017	56.706	0.000
3	0.058	0.061	81.655	0.000
4	0.054	0.044	103.42	0.000
5	0.017	0.011	105.62	0.000
6	-0.006	-0.010	105.88	0.000
7	0.040	0.037	117.85	0.000
8	0.060	0.050	144.81	0.000
9	0.032	0.024	152.27	0.000
10	-0.024	-0.031	156.71	0.000
11	-0.022	-0.026	160.25	0.000
12	-0.004	-0.010	160.39	0.000
13	0.055	0.056	182.57	0.000
14	-0.011	-0.017	183.47	0.000
15	0.019	0.024	186.14	0.000
16	0.033	0.019	194.23	0.000
17	-0.016	-0.024	196.07	0.000
18	-0.032	-0.027	203.59	0.000
19	-0.042	-0.037	216.60	0.000
20	0.019	0.022	219.31	0.000
21	0.034	0.030	228.09	0.000
22	0.011	0.011	229.06	0.000
23	-0.050	-0.052	247.86	0.000
24	-0.029	-0.027	254.24	0.000
25	-0.030	-0.028	260.81	0.000
26	-0.012	0.003	261.93	0.000
27	-0.025	-0.011	266.62	0.000
28	-0.042	-0.040	280.06	0.000
29	0.023	0.021	284.08	0.000
30	-0.024	-0.025	288.31	0.000
31	-0.026	-0.007	293.40	0.000
32	-0.030	-0.014	300.01	0.000
33	-0.006	0.001	300.29	0.000
34	0.003	0.006	300.35	0.000
35	0.014	0.019	301.80	0.000
36	0.017	0.018	303.86	0.000

	AR(p)	MA(q)	ARMA(p,q)
ACF	Tails off	Cuts off after lag q	Tails off
PACF	Cuts off after lag p	Tails off	Tails off

From the graph, we can not easy to choose AR(p), MA(q) or ARMA(p,q) model.

e. Use the AIC to search model.(Find the smallest AIC)

Dependent Variable: DCH
Method: ARMA Maximum Likelihood (OPG - BHHH)
Date: 11/26/19 Time: 01:27
Sample: 2 7424
Included observations: 7423
Convergence achieved after 29 iterations
Coefficient covariance computed using outer product of gradients

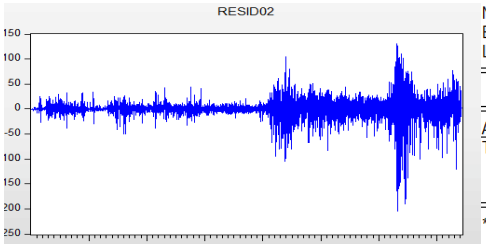
Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR(1)	0.086991	0.004138	21.02025	0.0000
SIGMASQ	319.9475	1.632518	195.9840	0.0000

R-squared	0.007429	Mean dependent var	0.213538
Adjusted R-squared	0.007295	S.D. dependent var	17.95510
S.E. of regression	17.88949	Akaike info criterion	8.606574
Sum squared resid	2374970.	Schwarz criterion	8.608436
Log likelihood	-31941.30	Hannan-Quinn criter.	8.607214
Durbin-Watson stat	1.997013		

Inverted AR Roots	.09

AR (1) has the smallest AIC that I found.

f. Check if the residual is white noise process.



Null Hypothesis: RESID02 has a unit root
Exogenous: Constant
Lag Length: 3 (Automatic - based on SIC, maxlag=35)

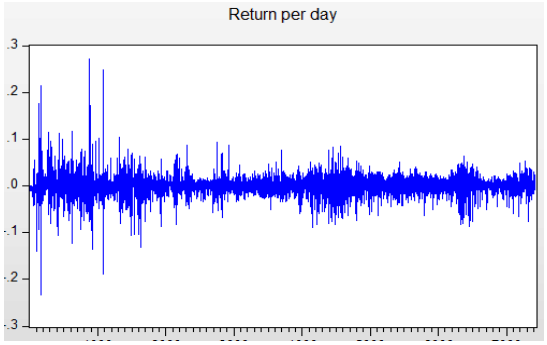
		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-39.79204	0.0000
Test critical values:	1% level	-3.431054	
	5% level	-2.861736	
	10% level	-2.566916	

*MacKinnon (1996) one-sided p-values.

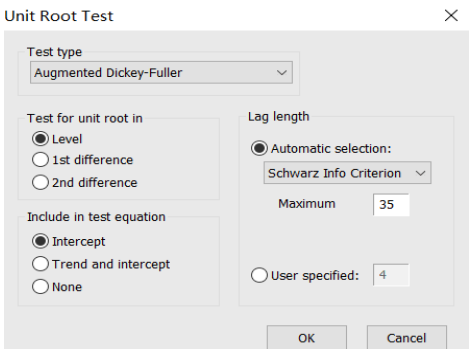
From the graph, we can see that residual is stationary. (From the ADF test, we can also know that residual does not have a unit root.)

g. Forecast.

2. a. Plot the data.



b. Check the data is stationary or not. (Unit root test)



Augmented Dickey-Fuller Unit Root Test on RETURN_PER_DAY

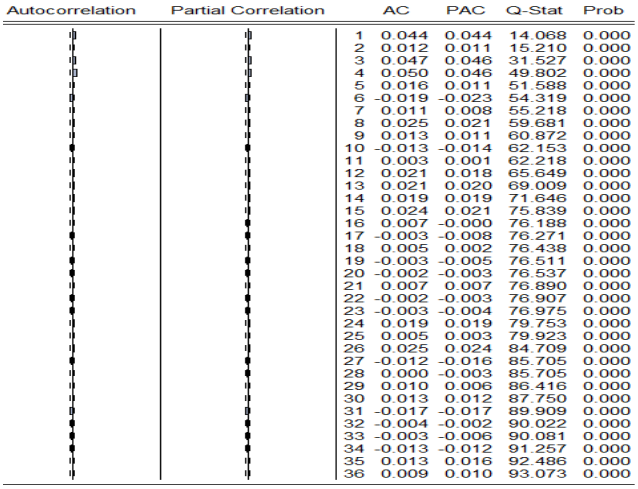
Null Hypothesis: RETURN_PER_DAY has a unit root
Exogenous: Constant
Lag Length: 3 (Automatic - based on SIC, maxlag=35)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-39.05245	0.0000
Test critical values:	1% level	-3.431054	
	5% level	-2.861736	
	10% level	-2.566916	

*MacKinnon (1996) one-sided p-values.

The return is stationary. (t-Statistic -39.05245 < critical values, reject null hypothesis, so return does not have a unit root.)

c. Examine the ACF/PACF.



Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.044	0.044	14.068	0.000
		2	0.012	0.011	15.210	0.000
		3	0.047	0.046	31.527	0.000
		4	0.050	0.046	49.802	0.000
		5	0.016	0.011	51.588	0.000
		6	-0.019	-0.023	54.319	0.000
		7	0.011	0.008	55.218	0.000
		8	0.025	0.021	59.681	0.000
		9	0.013	0.011	60.872	0.000
		10	-0.013	-0.014	62.153	0.000
		11	0.003	0.001	62.218	0.000
		12	0.021	0.018	65.649	0.000
		13	0.021	0.020	69.009	0.000
		14	0.019	0.019	71.646	0.000
		15	0.024	0.021	75.839	0.000
		16	0.007	-0.000	76.188	0.000
		17	-0.003	-0.008	76.271	0.000
		18	0.005	0.002	76.438	0.000
		19	-0.003	-0.005	76.511	0.000
		20	-0.002	-0.003	76.537	0.000
		21	0.007	0.007	76.890	0.000
		22	-0.002	-0.003	76.907	0.000
		23	-0.003	-0.004	76.975	0.000
		24	0.019	0.019	79.753	0.000
		25	0.005	0.003	79.923	0.000
		26	0.025	0.024	84.709	0.000
		27	-0.012	-0.016	85.705	0.000
		28	0.000	-0.003	85.705	0.000
		29	0.010	0.006	86.416	0.000
		30	0.013	0.012	87.750	0.000
		31	-0.017	-0.017	89.909	0.000
		32	-0.004	-0.002	90.022	0.000
		33	-0.003	-0.006	90.081	0.000
		34	-0.013	-0.012	91.257	0.000
		35	0.013	0.016	92.486	0.000
		36	0.009	0.010	93.073	0.000

From the graph, we can not easy to choose AR(p), MA(q) or ARMA(p,q) model.

d. Use the AIC to search model. (Find the smallest AIC)

Dependent Variable: RETURN_PER_DAY
Method: ARMA Maximum Likelihood (OPG - BHHH)
Date: 11/26/19 Time: 01:51
Sample: 1 7424
Included observations: 7424
Convergence achieved after 39 iterations
Coefficient covariance computed using outer product of gradients

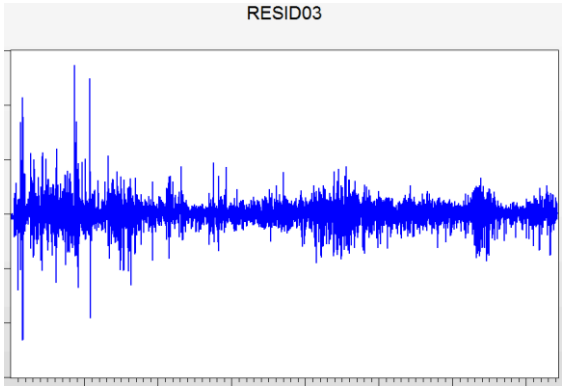
Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR(1)	0.043103	0.005533	7.790303	0.0000
MA(2)	0.008069	0.006001	1.344541	0.1788
SIGMASQ	0.000448	2.43E-06	184.0810	0.0000

R-squared	0.001670	Mean dependent var	0.000380
Adjusted R-squared	0.001401	S.D. dependent var	0.021175
S.E. of regression	0.021160	Akaike info criterion	-4.873002
Sum squared resid	3.322726	Schwarz criterion	-4.870209
Log likelihood	18091.58	Hannan-Quinn criter.	-4.872042
Durbin-Watson stat	2.000193		

Inverted AR Roots	.04	
Inverted MA Roots	-.00+.09i	-.00-.09i

ARMA (1,2) has the smallest AIC that I found.

e. Check if the residual is white noise process.



Null Hypothesis: RESID03 has a unit root
Exogenous: Constant
Lag Length: 3 (Automatic - based on SIC, maxlag=35)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-39.70935	0.0000
Test critical values:	1% level	-3.431054	
	5% level	-2.861736	
	10% level	-2.566916	

*MacKinnon (1996) one-sided p-values.

From the graph, we can see that residual is stationary. (From the ADF test, we can also know that residual does not have a unit root.)

f. Forecast.

Part 3

1. I think the average annual return lower than I would expect from financial theory. From the graph, the average price is decreasing after 2015. I estimate the annual return will go low. However, the price generally shows an upward trend. So people still will buy the stock. Then the price will rise again. My expect from financial theory will go high.
2.

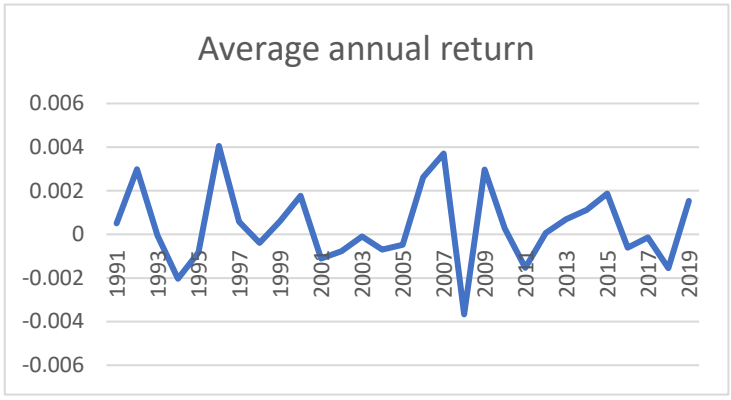
Series: RETURN_PER_DAY
Sample 1 7424
Observations 7424

Mean	0.000380
Median	0.000000
Maximum	0.272152
Minimum	-0.233607
Std. Dev.	0.021175
Skewness	0.446091
Kurtosis	19.35532
Jarque-Bera	82991.81
Probability	0.000000

Year	Average price	Return per day	Average annual return
1991	78.03846154	0	0.000505991
1992	214.3719847	-0.005012542	0.002984082
1993	277.419272	-0.004936789	-4.69864E-05
1994	167.3087846	-0.007909187	-0.00202786
1995	130.8206654	-0.007049117	-0.000832969
1996	220.3527557	-0.005242341	0.004052537
1997	387.1211609	-0.005062779	0.000583187
1998	384.6618199	-0.004880338	-0.000395982
1999	404.0798851	-0.009832715	0.000600356
2000	574.6632115	-0.004952853	0.001761032
2001	582.6687969	-0.004871343	-0.001109109
2002	455.5805057	-0.005001873	-0.000775283
2003	404.4523985	-0.004919796	-0.000101131
2004	380.1419542	-0.010021096	-0.000692393
2005	283.5687538	-0.005102883	-0.000480191
2006	396.3832462	-0.004910261	0.002618071
2007	1128.124153	-0.004934491	0.003702178
2008	903.8584466	-0.004958961	-0.003669321

The standard deviation of the daily returns is 0.021175. annualized s.d. in 1997 is 0.021175 * √0.000583187 = 0.00051138. The annualized s.d. is lower than I would expect from my reading of Section 4.4.

3.



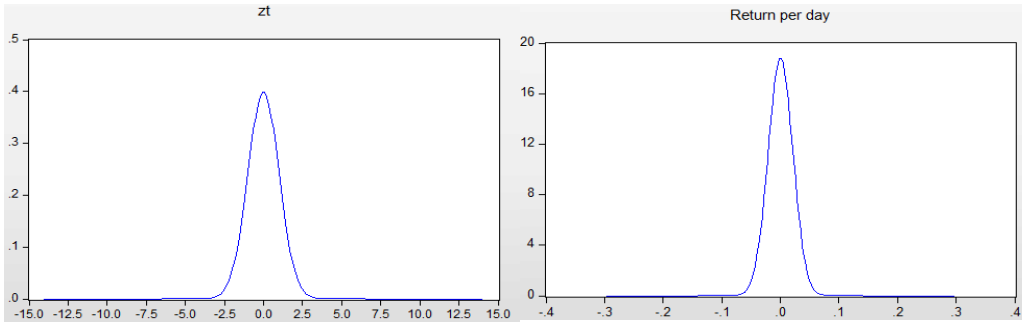
There is no obvious correlation in the different year.

4. 3 standard deviations above the mean = 3 * 0.021175 +0.00038 = 0.063905

3 standard deviations below the mean = 0.00038 – 3 * 0.021175 = -0.063145

I use excel to count them. There are no return more than 3 standard deviations above or below the mean.

5.



Compared with the standarded normal distribution, , estimates of the probability density of daily returns have:

More probability within half a standard deviation (s.d.) of the mean,

More probability three or more s.d. away from the mean.

Less overall probability elsewhere.

6.a). The autocorrelation of the returns

Date: 11/26/19 Time: 06:25
Sample: 1 7424
Included observations: 7424

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.044	0.044	14.068	0.000
		2	0.012	0.011	15.210	0.000
		3	0.047	0.046	31.527	0.000
		4	0.050	0.046	49.802	0.000
		5	0.016	0.011	51.588	0.000
		6	-0.019	-0.023	54.319	0.000
		7	0.011	0.008	55.218	0.000
		8	0.025	0.021	59.681	0.000
		9	0.013	0.011	60.872	0.000
		10	-0.013	-0.014	62.153	0.000
		11	0.003	0.001	62.218	0.000
		12	0.021	0.018	65.649	0.000
		13	0.021	0.020	69.009	0.000
		14	0.019	0.019	71.646	0.000
		15	0.024	0.021	75.839	0.000
		16	0.007	-0.000	76.188	0.000
		17	-0.003	-0.008	76.271	0.000
		18	0.005	0.002	76.438	0.000
		19	-0.003	-0.005	76.511	0.000
		20	-0.002	-0.003	76.537	0.000
		21	0.007	0.007	76.890	0.000
		22	-0.002	-0.003	76.907	0.000
		23	-0.003	-0.004	76.975	0.000
		24	0.019	0.019	79.753	0.000
		25	0.005	0.003	79.923	0.000

The autocorrelation of the absolute returns

Date: 11/26/19 Time: 06:26
Sample: 1 7424
Included observations: 7424

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.298	0.298	657.99	0.000
		2	0.289	0.220	1278.5	0.000
		3	0.257	0.144	1770.2	0.000
		4	0.254	0.124	2249.2	0.000
		5	0.213	0.063	2585.0	0.000
		6	0.189	0.039	2849.4	0.000
		7	0.172	0.031	3070.2	0.000
		8	0.165	0.032	3271.8	0.000
		9	0.142	0.014	3421.3	0.000
		10	0.175	0.067	3650.0	0.000
		11	0.158	0.039	3836.0	0.000
		12	0.123	-0.006	3948.2	0.000
		13	0.141	0.031	4096.3	0.000
		14	0.110	-0.007	4186.5	0.000
		15	0.162	0.068	4381.1	0.000
		16	0.121	0.009	4489.5	0.000
		17	0.139	0.034	4633.6	0.000
		18	0.148	0.044	4797.6	0.000
		19	0.145	0.030	4953.7	0.000
		20	0.162	0.049	5150.0	0.000
		21	0.139	0.010	5294.8	0.000
		22	0.119	-0.009	5400.0	0.000
		23	0.114	-0.006	5496.7	0.000
		24	0.132	0.030	5627.0	0.000
		25	0.170	0.071	5841.8	0.000

The autocorrelation of the squared returns

Date: 11/26/19 Time: 06:27
Sample: 1 7424
Included observations: 7424

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.199	0.199	293.13	0.000
		2	0.167	0.133	499.82	0.000
		3	0.164	0.115	699.28	0.000
		4	0.186	0.127	955.70	0.000
		5	0.116	0.035	1055.7	0.000
		6	0.086	0.010	1110.3	0.000
		7	0.077	0.011	1154.0	0.000
		8	0.049	-0.013	1171.7	0.000
		9	0.030	-0.014	1178.5	0.000
		10	0.069	0.043	1214.2	0.000
		11	0.061	0.029	1241.7	0.000
		12	0.034	0.002	1250.5	0.000
		13	0.034	0.006	1258.9	0.000
		14	0.022	-0.011	1262.5	0.000
		15	0.045	0.022	1277.8	0.000
		16	0.030	0.006	1284.7	0.000
		17	0.043	0.021	1298.2	0.000
		18	0.063	0.043	1327.6	0.000
		19	0.052	0.021	1347.5	0.000
		20	0.047	0.012	1363.7	0.000
		21	0.045	0.007	1378.9	0.000
		22	0.045	0.005	1393.9	0.000
		23	0.052	0.018	1414.3	0.000
		24	0.086	0.057	1469.2	0.000
		25	0.164	0.132	1670.2	0.000

There is almost no correlation between returns for different days. There is positive dependence between absolute returns on nearby days, and likewise for squared returns.

b). The Q-stat is shown on the graph.

The Q-stst is relatively small for the daily return, not reject null hypothesis. That the daily return is iid.

The Q-stst is relatively large for the absolute return, reject null hypothesis. That the absolute return is not iid.

The Q-stst is relatively large for the squared return, reject null hypothesis. That the squared return is not iid.

7. Yes. My returns data provide results that are compatible with the three stylized facts for daily returns.

a.The distribution of returns is not normal. (We can say of the distribution: It is approximately symmetric. It has fat tails.It has a high peak.)

b. There is almost no correlation between returns for different days.

c.There is positive dependence between absolute returns on nearby days, and likewise for squared returns.