

Etude d'article: Investigating tail-risk dependence in the cryptocurrency markets: A LASSO quantile regression approach

Yuan Li

Heimana Zhong

Steven Worick

1 Introduction

Le Bitcoin, la première monnaie virtuelle décentralisée a été créée au début de l'année 2009 par une personne utilisant le pseudonyme « Satoshi Nakamoto ». Elle est produite et échangée sans l'influence des gouvernements ou des institutions financières et utilise la technologie de « Blockchain ». C'est-à-dire l'ensemble des transactions depuis l'origine sont répertoriées dans un registre consultable par tous.

Le lundi 5 janvier 2015, Bitstamp a suspendu ses opérations à la suite d'une faille de sécurité par un groupe de pirates informatiques qui ont volé jusqu'à 5,2 millions de dollars de Bitcoin. L'incident a fait chuter le taux de change BTC/USD d'environ 21% le 14 janvier 2015. L'impact de la vente s'est répercuté sur l'ensemble des marchés de la crypto-monnaie, provoquant l'effondrement des prix d'autres crypto-monnaies populaires. Depuis 2015, il y a eu plusieurs autres épisodes similaires où les prix ont fortement évolué dans les deux sens. Ces événements extrêmes décrits dans l'exemple ci-dessus sont, par définition, des événements rares et à fort impact. Ils ont une faible probabilité de se manifester. Cependant, des preuves empiriques suggèrent que les événements extrêmes sur les marchés de la crypto-monnaie se produisent assez fréquemment. Les facteurs provoquant des sauts de prix sur les marchés de la crypto-monnaie sont principalement l'illiquidité, le déséquilibre des flux d'ordres et la domination des traders agressifs.

2 Objectif

Cet article a donc pour objectif d'étudier la connectivité du risque extrême sur les marchés des crypto-monnaies en construisant un réseau de répercussion de ce risque parmi les crypto-monnaies les plus populaires. Les auteurs ont aussi identifié les principaux conducteurs et les principaux récepteurs du risque parmi les crypto-monnaies étudiées. L'auteur a utilisé la technique de régression quantile de Least Absolute Shrinkage and Selection Operator (LASSO) pour analyser la manière dont le risque extrême est transmis, reçu et évolue au fil du temps.

3 LASSO quantile régression

3.1 Données

L'auteur a choisi les crypto-monnaies qui ont une capitalisation boursière de 50 millions de dollars minimale au 31 Décembre 2019 et possèdent au moins 3 ans d'historiques. Les stablecoins sont exclus. Au total, l'échantillon est composé de 21 crypto-monnaies. Les rendements sont calculés à partir des cours journaliers de la période du 1er novembre 2016 au 31 décembre 2019.

3.2 Méthodes

L'utilisation de la régression quantile pour expliquer le comportement des rendements des actifs aux différents quantiles de la distribution est idéale pour examiner le risque extrême. En effet, elle est l'un des principaux outils couramment utilisés par les chercheurs pour analyser le risque d'extrémité des rendements d'actifs.

Régression quantile :

$$\text{VaR}_{q,t}^i = \alpha^i + \theta^{i\top} E_t^{-i} + \omega^i X_{t-1}^i$$

où $\text{VaR}_{q,t}^i$ est la VaR (également appelée la fonction quantile conditionnelle) du rendement de la crypto-monnaie i au q ème quantile au temps t .

X_{t-1}^i est le rendement de la crypto-monnaie i au temps $t - 1$.

E_t^{-i} est un vecteur dont les éléments sont le dépassement des pertes de tous les autres crypto-monnaies du réseau à l'exception de la crypto-monnaie i au moment t .

Le dépassement de la perte de la crypto-monnaie j notée E_t^j , est définie comme suit:

$$E_t^j = \begin{cases} 0, & X_t^j \geq q \text{ ième quantile inconditionnel de } X^j \\ X_t^j, & \text{sinon.} \end{cases}$$

θ^i est le vecteur de rendement. Le j ème élément du vecteur, noté θ_j^i est le coefficient de répercussion de la crypto-monnaie j à la crypto-monnaie i où le rendement de la crypto-monnaie j est en-dessous du q ème quantile de la distribution du rendement. Plus la valeur de θ_j^i est grande, c'est-à-dire plus le choc subi par la crypto-monnaie j est grande, plus la VaR de la crypto-monnaie i augmente en valeur absolue.

Cependant, le modèle n'est pas consistant lorsqu'il a :

- un nombre important de régresseurs
- un nombre petit d'observation
- certains régresseurs ont un impact nul sur les variables

On adopte donc une régression quantile en LASSO. Le LASSO est une technique de sélection de variables pertinentes afin de réduire le nombre de régresseurs.

Étape 1 : L'approche suivie dans l'article consiste à estimer le vecteur de coefficient ξ^i en minimisant le programme d'optimisation suivant :

$$\frac{1}{T} \sum_{t=1}^T [q - I(X_t^i \leq \xi^{i\top} W_t^i)] (X_t^i - \xi^{i\top} W_t^i) + \lambda^i \frac{\sqrt{q(1-q)}}{T} \sum_{k=1}^K |\xi_k^i|$$

λ^i est le coefficient de pénalité du modèle, plus λ^i est grand, plus les régresseurs sont susceptibles d'être éliminés. Si $\lambda^i = 0$: cela revient à faire une régression quantile normale. Pour simplifier le modèle, on fixe λ^i à 17 pour toutes les crypto-monnaies. Il est préférable de normaliser les variables afin que le coefficient de pénalité agisse de manière homogène sur l'ensemble des coefficients de la régression.

X_t^i est le rendement de la crypto-monnaie i au temps t .

$I(X_t^i \leq \xi^{i\top} W_t^i)$ est la fonction d'indicateur qui prend la valeur de 1 si $X_t^i \leq \xi^{i\top} W_t^i$ et de 0 sinon.

K est le nombre de régresseurs dans W_t^i

ξ_k^i est le k ème élément du vecteur de coefficient ξ^i .

Étape 2 : Supprimer les variables du rang k tel que ξ_k^i au rang k inférieur au seuil prédéfini 0.0001

Étape 3 : Refaire la régression avec les variables sélectionnées :

$$\xi^{i\top} W_t^i = \alpha^i + \theta^{i\top} E_t^{-i} + \omega^i X_{t-1}^i$$

Les coefficients de répercussion estimés sont ensuite regroupés dans la matrice de connectivité du risque extrême notée A où l'élément A_j^i prend la valeur du coefficient estimé post-LASSO θ_j^i si la variable est sélectionnée et 0 sinon.

3.3 Résultats

La **Figure 1** présente le degré de connectivité du risque extrême estimé, calculé en tant que le nombre total de coefficients de répercussion pertinents de la matrice A avec VaR au seuil de 1%, 5%, 10% et 20% respectivement à droite et à gauche de la distribution.

Ce tableau met en évidence quelques caractéristiques intéressantes du réseau de risque extrême. Les crypto-monnaies semblent être fortement exposées au risque extrême. Au seuil de 10%, la VaR de la queue gauche d'une crypto-monnaie est affectée par les événements extrêmes provenant d'une moyenne de 7 autres cryptos du réseau tandis que la VaR de la queue droite est également affectée par au moins 8 autres cryptos en moyenne.

Deuxièmement, la connectivité du risque extrême a tendance à être plus forte aux seuils moins extrêmes, bien que cette relation ne soit pas strictement monotone. Aux queues de droite, le nombre de connexions parmi les crypto-monnaies au seuil de 20% est plus du double du nombre correspondant au seuil de 1% tandis qu'aux queues de gauche, le même ratio est presque le triple.

Ces résultats suggèrent que les événements extrêmes - à la fois positifs et négatifs - sont plus probablement dus par les chocs idiosyncratiques des crypto-monnaies que par les événements extrêmes provenant des autres crypto-monnaies du système. Troisièmement, en comparant le nombre de connexions à la queue de droite et à la queue de gauche, on observe qu'il y a plus de connectivité aux queues de droite qu'aux queues de gauche. Ces résultats mettent en évidence l'attrait des investissements dans les crypto-monnaies : les investisseurs profitent non seulement de la segmentation du risque de queue idiosyncratique sur les marchés des crypto-monnaies pendant les baisses, mais aussi des rallyes à l'échelle du marché lors des phases de reprise. Ces caractéristiques peuvent avoir contribué à l'émergence impressionnante des crypto-monnaies en tant que classe d'actifs alternative au cours des dernières années.

Le panneau A de la **Figure 2** présente les 5 principaux récepteurs du risque extrême dans le réseau, mesuré par leur «in-degrees». Le panneau B répertorie les 5 principaux facteurs/conducteur de risque extrême du réseau, capturés par leurs «out-degrees». Le panneau C répertorie les 5 principaux facteurs de risque extrême net du réseau, capturés par leurs «net-degrees».

Principaux récepteurs : Siacoin et NEM

Principal récepteur du risque extrême positif : Lisk

Principaux facteurs : Ethereum et Litecoin

Principal facteur du risque extrême positif : Bitcoin

Principaux facteurs de risque extrême net : Litecoin, Ethereum, NEM, Decred

Principal facteur du risque extrême positif net : Bitcoin

Principaux acteurs du réseau : Litecoin and Ethereum Classic

4 Portefeuille équipondéré contre Portefeuille mean-CVaR

4.1 Première approche

L'auteur a fait une analyse de performance entre un portefeuille équipondéré « buy-and-hold » et un portefeuille mean-CVaR « buy-and-hold » sur la période du 1^{er} mai 2017 au 1^{er} mai 2018. Pour le portefeuille mean-CVaR, les poids sont calculés de telle sorte que le compromis entre le risque et la performance soit optimisé.

Sur un horizon de 1 à 12 mois pour un seuil de VaR de 5%. Ils reportent dans le tableau les performances mensuelles de chaque portefeuille. Le portefeuille équipondéré surperforme le portefeuille mean-CVaR.

4.2 Deuxième approche

La deuxième approche consiste à construction les deux portefeuilles hypothétiques avec un rebalancement périodique des poids et des coûts de transactions.

4.3 Conclusion

La **Figure 3** présente les performances des deux portefeuilles, avec des coûts de transaction à l'ordre de 10bp, le premier portefeuille surperforme le portefeuille mean-CVaR avec respectivement 4900% et 3600% de performances. Avec 50bp, la performance du premier portefeuille est 9 fois meilleur que le portefeuille mean-CVaR. Alors que la stratégie d'optimisation mean-CVaR vise spécifiquement à minimiser le risque extrême, la pondération égale des poids est une approche naïve de diversification. On constate que la stratégie simple à pondération égale réussit à atteindre un risque similaire ou meilleur. La performance du portefeuille mean-CVaR cherchant activement à minimiser le risque extrême indique que la diversification peut être directement accomplie avec une pondération égale.

5 Quelques chiffres autour du sujet

Comme la crypto-monnaie est une création relativement récente (2009), le nombre des littératures portant sur ce sujet est aussi relativement faible. Mais on remarque qu'il y a de plus en plus d'études avec l'impressionnante envolée du bitcoin ces dernières années en analysant le résultat de recherche sur Google Scholar : **Figure 4**

Les résultats de cet article sont très intéressants mais on a trouvé une étude similaire portant sur le marché des crypto-monnaies en février 2020 avec l'article de Qiuhua Xua, Yixuan Zhang, Ziyang Zhang « Tail-risk spillovers in cryptocurrency markets ». Des résultats similaires ont été obtenus.

6 Limites

Quelques limites de l'article peuvent être soulignées. Comme l'auteur le dit, la période d'étude est relativement courte : trois ans d'observations. Donc les résultats doivent être interprétés avec précaution. Deuxièmement, une des limites du LASSO est qu'en cas de fortes corrélations entre les variables, LASSO en choisit une, masquant ainsi l'influence des autres. En observant la **Figure 5**, on peut remarquer qu'une corrélation importante entre ces cinq monnaies. Troisièmement, l'hypothèse de ne pas pouvoir vendre à découvert dans la construction des portefeuilles est aussi une des limites de l'analyse. Quatrièmement, l'article n'a pas précisé la différence de définition de la fonction du dépassement alors il paraît évident que pour la régression à la queue de droite, il faut utiliser la fonction de dépassement de gain au lieu de celle de perte.

7 Extensions

7.1 Ridge et Elastic-net

Une extension de l'article qu'on a décidé de faire est la régression Ridge et la régression Elastic-net. Si LASSO exerce une contrainte sur la norme l^1 des coefficients, Ridge met en place une contrainte sur la norme l^2 . Contrairement à LASSO, Ridge a tendance à réduire les variables sans les annuler (shrinkage). LASSO sélectionnera donc des modèles plus parcimonieux. La régression Elastic-net combine les avantages de Ridge et de LASSO : elle a la capacité de sélection de variables en annulant les coefficients et en même temps fait un partage de poids en cas de variables corrélées comme Ridge.

Les fonctions d'optimisations sont définies ci-dessous :

$$\text{LASSO}(l^1 \text{ pénalisation}) : \frac{1}{T} \sum_{t=1}^T [q - I(X_t^i \leq \xi^{iT} W_t^i)] (X_t^i - \xi^{iT} W_t^i) + \lambda \frac{\sqrt{q(1-q)}}{T} \sum_{k=1}^K |\xi_k^i|$$

$$\text{Ridge}(l^2 \text{ pénalisation}) : \frac{1}{T} \sum_{t=1}^T [q - I(X_t^i \leq \xi^{iT} W_t^i)] (X_t^i - \xi^{iT} W_t^i) + \lambda \frac{\sqrt{q(1-q)}}{T} \sum_{k=1}^K \xi_k^{i2}$$

$$\text{Elastic-net}(l^1 + l^2 \text{ pénalisation}) : \alpha l^1 + (1 - \alpha) l^2$$

La **Figure 6** résume les résultats de régression avec les 3 techniques de régularisation avec $\alpha=0.85$, le critère de sélection $=10^{-4}$ et $\lambda=17$. Nous observons qu'effectivement Ridge a tendance à garder plus de coefficient que LASSO et Elastic-net. Le résultat obtenu

avec LASSO et Ridge correspond au principal résultat de l'article : on observe qu'il y a plus de connectivité aux queues de droite qu'aux queues de gauche.

En choisissant un critère de sélection de 0.05, on obtient la **Figure 7** avec Ridge : cette fois-ci, on retrouve l'asymétrie de connexions entre les queues de droites et les queues de gauche.

7.2 Estimateur de Hill

Le phénomène de queues épaisses étant très marqué dans le marché des crypto-monnaies, nous avons décidé de calculer la VaR via l'estimateur de Hill, basée sur la théorie des Valeurs Extrêmes et prenant en compte les phénomènes de queues épaisses.

$$\xi_{k(n),n}^H = \frac{1}{k(n)} \sum_{i=n-k(n)+1}^n \log \left(\frac{X_{i:n}}{X_{n-k(n)+1:n}} \right)$$

$$\text{VaR}(p) = \left(\frac{k}{n(1-p)} \right)^{\xi^H} X_{n-k+1:n}$$

où ξ^H est l'estimateur de Hill du paramètre de la GEV.

La **Figure 8** résume les VaR simulées avec l'estimateur de Hill, en moyenne, la VaR à 99% est à 38.90% alors la VaR à 1% est à -2.55%. On observe que les VaR simulées indiquent une asymétrie de la distribution des crypto-monnaies. Dans l'article étudié, l'auteur a utilisé les quantiles des rendements historiques dans la définition de la fonction du dépassement de perte E_t^j , sachant que l'historique choisi est particulièrement important car aucune autre modélisation ne s'y ajoute. Afin de prendre en compte les phénomènes de queues épaisses et d'éviter l'utilisation de quantile historique, nous avons décidé d'utiliser la VaR estimée par l'estimateur de Hill pour définir le dépassement de perte et de gain de la crypto-monnaie j notée E_t^j , est redéfinie comme suit :

Pour les pertes :

$$E_t^j \begin{cases} 0, X_t^j \geq \text{VaR}^{\text{Hill}}(q) \text{ de } X^j \\ X_t^j \end{cases}$$

Pour les gains :

$$E_t^j \begin{cases} 0, X_t^j \leq \text{VaR}^{\text{Hill}}(q) \text{ de } X^j \\ X_t^j \end{cases}$$

La **Figure 9** indique que la relation est toujours vérifiée en utilisant des quantiles simulés avec l'estimateur de Hill

8 Conclusion

En utilisant la technique de la régression quantile LASSO, l'article met en évidence le caractère asymétrique de la dépendance du risque extrême : contrairement aux phénomènes généralement observés dans les autres classes d'actifs, la dépendance à la queue de droite est plus prononcée que celle à la queue de gauche. L'auteur a aussi identifié les principaux récepteurs et les principaux facteurs de risque dans le réseau en analysant leurs « In-Degree », « Out-Degree » et « Net-Degree ». En termes de stratégie de construction de portefeuille dans le marché des crypto-monnaies, la performance du portefeuille mean-CVaR cherchant activement à minimiser le risque extrême indique que la diversification peut être directement accomplie avec une pondération égale.

En termes d'extension, nous avons utilisé la régression Ridge et Elastic-net afin d'étudier l'impact des techniques de régularisation sur les résultats obtenus. La relation d'asymétrie est vérifiée pour Elastic-net au seuil de 10^{-4} et pour Ridge au seuil de 0.05. Nous avons redéfini la fonction de dépassement de perte avec les VaR simulées avec l'estimateur de Hill afin de modéliser le phénomène de queues épaisses. Nous retrouvons le phénomène d'asymétrie.

9 Annexe

FIGURE 1: Le risque de queue totale simultanée du réseau des crypto-monnaies. Échantillon complet analyse statique : 21 crypto-monnaies.

	Right tail	Left tail	Difference
1% VaR	73	39	34
5% VaR	182	85	97
10% VaR	174	140	34
20% VaR	162	106	56

FIGURE 2

Cryp- tocur- rency	1% left tail	Cryp- tocur- rency	5% left tail	Cryp- tocur- rency	10% left tail	Cryp- tocur- rency	20% left tail	Cryp- tocur- rency	20% right tail	Cryp- tocur- rency	10% right tail	Cryp- tocur- rency	5% right tail	Cryp- tocur- rency	1% right tail
Panel A: In-Degree															
Dogecoin	8	Siacoin	12	Waves	13	Siacoin	12	NEM	16	NEM	16	Bitcoin	13	Ethereum Classic	9
Ethereum Classic	7	Dash	11	Siacoin	12	Zcash	11	MonaCoin	14	Dogecoin	13	Dogecoin	13	DigiByte	8
Monero	7	Bytecoin	8	Bytecoin	10	Litecoin	10	Stellar	13	Ethereum	13	Lisk	13	Lisk	8
Augur	6	NEM	7	Litecoin	9	Monero	10	Bytecoin	11	MonaCoin	12	NEM	13	Verge	8
Neo	4	Litecoin	7	NEM	9	Ethereum Classic	9	Siacoin	11	Lisk	11	Siacoin	13	Augur	7
Panel B: Out-Degree															
NEM	7	Ethereum Classic	9	Ethereum	16	NEM	12	Ethereum	18	Lisk	17	Litecoin	16	Lisk	8
DigiByte	3	Decred	8	Ethereum Classic	12	Ethereum	11	Bitcoin	15	Bitcoin	15	Bitcoin	15	Bitcoin	7
Ethereum	3	Augur	7	Litecoin	12	Siacoin	8	Lisk	15	Litecoin	14	Dash	13	Waves	6
Ethereum Classic	3	Ethereum	7	NEM	12	Dash	7	Waves	14	Waves	14	Ethereum	12	Ethereum	5
Litecoin	3	Waves	7	Monero	8	Monero	7	Litecoin	13	Ethereum	13	Lisk	12	Ethereum Classic	5
Panel C: Net-Degree															
NEM	7	Ethereum Classic	9	Ethereum	11	Ethereum	11	Bitcoin	15	Bitcoin	13	Augur	11	NEM	5
DigiByte	3	Decred	8	Decred	7	NEM	11	Ethereum	13	Decred	7	Decred	9	Waves	5
Ethereum	3	Waves	7	Ethereum Classic	4	Waves	5	Litecoin	12	Ethereum Classic	6	DigiByte	5	Dash	4
Litecoin	3	Augur	4	Litecoin	3	Lisk	4	Lisk	11	Lisk	6	Litecoin	4	Bitcoin	3
Stellar	3	Neo	4	NEM	3	Neo	4	Dash	8	Dash	4	Zcash	3	Dogecoin	3

FIGURE 3

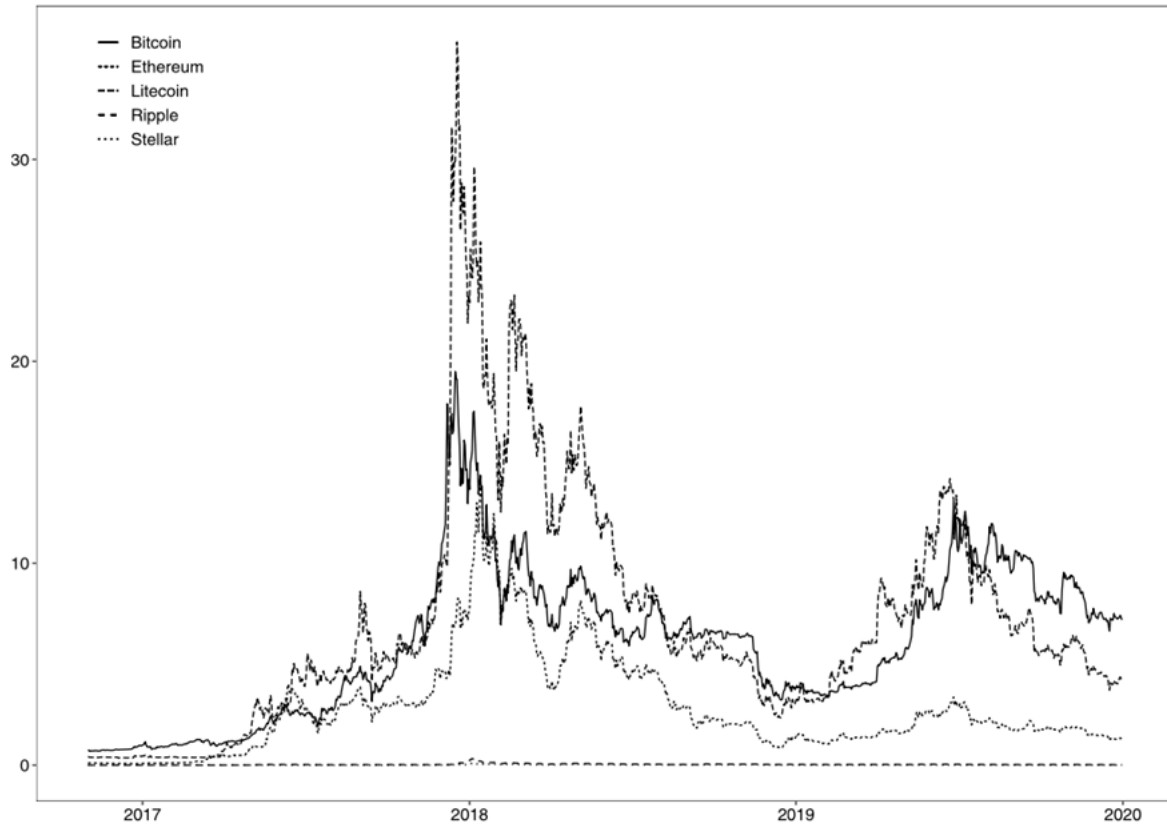
	1 month		3 months		6 months	
	Equal-Weighted	Mean-CVaR	Equal-Weighted	Mean-CVaR	Equal-Weighted	Mean-CVaR
<i>Transaction cost = 0 bps</i>						
Total return (%)	4,959.086	6,121.068	4,882.809	4,500.913	3,521.085	3,735.224
Average return (%)	0.576	0.586	0.596	0.564	0.575	0.566
Volatility (%)	5.916	5.726	6.284	5.890	6.465	6.235
CVaR (%)	-12.588	-12.212	-13.414	-12.491	-14.011	-13.357
Sharpe ratio	1.840	1.935	1.793	1.809	1.682	1.715
CVaR Sharpe ratio	0.865	0.907	0.840	0.853	0.776	0.801
<i>Transaction cost = 10 bps</i>						
Total return (%)	4,913.298	3,610.819	4,853.566	274.817	3,505.563	269.728
Average return (%)	0.575	0.708	0.595	0.337	0.575	0.333
Volatility (%)	5.915	8.442	6.285	6.445	6.465	6.378
CVaR (%)	-12.588	-15.815	-13.416	-13.275	-14.011	-13.023
Sharpe ratio	1.837	1.588	1.791	0.982	1.681	0.980
CVaR Sharpe ratio	0.863	0.848	0.839	0.477	0.776	0.480
<i>Transaction cost = 25 bps</i>						
Total return (%)	4,845.409	488.607	4,810.031	306.066	3,482.408	307.332
Average return (%)	0.573	0.395	0.594	0.348	0.574	0.350
Volatility (%)	5.915	6.602	6.285	6.481	6.465	6.496
CVaR (%)	-12.589	-13.097	-13.420	-13.052	-14.011	-13.098
Sharpe ratio	1.832	1.125	1.788	1.010	1.679	1.011
CVaR Sharpe ratio	0.861	0.567	0.838	0.501	0.775	0.501
<i>Transaction cost = 50 bps</i>						
Total return (%)	4,734.344	486.966	4,738.339	307.611	3,444.161	304.891
Average return (%)	0.571	0.396	0.593	0.349	0.573	0.349
Volatility (%)	5.915	6.628	6.286	6.483	6.466	6.499
CVaR (%)	-12.591	-13.118	-13.425	-13.056	-14.011	-13.099
Sharpe ratio	1.825	1.124	1.784	1.011	1.675	1.009
CVaR Sharpe ratio	0.857	0.568	0.835	0.502	0.773	0.501

FIGURE 4: Résultats de recherche sur Google Scholar

Mot-clé	cryptocurrency
2009-2014	2170 résultats
2015-2021	18800 résultats

Mot-clé	tail risk crypto
	9600 résultats

FIGURE 5



(a) Price

FIGURE 6

	LASSO		RIDGE		ELASTIC-NET	
	<i>Right Tail</i>	<i>Left Tail</i>	<i>Right Tail</i>	<i>Left Tail</i>	<i>Right Tail</i>	<i>Left Tail</i>
1% VaR	381	336	420	420	352	352
5% VaR	341	286	420	419	359	297
10% VaR	325	297	420	420	339	310
20% VaR	330	317	419	420	329	315

FIGURE 7

	RIDGE _{seuil=0.05}	
	<i>Right Tail</i>	<i>Left Tail</i>
1% VaR	268	223
5% VaR	316	246
10% VaR	280	230
20% VaR	224	204

FIGURE 8

	mean VaR Hill
99%	0.3890
95%	0.1999
90%	0.1519
80%	0.1162
20%	-0.0304
10%	-0.0294
5%	-0.0281
1%	-0.0255

FIGURE 9

	LASSO _{Hill}	
	<i>Right Tail</i>	<i>Left Tail</i>
1% VaR	373	340
5% VaR	340	288
10% VaR	312	304
20% VaR	328	313