Léa NAPASEUTH
Pierre NIETO

IF4

MARKET RISK REPORT

Question A (Ex2, part of Q1 and of Q2 of TD1)

a – From the time series of the daily prices of the stock Natixis between January 2015 and December 2016, provided with TD1, estimate a historical VaR on price returns at a one-day horizon for a given probability level (this probability is a parameter which must be changed easily). You must base your VaR on a non-parametric distribution (biweight Kernel, that is K is the derivative of the logistic function k.

We computed the simple return, Pt is the closing price for each days during 2015-2016

$$r_t = \frac{P_t - P_{t-1}}{P_{t-1}}$$

Then we established Standard Deviation of the return, the bandwidth h with the silverman's rule, n the total of days during 2015-2016 and we had chosen a fixed alpha quantile.

n	513
sigma	0,0244
Bandwidth	0,0074
alpha	0,05
dx	0,0005
min return	-0,0727
max return	0,2067
x	y-r/h

Then, we generated a grid between max and min return, it will be used to approximate the VaR with a 0,0005 accuracy. Then, we calculated the sum of weight with the Biweight Kernel function for each grid's parameter. We initiated x = y - ri / h with y is a grid's parameter and ri are dailies return values.

$$\sum_{i=1}^{n} K\left(\frac{y - r_i}{h}\right)$$

Then, the density function which is the Biweight Kernel function multiplied by 1/nh.

$$\hat{k}(y) = \frac{1}{nh} \sum_{i=1}^{n} K\left(\frac{y - r_i}{h}\right)$$

Thus, we've obtained the local probability with density multiplied by dx, the linespace in our grid.
With the sum of these local probabilities for each step of the grid, we got the Cumulative distribution function.
We looked for 0.95 CDF which is corresponding to the VaR value on the grid y. We got for the non parametric historical VaR : -3,67%. This represents the percentage of the potential loss of the total value in the 5% worst situation.

Historical VaR	-3,67%

b – Which proportion of price returns between January 2017 and December 2018 does exceed the VaR threshold defined in the previous question? Do you validate the choice of this non-parametric VaR?

The observed proportion of exceeding during the testing period (2017–2018) is around 1%, which is significantly lower than the theoretical threshold of 5% = alpha that we have chosen. Yes, we validate this model from a prudential perspective. The model overestimated the risk. This gap can be attributed to a shift in the volatility regime. The model was calibrated during a highly volatile period (2015–2016, marked by events such as Brexit), which inflated the VaR estimates. The testing period (2017–2018) was stable for the Natixis stock. Consequently, the model was overly conservative for this specific period.

n	510
Exceeding	5
Ratio of exceeding	1,0%

Question B (Ex2, Q5 of TD2)

Calculate the expected shortfall for the VaR calculated in question A. How is the result, compared to the VaR?

We have computed the probability-weighted return k(x)*dx strictly for values below the VaR threshold to isolate the distribution tail.

$$\mathrm{TC}_i = r_i\, w_i\, \mathbf{1}_{\{r_i \le \mathrm{VaR}_\alpha\}}$$

Then we have calculated the average expected loss conditional on exceeding the VaR computed as the sum of tail contributions divided by alpha

$$\mathrm{ES}_\alpha = \frac{1}{\alpha} \sum_{i=1}^{n} r_i\, w_i\, \mathbf{1}_{\{r_i \le \mathrm{VaR}_\alpha\}}$$

The expected shortfall is -4,85%, which is more negative and represents a greater loss than the VaR (-3.67%).

Expected ShortFall	-4,85%

This result is expected: while the VaR only provides the threshold of the worst 5% of cases, the ES calculates the average of all outcomes occurring beyond that threshold.
The gap between the two highlights that the distribution has a fat tail, when an extreme event occurs, it is likely to be significantly more severe than the VaR threshold suggests.

Question C (Ex2, Q1 and Q2 of TD3) With the dataset provided for TD1 on Natixis prices, first calculate daily returns. You will then analyse these returns using a specific method in the field of the EVT. a – Estimate the GEV parameters for the two tails of the distribution of returns, using the estimator of Pickands. What can you conclude about the nature of the extreme gains and losses? b – Calculate the value at risk based on EVT for various confidence levels, with the assumption of iid returns.

- <u>Calculation of daily returns</u>

In the course, we usually use log return formula : $r_t = \ln\left(\frac{P_t}{P_{t-1}}\right)$ We use log here to see both methods to compute the return.

- <u>Analyse of the results</u>

We first verify that the dates are sorted from the most ancient to the most recent. We can observe that there are missing weekends (normal), missing public holidays (normal), no duplicate date and no missing price values. That means that we have no issues. Our dataset is ready to calculate daily returns, the EVT (Extreme Value Theory) and block maxima.
We can see that the average of our price return is around 0 which is a good result. We know that losses are defined as $X_t^{(L)} = -r_t$, while gains are $X_t^{(G)} = r_t$. So we created 2 columns of each.

A	B	C	D	E	F	G	H
Date	Price	Daily returns	Loss	Gain	Month block	Max of the month	
02/01/2015	5,621				2015-01	0	
05/01/2015	5,424	-0,035676033	0,035676	-0,03568	2015-01	0	
06/01/2015	5,329	-0,017669947	0,01767	-0,01767	2015-01	0	
07/01/2015	5,224	-0,019900211	0,0199	-0,0199	2015-01	0	
08/01/2015	5,453	0,042902524	-0,0429	0,042903	2015-01	0	
09/01/2015	5,34	-0,020940263	0,02094	-0,02094	2015-01	0	
12/01/2015	5,264	-0,014334459	0,014334	-0,01433	2015-01	0	
13/01/2015	5,4	0,02550776	-0,02551	0,025508	2015-01	0	
14/01/2015	5,311	-0,016618812	0,016619	-0,01662	2015-01	0	
15/01/2015	5,42	0,020315674	-0,02032	0,020316	2015-01	0	
16/01/2015	5,492	0,013196672	-0,0132	0,013197	2015-01	0	
19/01/2015	5,635	0,02570466	-0,0257	0,025705	2015-01	0	

- <u>Choice of the methods</u>

In the extreme value analysis, we rely on the block maxima approach leading to a GEV modeling framework. The choice of block size is very important because it affects the validity of the asymptotic theory and the stability of parameter estimation. In TDs and mostly TD3, we worked with maxima of blocs, EVT asymptoticals, suppose observations are independent and identically distributed by blocs. So we need blocs that are big enough but not too many. Thus, the best choice was **monthly blocks of around 20 trading days.** Furthermore, ξ is more stable and the interpretation of Fréchet, Gumbel, Weibul are more accurate while using weekly blocks leads to more unstable estimates of the tail index.

- <u>a) Estimation of GEV parameters</u>

In a first sheet named Data, we created columns and calculated all the data that we needed : date, price, dairy returns, loss, gain, month and the maximum of the month . All of this was created in order to create another sheet named EVT max where we remove month date's duplicates and calculate the maximum loss of the month extracted from daily log-returns and who represents the left tail with the formula : =MAXIFS(Data!D:D, Data!F:F, A2) and

$$M_b^{(L)} = \max\{-r_t \text{ over month } b\}$$

Then we did the same with right tail : =MAXIFS(Data!E:E, Data!F:F, A2) and

$$M_b^{(G)} = \max\{r_t \text{ over month } b\}$$

Our results are coherent because we have 48 months, our loss and gain are positive, and are around 0.03 – 0.10.

Month block	loss maximal	gain maximal	loss maximal sorted
2015-01	0,035676033	0,042902524	0,010287043
2015-02	0,016540694	0,030378301	0,010287043
2015-03	0,020654512	0,030229002	0,010287043
2015-04	0,051416947	0,018143871	0,010287043
2015-05	0,039833568	0,02002679	0,010287043
2015-06	0,049459801	0,043323136	0,010287043
2015-07	0,042891565	0,048242136	0,018810349
2015-08	0,054303294	0,062744591	0,01945046
2015-09	0,046029413	0,036600283	0,020654512
2015-10	0,022637095	0,042387034	0,021046319
2015-11	0,02886508	0,050267174	0,022422464

We are now going to estimate the parameter of the tail ξ with the Pickands estimator. We sorted from the smallest to the biggest the loss maxima. We know that the Pickands estimator is based on the largest observations of the sample that we have and that a number k of extreme value has to be chosen. From the course, we know that the number k is supposed to be small as the sample size but large enough to reach the behaviour of the tail distribution and not skip some extreme values. Here, we have 48 blocks maxima so we can choose **k=6** in order to only keep the extreme observation and at the same time have a proportion k/48 small. This k provides us a stable estimation of the tail index by using the Pickands estimator. We choose Xn-k+1, Xn-2k+1, Xn-4k+1 with n=48. So we select X43, X37 and X25.

The formula of Pickands is :

$$\hat{\xi}_P = \frac{1}{\ln 2} \ln \left(\frac{X_{n-k+1:n} - X_{n-2k+1:n}}{X_{n-2k+1:n} - X_{n-4k+1:n}} \right)$$

Pickands Formula k =6
-0,83713068

With this result, we can conclude that ξ <0 so we are in **Weibull Domain** Since the estimated shape parameter is negative, the fitted extreme value distribution falls into the Weibull domain. This corresponds to a bounded tail, meaning that, within this modeling framework, extreme losses would be capped by a finite upper bound. This result should however be interpreted with caution given the relatively small number of block maxima (48 observations) and the sensitivity of tail estimators to the choice of k.

To check the robustness of the Pickands estimator, the tail index is also computed for nearby values of k with k = 5 and k=7 . The sign of the estimate remains unchanged, indicating that the result is not driven by a specific choice of k and that our result remains in the same domain.

E	F	G
Pickands Formula k =6	Pickands Formula k =5	Pickands Formula k =7
-0,83713068	-1,002726241	-0,403699941

We do the same process with the column gain maximal.

I	J	K
ickands Formula k =6	Pickands Formula k =5	Pickands Formula k =7
0,30555206	0,28618706	-0,735722177

The Pickands estimate is positive for k=5 and k=6, but becomes negative for k=7. So besides for k=7, here ξ belongs to the **Fréchet domain**. This indicates that the tail index estimation is sensitive to the choice of k, which is expected given the small sample size (48 monthly maxima). We can interpret the results with caution and keep k=6 as base.

➢ <u>b) Calculation of the value at risk</u>

We start by defining u = Xn-k+1 and Δ=u−Xn−2k+1 and create their columns with p and VaR_EVT(p). We know from the course that by combining EVT, Pickands estimators, hypothesis of independent identically distributed, static order of Xn-k+1, we obtain the formula :

$$\mathrm{VaR}(p) = \frac{\left(\frac{k}{n(1-p)}\right)^{\hat{\xi}} - 1}{1 - 2^{-\hat{\xi}}} (u - X_{n-2k+1:n}) + u$$

L	M	N	O
u	delta	p	VaR_EVT(p)
0,058734	0,009274	0,99	0,06910203
		0,995	0,06972865
		0,999	0,07031829

We obtain :
By using the EVT framework and the Pickands tail index estimated from monthly loss maxima $\xi \approx -0.84$, k=6, n=48, we compute extreme quantiles and obtain the following results VaR(99%) = 6.91%, VaR(99.5%) = 6.97%, and VaR(99.9%) = 7.03%.
The results are coherent because VaR increases moderately as p increases and differences are small. Thus, the VaR is consistent with a negative shape parameter and confirms that extreme losses do not explode in this modeling framework.

- <u>Conclusion</u>

By using Extreme Value Theory and the Pickands estimator on monthly block maxima, we studied the tail behavior of returns. Concerning losses, the estimated shape parameter is negative. This is typical of a Weibull-type tail and means that extreme monthly losses are bounded. For gains, the tail index is positive for the choice k=6 and shows right-tail behavior in the domain of Fréchet. However, the estimation is very sensitive to k. EVT-based Value-at-Risk increases slowly with the confidence level, which is consistent with the negative tail index. It confirms the coherence of the model. Finally, the results highlight the importance of EVT for extreme risk analysis, while putting at the same time attention to the need for caution concerning the limited sample size.

<u>Question D (Ex2, Q5 of TD4)</u> With the dataset and the framework provided for TD4, estimate all the parameters of Bouchaud's price impact model. Comment the obtained values. Is this model well specified?

We know from the course that the expression of Bouchaud's model is :

$$p_t = p_{-\infty} + \sum_{s<t} G(t-s)\,\varepsilon_s\, S_s\, V_s^{\,r}$$

We are taking the dataset of TD4 where transaction date, bid ask spread, volume of the transaction, sign and price are already filled. We calculate dp corresponding at dP = Pt - Pt-1. We have to create the lag of the sign of transactions so we choose a reasonable lag L= 10.
We calculate them with the excel function : OFFSET
The missing values at the beginning are due to the construction of lagged order signs and are discarded. So we are going to use from line 12 for the regression of dP and lag.

transaction date (1=1day=24 hours)	bid-ask spread	volume of the transaction (if known)	Sign of the transaction	Price (before transaction)	Dp	εt-1	εt-2	εt-3	εt-4	εt-5	εt-6	εt-7	εt-8	εt-9	εt-10
0,000202	0,11	8	-1	100											
0,00107	0,103		1	99,984	-0,016	-1	Sign of the tran	#REF!	#REF!	#REF!	#REF!	#REF!	#REF!	#REF!	#REF!
0,001496	0,1015		-1	100,029	0,045	1	-1	Sign of the	#REF!	#REF!	#REF!	#REF!	#REF!	#REF!	#REF!
0,003336	0,092		1	99,979	-0,05	-1	1	-1	Sign of the	#REF!	#REF!	#REF!	#REF!	#REF!	#REF!
0,003952	0,1106		1	100,06	0,081	1	-1	1	-1	Sign of the	#REF!	#REF!	#REF!	#REF!	#REF!
0,004014	0,1028		1	100,16	0,1	1	1	-1	1	-1	Sign of the	#REF!	#REF!	#REF!	#REF!
0,004074	0,1294	32	1	100,164	0,004	1	1	1	-1	1	-1	Sign of the	#REF!	#REF!	#REF!
0,005494	0,0981		-1	100,19	0,026	1	1	1	1	-1	1	-1	Sign of the	#REF!	#REF!
0,005559	0,1119		-1	100,084	-0,106	-1	1	1	1	1	-1	1	-1	Sign of the	#REF!
0,005749	0,0861		-1	100,021	-0,063	-1	-1	1	1	1	1	-1	1	-1	Sign of the transaction
0,008327	0,0969		1	99,989	-0,032	-1	-1	-1	1	1	1	1	-1	1	-1
0,011284	0,1147		1	100,058	0,069	1	-1	-1	-1	1	1	1	1	-1	1
0,011411	0,0812		1	100,116	0,058	1	1	-1	-1	-1	1	1	1	1	-1
0,011814	0,0878		-1	100,136	0,02	1	1	1	-1	-1	-1	1	1	1	1

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0,848955
R Square	0,720724
Adjusted R	0,717871
Standard E	0,039193
Observatio	990

ANOVA

	df	*SS*	*MS*	*F*	*ignificance F*
Regressior	10	3,881014	0,388101	252,6493	4,5E-263
Residual	979	1,503868	0,001536		
Total	989	5,384882			

	Coefficient	*tandard Err*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*ower 95,0%*	*Ipper 95,0%*
Intercept	0,00109	0,001246	0,874846	0,381872	-0,00135	0,003534	-0,00135	0,003534
-1	0,062555	0,00126	49,63051	1,5E-269	0,060081	0,065028	0,060081	0,065028
-1	-0,0038	0,001259	-3,01942	0,002598	-0,00627	-0,00133	-0,00627	-0,00133
-1	-0,00329	0,001258	-2,61338	0,009103	-0,00575	-0,00082	-0,00575	-0,00082
1	0,000588	0,001263	0,465936	0,641365	-0,00189	0,003066	-0,00189	0,003066
1	0,000918	0,001259	0,729151	0,466084	-0,00155	0,003388	-0,00155	0,003388
1	1,72E-05	0,001259	0,013637	0,989123	-0,00245	0,002488	-0,00245	0,002488
1	-0,00022	0,001263	-0,17467	0,861376	-0,0027	0,002257	-0,0027	0,002257
-1	-0,00014	0,001258	-0,11473	0,908683	-0,00261	0,002324	-0,00261	0,002324
1	-0,00236	0,001259	-1,87523	0,061059	-0,00483	0,00011	-0,00483	0,00011
-1	-0,00044	0,00126	-0,35115	0,725549	-0,00292	0,00203	-0,00292	0,00203
	G(l)							

We can see that R^2 ≈0.72 which means that a large part of price variations is explained by past order signs. The F-statistic is very high and the p-value is close to zero. That means that the model is globally significant. Moreover, the number of observations (990) is big enough to consider the regression reliable. The coefficients associated to εt−1 to εt−10 correspond to the estimators of the propagator G(l). The immediate impact G(1) is positive, large and highly significant, while the following coefficients decrease and quickly fade.

These results show that a trade has a strong immediate impact on prices however the impact decreases with time. Indeed, it is coherent with what we studied in class about transient price impact. Since the model explains price changes well without generating false trends, we can consider the model well-specified, even though it remains a simplified representation of real market dynamics. Some coefficients are not statistically significant so that shows the limits of the model and the fact that it does not capture all real life parameters.

Question E (Q2 and Q3 of TD5) a – With Haar wavelets and the dataset provided with TD5, determine the multiresolution correlation between all the pairs of FX rates, using GBPEUR, SEKEUR, and CADEUR (work with the average between the highest and the lowest price and transform this average price in returns on the smallest time step). Do you observe an Epps effect and how could you explain this?

We use the dataset of the TD5 and fill the mid price of GBPEUR, SEKEUR and CADEUR. The instruction indicates *transform this average price in returns on the smallest time step* so we calculate the return Rt= (Mid price t/ Mid price t-1) -1.
We use the Haar Wavelet formula of the course and apply it to returns.

$$A_1 = \frac{x_{2k-1} + x_{2k}}{\sqrt{2}}, \quad D_1 = \frac{x_{2k-1} - x_{2k}}{\sqrt{2}}$$

We do the calculus of A1 in a column named HaarApprox_GBPEUR, HaarApprox_SEKEUR and HaarApprox_CADEUR. Same for D1. As in the definition, we only do for paired coefficients. Then, we compute it for level 2 where we put a gap of 4. At the end, we compute their correlation using the detail formula at level 1 and 2.

A	B	C	D	E	F	G	H	I
GBPEUR Curncy					LEVEL 1		LEVEL 2	
Date	HIGH	LOW	MidPrice_GBPEUR	Return_GBPEUR	HaarApprox_GBPEUR	HaarDetail_GBPEUR	HaarApprox_GBPEUR	HaarDetail_GBPEUR
07/03/2016 09:00	1,2932	1,2917	1,29245					
07/03/2016 09:15	1,294	1,293	1,2935	0,000812411	0,000574461	-0,000574461	-0,000115717	0,000928127
07/03/2016 09:30	1,2943	1,2922	1,29325	-0,000193274				
07/03/2016 09:45	1,293	1,2913	1,29215	-0,00085057	-0,000738109	0,000464779		
07/03/2016 10:00	1,2931	1,2921	1,2926	0,000348257				
07/03/2016 10:15	1,2926	1,2921	1,29235	-0,000193409	0,000109494	0,000383015	-0,000367601	0,000522449
07/03/2016 10:30	1,293	1,2906	1,2918	-0,000425581				
07/03/2016 10:45	1,2917	1,2907	1,2912	-0,000464468	-0,00062936	2,74972E-05		
07/03/2016 11:00	1,292	1,2912	1,2916	0,000309789				
07/03/2016 11:15	1,2921	1,2907	1,2914	-0,000154847	0,000109561	0,000328547	-0,000116109	0,000271051
07/03/2016 11:30	1,2912	1,2905	1,29085	-0,000425894				
07/03/2016 11:45	1,2913	1,2905	1,2909	3,87342E-05	-0,000273764	-0,000328542		
07/03/2016 12:00	1,2919	1,2912	1,29155	0,000503525				
07/03/2016 12:15	1,2927	1,2916	1,29215	0,000464558	0,000684538	2,75535E-05	0,000561429	0,000406654
07/03/2016 12:30	1,2927	1,2918	1,29225	7,73904E-05				
07/03/2016 12:45	1,2928	1,2919	1,29235	7,73844E-05	0,000109442	4,23473E-09		

We repeat that for SEKEUR and CADEUR.

AC	AD	AE
	Correlation	
GBPEUR/SEKEUR	SEKEUR/CADEUR	CADEUR/GBPEUR
	Level 1	
0,128612822	0,200221954	0,290538197
	Level 2	
0,160816794	0,198341304	0,253396835

So yes, we observe an Epps effect. At level 1, correlations between FX returns are quite low, while in level 2, they become higher in general. This indicates that correlations increase as the time scale increases and it is the main characteristic of the Epps effect.

This effect can be explained by market microstructure effects, such as market noise and the fact that different currencies are not traded at exactly the same time.

b – Calculate the Hurst exponent of GBPEUR, SEKEUR, and CADEUR. Determine their annualized volatility using the daily volatility and Hurst exponents.

We worked with the Mid price of each currency and computed their logarithm. Then, we define their increments of 1 by : $\Delta_1 X_t = X_t - X_{t-1}$

where $X_t = \ln(\text{MidPrice}_t)$

We use the scaling of aggregated returns so we compute the return of the price, the aggregate return corresponding to the column R2, their moments M2 and the calculus of H based on these formulas : $M2 = \langle r_t^2 \rangle \quad \text{et} \quad M2^{(2)} = \langle (r_t^{(2)})^2 \rangle$

$$H = \frac{1}{2} \frac{\ln(M2^{(2)}/M2)}{\ln 2}$$

For the volatility, we compute the excel formula : STDEV.S() of the log return and for the annualized we use the formula : sigma annualized = sigma*(252^H).

MidPrice_GBPEUR	LogMid_GBP	d1_Log GBP	d1^2	M2	R2	R2^2	M2_2	Hurst		Volatility	Annualized Vol
1,29245	0,256539642			3,8903E-07			9,9803E-07	0,67961		0,00062	0,026729305
1,2935	0,257351723	0,00081208	6,5948E-07								
1,29325	0,25715843	-0,00019329	3,7362E-08		0,00061879	3,829E-07					
1,29215	0,256307498	-0,00085093	7,2409E-07		-0,00104422	1,0904E-06					
1,2926	0,256655694	0,0003482	1,2124E-07		-0,00050274	2,5274E-07					
1,29235	0,256462267	-0,00019343	3,7414E-08		0,00015477	2,3953E-08					
1,2918	0,256036595	-0,00042567	1,812E-07		-0,0006191	3,8328E-07					
1,2912	0,255572019	-0,00046458	2,1583E-07		-0,00089025	7,9254E-07					
1,2916	0,25588176	0,00030974	9,594E-08		-0,00015483	2,3974E-08					
1,2914	0,255726901	-0,00015486	2,3981E-08		0,00015488	2,3989E-08					
1,29085	0,255300916	-0,00042599	1,8146E-07		-0,00058084	3,3738E-07					
1,2909	0,25533965	3,8733E-05	1,5003E-09		-0,00038725	1,4996E-07					

We repeat that for SEKEUR and CADEUR.
Table of the results :

	GBPEUR	SEKEUR	CADEUR
Hurst exponent	0,679605	0,66759	0,66797
Volatility	0,00062	0,00033	0,00051
Annualized volatility	0,026729305	0,013118465	0,020347012

- For GBPEUR, we obtain H= 0,679605 which is higher than 0.5. This shows that the market has a long memory. That means that big price changes are more likely to be followed by more big changes, rather than being purely random.
- For SEKEUR, the Hurst exponent is around 0.67 and it indicates that there is a persistent behavior at the intraday scale. The volatility is lower than for GBPEUR, both at high frequency and after annualization. It is consistent with the lower volatility of the SEKEUR exchange rate.

- For CADEUR, the Hurst exponent is around 0.67 and shows persistent dynamics at the intraday scale. The volatility is more moderate compared to SEKEUR and GBPEUR. Moreover, the annualized volatility is higher than SEKEUR but lower than GBPEUR, which is economically consistent.

To conclude, the Hurst exponents of GBPEUR, SEKEUR and CADEUR are all above 0.5. That shows a persistent behavior at the intraday scale. With the annualized volatilities computed by using the Hurst exponent, we showed that GBPEUR is the most volatile pair, followed by CADEUR and then SEKEUR, which is consistent.