Financial Econometrics with R - Midterm

Gautier Pellerin, ID 86352

Introduction

For this assignment, we will be using the CAC40 index (ˆFCHI ticker), from 12th March 1990 to 31st December 2021. The CAC40 is the main index of Paris, France exchange. This makes 6076 trading days (after we remove days without data given, i.e. NaN), 285 months and 23 years.
We have computed summary statistics on our sample daily, monthly and yearly log-returns.

	daily	monthly	annual
Mean	0.0117	0.2160	2.6022
St.Deviation	1.4306	5.3158	22.1415
Diameter.C.I.Mean	0.0360	0.6172	9.0488
Skewness	-0.2150	-0.5962	-0.9301
Kurtosis	9.0324	4.3655	3.7751
Excess.Kurtosis	6.0324	1.3655	0.7751
Min	-13.0983	-19.2254	-55.6527
Quant.5%	-2.2662	-8.9167	-39.5328
Quant.25%	-0.6527	-2.7799	-4.8011
Min	-13.0983	-19.2254	-55.6527
Quant.5%	-2.2662	-8.9167	-39.5328
Quant.25%	-0.6527	-2.7799	-4.8011
Median.50%	0.0474	0.9961	7.1384
Quant.75%	0.7289	3.6220	16.6322
Quant.95%	2.1127	7.4425	25.4076
Max	10.5946	18.3312	41.2933
Jarque.Bera.stat	9257.8789	39.0244	3.8922
Jarque.Bera.pvalue	0.0000	0.0000	14.2827
Lillie.test.stat.D	0.0734	0.0672	0.1533
Lillie.test.pvalue.X100	0.0000	0.3420	17.3106
N.obs	6075	285	23

Table 1 – Summary statistics for log-returns

Stylized Fact 1 : Prices are not stationary

This fact states that mean, variance and Covariance between X_t and X_{t-k} are time-dependant. By applying the Auto-correlation function $\rho(k) = \frac{Cov(X_t,X_{t-k})}{\sqrt{V_{pt}}\sqrt{V_{pt-k}}}$
to our sample, we can see if the prices are time-dependent or not. If we compute the auto-correlations of the daily log prices, Cf 1, we can see our empirical ρ value is around 1, slowly decaying towards 0 while time increase. This means that prices have significant auto-correlation, matching the behavior of a non-stationary time series.The decrease towards 0 is more visible on the monthly-log prices.

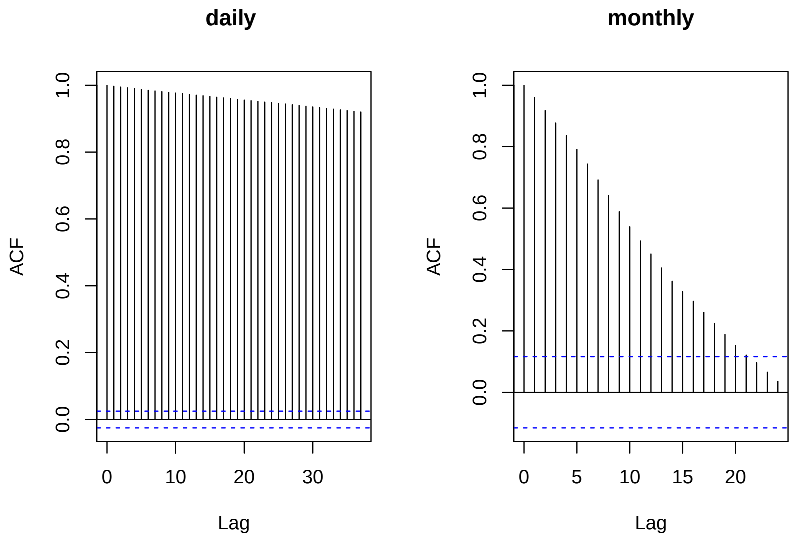

Figure 1 – Auto-correlation of log-prices. The computed values are vastly not in the acceptance region (i.e. the blue bands)

We could also compute the mean and variance from day 0 to day (x) from empirical values, cf 2. Only the mean variation graph is really useful to reinforce our previous analysis the mean value of return vary overtime. By opposition, we will see for stylized fact two the same graph for log-returns, with this time a clear showing of a stationary process.

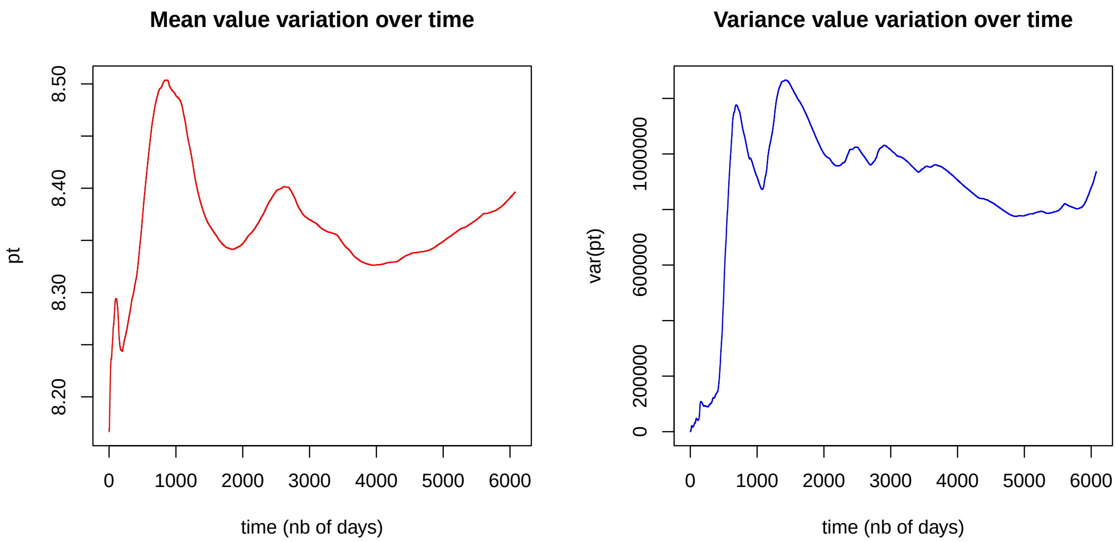

Figure 2 – Mean and Variance values of daily prices between day 0 and day x

Stylized Facts 6 and 2 : returns are not auto-correlated and therefore, stationary

To test if the auto-correlation between daily and monthly return are insignificant (i.e. 0), we compute the Ljung-Box test for lag from 1 to 15. The null hypothesis is the following : $H_0 : \rho = 0$, meaning there is no auto-correlation. We don't test the yearly frequency as the sample size would be too small to be relevant (LB is an asymptotic test).

	daily		monthly		H0 Acceptance		
lag (k)	Ljung-Box stat	LB pvalue	Ljung-Box stat	LB pvalue	$\chi^2_{k,0.95}$	Daily	Monthly
1	1.6460	0.1990	1.8680	0.1720	3.8410	Accept	Accept
2	4.1560	0.1250	2.3440	0.3100	5.9910	Accept	Accept
3	15.0750	0.0020	3.0240	0.3880	7.8150	Reject	Accept
4	17.4060	0.0020	3.4080	0.4920	9.4880	Reject	Accept
5	27.4520	0.0000	4.0330	0.5450	11.0700	Reject	Accept
6	32.4280	0.0000	4.4340	0.6180	12.5920	Reject	Accept
7	35.8980	0.0000	4.5220	0.7180	14.0670	Reject	Accept
8	36.3720	0.0000	5.0880	0.7480	15.5070	Reject	Accept
9	39.7680	0.0000	5.3870	0.7990	16.9190	Reject	Accept
10	40.7570	0.0000	5.4030	0.8630	18.3070	Reject	Accept
11	40.8120	0.0000	5.9330	0.8780	19.6750	Reject	Accept
12	40.8440	0.0000	6.4640	0.8910	21.0260	Reject	Accept
13	41.5410	0.0000	6.8030	0.9120	22.3620	Reject	Accept
14	43.8270	0.0000	7.2330	0.9250	23.6850	Reject	Accept
15	45.4890	0.0000	8.5940	0.8980	24.9960	Reject	Accept

Table 2 – Ljung-Box test for return and result for given lag between time series

The tab display great evidence that monthly return have an unsignificant auto-correlation. On the contrary, the daily frequency only display poor evidence. We can have another analysis before jumping to conclusion. Below are the result of auto-correlation function applied to our samples. For monthly and yearly there is a clear evidences of the unsignificant auto-correlation. The result for the daily frequency this time show a better evidence of an unsignificant auto-correlation, as most values are located inside the acceptance values, meaning $\rho = 0$.

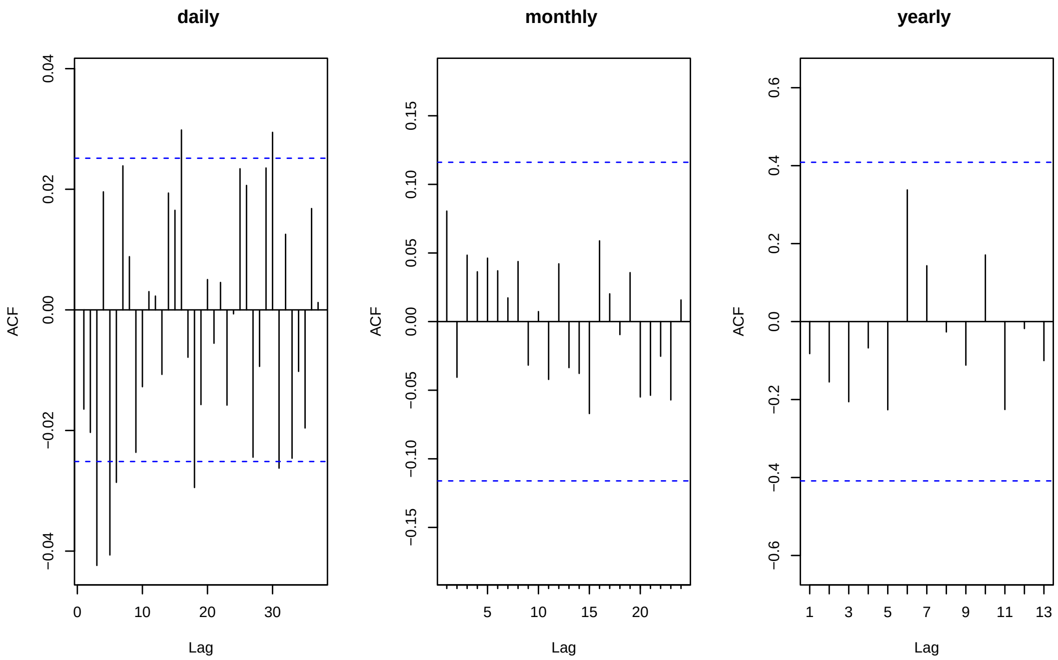

Figure 3 – Auto-correlation for our 3 frequencies.There is clear evidence that the ρ values are in vast majority located in the acceptance region. Please note that we have remove to first auto-correlation (lag 0) as it wouldn't make sense to check if a time series is auto-correlated to itself (0 lag = 0 shift)

We can therefore say that returns are not auto-correlated. This conclusion leads us to stylized fact 2 : by definition of a stationary process,the definition of the auto-correlation function mentioned in stylized fact1 and as the returns are not auto-correlated, we can conclude that the returns are stationary.

There is also empirical evidence that mean and variance are finite and not time dependent.

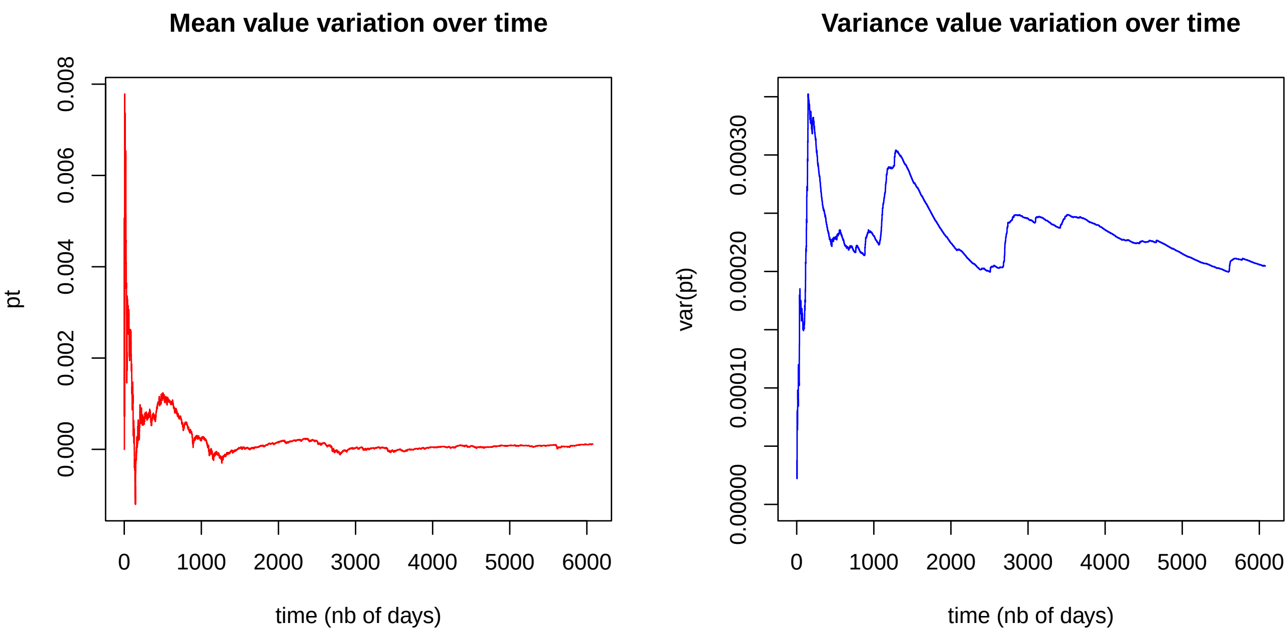

Figure 4 – Mean and Variance values of daily return between day 0 and day x. As t increases, the mean and variances both stabilize. This phenomena is more visible on the mean and would probably be more visible for variance if we had more data available.

.

Stylized Facts 3 and 4 : returns have negative skewness and excess kurtosis

Returns tend to have negative skewness. This mean they are likely to witness negative large negative returns occur more often than large positive ones. On a histogram, the negative skewness will present the following characteristics :
- The left tail is longer
- The mass of the distribution is concentrated on the right side.

Return tend to have excess kurtosis. Thus means large (and small) returns occur more often than expected. On a histogram, we would therefore observe a fat-tailed and peaked distribution.
Below are the histograms for daily and monthly log return, together with a line representing the Normal probability density curve.

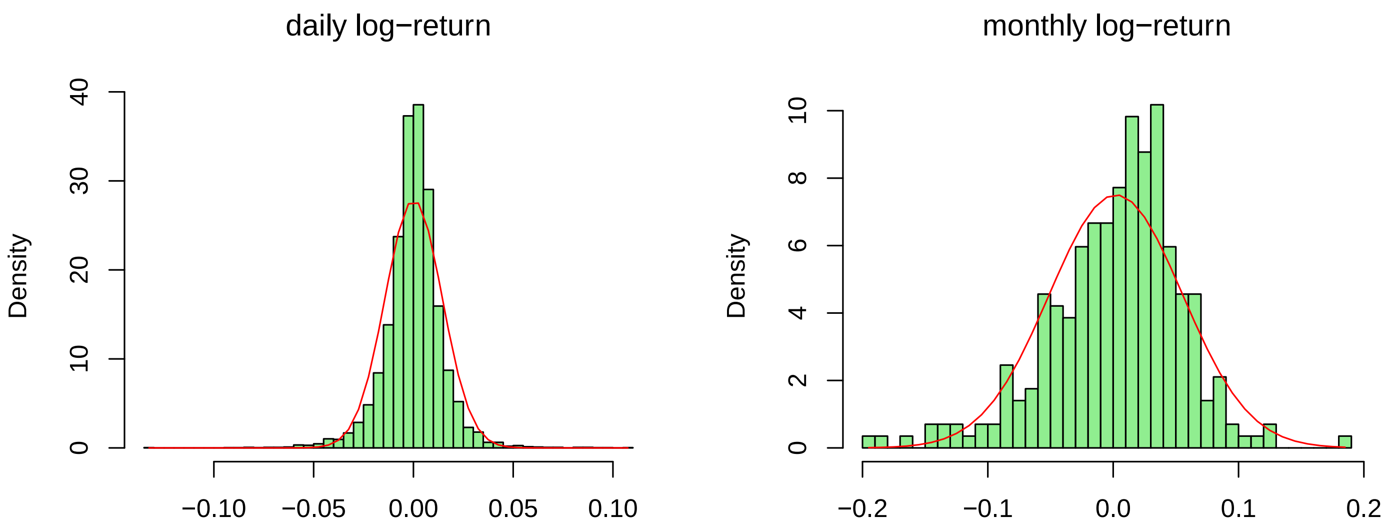

Figure 5 – Histogram distribution of daily and monthly log-returns

We can see that the returns disclose a higher peak (excess kurtosis), the mass is more concentrated on the right side (asymetric). We can also observe that more mass is located in the tails (compared to Normal distribution) : this is called a leptokurtic distribution.
To add more proof of fact 4, we compute the Q-Q plot. The redline represent a Standard Normal distribution and we can see the differences between our sample and the Normal on the tails.

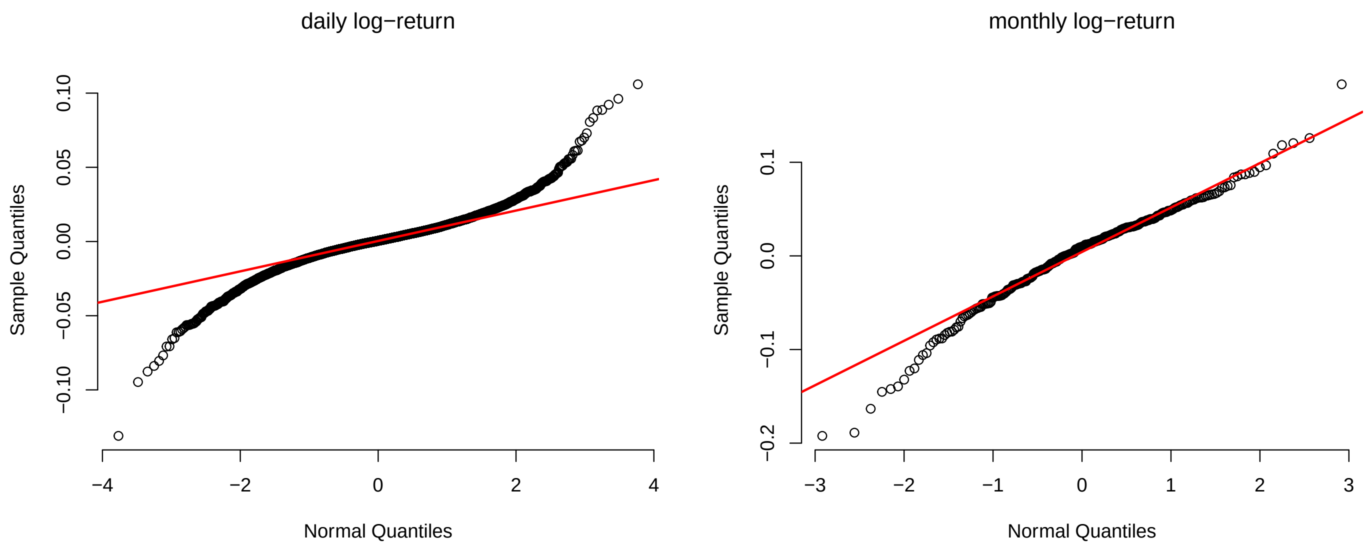

Figure 6 – Quantile-Quantile plot of daily log returns

Stylized Fact 5 : Aggregational Gaussianity

This fact states that the distribution of lower frequencies (yearly) returns tends to be Gaussian, even if the distribution of higher frequencies returns (daily) are not. The negative skewness, large kurtosis and fat tails showed in stylized 3 and 4 actually already gives us information about a departure from Gaussianity for those frequencies. To have further proof, we'll have a look at both the Jarque-Bera and Goodness of fit (Lilliefors) tests.

	daily	montly	annual
Jarque.Bera.stat.X-squared	9257.8789	39.0244	3.8922
Jarque.Bera.pvalue.X100	0.0000	0.0000	14.2827
Lillie.test.stat.D	0.0734	0.0672	0.1533
Lillie.test.pvalue.X100	0.0000	0.3420	17.3106
N.obs	6075	285	23

Table 3 – Jarque Bera and Lillieforst test on returns

The $\chi^2_{1-\alpha}(2)$ value, with $\alpha = 0.05$, is equal to 5.99. As we can see on 3, the Jarque Bera values for daily and monthly are superior to 5.99. We can therefore reject the null hypothesis at the 5% level for those two frequencies, meaning the distribution for daily and monthly are not Gaussian. For yearly, we have to look at Lilliefors test as JB test isn't a normality test. As $0.1533 < \frac{0.886}{\sqrt{23}} = 0.185$, we cannot reject Lilliefors null hypothesis, implying the yearly return are a normal distribution.

Stylized Fact 7 : Volatility clustering

Volatility clustering means that large price changes, i.e. returns with large absolute values or large squares, occur in clusters. This should be reflected by significant auto-correlation of the squared returns and absolute returns. As we did for Stylized Fact 6, we'll compute the Ljung-Box test, both on squared and absolute returns but only for the daily frequency to remain concise.

	Squared		Absolute	
lag (k)	Ljung-Box stat	LB pvalue	Ljung-Box stat	LB pvalue
1	164.4028	0.0000	254.4263	0.0000
2	487.2420	0.0000	701.5208	0.0000
3	927.3334	0.0000	1182.8761	0.0000
4	1252.6494	0.0000	1586.0534	0.0000
5	1592.9798	0.0000	1986.4048	0.0000
6	1818.1339	0.0000	2385.4392	0.0000
7	2003.2333	0.0000	2733.2266	0.0000
8	2250.9570	0.0000	3107.7213	0.0000
9	2513.5427	0.0000	3459.6795	0.0000
10	2748.5650	0.0000	3783.6696	0.0000
11	2960.1238	0.0000	4096.0255	0.0000
12	3159.5214	0.0000	4376.3924	0.0000
13	3322.6025	0.0000	4636.3413	0.0000
14	3425.9654	0.0000	4872.6783	0.0000
15	3541.8451	0.0000	5110.2345	0.0000

Table 4 – Ljung-Box test on squared and absolute returns

As all Ljung-Box pvalue in the table are equal to 0, the null hypothesis is accepted and we have strong evidence that daily squared and daily absolute returns have significant auto-correlation. The empirical auto-correlation function, applied to our sample, confirms our thinking and indicates possible long-memory properties in higher moments of returns. These auto-correlations are then a lot smaller when increasing the sample interval from a day to a month, which is why we do not display them here.

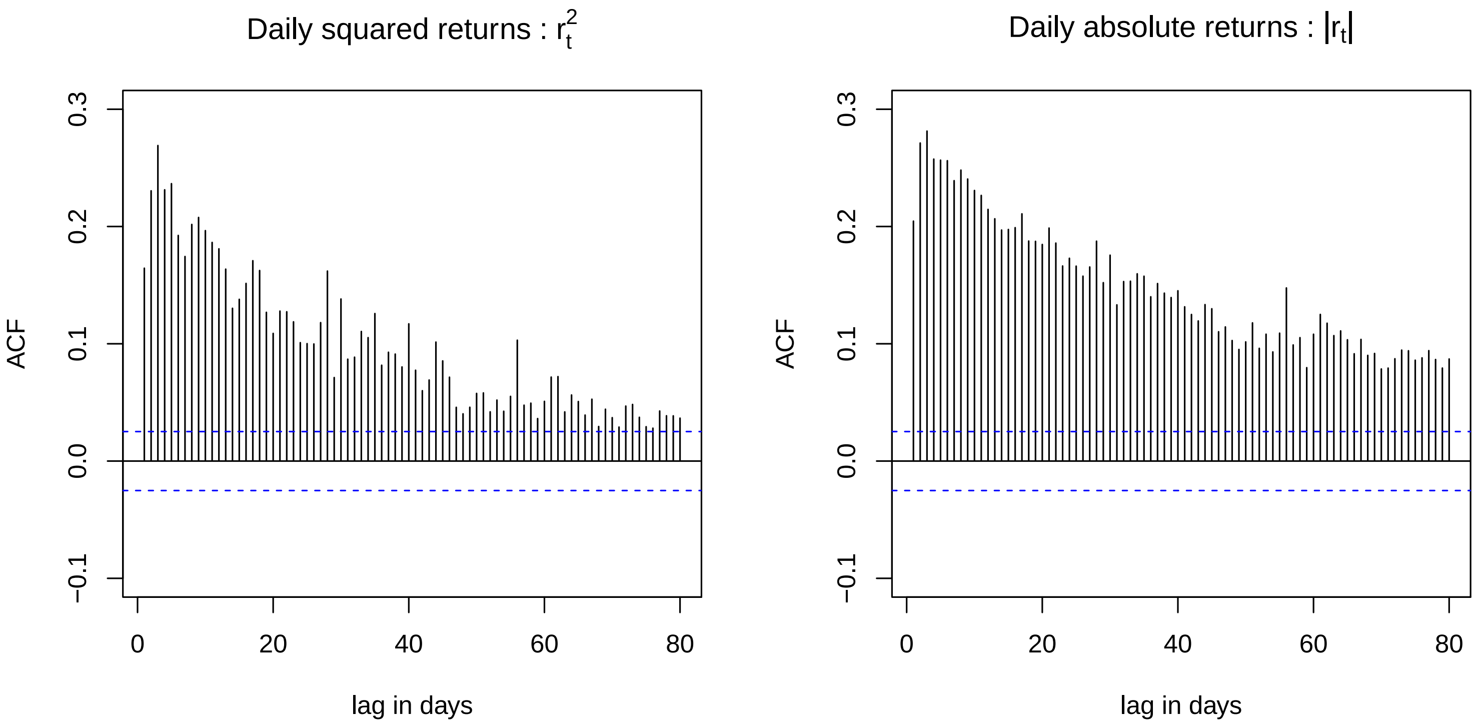

Figure 7 – Auto-correlation of daily squared absolute returns

Stylized Fact 8 : Leverage effect

This fact states that asset returns are negatively correlated with the changes of their volatilities : this negative correlation is called leverage effect. In particular there is a supposed strong empirical evidence that asset-return volatility rises after price declines, with larger declines inducing greater volatility spikes. This is translated by :

$$corr(r_{t-j}, r_t^2) < 0\ ,\ for\ j > 0$$

We will provide evidence of the leverage effect by looking at the cross-correlations between r_{t+j} and r_t^2)

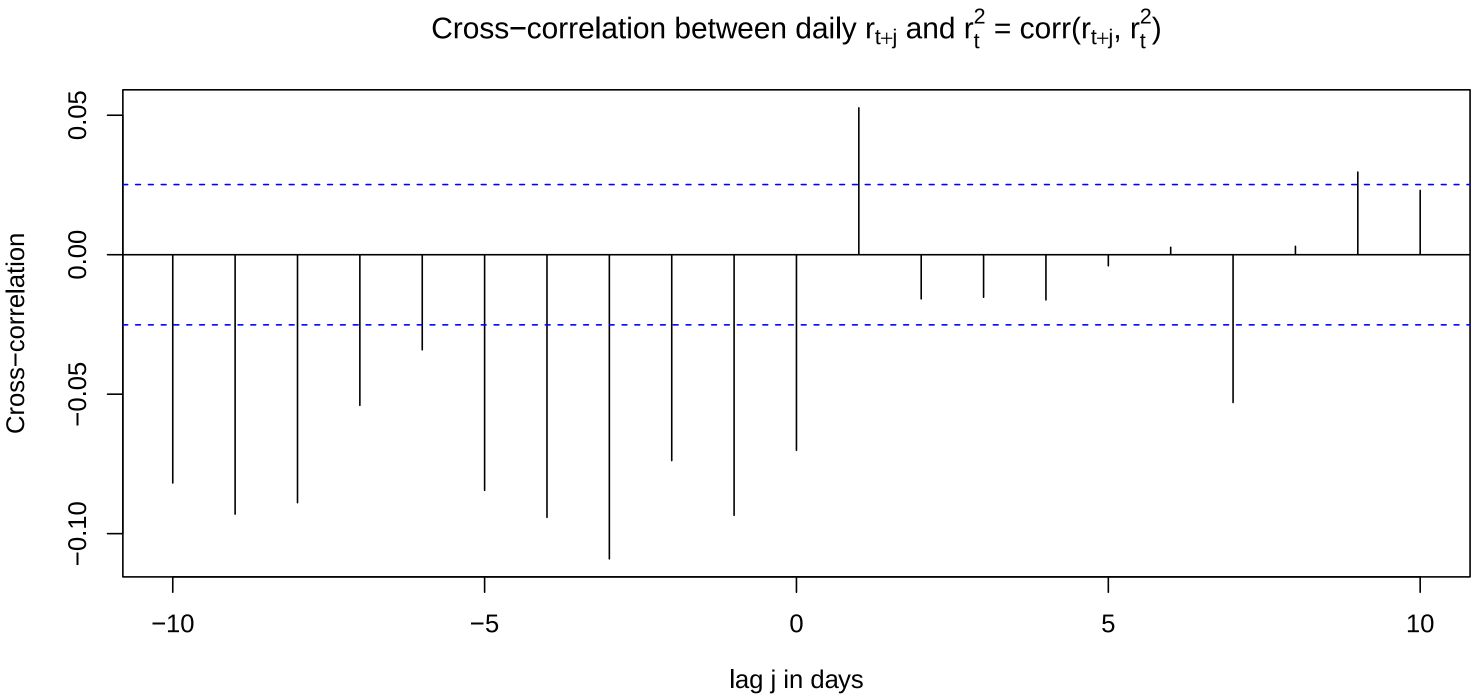

FIGURE 8 – Empirical cross-correlation of daily lagged CAC40 log-returns, with squared returns : $corr(r_{t-j}, r_t^2)$ for j = -10,.., +10. The Dashed blue lines are the (asymptotic) bounds for the rejection region a significance test of each cross-correlation. A line above or below the blue dashed line represent a significant cross-correlation.

As the vast majority of the rejected cross-correlations (outside the blue bands) are negative, we have some strong evidence that the CAC 40 Index has indeed a leverage effect. The CAC 40 Index doesn't have an index with it's options prices, so we cannot compute the cross-correlation between the daily returns and the changes in the implied volatility (the option index).

Conclusion

Throughout this document, we have demonstrated that the CAC 40 Index, over the period from 12th March 1990 to 31st December 2021, present features in line with all the Stylized Facts described in the document.

APPENDIX : R code for Question1