

Cointégration et construction de stratégie action

Adrien Cortes
Louis-Marie Lovichi

April 4, 2024

Abstract

L'objet de ce travail est l'étude de l'article *The Cointegration Alpha: Enhanced Index Tracking and Long-Short Equity Market Neutral Strategies* dans le cadre *Allocation d'actifs et arbitrage multi-asset*.

Cet article propose différentes applications de stratégies d'investissement qui se basent sur la cointégration, on note :

- une stratégie classique de suivi d'indices;
- une stratégie d'actions long-short market neutral;
- une combinaison des deux approches précédentes.

L'idée de la première stratégie est de reproduire les rendements et la volatilité d'un indice de référence spécifique (exemple du CAC40), c'est-à-dire un portefeuille ayant de propriétés similaires à celui de l'indice en termes de rendements, volatilité et corrélation; tandis que les deux stratégies suivantes visent à minimiser la volatilité et générer des rendements stables quelles que soient les conditions de marché.

L'avantage de se baser sur la cointégration pour l'optimisation de portefeuille est qu'elle permet d'utiliser toutes les informations comprises dans les prix et de construire des poids sur le comportement long terme des actions. Par ailleurs, l'estimation des poids du portefeuille basée sur la cointégration permet une certaine stabilité, cela tend à ignorer les mouvements de prix à court-terme tels que le bruit.

L'objet de ces expériences numériques est de montrer que les techniques de cointégration produisent des résultats prometteurs dans différentes stratégies d'investissement que cela soit le tracking d'indice ou les stratégies long-short market neutral.

Nous passons en revue les méthodes et les backtests introduites dans l'article et essayons de transposer l'approche de l'article au marché français. Nous en commenterons notamment les similitudes et les différences.

1 Index Tracking

L'article étudiait le cas de l'indice DJIA. Nous tentons cette approche sur les données du CAC40.

Data Les données à disposition sont toutes les actions du CAC 40 en données journalières depuis janvier 2001 ainsi que les pondérations dans l'indice. Les changements les plus significatifs ont lieu lors de la crise des subprimes ainsi que lors de la crise du COVID.

Principe de l'Index Tracking L'Index Tracking présenté par l'article repose sur un suivi basé sur la cointégration plutôt que la corrélation, cela est justifié par plusieurs caractéristiques importantes des systèmes cointégrés :

- La réversion à la moyenne de l'erreur de suivi : intrinsèquement, l'erreur de suivi est réversible vers la moyenne, ce qui contribue à sa stabilité.
- La stabilité des pondérations des actions : les pondérations des actions dans le portefeuille restent stables, ce qui réduit le besoin de transactions de rééquilibrage fréquentes et les coûts associés.

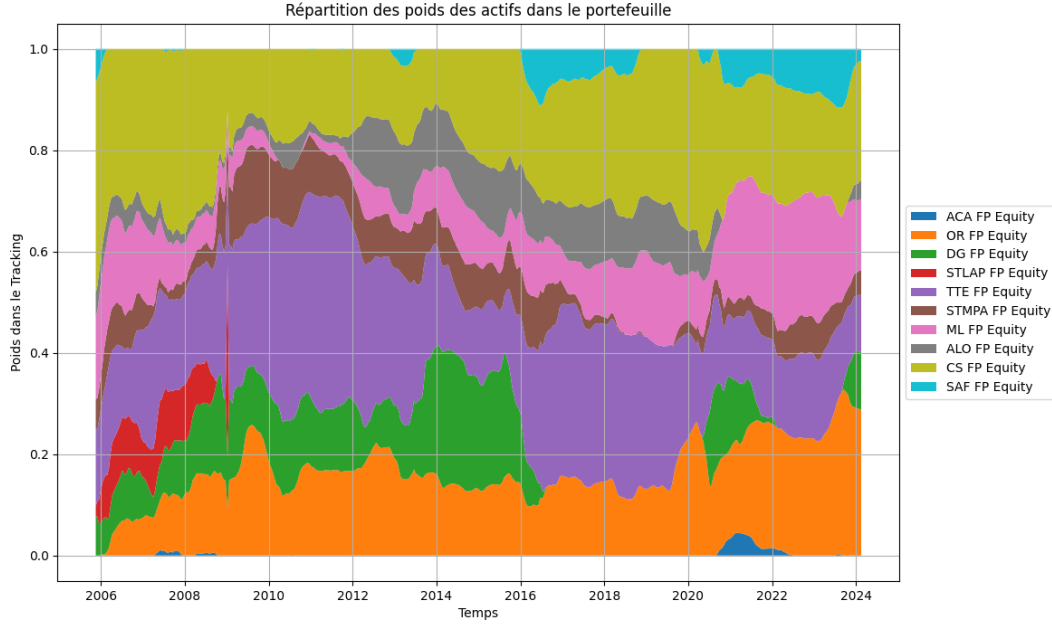


Figure 1: Poids de nos 10 assets dans notre portefeuille.

- Utilisation améliorée de l'information : la cointégration permet d'utiliser de manière plus efficace l'information contenue dans les prix.

L'Index Tracking repose sur deux étapes principales :

- **La sélection des actions** qui est souvent effectuée à la discrétion de l'expertise du portfolio manager. Cette sélection impacte grandement la performance du suivi que l'on peut obtenir.
- **La détermination des positions du portefeuille** qui s'effectue en estimant les coefficients obtenus par moindres carrés ordinaires (MCO). Il s'agit de régresser le prix logarithmique de l'indice sur les prix logarithmiques des actions sélectionnées sur une certaine période de calibration. La transformation logarithmique est appliquée afin de rendre les séries plus homogènes (par ailleurs, les logarithmes des variables cointégrées restent cointégrés). Cette régression s'écrit mathématiquement comme suit :

$$\log(P_{\text{index},t}) = \alpha + \sum_{i=1}^n \beta_i \log(P_{i,t}) + \varepsilon$$

où :

- $\log(P_{\text{index},t})$ est le logarithme du prix de l'indice,
- $\log(P_{i,t})$ est le logarithme du prix de l'action i sélectionnée dans le portefeuille,
- α est l'ordonnée à l'origine,
- β_i sont les coefficients des actions sélectionnées dans le portefeuille,
- ε est le terme d'erreur.

Une remarque importante est que les résidus ε sont stationnaires si et seulement si $\log(P_{\text{index},t})$ et $\sum_{i=1}^n \beta_i \log(P_{i,t})$ sont cointégrés.

Back-test Le Back-test consiste à effectuer une période de calibration de 5 années et à effectuer la régression tous les 10 jours glissants. Les performances de ce back-test se basent sur plusieurs critères, par souci de concision, nous évoquerons uniquement ces critères à la partie suivante sur les stratégies long-short. Pour ces backtests, nous nous sommes limités à 10 assets parmi les 40.

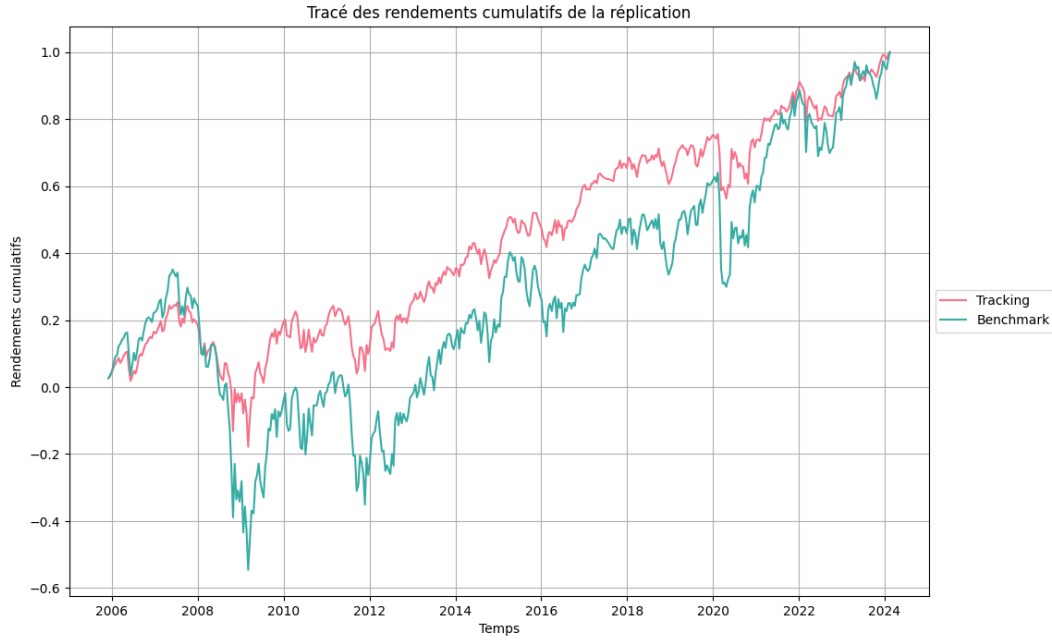


Figure 2: Réplication des rendements cumulés du CAC40.

Rapide commentaire des résultats Dans le cas d'un tracking basé sur 10 assets, on observe sur la figure 1 que le portefeuille ne présente effectivement que peu de variations abruptes conformément à ce qui était remarqué dans l'article. L'augmentation à respectivement 20 et 25 assets parmi les 40 conduit aux pondérations suivantes 3 et 4, ces pondérations commencent à être stables à partir de 2021 lorsque l'intégralité des assets sont présent dans l'indice. La figure 2 fournit la réplcation de l'indice dans le cas de 10 assets utilisés. Cela souligne comme indiqué dans l'article que 10 assets parmi 40 présentent des résultats convenables mais qui doivent être parfaits en ajoutant davantage d'assets au tracking.

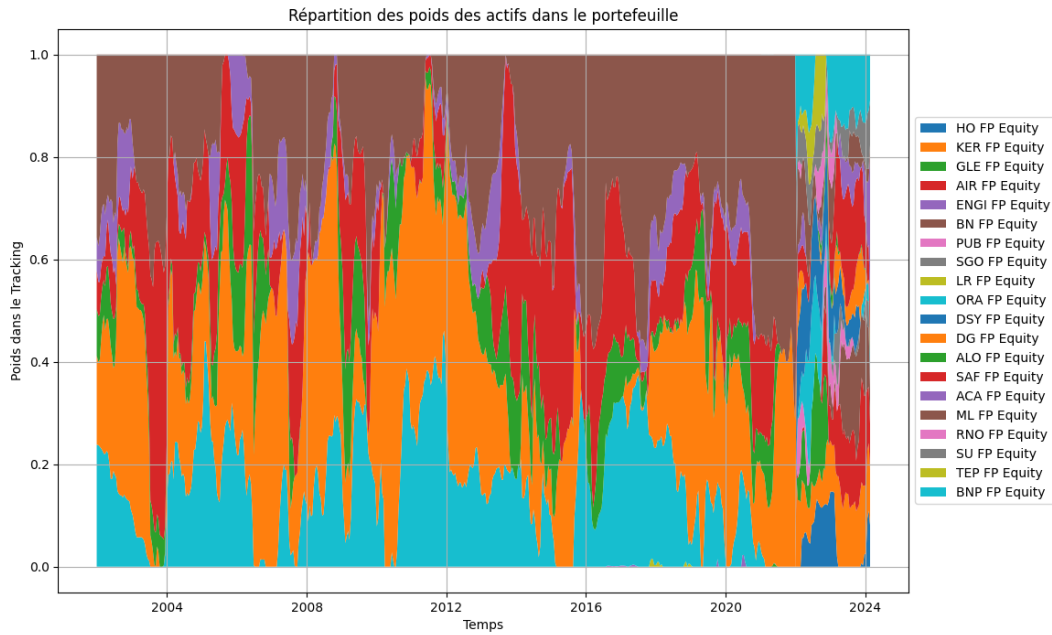


Figure 3: Poids de nos 20 assets dans notre portfolio.

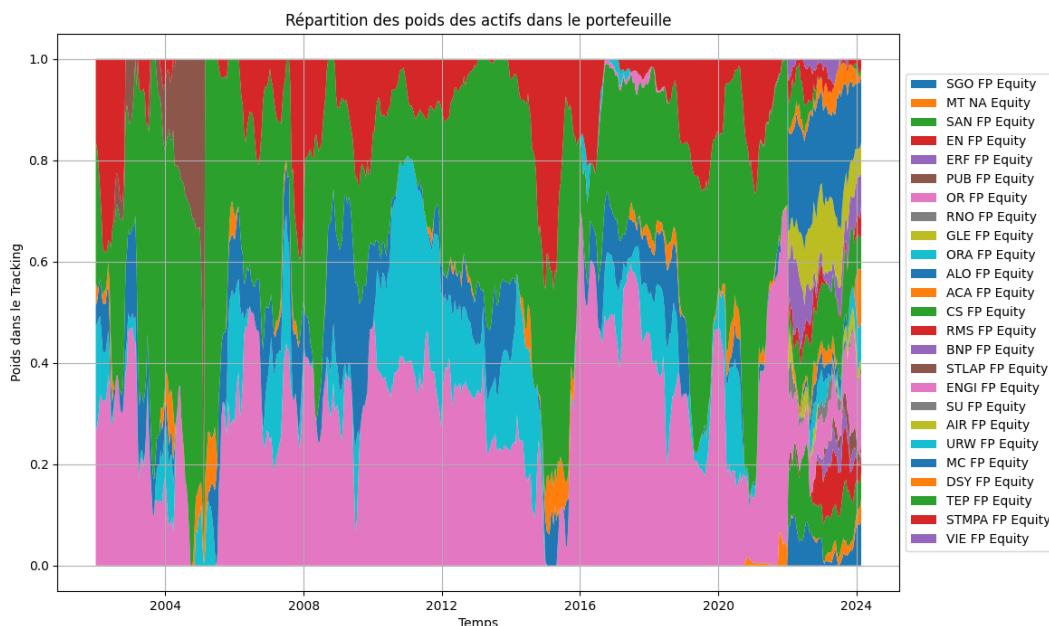


Figure 4: Poids de nos 25 assets dans notre portefeuille.

2 Long-short market neutral strategy

Présentation du problème Après avoir établi une stratégie de suivi simple, nous explorons l'extension naturelle de cette approche à travers la création de benchmarks "plus" et "moins". Attendues pour produire des rendements basés sur l'écart "plus"/"moins" avec une volatilité modérée, ces stratégies ne garantissent pas explicitement une corrélation nulle avec les rendements du marché. Toutefois, étant donné la forte corrélation des portefeuilles "plus" et "moins" avec le benchmark initial, leur différence présente une faible corrélation avec le benchmark, à condition que les erreurs de suivi ne soient pas corrélées avec le marché.

Génération des benchmarks plus et moins Les benchmarks "plus" et "moins" sont générés en ajustant les rendements du benchmark par un excédent ou un déficit annuel de $x\%$, distribué uniformément sur les rendements quotidiens.

Les régressions de cointégration peuvent être formalisées comme dans l'article. La figure 5 montre nos indices minis et max que nous avons pu construire, sur un cas particulier.

Ici, nous n'avons ajouté un rendement journalier de manière à ce que le rendement annuel soit augmenté de 5% par rapport au CAC40, et nous avons fait ça sur 16 ans.

Données utilisées Une procédure de rééquilibrage a été appliquée à notre stratégies long-short : les coefficients de cointégration ont été réestimés tous les 10 jours de bourse, et le nombre d'actions a été maintenu constant entre chaque rééquilibrage. Nous avons pris 10 actifs du CAC40, que nous détaillons dans notre notebook. Ci-dessous les actifs utilisés.

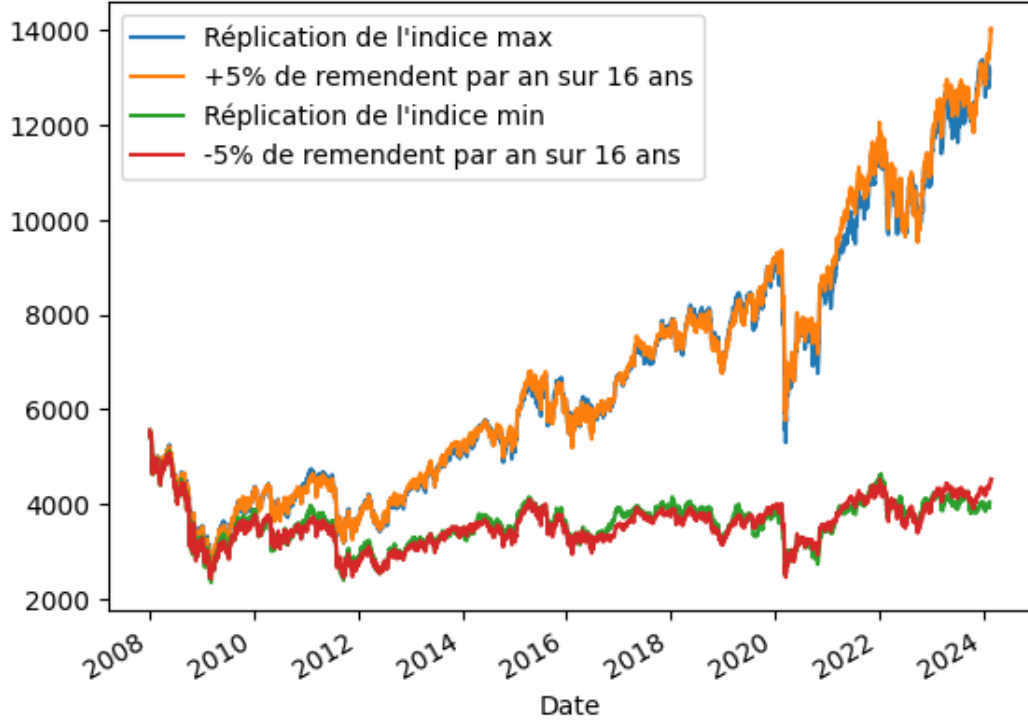


Figure 5: Construction des indices minus et plus.

Numéro	Code de l'actif
1	ACA FP Equity
2	OR FP Equity
3	DG FP Equity
4	STLAP FP Equity
5	TTE FP Equity
6	STMPA FP Equity
7	ML FP Equity
8	ALO FP Equity
9	CS FP Equity
10	SAF FP Equity

Coût de notre stratégie Notre stratégie tient compte d'un coût. En effet, changer l'allocation du portefeuille est payant. Nous avons été en mesure de calculer les coûts de réallocation de notre stratégie :

$$TC_T = 0.002 \sum_{k=1}^n \text{abs} \left((w_{\text{plus } k, T} - w_{\text{minus } k, T}) - (w_{\text{plus } k, T-10} - w_{\text{minus } k, T-10}) \right) P_{k, T}$$

Éléments supplémentaires importants Nous avons aussi plusieurs éléments intéressants que nous pouvons mentionner. Le papier d'origine donne une expression explicite pour les rendements de l'indice plus et minus. Nous les laissons au titre informatif au lecteur, et nous renvoyons sur le papier d'origine pour avoir le détail des quantités évoquées ci-dessous.

$$R_{+,t} = \rho_+ \frac{\sigma_+}{\sigma_{B+}} R_{B+,t} + \varepsilon_{+,t}$$

$$R_{-,t} = \rho_- \frac{\sigma_-}{\sigma_{B-}} R_{B-,t} + \varepsilon_{-,t}$$

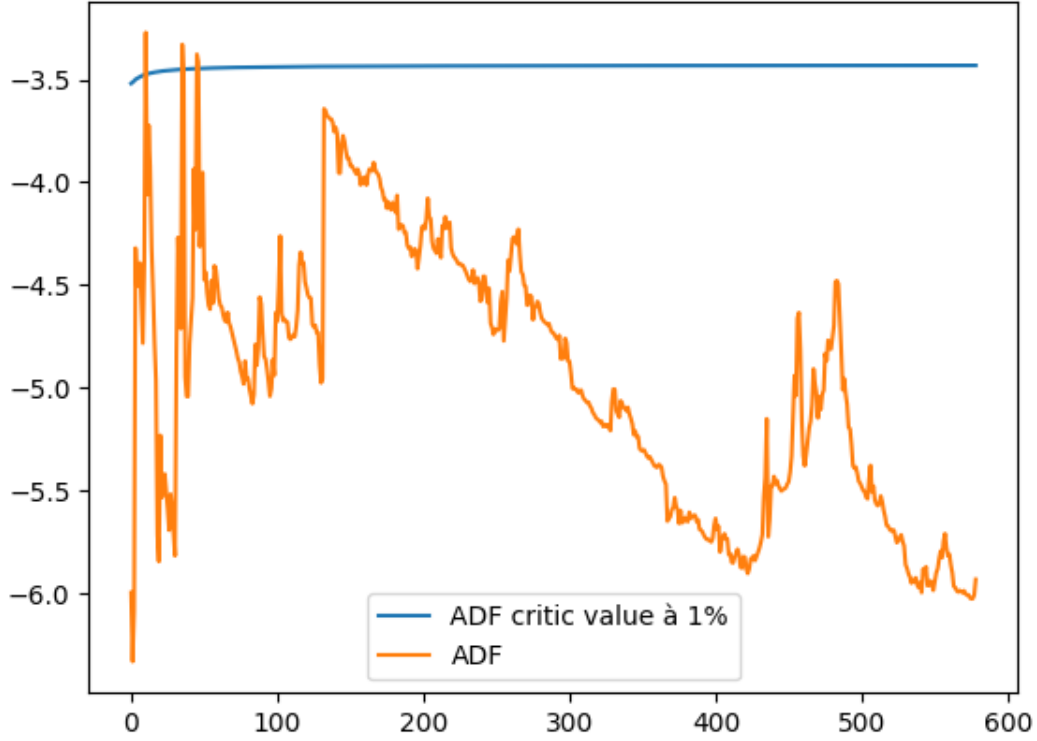


Figure 6: Test ADF pour vérifier que la cointégration s'applique bien.

$$R_{+/-,t} = R_{+,t} - R_{-,t} = R_{M,t} \left(\rho_+ \frac{\sigma_+}{\sigma_{B+}} - \rho_- \frac{\sigma_-}{\sigma_{B-}} \right) + \mu_+ \rho_+ \frac{\sigma_+}{\sigma_{B+}} + \mu_- \rho_- \frac{\sigma_-}{\sigma_{B-}} + \varepsilon_{+,t} - \varepsilon_{-,t}$$

2.1 Back-test results

Test ADF On commence par vérifier que nous arrivons bien à cointégrer. Pour cela, nous allons réaliser un test ADF. Nous laissons le résultat en figure 6. Le test de Dickey-Fuller augmenté (ADF) sert à vérifier si une série temporelle est stationnaire, condition nécessaire pour appliquer la cointégration. Supposons une série temporelle Y_t . Le test ADF examine l'hypothèse nulle que Y_t possède une racine unitaire, signifiant qu'elle n'est pas stationnaire. Le modèle testé est le suivant :

$$\Delta Y_t = \alpha + \beta t + \gamma Y_{t-1} + \sum_{i=1}^P \delta_i \Delta Y_{t-i} + \varepsilon_t$$

où ΔY_t est le changement de Y_t au temps t , α est une constante, βt représente une tendance temporelle, γ est le coefficient de Y_{t-1} (dont la signification est principalement testée pour la présence de racine unitaire), p est le nombre de retards sélectionnés pour le modèle, et ε_t est le terme d'erreur. Une valeur p faible pour le coefficient γ rejette l'hypothèse nulle, suggérant ainsi que la série est stationnaire et, par extension, susceptible d'être cointégrée avec d'autres séries, sous réserve de tests complémentaires.

Vérification que l'on réplique bien l'indice minus et l'indice plus Le deuxième point important était de visualiser que l'indice minus et l'indice plus étaient bien répliqués par notre stratégie, qui recalcule les poids tous les 10 jours avec le modèle de régression. La figure 7 montre bien que notre modèle fonctionne correctement.

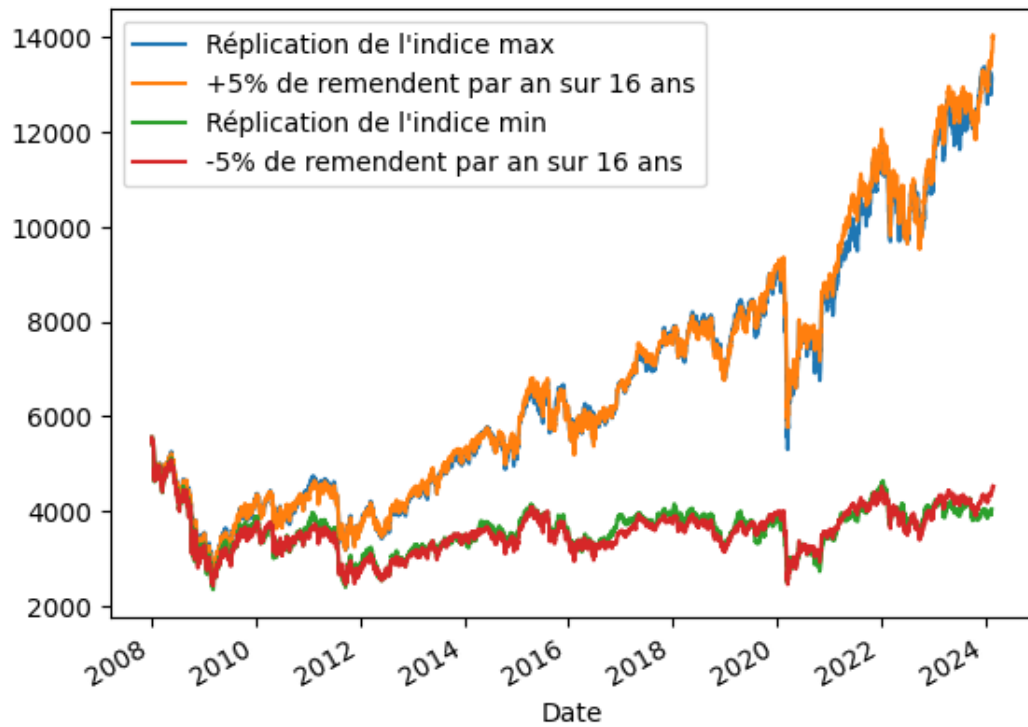


Figure 7: Nous vérifions que notre stratégie permet de bien répliquer les deux indices d'intérêt.

2.2 Returns on the long-short strategy

Analyse des rendements des différentes stratégies On propose de regarder les rendements mensuel de notre stratégie dans différents cas, selon que l'on ajoute 5, 10 ou 15% à la performance du CAC40 pour construire les indices min et au plus. La figure 8 présente ces résultats. On obtient ici des résultats proche de ceux obtenus par les auteurs du papier original, bien que nous travaillons ici avec le CAC40. Selon l'auteur, *Parmi les stratégies long-short analysées, celles qui produisent les résultats positifs les plus constants sont celles qui suivent de petits écarts, même si l'ampleur de ces retours est réduite.* Nous n'observons pas ceci de manière claire sur nos données, nous trouvons que les rendements sont très sensiblement proches pour les différentes stratégies en mensuel.

Histogramme des rendements Pour venir appuyer cela, nous regardons l'histogramme des rendements mensuels de nos différentes stratégies, en figure 9. Ainsi, pour les données du CAC40, nos stratégies semblent pour l'instant comparables. C'est pour cela que nous avons besoin de plus d'indicateurs.

2.3 Corrélations des rendements avec les rendements du CAC40

Les matrices de corrélations ci-dessous montrent les relations entre les rendements des stratégies ajustées de 5%, 10%, et 15% et les rendements du CAC40. Ces résultats sont en accord avec les attentes basées sur la stratégie mise en place, reflétant une faible corrélation avec le marché global :

Stratégie Corrélation avec CAC40

5%	0.1458
10%	0.1560
15%	0.1644

On remarque que les stratégies avec les plus faibles

écarts sont ici les meilleurs, en adéquation avec ce qu'avaient indiqués les auteurs.

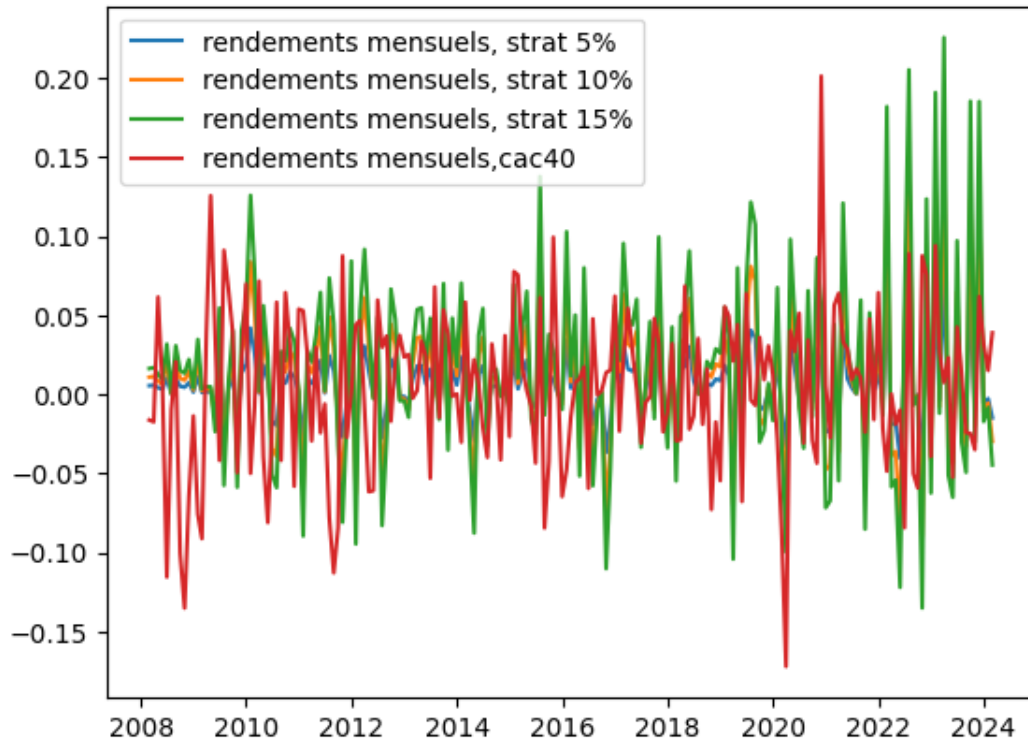


Figure 8: Rendements mensuel de la stratégie vs rendements du CAC40.

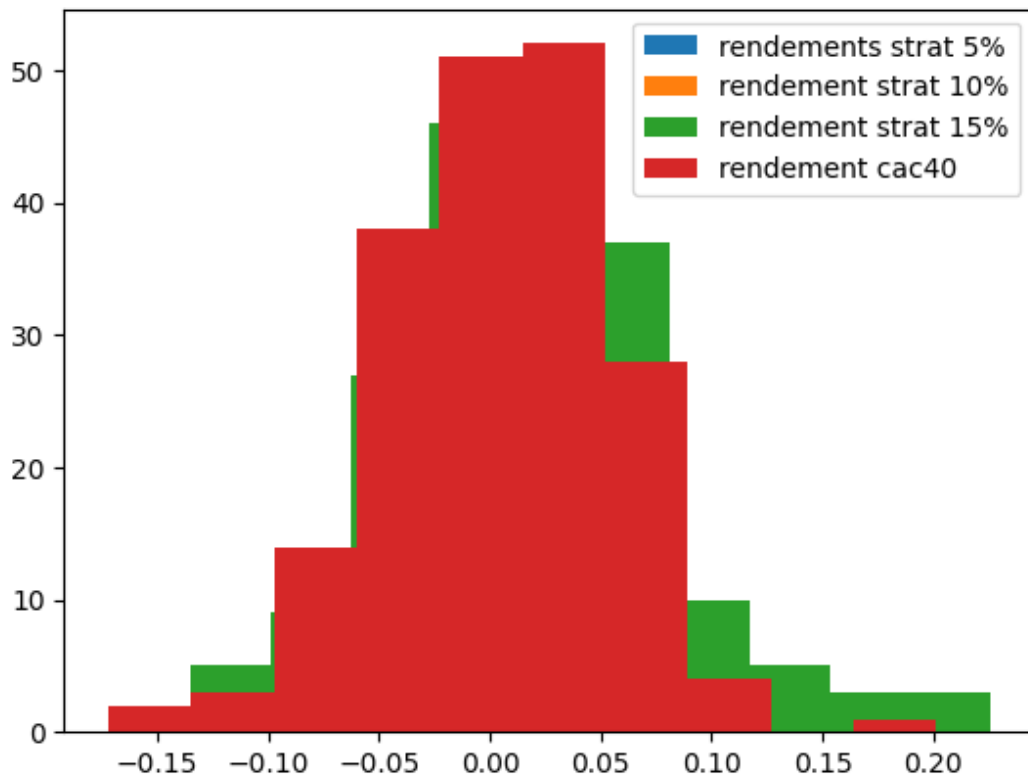


Figure 9: Rendements mensuel de la stratégie vs rendements du CAC40, histogramme.

2.4 Résultats plus globaux

Les résultats suivants résument les performances des stratégies basées sur des ajustements de 5%, 10%, et 15% pour la création des indices "plus" et "moins", évaluées selon la skewness (asymétrie), le ratio de Sharpe, et le kurtosis (tels que définis dans l'article):

Ajustement	Skewness	Ratio de Sharpe	Kurtosis
5%	1.064	1.306	2.509
10%	1.244	1.206	2.807
15%	1.370	1.110	3.005

La kurtosis (ou aplatissement) mesure l'intensité des queues de la distribution des rendements d'un actif par rapport à une distribution normale. Une kurtosis supérieure à 3 (le kurtosis d'une distribution normale) indique une distribution avec des queues plus épaisses, signalant un risque plus élevé d'événements extrêmes (rendements très élevés ou très bas).

Le ratio de Sharpe évalue la performance ajustée au risque d'un investissement en divisant l'excès de rendement (le rendement de l'investissement moins le taux de rendement sans risque) par l'écart-type des rendements de cet investissement. Un ratio de Sharpe plus élevé indique une meilleure performance ajustée au risque.

La skewness (ou asymétrie) mesure le degré d'asymétrie de la distribution des rendements par rapport à une distribution normale. Une skewness positive signifie que la distribution a une queue plus longue vers la droite, indiquant une probabilité plus élevée de rendements extrêmes positifs. Inversement, une skewness négative indique une queue plus longue vers la gauche, signifiant une probabilité plus élevée de rendements extrêmes négatifs.

Conclusion : En se basant sur ces métriques, la stratégie ajustant les indices de 5% semble la plus efficace. Bien que l'ampleur des retours soit potentiellement réduite, cette stratégie présente le ratio de Sharpe le plus élevé, indiquant une meilleure performance ajustée au risque. La skewness est la plus faible des trois, ce qui peut signaler un risque plus bas de rendements extrêmes négatifs, et le kurtosis est également le plus bas, suggérant moins de risque d'événements extrêmes par rapport aux autres stratégies. En résumé, la stratégie à 5% offre un équilibre attrayant entre la recherche de rendements et la gestion du risque, la rendant préférable dans une optique de prudence et d'efficacité ajustée au risque.

3 Pour aller plus loin

Dans les sections précédentes, nous avons introduit des stratégies de suivi d'indice et des stratégies long-short neutres au marché qui, loin d'être exclusives l'une de l'autre, peuvent être combinées pour améliorer certains aspects des stratégies neutres au marché ou pour transférer l'alpha généré par ces dernières vers l'indice du marché.

Une première amélioration des stratégies long-short neutres au marché consiste en une approche "fonds de fonds". Cette méthode, en investissant dans plusieurs stratégies similaires, vise à réduire la volatilité des rendements, similairement à la construction d'un indice présentant des propriétés statistiques plus attractives pour les investisseurs que les séries de rendements individuels.

Ces observations ouvrent des pistes supplémentaires qui pourraient enrichir et renforcer nos travaux, en explorant notamment la synergie entre les stratégies de suivi d'indice et long-short neutres au marché dans un cadre "fonds de fonds". Cette combinaison stratégique promet non seulement d'améliorer la robustesse de nos approches existantes mais aussi de proposer des alternatives novatrices pour la gestion de portefeuille.