THESE DE DOCTORAT

DE

L'Université de Paris XIII-Sorbonne Paris Cité

Ecole Doctorale Erasme - Sciences économiques et de gestion

CEPN (UMR-CNRS 7234)

M. Sébastien VALEYRE

Modélisation fine de la matrice de covariance/corrélation des actions

Thèse présentée et soutenue le 29 Mai 2019, à 10h, à la Maison du Poumon, 66 boulevard Saint-Michel, Paris. Composition du Jury:

M. Prigent, Jean-Luc	Professeur, Université de Cergy-Pointoise	Président du jury
M. Malevergne, Yannick	Professeur, Université Paris 1	Rapporteur
M. Maillet, Bertrand	Professeur, EM Lyon	Rapporteur
M. Courtault, Jean-Michel	Professeur, Université Paris 13	Examinateur
Mme. Rivals, Isabelle	Maitre de conférence, E.S.P.C.I.	Examinateur
M. Aboura, Sofiane	Professeur, Université Paris 13	Directeur de thèse

Résumé

Une nouvelle méthode a été mise en place pour débruiter la matrice de corrélation des rendements des actions en se basant sur une analyse par composante principale sous contrainte en exploitant les données financières. Des portefeuilles, nommés "Fundamental Maximum variance portfolios", sont construits pour capturer de manière optimale un style de risque défini par un critère financier ("Book", "Capitalization",etc.). Les vecteurs propres sous contraintes de la matrice de corrélation, qui sont des combinaisons linéaires de ces portefeuilles, sont alors étudiés. Grâce à cette méthode, plusieurs faits stylisés de la matrice ont été mis en évidence dont: i) l'augmentation des premières valeurs propres avec l'échelle de temps de 1 minute à plusieurs mois semble suivre la même loi pour toutes les valeurs propres significatives avec deux régimes; ii) une loi "universelle" semble gouverner la composition de tous les portefeuilles "Maximum variance". Ainsi selon cette loi, les poids optimaux seraient directement proportionnels au classement selon le critère financier étudié; iii) la volatilité de la volatilité des portefeuilles "Maximum Variance", qui ne sont pas orthogonaux, suffirait à expliquer une grande partie de la diffusion de la matrice de corrélation; iv) l'effet de levier (augmentation de la première valeur propre avec la baisse du marché) n'existe que pour le premier mode et ne se généralise pas aux autres facteurs de risque. L'effet de levier sur les beta, sensibilité des actions avec le "market mode", rend les poids du premier vecteur propre variables.

Mots clefs: corrélation, filtre, diagonalisation sous contrainte, modèle multifactoriel, portefeuilles optimaux, gestion d'actifs, diffusion

A mon épouse, Léa, A ma fille Gabrielle, A mes parents, Françoise et Dominique

Remerciements

Je souhaite ici rendre hommage et exprimer ma profonde gratitude à tous ceux qui, de près ou de loin, ont contribué à la réalisation de cette thèse. Mes remerciements s'adressent tout d'abord à mon Directeur de thèse, le Professeur Sofiane Aboura. Tout au long de ce travail, il a su m'apporter une disponibilité, une écoute, une confiance, des conseils précieux et avisés à la hauteur de ses compétences et de ses qualités humaines. J'adresse également tous mes remerciements aux Professeurs Bertrand Maillet et Yannick Malevergne, de l'honneur qu'ils m'ont fait en acceptant d'être rapporteurs de cette thèse ainsi qu'aux Professeur Jean-Michel Courtault, Jean-Luc Prigent et Isabelle Rivals de l'honneur qu'ils m'ont fait en acceptant d'être examinateur. Je remercie aussi mes collègues, l'équipe de recherche de John Locke, Denis Grebenkov, Stanislav Kuperstein, Maxime Beucher, Edouard Limouse et Frédéric Herbette. Sans eux, ce travail n'aurait probablement pas vu le jour. Leur aide et soutien ont été déterminants. J'ai aussi une pensée pour Thomas Dionysopoulos, avec lequel, suite à nos discussions en janvier 2015, j'ai commencé à réfléchir à l'opportunité de publier des articles sur la matrice de corrélation.

Enfin, sur un plan plus personnel, l'amour que je porte à mon épouse, Léa, et à ma fille, Gabrielle, m'a permis de surmonter les étapes difficiles. J'ai aussi une pensée pour les personnes qui ont été déterminantes dans ma reconversion professionnelle il y a 12 ans de l'énergie nucléaire à la finance de

marché dont cette thèse est l'un des jalons. Je me souviens notamment de ma discussion avec le Professeur Jean-Philippe Bouchaud, en mai 2006, au jardin du Luxembourg, qui a été un moment crucial. Il a été très bienveillant en me donnant des bons conseils: il ne fallait surtout pas reprendre mes études dans les produits dérivés, les calculs stochastiques et les équations différentielles, il fallait au contraire reprendre mes études axées sur la recherche, l'empirisme et les inefficiences de marché, sujet qui n'était pas encore très à la mode avant la crise financière. Je me souviens, par sa clairvoyance, qu'il avait déjà anticipé la crise des subprimes de 2007-2008 en me disant que selon lui les hypothèses liées à la valorisation des produits financiers dérivés complexes étaient devenues surréalistes. J'ai toujours été impressionné par sa simplicité, sa disponibilité et sa gentillesse vu sa réussite exceptionnelle. Je me souviens à cette époque du nombre de journées que j'ai passées à lire son livre et ses nombreux articles de recherche qui m'ont permis de former mes premières intuitions par rapport au fonctionnement des marchés financiers. J'ai aussi une pensée émue pour les Professeurs Jacques Prost et Pierre-Gilles de Gennes, tout aussi simples, bienveillants et exceptionnels, qui ont permis probablement en me recommandant, que l'université dauphine m'accepte en tant qu'étudiant, à 28 ans, avec mon parcours atypique au master de finance. J'ai aussi une pensée pour Joel Benarroch et François Bonnin qui m'ont appris le métier de gérant de fonds discrétionnaire et systématique. J'ai aussi une pensée pour les seuls professeurs de l'ESPCI qui ont vraiment cru en moi, Isabelle Rivals et Léon Personnaz, qui m'ont formé aux statistiques pendant un an dans leur laboratoire. Ils ont eu beaucoup d'influence sur mes travaux. J'ai aussi une pensée pour mes amis les plus proches Pierre,

Alexandre, Charles-Henri, Olivier, Emmanuel et Antoine et pour mes parents et mes frères et sœurs, qui ont cru en moi, et m'ont toujours soutenu dans cette reconversion professionnelle et je les remercie aussi.

Avant-Propos

Cette thèse a débuté en janvier 2016 dans le laboratoire CEPN (UMR 7234, CNRS) de l'université Paris-XIII. Le travail de recherche a été réalisé chez John Locke Investments, société de gestion indépendante et à taille humaine (15 salariés), pour laquelle j'ai continué à travailler à pleins temps en tant que chercheur et gérant des fonds systématiques John Locke Equity Market Neutral et John Locke Smart Equity. J'ai ainsi pu profiter de mon expérience concrète des marchés financiers pour adapter mes modèles à la réalité. Aussi j'ai dû me concentrer sur des modèles qui devaient avoir un intérêt certain pour la gestion d'actifs et les deux fonds que je gère. Modéliser la matrice de corrélation des actions est clef chez John Locke Investments. Ainsi les portefeuilles optimaux pour faire du trend following se basent uniquement sur l'exploitation de la matrice de corrélation qu'il faut maitriser, nettoyer, inverser, modéliser très proprement pour pouvoir amplifier les faibles autocorrélations en performances robustes. Ainsi les compétences de certains gérants peuvent très bien se limiter à la bonne modélisation de la matrice de corrélation. Les papiers de recherche devaient aussi être pratiques et constituer un support intellectuel pour convaincre les clients des fonds du fondement scientifique de mes modèles de gestion.

Trois papiers ont été présentés lors des conférences en 2016 (Liège, Belgique), en 2017 (Valence) et en 2018 (Paris) de l'AFFI et un quatrième sera présenté lors de la conférence en Juin 2019 (Laval, Quebec):

– le papier “Emergence of Correlation between Securities at Short Time

Scales" a été présenté au 35th International Conference of the French Finance Association à l'ESCP à Paris du 20 au 24 mai 2018. Le papier est présenté en premier au chapitre 1 car il explique l'origine physique des corrélations entre actions;

- le papier "Fundamental Market Neutral Maximum Variance Portfolios" a été soumis en janvier 2019 au 36th International Conference of the French Finance Association à Laval au Quebec. Le papier justifie la méthodologie utilisée dans la thèse pour débruiter la matrice de correlation. Le papier est à cheval entre plusieurs spécialités (model factoriels, matrice aléatoires, Asset Pricing) et doit être restructuré et découpé en plusieurs projets pour être publiable. Le papier est présenté au chapitre 2.
- le papier "The Reactive Beta Model" a été présenté au 34th International Conference of the French Finance Association à Valence le 31 mai et 1er et 2 juin 2017. Le papier est présenté au chapitre 4;
- le papier "Should Employers Pay Better their Employees? An Asset Pricing Approach" a été présenté au 33rd International Conference of the French Finance Association à HEC-Management School of the University of Liege du 23 au 25 mai 2016. Le papier est présenté dans le manuscrit au chapitre 6 comme une application.

0. Table des matières

I Introduction Générale **9**

1 Introduction **11**

2 Revue de la littérature **21**

2.1 Gestion de portefeuille 21

2.2 Econophysique 24

2.3 Modèles multifactoriels 25

2.4 Analyse par composante principale 27

2.5 Asset pricing 29

2.6 Processus stochastique 31

3 Contributions principales **35**

II Dissertation doctorale 41

1 Emergence of Correlation between Securities at Short Time Scales 43

2 Fundamental Market Neutral Maximum Variance Portfolios 63

3 Time Scale Effect on Correlation between Securities at Long Time Horizon 125

4 The Reactive Beta Model 133

5 The Model of Diffusion of Correlations between Securities 183

6 Should Employers Pay Better their Employees? An Asset Pricing Approach 209

III Conclusion Générale 259

I Introduction Générale

1. Introduction

La matrice de corrélation des rendements des actions est nécessaire à l'analyse du risque d'un portefeuille. Une modélisation fine est nécessaire pour construire les portefeuilles optimaux robustes (maximiser le gain potentiel tout en minimisant le risque). Les mesures empiriques de la matrice de corrélation sont bruitées du fait d'un nombre trop faible de rendements indépendants et homoscedastiques disponibles et d'un nombre trop grand d'actions. Ainsi il est courant de devoir mesurer les corrélations entre 500 actions ou plus avec beaucoup moins d'un an d'historique[1] afin de pouvoir supposer que les corrélations restent à peu près constantes sur cette période. L'échantillon se réduit encore lorsqu'on s'intéresse aux corrélations des rendements mensuels voire annuels, qui importent le plus pour les investisseurs. Les autocorrélations des rendements sont faibles mais suffisantes pour déformer la matrice selon l'horizon de temps et transformer des facteurs de risque

1. ce qui correspond à moins de 250 rendements journaliers qui ne sont que très approximativement gaussiens

négligeables à l'horizon de la journée en significatif à l'horizon du mois pour un gérant. Les mesures empiriques, qui se basent sur un échantillon trop petit, capturent des corrélations fallacieuses. Ces mesures fallacieuses peuvent résulter en portefeuilles qui semblent sans risque dans l'échantillon utilisé pour mesurer la matrice mais risqués dans un autre échantillon. La moindre optimisation de portefeuille, qui cherche à minimiser le risque pour une même rentabilité espérée, va forcément privilégier les portefeuilles qui semblent sans risque ou très peu risqués "in the sample" et il en résulte un manque de robustesse. Les vraies corrélations sont réputées être de plus très variables en fonction du temps. Ainsi quand le marché est stressé, les investisseurs se mettent à paniquer et les corrélations ont tendance à augmenter, si bien que toutes les actions sont entrainées par les mouvements des indices. Lorsqu'un facteur de risque devient majeur, lorsqu'un évènement inattendu survient, alors toutes les actions qui capturent ce facteur de risque vont se corréler brusquement. Lorsqu'une action sous-performe ou surperforme, ses corrélations avec les autres actions vont changer. Ainsi on peut parler de corrélations non linéaires car les corrélations dépendent des trajectoires de chaque action mais une grande partie des variations semblent complètement stochastiques. Une grande difficulté est donc de mesurer les corrélations de "population" sans erreur et sans retard. Une autre difficulté est aussi de prévoir comment la matrice risque de varier. Malgré les enjeux, de nombreux faits stylisés de la matrice de corrélation, qui sont noyés dans le bruit, restent pourtant encore à découvrir. Cette thèse, qui cherche à mettre en évidence plusieurs faits stylisés, remplit un vide dans la littérature académique et fait le lien entre plusieurs disciplines entre mathématiques (processus stochastique), finance (modèles multifactoriels), gestion d'actifs (portefeuilles optimaux), econophysique (matrice aléatoire) et économie (Asset Pricing).

Je me suis d'abord intéressé à l'origine des corrélations des rendements des actions: j'ai ainsi modélisé l'émergence des corrélations des rendements des actions européennes et américaines de 2000 à 2017 sur des échelles courtes (de 1 minute à 1 jour) grâce à un modèle de retard inspiré de la microstructure. A des échelles très courtes, de l'ordre de la seconde, les corrélations sont nulles puis elles apparaissent et augmentent avec l'échelle de temps. L'émergence des corrélations est en fait la conséquence de l'impact des transactions, qui se matérialisent entre les actions similaires via des algorithmes de trading. J'ai mis en place, dans le chapitre 1, un modéle de retard qui reproduit très bien l'effet d'échelle mesuré qui permet d'extrapoler la vision de la matrice des rendements 1 minute à la journée.

Pour identifier la structure et la dynamique des valeurs propres et vecteurs propres, j'ai mis en place, dans le chapitre 2, une méthodologie basée sur l'analyse par composante principale contrainte qui permet de débruiter la matrice en tirant bénéfice des informations financières, comme, par exemple, le ratio entre la valeur comptable et la valeur de marché de l'action ("Book"), la valeur capitalistique de l'action ("Capitalization") ou de nombreux autres ratios financiers. L'analyse par composante principale appliquée aux rendements de l'ensemble des actions revient à diagonaliser la matrice de corrélation des rendements. La matrice de corrélation est préférable à la matrice de covariance pour éviter un biais vers les actions les plus volatiles. La diagonalisation permet d'identifier les portefeuilles d'actions decorrélés les uns des autres qui génèrent le plus de volatilité pour un même investissement (mesurée exactement par la volatilité du portefeuille obtenue sans tenir compte des corrélations entre les actions). Les rendements de ces portefeuilles particuliers permettent de modéliser simplement les mouvements principaux du marché: ces portefeuilles particuliers sont réputés proches des combinaisons très bruitées de stratégies de base (les indices pondérés par les capitalisations, les indices sectoriels, les indices investis sur les petites capitalisations, les indices investis sur les entreprises de croissance, les indices "Min Variance" investis sur les actions peu volatiles, les indices investis sur les entreprises "Value", etc.). Les variances de ces portefeuilles lorsqu'ils sont normalisés sont proportionnelles aux valeurs propres. Les corrélations entre actions sont quasiment toutes positives et rend l'identification du premier vecteur propre plutôt aisée: le premier vecteur propre est très significatif. Il reste proche du portefeuille investi sur chacune des actions avec une valeur propre de l'ordre de 100[2]. Les autres vecteurs propres ont des valeurs propres beaucoup plus petites (inférieure à 20) et représentent des portefeuilles "long/short" et "market neutre" d'abord plutôt sectoriels puis plutôt de style. Cependant l'instabilité de la matrice de corrélation associée au bruit de mesure rend difficile l'interprétation des vecteurs propres "long/short" mesurées. Ainsi d'un côté, on devrait réduire la profondeur sur laquelle on mesure la corrélation pour espérer une certaine stabilité des corrélations sur la période de mesure, mais, de l'autre coté, on devrait augmenter la période et la fréquence pour réduire le bruit de mesure.

Pour filtrer le bruit de mesure, inhérent à l'analyse par composante principale, j'ai contraint l'analyse au sous espace des facteurs de risque princi-

2. proche de la corrélation moyenne de l'ordre de 0.4 au carré multipliée par le nombre d'action de l'ordre de 500 dans mon cas

paux, déjà identifiés dans la littérature, dont j'ai optimisé la construction. J'ai inclus les facteurs de styles principaux (“Momentum”, “Capitalization”, “Quality”, etc.) et les facteurs de risque sectoriel. Les facteurs principaux optimisés ont été nommés “Fundamental Maximum variance market neutral portfolios”, car la variance de leurs rendements a été optimisée par construction. Ces facteurs peuvent aussi être directement utiles dans l'industrie de la gestion d'actifs, car ils optimisent théoriquement le gain ajusté du risque des primes de risque alternatives, qui sont devenues des véhicules d'investissement très populaires. Aussi grâce à l'optimisation, j'ai pu relier les valeurs propres sous contraintes débruitées aux valeurs propres bruitées de la matrice. Cela m'a permis de débruiter la matrice de corrélation et de caractériser finement une loi universelle, selon laquelle, les poids optimaux des facteurs de risque seraient uniformément distribués pour tous les critères financiers ce qui est particulièrement intriguant (on aurait attendu une distribution gaussienne et non uniforme plus logique pour obtenir des vecteurs propres aléatoires). Cette loi universelle a des conséquences importantes dans l'Asset Pricing: la norme dans cette discipline est de construire des portefeuilles “long/short” investis à l'achat sur le premier quintile selon le critère financier étudié et à la vente sur le dernier quintile. Si le portefeuille capture une performance significativement différente de zéro alors une anomalie de marché est identifiée. Une construction plus optimale du portefeuille avec une règle d'investissement linéaire au lieu de la marche en escalier peut aider à obtenir des performances plus significatives pour les petites anomalies.

Le filtre du bruit de mesure m'a aussi permis de caractériser finement la dynamique de la matrice de corrélation et d'identifier notamment les violents changements des valeurs propres (la valeur propre du premier mode peut passer de 200 à 30 soit une variation de corrélation moyenne de 0.5 à 0.05 en quelques mois seulement). Ce violent changement peut être modélisé par l'effet de levier (corrélation négative entre les rendements et les volatilités) pour la premiére valeur propre. J'ai aussi vérifié que les premiers vecteurs propres s'investissaient sur les facteurs fondamentaux de risque les plus risqués qui sont différents selon les périodes. Selon les crises, il peut s'agir du secteur IT, du secteur de la finance, du secteur de l'énergie, des REITs ou des entreprises exposées à la dette. Les entreprises qui sont peu sensibles aux variations de l'indice, et celles qui constituent les composants du facteur "Momentum", restent très représentées dans le deuxième et troisième vecteurs propres. Les facteurs "Capitalization" et "Book" de Fama et French sont très peu représentés dans les premiers vecteurs propres de la matrice de corrélation.

La première application de la méthodologie que j'ai introduite et qui permet de débruiter la matrice de corrélation a consisté à étendre l'étude de l'effet d'échelle sur les valeurs propres de 1 minute à 1 journée sur des échelles de temps plus longues entre 1 jour et plusieurs mois. Les corrélations continuent d'augmenter avec l'échelle de temps. Cela explique par exemple que la norme dans l'Asset Pricing est de se baser sur les rendements mensuels pour estimer les corrélations. En effet, même si l'utilisation des rendements journaliers donnerait des résultats plus robustes, les chercheurs dans l'Asset Pricing préfèrent travailler avec les rendements mensuels, car les corrélations sont réputées plus fortes lorsqu'elles sont mesurées à partir de rendements mensuels qu'à partir des rendements journaliers à cause d'un effet d'echelle que les chercheurs redoutent. La réduction de la matrice de corrélation dans le sous espace généré par les portefeuilles fondamentaux "Maximum variance" permet de confirmer cette crainte avec des mesures significatives. Les corrélations ont tendance à continuer à augmenter sur des horizons de temps plus long. Ce phénomène est expliqué grâce à un modèle d'autocorrélation, qui permet de reproduire l'effet de manque de liquidité du marché. L'illiquidité crée de l'inertie et fait qu'un mouvement de marché dure et peut être prolongé par le comportement moutonnier des investisseurs. Les autocorrélations, introduites dans le chapitre 3, apparaissent plus robustes que les anomalies non conditionnelles pas toujours significatives telles qu'identifiées dans l'Asset Pricing. Ces anomalies se matérialisent par des primes de risque alternatives pour justifier les incohérences avec le modèle d'évaluation des actifs financiers (MEDAF ou CAPM en anglais), selon lequel, les primes de risque ne doivent dépendre que du beta, sensibilité de l'action avec les variations de l'indice.

La deuxième application de de la méthodologie que j'ai introduite et qui permet de débruiter la matrice de corrélation a consisté à caractériser la dynamique de la matrice de corrélation qui est importante à modéliser pour estimer les risques. En effet la matrice de corrélation de population peut changer et cela peut représenter un risque. Le problème est que la matrice est déjà tellement bruitée qu'espérer mesurer ces changements est illusoire, si bien que les modèles stochastiques théoriques ne peuvent pas facilement être validés empiriquement. Dans le chapitre 5, j'ai réussi à faire plusieurs mesures grâce à la méthodologie qui permet d'utiliser les informations financières pour réduire la taille de la matrice de corrélation et grâce à l'utilisation des rendements 5 minutes. J'ai ainsi mis en évidence certains faits stylisés de la diffusion de la matrice de corrélations très mal connus et très mal reproduits par les modèles standard issus de Wishart, comme la distribution des valeurs propres des incréments de la matrice de corrélations des actions (il s'agit ici précisément de la matrice de corrélation des actions sous sa forme réduite dans le sous espace des 24 facteurs fondamentaux Maximum Variance pour éliminer le bruit de mesure). L'étude de cette distribution permet de caractériser la diffusion des vecteurs propres de la matrice de corrélation des actions. Cette distribution ne suit pas une loi demi-cercle de Wigner mais une distribution avec des queues, qui peuvent être interprétées par la présence de valeurs propres extrêmes. Ces valeurs propres extrêmes expliquent que des corrélations entre actions peuvent changer beaucoup plus brutalement que les modèles classiques ne peuvent le prévoir. J'ai ainsi modélisé l'instabilité de la matrice de corrélation avec un processus empirique plus réaliste. La diffusion dans la composition des premiers vecteurs propres explique en grande partie la distribution des valeurs propres des incréments de la matrice de corrélation. Cette diffusion s'explique quasiment entièrement par la volatilité de la volatilité des portefeuilles fondamentaux "Maximum variance". Les portefeuilles n'étant pas orthogonaux, la volatilité de la volatilité permet de répliquer la diffusion des vecteurs propres tout en supposant les corrélations entre portefeuilles fondamentaux fixes.

Une composante particulière de la diffusion de la matrice a aussi fait l'objet d'une grande attention: la dynamique des poids du premier vecteur propre qui sont liés aux beta, qui est la sensibilité des rendements d'une action avec les indices boursiers, a été analysée en profondeur. Les beta constituent par ailleurs une mesure du risque qui est capitale car ils forment une indication d'un risque systématique qui ne peut pas se diversifier ou s'éliminer pour un investisseur classique. Cela justifie intuitivement que les actions à fort beta doivent rémunérer plus les actionnaires et que les primes de risque doivent être proportionnelles au beta qui est à la base du MEDAF. Aussi les fonds alternatifs, qu'on appelle aussi "hedge funds", ont la capacité à prendre des positions vendeuses avec des ventes à découvert pour neutraliser l'exposition de leur investissement aux variations des indices boursiers. Cela permet de mieux contrôler le risque et de proposer des investissements diversifiant aux épargnants. Pour construire des portefeuilles immunisés contre les variations de la bourse, qu'on appelle "beta neutre", il est extrêmement important de se baser sur des mesures fiables et sans biais des betas d'autant plus que certaines stratégies très populaires ont tendances à amplifier les biais de mesure du beta. Cela m'a motivé à modéliser finement l'effet de levier et l'élasticité des beta, qui décrivent aussi la composante du premier vecteur propre de la matrice. Par exemple, lorsqu'une action sous performe, son beta va augmenter. Lorsque la volatilité de l'action augmente plus que les autres, son beta va augmenter aussi. J'ai mis au point, dans le chapitre 4, une méthode réactive de la mesure des beta nécessaire pour construire des facteurs fondamentaux beta et secteur neutre moins biaisés et potentiellement mieux valider le MEDAF. Des tests montrent l'intérêt d'un tel modèle par rapport à des méthodes standards (OLS, régression par quantile, DCC GARCH).

Enfin une application concrète de mes travaux, dont la portée peut ne pas se limiter à la gestion d'actifs, met en avant l'intérêt de la méthode que j'ai introduite en se révélant assez fine pour distinguer le facteur "Rémuneration" du bruit. Dans le chapitre 6, je montre que facteur “Rémuneration” s’avère être un facteur de risque commun significatif. Les entreprises qui rémunèrent mieux leurs employés ont un risque en commun. Ces entreprises ont aussi tendance à avoir des meilleures performances. J’ai ainsi découvert une nouvelle anomalie par rapport au MEDAF et aux facteurs de Fama et French qui pourrait avoir une portée managériale voire politique.

Cette thèse peut donc avoir de multiples applications: une meilleure analyse du risque, une optimisation plus robuste d’un portefeuille, une meilleure modélisation des autocorrélations qui sont exploitées par les programmes de trading d’arbitrage de style, une meilleure mesure des anomalies dans l’Asset Pricing, une modélisation plus réaliste de la dynamique de la matrice de corrélations pour évaluer des produits dérivés. Elle peut aussi avoir des implications très concrètes en économie et en management car, par exemple, elle permet de montrer que les entreprises qui rémunèrent bien leurs employés partagent un risque significatif en commun et ont aussi tendance à mieux performer.

2. Revue de la littérature

Ce travail de recherche s'est articulé autour de six champs disciplinaires relativement cloisonnés entre plusieurs disciplines finance, économie, éconophysique et mathématiques appliquées.

2.1 Gestion de portefeuille

La bonne estimation des corrélations des rendements des actions est nécessaire pour l'analyse de risque d'un portefeuille et pour son optimisation. La bonne compréhension des variations temporelles des corrélations en cours ou potentielles est aussi critique pour la gestion d'un portefeuille et notamment d'un fonds "market neutre" qui utilise un fort effet de levier financier et dont l'arbitrage de style est un des moteurs de performance. L'arbitrage de style est une stratégie de trading qui consiste à investir sur les styles de gestion porteurs. Par exemple si le style de gestion qui consiste à acheter des petites capitalisations et à vendre des grosses capitalisations est profitable, la stratégie va acheter les petites capitalisations et vendre les grandes. Dans le cas inverse, la stratégie va acheter les grosses et vendre les petites. Plus de 240 styles de gestion ou facteurs de risques profitables ont été publiés dans la littérature scientifique décrites dans la section 2.5. L'intérêt du market timing ne fait pas consensus (Lee (2017); Bender et al. (2018); Bass et al. (2017)) et certains préfèrent bénéficier simplement de la diversification. DeMiguel et al. (2017) montrent qu'en pratique il suffit, pour construire un portefeuille, de sélectionner 15 critères financiers significatifs sur plus de 100. Les stratégies de market timing peuvent être complexes. Elles s'appuyent sur des modèles de prévision. Hodges et al. (2017) cherchent des prédicteurs des facteurs dans diffèrent régimes économiques et différentes conditions de marché. Ils trouvent que l'utilisation d'une combinaison d'indicateurs sur le cycle économique, la valorisation, la tendance et la dispersion serait plus efficace que l'utilisation d'indicateur individuel. Ainsi Dichtl et al. (2018) fabriquent un portefeuille "long/short" grâce à la méthode d'optimisation des paramètres introduite par Brandt et al. (2009) en utilisant plusieurs indicateurs de valorisation et de tendance et ils montrent que le market timing permet de surperformer le portefeuille investi équitablement sur les différents styles de gestion dont les primes de risque sont positives. La fragilité de ces résultats vient du risque de surapprentissage. De plus ces stratégies en général peuvent souffrir de chocs de corrélations entre les différents styles de gestion qui peuvent survenir et générer des pics de volatilités. Ainsi les stratégies quantitatives d'habitude non corrélées peuvent se corréler fortement de manière brutale. Cela s'est passé du 8 au 9 août 2007, quand la plupart des fonds d'arbitrage de style ont subi des pertes très significatives brutalement en même temps (Stein (2009)). Lors de cet évènement, nommé plus tard "quant crash", la plupart des fonds touchés employaient des stratégies "market neutre" quantitatives sans exposition au marché ce qui remet en question leur statut "market neutre" (Khandani and Lo (2011)). Il semble en fait que trop de gérants étaient investis sur les même "crowded" stratégies avec trop de levier et qu'ils ont tous voulu réduire leurs positions en même temps au même signal. De tels risques de "crowding" affectent une grande variété de stratégies, comme le style "Momentum" (acheter les actions qui ont surperformé et vendre les actions qui ont sous performé) car ils ne dépendent pas d'estimation indépendante des valeurs fondamentales des entreprises (Hong and Sraer (2016); Stein (2009)).

Des centaines de milliards de dollars sont aussi gérées directement en utilisant l'optimisation Mean-Variance introduite par Markowitz (1952) en préférant se baser sur des hypothèses simples concernant les espérances des rendements. La valeur ajoutée des gérants viendrait seulement d'une modélisation plus adaptée de la matrice de corrélation et d'une bonne capacité à exécuter les ordres en minimisant l'impact de marché. Ainsi le portefeuille "Min Variance" suppose que les espérances des rendements sont toutes identiques et que la matrice de corrélation peut se modéliser simplement, par exemple, avec un modèle à un facteur (Clarke et al. (2013)). Le portefeuille "Max Diversification" introduit par Choueifaty and Coignard (2008) suppose que les espérances sont proportionnelles au risque. Ces deux derniers portefeuilles nécessitent d'inverser la matrice de corrélation ce qui peut poser problème si la matrice n'est pas proprement modélisée. Le portefeuille "Equal-

Risk Contribution" introduit par Maillard et al. (2010) est moins sensible aux bruits de mesure et est donc plus robuste mais n'est plus théoriquement optimal. Benichou et al. (2017) introduisent le portefeuille "Agnostic Risk Parity". Les espérances ne sont plus forcément positives mais dépendent des rendements passés. Le portefeuille dépend alors de l'inverse de la racine carré de la matrice de corrélation multipliée par des signaux qui représentent des indicateurs techniques des tendances. Ce portefeuille "trend following" alloue le même risque sur chaque vecteur propre de la matrice de corrélation.

2.2 Econophysique

Aujourd'hui les performances des différents styles de gestion et les performances sectorielles sont très suivies par tous les acteurs du marché qui ne se contentent plus d'avoir une vue binaire (le marché va-t-il monter ou baisser?) et s'auto-alimentent par un phénomène d'effet moutonnier très bien décrit dans la littérature (Guedj and Bouchaud (2005); Michard and Bouchaud (2005); Cont and Bouchaud (2000); Wyart and Bouchaud (2007); Lux and Marchesi (1984)). Ainsi quand tel ou tel style de gestion chute, les acteurs vont le vendre en même temps et accentuer sa chute. La moindre nouvelle macroéconomique va impacter les indices mais aussi les autres facteurs de risque. Quand la Réserve fédérale des États-Unis se dit prête à augmenter les taux d'intérêt, le facteur levier (vente d'actions endettées, achat d'actions peu endettées) sera joué puis d'autres facteurs seront entrainés. Benzaquen et al. (2017) mettent en évidence le lien entre le trading et la matrice de corrélation en partant de la microstructure et du cross impact des transactions sur les prix. Les corrélations ne décriraient que l'interaction entre actions par le jeu des traders. Les corrélations sont aussi réputées pour augmenter avec l'échelle de temps: les rendements mensuels sont plus corrélés que les rendements journaliers qui sont plus corrélés que les rendements 1 minutes (Epps (1979)). Bouchaud and Potters (2018) avaient déjà proposé dans la partie "some open problem" une piste (les rendements de l'action i n'impactent pas instantanément les rendements de l'action j mais avec un certain retard) pour expliquer la dépendance des corrélations à la fréquence mais ne l'avait pas développé.

2.3 Modèles multifactoriels

Depuis l'article majeur de Markowitz (1952), l'optimisation « Mean Variance » est devenue une méthode rigoureuse pour construire un portefeuille d'investissement. Deux ingrédients fondamentaux sont nécessaires: les espérances des rendements de chaque action et la matrice de covariance des rendements. L'estimation de la matrice de covariance a toujours été un sujet important. La méthode de base se contente d'agréger les rendements historiques et de calculer leurs covariances historiques. Malheureusement cela crée des problèmes bien documentés (Jobson and Korkie (1980)). Pour l'expliquer simplement, quand le nombre d'actions est grand devant le nombre d'observations disponibles, ce qui est généralement le cas, la matrice de corrélation historique comporte beaucoup d'erreurs. Cela implique que les coefficients les plus extrêmes prennent des valeurs extrêmes non pas à cause de la réalité mais à cause d'erreurs extrêmes. Invariablement les optimisations de portefeuille vont miser leurs plus gros paris sur ces erreurs extrêmes ce qui rendra l'optimisation extrêmement non fiable. Michaud (1989) appelle ce phénomène "error-optimization". De manière alternative on peut considérer une estimation avec beaucoup de contraintes, comme le « single-factor model » de Sharpe (1963). Ces estimateurs de la matrice de corrélation contiennent d'un côté peu d'erreurs mais de l'autre beaucoup d'erreurs de spécification et de biais. Une alternative est le "Shrinkage" qui consiste à un mélange entre l'estimation sans contrainte et l'estimation avec la contrainte (Ledoit and Wolf (2003, 2012)). L'APT ("Arbitrage Pricing Theory") de Ross (1976) a généré un intérêt croissant dans les modèles multifactoriels. Ainsi le standard de l'industrie de la gestion d'actifs est d'utiliser des modèles multifactoriels. Quelques entreprises, comme APT, Barra et Axioma (Barra (1998)) qui sont devenues incontournables dans l'industrie de la gestion d'actifs, proposent à leurs clients des matrices de covariances qui s'adaptent mieux aux optimisations de portefeuille. Ces sociétés ont été accusées d'être à l'origine du "quant crash" de 2007, déjà mentionné dans la section 2.1, car elles favorisaient le "crowding" en fournissant les même facteurs de risque à tous les gérants. Ces méthodes se basent sur des modèles multifactoriels fondamentaux combinant une cinquantaine de facteurs sectoriels et d'autres risques. Ces facteurs utilisent le rendement des portefeuilles associés à certains critères financiers observables tel que le "Dividend Yield", le "Book to Market" ratio ou les secteurs d'appartenance. Une autre approche est d'utiliser les facteurs statistiques issus de l'analyse par composante principale, qui est décrite dans la section 2.4, avec un nombre total de facteurs de l'ordre de 5. Connor (1995) montre que les modèles multifatoriels "fondamentaux" permettent d'expliquer 42% ($R^2 = 42\%$ étant le pouvoir explicatif du modèle) des rendements alors qu'une simple analyse par composant principale sur 5 facteurs explique deja

39%. Connor (1995) trie les facteurs selon leur pouvoir explicatif. Les secteurs permettent d'augmenter de 18%, puis le facteur "Low Volatility" (proche du facteur "Low Beta") augmente le R^2 de 0.9% puis les facteurs "Momentum", "Capitalization", "Liquidity", "Growth", "Earning", augmentent de moins de 0.8%. Puis il reste par ordre d'importance décroissant des facteurs plutôt mineurs: le "Book to market", le "Earning Variability", le "Leverage", l'investissement à l'étranger, le coût du travail et enfin le "Dividend Yield". Toutefois la sélection des facteurs nécessaires et le choix du nombre a fait l'objet de nombreuses controverses (Roll and Ross (1980, 1984); Dhrymes et al. (1984); Luedecke (1984); Trzcinka (1986); Conway and Reinganum (1988); Brown (1989)). Connor and Korajczyk (1993) proposent une méthodologie simple pour estimer le nombre de facteurs significatifs: si le rajout d'un facteur ne réduit pas significativement le carré du résidu alors le facteur n'est pas considéré comme significatif. La plupart des études académiques se base sur une analyse historique depuis 1967 en exploitant la base de données du centre de recherche des prix des actions (Center of Research in Security Prices). Cette base de données regroupe principalement les actions cotées à la bourse du New-York Stock Exchange depuis 1926.

2.4 Analyse par composante principale

L'analyse par composante principale (ACP) prend sa source dans un article de Karl Pearson publié en 1901. Encore connue sous le nom de transformée de Karhunen-Loeve ou de transformée de Hotelling, l'ACP a été de nouveau développée et formalisée dans les années 1930 par Harold Hotelling. La puissance mathématique de l'économiste et statisticien américain le conduira aussi à développer l'analyse canonique, généralisation des analyses factorielles dont fait partie l'ACP. Les champs d'application sont aujourd'hui multiples, allant de la biologie à la recherche économique et sociale, et plus récemment le traitement d'images.

La théorie de la matrice aléatoire, dont la distribution des valeurs propres obtenues par l'ACP suit la loi de Marčenko-Pastur pour les grandes matrices, modélise les bruits de mesure des corrélations et montre que les petites valeurs propres en dessous d'une valeur propre critique sont sous estimées et ne sont pas significatives (Laloux et al. (1999); Plerou et al. (1999, 2002); Potters et al. (2005); Wang et al. (2011)). Bun et al. (2016) appliquent une méthode théorique introduite par Ledoit and Péché (2011), qu'ils appellent "Rotationnaly invariant estimator", pour debiaiser de manière continue les valeurs propres empiriques et ils montrent que la méthode semble plus robuste que celles du "Clipping" ou du "Shrinkage" qui sont bien documentées par Ledoit and Wolf (2004, 2003)). Allez and Bouchaud (2012) modélisent l'impact du bruit sur le premier vecteur propre et montre que ce dernier tourne légèrement autour d'un vecteur fixe. L'angle de rotation dépend du ratio entre la première valeur propre et les autres. En appliquant ce modèle aux autres vecteurs propres, on comprend qu'ils tournent aussi autour d'axes fixes mais avec un angle de rotation bien plus important. Ils sont ainsi très bruités ce qui explique la difficulté à les interpréter.

L'ACP avec une contrainte linéaire est une alternative aux filtres issus de la théorie de la matrice aléatoire pour éliminer le bruit de mesure et est entièrement résolu depuis longtemps (Golub (1973)). Dans ce cas les vecteurs propres sous contrainte appartiennent tous au sous espace solution de la contrainte : les vecteurs propres sous contrainte sont simplement les vecteurs propres d'une matrice qui a été réduite et débruitée. Toute la difficulté est de définir les facteurs formant le sous espace contraint pour que les contraintes n'impactent principalement que le bruit des valeurs propres. Pour cela, il est possible de s'inspirer de la littérature de l'Asset Pricing décrite dans la section 2.5) et des modèles multifactoriels décrite dans la section 2.3).

2.5 Asset pricing

Fama (1965) a abouti à la théorie des marchés efficients, selon laquelle, les prix suivent des marches aléatoires. Puis Sharpe (1964) dérive le MEDAF à partir d'hypothèses plus ou moins réalistes, comme l'absence de coût de transaction et la rationalité des investisseurs. Selon le MEDAF, l'espérance des rendements doit être théoriquement proportionnel au beta, seul risque qui n'est pas diversifiable et qui doit être rémunéré. Depuis 1970, différentes anomalies ont été observées par rapport à cette théorie. Les facteurs classiques de Fama and French (1992, 1993) sont investis à l'achat sur le top 20 %, selon le critère financier étudié, et investis à la vente sur le bottom 20%. Ces facteurs peuvent capturer une anomalie par rapport à la théorie des marchés efficients s'ils génèrent des gains significativement différents de zéros. La construction top 20 % bottom 20% est clairement sous optimale, selon Asness et al. (2013), mais reste paradoxalement la référence dans le domaine de l'Asset Pricing. La régression de Fama and MacBeth (1973) est la méthode la plus utilisée pour mettre en évidence des anomalies par rapport au MEDAF. Plusieurs modèles ont été développés pour fournir une interprétation économique aux nombreuses anomalies et pour améliorer le MEDAF.

Fama and French (1993) ont proposé un modèle à trois facteurs pour modéliser les espérances des rendements. Harvey and Liu (2018) ont listé 316 facteurs potentiels censés capturer une anomalie à partir de 313 articles depuis 1967. Selon eux, la plupart des facteurs peuvent être le fruit du data mining et ne seraient pas robustes. La plupart de ces facteurs se recoupent c'est pourquoi une vingtaine peut suffire mais le niveau de significativité pour caractériser les anomalies ne fait pas consensus. Les travaux académiques ont d'abord retenu les critères financiers tels que la "Capitalization", le "Price Earning Ratio", le "Cash Flow", le "Book to Market", la croissance et le "Momentum". Par exemple les actions de petites capitalisations tendent à surperformer (Banz (1981)). La volume moyen semble plus adéquate que la taille pour Ciliberti et al. (2017). Une autre anomalie importante est la prime "Value": les entreprises "Value" tendent à surperformer les entreprises de croissance (Fama and French (1998)). La profitabilité proche du "Cash Flow" est aussi une variable explicative significative de l'espérance des rendements (Fama and French (2015)). L'anomalie "Low Volatility" ou "Low Beta" ont aussi été révélées (Jordan and Riley (2013); Fu (2009); Ang et al. (2006)). L'anomalie la plus populaire reste le "Momentum": les actions qui ont surperformé auront tendance à continuer à surperformer (Jegadeesh and Titman (1993)). Les anomalies sont directement exploitées dans la gestion d'actifs dont les strategies sont décrites dans la section 2.1. Asness et al. (2013) expliquent ainsi qu'une stratégie de base d'investissement et très populaire simplement allouée en partie sur le "Momentum" et sur l'anomalie "Value" permet d'atteindre un Sharpe "in the sample" supérieur à 1.

Les théories financières pour justifier de telles primes de risque alternatives (manque de liquidité, asymétrie) sont remises en cause car les anomalies ont tendance à disparaitre une fois publiées. McLean and Pontiff (2016) ont plusieurs explications alternatives: le biais “in the sample” avec le problème de la suroptimisation ou l’adaptation des marchés.

A ma connaissance aucune étude ne s’est encore intéressée à la mise en évidence des autocorrélations des rendements des facteurs de risque qui pourrait constituer une inefficience plus subtile et plus robuste des marchés financiers. Une explication est que les autocorrélations sont trop difficiles à caractériser de manière significative. Des articles existent mais la significativité et la robustesse de leurs résultats ne sont pas convaincants. Ainsi Hodges et al. (2017) cherchent des prédicteurs des facteurs dans diffèrents régimes économiques et différentes conditions de marché. Ils trouvent que l’utilisation d’une combinaison d’indicateurs sur le cycle économique, la valorisation, la tendance et la dispersion serait plus efficace que l’utilisation d’indicateurs individuels.

2.6 Processus stochastique

L’instabilité de la matrice de corrélation de population a d’abord été modélisée par des modèles de diffusion pour évaluer des produits dérivés (Possamai and Gauthier (2011)). Les modèles théoriques ont été ajustés pour retrouver les prix des produits dérivés sans chercher à connaitre la réalité de la dynamique de la matrice de corrélation empirique car cette dernière est difficilement mesurable avec la précision recherchée: les modèles ARCH ont été initialement développés pour décrire l’heteroscedasticité des variations de l’inflation (Engle (1982)) mais ont ensuite été utilisés pour modéliser la dynamique de la volatilité des actions pour évaluer des options (Duan (1995)). Des modèles de type “Dynamic Conditional Correlation” (DCC GARCH, Engle (2002, 2016)) ont étendu le modèle GARCH à une dimension et ont été développés pour modéliser la dynamique des corrélations et des volatilités. De la même façon le processus introduit par Cox et al. (1985), qui est très populaire en finance pour décrire la dynamique des taux d’intérêt et de la volatilité des actions pour évaluer des produits derivés, a aussi été étendu à partir de la diffusion Feller pour modéliser la dynamique des covariances: les processus de Wishart généralisent à plusieurs dimensions la diffusion de Feller. Gourieroux (2006) introduit ainsi un terme de retour vers la moyenne au processus de Wishart en le rendant stationnaire et généralise le processus de Cox et al. (1985). Da Fonseca et al. (2007) généralisent de la même manière le modèle d’Heston (1993) pour valoriser les options multi asset. Un processus de Wishart peut être vu comme le carré de Browniens ou dans sa version stationnaire d’Ornstein-Uhlenbeck. Cuchiero et al. (2011) analysent les fondations des processus stochastiques affines continus sur l’univers des matrices de covariance motivé par l’utilisation de tels modèles pour valoriser des options multi-asset ou pour décrire les intensités de défauts. Bru (1991) dérive les équations stochastiques pour décrire la dynamique de la matrice et la dynamique des valeurs propres. D’autres matrices aléatoires sont aussi très étudiées, comme les matrices gaussiennes dont la distribution des valeurs propres suit la loi circulaire de Wigner. Ahdida and Alfonsi (2013) s’intéressent à des matrices de corrélations aléatoires à travers la diffusion de Wright-Fisher pour modéliser les corrélations des actions. Des algorithmes ont aussi été implémentés pour générer des marches aléatoires parmi les matrices de rotation. Cela permet de décrire la diffusion des vecteurs propres de la matrice de corrélation. Ainsi la marche aléatoire de Kac (1959) est un algorithme assez efficient mais il ne contient pas de retour vers la moyenne, si bien qu'au bout d'un certain temps la matrice n'a plus aucun lien avec la matrice initiale.

D'autres phénomènes assez fins, comme l'effet de levier restent mal modélisés par les modèles de la littérature. Ainsi des versions asymétriques des modèles type DCC GARCH ont été développées pour tenir compte de l'effet de levier. Malgré une littérature conséquente sur l'effet de levier (quand les prix baissent, la volatilité augmente, selon Black (1976); Christie (1982); Campbell and Hentschel (1992); Bekaert and Wu (2000); Bouchaud et al. (2001)), aucun ne s'intéresse à la réalité et la complexité du phénomène bien décrite dans Bouchaud et al. (2001). De nombreux papiers rapportent que les beta, sensibilité des prix des actions aux variations de l'indice, peuvent varier (Blume (1971); Fabozzi and Francis (1978); Jagannathan and Wang (1996); Fama and French (1997); Bollerslev et al. (1988); Lettau and Ludvigson (2001); Lewellen and Nagel (2006); Ang and Chen (2007)) sans établir de relation précise entre entre l'effet de levier et l'augmentation des beta. Les actions à fort effet de levier sont plus exposées à un beta instable (Galai and Masulis (1976); DeJong and Collins (1985)). Bien tenir compte de la variabilité des beta est important aussi pour bien tester les modèles d'Asset Pricing. Ainsi Bali et al. (2017) prétendent qu'une fois que les beta sont bien estimés à partir d'un modèle DCC GARCH, alors l'anomalie "Low Beta" disparait et le MEDAF est alors enfin vérifié empiriquement (le rendement espéré serait bien proportionnel au beta quand il est bien mesuré).

3. Contributions principales

Le travail de recherche s'est focalisé sur six sujets pointus et s'est décliné sous la forme de six projets d'articles. Les contributions principales pour chacun des six sujets sont les suivantes:

- "Emergence of Correlation of Securities at Short Time Scales" (chapitre 1) : l'article introduit un modèle multifactoriel de retard, qui reproduit assez fidèlement les mesures de l'effet d'échelle sur les valeurs propres. Le modèle s'inspire du modèle d'impact de Kyle (1985). Le modèle suppose que les transactions sur les facteurs de risque, impactent le prix des actions avec un certain retard. Je dérive, sous certaines hypothèses, une formule simple pour décrire la dépendance des valeurs propres avec l'échelle de temps. La formule contient deux paramètres pour chaque valeur propre: la valeur propre asymptotique et un temps de relaxation de l'ordre d'1 minute qui traduit un retard moyen de l'ordre de quelques minutes entre les actions et les facteurs de risque. Ainsi les corrélations apparaissent à partir d'une minute. Toutefois ce retard de quelques minutes continue d'impacter les valeurs propres de la matrice de corrélation des rendements 20 minutes et au-delà à cause d'une loi en puissance qui s'explique par un mécanisme relativement subtile bien que le phénomène sature. L'article identifie donc une inefficience significative du marché, qui pourrait générer des gains dans le cas théorique où les coûts de transactions sont nuls.

- "The Fundamental Market Neutral Maximum Variance Portfolios" (chapitre 2): l'article introduit le "FCL" d'un portefeuille (ratio entre la variance du portefeuille et la variance du portefeuille dans le cas où les corrélations entre actions seraient nulles). Le "FCL" est un concept proche des valeurs propres et a l'avantage de s'appliquer non seulement aux vecteurs propres mais aussi à n'importe quel facteur de risque. Le "FCL" serait une mesure idéale pour caractériser la significativité d'un facteur de risque. J'introduis aussi le portefeuille "fundamental Max variance" qui optimise le "FCL" et qui peut être interprété comme un vecteur propre de la matrice de corrélation sous contrainte pour capturer au mieux un style donnée defini par un critère financier. Je montre que les poids optimaux dépendent directement des classements des actions en fonction de ce critère et suivent une même loi universelle qui s'applique à tous les critères financiers. Je montre que cette optimisation permet de répliquer au mieux la matrice de corrélation à partir de quelques facteurs ainsi que sa dynamique en filtrant le bruit. Je fais le lien entre les différents "FCL", les valeurs propres sous contraintes et les valeurs propres empiriques. Je montre enfin que les vecteurs propres principaux de la matrice de corrélation s'investissent sur les facteurs qui ont les "FCL" les plus élevés. Les "FCL" sont volatiles et sont bien modélisés par des processus d'Orstein-Ulhenbeck avec un temps de relaxation de 60 jours. La composition des vecteurs propres est donc très variable ce qui explique pourquoi leur interprétation est difficile à l'exception du premier. Je montre aussi sous certaines hypothèses que le Sharpe des portefeuilles "maximum variance" est optimal théoriquement. Les résultats de ce chapitre ont été obtenus en collaboration avec Stanislav Kuperstein.

– "Time Scale Effect on Correlation at Long Time Horizon" (chapitre 3): l'article décrit une forme plus subtile mais plus robuste d'inefficience des marchés financiers que les écarts entre les espérances non conditionnelles des rendements et les prédictions du MEDAF. Il s'agit de l'autocorrélation des rendements des facteurs de risque qui s'explique par l'illiquidité des marchés financiers et par le comportement moutonnier des investisseurs qui ont tendance à acheter les produits qui ont marché. Cette autocorrélation qui n'est pas décrite dans la littérature va rendre les vecteurs propres et les valeurs propres de la matrice de corrélation sensibles à l'échelle de temps.

– "The Reactive Beta Model" (chapitre 4): l'article décrit le modèle de levier systématique (la corrélation augmentent lorsque l'indice baisse), spécifique (le beta d'une action augmente lorsque elle sous performe) et d'élasticité (lorsque la volatilité relative augmente le beta augmente). Il ressort qu'une grande partie de la variabilité des beta s'explique par ces phénomènes. L'approche qui consiste à normaliser les rendements pour corriger ces petits phénomènes permet de réduire le bais de certains facteurs (“Momentum” et “Low Beta”) par rapport à la régréssion linéaire directe sur les rendements. Des tests empiriques montrent la supériorité du modèle par rapport à la simple régréssion linéaire. Des simulations Monte-Carlo montrent aussi l’avantage d’un tel modèle par rapport aux méthodes robustes telles que les régressions par quintiles et les modèles de type DCC GARCH symétriques ou asymétriques. Je montre que mon modèle semble le plus adapté à la réalité des marchés car il a été conçu pour s’adapter à des phénomènes bien caractérisés et mesurés.

– “The Model of Diffusion of the Correlation between Securities” (chapitre 5): l’article identifie quelques faits stylisés qui caractérisent la diffusion des vecteurs propres empiriques des marchés. Les vecteurs propres de la matrice à l’instant t voient leur corrélation en utilisant la matrice à l’instant t $+\tau$ augmenter très légèrement avec τ. Je m’intéresse à la distribution des valeurs propres des incréments de la matrice de corrélation qui est différente de la loi demi-cercle de Wigner et de la distrution qui ressemble à un chapeau pointu. Les équations stochastiques standard (Wright-Fisher, Feller) qui simulent directement la matrice de corrélation ainsi que d’autres méthodes simples qui simulent des trajectoires aléatoires de la matrice de rotation autour de la matrice identité avec un terme de retour vers la moyenne pour simuler la diffusion des vecteurs propres ne permettent pas de reproduire la distribution empirique des valeurs propres. La diffusion des FCL, définies dans le chapitre 2, permet de generer simplement cette distribution. Les résultats de ce chapitre ont été obtenus en collaboration avec Stanislav Kuperstein.

- "Should Employers Pay Better their Employees? An asset Pricing Approach" (chapitre 6) : le facteur rémunération est identifié comme un facteur de risque commun significatif grâce au "FCL" mesuré qui est significativement supérieur à 1. Le facteur est ainsi aussi significatif que le facteur "Book" de Fama et French. Le facteur rémunération révèle aussi une faible anomalie de marché: les entreprises qui payent bien leurs employés ont un risque en commun et tendent à surperformer les autres. L'article remet en cause la méthodologie de Fama et French qui ne serait pas assez fine pour caractériser une telle anomalie. Ainsi il semble très important de maintenir à chaque instant le facteur beta neutre et pas seulement en moyenne seulement pour mesurer l'anomalie.

II Dissertation doctorale

1. Emergence of Correlation between Securities at Short Time Scales

Emergence of correlations between securities at short time scales

Abstract

The correlation matrix is the key element in optimal portfolio allocation and risk management. In particular, the eigenvectors of the correlation matrix corresponding to large eigenvalues can be used to identify the market mode, sectors and style factors. We investigate how these eigenvalues depend on the time scale of securities returns in the U.S. market. For this purpose, one-minute returns of the largest 533 U.S. stocks are aggregated at different time scales and used to estimate the correlation matrix and its spectral properties. We propose a simple lead-lag factor model to capture and reproduce the observed time-scale dependence of eigenvalues. We reveal the emergence of several dominant eigenvalues as the time scale increases. This important finding evidences that the underlying economic and financial mechanisms determining the correlation structure of securities depend as well on time scales.

1 Introduction

How do the eigenvalues of securities correlation matrices emerge at different time scales? This fundamental question is important because cross-correlations change over different investment horizons while a reliable empirical determination of the correlation matrix remains difficult due to its time and frequency dependence. This was first evidenced by Epps, who demonstrated the decay of correlations among U.S. stocks when shifting from daily to intra-daily time scales (or frequencies) [1]. In other words, the price correlation decreases with the duration of the time interval over which price changes are measured. The economic argument behind the Epps effect is that the information is not instantaneously transmitted at shorter time intervals, where the average adjustment lag in response of prices lies approximately between 10 and 60 minutes. This appears to reduce the scope of the Efficient Market Hypothesis [2] at short time scales given that tick data prices seem to adjust to new information only after a lag time, thus do not reflect all available information. Since its inception, the Epps effect has been confirmed by several studies, although its impact has been progressively declined in the NYSE, indicating that the market becomes increasingly more efficient [3].

The dependence of securities cross-correlations on time scales can be captured via the eigenvalues of the correlation matrix. In particular, the largest eigenvalue reflects changes in the average correlation between stocks, whereas the corresponding eigenvector is associated to the "market mode". Kwapien *et al.* showed a significant elevation of the largest eigenvalue with increasing time scale using data from 1 minute to 2 days from NYSE, NASDAQ and Deutsche Börse (1997-1999) [4]. Using high-frequency stock returns from NYSE, AMEX and NASDAQ (1994-1997), Plerou *et al.* supported the idea that the largest eigenvalue and its eigenvector reflect the collective response of the entire market to stimuli such as certain news breaks (e.g., central bank interest rates hikes) [5]. This is particularly true during periods of high volatility when the collective behavior is enhanced. Coronnello *et al.* confirmed that the largest eigenvalue, computed from 5-minute data, describes the common behavior of the stocks composing the LSE stock index (2002) [6].

As firms having similar business activities are correlated, some other eigenvectors can economically be interpreted as business sectors [7]. So, Gopikrishnan *et al.* computed the eigenvectors of cross-correlation matrices of 1000 U.S. stocks at a 30-minute scale (1994-1995) and a 1-day scale (1962-1996) [7]. They found that the correlations in a business sector, captured via an eigenvector, were stable in time and could be used for the construction of optimal portfolios with a stable Sharpe ratio. In the same vein, as similar trading strategies induce cross-correlations in stocks, some eigenvectors can be financially interpreted as style factors. The corresponding eigenvalues are thus expected to exhibit non-trivial dependence on time scales. However, an accurate statistical analysis of multiple eigenvalues at different time scales is challenging due to measurement noises. In fact, as the correlation matrix is estimated from time series of stocks' returns, its elements are unavoidably random and thus prone to fluctuations. These fluctuations become larger as the length of time series is reduced, i.e., when the time scale is increased. While the largest eigenvalue typically exceeds the level of fluctuations by two orders of magnitude, the other eigenvalues rapidly reach this level and become non-informative. Several researchers employed the random matrix theory to distinguish economically significant eigenvalues from noise [8, 9, 10, 11, 12]. In particular, Laloux *et al.* showed that only 6% of the eigenvalues carried some information of the S&P 500 (1991-1996), while the remaining 94% eigenvalues were hidden by noise [8]. Guhr and Kalber proposed an alternative statistical approach to reduce noise that they called "power mapping" [13]. Andersson *et al.* extended this work by comparing the power mapping approach to a standard filtering method discarding noisy eigenvalues for Markowitz portfolio optimization using daily Swedish stock market returns (1999-2003) [14].

In this paper, we consider the correlation matrix of financial securities and investigate the emergence of its eigenvalues at small time scales. As the financial literature on this critical issue remains sparse, this research fills the gap by investigating the eigenvalues at intraday time scales using 1-min returns. We propose a simple model, coined the "lead-lag factor model", as an adaptation of the well-known "one-factor marker model" [15] to smaller time scales and to multiple sectors and style factors. In this model, stock returns are correlated to the returns of selected factors at earlier time steps. A detailed description of the eigenvalues as functions of the time scale is then derived. An empirical validation is performed on long time series of 1-min returns of a large universe of U.S. stocks. To get several significant eigenvalues at time scales from 1 minute to 2 hours, the correlation matrix was estimated over the whole available period (2013-2017) so that variations of cross-correlations over time were ignored (note that the dynamics of the eigenvalues and eigenvectors over time has been investigated elsewhere [16, 17, 18]). In spite of its simple character, the lead-lag factor model is shown to be able to reproduce the dependence of large eigenvalues on the time scale.

The paper is organized as follows. In Sec. 2, we estimate the correlation matrix of U.S. stocks' returns at different time scales and present the empirical dependence of large eigenvalues on the time scale. To rationalize the observed behavior, we develop in Sec. 3 the lead-lag factor model and compare it to empirical results. Section 4 summarizes and concludes. Some derivations and more technical analysis of the lead-lag factor model are presented in Appendices.

2 Empirical results

2.1 Data description

We study the correlation structure of a universe that includes 533 U.S. stocks whose capitalization exceeded 1 billion dollars in 2013. For the considered period from 1st of January 2013 to 28th of June 2017, our database contains 338 176 1-min returns for each stock. We have also verified that the arithmetic aggregation of returns, $r_i(1) + \ldots + r_i(\tau)$, is almost identical to considering the product $(1+r_i(1)) \ldots (1+r_i(\tau))-1$, given that the 1-min returns $r_i(t)$ are very small.

From the time series of 1-min returns, we estimate the correlation matrix over the whole available period, and then compute its eigenvalues. Then we aggregate the returns into 2-min, 4-min, ..., 128-min returns, producing time series with 169 088, 84 544, ..., 2 642 points, respectively. At each time scale τ, we repeat the computation to investigate the dependence of the eigenvalues on τ.

2.2 Empirical results

Figure 1a shows the four largest eigenvalues of the covariance matrix of 533 U.S. stocks' returns, computed by aggregating 1-min returns with the time scale τ, ranging from 1 minutes to 128 minutes (2 hours). The first two eigenvalues exhibit almost linear growth with τ, the others show minor deviations from linearity at small τ but scale linearly with τ at large τ. This behavior reflects the diffusion-like growth of the variance of aggregated returns; in particular, if the returns were independent, the eigenvalues of the corresponding covariance

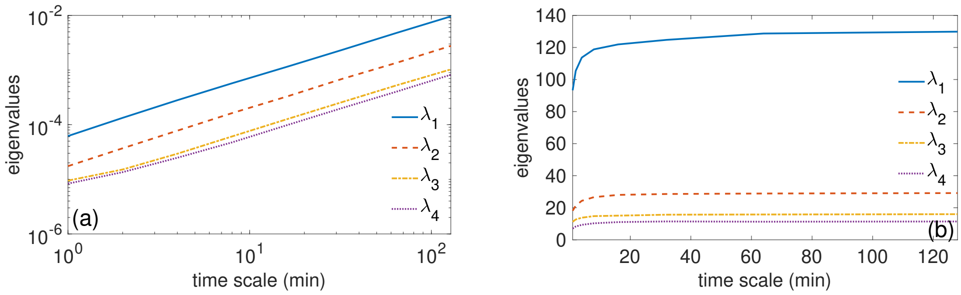

Figure 1: Four largest eigenvalues of the covariance matrix **(a)** and of the correlation matrix **(b)** for returns of 533 U.S. stocks, computed by aggregating 1-min returns with the time scale τ, varying from 1 minute to 128 minutes (2 hours).

matrix, $C_{ij} = \tau\sigma_i^2\delta_{ij}$, would be just $\lambda_i = \tau\sigma_i^2$, and thus proportional to τ. Although correlations affect this linear growth, their effect is subdominant, at least for large eigenvalues, as witnessed by Fig. 1a. To highlight the effect of correlations, we focus on the eigenvalues of the *correlation* matrix. This choice is also justified from the financial point of view to level off the variability of stocks volatilities.

Figure 1b shows the four largest eigenvalues of the correlation matrix of the same 533 U.S. stocks' returns. If the returns were independent, the correlation matrix would be the identity, and thus all its eigenvalues would be equal to 1. The growth of these eigenvalues with the time scale τ indicates strong cross-correlations between stocks. The largest eigenvalue can be naturally attributed to the market mode, whereas the next eigenvalues correspond to different sectors and style factors.

After a sharp growth at short time scales (few minutes), the eigenvalues slowly approach to their long-time limits. The existence of these upper bounds is expected because the sum of eigenvalues of a correlation matrix is equal to its size (i.e., to the number of stocks, N). This saturation effect contrasts with the unlimited growth of eigenvalues of the covariance matrix (Fig. 1a). Finding the functional form of this approach and identifying its characteristic time scales present the main aim of our work. Recently, Benzaquen *et al.* proposed a multivariate linear propagator model for dissecting cross-impact on stock markets and revealing their dynamics [19]. Due to its very general form accounting for both cross-correlations and auto-correlations of stocks, the proposed model contains too many parameters, while the resulting formulas are not explicit. Our ambition is rather the opposite and consists in suggesting an explicit model, as simple as possible, that would capture the empirical results shown in Fig. 1b and thus provide a minimalistic framework for their financial interpretation.

3 The lead-lag factor model

3.1 Basic lead-lag one-factor model

We consider a trading universe with N assets. In a conventional one-factor model, the return of the i-th asset at time t, $r_i(t)$, is modeled as a combination of a specific, asset-dependent random fluctuation, $\varepsilon_i(t)$, and an overall market contribution, $R(t)$,

$$r_i(t) = \varepsilon_i(t) + \beta R(t), \tag{1}$$

with a market sensitivity β (that we generalize below to other factors). The asset-specific random fluctuations $\varepsilon_i(t)$ are typically modeled as independent centered Gaussian variables with volatilities σ_i.

We propose a modification of this conventional model by incorporating the lead-lag effect, in which the i-th asset return at time t is influenced by a common factor $R(t-k)$ at earlier times $t-k$, with progressively decaying weights:

$$r_i(t) = \varepsilon_i(t) + \beta \sum_{k=0}^{\infty} \alpha^k R(t-k), \tag{2}$$

where $0 \le \alpha < 1$ characterizes the relaxation time of the memory decay. Note that the upper limit of the sum in Eq. (2) is formally extended to infinity, bearing in mind that contributions for very large k are exponentially small. We will analyze the model in the stationary regime as $t \to \infty$ in order to eliminate transient effects.

The common term $R(t)$ can be interpreted as an idealized factor without auto-correlations in an efficient market that most stocks follow with a lead lag delay. We model therefore $R(t)$ by independent centered Gaussian variables with volatility Σ. The term $R(t)$ can represent the market mode but also sectors or style factors, or any popular trading portfolio. Moreover, $R(t)$ can also be interpreted as being linked to the market order transactions for a particular strategy (market, sector or styles). In this light, our model can be seen as an extension of the Kyle model [20] that explains the impact of transactions on price for a single stock and without delay. Here, we consider multiple stocks and include an exponential decay of the impact. While more sophisticated models with a power law decay of the impact were proposed [19, 21], we will show that our minimalistic model is enough to reproduce a slow growth of the eigenvalues of the correlation matrix. For the sake of clarify, we first analyze this basic lead-lag one-factor model and then discuss its several straightforward extensions.

The one-factor relation (2) is the basic model for returns at the smallest time scale. We then consider the returns aggregated on the time scale τ:

$$r_i^\tau(t) = \sum_{\ell=0}^{\tau-1} r_i(t-\ell), \tag{3}$$

with t being a multiple of τ. Under the former Gaussian assumptions, the covariance function of the aggregated returns reads (see A):

$$C_{ij}^{\tau} = \langle r_i^{\tau}(t) r_j^{\tau}(t) \rangle = \tau\sigma^2\delta_{ij} + \frac{\beta^2\Sigma^2\big(\tau(1-\alpha^2) - 2\alpha(1-\alpha^{\tau})\big)}{(1-\alpha^2)(1-\alpha)^2}\,, \tag{4}$$

where $\langle \cdots \rangle$ denotes the expectation, and $\delta_{ij} = 1$ for $i = j$, and 0 otherwise. Note that we set here $\sigma_i = \sigma$ for all assets for simplicity (this simplification will be relaxed below). As we consider the stationary regime, the covariance function does not depend on time t.

Denoting

$$\kappa_\alpha(\tau) = \frac{\tau(1-\alpha^2) - 2\alpha(1-\alpha^{\tau})}{1-\alpha^2}\,, \tag{5}$$

one gets the correlation matrix

$$\mathcal{C}_{ij}^{\tau} = \frac{C_{ij}^{\tau}}{\sqrt{C_{ii}^{\tau} C_{jj}^{\tau}}} = \begin{cases} 1 & i = j, \\ \rho^2(\tau) & i \neq j, \end{cases} \tag{6}$$

with

$$\rho(\tau) = \big(1 + \eta(\tau)/\gamma\big)^{-1/2}, \tag{7}$$

where

$$\gamma = \frac{\Sigma^2\beta^2}{\sigma^2} \tag{8}$$

and

$$\eta(\tau) = \frac{\tau}{\kappa_\alpha(\tau)/(1-\alpha)^2} = \frac{(1-\alpha)^2}{1 - \frac{2\alpha}{1-\alpha^2}(1-\alpha^{\tau})/\tau}\,. \tag{9}$$

The function $\eta(\tau)$, that will play the central role in our analysis, monotonously decreases from $\eta(1) = 1 - \alpha^2$ to $\eta(\infty) = (1-\alpha)^2$.

Since the matrix $\mathcal{C}^{\tau} - (1 - \rho^2(\tau))I$ has rank 1 (I being the identity $N \times N$ matrix), there are $N-1$ eigenvalues $\lambda_i = 1 - \rho^2(\tau)$. In turn, the single largest eigenvalue of the correlation matrix $\mathcal{C}^{\tau}$ can be obtained as follows: $N = \mathrm{Tr}(\mathcal{C}^{\tau}) = \lambda_1 + (N-1)\lambda_i$, from which $\lambda_1 = 1 + (N-1)\rho^2(\tau)$. We get thus the complete description of the eigenvalues as functions of the time scale τ:

$$\lambda_1 = 1 + (N-1)\rho^2(\tau), \tag{10}$$

$$\lambda_i = 1 - \rho^2(\tau) \qquad (i = 2, 3, \ldots, N). \tag{11}$$

In the limit of very large τ, one finds

$$\rho^2(\infty) = \big(1 + (1-\alpha)^2/\gamma\big)^{-1}. \tag{12}$$

This simplest lead-lag one-factor model predicts a monotonous growth of the largest eigenvalue (corresponding to the market mode) with the time scale τ, up to a saturation plateau. In turn, the other eigenvalues exhibit a monotonous decrease to a plateau. In spite of the exponential decay of the lead-lag memory effect in Eq. (2), the approach to the plateau is governed by a slow, $1/\tau$ power law, in a qualitative agreement with the empirical observation (see Sec. 3.5 for quantitative comparison). In particular, this approach has no well-defined time scale.

While the basic model can potentially capture the behavior of the largest eigenvalue, it clearly fails to distinguish other eigenvalues. One needs therefore to relax some simplifying assumptions to render the model more realistic.

3.2 General lead-lag one-factor model

We start by introducing arbitrary volatilities σ_i and sensitivities β_i of the i-th asset to the common factor $R(t)$:

$$r_i(t) = \varepsilon_i(t) + \beta_i \sum_{k=0}^{\infty} \alpha^k R(t-k). \tag{13}$$

In this case, the computation is precisely the same, the only difference is that

$$C_{ij}^{\tau} = \tau\sigma_i^2\delta_{ij} + \Sigma^2\beta_i\beta_j\kappa_\alpha(\tau). \tag{14}$$

As a consequence, the structure of the correlation matrix is fully determined by β_i, whereas the dependence on the time scale τ is still represented by $\kappa_\alpha(\tau)$. The correlation matrix reads

$$\mathcal{C}_{ij}^{\tau} = \begin{cases} 1 & (i=j), \\ \rho_i(\tau)\rho_j(\tau) & (i \neq j), \end{cases} \tag{15}$$

with

$$\rho_i(\tau) = \big(1 + \eta(\tau)/\gamma_i\big)^{-1/2}, \qquad \gamma_i = \frac{\Sigma^2\beta_i^2}{\sigma_i^2}\,. \tag{16}$$

The eigenvalues of this correlation matrix can be computed as follows.

If all γ_i are distinct[1], the components of an eigenvector are

$$v_i = \frac{\rho_i Q}{\lambda - 1 + \rho_i^2} \quad (i = 1,\dots,N), \qquad \text{with } \; Q = \sum_{i=1}^{N} \rho_i v_i\,, \tag{17}$$

[1] When some γ_i are identical, the analysis of eigenvalues becomes more involved (see B), but the largest eigenvalue still satisfies Eq. (18) and can thus be approximated by Eq. (19). In particular, if all $\gamma_i = \gamma$, one gets $\lambda_1 \approx N\rho^2(\tau)$, which is close to the exact solution (10).

from which one gets the equation on the eigenvalues λ

$$\sum_{i=1}^{N} \frac{\rho_i^2}{\lambda - 1 + \rho_i^2} = 1. \tag{18}$$

This equation has N distinct solutions that can be characterized in terms of ρ_i^2 (see B). When N is large, the largest eigenvalue is expected to be large, and the asymptotic expansion of Eq. (18) yields

$$\lambda_1 \approx \sum_{i=1}^{N} \rho_i^2 = \sum_{i=1}^{N} \big(1 + \eta(\tau)/\gamma_i\big)^{-1}. \tag{19}$$

In turn, the other eigenvalues are below 1 (see B). As a consequence, such a lead-lag one-factor model cannot reproduce several eigenvalues larger than 1. For this purpose, one needs to consider multiple factors.

3.3 General lead-lag multi-factor model

Now we consider a general lead-lag multi-factor model

$$r_i(t) = \varepsilon_i(t) + \sum_{k=0}^{\infty} \alpha^k \sum_{f=1}^{F} \beta_{i,f}\, R_f(t-k), \tag{20}$$

where $\varepsilon_i(t)$ are independent centered Gaussian variables (representing random fluctuations specific to the stock i) with variance σ_i^2, F is the number of factors, $R_f(t)$ are independent centered Gaussian returns of the factor f with variance Σ_f^2, $\beta_{i,f}$ is the sensitivity of the stock i to the factor f, and α sets the relaxation time. Repeating the computation from A, one gets

$$\mathcal{C}_{ij}^{\tau} = \delta_{ij} + (1 - \delta_{ij}) \sum_{f=1}^{F} \rho_{i,f}\rho_{j,f}, \tag{21}$$

where

$$\rho_{i,f}(\tau) = \frac{\Sigma_f \beta_{i,f}}{\beta_i}\, \rho_i(\tau) \tag{22}$$

and

$$\rho_i(\tau) = \big(1 + \eta(\tau)/\gamma_i\big)^{-1/2}, \qquad \gamma_i = \frac{\beta_i^2}{\sigma_i^2}, \qquad \beta_i^2 = \sum_{f=1}^{F} \Sigma_f^2\, \beta_{i,f}^2\,. \tag{23}$$

Considering $\rho_{i,f}$ as the elements of an $N \times F$ matrix ρ, one can rewrite Eq. (21) in a matrix form

$$\mathcal{C}^{\tau} = (I - P) + \rho\rho^{\dagger}, \tag{24}$$

where P is the diagonal matrix formed by ρ_i^2, and $\dagger$ denotes the matrix transpose.

The matrix ρ of size $N \times F$ plays the central role in the following analysis. As the elements of the matrix ρ are real, $\rho\rho^\dagger$, as well as $\rho^\dagger\rho$, are positive semi-definite matrices which have nonnegative eigenvalues. The rank of the matrix ρ is equal to that of matrices $\rho\rho^\dagger$ and $\rho^\dagger\rho$ and thus cannot exceed $\min\{F, N\}$. Given that $F \ll N$, the correlation matrix $\mathcal{C}^\tau$ appears as the perturbation of a diagonal matrix by a low-rank matrix.

The eigenvalues of the correlation matrix are the zeros of the determinant

$$0 = \det\bigl(\lambda I - \mathcal{C}^\tau\bigr) = \det\bigl(\lambda I - I + P - \rho\rho^\dagger\bigr). \tag{25}$$

Since $\rho\rho^\dagger$ is a low-rank perturbation, one can expect, as in the one-factor case of B, that most eigenvalues coincide with that of the unperturbed diagonal matrix $I - P$, i.e., they are given by $1 - \rho_i^2$ for some indices i. These eigenvalues are essentially hidden by noise and non-exploitable in practice. We are interested in large eigenvalues that (significantly) exceed 1.

If λ exceeds 1, it cannot be equal to $1 - \rho_i^2$ for all i, the matrix $\lambda I - I + P$ is nonsingular, its inverse exists, so that one can rewrite Eq. (25) as

$$0 = \det\bigl(\lambda I - I + P\bigr)\det\bigl(I - \rho^\dagger(\lambda I - I + P)^{-1}\rho\bigr), \tag{26}$$

from which one gets a new equation on eigenvalues:

$$0 = \det\bigl(I - \underbrace{\rho^\dagger(\lambda I - I + P)^{-1}\rho}_{\phi(\lambda)}\bigr). \tag{27}$$

(here we used a general property: if $A \in \mathbb{C}^{m\times m}$ is nonsingular matrix and $U, V \in \mathbb{C}^{m\times r}$, then $\det(A + UV^*) = \det(A)\det(I + V^*A^{-1}U)$, see [22]). Denoting the $F \times F$ matrix in the determinant as $\phi(\lambda)$, one can write explicitly its elements as

$$\phi_{f,g}(\lambda) = \sum_{i=1}^{N} \frac{\rho_{i,f}\,\rho_{i,g}}{\lambda - 1 + \rho_i^2}\,. \tag{28}$$

The solutions of Eq. (27) determine some eigenvalues λ of the correlation matrix in Eq. (21). As one typically deals with the situation $N \gg F$, the reduction of the original determinant equation (25) for a matrix of size $N \times N$ to Eq. (27) for a matrix of size $F \times F$ is a significant numerical simplification of the problem. Most importantly, this formal solution allows one to get analytical insights onto the eigenvalues, as we did in the one-factor case in B. Note that in the one-factor case ($F = 1$), the determinant equation (27) is simply reduced to

$$0 = \det(I - \phi(\lambda)) = 1 - \phi_{1,1}(\lambda) = 1 - \sum_{i=1}^{N} \frac{\rho_i^2}{\lambda - 1 + \rho_i^2}\,, \tag{29}$$

i.e., we retrieve Eq. (18).

If one searches for large eigenvalues, $\lambda \gg 1$, one can neglect the matrix $P-I$ in comparison to λI in Eq. (27), that yields

$$\det\bigl(\lambda I - \rho^\dagger \rho\bigr) = 0. \tag{30}$$

In other words, the large eigenvalues of the correlation matrix can be approximated by the eigenvalues of the matrix $\rho^\dagger\rho$ of size $F \times F$. This symmetric positive semi-definite matrix has F nonnegative eigenvalues that correspond to F factors.

3.4 Practical approximation

As we will discuss in detail in Sec. 3.5, empirical data exhibit the short-range memory effect (α is small) and the relatively small impact of the factors onto the variance of individual stocks as compared to the stock-specific fluctuations (γ_i are small). In this situation, which is particular to the time series of securities returns at the considered time scales, one has $\eta(\tau)/\gamma_i \gg 1$ so that $\rho_i(\tau)$ in Eq. (23) can be approximated as

$$\rho_i^2(\tau) \simeq \frac{\gamma_i}{\eta(\tau)}. \tag{31}$$

This approximation greatly simplifies the elements of the matrix $\rho^\dagger\rho$:

$$(\rho^\dagger\rho)_{f,g} = \sum_{i=1}^{N} \underbrace{\frac{\Sigma_f \beta_{i,f}\rho_i(\tau)}{\beta_i}}_{=\rho_{i,f}} \underbrace{\frac{\Sigma_g \beta_{i,g}\rho_i(\tau)}{\beta_i}}_{=\rho_{i,g}} \approx \frac{N}{\eta(\tau)}\,\Gamma_{f,g}\,, \tag{32}$$

where the matrix elements $\Gamma_{f,g}$ do not depend on the time scale:

$$\Gamma_{f,g} = \frac{\Sigma_f\Sigma_g}{N}\sum_{i=1}^{N}\frac{\beta_{i,f}\beta_{i,g}}{\sigma_i^2}\,. \tag{33}$$

As a consequence, all the elements of the matrix $\rho^\dagger\rho$ and thus its eigenvalues exhibit the same dependence on the time scale τ, expressed via the explicit function $\eta(\tau)$ given by Eq. (9). Denoting the eigenvalues of the matrix Γ as γ_f ($f = 1,\dots,F$), one gets the following approximation for large eigenvalues of the correlation matrix:

$$\lambda_f \approx \frac{N\gamma_f}{\eta(\tau)} \qquad (f = 1,\dots,F)\,. \tag{34}$$

From the explicit form (16) of $\eta(\tau)$, one deduces a slow, $1/\tau$, power law approach of the eigenvalue to the saturation level as the time scale τ increases. Within this approximate computation, all large eigenvalues exhibit the same dependence on the time scale.

f	1	2	3	4
γ	0.17	0.03	0.02	0.01
α	0.16	0.25	0.18	0.26
t_α (min)	0.55	0.72	0.58	0.74

Table 1: Two adjustable parameters of the fitting formula (34) applied to four largest eigenvalues of the correlation matrix of $N = 533$ U.S. stocks' returns. The corresponding relaxation time t_α in minutes is obtained as 1 min/ $\ln(1/\alpha)$.

In practice, one aims at constructing the factors R_f to capture independent features of cross-correlations in the market. The sensitivies $\beta_{i,f}$ and $\beta_{i,g}$ of the stock i to factors R_f and R_g are thus expected to be "orthogonal", and this property can be formally expressed by requiring that the nondiagonal elements of the matrix Γ are negligible. In this case, the eigenvalues γ_f are given by the diagonal elements

$$\gamma_f = \Gamma_{f,f} = \frac{1}{N} \sum_{i=1}^{N} \frac{\Sigma_f^2 \beta_{i,f}^2}{\sigma_i^2} . \tag{35}$$

This is a kind of empirical mean of the squared sensitivities $\beta_{i,f}^2$, normalized by the squared volatilities σ_i^2.

3.5 Application to empirical data

We aim at applying the lead-lag factor model to fit the eigenvalues of the empirical correlation matrix of U.S. stocks' returns. The fitting formula (34) has two adjustable parameters: the relaxation time α in the function $\eta(\tau)$ and the amplitude $N\gamma_f$. Using the least square fitting algorithm implemented as the routine `lsqcurvefit` in Matlab, we apply the formula (34) separately to each empirical eigenvalue.

Figure 2 shows the fitting of the four largest eigenvalues. The good quality of the fit by the lead-lag factor model indicates that, in spite of numerous simplifying assumptions on which the model was built, it captures the overall behavior qualitatively well. In particular, the eigenvalues converge to limiting values, at least for the considered short-time scales (up to 2 hours). Moreover, this saturation level is approached slowly, with the characteristic $1/\tau$ power law dependence. The adjustable parameters are summarized in Table 1. Rewriting the attenuation factor α^k in the lead-lag factor model (2) as $\exp(-t/t_\alpha)$ with $t = k\tau_0$ and $t_\alpha = \tau_0/\ln(1/\alpha)$, where $\tau_0 = 1$ min is the finest time scale of the time series used, one gets the relaxation time t_α in minutes. One can see that the relaxation times α (or t_α) for four eigenvalues are close to each other. In other words, all the dominant eigenmodes evolve at comparable time scales. This is an important conclusion which refutes a common

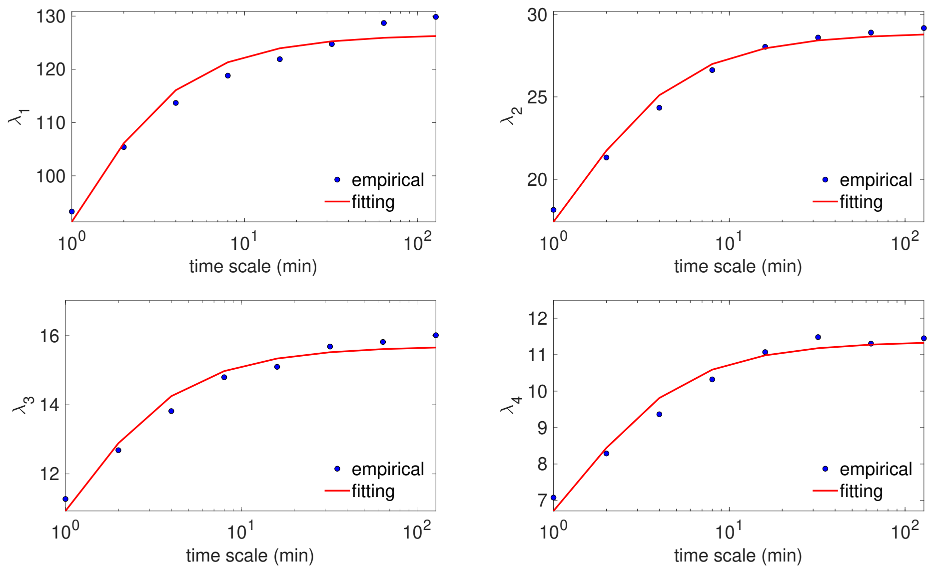

Figure 2: Fitting by Eq. (34) of the four largest eigenvalues of the correlation matrix of $N = 533$ U.S. stocks' returns, computed by aggregating 1-min returns with the time scale τ. The adjustable parameters α and γ are summarized in Table 1.

belief that the market mode (corresponding to the largest eigenvalue) evolves at a time scale that is significantly different from other modes (sectors and style factors). The values of t_α are of the order of one minute, in agreement with predictions by Benzaquen *et al.* [19]. Remarkably, while the lead-lag memory effects vanish so rapidly, they impact the behavior of the eigenvalues at much longer time scales. In particular, if the lead-lag was ignored (by setting $\alpha = 0$), the largest eigenvalue would be $\simeq N\gamma_1$ and independent of the time scale τ. For instance, using the estimated value $\gamma_1 = 0.17$ and setting $\alpha = 0$, one would get the largest eigenvalue to be 90, which is significantly smaller than the expected limit 128 for $\alpha = 0.16$ or the observed value 130 at $\tau = 128$ min.

4 Conclusion

We investigated the dependence of the eigenvalues of the correlation matrix on the time scale τ. Aggregating 1-min returns of the largest 533 U.S. stocks (2013-2017) to estimate the correlation matrix at different time scales, we showed that its large eigenvalues grow with τ and apparently saturate to limiting values. This growth reflects the important phenomenon that inter-stock correlations accumulate over time scales.

To rationalize this phenomenon and to interpret empirical observations, we developed the lead-lag factor model. In the one-factor case, each stock is considered to be partly correlated to a given lead-lag factor. Under several simplifying assumptions, we derived a simple formula for large relevant eigenvalues. This formula containing just two easily interpretable adjustable parameters, was then validated on empirical data.

The relaxation time of the stock market was estimated to be around 1 minute. A possible interpretion of this observation can be that a transaction can generate a cascade of transactions that decays in 1 minute so that the impact of transaction on price decays in 1 minute. As correlations emerge from the cross-impact of transactions on prices, we model this effect by extending the Kyle model to the impact of transaction on preferential portfolios with a lead lag effect.

The small value of the observed relaxation time suggests that correlation measurements based on 5 minutes returns should provide a good proxy of correlation of daily returns for risk management, in line with the conclusion by Liu *et al.* on volatility estimation [23]. However, other phenomena are likely to occur at much larger time scales (from day to month), e.g., autocorrelations of returns of financial factors (book, size, momentum) due to herding effect, or lack of liquidity. An accurate estimation of correlations at larger time scales remains a challenging problem because of a limited number of the available returns and thus higher impact of noise in the estimated correlation matrix. To overcome this limitation, one can either consider time horizons over several decades (in which case neglecting variations of corrections over time becomes debatable), or reduce the number of considered securities and thus the dimension of the correlation matrix (in which case financial meaning of estimated correlations may be debatable). A possible solution consists in constructing relevant financial factors and investigating how their correlations change with the time scale, as suggested by our factor-based model.

A Computation of the covariance matrix

The covariance matrix of aggregated centered Gaussian returns $r_i^\tau(t)$ defined by Eq. (3) is

$$\begin{aligned} C_{ij}^\tau &= \langle r_i^\tau(t) r_j^\tau(t)\rangle \\ &= \tau\sigma^2\delta_{ij} + \beta^2 \sum_{\ell_1,\ell_2=0}^{\tau-1} \sum_{k_1,k_2=0}^{\infty} \alpha^{k_1}\alpha^{k_2}\langle R(t-\ell_1-k_1)R(t-\ell_2-k_2)\rangle . \end{aligned} \tag{36}$$

The first term in this expression comes from the uncorrelated stock-dependent fluctuations. The independence of returns $R(k)$ implies

$$C_{ij}^\tau = \tau\sigma^2\delta_{ij} + \beta^2\sigma_m^2 \sum_{\ell_1,\ell_2=0}^{\tau-1} \sum_{k_1,k_2=0}^{\infty} \alpha^{k_1}\alpha^{k_2}\delta_{\ell_1+k_1,\ell_2+k_2}. \tag{37}$$

To calculate these four sums, it is convenient to consider separately various terms depending on ℓ_1 and ℓ_2:

- there are τ terms with $\ell_1 = \ell_2$ that implies $k_1 = k_2$, whose contribution is

$$\tau \sum_{k=0}^{\infty} \alpha^{2k} = \frac{\tau}{1-\alpha^2} \,; \tag{38}$$

- there are $\tau - 1$ terms with $\ell_1 = \ell_2 + 1$ that implies $k_1 = k_2 - 1$, whose contribution is

$$(\tau - 1) \sum_{k=0}^{\infty} \alpha^{2k+1} = \frac{(\tau-1)\alpha}{1-\alpha^2} \,. \tag{39}$$

Moreover, the same contribution comes from $\ell_1 = \ell_2 - 1$ and $k_1 = k_2 + 1$.

- similarly, there are $\tau - j$ terms with $\ell_1 = \ell_2 + j$ that implies $k_1 = k_2 - j$, whose contribution is

$$(\tau - j) \sum_{k=0}^{\infty} \alpha^{2k+j} = \frac{(\tau-j)\alpha^j}{1-\alpha^2} \,, \tag{40}$$

and this contribution is doubled by the symmetry argument.

- finally, there is one term with $\ell_1 = \ell_2 + (\tau - 1)$ and thus $k_1 = k_2 - (\tau - 1)$ whose contribution is $\alpha^{\tau-1}/(1-\alpha^2)$.

Combining all these terms, one gets after simplifications Eq. (4).

B Analysis of the lead-lag one-factor model

We study in more detail the model (15) of the correlation matrix $\mathcal{C}$, with $\rho_i(\tau)$ given by Eq. (16). This matrix is a perturbation of the identity matrix by a rank one matrix, for which many spectral properties are known (see, e.g., [24]). This matrix combines both effects: the correlation coefficient ρ and the impact of the exponential moving average (with the coefficient α). We search for an eigenvector of this matrix as $v = (v_1, v_2, \ldots, v_n)^\dagger$. Writing explicitly $\mathcal{C}v = \lambda v$, we get

$$v_i(1-\rho_i^2) + \rho_i Q = \lambda v_i \qquad (i = 1, \ldots, N), \tag{41}$$

where

$$Q = \sum_{i=1}^{N} v_i \rho_i. \tag{42}$$

First, we note that if $\rho_i = 0$ for some i, then the above equation is reduced to $v_i = \lambda v_i$ that has two solutions: either $\lambda = 1$ and v_i can be arbitrary; or $v_i = 0$. One can check that if $\rho_{i_1} = \ldots = \rho_{i_k} = 0$ for k stocks, then the correlation matrix has the eigenvalue $\lambda = 1$ with the multiplicity k. The corresponding eigenvectors can be chosen as an orthogonal basis in the subspace $\mathbb{R}^k$. In turn, the remaining $n-k$ eigenvalues are nontrivial, and can be determined as discussed below. In what follows, we focus on these nontrivial eigenvalues, i.e., we assume that all $\rho_i \neq 0$.

The equation (41) has two solutions:

(i) either $\lambda = 1-\rho_i^2$ and $Q=0$; or

(ii) $\lambda \neq 1-\rho_i^2$ and

$$v_i = \frac{\rho_i Q}{\lambda - 1 + \rho_i^2}\,. \tag{43}$$

In the latter case, one can substitute this expression into Eq. (42) to get an equation on the eigenvalue λ:

$$\sum_{i=1}^{N} \frac{\rho_i^2}{\lambda - 1 + \rho_i^2} = 1. \tag{44}$$

This equation can be seen as a polynomial of degree N which has N (*a priori* complex-valued) zeros. Finally, Q can be fixed by setting the normalization condition on v:

$$1 = \sum_{i=1}^{N} v_i^2 = Q^2 \sum_{i=1}^{N} \frac{\rho_i^2}{(\lambda - 1 + \rho_i^2)^2}\,. \tag{45}$$

This is a generic situation.

Let us return to the first option, namely, we suppose that $\lambda = 1-\rho_k^2$ for some index k that implies that $Q=0$. If all ρ_i are distinct, i.e., $\rho_1 \neq \rho_2 \neq \ldots \neq \rho_N$, so that $v_i = 0$ for all $i \neq k$, but, due to $Q=0$, it would also imply that $v_k = 0$. As a consequence, $v=0$ but this is not an eigenvector. We conclude that, if all ρ_i are distinct, then λ cannot be given by $1-\rho_i^2$, and this option is excluded.

Now, we consider the case when two or more values ρ_i are identical. For instance, let us assume that $\rho_1 = \rho_2 \neq \rho_3 \neq \ldots \neq \rho_N$. In this case, $\lambda = 1-\rho_1^2$ is indeed an eigenvalue. In fact, one gets $Q=0$ and thus $v_i = 0$ for $i>2$. However, one has $Q = \rho_1 v_1 + \rho_2 v_2 = \rho_1(v_1+v_2) = 0$, implying that $v_1 = -v_2$. The normalization condition implies thus $v_1 = -v_2 = 1/\sqrt{2}$. We conclude that $\lambda = 1-\rho_1^2$ is then a single eigenvalue. More generally, if $\rho_1 = \rho_2 = \ldots = \rho_k \neq \rho_{k+1} \neq \ldots \neq \rho_N$, then the eigenvalue $\lambda = 1-\rho_1^2$ has the multiplicity $k-1$.

In general, it is convenient to denote $z_i = 1-\rho_i^2$ and to order them in an increasing order:

$$z_1 \leq z_2 \leq z_3 \leq \ldots \leq z_N \tag{46}$$

or, equivalently, by grouping the eventual identical values:

$$\begin{aligned} z_1 = z_2 = \ldots = z_{i_1} < z_{i_1+1} = z_{i_1+2} = \ldots = z_{i_1+i_2} \\ < \ldots < z_{i_1+\ldots+i_m} = z_{i_1+\ldots+i_m+1} = \ldots = z_N. \end{aligned} \tag{47}$$

In other words, there are i_1 identical values $z_1 = \ldots = z_{i_1}$; i_2 identical values $z_{i_1+1} = \ldots = z_{i_1+i_2}$, etc. (note that when all z_i are distinct, one has $i_1 = i_2 = \ldots = 1$). In this configuration, the correlation matrix has: the eigenvalue z_1 with the multiplicity $i_1 - 1$ (if $i_1 > 1$); the eigenvalue z_2 with the multiplicity $i_2 - 1$ (if $i_2 > 1$); etc. If for some k, $z_{i_k} = 1$, then this eigenvalues has the multiplicity i_k. Finally, the remaining eigenvalues are determined as solutions of Eq. (44) that can be written as $f(z) = 1$, with

$$f(z) = \sum_{i=1}^{N} \frac{\rho_i^2}{z - 1 + \rho_i^2} = \sum_{i=1}^{N} \frac{1 - z_i}{z - z_i}\,. \tag{48}$$

The terms with $z_i = 1$ (resulting in the eigenvalue $\lambda = 1$) are excluded from this sum. Moreover, if some z_i are identical, the corresponding terms are just grouped together. As a consequence, the equation $f(z) = 1$ is reduced to a polynomial of degree at most N (the degree N corresponding to the case when all z_i are distinct).

It is worth noting that the function $f(z)$ is decreasing everywhere:

$$f'(z) = -\sum_{i=1}^{N} \frac{\rho_i^2}{(z - z_i)^2} < 0. \tag{49}$$

As a consequence, one gets immediately that each interval (z_i, z_{i+1}) (with $z_i < z_{i+1}$ and $z_{N+1} = \infty$) has exactly one solution of the equation $f(z) = 0$, i.e., one eigenvalue. In particular, one gets the following bounds for the smallest eigenvalue

$$z_1 \leq \min_{1\leq i\leq N} \{\lambda_i\} \leq z_2. \tag{50}$$

We conclude that all eigenvalues are positive if and only if $z_1 \geq 0$, i.e., $\rho_i^2 \leq 1$ for all i. In other words, the inequalities $\rho_i^2 \leq 1$ for all i present the necessary and sufficient condition for the positive definiteness of the matrix. These conditions are evidently satisfied in our setting.

Since $f(1) \geq 1$, one also gets the following bound for the largest eigenvalue

$$\lambda_1 = \max_{1\leq i\leq N} \{\lambda_i\} \geq 1 \tag{51}$$

(note that the eigenvalues are ordered in descending order, $\lambda_1 \geq \lambda_2 \geq \ldots$, in contrast to z_k). However, this bound is rather weak. In turn, since $\lambda_2 \leq z_N = 1 - \rho_N^2 < 1$, all other eigenvalues are below 1:

$$\lambda_i < 1 \qquad (i = 2, 3, \ldots, N). \tag{52}$$

References

References

[1] T. W. Epps, Comovements in Stock Prices in the Very Short Run, J. Am. Stat. Asso. **74**, 291-298 (1979).

[2] E. F. Fama, Efficient capital markets: A review of theory and empirical work, J. Finance **25**, 383-417 (1970).

[3] B. Tóth and J. Kertész, Increasing market efficiency: Evolution of cross-correlations of stock returns, Physica A **360**, 505-515 (2006).

[4] J. Kwapien, S. Drozdz, and J. Speth, Time scales involved in emergent market coherence, Physica A **337**, 231-242 (2004).

[5] V. Plerou, P. Gopikrishnan, B. Rosenow, L. A. Nunes Amaral, T. Guhr, and H. E. Stanley, Random matrix approach to cross correlations in financial data, Phys. Rev. E **65**, 066126 (2002).

[6] C. Coronnello, M. Tumminello, F. Lillo, S. Micciché, R. N. Mantegna, Sector Identification in a Set of Stock Return Time Series Traded at the London Stock Exchange, Acta Phys. Pol. B **36**, 2653 (2005).

[7] P. Gopikrishnan, B. Rosenow, V. Plerou, and H. E. Stanley, Quantifying and interpreting collective behavior in financial markets, Phys. Rev. E **64**, 035106(R) (2001).

[8] L. Laloux, P. Cizeau, J.-P. Bouchaud, and M. Potters, Noise Dressing of Financial Correlation Matrices, Phys. Rev. Lett. **83**, 1467 (1999).

[9] S. Pafka, M. Potters, and I. Kondor, Exponential Weighting and Random-Matrix-Theory-Based Filtering of Financial Covariance Matrices for Portfolio Optimization, ArXiv cond-mat 0402573v1 (2004).

[10] Z. Burda, A. Gorlich, A. Jarosz, and J. Jurkiewicz, Signal and noise in correlation matrix, Physica A **343**, 295-310 (2004).

[11] M. Potters, J.-P. Bouchaud, and L. Laloux, Financial applications of random matrix theory: Old laces and new pieces, Acta Phys. Pol. B **36**, 2767-2784 (2005).

[12] T. Conlon, H. J. Ruskin, and M. Crane, Random matrix theory and fund of funds portfolio optimisation, Physica A **382**, 565-576 (2007).

[13] T. Guhr and B. Kälber, A new method to estimate the noise in financial correlation matrices, J. Phys. A **36**, 3009 (2003).

[14] P.-J. Andersson, A. Öberg, and T. Guhr, Power Mapping and Noise Reduction for Financial Correlations, Acta Phys. Pol. B **36**, 2611 (2005).

[15] S. A. Ross, The Arbitrage Theory of Capital Asset Pricing, J. Econ. Theory **13**, 341-360 (1976).

[16] T. Conlon, H.J. Ruskin, and M. Crane, Cross-correlation dynamics in financial time series, Physica A **388**, 705-714 (2009).

[17] R. Allez and J.-P. Bouchaud, Eigenvector dynamics: General theory and some applications, Phys. Rev. E **86**, 046202 (2012).

[18] G. Buccheri, S. Marmi, and R. N. Mantegna, Evolution of correlation structure of industrial indices of U.S. equity markets, Phys. Rev. E **88**, 012806 (2013).

[19] M. Benzaquen, I. Mastromatteo, Z. Eisler, and J.-P. Bouchaud, Dissecting cross-impact on stock markets: an empirical analysis, J. Stat. Mech. 023406 (2017).

[20] A. S. Kyle, Continuous Auctions and Insider Trading, Econometrica **53**, 1315-1336 (1985).

[21] J.-P. Bouchaud, J. D Farmer, and F. Lillo, How markets slowly digest changes in supply and demand, in "Handbook of financial markets: dynamics and evolution", ed. by T. Hens and K. Reiner Schenk-Hoppe, (Elsevier, North-Holland, 2009), pp. 57-160.

[22] Z. Xiong and B. Zheng, On the eigenvalues of a specially rank-r updated complex matrix, Comput. Math. Appl. **57**, 1645-1650 (2009).

[23] L. Liu, A. Patton, and K. Sheppard, Does anything beat 5-minute RV? A comparison of realized measures across multiple asset classes, J. Econometr. **187**, 293-311 (2015).

[24] G. Golub, Some modified matrix eigenvalue problems, SIAM Rev. **15**, 318-334 (1973).

2. Fundamental Market Neutral Maximum Variance Portfolios

[1]

1. The results of this chapter were obtained in collaboration with Stanislav Kuperstein.

The Fundamental Market Neutral Maximum-Variance Portfolios

January 17, 2019

Abstract

We introduce *Maximum-Variance* portfolio that maximises the exposure to a given fundamental signal while remaining market neutral. Using real stock data we show that the Maximum-Variance portfolio weights are proportional to the stock rankings with respect to the signal, implying that the signal sensitivities are uniformly distributed among the stocks. Those signals are derived from financial factors, like *Book*, *Size*, *etc.* and the portfolio constructions are performed independently for each factor. We argue that this results in a large overlap between the subspaces spanned by the Maximum-Variance portfolios and the leading eigenvectors of the sample correlation matrix. Reducing the initial space allows us to reproduce the eigenvalues of the sample correlation matrix with remarkable accuracy. We thus can replicate any alternative beta strategy more efficiently than by using the mainstream 20% top-bottom approach. Moreover, our method permits to mimic the eigenvalue dynamics. The empirical analysis is carried out on the 500 largest U.S. stocks within different time scales and with 24 most popular factors, both fundamental and sectoral. Under certain hypotheses, the Maximum-Variance portfolio also optimises the Sharpe ratio, although our 18 years of data are insufficient for statistically significant backtesting results.

Keywords: Portfolio Management, Factor Investing, Alternative Risk Premia, Correlation.
JEL classification: C5, C61, G11, G12, G23, G4.

1 Introduction

This paper develops a theoretical portfolio, coined "Maximum-Variance" as it is obtained when optimising the variance correlated to a signal while minimising the specific risk, to provide a sound rule for determining optimal alternative beta portfolios. Alternative beta factors refer to the traditional Fama and French ([1], [2], [3]) setting where the long/short 20% stocks are ranked by their factor signal (i.e., size, book-to-market etc.). The fact that such long/short positions are not optimal is a crucial issue for investors. This issue is even no addressed by the classical "long-only" optimal management because it optimises the Sharpe ratio, which is not adopted at all for alternative beta. To that end, our solution, hereafter "Maximum-Variance", that intends to improve the 20% top/bottom approach, formulates this important problem in terms of maximising the risk exposure to the targeted factor, while minimising specific risk and other systematic risks.

The objective of the Maximum-Variance approach is to provide an optimal solution for the alternative beta portfolios. The risk-based "long-only" portfolios,[1] namely the "Minimum Variance", the "Risk Parity",

[1]By opposition to the (non-optimal) heuristic rule-based portfolios, Value-Weighted and the Equally-Weighted. A drawback is that it has no theoretical grounds in contrast to Mean-Variance theory. It is also the case for the most popular Equally Weighted heuristic portfolio, known as "$1/n$-portfolio", which assigns equal weights to all constituents (see e.g. [4] and [5]). It is relevant in the absence of any information on expected returns and on the covariance matrix. It is Mean-Variance optimal only if asset classes have the same expected returns and covariances [6]. Such an equal weighting scheme, expressed in terms of market value, tends to produce a systematically higher allocation to undervalued stocks at the expense of overvalued ones, which explains its high performance [7]. A drawback is that diversification is not optimal in terms of risks.

or the “Maximum Diversification” provided solutions to overcome the inefficiency of the market capitalisation weighted indices but cannot help for the alternative beta issues. The properties of the “Minimum Variance”, the “Risk Parity”, or the “Maximum Diversification” have been well documented in asset management literature (See e.g. [8], [9], [10], [11], [6]). Such risk budgeting allocation approaches are known to be robust because they do not require any return forecasts. It is the case for the well-known Minimum Variance portfolio [8] that has the lowest risk of all portfolios as it is located in the left-most part of the efficient frontier. It has the unique property that optimal security weights are solely dependent on security covariance matrix without regards to the expected returns; hence, it does not rely on any specific expected return estimate (see e.g. [12]), which makes it appear more robust than the mean-variance framework. Minimum-Variance strategies have gained popularity notably due to the empirical finding that low-volatility stocks tend to have returns that tend to exceed in average the market returns [13]. A drawback is that the portfolios concentration around is very sensitive to the covariance matrix noises, while it could be more equally distributed. This trouble is solved through the Risk Parity approach that induces a more conservative way of allocating assets according to their risk contribution to the portfolio. The Maximum Diversification relies on the concept of diversification ratio as the ratio of the weighted average of volatilities divided by the portfolio volatility [11]. The maximisation of the diversification ratio is equivalent to the minimisation of the variance in a universe where all stocks have the same expected volatility. In this case where all stocks have the same volatility, the Most Diversified portfolio becomes equal to the global Minimum Variance portfolio. The objective function is motivated by maximising the Sharpe ratio where expected asset returns are assumed to be proportional to asset volatility. All these portfolios have optimal risk-based weights equations determined with semi-closed-form analytical solutions, where Minimum-Variance weights are generally proportional to inverse variance, while Maximum Diversification and Risk Parity weights are generally proportional to inverse volatility [10]. It is also the case for the market neutral Maximum-Variance, we introduced in the paper. The difference is that the Maximum-Variance portfolio optimises under a market neutral constraint a different ratio than the Sharpe ratio that appears not adapted to the alternative beta issue. To do so, we introduced the concept of Factor “Correlation Level” (FCL) as the ratio between the variance of the portfolio and the variance of the portfolio as if the correlation between single stocks were zero. FCL can also be simply interpreted as the average correlation between stocks within a given portfolio, or as a weighted average of the eigenvalues of the correlation matrix with the weights given by the squares of the eigenvectors projections of the portfolio. In fact, the FCL ratio is closely related to the diversification ratio that can also be defined as the square root of the ratio between the variance of the portfolio as if correlations were one and the variance of the portfolio. We argue that the FCL plays an equivalent role of the Sharpe ratio in the context of alternative beta portfolios but could be measured more accurately (theoretical expected Sharpe ratio for a market neutral portfolio is null with the efficient market hypothesis, and empirically the expected returns are rarely significantly different from zero and remain very controversial in the literature [14]. We also show that the Maximum-Variance portfolio has a robust and theoretically optimal Sharpe ratio under some hypotheses that are different from the efficient market ones.

The Maximum-Variance portfolio is defined as the market neutral portfolio that optimises the FCL, when using a Two-Factor model for the returns. The Two-Factor model including the dominant factor and the style of interest is justified as the effect of additional orthogonal factors is small [15]. We argue that the parameters of the model could be fitted via a linear law depending on the ranking of the stocks according to the signal.

In fact, this weighting scheme has been already implemented in [16] for the value and momentum factors, as a means to reduce the influence of outliers (lowest and highest ranked stocks), but without further theoretical and empirical justification.

Remarkably, the law seems to be universal for any signal. To model the correlation matrix in that way introduces a constraint and avoid getting the sample eigenvectors as the optimal portfolios. The style of interest could be sectoral risk or include fundamental factors that have been mainly formed on firm characteristics such like capitalisation [17, 1, 2], book-to-market [18, 1, 2], low volatility/beta [19, 13] or momentum [20, 21], to quote only the most popular.

The authors of [22] study more than 330 return predictive signals that are mainly accounting based and show the large diversification benefits by suitably combining these signals. In[23] an out-of-sample approach is used to study the post-publication bias of discovered anomalies. The overall finding of this literature is that many discovered factors are likely false. In [24] more than 300 different factors were listed, claimed to capture an alternative risk premia but they argue that most claimed research findings are likely false. In [14] 14 major factors were selected and it was showed that the original market factor is by far the most important factor in explaining the cross-section of expected returns.

Market participants have extended this risk-based investment style with what is generally named the *alternative risk premia*, which corresponds to any risk premia that can be earned by building long-short portfolios exposed to common risk factors as opposed to the long-only exposure of the so-called "traditional risk premia" approach. Alternative risk premia can be aggregated to build multi-factor portfolios likely to extend the horizon of assets for a better diversification, while providing, not only a more economically meaningful investment opportunity, but also a more transparent systematic risk exposure to investors. Generally, risk premia factors constructions are not optimal as they have a top-down approach to gain access to the interest factor and they are beta market neutral. In the classical setting [1, 2, 3], factor construction is generally based on a long-short basket approach that is long/short the top 20% stocks ranked by the factor signal (i.e., size, book-to-market etc.). Other approaches from the industry, that are not necessarily disclosed, are often of a bottom-up type. Additionally, the methodology applies constraints on region-sector exposure, maximum constituent weights, liquidity and turnover and did little effort to discuss the problems of optimisation. The optimised risk premia factors, through the Maximum-Variance approach, constitute a set of portfolios that allows replicating easier any alternative beta strategies than a set of 20% top/bottom risk premia factors would do. Moreover, the conventional way to identify factors to model the correlation matrix is done through fundamental[2] multi-factor models.[3] These models specify that expected returns are linearly related to the weights of the common factors, but remained generally silent on the number of factors, which has induced some controversy[4], in the financial economics literature. We argue in our paper that using the optimised factors should improve the explanation power of the cross-section of single stock returns and that the FCL should be a criterion of the importance of a factor used to decide to select it or not into the model.

Furthermore, when using enough factors, the optimised factors could also be used as an efficient filter-

[2]Indeed, the multi-factor models of security market returns can be divided into three types (macroeconomic, fundamental and statistical), but the fundamental model remains more suited to ensure whether those factors are associated with risk premia. The fundamental model slightly outperforms the statistical model as it explains 42.6% of the total explanatory power, against 39% for statistical models and only 10.9% for macroeconomic model [25].

[3]Recall that the multi-factor model, initially formulated in [26], through the Arbitrage Pricing Theory (APT) offers a testable alternative to the Capital Asset Pricing Model (CAPM) introduced initially in [27]. The two major differences between the APT and the CAPM model is that the APT allows for a more than just one generating factor and demonstrates that every equilibrium will be characterized by a linear relationship between each asset's expected return and its return's loading on the common factors [28]. Notice that while [28] suggests that APT results are invariant to rotation of the original factors, [29] brings a nuance by stating that the statistical tests for the number of priced factors are not invariant to rotation.

[4]Some early influential papers were written on that particular topic in the financial economic literature. [30] explain that as one increases the number of securities, the number of factors increases. [31] demonstrate that returns can be linearly related to n factors if n eigenvalues of the covariance matrix of returns become large as the number of securities increases. [32] finds evidence that one eigenvalue at most dominates the covariance matrix indicating that a one-factor model may describe security pricing.[33] obtain that the number of factors common across securities was limited to three or five.[34] find the existence of only one dominant factor and suggest the effect of adding orthogonal factors is small. [35] shows that one market factor explains the major part of security returns. [36] explain if n is the correct number of pervasive factors, then there should be no significant decrease in the cross-sectional mean square of idiosyncratic returns in moving from n to $n+1$ factors. A wide collection of influential papers has also addressed this issue in the econo-physics literature to find roughly comparable conclusions around the idea of one large eigenvalue (see e.g.,[37],[38],[39])

ing method of the correlation matrix that allows obtaining constrained eigenvalues as close as possible to eigenvalues of the correlation matrix. We highlighted two interesting applications. First, we argue that the optimised factor enables to replicate better any strategy, which is the purpose of alternative beta products. Second, we argue that filtering the correlation matrix is a key issue in asset management and portfolio optimisation in general, because sample correlation matrix may become too noisy and the optimisation could be over-fitted by noises and fallacious hedges. This can be related to a well-known problem in portfolio optimisation typically called "error maximisation" by [40] who explains that optimisation scheme tends to maximise the effects of errors in the input assumptions as these inputs are not without statistical error.

Our filtering method could help to understand better the dynamics of the correlation matrix and their eigenvectors that is not well identified [41]. So far, it seems to be an unequivocal empirical support that very few eigenvalues dominate the correlation matrix but the interpretation of the corresponding eigenvectors remain difficult except for the market mode. We argue that the constrained[5] eigenvectors into the subspace of the optimised factors could be a very good approximation of the true eigenvectors. The main constrained eigenvectors are combinations of Maximum-Variance portfolios which have the highest FCL. As FCL can change in a brutal way depending on market issues, the eigenvectors can also change brutally that explains the difficulty to interpret directly different eigenvectors. The Maximum-Variance finally enables to filter the correlation matrix and relies on a diagonalisation method that enables to obtain orthogonal factors as close as possible to eigenvectors of the correlation matrix. It corresponds to a dimension reduction introduced by economic constraints to filter noises. Roughly speaking, those economic constraints applied to these factors act like "constrained eigenvectors" and constitute a kind of filter of the correlation matrix that consists in a reduction of the traded universe dimension from several hundred single stocks or more to few tens factors. The filter introduced by economical constraints helps to interpret the first constrained eigenvectors and also allows one to capture small eigenvalues that are typically hidden by noises.[6] Maximum-Variance enables to show that small eigenvalues are mainly coming from combination of styles like capitalisation and book. The filter is complementary to the standard approach based on the Random Matrix theory, which makes statements about the density of the eigenvalues of large random matrices. The empirical correlation matrix computed on a given realisation must be distinguished from the true correlation matrix of the underlying statistical process as the large number of simultaneous noisy variables creates important systematic errors in the computation of the matrix eigenvalues [46].

To our knowledge, there has so far been just no published study on how to build optimal portfolios for alternative beta portfolios despite the current major issues at stake with the emergence of the alternative risk premia vehicles as the low cost and efficient solution of Hedge Fund that are made available to any investor to manage assets. Building upon the body of evidence discussed above, this paper provides a novel methodology with a closed-form solution to compute optimal risk premia portfolios. To test the Maximum-Variance methodology, we use stock returns at multiple time scales (from 5 minutes to 100 days) from a panel of the largest 500 U.S. stock since 2000 to compute the different Maximum-Variance portfolios and to test the improvement brought by the optimisation in the capture of the eigenvalues and their dynamics. This shows that optimising factors helps to increase the explanation power of the cross-section of single stocks returns and to replicate better any alternative beta strategy. We also test improvement of the Sharpe ratio but, as expected, measurements are not significant. We have restricted ourselves to the 24 most popular factors according to the literature (market mode, dividend yield, capitalisation, volume/capitalisation, STR,

[5][42] already introduced in 1971 the concept of constrained eigenvalues and eigenvectors into a subspace. They correspond to the eigenvalues and eigenvectors of a transformed matrix projected into the subspace

[6]The need to 'clean' the empirical correlation matrix requires a device for distinguishing signal from noise [43]; it requires distinguishing meaningful eigenvalues (beyond the edge) from noisy ones (inside the bulk) given that all eigenvalues in the bulk of the Marčenko-Pastur spectrum are deemed as noise; to that end, [44] define a threshold that separates only those eigenvalues that are outside the noise band. The standard approach to clean up the empirical correlation matrix requires separating the largest eigenvalue, economically interpreted as the "market mode", from the bulk where all other eigenvalues reside and are buried under the noise [45].

momentum, beta, leverage, sales to price, book to price, cash to price, price to earning, growth of earning, sensitivity to Euro dollar, sensitivity to 10 years rates, energy, finance, IT, utilities, consumer, industry, pharmacy, consumer discretionary vs. staple, REITs). This selection is similar to the one of [14].

The paper is organized as follows. Section 2 develops the theoretical framework. Section 3 discusses the practical implementation of the Maximum-Variance portfolio. Section 4 summarizes the empirical tests and compares performances and the ability to capture eigenvalues with the classical 20% top-bottom approach. Sections 5 describes remaining open problems around the correlation matrix that the Maximum-Variance approach should help to answer.

2 The Maximum-Variance Factor

The main objective of this paper is to show how one can "clean" a correlation matrix by using supplementary information regarding the time-series at hand. Two series having similar *characteristics* according to the new input information should have close correlations with the other time-series. Our method allows to reduce the initial N-dimensional space to a smaller K-dimensional subspace, where N and K are the number of the time-series and the characteristics respectively. The method consists of two steps. First, we find K one-dimensional subspaces for each characteristic independently. Second, we determine the optimal eigenvectors in the K-dimensional subspace, which is the sum of the smaller K 1-dimensional subspaces.

We then use our method to noise-filter a correlation matrix of single stocks returns. In this application the first eigenvalue is much larger than the others, and plenty of financial information data is available (*Book*, *Capitalisation*, etc.).

On three occasions in this section we used the same strategy that consists to reduce the number of the optimization parameters (the space dimension) either from N (the number of single stocks) to $N-1$ (the dimension of the subspace of returns orthogonal to the stock index that is highly correlated to the first eigenvector), either from N to K (the number of characteristics) or from N to Q (the number of quantiles to group stocks with similar characteristics). Our approach relies on the fact that constrained eigenvectors, namely those forced to belong to a given subspace, are also the eigenvectors of the matrix reduced to this subspace. If the choice of the reduced space (for example, the K Maximum-Variance or the Q quantile portfolios) has a sufficient overlap with the space spanned by the leading eigenvectors then the constrained eigenvalues will be close to the unconstrained eigenvalues of the correlation matrix.

In Section 2.1 we define the covariance and the overlap matrices of N elementary single stocks. The covariance and overlap matrices between any K or Q portfolios can be generalized. We define then the Factor Correlation Level (FCL) that measures the variance of a portfolio when it is normalized by the overlap matrices.

In Section 2.2 we introduce the Two-Factor model which can be implemented independently for each financial characteristic. It helps us to generate one solution per characteristic, that is implemented independently from other characteristics, that optimizes the FCL and that is correlated to the characteristic of interest.

In Section 2.3 we define the Fundamental Maximum-Variance portfolio as such a solution and we derive an explicit formula from the Two-Factor Model parameters. In Section 2.4 we demonstrate that the Maximum-Variance portfolios are the portfolios replicating as best as possible the entire correlation matrix given a selection of characteristics. In Section 2.5 we describe a methodology to measure the parameters of the Two-Factor model, the factor loadings.

In Section 2.6 we comment on on empirical universal law that we found for the optimal factor loadings and in Section 2.7 we finally compare our methodology with the usual ordinary least square (OLS) and the Fama-MacBeth regressions. In Section 2.8 we provide some plausible theoretical patterns of the correlation matrix of single stocks returns that explain why the overlap between the K Maximum-Variance and the space spanned by the leading eigenvectors is higher than we could have expected if the eigenvectors were randomly generated.

2.1 Notations and Definitions

Let us first introduce notations that we will use throughout the paper. $r_i(t)$ are the time series of N single stock returns with $i = 1, \ldots, N$. For a given functional $\varphi(r_i(t))$ we will denote by $\mathbb{E}_{t-1}(\varphi(r_i(t)))$ the *conditional* expectation of $\varphi(r_i(t))$ based on information available from the entire previous period. For instance, we will assume that $\mathbb{E}_{t-1}(r_i(t)) = 0$, unless otherwise mentioned. With this assumption $\boldsymbol{\Sigma}$ is the volatility vector[7] defined by $\Sigma_i(t) = \sqrt{\mathbb{E}_{t-1}(r_i^2(t))}$. With these conventions the conditional covariance and correlation matrices of returns are given by:

$$H_{ij}(t) = \mathbb{E}_{t-1}(r_i(t) r_j(t)) \qquad \text{and} \qquad C_{ij}(t) = \frac{H_{ij}(t)}{\Sigma_i(t)\Sigma_j(t)} \,. \tag{1}$$

Notice that $\Sigma_i = \sqrt{H_{ii}}$ and so $C_{ii}(t) = 1$ for any t as it should be for a correlation matrix. For the upcoming definitions we will also define $\boldsymbol{\Gamma}$ as the covariance matrix of positions. In the special case where the constituents of the base do not have any common position as it is for the elementary single stocks, the matrix $\boldsymbol{\Gamma}$ will correspond to the diagonal matrix of variances $\Gamma_{ij} \equiv \Sigma_i^2 \delta_{ij}$, from which

$$\boldsymbol{C} = \boldsymbol{\Gamma}^{-\frac{1}{2}} \boldsymbol{H} \boldsymbol{\Gamma}^{-\frac{1}{2}} \,. \tag{2}$$

Since the correlation matrix is less biased towards high-volatility stocks and therefore is more adapted for the application considered in this paper than the covariance matrix, the $\boldsymbol{\Gamma}^{-1}$ matrix will be often used to calculate vector products. For example, we will refer to two vectors $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$ as $\boldsymbol{\Gamma}^{-1}$-orthogonal if they satisfy $\boldsymbol{u}_1^{\mathrm{T}} \boldsymbol{\Gamma}^{-1} \boldsymbol{u}_2 = 0$. In financial terms it reduces the weights of companies with large volatilities. As volatilities change with time, thus generating heteroscedasticity, it makes another reason to prefer the correlation over the covariance matrix.

A generic portfolio $\mathbb{p}$ is determined by N (time-dependent) weights $\omega_i^{\mathbb{p}}(t)$ and its return is then given by $r^{\mathbb{p}}(t) = \sum_i \omega_i^{\mathbb{p}}(t) r_i(t)$. Among all possible $\boldsymbol{\omega}^{\mathbb{p}}$'s the *market-mode* portfolio plays a special role. We will denote it by $\boldsymbol{\omega}^{\mathbf{m}}(t)$ and the corresponding return by $r^{\mathbf{m}}(t)$. In the paper, $r^{\mathbf{m}}$ is taken to be the *stock index* return, which is very close to the value-weighted portfolio return. In the latter the weight of each stock is proportional to its market capitalisation, which we will denote by $\mathrm{Cap}_i(t)$. The market-mode portfolio weights are close, therefore, to the principal component of the matrix $\sqrt{\mathrm{Cap}_i} H_{ij} \sqrt{\mathrm{Cap}_j}$ (with neither i nor j summations). $r^{\mathbf{m}}(t)$ provides a good proxy of the accumulated gain or loss for all investors. It thus corresponds to the real systematic risk that all investors want to avoid for alternative beta vehicles in order to diversify and optimise their investments. The value weighted portfolio differs slightly from the principal component of $\mathbf{C}$. The two portfolios produce highly correlated returns but the value-weighted portfolio is invested primarily in companies with large capitalisation whereas the principal component of $\mathbf{C}$ is invested mainly in the small firms as they are better represented in portfolios with large N.

In our conventions a given stock beta describes the conditional sensitivity of the stock return to the stock index:

$$\beta_i(t) \equiv \frac{\mathbb{E}_{t-1}(r_i(t) r^{\mathbf{m}}(t))}{\mathbb{E}_{t-1}\left((r^{\mathbf{m}}(t))^2\right)} \qquad \text{or} \qquad \boldsymbol{\beta} \equiv (\Sigma^{\mathbf{m}})^{-2} \cdot \boldsymbol{H}\boldsymbol{\omega}^{\mathbf{m}} \tag{3}$$

[7] With a few obvious exceptions we will use the bold font for matrices and vectors with no explicit indices.

in the matrix notations. We will also need an additional time-series of returns derived from $\boldsymbol{\beta}$:

$$r^{\mathbf{m}}_{\star}(t) = \frac{\sum\limits_{j=1}^{N} r_j \beta_j \Sigma_j^{-2}}{\sum\limits_{k=1}^{N} \left(\beta_k \Sigma_k^{-1}\right)^2}\,, \tag{4}$$

where the time-dependence on the right-hand side is, again, implicit. For reasons to be clarified below, we will refer to $r^{\mathbf{m}}_{\star}(t)$ as the return of *Maximum-Variance market-mode* portfolio. This portfolio is optimal in a sense to be clarified shortly. The relation between $r^{\mathbf{m}}_{\star}(t)$ and $r^{\mathbf{m}}(t)$ is straightforward. Starting from the stock index return $r^{\mathbf{m}}(t)$ one can compute the betas as in (3). On the other hand, knowing the full set of $\beta_{i=1,\dots,N}(t)$ and the stock returns $r_i(t)$, the index return might be *reproduced* (or, to be more precise, approximated) by $r^{\mathbf{m}}_{\star}(t)$. Importantly, if beta's time-dependence is negligible, we may interpret both (3) and (4) as outputs of the weighted least-squares (WLS) and the ordinary least-squares (OLS) regressions respectively, both based on $r_i(t) = \beta_i r^{\mathbf{m}}(t)$ relation. In Appendix A we show that for a fixed time-period T computing iteratively the (constant) betas from the market mode and the market return from the betas leads (after a sufficiently large number of iterations) to fixed-point values, where the identity $r^{\mathbf{m}}_{\star}(t) = r^{\mathbf{m}}(t)$ holds and the product $b_i r^{\mathbf{m}}(t)$ is merely the principal component (PC) of the correlation matrix.

It will be later on useful to introduce two additional notations: the ρ_H-correlation of two portfolios $\boldsymbol{\omega}^{\mathbb{p}_1}$ and $\boldsymbol{\omega}^{\mathbb{p}_2}$, which denotes the conditional correlation between the corresponding returns, $r^{\mathbb{p}_{1,2}}(t) = \sum_i \omega_i^{\mathbb{p}_{1,2}}(t) r_i(t)$, and the similarly defined *position overlap* ρ_Γ:

$$\begin{aligned}
\rho_H\left(\boldsymbol{\omega}^{\mathbf{p}_1}, \boldsymbol{\omega}^{\mathbf{p}_2}\right) \equiv & \frac{\boldsymbol{\omega}^{\mathbb{p}_1\,\mathrm{T}} \boldsymbol{H} \boldsymbol{\omega}^{\mathbb{p}_2}}{\left(\boldsymbol{\omega}^{\mathbb{p}_1\,\mathrm{T}} \boldsymbol{H} \boldsymbol{\omega}^{\mathbb{p}_1}\right)^{\frac{1}{2}} \left(\boldsymbol{\omega}^{\mathbb{p}_2\,\mathrm{T}} \boldsymbol{H} \boldsymbol{\omega}^{\mathbb{p}_2}\right)^{\frac{1}{2}}} \\
\rho_\Gamma\left(\boldsymbol{\omega}^{\mathbf{p}_1}, \boldsymbol{\omega}^{\mathbf{p}_2}\right) \equiv & \frac{\boldsymbol{\omega}^{\mathbb{p}_1\,\mathrm{T}} \boldsymbol{\Gamma} \boldsymbol{\omega}^{\mathbb{p}_2}}{\left(\boldsymbol{\omega}^{\mathbb{p}_1\,\mathrm{T}} \boldsymbol{\Gamma} \boldsymbol{\omega}^{\mathbb{p}_1}\right)^{\frac{1}{2}} \left(\boldsymbol{\omega}^{\mathbb{p}_2\,\mathrm{T}} \boldsymbol{\Gamma} \boldsymbol{\omega}^{\mathbb{p}_2}\right)^{\frac{1}{2}}}
\end{aligned} \tag{5}$$

A portfolio $\mathbb{p}$ is market-neutral if its return is uncorrelated with $r^{\mathbf{m}}(t)$, that is $\mathbb{E}_{t-1}\left(r^{\mathbb{p}}(t) r^{\mathbf{m}}(t)\right) = 0$. In terms of the weights it amounts to $\rho_H\left(\boldsymbol{\omega}^{\mathbf{m}}(t), \boldsymbol{\omega}^{\mathbb{p}}(t)\right) = 0$. Using (3) this is also equivalent to

$$\boldsymbol{\beta}^{\mathrm{T}}(t) \boldsymbol{\omega}^{\mathbb{p}}(t) = 0\,. \tag{6}$$

Employing these notations, we define the *factor correlation level* (FCL) mentioned earlier in Introduction:

$$\lambda^{(0)}\left(\boldsymbol{\omega}^{\mathbb{p}}\right) \equiv \frac{\left(\boldsymbol{\omega}^{\mathbb{p}}\right)^{\mathrm{T}} \boldsymbol{H} \boldsymbol{\omega}^{\mathbb{p}}}{\left(\boldsymbol{\omega}^{\mathbb{p}}\right)^{\mathrm{T}} \boldsymbol{\Gamma} \boldsymbol{\omega}^{\mathbb{p}}}\,, \tag{7}$$

where for the sake of shortness we omitted the time-dependence. It follows from (2) that, once $\sqrt{\boldsymbol{\Gamma}} \boldsymbol{\omega}^{\mathbb{p}}$ is chosen to be an eigenvector of the correlation matrix (2), the FCL is equal to the corresponding eigenvalue. Otherwise it could be interpreted as the weighted average of the eigenvalues of $\boldsymbol{C}$ with the weights given by the squares of the eigenvectors projections on $\boldsymbol{\omega}^{\mathbb{p}}$. Indeed, if $\boldsymbol{e}_{i=1,\dots,N}$ are the eigenvectors of the correlation matrix in (2) and $\ell_{i=1,\dots,N}$ are the corresponding eigenvalues, then $\boldsymbol{C} = \sum_{i=1}^{N} \ell_i \boldsymbol{e}_i \boldsymbol{e}_i^{\mathrm{T}} / \left(\boldsymbol{e}_i^{\mathrm{T}} \boldsymbol{e}_i\right)$ and

$$\lambda^{(0)}\left(\boldsymbol{\omega}^{\mathbb{p}}\right) = \sum_{i=1}^{N} \left(a_i^{\mathbb{p}}\right)^2 \ell_i\,, \qquad \text{where} \qquad a_i^{\mathbb{p}} \equiv \rho_\Gamma\left(\boldsymbol{\Gamma}^{-\frac{1}{2}} \boldsymbol{e}_i, \boldsymbol{\omega}^{\mathbb{p}}\right) \quad \text{with} \quad \sum_{i=1}^{N} \left(a_i^{\mathbb{p}}\right)^2 = 1\,. \tag{8}$$

As we explained above, the main goal of this paper is to optimise the FCL (7) in a reduced subspace that contains the leading eigenvectors. To identify the subspace we use the financial information coming in form of signals as we review in the next section.

2.2 The Two-Factor Model

As it was briefly outlined in Introduction we want to optimise the FCL in (7) separately for each of the K factors/styles. To estimate $\lambda^{(0)}$ we will *approximate* the returns using a Two-Factor model pertinent for a given factor as will be described below.

Each of these K models employs the market mode $r^{\mathbf{m}}(t)$ and an additional factor mode $r^{\mathbf{f}}_a(t)$ for $a = 1, \ldots, K$ that captures the relevant style information encoded in a time-dependent signal $s_{a,i}(t)$. These K different signals correspond to *book*, *capitalisation* and other financial criteria we list in Table 14 of Appendix B. The signals determine (a priori time-dependent) rankings of the stocks. We will denote the rankings by $q_{a,i}(t)$ and we will drop the a-index all the way up to Section 2.4 focusing meanwhile on a single factor/style. The rankings take values among $1, \ldots, N$ and by definition $q_i(t) < q_j(t)$ if and only if $s_i(t) < s_j(t)$.[8] For example, if $N = 3$ and we have $(s_1(t), s_2(t), s_3(t)) = (5., 9., -3.)$ then $(q_1(t), q_2(t), q_3(t)) = (2, 3, 1)$.

The usual practice is to model $r^{\mathbf{f}}(t)$ by the so-called *benchmark* portfolio return, $r^{\mathbf{b.m.}}(t)$. The latter is built by buying the top 20% of the stocks and shorting the bottom 20%, while sizing the two legs to keep the portfolio beta market-neutral. The weights of the benchmark portfolio may appear either in the *equal-weighted* or the *equal-risk-weighted* version:

$$\omega_i^{\mathbf{b.m.}}(t) = \left(\begin{array}{c|c|c} \begin{array}{c}\text{Equal}\\ \text{weighted}\end{array} & \begin{array}{c}\text{Equal-risk}\\ \text{weighted}\end{array} & q_i(t) \\ \hline \omega_+(t) & \omega_+(t)/\Sigma_i(t) & q_i(t) \geqslant 0.8 \\ 0 & 0 & 0.2 < q_i(t) < 0.8 \\ \omega_-(t) & \omega_-(t)/\Sigma_i(t) & q_i(t) \leqslant 0.2 \end{array} \right) \qquad r^{\mathbf{b.m.}}(t) = \sum_{i=1}^{N} \omega_i^{\mathbf{b.m.}}(t) r_i(t)\,. \tag{9}$$

Here $\omega_+(t)$ and $\omega_-(t)$ are fixed by the benchmark portfolio market-neutrality condition, see (6):

$$\sum_{i=1}^{N} \beta_i(t) \omega_i^{\mathbf{b.m.}}(t) = 0\,, \tag{10}$$

and by the over-all normalisation $\max(\omega_+(t), \omega_-(t)) = 1/N$. Notice that the solutions for $\omega_+(t)$ and $\omega_-(t)$ depend on the choice of the version in (9).

One may extend the benchmark portfolio construction to include as well stocks which are not part of our original selection of N stocks. If the number of stocks is sufficiently large the specific risk will vanish and the new benchmark portfolio, $\boldsymbol{\omega}^{\mathbf{f}}(t)$, will be an exogenous factor. We will denote by $r^{\mathbf{f}}(t)$ the return of $\boldsymbol{\omega}^{\mathbf{f}}(t)$. The Two-Factor model then relates the residual returns $r_i - \beta_i r^{\mathbf{m}}_\star$ and the factor return $r^{\mathbf{f}}(t)$ as follows:[9]

$$r_i(t) - \beta_i(t) r^{\mathbf{m}}_\star(t) = b_i r^{\mathbf{f}}(t) + \epsilon_i(t) \qquad \text{for} \qquad i = 1, \ldots, N\,, \tag{11}$$

where the last term stands for the *idiosyncratic* returns. We will refer to b_i as *factor loadings*. The model has to satisfy two *orthogonality* conditions:

- All of $\epsilon_i(t)$'s are uncorrelated with the factor return, $r^{\mathbf{f}}(t)$:

$$\mathbb{E}_{t-1}\left(r^{\mathbf{f}}(t)\epsilon_i(t)\right) = 0\,. \tag{12}$$

- The vector $\boldsymbol{\Gamma}^{-\frac{1}{2}}\boldsymbol{b}$ is an eigenvector of $\boldsymbol{\Gamma}^{-\frac{1}{2}}\boldsymbol{H}^{\epsilon}\boldsymbol{\Gamma}^{-\frac{1}{2}}$, where $\boldsymbol{H}^{\epsilon}(t) = \mathbb{E}_{t-1}\left(\epsilon_i(t)\epsilon_j(t)\right)$ is the covariance matrix of the idiosyncratic returns:

$$\boldsymbol{H}^{\epsilon}(t)\boldsymbol{\Gamma}^{-1}(t)\boldsymbol{b} = \ell^{\epsilon,\boldsymbol{b}}(t) \cdot \boldsymbol{b}\,, \tag{13}$$

[8] In this section we assume that the signals do not coincide.

[9] Notice that we have $r^{\mathbf{m}}_\star$ rather than $r^{\mathbf{m}}$ in this formula.

while for any t the eigenvalue $\ell^{\epsilon,\boldsymbol{b}}$ remains much smaller than the variance of the factor return:

$$\ell^{\epsilon,\boldsymbol{b}} \ll \left(\Sigma^{\mathbf{f}}\right)^2 \qquad \text{where} \qquad \left(\Sigma^{\mathbf{f}}\right)^2 \equiv \mathbb{E}_{t-1}\left(\left(r^{\mathbf{f}}(t)\right)^2\right) \tag{14}$$

We will clarify later on in this section what do we mean by sufficiently "small".

Importantly, both requirements above are motivated by the principal component analysis of the correlation matrix. If the $\beta_i r^{\mathbf{m}}_\star(t)$ and the $b_i r^{\mathbf{f}}$ terms were respectively the leading and the sub-leading terms in the PCA expansion of the correlation matrix $\boldsymbol{C}$, the orthogonality of the modes (12) would not be an additional constraint but rather a direct consequence of the expansion. Moreover, it would as well enforce $\ell^{\epsilon,\boldsymbol{b}} = 0$. We show it in Appendix A. In our case neither the betas nor the market return are principal components of $\boldsymbol{C}$, but we may adopt the orthogonality property as a good approximation for our model (see the previous subsection).

Before closing the section we should stress that (11) is conceptually different from the multivariate factor model well explored in the literature. In this model there would be a single equation combining the market return, K additional factors returns and one series of the idiosyncratic returns. In our approach there are instead K equations for each factor with a different $\epsilon_i(t)$ term. This will allow us to optimise FCL separately within each one of the K Two-Factor models, thus reproducing eventually the largest eigenvalues of the correlation matrix. We come back to this issue in Appendix H.

2.3 The Maximum-Variance Market-Neutral Portfolio

Given a fundamental signal, a corresponding Two-Factor model (11) generates the matrices $\boldsymbol{H}(t)$, $\boldsymbol{\Gamma}(t)$ and $\boldsymbol{C}(t)$ in terms of $r^{\mathbf{m}}(t)$, $\beta_i(t)$, $r^{\mathbf{f}}(t)$, b_i and $\epsilon_i(t)$. Modelling $\boldsymbol{H}(t)$ and $\boldsymbol{\Gamma}(t)$ by means of the Two-Factor model (11) helps to isolate among the $N-1$ market-neutral portfolios whose weights $\omega_i(t)$ optimise (locally) the FCL (7), a single optimal portfolio that is highly correlated with the factor return $r^{\mathbf{f}}$. We call this portfolio/vector the *maximum-variance market-neutral factor* and denote it by $\boldsymbol{\omega}^{(0)}_\star$ and the optimal of FCL value by $\lambda^{(0)}_\star$ (as was already mentioned earlier we omit the factor index a).[10] To recapitulate:

$$\lambda^{(0)}_\star = \lambda^{(0)}\left(\boldsymbol{\omega}^{(0)}_\star\right) \qquad \text{with} \qquad \left.\frac{\delta\lambda^{(0)}(\boldsymbol{\omega})}{\delta\boldsymbol{\omega}}\right|_{\boldsymbol{\omega}=\boldsymbol{\omega}^{(0)}_\star} = 0 \qquad \text{and} \qquad \boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{\omega}^{(0)}_\star = 0\,, \tag{15}$$

and for any other market-neutral portfolio $\boldsymbol{\omega}'$ satisfying the optimisation condition in (15), yet different from $\boldsymbol{\omega}^{(0)}_\star$, one necessarily has[11]

$$\rho_H\left(\boldsymbol{\omega}^{\mathbf{f}},\boldsymbol{\omega}'\right) < \rho_H\left(\boldsymbol{\omega}^{\mathbf{f}},\boldsymbol{\omega}^{(0)}_\star\right) \tag{16}$$

with the short-cut notations (5). Again, we dropped above the explicit time-dependence on order to keep the formulae short and readable. The optimisation, in fact, takes place for any t and thus $\lambda^{(0)}_\star = \lambda^{(0)}_\star(t)$ is also time-dependent.

The weights $\boldsymbol{\omega}^{(0)}_\star$ are called the *maximum-variance market-neutral* portfolio with the first part of the name coming from the fact that it maximises (optimises) the variance (the numerator in the FCL definition (7)), while maintaining market-neutrality and the correlation with the given signal. Importantly, the conditions

[10] Here the $\star$ alludes to the FCL optimisation and the $^{(0)}$ superscript will be clarified in subsection 2.4.

[11] In fact the definition can be presented mathematically as

$$\boldsymbol{\omega}^{(0)}_\star \equiv \underset{\substack{\text{opt} \quad (\lambda^{(0)}(\boldsymbol{\omega})) \\ \text{s.t.}\boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{\omega}=0}}{\operatorname{argmax}} \left(\rho_H\left(\boldsymbol{\omega}^{\mathbf{f}},\boldsymbol{\omega}\right)\right)$$

but this definition is too cumbersome to be exploited.

(15) and (16) fix the weights only up to an overall rescaling, $\omega_i \to \text{const} \cdot \omega_i$. The ambiguity might be easily eliminated, for instance, by setting to one the $\boldsymbol{\omega}_\star^{(0)}$-portfolio variance.

It is worth emphasising here that $\boldsymbol{\omega}_\star^{(0)}$ does not necessarily provide a maximum of the FCL. In Appendix C we argue that the signature of the Hessian matrix obtained from an FCL-like Lagrangian at one of its optimal points has both positive and negative directions, thus describing a saddle point, rather than a (local) maximum, which, in turn, appears only if we pick up the highest eigenvalue as the optimisation solution.

Despite the somewhat verbose definition of $\lambda_\star^{(0)}$ it apparently has a simple straightforward interpretation. First, notice that the *unconstrained* optimisation of (7) is equivalent to solving for eigenvalues of the matrix $\boldsymbol{\Gamma}^{-\frac{1}{2}}\boldsymbol{H}\boldsymbol{\Gamma}^{-\frac{1}{2}}$. To see this, one can simply redefine vector $\boldsymbol{\omega}$ by $\boldsymbol{\omega} \to \boldsymbol{\Gamma}^{\frac{1}{2}}\boldsymbol{\omega}$. Second, as we explain in details in Appendix D, finding *constrained* eigenvectors $\boldsymbol{v}_i$ of a square matrix $\boldsymbol{M}$ subject to an additional requirement $\boldsymbol{c}^{\mathrm{T}}\boldsymbol{v}_i = 0$ is equivalent to the unconstrained diagonalisation of $\boldsymbol{P^c}\boldsymbol{M}\boldsymbol{P^c}$, where $\boldsymbol{P^c}$ is a projection operator defined by $\boldsymbol{P^c}\boldsymbol{c} = 0$ and $(\boldsymbol{P^c})^2 = \boldsymbol{P^c}$. To be more specific, if $l_i \neq 0$ and $\boldsymbol{u}_i$ are an eigenvalue/eigenvector pair of $\boldsymbol{P^c}\boldsymbol{M}\boldsymbol{P^c}$, then $\boldsymbol{v}_i = \boldsymbol{P^c}\boldsymbol{u}_i$ is a *constrained* eigenvector of $\boldsymbol{M}$.

In our case $\boldsymbol{M} = \boldsymbol{\Gamma}^{-\frac{1}{2}}\boldsymbol{H}\boldsymbol{\Gamma}^{-\frac{1}{2}}$ and the constraint vector is $\boldsymbol{c} = \boldsymbol{\Gamma}^{-\frac{1}{2}}\boldsymbol{\beta}$, where the $\boldsymbol{\Gamma}$-factor comes from the aforementioned redefinition of $\boldsymbol{\omega}$.

To summarize, the maximum-variance market-neutral portfolio $\boldsymbol{\omega}_\star^{(0)}$ can be found following these four steps:

1. To find all eigenvectors with non-vanishing eigenvalues of the $\boldsymbol{P^c}$-projection of the correlation matrix $\boldsymbol{P^c}\boldsymbol{\Gamma}^{-\frac{1}{2}}\boldsymbol{H}\boldsymbol{\Gamma}^{-\frac{1}{2}}\boldsymbol{P^c}$, where $\boldsymbol{c} = \boldsymbol{\Gamma}^{-\frac{1}{2}}\boldsymbol{\beta}$ and thus

$$(\boldsymbol{P^c})_{ij} = \left(\delta_{ij} - \frac{\left(\beta_i \Sigma_i^{-1}\right)\left(\beta_j \Sigma_j^{-1}\right)}{\sum\limits_{l=1}^{N} \left(\beta_l \Sigma_l^{-1}\right)^2} \right) . \tag{17}$$

   Notice that the maximum number of such vectors is $(N-1)$.

2. To calculate the $\boldsymbol{P^c}$-projection of each of these eigenvectors with $\boldsymbol{P^c}$ given in (17).

3. To multiply the projected eigenvectors by $\boldsymbol{\Gamma}^{-\frac{1}{2}}$.

4. To arrive at $\boldsymbol{\omega}_\star^{(0)}$, the vector/portfolio producing the strongest correlation to the factor return, see (16), should be selected among the available vectors.

Let us now rewrite (17) in terms of the returns defined by the Two-Factor model (11). Upon (11) the matrix $\boldsymbol{P^c}\boldsymbol{\Gamma}^{-\frac{1}{2}}\boldsymbol{H}\boldsymbol{\Gamma}^{-\frac{1}{2}}\boldsymbol{P^c}$ becomes:

$$\left(\boldsymbol{P^c}\boldsymbol{\Gamma}^{-\frac{1}{2}}\boldsymbol{H}\boldsymbol{\Gamma}^{-\frac{1}{2}}\boldsymbol{P^c}\right)_{ij} = \Sigma_i^{-1}\Sigma_j^{-1}\mathbb{E}_{t-1}\left(\left(b_i r^{\mathbf{f}} + \epsilon_i\right)\left(b_j r^{\mathbf{f}} + \epsilon_j\right)\right) = \left(\boldsymbol{\Gamma}^{-\frac{1}{2}}\left(\left(\Sigma^{\mathbf{f}}\right)^2 \cdot \boldsymbol{b}\boldsymbol{b}^{\mathrm{T}} + \boldsymbol{H}^{\epsilon}\right)\boldsymbol{\Gamma}^{-\frac{1}{2}}\right)_{ij} . \tag{18}$$

Here $\left(\Sigma^{\mathbf{f}}\right)^2 = \mathbb{E}_{t-1}\left(\left(r^{\mathbf{f}}(t)\right)^2\right)$ is the conditional variance of the factor return. In deriving this result we used the first property of the Two-Factor model (12) and the relation

$$\sum_j \boldsymbol{P}^{\boldsymbol{c}}_{ij}\left(\frac{r_j}{\Sigma_j}\right) = \frac{1}{\Sigma_i}\left(r_i - \beta_i r_\star^{\mathbf{m}}\right) , \tag{19}$$

which follows directly from (17), as well as the definitions (3) and (4).

Notice now that by virtue of the last property of the Two-Factor model, see (13), the vector $\boldsymbol{\Gamma}^{-\frac{1}{2}}\boldsymbol{b}$ is an eigenvector of the matrix $\boldsymbol{P^c}\boldsymbol{\Gamma}^{-\frac{1}{2}}\boldsymbol{H}\boldsymbol{\Gamma}^{-\frac{1}{2}}\boldsymbol{P^c}$. Indeed, this is an eigenvector of all terms on the right-hand side of (18), so it should also be an eigenvector of their sum.

Finally, $\boldsymbol{\Gamma}^{-\frac{1}{2}}$ times the $\boldsymbol{P}^{\mathbf{c}}$-projection of the vector $\boldsymbol{\Gamma}^{-\frac{1}{2}}\boldsymbol{b}$ is equal to:[12]

$$\left(\omega_\star^{(0)}\right)_i(t) \sim \Sigma_i^{-2}(t)\left(b_i - \frac{\sum_{j=1}^N b_j\beta_j(t)\Sigma_j^{-2}(t)}{\sum_{k=1}^N \beta_k^2(t)\Sigma_k^{-2}(t)}\beta_i(t)\right) \quad \text{or} \quad \omega_\star^{\mathbf{0}}(t) \sim \boldsymbol{\Gamma}^{-1}(t)\left(\boldsymbol{b} - \frac{\boldsymbol{b}^{\mathrm{T}}\boldsymbol{\Gamma}^{-1}(t)\boldsymbol{\beta}(t)}{\boldsymbol{\beta}^{\mathrm{T}}(t)\boldsymbol{\Gamma}^{-1}(t)\boldsymbol{\beta}(t)}\cdot\boldsymbol{\beta}(t)\right). \tag{20}$$

This is precisely the maximum-variance market-neutral portfolio that we defined in this section. All we have to do in order to verify it, is to calculate the correlation between the factor return $r^{\mathbf{f}}(t)$ and the return of the portfolio (20). Using the definition of $r_\star^{\mathbf{m}}(t)$ in (4) and the Two-Factor model (11) we obtain:

$$\sum_{i=1}^N \left(\omega_\star^{(0)}\right)_i r_i(t) \sim \sum_{i=1}^N \frac{b_i}{\Sigma_i^2(t)}\left(r_i(t) - \beta_i(t) r_\star^{\mathbf{m}}(t)\right) = \sum_{i=1}^N \frac{b_i^2}{\Sigma_i^2(t)}\cdot r^{\mathbf{f}}(t) + \sum_{i=1}^N \frac{b_i\epsilon_i(t)}{\Sigma_i^2(t)}. \tag{21}$$

The property (12) then guarantees that

$$\begin{aligned}
\mathbb{E}_{t-1}\left(r^{\mathbf{f}}\cdot\sum_{i=1}^N\left(\omega_\star^{(0)}\right)_i r_i\right) &\sim \boldsymbol{b}^{\mathrm{T}}\boldsymbol{\Gamma}^{-1}\boldsymbol{b}\cdot\left(\Sigma^{\mathbf{f}}\right)^2 \\
\mathbb{E}_{t-1}\left(\left(\sum_{i=1}^N\left(\omega_\star^{(0)}\right)_i r_i\right)^2\right) &\sim \left(\boldsymbol{b}^{\mathrm{T}}\boldsymbol{\Gamma}^{-1}\boldsymbol{b}\right)^2\cdot\left(\Sigma^{\mathbf{f}}\right)^2 + \boldsymbol{b}^{\mathrm{T}}\boldsymbol{\Gamma}^{-1}\boldsymbol{H}^{\epsilon}\boldsymbol{\Gamma}^{-1}\boldsymbol{b}.
\end{aligned} \tag{22}$$

Using (13) and the definition of ρ_H from (5) we thus have:

$$\rho_H\left(\boldsymbol{\omega}^{\mathbf{f}}, \boldsymbol{\omega}_\star^{(0)}\right) = \left(1 + \frac{\ell^{\epsilon,\boldsymbol{b}}}{\left(\Sigma^{\mathbf{f}}\right)^2\left(\boldsymbol{b}^{\mathrm{T}}\boldsymbol{\Gamma}^{-1}\boldsymbol{b}\right)}\right)^{-1/2}. \tag{23}$$

We see, finally, that (14) ensures a strong correlation between the return of the portfolio (20) and the factor return $r^{\mathbf{f}}(t)$, and therefore for sufficiently small $\ell^{\epsilon,\boldsymbol{b}}\left(\Sigma^{\mathbf{f}}\right)^{-2}$ all other portfolios satisfying (15) will also fulfil the second part of the maximum-variance market-neutral portfolio definition, the inequality (16).

In Section 2.5 we will present a robust method to estimate the factor loadings b_i. We will see that the ratios b_i/Σ_i can be well modelled empirically by the ranking of stock according to the signal. The estimate works well for most of the factors/styles confirming the universality of the Two-Factor model.

Before closing this subsection let us make two comments:

- There is a connection between the Two-Factor model (11) and the principal component analysis (PCA). In the paragraph following (4) we have already mentioned the way $r^{\mathbf{m}}$, $r_\star^{\mathbf{m}}$ and $\boldsymbol{\beta}$ are related to the PC of the correlation matrix, $\boldsymbol{C}$. Similarly, the right-hand side of (11) alludes to the sub-leading term in the PCA. In Appendix A we present an expression for b_i as *if* it were derived from the WLS regression of the left-hand side of (11) with respect to $r^{\mathbf{f}}$. In particular, this expression leads to the second eigenvector of $\boldsymbol{C}$ (and so $\ell^{\epsilon,\boldsymbol{b}} = 0$), provided $r^{\mathbf{f}}$ is replaced by the sub-leading term in the PCA expansion of the correlation matrix. It is very important to stress, nevertheless, that we *do not* expect $\boldsymbol{b}$ to resemble *any* of $\boldsymbol{C}$'s eigenvectors, as it would have implied, among other things, that $\boldsymbol{b}$ is the same for all factors. Instead, the procedure described in Section 2.5 leads to different results for b_i.

[12] We use the $\sim$ sign to keep in mind the overall rescaling.

- The weights of the *Maximum-Variance market-mode* portfolio, that was first mentioned below (4), are

$$\left(\omega_\star^{\mathbf{m}}\right)_j(t) = \frac{\beta_j \Sigma_j^{-2}}{\sum\limits_{k=1}^{N}\left(\beta_k \Sigma_k^{-1}\right)^2}\,, \tag{24}$$

In Appendix E we argue that $\boldsymbol{\omega}_\star^{\mathbf{m}}$ optimizes the FCL but without any market-neutral constraint treating instead the market mode as yet another factor and replacing b_i and $r^{\mathbf{f}}$ in (11) by β_i and $r_\star^{\mathbf{m}}$ respectively and drop the $\beta_i r_\star^{\mathbf{m}}$ term on the left-hand side. The Maximum-Variance market-mode portfolio could be interpreted as the complementary portfolio of the *Minimum Variance* and *Maximum Diversification* portfolios. Contrary to these two popular portfolios, the weights (24) are higher for higher-beta stocks.

2.4 Reproducing the Eigenvalues of the Empirical Correlation Matrix

Once the K styles are selected and the corresponding K different Maximum-Variance portfolios $\left(\boldsymbol{\omega}_\star^{(0)}\right)_a$ ($a = 1,\ldots,K$) are derived, we may also determine their best combinations to match the unconstrained eigenvectors of the full $N\times N$ correlation matrix. In a very idealistic (and highly unrealistic) scenario each factor corresponds to a different PC of the stock returns. In practice, the vectors $\left(\boldsymbol{\omega}_\star^{(0)}\right)_a$ are *not* eigenvectors of neither $\boldsymbol{H}$ nor $\boldsymbol{\Gamma}^{-1/2}\boldsymbol{H}\boldsymbol{\Gamma}^{-1/2}$, and they are by no means expected to be orthogonal. Therefore, in order to mimic the (largest) eigenvalues of the correlation matrix, we have to diagonalize $\boldsymbol{H}$ in the K-dimensional subspace spanned by $\left(\boldsymbol{\omega}_\star^{(0)}\right)_1,\cdots,\left(\boldsymbol{\omega}_\star^{(0)}\right)_K$ subject to the $\boldsymbol{\Gamma}$-product normalisation. In other words, we have to find a new base of K vectors $\left(\boldsymbol{\omega}_\star^{(1)}\right)_a = R_a^{\ b}\left(\boldsymbol{\omega}_\star^{(0)}\right)_b$, which maximise $\left(\boldsymbol{\omega}_\star^{(1)}\right)_a^{\mathrm{T}}\boldsymbol{H}\left(\boldsymbol{\omega}_\star^{(1)}\right)_a$ for each a with the constraint $\left(\boldsymbol{\omega}_\star^{(1)}\right)_a^{\mathrm{T}}\boldsymbol{\Gamma}\left(\boldsymbol{\omega}_\star^{(1)}\right)_b = \delta_{ab}$.

This, in turn, is equivalent to the optimisation of $\left(\boldsymbol{R}\boldsymbol{h}\boldsymbol{R}^{\mathrm{T}}\right)_{aa}$ for each a with the extra constraint of $\boldsymbol{R}\boldsymbol{\gamma}\boldsymbol{R}^{\mathrm{T}} = \boldsymbol{I}_K$, where $\gamma_{ab} \equiv \left(\boldsymbol{\omega}_\star^{(0)}\right)_a^{\mathrm{T}}\boldsymbol{\Gamma}\left(\boldsymbol{\omega}_\star^{(0)}\right)_b$ and $h_{ab} \equiv \left(\boldsymbol{\omega}_\star^{(0)}\right)_a^{\mathrm{T}}\boldsymbol{H}\left(\boldsymbol{\omega}_\star^{(0)}\right)_b$. This is equivalent to saying that $\boldsymbol{O} = \boldsymbol{R}\boldsymbol{\gamma}^{1/2}$ is the diagonalisation matrix of $\boldsymbol{h}$:

$$\boldsymbol{O}\boldsymbol{\gamma}^{-1/2}\boldsymbol{h}\boldsymbol{\gamma}^{-1/2}\boldsymbol{O}^{T} = \begin{pmatrix} \lambda_{\star\ 1}^{(1)} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_{\star\ K}^{(1)} \end{pmatrix}, \tag{25}$$

where $\lambda_{\star\ a}^{(1)}$'s are the eigenvalues to be compared to those of the empirical correlation matrix. This formula reduces to the diagonalisation of the full correlation matrix for $K = N$, since in this case $\boldsymbol{\gamma} = \boldsymbol{\Gamma}$, $\boldsymbol{h} = \boldsymbol{H}$. Crucially, for any other K the matrix $\boldsymbol{\gamma}$ is not diagonal as it is no longer related to the covariance matrix by $\boldsymbol{\gamma} = \mathrm{diag}\left(\boldsymbol{h}\right)$. Indeed, a necessary condition for $\boldsymbol{\gamma} = \mathrm{diag}\left(\boldsymbol{h}\right)$ to happen is to have zero overlap between the factor portfolios $\left(\boldsymbol{\omega}_\star^{(0)}\right)_a$, which is very unrealistic. Consequently, $\boldsymbol{\gamma}^{-1/2}\boldsymbol{h}\boldsymbol{\gamma}^{-1/2}$ does not have ones along its main diagonal, and so the trace is not equal to K.

The matrices $\boldsymbol{\gamma}$ and $\boldsymbol{h}$ have a simple relation to the FCLs of Maximum-Variance portfolios:

$$\lambda_{\star\ a}^{(0)} \equiv \lambda^{(0)}\left(\boldsymbol{\omega}_{\star\ a}^{(0)}\right) = \frac{h_{aa}}{\gamma_{aa}} \qquad \text{for} \qquad a = 1,\ldots,K\,, \tag{26}$$

where we used the FCL definition (7). Alternatively, one can say that $\boldsymbol{\lambda}_\star^{(0)}$ are the eigenvalues of the diagonal matrix $\mathrm{diag}\left(\boldsymbol{\gamma}\right)^{-\frac{1}{2}}\mathrm{diag}\left(\boldsymbol{h}\right)\mathrm{diag}\left(\boldsymbol{\gamma}\right)^{-\frac{1}{2}}$.

We interpret ${\boldsymbol{\omega}_\star^{(1)}}_a$ and ${\lambda_\star^{(1)}}_a$ as the *constrained* eigenvectors and eigenvalues of the correlation matrix inside the subspace generated by all of the Maximum-Variance factors. Recall that we have already encountered the constrained diagonalisation. The market-neutrality constraint is replaced now by the requirement to remain in the K-dimensional space spanned by $\left(\boldsymbol{\omega}_\star^{(0)}\right)_a$ for $a = 1, \ldots, K$. The analogue of $\boldsymbol{P}^{\mathbf{c}}$ would be now the projection matrix from the N-dimensional into the K-dimensional space spanned by the vectors $\left(\boldsymbol{\omega}_\star^{(0)}\right)_a$, which is by definition the same space as the one spanned by $\left(\boldsymbol{\omega}_\star^{(1)}\right)_a$:

$$\boldsymbol{P}^{\boldsymbol{\omega}_\star^{(0)}} \equiv \sum_{a=1}^{K} \boldsymbol{P}_a\,, \qquad \text{where} \qquad \boldsymbol{P}_a \equiv \left(\boldsymbol{\omega}_\star^{(1)}\right)_a \left(\boldsymbol{\omega}_\star^{(1)}\right)_a^T \boldsymbol{\Gamma}\,. \tag{27}$$

Here we used the fact that $\left(\boldsymbol{\omega}_\star^{(1)}\right)_a$ are already $\boldsymbol{\Gamma}$-orthonormalised (see the end of this section's first paragraph).

Our approach is, in fact, similar to the model of *linear combination* of atomic orbitals in quantum chemistry (LCAO). Since the Schrödinger wave equation is hard to solve for a system with many electrons, the model limits the molecule orbital wave function to a linear combination of atom orbitals. The Hartree-Fock method is then used to determine the coefficients. The analogue of the molecular and atomic orbitals are the $\boldsymbol{\omega}_\star^{(0)}$ and $\boldsymbol{\omega}_\star^{(1)}$ vectors respectively, and the FCL optimisation can be seen as the energy level minimisation of the molecular and atomic Hamiltonians: the former corresponds to $\lambda_\star^{(1)}$ and the latter to $\lambda_\star^{(0)}$.

Overall the procedure of finding the eigenvalues of $\boldsymbol{C}$ can be summarised in the diagram:

$$\boldsymbol{\lambda}^{(0)} \xrightarrow[\text{see (15)}]{\text{Optimisation}} \boldsymbol{\lambda}_\star^{(0)} \xrightarrow[\text{see (25)}]{\text{Diagonalisation}} \boldsymbol{\lambda}_\star^{(1)} \overset{?}{\approx} \boldsymbol{\lambda}^{\text{Emp}} \tag{28}$$

where the latter denotes the first K eigenvalues of the empirical correlation matrix $\boldsymbol{C}$.

In Appendix G we prove that $\boldsymbol{\lambda}^{(1)}$ are maximised in the Maximum-Variance portfolio (the first arrow in the diagram) assuming that the optimisation does not impact too much the off-diagonal elements of $\boldsymbol{h}$. To be more explicit, we show that the ordered $\lambda_i^{(1)}$'s derived from $\boldsymbol{DhD}$ are smaller than those obtained from the diagonalisation of $\boldsymbol{h}$ provided $\boldsymbol{D}$ is diagonal matrix with entries smaller than one. This shows that the FCL optimisation (15) is the right way to reproduce the eigenvalues of the correlation matrix, $\boldsymbol{\lambda}^{\text{Emp}}$. That is to say that the Maximum-Variance factors allow one to replicate as good as possible any alternative beta strategy.

In Section 4 we show that $\boldsymbol{\lambda}_\star^{(1)}$ captures well $\boldsymbol{\lambda}^{\text{Emp}}$ and its dynamics, meaning that the financial and economic constraints help to withdraw some noises in the measurement of $\boldsymbol{\lambda}^{\text{Emp}}$. We also show that the constrained eigenvectors ${\boldsymbol{\omega}_\star^{(1)}}_a$ appear to be unstable portfolios. Indeed, they are invested mainly in the factors with the highest $\left(\lambda_\star^{(0)}\right)_a$, which in turn vary strongly with time explaining the portfolio instability.

2.5 Measurement of the optimal weights

The goal of this section is to estimate the sensitivities b_i's appearing in the Two-Factor model (11). As we briefly mentioned below (11), knowing the covariance matrix of the residual returns, the factor return $r^{\mathbf{f}}$, the sensitivities as well as the market stock index and the betas, one may readily model the correlation matrix of the returns, $\mathbf{C}$.

To use (11) as a definition of b_i, one has to provide the $r^{\mathbf{f}}(t)$ time-series. We may replace $r^{\mathbf{f}}(t)$ by the benchmark return $r^{\mathbf{b.m.}}(t)$, see (9), which should be a reasonable approximation for sufficiently large N. In general, one may start using $r^{\mathbf{b.m.}}$ to extract b_i from (11) using either OLS or WLS linear regression. This should be equivalent to picking an eigenvector b_i of the market-neutral projection of the correlation matrix,

$\boldsymbol{P}^{\mathbf{c}}$-projection of the correlation matrix $\boldsymbol{P}^{\mathbf{c}}\boldsymbol{\Gamma}^{-\frac{1}{2}}\boldsymbol{H}\boldsymbol{\Gamma}^{-\frac{1}{2}}\boldsymbol{P}^{\mathbf{c}}$, whose respective Maximum-Variance portfolio return $\sum_i \left(\omega_\star^{(0)}\right)_i r_i(t)$ has the strongest correlation with $r^{\mathbf{b.m.}}(t)$ among all eigenvectors. This method is, however, not practical for two reasons. First, most eigenvectors are determined with a lot of noises. Second, the correlation between the most correlated eigenvector and the signal could be insignificant.

Let us first discuss how to reduce the noise. The simplest way to achieve this goal, is to group (aggregate) stocks whose signals, and so the rankings, are sufficiently close to each other with respect to the given style F, and to follow the analysis of the previous paragraph for *groups* rather than for single stocks. If two stocks fall into two different groups with respect to the signal associated with factor F, but into the same groups with respect to a different factor F', then the impact of the F' signal on our eigenvector analysis will be significantly suppressed. As we want the groups to be of the same size, the aggregation is equivalent to grouping the stocks into *quantiles*. The new parameter Q should be sufficiently small in order to reduce the noise, but still large enough in order for the regression/eigenvector analysis results to be reliable. We denote the overall number of these quantiles by Q, meaning that at any time there are N/Q stocks in every group/quantile. We will employ a new notation $\mathfrak{q}_i(t)$ for the quantiles. That is $\mathfrak{q}_i(t) \in [1, \ldots, Q]$ for any t, in contrast to $q_i(t) \in [1, \ldots, N]$[13] and[14]

$$\mathfrak{q}_i = \left[q_i \cdot \frac{Q}{N}\right] . \tag{29}$$

To go on with the grouping idea we have to redefine the sensitivities (factor loadings) b_i's. Two stocks i and j belonging to the same quantile should now have identical sensitivities: $b_i = b_j$. This identification, however, seems to be far-fetched for stocks of different size classes. In view of (11) it makes sense to normalise the factor loadings by the corresponding stock volatilities: $\Sigma_i^{-1} b_i = \Sigma_j^{-1} b_j$. This normalisation is yet another manifestation of the fact that the fundamental object describing the stocks dynamics is the correlation, rather the covariance, matrix. We thus have to redefine the factor loadings as functions of the quantile rankings:

$$b_i \to \Sigma_i(t) \mathfrak{B}\left(\mathfrak{q}_i(t)\right) . \tag{30}$$

This is one of the central formulae in this paper, and it is worth making two important comments. First, contrary to the original theoretical Two-Factor model the factor loadings are now time-dependent. This time-dependence, however, comes from the ranks and the volatilities, both varying much slower than the other functions, $r^{\mathbf{m}}(t)$ for instance. We thus do not depart too far form (11). Second, we are about to argue that the function $\mathfrak{B}(\mathfrak{q})$ is surprisingly the same for most of the factors. This is one of our main observations. We will refer to $\mathfrak{B}(\mathfrak{q})$ as the quantile factor loading (or quantile sensitivity) of the $\mathfrak{q}$-quantile.

Here we determine the Market neutral Maximum-Variance portfolio that optimizes the FCL in the sub-space generated by the Q equal-risk weighted quantile portfolios defined by:

$$\left(\omega^{(\mathfrak{q})}\right)_i (t) \equiv \begin{cases} \Sigma_i^{-1}(t) & \text{if} \quad \mathfrak{q}_i(t) = \mathfrak{q} \\ 0 & \text{otherwise} \end{cases} . \tag{31}$$

The Maximum Variance portfolio will be the first market neutral constrained eigenvector of the projected matrix $\widetilde{\boldsymbol{\gamma}}^{-\frac{1}{2}}\widetilde{\boldsymbol{h}}\widetilde{\boldsymbol{\gamma}}^{-\frac{1}{2}}$ of $\boldsymbol{\Gamma}^{-\frac{1}{2}}\boldsymbol{H}\boldsymbol{\Gamma}^{-\frac{1}{2}}$ into the subspace of dimension Q. The $\widetilde{\boldsymbol{\gamma}}$ and $\widetilde{\boldsymbol{h}}$ are the covariance matrix of positions and the covariance matrix of returns between the Q quantile portfolios $\omega^{(\mathfrak{q})}(t)$ and are obtained in the same way as the $\boldsymbol{\gamma}$ and $\boldsymbol{h}$ matrices of Section 2.4. The first difference is that instead of reducing the dimension from N single stocks to K Maximum-Variance portfolios, we reduce here the dimension from N single stocks to the Q quantile portfolios. The second difference is that there is no overlap between the Q quantile portfolios so that $\widetilde{\boldsymbol{\gamma}}$ is diagonal. The important point is that there is no need to use the unknown

[13] We skip the factor superscript $\mathbf{f}$, because until late in the section we consider only one factor at a time.
[14] $[x]$ sands for the integer part of x.

two-factors model in (11) to model the matrix $\widetilde{\boldsymbol{\gamma}}^{-\frac{1}{2}}\widetilde{\boldsymbol{h}}\widetilde{\boldsymbol{\gamma}}^{-\frac{1}{2}}$ to be sure that one of the constrained eigenvector captures well the signal. Indeed, if Q is small enough, the first constrained eigenvector is very likely to capture the signal, if it is strong enough, as we will see later. By identification with the theoretical portfolio (20) we can therefore determine the $\mathfrak{B}(\mathfrak{q})$ as the market neutral constrained first eigenvector of $\widetilde{\boldsymbol{\gamma}}^{-\frac{1}{2}}\widetilde{\boldsymbol{h}}\widetilde{\boldsymbol{\gamma}}^{-\frac{1}{2}}$.

We determine the beta reduced in Q dimension used in the market neutral condition as the beta of each quantile portfolio through (32).

$$\widetilde{\beta}_{\mathfrak{q}}(t) \equiv \sum_{\mathfrak{q}_i(t)=\mathfrak{q}} \Sigma_i^{-1}(t)\beta_i(t)\,. \tag{32}$$

By analogy with the "original" covariance matrix of returns and of positions of single stocks, we define the covariance matrix of returns and of positions of the Q quantile portfolios (see (1) and (2)):

The covariance matrix	The position overlap matrix	The correlation matrix	Dimensions	Indices
$\boldsymbol{H}$	$\boldsymbol{\Gamma} = \mathrm{diag}\,(\boldsymbol{H})$	$\boldsymbol{C} = \boldsymbol{\Gamma}^{-\frac{1}{2}}\boldsymbol{H}\boldsymbol{\Gamma}^{-\frac{1}{2}}$	$N \times N$	$i, j, \cdots$
$\widetilde{\boldsymbol{h}}$	$\widetilde{\boldsymbol{\gamma}}$	$\widetilde{\boldsymbol{C}} = \widetilde{\boldsymbol{\gamma}}^{-\frac{1}{2}}\widetilde{\boldsymbol{h}}\widetilde{\boldsymbol{\gamma}}^{-\frac{1}{2}}$	$Q \times Q$	$\mathfrak{q}, \mathfrak{q}', \cdots$

(33)

It is important to remind here, that while the first line in Table (33) has no reference to any particular style/factor, the quantities in the second line are different for each factor. If the relevant signal is sufficiently strong and/or the Q parameter is correctly chosen, the first market neutral constrained eigenvector $\widetilde{\boldsymbol{C}}$ would be a good proxy of the Market neutral Maximum Variance portfolio derived from the correlation matrix (18) computed from the returns modelled by (11). Therefore by simple term identification, we may expect $\mathfrak{B}(\mathfrak{q})$ [15] to be the *first* market neutral constrained eigenvector of $\widetilde{\boldsymbol{C}}$. There are three important obstacles:

- First, the estimation of $\widetilde{\boldsymbol{C}}$ of the previous sections was based on the conditional expectations, see (1), which is a purely theoretical concept. In practice we have a single length-T time-series for each stock. The best way to built a "smoothed" covariance matrix $\widetilde{\boldsymbol{h}}$ from this data is to use the exponential moving average (EMA) with a parameter α satisfying $1 \ll \alpha^{-1} \ll T$. In Appendix I we explain in details the EMA of all the matrices leading to our estimate of $\widetilde{\boldsymbol{C}}(t)$. In what follows we denote by $\left\langle \widetilde{\boldsymbol{C}}(t) \right\rangle$ the average of this matrix over the entire period T.

- Second, $\widetilde{\beta}_{\mathfrak{q}}(t)$ could depend on time, and we believe that subtracting $\beta_i(t)r^{\mathbf{m}}(t)$ from the stock returns used to estimate $\widetilde{\boldsymbol{h}}$ could lead to a very minor improvement of the estimation of market neutral FCL. As the constrained eigenvectors of $\left\langle \widetilde{\boldsymbol{C}}(t) \right\rangle$ would be only market neutral on average and not at any time, we believe that subtracting $\beta_i(t)r^{\mathbf{m}}(t)$ would simulate an hedge with the stock index to maintain the eigenvectors returns hedged against the index at any time and not only on average so that there is no contribution of the market mode into the FCL.

- Third, to estimate the market neutral constrained eigenvector we have to consider, once more, the *constrained* eigenvectors of $\widetilde{\boldsymbol{C}}$. Recall that the latter is a $Q \times Q$ matrix, while the original $\boldsymbol{P}^{\mathbf{c}}$ projection operator in (18) is $N \times N$, as $\boldsymbol{c} = \boldsymbol{\Gamma}^{-\frac{1}{2}}\boldsymbol{\beta}$ is an N-vector. The new projection vector is determined through $c = \widetilde{\boldsymbol{\gamma}}^{-\frac{1}{2}}\widetilde{\beta}$ defined in (32).

  Contrary to the correlation matrix $\boldsymbol{C}$, however, it will be difficult to project out the market mode independently for each t as we did for the Two-Factor model around (18). This is so because we use

[15] We estimate a market neutral form of $\mathfrak{B}(\mathfrak{q})$ that is to say it is possible that in reality the weights correspond to $\mathfrak{B}(\mathfrak{q}) + k\widetilde{\boldsymbol{\gamma}}^{-\frac{1}{2}}\widetilde{\beta}$

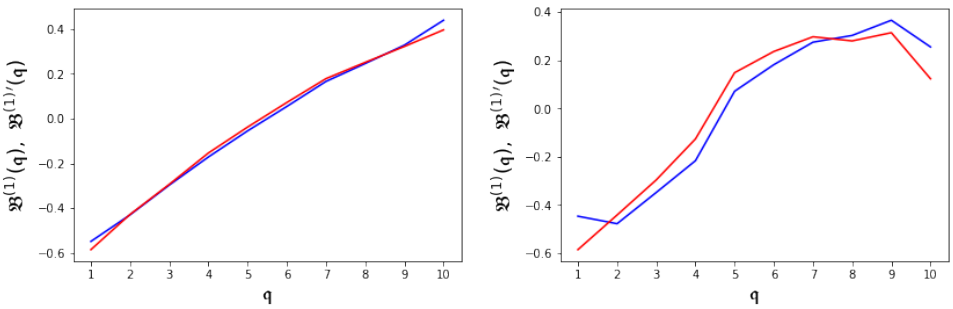

Figure 1: The graphs show $\mathfrak{B}(\mathfrak{q})$ (blue line) and $\mathfrak{B}'(\mathfrak{q})$ (red line) for two factors: *Beta* (left) and *Cash* (right) with $Q = 10$. In both cases the adjustment of the second step is quite small for all quantiles $\mathfrak{q} = 0, \ldots, 9$. The same holds for the remaining eleven factors.

EMA to compute the *empiric* correlation matrix $\widetilde{\boldsymbol{C}}$. To make sense of the projection one then has to "smooth" as well the corresponding $\widetilde{\boldsymbol{P}^{\boldsymbol{c}}}$ matrix built from $\widetilde{\beta}_{\mathfrak{q}}(t)$, ending up with a projection matrix that does not satisfy the basic property $\boldsymbol{P}^2 = \boldsymbol{P}$. We will prefer instead to diagonalise $\left\langle \widetilde{\boldsymbol{C}}(t) \right\rangle$ with a constant projection defined by:

$$\widetilde{\boldsymbol{P}^{\boldsymbol{c}}} = \boldsymbol{I}_{\mathfrak{q}} - \frac{\widetilde{\boldsymbol{\beta}}\widetilde{\boldsymbol{\gamma}}^{-1}\widetilde{\boldsymbol{\beta}}^{\mathrm{T}}}{\widetilde{\boldsymbol{\beta}}^{\mathrm{T}}\widetilde{\boldsymbol{\gamma}}^{-1}\widetilde{\boldsymbol{\beta}}} \qquad \text{where} \qquad \widetilde{\boldsymbol{\beta}} \equiv \langle \boldsymbol{\beta}(t) \rangle \tag{34}$$

is the mean of the betas over the T-period and $\boldsymbol{I}$ is the $Q \times Q$ identity matrix. As we discussed in this section on two different occasions, the constrained diagonalisation of $\widetilde{\boldsymbol{C}}$ is equivalent to the regular diagonalisation of $\widetilde{\boldsymbol{P}^{\boldsymbol{c}}}\widetilde{\boldsymbol{C}}\widetilde{\boldsymbol{P}^{\boldsymbol{c}}}$. This additional subtraction of the market-mode should provide a better estimation of the factor sensitivities.

With these three points in mind we propose a two-step procedure to evaluate the factor loadings $\mathfrak{B}(\mathfrak{q})$.

First step. We calculate the EMA version of the correlation matrix $\widetilde{\boldsymbol{C}}(t)$ from the Σ-normalised quantile portfolio $\omega_{\mathfrak{q}}(t)$ defined in (31) while subtracting the part of the portfolio returns explained by the stock index, and use $\widetilde{\boldsymbol{P}^{\boldsymbol{c}}}$ to find the $(Q-1)$ constrained eigenvectors of $\left\langle \widetilde{\boldsymbol{C}}(t) \right\rangle$ orthogonal to $\widetilde{\boldsymbol{\gamma}}^{-\frac{1}{2}}\widetilde{\beta}_{\mathfrak{q}}(t)$. Each of the eigenvectors, which we will denote by $\boldsymbol{\mathfrak{B}}^{(\mathfrak{p})}$ with $\mathfrak{p} = 1, \ldots, Q-1$ and their eigenvalues by $\lambda^{\mathfrak{B}}_{\mathfrak{p}} \equiv \lambda\left(\mathfrak{B}^{(\mathfrak{p})}\right)$, gives rise to a portfolio $\boldsymbol{v}^{(\mathfrak{p})}$ defined by (20) with factor loadings determined by

$$b_i(t) = \Sigma_i(t)\mathfrak{B}^{(\mathfrak{p})}\left(\mathfrak{q}_i(t)\right) . \tag{35}$$

These $(Q-1)$ portfolios are by construction market-neutral and, as we commented above the portfolio built from $\mathfrak{B}_{\mathfrak{q}_i(t)}$, that is $\boldsymbol{v}$, has a good chance to mimic the Maximum-Variance portfolio $\boldsymbol{\omega}^{(0)}_{\star}$ provided the signal is strong enough and the number of quantiles is properly chosen to capture it. Accordingly, the first constrained eigenvalue, $\lambda^{\mathfrak{B}}_1$, associated with $\boldsymbol{\mathfrak{B}}_1$ is a good estimate of $\lambda^{(0)}_{\star}$, the optimal FCL from (15). The signal strength is directly translated into $\lambda^{\mathfrak{B}}_1$'s value.

Second step. At the final stage of the first step we get $(Q-1)$ portfolios. We used (20) to ensure their market-neutrality for any t. This adjustment does not necessarily respect the optimisation involved in finding $\mathfrak{B}(\mathfrak{q})$. Therefore, these portfolios are not a priori optimal. The second step will correct the $\boldsymbol{v}$ portfolios slightly so they become optimal while staying market neutral at any time. We may consider the covariance matrices $(\widetilde{\boldsymbol{h}}', \widetilde{\boldsymbol{\gamma}}')$ of positions and of returns of the $\boldsymbol{v}^{\mathfrak{p}}$ portfolios for $\mathfrak{p} = 1, \ldots, (Q-1)$. To obtain this matrix we proceed as in the previous step (including the EMA and the time average), but replacing the normalised quantile portfolios $\omega^{(\mathfrak{q})}(t)$ with the $(Q-1)$ $\boldsymbol{v}^{(\mathfrak{p})}$ portfolios. The covariance matrices $\widetilde{\boldsymbol{h}}', \widetilde{\boldsymbol{\gamma}}'$ should be very close to be diagonal, since the portfolios were by construction orthogonal to each other before being slightly adjusted. Let us denote by $\boldsymbol{\mathcal{O}}$ the rotation matrix that diagonalises the covariance matrix $\widetilde{\boldsymbol{\gamma}}^{-\frac{1}{2}}\widetilde{\boldsymbol{h}}\widetilde{\boldsymbol{\gamma}}^{-\frac{1}{2}}$ and $\lambda_{\mathfrak{p}}^{\mathfrak{B}\,\prime}$ the eigenvalues. Then the improved sensitivities $\mathfrak{B}^{(\mathfrak{p})\,\prime}$ are:

$$\mathfrak{B}^{(\mathfrak{p})\,\prime} = \sum_{\bar{\mathfrak{p}}=1}^{Q-1} \mathcal{O}_{\bar{\mathfrak{p}}}^{\mathfrak{p}}\, \mathfrak{B}^{(\bar{\mathfrak{p}})} \tag{36}$$

As above for each vector $\mathfrak{B}^{(\mathfrak{p})\,\prime}$ we have a corresponding market-neutral portfolio, $\boldsymbol{v}^{(\mathfrak{p})\,\prime}$, and $\boldsymbol{v}^{(1)\,\prime}$ should provide even a better approximation for Maximum-Variance market-neutral portfolio (20), than $\boldsymbol{v}^{(1)}$ of the first step. The first eigenvalue $\lambda_1^{\mathfrak{B}\,\prime}$ associated with $\mathcal{O}$ is expected to be a better estimation of the optimal $\lambda_\star^{(0)}$ than the previous $\lambda_1^{\mathfrak{B}}$. This step may be repeated again, but we will see that there is no need in doing so, as the portfolio stabilises after the very first adjustment.

The two-step procedure can be summarized in the following diagram:

$$\begin{pmatrix} r_i(t) \\ \mathfrak{q}_i(t) \end{pmatrix} \overset{(31)}{\mapsto} \omega^{(\mathfrak{q})}(t) \overset{(33)(32)}{\mapsto} \begin{pmatrix} \widetilde{\boldsymbol{\gamma}} \\ \widetilde{\boldsymbol{h}} \\ \widetilde{\boldsymbol{\beta}} \end{pmatrix} \mapsto \begin{pmatrix} \mathfrak{B}^{(\mathfrak{p})} \\ \lambda_{\mathfrak{p}}^{\mathfrak{B}} \end{pmatrix} \overset{(35)}{\mapsto} b_i \overset{(20)}{\mapsto} v_i^{(\mathfrak{p})} \mapsto \begin{pmatrix} \mathcal{O}_{\mathfrak{p}_2}^{\mathfrak{p}_1} \\ \lambda_{\mathfrak{p}}^{\mathfrak{B}\,\prime} \end{pmatrix} \overset{(36)}{\mapsto} \mathfrak{B}_{\mathfrak{p}}{}' \overset{(35)}{\mapsto} b_i' \overset{(20)}{\mapsto} v_i^{(\mathfrak{p})\,\prime} = \left(\omega_\star^{(0)}\right)_i \tag{37}$$

with

$$i \in [1, N] \qquad \mathfrak{q} \in [1, Q] \qquad \mathfrak{p} \in [1, Q-1]\,. \tag{38}$$

It is crucial to emphasize here that the definition of the Maximum-Variance market-neutral portfolio included the maximum correlation condition (16) which we have not yet mentioned here. Instead we select the *first* constrained eigenvector and use it to construct $\boldsymbol{\omega}_\star^{(0)}$. To verify the consistency of this approach we will find the correlation between the returns of the portfolio obtained in the end of (37) and the benchmark portfolio.

In Table 14 of Appendix B we list the fourteen factors used to estimate the quantile sensitivities. We set $Q = 10$. We summarise technical details in Appendix I, with the first two lines of Table 18 being the reference for the relevant calculation. Our results show that the difference between $\mathfrak{B}^{(\mathfrak{p})\,\prime}$ and $\mathfrak{B}^{(\mathfrak{p})}$ is insignificant for all the factors. On Figure 1 we present both functions for two randomly selected factors with $\mathfrak{p} = 1$ to illustrate this fact. We thus may assume that the eigenvalues $\lambda_{\mathfrak{p}}^{\mathfrak{B}}$ do not change much after the second step adjustment. In order to avoid clustering of indices and superscripts we will omit the $'$ in $\mathfrak{B}^{(\mathfrak{p})}$ and $\lambda_{\mathfrak{p}}^{\mathfrak{B}}$.

On Figure 2 we demonstrate the $\mathfrak{B}^{(1)}$ results for all fourteen factors. We see that with a common overall normalisation for nine factors this function is very well approximated by:

$$\mathfrak{B}_\star(\mathfrak{q}, Q) \approx \frac{1}{Q-1} \cdot (\mathfrak{q} - 1) - \frac{1}{2} \qquad \text{for} \qquad \mathfrak{q} = 1, \ldots, Q\,. \tag{39}$$

This confirms the universal nature of our Ansatz (35). This is one of our main observations. As for the remaining factors, we will argue below that their signals capture a risk factor that is too weak, explaining the deviation of their $\mathfrak{B}(\mathfrak{q})$ functions from (39).

On Figure 4 we compare the double Heaviside function used to construct a benchmark portfolio (9), $\mathfrak{B}_\star(\mathfrak{q}, Q)$ of (39) and a random form of $\mathfrak{B}^{(\mathfrak{p})}(\mathfrak{q})$. Since for $Q = 10$ (and in fact for any other multiple

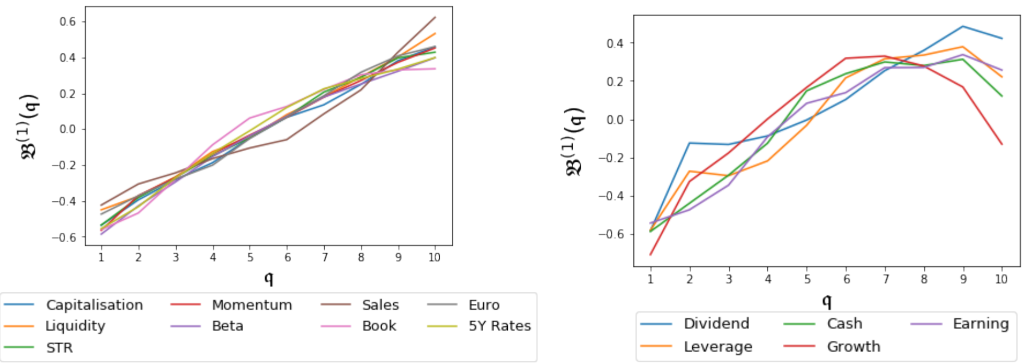

Figure 2: For nine factors (left) the function $\mathfrak{B}(\mathfrak{q})$ is very close to (39). For the remaining five factors these functions have a different shape that could be explained either by noises or by convergence to another factor/risk. We will see below that five signals capture risk factors that are too weak to measure $\mathfrak{B}(\mathfrak{q})$ properly. The normalisation is identical for all factors.

of 5) the Heaviside function of the benchmark portfolio might be represented as a linear combination of Q generic vectors $\mathfrak{B}^{(\mathfrak{p})}(\mathfrak{q})$, we may assume that the benchmark portfolio is, in turn, a linear combination of the Q market-neutral quantile portfolios $\boldsymbol{v}^{(\mathfrak{p})}$'s. This simple observation allows us to use (8) in the $(Q-1)$ dimensional subspace spanned by $\boldsymbol{v}^{(\mathfrak{p})}$'s, in order to find the benchmark portfolio FCL as a weighted sum of the eigenvectors $\lambda_{\mathfrak{p}}^{\mathfrak{B}}$:

$$\lambda^{(0)}\left(\boldsymbol{\omega}^{\mathbf{b.m.}}\right)=\sum_{\mathfrak{p}=1}^{Q-1}\left(a_{\mathfrak{p}}^{\mathbf{b.m.}}\right)^{2}\lambda_{\mathfrak{p}}^{\mathfrak{B}} \quad \text{with} \quad a_{\mathfrak{p}}^{\mathbf{b.m.}}\equiv\rho_{\Gamma}\left(\widetilde{\boldsymbol{\gamma}}^{-\frac{1}{2}}\boldsymbol{v}^{(\mathfrak{p})},\boldsymbol{\omega}^{\mathbf{b.m.}}\right) \quad \text{and} \quad \sum_{\mathfrak{p}=1}^{Q-1}\left(a_{\mathfrak{p}}^{\mathbf{b.m.}}\right)^{2}=1\,. \tag{40}$$

Though this equation is only an approximation, we may use it to evaluate the strength signal for different factors. Ordering the eigenvalues $\lambda_{\mathfrak{p}}^{\mathfrak{B}}$, we see that:

$$\left(a_{1}^{\mathbf{b.m.}}\right)^{2}\geqslant\frac{\lambda^{(0)}\left(\boldsymbol{\omega}^{\mathbf{b.m.}}\right)-\lambda_{2}^{\mathfrak{B}}}{\lambda_{1}^{\mathfrak{B}}-\lambda_{2}^{\mathfrak{B}}}\,. \tag{41}$$

The right-hand side of this inequality may serve as indications for the signal strength. The value of $\lambda_{1}^{\mathfrak{B}}$ may be seen as the optimised FCL in the $(Q-1)$ dimensional space of market-neutral quantile portfolios. On the other hand, $\boldsymbol{\omega}^{\mathbf{b.m.}}$ is our proxy for the factor portfolio $\boldsymbol{\omega}^{\mathbf{f}}$. Thus for a strong signal (strong means that the signal captures a risk that is high and that both $\lambda^{(0)}\left(\boldsymbol{\omega}^{\mathbf{b.m.}}\right)$ and $\lambda_{1}^{\mathfrak{B}}$ should be high) and a properly chosen Q, one should get $\lambda^{(0)}\left(\boldsymbol{\omega}^{\mathbf{b.m.}}\right)$ closer to $\lambda_{1}^{\mathfrak{B}}$ than to $\lambda_{2}^{\mathfrak{B}}$, implying that the right-hand side of (41) is greater than $\frac{1}{2}$. That implies that $a_{1}^{\mathbf{b.m.}}$, that measures the position overlap between the benchmark portfolio and the first constrained eigenvector, is higher than $\frac{1}{2}$. This makes it very likely for the first eigenvector to be the most correlated to the signal. For weaker signals with low $\lambda^{(0)}\left(\boldsymbol{\omega}^{\mathbf{b.m.}}\right)$ one should get $\lambda^{(0)}\left(\boldsymbol{\omega}^{\mathbf{b.m.}}\right)$ closer to $\lambda_{2}^{\mathfrak{B}}$ than to $\lambda_{1}^{\mathfrak{B}}$, meaning that the right-hand side of (41) is below $\frac{1}{2}$. For weak signals it is very likely that the first eigenvector captures some noise or a risk that has nothing to do with the original signal. We believe that $\lambda_{2}^{\mathfrak{B}}$ is increasing with Q so that for weaker signal we should set a smaller Q in order to increase the possibility for the first constrained eigenvector to capture well its signal.

On Figure 3 we show the dependence between the time average of the two sides of (41). The inequality holds for all fourteen factors, and even more impressively, the values are close to one exactly for the those

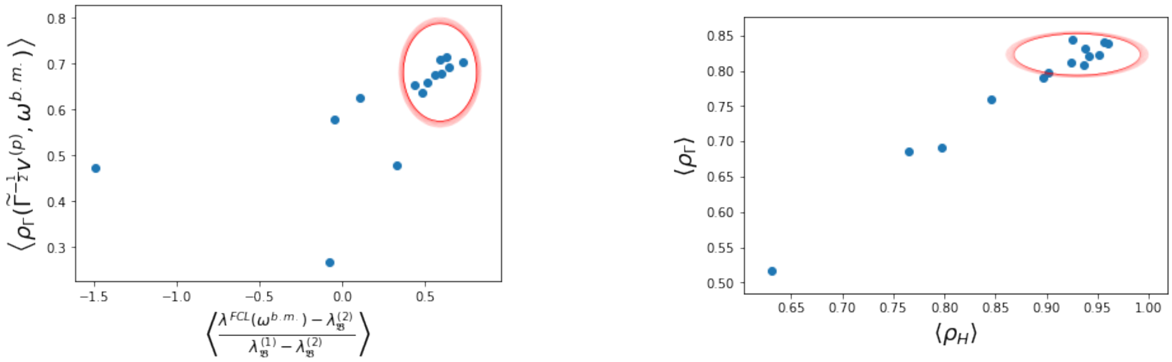

Figure 3: For the left graph the (x, y)-values are the right and the left-hand sides of the inequality (41) respectively for different factors averaged over the T-time period. While the inequality holds for all factors, the values are close to the point $(.5, .75)$ (surrounded by the red ellipse) only for the "strong" factors with measurements that resemble $\mathfrak{B}^{(1)}(\mathfrak{q})$ of (39). On the right graph we plotted the same position overlaps versus the H-correlation coefficients (5). The red ellipse encircles the same factors as on the left graph.

factors whose vector $\mathfrak{B}$ is well-approximated by (39). On the same figure we show the connection between the position overlap $\rho_\Gamma\left(\widetilde{\gamma}^{-\frac{1}{2}}\boldsymbol{v}, \boldsymbol{\omega}^{\mathbf{b.m.}}\right)$ and the ρ_H-correlation between the same portfolios, as it was defined in (5). The two quantities are closely related. For two portfolios with similar positions, the corresponding returns will be strongly correlated. As the graph shows, the points standing for the factors producing (39) are indeed close to $(\rho_\Gamma, \rho_H) = (1, 1)$.

2.6 Validation of the empirical universal law

In the previous subsection we showed how to estimate the factor loadings b_i's by grouping the N stocks into Q quantiles. The final result

$$b_i(t) = \Sigma_i(t)\mathfrak{B}_\star(\mathfrak{q}_i(t)) \qquad \text{with} \qquad \mathfrak{B}_\star(\mathfrak{q}) = \frac{1}{Q-1}\cdot(\mathfrak{q}-1) - \frac{1}{2}\,. \tag{42}$$

is universal for nine stronger factors out of fourteen.

We find this result very surprising. Ignoring the impact of different Σ_i's, the linear form of $\mathfrak{B}_\star(\mathfrak{q})$ guarantees that for a fixed range of factor loadings, between $\bar{b}$ and $\bar{b} + \delta b$, there is the same number of quantiles with loading in this range regardless of $\bar{b}$. Extrapolating to the full $1, \ldots, N$ ranking, it means that the sensitivities have a uniform distribution. This is in contrast to our expectations of getting a distribution close to the Gaussian one. This would lead to higher values of $\mathfrak{B}(\mathfrak{q})$ for lower quantiles, and smaller $\mathfrak{B}(\mathfrak{q})$ for larger $\mathfrak{q}$'s. Instead we observe a very good straight line approximation.

By its very construction the estimate forces a sensibility $b_i(t)$ to depend on the relevant quantile ranking $\mathfrak{q}_i \in [1, \ldots, Q]$ rather than the standard ranking $q_i(t) \in [1, \ldots, N]$. Once we have arrived at the final estimate, however, we may take one step forward and adapt (39) for the full rankings $q_i(t)$'s:

$$b_i(t) = \Sigma_i(t)\mathfrak{B}_\star(q_i(t)) \qquad \text{with} \qquad \mathfrak{B}_\star(q) = \frac{1}{N-1}\cdot(q-1) - \frac{1}{2}\,. \tag{43}$$

Here the expression for $\mathfrak{B}_\star(q, N)$ follows from (42) by replacing $(\mathfrak{q}, Q) \to (N, q)$. In what follows we will sometimes refer to $\mathfrak{B}_\star(\mathfrak{q})$ and $\mathfrak{B}_\star(q)$ as the "multi-step" and (simply) the linear $\mathfrak{B}$-functions. They appear

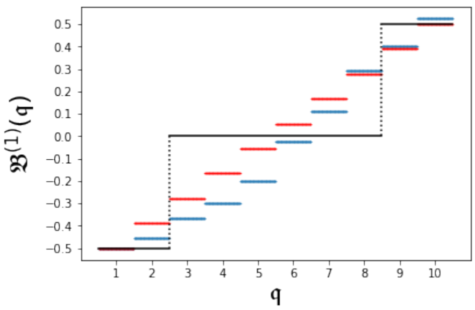

Figure 4: the schematic comparison between the benchmark weights (9) with the 20% bottom/top double Heaviside function shown in black, the linear $\mathfrak{B}_\star(\mathfrak{q}, Q)$ function of (39) shown in red and finally a generic form of $\mathfrak{B}^{(\mathfrak{p})}(\mathfrak{q})$ for $\mathfrak{p} > 1$ in light blue. For simplicity we use the same normalisation for the two functions and ignore the Σ-factors in (9). The double Heaviside function may be approximated by a linear combination of the $Q-1$ functions $\mathfrak{B}^{(\mathfrak{p})}(\mathfrak{q})$.

on Figure 5. As we have mentioned in Introduction, a similar version of this weighting (without the volatility factor) was used in [16] for the *Value* (*Book* in our conventions) and *Momentum* factors.

With the factor loadings (42) and (43) at hand we may now construct two portfolios using (20). From the discussion of the previous subsection, it is clear that the former is the Maximum-Variance market-neutral portfolio of the pair $\left(\widetilde{\boldsymbol{h}}, \widetilde{\boldsymbol{\gamma}}\right)$. It is thus natural to expect that the portfolio built from (43) is a good approximation for the Maximum-Variance market-neutral portfolio of the original pair $(\boldsymbol{H}, \boldsymbol{\Gamma})$. With this assumption in mind, we denote by $\lambda_\star^{(0)}$ the (optimised) FCL of this portfolio. We already referred to the (optimised) FCL of (42) as $\lambda_1^{\mathfrak{B}}$. We summarise the values $\lambda_\star^{(0)}$ and $\lambda_1^{\mathfrak{B}}$ on Table 1 for all fourteen factors.

2.7 Comparison with the Fama-MacBeth Regression

In this section we would like to compare our findings with the most popular approaches known in the literature.

In the previous section we argued that a straightforward linear regression between $r_i(t)$ and one of the benchmark returns $r^{\mathbf{b.m.}}(t)$ is not the best tool to estimate the factor loadings. Instead we adopted Ansatz (35) and avoided almost completely the use of $r^{\mathbf{b.m.}}(t)$. Nevertheless, it is interesting to compare the FCL with the classical R^2 that is optimized in the Fama-MacBeth regression [47] when using the linear regression analysis. $R^2(\mathfrak{q})$ for each quantile portfolio is estimated for the market model (CAPM) and the Two-Factor model:

- $R^2_{\text{2FM}}(\mathfrak{q})$ obtained from a multi-linear regression of the Two-Factor model (11) and
- $R^2_{\text{CAPM}}(\mathfrak{q})$ obtained from a simple linear regression of the pair $(r_i(t), r^{\mathbf{m}}(t))$

The difference between the two coefficients may measure the improvement in the replication of asset return by the two returns, $r^{\mathbf{m}}(t)$ and $r^{\mathbf{f}}(t)$, of (11), over that with the single $r^{\mathbf{m}}(t)$ return of the CAPM within the given quantile $\mathfrak{q}$.

This approach was carried out by Fama and French in [2] for the celebrated Three-Factor Model. Table 4 of this paper summarizes the R^2 coefficients for the CAPM fitting within each group of the Small-Minus-Big (SMB) classification of market capitalizations *and* the High-Minus-Low (HML) classification of book-to-market ratios. The coefficients in this table are significantly smaller as compared to their counterparts in Table 7a, where the multi-linear regression now includes two additional returns: those of the SMB and the HML benchmark portfolios. This increase confirms the predictive power of the Three-Factor Model.

In Figure 6 we plot the optimised FCL $\lambda_\star^{(0)}$ for all our factors as a function of $\langle R^2_{2FM} - R^2_{CAPM}\rangle$, with the average over the Q quantiles. We see that the higher the FCL is the higher the improvement in R^2 could be expected. In Appendix H we provide more details about the similarities and differences between the classical

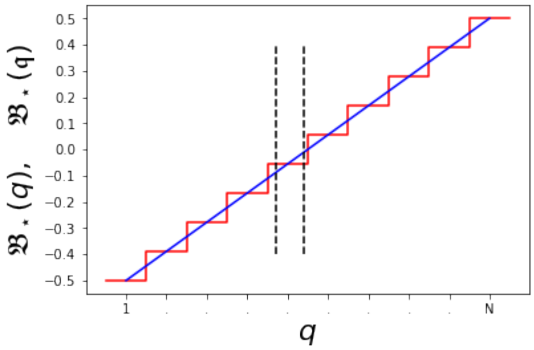

Figure 5: The comparison between the linear (blue) function $\mathfrak{B}_\star(\mathfrak{q})$ and the multi-step (red) function $\mathfrak{B}_\star(q)$ of (42) and (43) respectively. The relation between q and $\mathfrak{q}$ appears in (29). The black dashed lines correspond to two different values of q. They intersect $\mathfrak{B}_\star(\mathfrak{q})$ on the same "step", and yet yield different values of $\mathfrak{B}_\star(q)$. This may explain the results in Table 1.

approach of [47, 2] and the one of our paper. We also argue in Appendix D that the FCL optimisation might be interpreted as a constrained WLS regression analysis, in contrast to the unconstrained OLS of [47, 2]. Another important point is that the original Fama-MacBeth approach does not yield optimal factors, unless one repeats the regression analysis iteratively. Nevertheless, successive iterations will eventually converge to the eigenvectors of the empirical covariance matrix (either in its full version or reduced version). In the best case scenario these eigenvectors are, in turn, a noisy mixture of the initial signals. In the worst case, they will have no relation to the initial signals. Moreover, the number Q of quantiles in the Fama-MacBeth approach is proportional to K^2 (the number of factors squared), making the dimensional reduction very difficult in the regime, where $Q \ll N$ does not hold any more. That explains why the 20% top-bottom approach has remained a standard tool in finance even though it is not optimal.

2.8 Random factors and the robustness check

Our method of finding the eigenvalues of the correlation population matrix consists of two main steps. First, we optimise the FCL for each factor separately. This yields K Maximum-Variance portfolios $\left(\boldsymbol{\omega}_\star^{(0)}\right)_a$ for $a = 1, \ldots, K$. We denoted optimal FCLs values by $\left(\lambda_\star^{(0)}\right)_a$. At the second step we optimise the FCL in the K-dimensional subspace spanned by the K portfolios. This results in yet another K pairs $\boldsymbol{\omega}_\star^{(0)}{}_a$, $\lambda_\star^{(0)}{}_a$.

The goal of this section is to test the second step using *randomly* generated factor portfolios $\left(\boldsymbol{\omega}_\star^{(0)}\right)_a$. This is a robustness check of our scheme.

We start with a random correlation matrix $\boldsymbol{C}^{\mathbf{r}\cdot}$. We denote its eigenvectors/eigenvalues pairs by $(\boldsymbol{e}_i^{\mathbf{r}\cdot}, \ell_i^{\mathbf{r}\cdot})$ for $i = 1, \ldots, N$ as in (8). Then the ith component of the ath factor portfolio is generated by:

$$(v_a^{\mathbf{r}\cdot})_i = \sum_{j=1}^{N} \left(\ell_j^{\mathbf{r}\cdot}\right)^{\mu/2} \left(\boldsymbol{e}_j^{\mathbf{r}\cdot}\right)_i \mathcal{E}_{ja}\,, \tag{44}$$

where $\boldsymbol{\mathcal{E}}$ is a $N \times K$ matrix of standard normal random variables simulating our returns, and μ is a free positive parameter to be fixed soon. The intuition behind this Ansatz is that for a given a the portfolio $v_a^{\mathbf{r}\cdot}$ is a random linear mixture of the correlation matrix eigenvectors with weights determined by their eigenvalues. To put it differently, the linear combination is dominated by the leading eigenvectors, and this dominance is controlled by μ. In particular, in the limit $\mu \to \infty$ all of the vectors defined in (44) are proportional to the first eigenvector, while for $\mu = 0$ there is no preference to any particular eigenvector.

We identify (44) with the Maximum-Variance market-neutral portfolios $\boldsymbol{\omega}_\star^{(0)}$:

$$\boldsymbol{v}_a^{\mathbf{r}\cdot} \sim \boldsymbol{\omega}_\star^{(0)}{}_a \tag{45}$$

Factor	$\lambda_\star^{(0)}$	$\lambda_1^{\mathfrak{B}}$	
Beta	6.35	6.10	•
Momentum	5.87	5.34	•
5Y Rates	5.38	5.35	•
Capitalisation	4.43	4.39	•
STR	4.10	4.14	•
Dividend	3.86	5.64	
Euro	3.72	3.60	•
Sales	2.88	2.96	•
Liquidity	2.82	2.88	•
Book	2.49	2.60	•
Leverage	2.08	2.32	
Earning	1.89	2.09	
Cash	1.52	1.64	
Growth	1.47	1.68	

Table 1: Summary of two FCLs $\lambda_\star^{(0)}$ and $\lambda_1^{\mathfrak{B}}$, corresponding to the Maximum-Variance portfolios based on (43) and (42) respectively. We see that $\lambda_\star^{(0)}$ outperforms $\lambda_1^{\mathfrak{B}}$ for the three leading factors, while the situation is different for the smallest values. It can be explained by the fact that the multi-step function fails to capture the impact of two stocks that fall into the same quantile, but nevertheless have significantly different correlations with the factor return, see Figure 5. This is obviously more relevant for larger FCLs. For smaller ones $\lambda_1^{\mathfrak{B}} > \lambda_\star^{(0)}$ instead, because $\lambda_1^{\mathfrak{B}}$ comes from a genuine optimisation rather than from the "educated guess" (43). The bullet on the right-hand side denotes factors (styles) with strong signals, see Figure 3.

It is now straightforward to compute the FCL associated with $\boldsymbol{v}_a^{\mathbf{r}\cdot}$: the first formula in (8) is precisely what we need. We obtain:

$$\lambda^{\mathbf{r}\cdot}{}_a^{(0)} = \frac{\sum\limits_{j=1}^{N} \left(\ell_j^{\mathbf{r}\cdot}\right)^{1+\mu} \mathcal{E}_{ja}^2}{\sum\limits_{j=1}^{N} \left(\ell_j^{\mathbf{r}\cdot}\right)^{\mu}} \,. \tag{46}$$

Continuing along the lines of Section 2.5 we can also find the constrained eigenvalues $\left(\lambda^{\mathbf{r}\cdot(1)}\right)_a$. As we described in details in Section 2.5, they are the optimised FCLs in the K-dimensional space spanned by the vectors $\boldsymbol{v}_a^{\mathbf{r}\cdot}$, and one can easily compute them from the diagonalisation of the matrix $\boldsymbol{\gamma}^{\mathbf{r}\cdot -\frac{1}{2}} \boldsymbol{h}^{\mathbf{r}\cdot} \boldsymbol{\gamma}^{\mathbf{r}\cdot -\frac{1}{2}}$, where

$$h_{ab}^{\mathbf{r}\cdot} = \boldsymbol{v}_a^{\mathbf{r}\cdot} \boldsymbol{C}^{\mathbf{r}\cdot} \boldsymbol{v}_b^{\mathbf{r}\cdot} \qquad \text{and} \qquad \gamma_{ab}^{\mathbf{r}\cdot} = \sum_{i=1}^{N} \left(v_a^{\mathbf{r}\cdot}\right)_i C_{ii}^{\mathbf{r}\cdot} \left(v_b^{\mathbf{r}\cdot}\right)_i \,. \tag{47}$$

Following previous notations we denote the constrained eigenvalues by $\lambda^{\mathbf{r}\cdot}{}_i^{(1)}$.

To summarize so far, we showed how to derive the $\lambda^{\mathbf{r}\cdot}{}_a^{(0)}$ and $\lambda^{\mathbf{r}\cdot}{}_i^{(1)}$ eigenvalues from the random correlation matrix $\boldsymbol{C}^{\mathbf{r}\cdot}$ and a Gaussian matrix $\boldsymbol{\mathcal{E}}$. To finally test our method we still have to decide how to generate the correlation matrix. For sufficiently large N, the eigenvalues of a random covariance matrix Ω_{ij} with trace equal to N ($\mathrm{Tr}\,(\boldsymbol{\Omega}) = N$) and of the corresponding correlation matrix ${\Omega_{ii}}^{-\frac{1}{2}} \Omega_{ij} {\Omega_{ii}}^{-\frac{1}{2}}$ are very close to each other. We therefore will construct $\boldsymbol{H}^{\mathbf{r}\cdot}$ instead of $\boldsymbol{C}^{\mathbf{r}\cdot}$. Furthermore, we will set the eigenvectors $\boldsymbol{e}_i^{\mathbf{r}\cdot}$ to be a random orthonormal basis (rows of an orthonormal matrix), while for the eigenvalues $\ell_i^{\mathbf{r}\cdot}$ we will take the empiric eigenvalues in the penultimate column of Table 5. Recall that these values were obtained from 5-minutes returns over almost six-years long period. The Marčenko-Pastur ratio N/T is therefore very small, meaning that these empiric eigenvalues should not be different from the eigenvalues of the "theoretical"

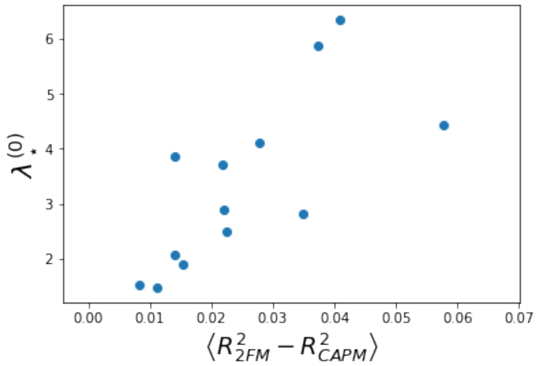

Figure 6: The FCL $\lambda_\star^{(0)}$ (the first column of Table (1)) as a function of the mean difference $R^2_{2FM}(\mathfrak{q}) - R^2_{CAPM}(\mathfrak{q})$. We see that the difference is always positive, indicating a better regression fit. Moreover, the improvement clearly works much better for factors with larger FCL. It means that the stronger the signal, the better works the Two-Factor model (11). We used here the equal-weighted version of the benchmark portfolio, see (9), to be closer to the original computations of [2].

population correlation matrix. Notice as well that for $i > K$ these eigenvalues are very noisy. It makes then to set all $\ell_i^{\mathbf{r}\cdot}$ for $i > K$ to the same constant value $\ell_{\mathbf{c}.}$, which is fixed by the trace condition. To summarize:

$$\boldsymbol{H}^{\mathbf{r}\cdot} = \sum_{a=1}^{K} \lambda_a^{\text{Emp}} \boldsymbol{e}_a^{\mathbf{r}\cdot} \boldsymbol{e}_a^{\mathbf{r}\cdot\text{T}} + \ell_{\mathbf{c}.} \sum_{j=K+1}^{N} \boldsymbol{e}_j^{\mathbf{r}\cdot} \boldsymbol{e}_j^{\mathbf{r}\cdot\text{T}} \qquad \text{with} \qquad \ell_{\mathbf{c}.} = 1 - \frac{\sum_{a=1}^{K} \lambda_a^{\text{Emp}}}{N-K} \,. \tag{48}$$

Overall our test follows the following diagram:

$$\begin{pmatrix} \ell_i^{\mathbf{r}\cdot} = \left(\lambda_a^{\text{Emp}}, \ell^{\mathbf{c}\cdot}\right) & \text{Table 5} \\ \boldsymbol{e}_i^{\mathbf{r}\cdot} & \text{random}, \, \boldsymbol{e}_k^{\mathbf{r}\cdot\text{T}} \boldsymbol{e}_l^{\mathbf{r}\cdot} = \delta_{kl} \end{pmatrix} \to \begin{pmatrix} \lambda^{\mathbf{r}\cdot}{}_a^{(0)} \\ \lambda^{\mathbf{r}\cdot}{}_a^{(1)} \end{pmatrix} . \tag{49}$$

Here the a-index reminds that we are interested in matching the first K eigenvalues only. We present our results on Figure 7. We found that the best matching between $\lambda^{\mathbf{r}\cdot}{}_a^{(1)}$ and λ_a^{Emp} occurs for $\mu = 1.4$. This enabled us to reproduce the empirical link between the ordered FCLs and the ordered (constrained and unconstrained) eigenvalues (the blue line on Figure 7 is close to the yellow points).

Notice that so far, we have had no need to generate the returns. This can be done by

$$r_i(t) = \sum_{j=1}^{N} \sqrt{\ell_j^{\mathbf{r}\cdot}} \left(\boldsymbol{e}_j^{\mathbf{r}\cdot}\right)_i \varepsilon_j(t) \,, \tag{50}$$

where both $\boldsymbol{e}_j^{\mathbf{r}\cdot}$ are $\ell_j^{\mathbf{r}\cdot}$ are defined as in (48) and $\varepsilon_j(t)$ are TN standard normal random variables. For $T \ll N$ the covariance matrix of these returns reduces to (48), but away from this regime these matrices are different. Starting from these returns we calculate the new correlation matrix, $\boldsymbol{C}^{\mathbf{r}\cdot}$, as well as its N unconstrained and K constrained eigenvalues with the $\boldsymbol{h}^{\mathbf{r}\cdot}$ and $\boldsymbol{\gamma}^{\mathbf{r}\cdot}$ matrices computed from (47). We demonstrate the output of this calculation on Figure 7.

From the bottom-left graph of Figure 7 we learn that the projection into the subspace generated by the K random factors helps to reduce the noises and the bias for $T \ll N$. That is also interesting (and somewhat intriguing) to see from the bottom-right plot that for $T \gg N$ the matching between the simulated constrained eigenvalues and the unconstrained eigenvalues of the population correlation matrix is much weaker than we could have expected from the empirical results of Table 5 of Section 4.2.

The discrepancy between our simulation and empirical measurements could be explained by the universal linear law that we could not have respected. To do so one needs to generate a random rotation matrix e_{ij} (44) whose j-row elements have a uniform distribution instead of the more natural "bell-like" distribution.

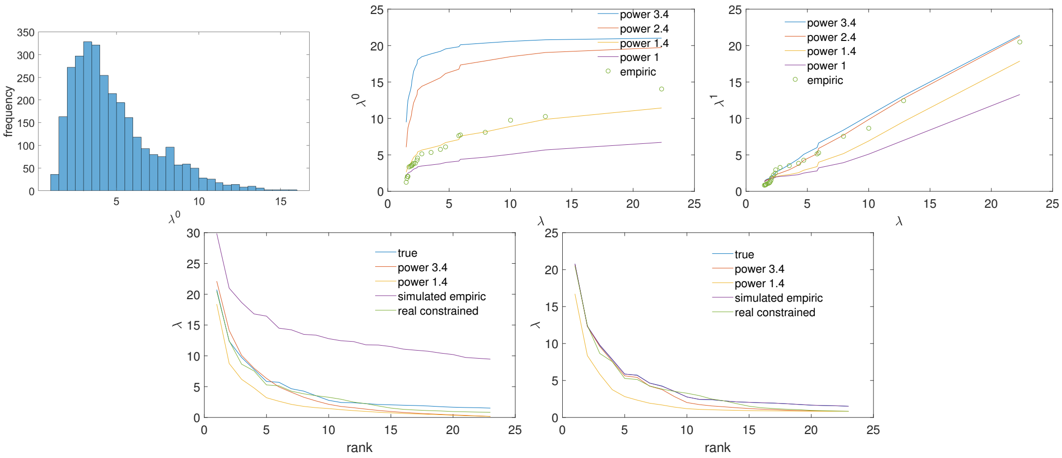

Figure 7: Top left: theoretical distribution of $\lambda_\star^{(0)}$ with $\mu = 1.4$. We simulate the $\lambda_\star^{(0)}$ distribution based on a random selection of signals. We suppose that the angle between the random factor and any unconstrained eigenvector is a Gaussian random variable with a standard deviation proportional to the square root of the unconstrained eigenvalue power μ. We apply a Monte Carlo simulation with $3,000$ trajectories. Top middle: within this model, we explain the relationship between the ordered values of $\lambda_\star^{(0)}$ and the true λ^{Emp} without the measurement noises. $\mu = 1.4$ is the best fit among $\mu = 1$, $\mu = 2.4$ and $\mu = 3.4$. Top right: within the model, we explain the relationship between the ordered values of $\lambda_\star^{(1)}$ and the true λ^{Emp} without the measurement noise. Surprisingly $\mu = 3.4$ and $\mu = 2.4$ are the best fit among $\mu = 1$, $\mu = 1.4$. $T = 50$ (Bottom left) and $T = 80,000$ (Bottom right) are used to replicate the measurement noise of the simulated empirical eigenvalues. $T = 80,000$ is the number of 5 minutes returns used in Table 5. In both cases ($T = 50$ and $T = 50,000$) the simulated constrained eigenvalues with $\mu = 3.4$ are the best estimation of the true eigenvalues. For $T = 50$ the simulated empirical eigenvalues overestimate the true eigenvalues very significantly whereas the simulated constrained eigenvalues with $\mu = 3.4$ remains close to the true eigenvalues until the 10th rank. For $T = 80,000$ the measurement noise is considerably reduced. The real constrained eigenvalues perform better than the simulated ones as they fit until the 15th rank. We suspect that the inconsistency ($\mu = 1.4$ fits the $\lambda_\star^{(0)}$ in the top middle plot but does not fit the $\lambda_\star^{(1)}$ whereas $\mu = 3.4$ does not fit the $\lambda_\star^{(0)}$ in the top middle plot but fits the other plots) is coming from a drawback of the random generation of factors that does not take into account the Maximum-Variance optimization and the 'universal law' that is an important pattern of the empirical correlation matrix and that should play an important role.

3 Empirical improvements

Here we describe different variations of the Maximum-Variance and benchmark portfolios tested in the paper. In the last subsection we summarise all of the proposed improvements in a single table.

3.1 Optimisation of the Sharpe ratio

We argue that under certain additional assumptions regarding the Two-Factor model (11), the Maximum-Variance portfolio (20) has also the largest Sharpe ratio under the market-neutral constraint that captures the signal.

We adopt the following conjectures:

- We suppose that the signal captures a positive risk premium and we thus have $\mathbb{E}_{t-1}\left(r^{\mathbf{f}}\right) > 0$, while residual returns are expected to have mean zeros, $\mathbb{E}_{t-1}\left(\epsilon_i(t)\right) = 0$.
- The covariance matrix of the residual returns takes the form $\boldsymbol{H}^{\epsilon} = k\boldsymbol{\Gamma}$, where k is a positive constant.[16] In other words, different ϵ_i's are uncorrelated, $\mathbb{E}_{t-1}\left(\epsilon_i\epsilon_j\right) = 0$, and their volatilities are proportional to the corresponding stock volatilities, $\mathbb{E}_{t-1}\left(\epsilon_i^2\right) = k\Sigma_i^2$.

For the sake of simplicity we will also ignore the time-dependence of the portfolio weights. With these assumptions and focusing meanwhile exclusively on a market neutral portfolio with weights ω_i, we find that:

$$\mathbb{E}_{t-1}\left(\sum_{i=1}^{N}\omega_i r_i\right) = \left(\boldsymbol{b}^{\mathrm{T}}\boldsymbol{\omega}\right)\mathbb{E}_{t-1}\left(r^{\mathbf{f}}\right) \quad \text{and} \quad \mathbb{E}_{t-1}\left(\left(\sum_{i=1}^{N}\omega_i r_i\right)^2\right) = \left(\boldsymbol{b}^{\mathrm{T}}\boldsymbol{\omega}\right)^2\left(\Sigma^{\mathbf{f}}\right)^2 + k\left(\boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{\Gamma}\boldsymbol{\omega}\right). \tag{51}$$

The Sharpe ratio is therefore:

$$S^{\mathbf{f}} \equiv \frac{\mathbb{E}_{t-1}\left(\sum_{i=1}^{N}\omega_i r_i\right)}{\sqrt{\mathbb{E}_{t-1}\left(\left(\sum_{i=1}^{N}\omega_i r_i\right)^2\right)}} = \frac{\mathbb{E}_{t-1}\left(r^{\mathbf{f}}\right)}{\sqrt{\left(\Sigma^{\mathbf{f}}\right)^2 + k\cdot\dfrac{\left(\boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{\Gamma}\boldsymbol{\omega}\right)}{\left(\boldsymbol{b}^{\mathrm{T}}\boldsymbol{\omega}\right)^2}}}. \tag{52}$$

Maximising this expression with respect to $\boldsymbol{\omega}$ subject to the market-neutrality condition $\boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{\omega} = 0$, we arrive at (20). We can conclude therefore that finding the highest Sharpe ratio of a market-neutral portfolio is equivalent to the FCL optimisation.

The same approach might be used as well if we weaken the second assumption above. We may presume instead that the covariance matrix of the residual returns has a small rank-1 off-diagonal market-neutral term: $\boldsymbol{H}^{\epsilon} = k\left(\boldsymbol{\Gamma} + \varepsilon\boldsymbol{u}^{\epsilon}\boldsymbol{u}^{\epsilon\mathrm{T}}\right)$. Requiring $\boldsymbol{\omega}$ both to be market neutral *and* orthogonal to $\boldsymbol{u}^{\epsilon}$, $\boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{u}^{\epsilon} = \boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{\beta} = 0$, the new solution is

$$\boldsymbol{\omega}_{\star} \sim \boldsymbol{\Gamma}^{-1}\left(\boldsymbol{b} - \frac{\boldsymbol{b}^{\mathrm{T}}\boldsymbol{\Gamma}^{-1}\boldsymbol{\beta}}{\boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{\Gamma}^{-1}\boldsymbol{\beta}}\cdot\boldsymbol{\beta} - \frac{\boldsymbol{b}^{\mathrm{T}}\boldsymbol{\Gamma}^{-1}\boldsymbol{u}^{\epsilon}}{\boldsymbol{u}^{\epsilon\mathrm{T}}\boldsymbol{\Gamma}^{-1}\boldsymbol{u}^{\epsilon}}\cdot\boldsymbol{u}^{\epsilon}\right). \tag{53}$$

Here the first two terms appear already in the Maximum-Variance market-neutral portfolio (20), while the last one guarantees the $\boldsymbol{u}^{\epsilon}$-orthogonality. This is a Maximum-Variance portfolio which is both market and $\boldsymbol{u}^{\epsilon}$-neutral, as it optimises the FCL under these two constraints.

3.2 Parameter ν

The formula (43) for $b_i(t)$ might be generalised to

$$b_i(t) = \Sigma_i^{\nu}(t)\mathfrak{B}_{\star}\left(q_i(t)\right), \tag{54}$$

where the original assumption (30) corresponds to $\nu = 1$, and $\mathfrak{B}_{\star}(q)$ is the same as in (43). For $\mathbb{E}(r^{\mathbf{f}}) > 0$, the Two-Factors model (11) implies that the excess expected return satisfies $\left(\mathbb{E}(r_i) - \beta_i\mathbb{E}(r^{\mathbf{m}})\right) \sim \Sigma_i^{\nu}\mathfrak{B}_{\star}\left(q_i(t)\right)$.

We will consider the following variations:

- $\nu = 0$. This is equivalent to assuming that the excess expected return of a single stock does not depend on its volatility. Alternatively, one may say that the risk is not fairly rewarded.

[16] The model would be called a *scalar* strict factor model, if $\boldsymbol{H}^{\epsilon}$ were proportional to the identity matrix.

- $\nu = 1$. This choice makes sense from the economic point of view and was also briefly justified in the paragraph above (30). It implies that the reward is proportional to the single stock volatility. There is a direct link between the $\nu = 1$ choice and the Maximum-Diversification portfolio of [11]. The main hypothesis behind the construction in [11] is the proportionality between the expected return of a stock and its volatility, $\mathbb{E}(r_i) \sim \Sigma_i$. Together with the central CAPM result, $\mathbb{E}(r_i) = \beta_i \mathbb{E}(r^{\mathbf{m}})$, it leads to $\beta_i \sim \Sigma_i$. The analogue of the Maximum-Diversification hypothesis for the excess return should be $(\mathbb{E}(r_i) - \beta_i \mathbb{E}(r^{\mathbf{m}})) \sim \Sigma_i \mathfrak{B}_\star (q_i(t))$, and according to the Arbitrage Pricing theory [26] the same excess return is proportional to the factor sensitivity, $(\mathbb{E}(r_i) - \beta_i \mathbb{E}(r^{\mathbf{m}})) \sim b_i$ provided $\mathbb{E}(r^{\mathbf{f}}) > 0$. Combining the two formulae we see that $b_i \sim \Sigma_i \mathfrak{B}_\star (q_i(t))$. We also argue that the Maximum Variance portfolio with $\nu = 1$ is equivalent to the *constrained* WLS regression (see Appendix D).

- $\nu = 2$. It describes the situation where most volatile stocks generating more reward than expected for $\nu = 1$. It also corresponds to the common-practice OLS regression but with *constraints*: comparing the two last columns of Table 17, we see that the $\nu = 2$ Maximum Variance portfolio is close to the *constrained* OLS regression with $\nu = 0$.

3.3 The "residual" as the alternative method of orthogonalisation

In Section 3.1 we presented (53) that is an extension of the Maximum Variance portfolio that is not only market neutral but also neutral to another main factor of risk, the control variable, that remains to define. It is usually one of the Fama and French popular *Book* or *Size* factors. The extended portfolio was derived to get the optimal Sharpe ratio. That motivates to develop a method to make factors as decorrelated as possible. We presented a methodology, in the Section 2.4, that transforms the initial risk premia factors into the constrained eigenvectors of the correlation matrix, that would be the natural orthogonalized risk premia. Unfortunately we show in the empirical validation part that these constrained eigenvectors present an unstable combination of initial risk premia factors. We believe that this instability is intrinsic to any methodology that claims to orthogonalize the factors. Here we present the alternative methodologies, one of which will be implemented and tested in Section 4.

Multiply-sorted portfolios are implemented in [3]. More complex schemes are used in Asset Pricing models to withdraw bias, when different characteristics are correlated [48, 49]. To solve the dependency problem betweens factors, an optimal procedure is proposed in [50] to find orthogonalized risk premia inspired from the methodology attributed to [51]. Normalising returns of factors by the square root of the covariance matrix is also used in agnostic risk parity [52]. The new orthogonalized risk premia may diverge significantly from the original ones. Many "orthogonalisation" methods, though popular in the Asset Management industry, are yet to be documented. One of the most famous ones is known as the *residual* method, since it tries to withdraw the common part directly from signals. Suppose that for any time t and stock i we are given the raw ranking $q_i^r(t)$ of the variable to be priced and $q_i^c(t)$ of the control variable. For fixed t, we regress $q_1^r(t), \ldots, q_N^r(t)$ against $q_1^c(t), \ldots, q_N^c(t)$ treating the regression residual as the new signal of the variable to be priced. The "orthogonalized" signal is therefore the residual ranking that is not explained by the control variable. This is ideologically close to the so-called *residual momentum strategy*, first presented in [53], where the regression analysis is performed instead on the *residual* returns rather than the rankings.

We select *Book*, *Size* and *Beta* as the three control variables. The *Book* and *Size* are justified, since according to Fama and French they are the best candidates to explain the cross-section of the expected returns. The *Beta* signal was selected because according to its empirical FCL it appears to be the most important factor. As we will see later, the empirical results are very disappointing. Our interpretation is that there is no way to determine which control variable to select as it requires to prefer some signals over the others. As we briefly discuss in 4, the hierarchy could be determined based on raw FCL though it may change very rapidly with time.

Notation	Model	ν	Number of clusters	Number of sectoral factors	Residual
MaxVar(1,6,9)	Maximum-Variance	1	6	9	
MaxVar(0,6,9)	—"—	0	6	9	-
MaxVar(2,6,9)	—"—	2	6	9	-
MaxVar(1,6,30)	—"—	1	6	30	-
MaxVar(1,1,9)	—"—	1	1	9	-
MaxVar(1,6,9,*Beta*)	—"—	1	6	9	*Beta*
MaxVar(1,6,9,*Book*)	—"—	1	6	9	*Book*
MaxVar(1,6,9,*Size*)	—"—	1	6	9	*Size*
BM-ERW(6,9)	Benchmark (equal-risk weighted)	-	6	9	-
BM-EW(6,9)	Benchmark (equal-weighted)	-	6	9	-

Table 2: Different methods of the portfolio construction tested in the paper. The equal-risk-weighted and equal-weighted benchmark portfolios were introduced in (9) and ν is the model parameter in (54), which is relevant for the Maximum-Variance models only. The next two columns correspond to the number of industries and sectoral factors used for signal sorting. The last column specifies the residual choice for the orthogonalisation method discussed in Section 3.3. It will be tested only for three factors listed in this table.

3.4 Sectoral constraints and sectoral factors

Most academic papers do not take into account the sectors as control variable, even though it may reduce noises of measured risk premia and increase Sharpe ratios. As we will show in the next section, common factors extracted from industry returns explain significant cross-sectional returns, even surpassing the explanatory power of *Book* and *Size*. We estimate that more than 40% of the unconstrained factors variance is explained by sectoral risk (difference between the FCL with or without sectoral constraint). This is in line with [54] which showed that the Sharpe ratio of the *Value* (*Book* in our conventions) factor is higher if the sectoral risk is completely withdrawn. Conversely, the *Momentum* premium is found to be better explained when the sectors are taken into account [55].

We thus decided to investigate sectoral factors (either 9 or 30), while maintaining our fourteen styles (*Book*, *Size*, etc.) sector-neutral. An additional incentive for our choice is the fact that most alternative risk premia vehicles are marketed as sector-neutral portfolios. Since the signals for the sector factors can only be binary (a company at question either belongs or does not belong to the sector) we can use neither the linear function $\mathfrak{B}_\star(q)$ of (43)) nor the step-function $\mathfrak{B}_\star(\mathfrak{q})$ of (42). Instead we adopt the following weight function for a given sector s and a company i:

$$\mathfrak{B}^s(i) = \begin{cases} 1 - \dfrac{n_s}{N} & \text{if the company } i \text{ belongs to the sector } s \\ -\dfrac{n_s}{N} & \text{if it does not} \end{cases}, \tag{55}$$

where n_s is the sector s size and $N = \sum_s n_s$ is the total number of stocks. Notice that the sum $\sum_i \mathfrak{B}^s(i)$ vanishes for all s, exactly as it does for $\mathfrak{B}_\star(q)$ and $\mathfrak{B}_\star(\mathfrak{q})$.

Our style portfolio construction is based on sorting with respect to the relevant signal. To impose sector-neutrality we rank the stocks separately within 6 industry clusters (more on this below) and then combine the rankings to get a single sorting list. In other words, six first-ranked stocks are to be followed by six second-ranked stocks, *etc.* Since we deal with a sufficiently large number of stocks, the way the stocks with identical rankings are ordered between them, is unimportant. These six clusters are explained in Appendix F. They are not optimised but rather inspired by the Global Industry Classification Standard (GICS), which also was the basis for our selection of 9 sectoral factors presented in the same appendix.

3.5 Financing, liquidity, turnover and leverage constraints

In this paper we have not considered any liquidity, turnover and leverage constraints. *Liquidity* is important in order to examine the full impact of trading. Turnover incurs brokerage fees and slippage. Leverage generates financial cost because of the gap between the borrowing and lending rates. All the more, we assume that *all* interest rates are zero, and so all stocks could be borrowed and be shorted with the zero interest rate. The Sharpe ratio of those factors that require more financing for the long legs than for the short legs, for instance the Low-Beta (*Beta* below) factor, should be sensitive to this assumption.

3.6 Summary of the improvements

In Table 7 we summarize different modifications of the Maximum-Variance and benchmark portfolios. For the former we test three values of the ν-parameter, 6 versus a single industry cluster, either 9 or 30 sectoral factors, and, finally, three different control factor selections. For the benchmark portfolios we explore the equal-risk and the equal-risk-weighted schemes, while keeping 6 clusters and 9 sectoral factors.

4 Empirical validation

The goal of this section is to test the ideas discussed so far with empirical data on stock returns.

4.1 Data

Our universe consists of 611 USA stocks selected in 2013, all of whom had capitalisation above one billion dollars back in the period. Between 2000 to 2013 the universe suffered from the *survivor bias*. It impacts only the daily data and the measured Sharpe ratio of the *Capitalisation* factor, overestimating it by 1. At the same time, it underestimates the *Momentum* Sharpe ratio by 0.2 and has negligible influence on market-neutral factors. By the end of November we expect to have the unbiased results for a rolling universe that replicates the historical constituents of the SP500 for the daily data.

We have two types of data:

- *Five*-minutes returns for the period 2013-2018 provided by *John Locke Investments*. The market relaxation time is believed to be around one minute [56], and thus the correlation should be sufficiently stable on the scale between five minutes and one day. Nevertheless, we detect a weak autocorrelation that changes the correlation matrix measurements as we move between the 1-day and the 100-days time-scales.
- *Daily* returns for the period 2000-2018 provided by *FactSet*.

As we have already explained in the previous section, stock's companies are divided into 30 industries according to their GICS classification, see Appendix F. They are further grouped into 6 clusters, as shown in the same appendix.

Finally, we have 14 signals for the styles summarised in Appendix B. The financial information is provided on daily basis by *FactSet*. Importantly, the data is accessible with a one-day shift. For all but one case in Table 2, the 14 styles are accompanied by exactly 9 sectoral factors. Since we always take into account the market mode, it amounts to capturing the first 24 eigenvectors of the empirical correlation matrix.

	Five-minutes returns		Daily returns	100 days returns
	conditional $T = 72 \cdot 5$	unconditional $T = 72 \cdot 5 \cdot 52 \cdot 6$	conditional $T = 255 \cdot 18$	unconditional $T = (255 \cdot 18)/100$
$\boldsymbol{\lambda}^{\mathrm{Emp}}$ $(N = 610)$	5.30	1.15	1.86	21.58
$\boldsymbol{\lambda}^{(1)}$ $(N = 24)$	1.58	1.03	1.15	2.97

Table 3: The Marčenko-Pastur threshold $\ell_+ = \ell_+(N,T)$ for $\boldsymbol{\lambda}^{\mathrm{Emp}}$ and $\boldsymbol{\lambda}^{(1)}$ with the number of available returns. Notice the high value of the threshold for the unconditional empirical estimation based on 100 days. The table may be easily extended to other time-scales of Tables 10 and 11, but here we report only the extreme values.

4.2 Measurements method

In the theoretical part of the paper we used the conditional expectations $\mathbb{E}_{t-1}(\cdots)$ to construct the covariance and other matrices. In practice we can only estimate these conditional expectations. A common practise is to use the $[t, t-\Delta t]$ period average of a given time-series as an estimate of the conditional expectation of the relevant quantity at time t. To get a better result one can use an Exponential Moving Average (EMA) with the parameter $\alpha^{-1} = 1/\Delta t$. As we explain in Appendix I for the 5-minutes returns the optimal averaging period is one week (5 days). As for the daily returns, we do not possess sufficiently long historic data to estimate conditional expectations of various eigenvalues. Nonetheless, we can use the daily data to find multi-day returns, which, as we explain below, lead to different results for the same eigenvalues. In what follows we will occasionally refer to conditional results as time-dependent, and to unconditional as time-independent.

We are interested in three different sets of eigenvalues: $\boldsymbol{\lambda}^{(0)}$, $\boldsymbol{\lambda}^{(1)}$ and $\boldsymbol{\lambda}^{\mathrm{Emp}}$. The first two are determined by the $\boldsymbol{h}$ and $\boldsymbol{\gamma}$ matrices as in (26) and (25) respectively. In Appendix I we explain in details how to estimate the conditional (time-dependent) and unconditional (constant) matrices $\boldsymbol{h}$ and $\boldsymbol{\gamma}$ based either on the 5-minutes or the (multi-)daily returns. In the last subsection of the same appendix we describe the derivation of the *empiric* eigenvalues $\boldsymbol{\lambda}^{\mathrm{Emp}}$. Our ultimate goal is to compare $\boldsymbol{\lambda}^{(1)}$ and $\boldsymbol{\lambda}^{\mathrm{Emp}}$. As we will explain shortly, the latter are our best shot at the eigenvalues of the "true" (also called population) correlation matrix. Obviously, the closer $\boldsymbol{\lambda}^{(1)}$ to $\boldsymbol{\lambda}^{\mathrm{Emp}}$, the better our approach. To avoid further confusion we will stick to the dictionary:

$$\boldsymbol{\lambda}^{(0)} : \text{FCLs} \qquad \boldsymbol{\lambda}^{(1)} : \text{Constrained eigenvalues} \qquad \boldsymbol{\lambda}^{\mathrm{Emp}} : \text{Empiric eigenvalues} \tag{56}$$

Here *constrained* refers to the fact that $\boldsymbol{\lambda}^{(1)}$ were obtained from the constrained diagonalisation, see Section 2.4.

Before proceeding to other quantities investigated in this section let us make an important comment. According to the celebrated paper [57] by Marčenko and Pastur, the eigenvalues ℓ_i of the matrix $\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}$, an $T \times N$ matrix of IID standard normal variables, are non-uniformly distributed between two parameters, ℓ_- and ℓ_+:

$$\ell_\pm = \left(1 \pm \sqrt{\frac{N}{T}}\right)^2 \tag{57}$$

provided both $N, T \to \infty$ with N/T hold fined. The larger this ratio, the stronger ℓ_i deviate from the "true" value of 1. In spite of the fact that our "true" eigenvalues are very different from 1 (in fact, we will see that the largest eigenvalue is of order $N/5$) and the distribution is far from the Gaussian one, we might still adopt

	Five-minutes returns		Daily and multi-daily returns	
	conditional	unconditional	conditional	unconditional
$\boldsymbol{\lambda}^{\text{Emp}}$	Figure 8	Table 5	-	-
$\boldsymbol{\lambda}^{(0)}$	Figure 8	Tables 6 and 7	-	Tables 8 and 9
$\boldsymbol{\lambda}^{(1)}$	Figure 8	Table 5	-	Tables 10 and 11
x_{ab}	Figure 9	Table 12	-	-
Sharpe ratio	-	-	Figure 10	Table 13

Table 4: Summary of the figures and tables of this section according to the data used (either 5-minutes or daily returns) and the time-dependence: the figures present conditional (time-dependent) variables and the tables are reserved for unconditional (constant) results.

ℓ_+ as the threshold between the significant and noisy eigenvalues. In what follows we refer to this value as the Marčenko-Pastur threshold.

For $\boldsymbol{\lambda}^{\text{Emp}}$ we always use the full collection of stocks, and so $N = 611$ as we already mentioned in the beginning of this section. For $\boldsymbol{\lambda}^{(1)}$, on the other hand, $N = K$ since we obtain these eigenvalues from the $\boldsymbol{\gamma}^{-\frac{1}{2}}\boldsymbol{h}\boldsymbol{\gamma}^{-\frac{1}{2}}$ diagonalisation in the subspace spanned by K Maximum-Variance portfolios. In this paper $K = 14\,(\text{styles}) + 9\,(\text{sectoral factors}) + 1\,(\text{market}) = 24$. The situation is slightly more complicated for T. For instance, for the unconditional measurement based on the 5-minutes returns we have on average 72 daily returns, 5 working days, 52 weeks per year and, finally, 6 six years of data. Thus $T \approx 1.1 \cdot 10^5$. At the same time, for the respective conditional calculation $T = 72 \cdot 5 = 360$, if we take one week as the averaging period to estimate the expectation. Conversely, for the daily returns $T = 255 \cdot 18$, because now we have 18 years of data. In Table 3 we summarise the Marčenko-Pastur (MP) thresholds for the empiric and constrained eigenvalues, and different evaluation schemes. The thresholds are important, as they indicate how many eigenvalues we should consider as non-noisy. There are two lessons we might learn from this table. First, the MP threshold for the unconditional empiric eigenvalues on the time-scale of 100 days is very large. We will see that it leaves out only the market mode. Second, the one-day threshold, though much lower, is still well above the unconditional threshold for the 5-minutes returns. We conclude therefore it is worth to evaluate the empiric eigenvalues *only* from the 5-minutes returns that will serve as a reference to evaluate how well our method reproduces the "true" eigenvalues. Consequently we will investigate the time-scale dependence of the eigenvalues using only the constrained eigenvectors method, see the second row of Table 3.

Apart from the eigenvalues we consider two additional measurements: the transformation matrix $\boldsymbol{O}$ from Maximum-Variance portfolio to constrained eigenvectors introduced in (25), and the Sharp ratio. Since the factors may capture different level of risk, it will be more illustrative to compare the rescaled components, $x_{ab} = O_{ab}\sqrt{\lambda_b^{(1)}}\sqrt{\lambda_a^{(0)}}$. In this paper the Sharpe ratio is measured as the ratio between the average and the standard deviation of the daily-returns based portfolio. We also multiply the expression by $\sqrt{255}$ in order to annualise the final result. To compare the Sharpe ratios for different factors we exploit only daily data from 2000 to 2017. We stick with the constant gross-investment normalisation, $\sum_i |\omega_i(t)| = \text{const}$, as it is commonly accepted in Asset Pricing. It was shown in [58] that due to the non-Gaussian distribution of returns, the normalisation by monthly look-back volatilities leads to the anomalies and the Sharpe ratios which are different from those derived with the constant gross-investment normalisation. We verified that all our observations regarding the Sharpe ratio still hold under the normalisation method of [58].

#	$\lambda_\star^{(1)}$, Maximum-Variance						$\lambda^{(1)}$		$\lambda_\star^{(1)}$	λ^{Emp}	%
	(1,6,9)	(0,6,9)	(2,6,9)	(1,1,9)	(1,6,9, *Beta*)	(1,6,9, *Book*)	BM-ERW (6,9)	BM-EW (6,9)	MaxVar (1,6,30)		
1	100.35	100.27	100.17	102.33	100.07	99.63	102.93	100.53	101.06	109.02	-3
2	20.49	20.17	20.43	21.37	20.42	20.33	18.40	17.68	20.88	22.32	11
3	12.44	10.94	11.69	13.31	12.40	12.25	11.76	10.45	12.62	12.84	6
4	8.65	7.92	8.56	9.12	8.62	8.52	8.43	8.24	9.19	10.01	3
5	7.55	7.48	7.55	7.74	7.55	7.39	7.02	6.78	7.63	7.94	8
6	5.28	4.83	5.19	6.38	5.28	4.87	4.97	4.52	5.62	5.92	6
7	5.14	4.69	5.06	5.67	5.13	4.73	4.78	4.20	5.32	5.79	8
8	4.27	4.04	4.25	5.24	4.27	4.22	3.64	3.53	4.37	4.70	17
9	3.84	3.73	3.82	4.88	3.83	3.45	3.24	2.90	4.00	4.29	19
10	3.52	3.40	3.47	4.36	3.48	3.13	2.69	2.67	3.85	3.53	31
11	3.28	3.20	3.28	3.69	3.22	2.98	2.51	2.38	3.30	2.77	30
12	2.95	2.91	2.96	3.17	2.92	2.40	2.30	2.22	3.10	2.42	28
13	2.49	2.42	2.49	2.89	2.49	2.34	2.08	1.98	2.63	2.38	20
14	2.24	2.28	2.21	2.47	2.22	2.09	1.79	1.68	2.42	2.23	25
15	1.86	1.90	1.84	2.07	1.98	1.86	1.58	1.51	2.16	2.10	18
16	1.52	1.44	1.50	1.96	1.62	1.62	1.35	1.34	1.99	2.03	13
17	1.31	1.28	1.30	1.69	1.47	1.50	1.19	1.18	1.85	1.96	11
18	1.20	1.16	1.18	1.44	1.25	1.26	1.03	1.05	1.83	1.92	16
19	1.11	1.10	1.09	1.40	1.11	1.17	0.97	0.96	1.74	1.84	14
20	1.05	1.04	1.04	1.23	1.05	1.05	0.92	0.94	1.70	1.73	14
21	0.95	0.94	0.94	1.12	0.96	0.95	0.89	0.89	1.64	1.64	7
22	0.91	0.88	0.90	1.08	0.91	0.93	0.86	0.85	1.49	1.60	6
23	0.87	0.87	0.87	1.04	0.87	0.92	0.83	0.82	1.45	1.56	5
24	0.84	0.83	0.82	0.92	0.83	0.88	0.80	0.81	1.38	1.50	5

Table 5: The 2013-2018 conditional eigenvalues obtained with different methods of Table 2. λ^{Emp}-column are the sample eigenvalues. The last column shows the improvement of the MaxVar(1,6,9) method compared to the standard benchmark equal-risk weighted portfolio. We see that Maximum-Variance increases all but the first eigenvalue, which corresponds to the market mode. Recall that to construct the Maximum-Variance market-neutral portfolio we used the stock index $r^{\mathbf{m}}(t)$ and its proxy $r_\star^{\mathbf{m}}(t)$, rather than the first principal component of the returns matrix. We believe owing to this approximation the equal-risk-weighted portfolio performs better than $\boldsymbol{\omega}^{\mathbf{m}.}{}_\star^{(0)}$. For λ_i^{Emp} $(i = 2, \ldots, 5)$ the Maximum-Variance improvement comes largely from the sectoral factors. On the other hand, the Sharpe ratio increases in the last column for eigenvalues ranging between 2 and 6 is due to the style factors, mostly *Book*. These issues are discussed in more details in Section 4.3.1. Among all Maximum-Variance methods, MaxVar(1,1,9) has the best results.

In Table 4 we list all figures and tables containing the measurements of the eigenvalues, the transformation matrix elements and the Sharpe ratio both conditional and unconditional and based on the 5-minutes or (multi-)daily returns.

Sectoral factor	$\lambda_\star^{(0)}$, Maximum-Variance				$\lambda^{(0)}$, Benchmark	
	(1,6,9)	(0,6,9)	(2,6,9)	(1,1,9)	BM-ERW(6,9)	BM-EW(6,9)
Utilities	14.04	13.65	14.03	14.04	12.99	12.74
Energy	10.27	9.08	9.25	10.27	9.81	8.67
Reits	9.76	9.43	9.69	9.75	9.59	9.53
Finance	7.74	7.09	7.64	7.74	7.87	7.70
Pharmacy	5.75	5.18	5.62	5.76	5.62	5.39
IT	5.33	4.97	5.22	5.33	5.22	5.02
Consumer	4.56	3.98	4.47	4.56	4.61	3.71
Discretionary vs Staples	3.83	3.68	3.84	3.83	3.47	3.03
Industry	3.53	3.56	3.57	3.53	3.14	2.97

Table 6: Different unconditional FCL ($\lambda^{(0)}$) for the sectoral factors obtained for the 2013-2018 period with different methods of Table 2. Improvement from the Maximum-Variance optimisation is very limited for sectoral factor as the signal is binary. The sectoral factors have higher FCL than the style factors. The major sectoral factor is *Utilities*, that is expected to be a highly leveraged sector, despite its small size, as it is used by traders to speculate on the FED policy that was the major issue during the period. Energy and REITS were also very volatile due to the oil price decline and the crisis of Malls.

4.3 Reproducing the true eigenvalues

From Table 5 we see that the Maximum-Variance portfolios capture well the "true" empirical eigenvalues measured with 5-minutes returns of the average correlation matrix: $\boldsymbol{\lambda}_\star^{(1)}(\mathrm{MaxVar}(1,6,9))$ are very close to the $\boldsymbol{\lambda}^{\mathrm{Emp}}$. The 24 first unconstrained eigenvectors explain $R^2 = 41.48\%$ of the cross-section regression of normalized returns ($R^2 = \left(\sum_{k=1}^{K} \lambda_k^{\mathrm{Emp}}\right)/N = 41.48\%$) whereas the 24 first constrained eigenvectors $\boldsymbol{\omega}_\star^{(1)}(\mathrm{MaxVar}(1,6,9))$ obtained with the Maximum-Variance optimisation (MaxVar(1,6,9)) explain 37.62% (40.03% when sector constraints are withdrawn, MaxVar(1,1,9)). We find also a significant difference between $\lambda_\star^{(0)}(\mathrm{MaxVar}(1,6,9))$ and $\lambda_\star^{(1)}(\mathrm{MaxVar}(1,6,9))$ (see Tables 7 and 6). This is the impact of the strong interaction between the economic factors (see Section 2.4). Moreover, all market-neutral factors have FCLs much smaller than the FCL of the market mode (70), which is close to 100. This is consistent with our choice to have K Two-Factor models for all the factors different from the market rather using a single multi-factor model, see the last paragraph of Section 2.2.

We observe as well that the major market-neutral factors $\boldsymbol{\omega}_\star^{(0)}(\mathrm{MaxVar}(1,6,9))$ are mostly the sectoral factors: *Utilities*, a small size sector, has surprisingly the highest $\lambda_\star^{(0)}(\mathrm{MaxVar}(1,6,9))$. This could be specific to the sample period when the monetary policy change played a crucial role. Traders used to speculate on shorting the *Utilities*, believed to be highly indebted, in the goal to take positions on an upward interest move.

Presence of *Utilities* in the first eigenvector was confirmed by [59] in which *Finance*, *Oil* and *Utilities* were found to be the main components of this eigenvector. We also see *Reits* as an important factor most certainly due to the Malls crisis in 2017. We find that the two most popular risk premia, *Momentum* and

Style factor	$\lambda_\star^{(0)}$, Maximum-Variance						$\lambda^{(0)}$, Benchmark	
	(1,6,9)	(0,6,9)	(2,6,9)	(1,1,9)	(1,6,9,*Beta*)	(1,6,9,*Book*)	ERW(6,9)	EW(6,9)
10Y Rates	8.10	8.58	8.42	14.62	7.76	6.94	4.38	3.76
Beta	7.63	7.71	7.76	12.72	7.63	7.52	4.82	4.47
Momentum	6.10	5.69	5.81	9.39	6.09	4.29	4.39	3.95
Capitalisation	5.14	4.79	4.99	6.08	5.14	2.21	4.00	3.65
Dividend	4.28	4.59	4.45	6.01	4.27	4.27	2.47	2.16
Euro	3.85	3.97	3.89	6.82	3.86	3.36	2.49	2.27
Liquidity	3.70	3.23	3.17	6.24	7.03	6.73	3.13	3.00
Book	3.53	3.27	3.31	5.74	3.54	3.54	2.75	2.59
STR	3.44	3.46	3.42	5.51	3.39	3.05	2.29	2.09
Sales	3.35	3.04	3.28	4.58	3.36	1.53	2.71	2.45
Leverage	2.05	2.25	2.12	2.99	2.05	2.31	1.27	1.22
Earning	2.04	2.01	2.06	3.12	2.04	2.03	1.48	1.45
Cash	1.79	1.92	1.85	2.50	1.78	1.71	1.19	1.14
Growth	1.24	1.16	1.17	1.84	1.24	1.24	1.06	1.05

Table 7: The 2013-2018 unconditional FCL obtained with different methods of Table 2 and for all styles. The six $\lambda_\star^{(0)}$-columns correspond to the Maximum-Variance method and so we omit the common part in the method references. Similarly, we save space for the last two benchmark columns. Comparing the MaxVar(1,6,9) and the BM-ERW(6,9) columns, one might notice an improvement of 45% thanks to the Maximum-Variance optimisation. We also see that the sectoral constraints are suboptimal: MaxVar(1,1,9) appears to be the best method. But to maintain the sector constraints, on average MaxVar(1,6,9) appears the best. The model $\nu = 1$ appears to be most realistic for most factors even if some exception could be real. The most important style factors are the 10Y Rates (stocks that are the most sensitive positively to the interest rate increase vs stock that are less) as the FED policy was a major issue on the period. The *Beta* and *Momentum* factors are the two other major factors whereas the traditional Fama & French factors, *Capitalisation* and *Book*, seem to be far less important.

Beta, are the 6th and the 7th risk factors in $\lambda_\star^{(0)}(\text{MaxVar}(1,6,9))$. The *Book* factor has a surprising low $\lambda_\star^{(0)}(\text{MaxVar}(1,6,9))$ at a 5-minutes time scale whereas it is supposed to be an important factor at least at a longer time scale based on the known results in the literature. The *Growth* factor is very close to noise as its FCL is just slightly above an FCL of a random signal. We see that Maximum-Variance approach manages nevertheless to improve its FCL.

4.3.1 Comparison between the Maximum-Variance and the benchmark portfolios

For the 14 styles the Maximum-Variance optimisation allows to improve $\lambda^{(0)}$ by 45% compared to the benchmark portfolios (see Table 7). As we have already mentioned above the improvements for the sectoral factors are significantly weaker. This is actually not too surprising if we think of the Maximum-Variance function $\mathfrak{B}_\star(\mathbf{q})$ as the "smoothed" version of the double Heaviside function used for the benchmark portfolio construction (9). No such "smoothing" option exists for a sectoral factor, whose signals are strictly binary. Thus the linear function optimisation has a weaker impact.

Factor	1 day	5 days	10 days	20 days	40 days	80 days	100 days
Beta	6.35	7.99	8.77	9.34	10.37	11.64	14.45
STR	4.10	4.73	5.00	5.51	6.72	8.45	12.21
Momentum	5.87	7.09	7.64	8.20	9.00	9.60	10.80
Capitalisation	4.43	4.95	5.24	5.66	6.48	7.53	9.54
Book	2.49	2.94	3.40	3.99	4.89	5.50	6.40
Sales	2.88	3.47	3.85	4.19	4.85	5.26	6.30
Dividend	3.86	4.44	4.90	5.28	5.77	5.83	5.79
10Y Rates	5.38	6.24	6.26	6.06	5.80	5.80	5.42
Liquidity	2.82	2.93	3.12	3.34	3.77	4.11	4.91
Euro	3.72	4.12	4.15	4.08	3.82	3.74	3.55
Leverage	2.08	2.30	2.54	2.72	2.93	3.10	3.41
Earning	1.89	2.09	2.20	2.24	2.40	2.59	3.17
Cash	1.52	1.66	1.72	1.73	1.85	2.04	2.46
Growth	1.47	1.67	1.76	1.78	1.92	2.00	2.00

Table 8: The FCL of the Maximum Variance MaxVar(1,6,9) at different time scales are optimised by 47% as compared to the benchmark BM-ERW(6,9). Improvement are 41% for 1 day scale to 49% to 100 days scale. Based on daily data from 2000 to mid 2018.

This observation provides a good explanation for the last column of Table 5. The sectoral factors contribution is more significant for the leading (the first four) eigenvalues, while it decreases for the intermediate ones. As a result our Maximum-Variance approach is more successful for $\lambda_i^{(1)}$, with $i > 4$. The $\lambda^{(0)}$ improvement is more significant for longer time scales as one can see from Tables 8 and 9.

Surprisingly, the first constrained eigenvalue for the benchmark portfolios, $\lambda^{(1)}\,(\text{BM-ERW}(6,9))_1$, is slightly higher than the first Maximum-Variance eigenvalue $\lambda_\star^{(1)}(\text{MaxVar}(1,6,9))_1$ meaning that it is slightly better to model the market-mode portfolio as an equal-risk-weighted portfolio (9) than the Maximum-Variance market-mode portfolio (70). This is also evident from the fact that the FCL of the Maximum-Variance market-mode portfolio $\lambda^{\mathbf{m}}{}_\star^{(0)}(\text{MaxVar}(1,6,9)) \approx 92.12$ is slightly lower than the FCL of the equal-risk-weighted portfolio $\lambda^{\mathbf{m}(0)}\,(BM - ERW(6,9)) \approx 92.44$. We believe that it is related to our choice to stick with the β's derived from the stock index $r^{\mathbf{m}}(t)$, rather than the Maximum-Variance return $r_\star^{\mathbf{m}}(t)$ (4). We discussed in details the difference between the two available indices in the paragraphs preceding (3).

The upshot of the last two paragraphs is that R^2 of the cross-section regression with the 24 factors is improved only slightly from 36.23% with the benchmark to 37.62% with the Maximum-Variance portfolios. With longer time scale the improvement is even more significant (10% for 100 days) and the R^2 is higher as well (59.44% for 100 days) (Tables 10 and 11). This is thanks to the quarter returns being stronger correlated than the daily returns. We see that FCL for the styles are always optimised for the Maximum-Variance portfolios compared to the benchmark portfolios for any time scale (Tables 8 and 9) and interestingly the FCLs are increasing with the time scale. The *Book* factor becomes the fifth most important factor after *Beta*, *Capitalisation* and *Momentum* at 100-days time scale.

Factor	1 day	5 days	10 days	20 days	40 days	80 days	100 days
Beta	4.13	4.98	5.44	5.94	6.82	7.90	10.12
STR	2.94	3.23	3.37	3.64	4.29	5.29	7.55
Momentum	4.27	4.91	5.18	5.49	5.97	6.39	7.31
Capitalisation	3.25	3.61	3.89	4.29	4.98	5.86	7.24
Book	1.80	1.97	2.22	2.58	3.11	3.39	3.79
Sales	2.36	2.80	3.13	3.45	3.98	4.26	5.10
Dividend	2.51	2.75	2.99	3.16	3.47	3.50	3.35
10Y Rates	3.07	3.42	3.35	3.26	3.17	3.12	2.73
Liquidity	2.18	2.19	2.31	2.49	2.85	3.03	3.39
Euro	2.47	2.64	2.67	2.62	2.47	2.48	2.49
Leverage	1.43	1.52	1.62	1.77	1.95	2.12	2.45
Earning	1.39	1.49	1.57	1.59	1.69	1.81	2.18
Cash	1.13	1.17	1.20	1.23	1.36	1.51	1.79
Growth	1.17	1.26	1.32	1.35	1.50	1.57	1.52

Table 9: The FCL of the benchmark BM-ERW(6,9) at different time scales. Based on daily data from 2000 to mid 2018.

Different versions of the Maximum-Variance portfolios MaxVar(1,6,9), MaxVar(0,6,9) and MaxVar(2,6,9) give similar results as volatilities Σ_k are not heterogeneous enough.

We see that the sectoral constraints appear to be highly suboptimal as the solution without any sectoral constraints MaxVar(1,1,9) gives higher FCL $\lambda_\star^{(0)}(MaxVar(1,1,9))$ and higher constraint eigenvalues $\lambda_\star^{(1)}(MaxVar(1,1,9))$ (Tables 7 and 5).

The application of the residual methods (columns MaxVar(1,6,9,*Beta*) and MaxVar(1,6,9,*Book*) of Table 7) is surprisingly efficient only for the following three factors: *Liquidity*, *Momentum* and *Capitalisation*. The immediate interpretation is that the signals of these three factors are strongly correlated to the *Book* signal. As for the $\lambda_\star^{(1)}$ eigenvalues in Table 10, the variations MaxVar(1,6,9,*Beta*), MaxVar(1,6,9,*Book*) and MaxVar(1,6,9,*Size*) produce almost the same results as the Maximum-Variance method MaxVar(1,6,9). We conclude, therefore, that the *residual* method does not bring any major improvement. Apart from that, as we discussed at the end of Section 3.3, the control variable selection has the drawback of breaking the rotational symmetry between the K factors.

4.3.2 Capturing the dynamics of eigenvalues and eigenvectors

Up to this point we discussed only the means of the empirical and modelled correlation matrices and their eigenvalues. In this section we would like to focus instead on the dynamics (time-dependence) of the eigensystems.

The Maximum-Variance FCL, $\lambda_\star^{(0)}$, resonates stronger with the factor volatility jumps than the relevant benchmark portfolio. We demonstrate this phenomenon on the top three graphs of Figure 9. The strong

Eigenvalue	1 day	5 days	10 days	20 days	40 days	80 days	100 days
1	152.08	164.58	163.63	166.58	164.34	165.96	174.74
2	17.05	20.46	22.37	23.48	24.45	25.08	27.76
3	12.67	13.41	14.09	14.80	15.91	17.00	21.24
4	8.68	10.12	10.60	11.59	12.57	14.16	15.78
5	7.31	7.76	7.97	8.07	8.67	9.70	11.72
6	5.78	6.67	6.93	7.38	7.86	8.36	9.41
7	5.21	5.76	6.12	6.53	7.14	8.08	8.04
8	4.67	5.00	5.51	5.71	6.37	6.60	6.20
9	4.07	4.28	4.50	5.01	5.27	4.92	4.56
10	3.88	4.21	4.12	4.13	4.47	4.51	4.39
11	3.41	3.62	3.61	3.53	3.64	3.75	3.64
12	2.96	3.12	3.02	3.18	3.59	3.62	3.30

Table 10: $\lambda_\star^{(1)}(\text{MaxVar}(1,6,9))$ from the Maximum-Variance is a good proxy for true eigenvalue of the correlation matrix. We see that correlation increase sharply with time scale and that the optimisation is working for any time scale. Indeed the (MaxVar(1,6,9)) optimised by an average of 23% compared to BM-ERW(6,9). The improvement is 16% in daily scales to 27% in 100-days scales. The trace is increased by 10% at the longer time scale. Daily data from 2000 to mid 2018

resonance helps the constrained eigenvalues $\lambda_\star^{(1)}$ to capture well the dynamics of the first empirical eigenvalues λ^{Emp}, see the bottom three graphs of Figure 9.

For large dimensions the first eigenvalue of a correlation matrix is linked to the average of its off-diagonal elements. In financial terms it implies that the high volatility of the first (market-mode) eigenvalue might be interpreted as the increased correlation between single stocks. Both the empirical and Maximum-Variance eigenvalues reproduce this behaviour as one can see on Figure 9. Moreover, the dynamics of the two eigenvalues (the constrained and the unconstrained) are very close to each other. The same holds for the second and the third eigenvalues.

The first eigenvector is well known as the market mode but the second eigenvector has always remained difficult to interpret according to the literature. Our findings clarify the origin of this problem. The largest components of the 2nd eigenvector come from the factors with the highest FCLs. The latter, however, are very volatile, and so are the components of the second eigenvector. In general the eigenvalues of a large random matrix are expected to be repulsive and so no crossover phenomenon usually happens. What we observe nevertheless is the crossover of the factor FCLs and their components in the second and third eigenvectors. For example, as soon as the *Financials* sector FCL exceeds *Utilities*'s FCL, the components of the second eigenvector change accordingly. This is shown on Figure 9. *Beta*, *Rates* and *Utilities* are the main contributors to the second eigenvector in the period from 2013 to 2018 (Table 12 and Figure 9).

4.4 Evidence of alternative risk premia

4.4.1 Improvement of the Sharpe ratio

The measured Sharpe ratio is improved thanks to the Maximum-Variance optimisation as expected theoretically for the factors with the most significant risk premia within the sample period. In other words, for

Eigenvalue	1 day	5 days	10 days	20 days	40 days	80 days	100 days
1	155.83	166.85	165.11	167.72	165.34	167.39	176.27
2	14.56	16.76	18.10	18.82	19.42	19.60	20.95
3	12.29	12.59	13.14	13.76	14.26	14.50	15.69
4	8.17	9.04	9.15	9.55	10.27	11.60	13.47
5	6.58	6.66	6.83	6.91	7.52	8.42	10.12
6	4.92	5.21	5.40	5.84	6.40	6.84	6.98
7	4.69	5.03	5.31	5.49	5.52	6.11	6.80
8	4.08	4.35	4.70	4.87	5.33	5.42	5.14
9	3.51	3.48	3.67	4.02	4.15	4.05	3.70
10	2.93	2.90	2.96	3.12	3.51	3.49	3.29
11	2.67	2.80	2.70	2.69	2.89	2.86	2.81
12	2.43	2.34	2.28	2.49	2.75	2.79	2.47

Table 11: $\lambda^{(1)}$ (BM-ERW$(6, 9)$) calculated for the daily returns from 2000 to mid 2018.

those factors the value $\mathbb{E}(r^{\mathbf{f}})$ in our Sharpe ratio (52) is sufficiently large. These factors are *STR*, *Liquidity* and *Cash*. We present the results in Table 13 and Figure 10. On the other hand, for the three most popular factors, *Momentum*, *Book* and *Capitalisation*, the value $\mathbb{E}(r^{\mathbf{f}})$ is too small, the benchmark Sharpe ratio is weak (or even negative) and therefore we cannot test our optimisation method for this period. Indeed, the Sharpe ratio during these 18 years is only 0.51 annualized for *Book* and so is not even actually statistically significant (t-statistics is $0.51\sqrt{18} \approx 2$). The Sharpe ratio for *Beta* is only 0.34 and is even negative for *Momentum*. The three factors are nevertheless the most popular risk premia: *Beta* for quality, *Momentum* for trend and growth and *Book* for value. According to most of the references on market anomalies and asset pricing in Table 14 these factors are substantially profitable but on a much longer period (usually since 1960), though even this claim is controversial. The 2008 crisis generated exceptional losses to the *Beta* and *Momentum* factors, although these losses are not representative for a longer historical period (see Figure 10). The factor returns are highly non-Gaussian with extreme losses accumulating into a short period and t-statistics that are common test in asset pricing should be interpreted with caution. Consequently, the Sharpe ratio estimation is very sensitive to the portfolio normalisation. At the end of Section 4.2 we already mentioned two possible ways to normalise the portfolios.

To summarize, the theoretical Sharpe ratio improvement in the framework of the Maximum-Variance optimisation could not be confirmed empirically in a conclusive manner, since the 20 years period is too short of a sample to produce statistically significant results. This is not really disappointing. To verify empirically a (very optimistic) annual Sharp ratio improvement of 0.2 with $t_{\text{stat}} > 2$, we would need at least 100 years of daily data. Despite all this we firmly believe that upon the assumption that there is a substantial alternative risk premium, the Maximum-Variance portfolio has a higher expected Sharpe ratio than the standard 20% top-bottom portfolio.

4.4.2 Skewness, Leverage effect and alternative risk premia

Based on our measurements the daily returns skewness is not necessarily significant. It is negative (-0.44) for the market mode. *Momentum*, *Beta* and *Liquidity* exhibit negative skewness (-0.35, -0.39, -0.08 respectively)

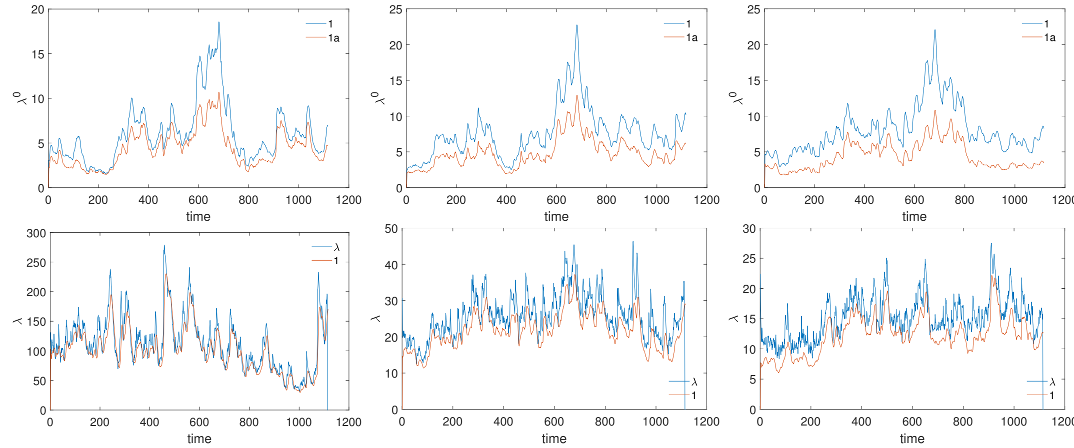

Figure 8: Top: measure from 2013 to 2018 of the $\lambda_\star^{(0)}(\mathrm{MaxVar}(1,6,9))$ and $\lambda^{(0)}$BM-ERW$(6,9)$ for *Momentum*, *Beta* and *10Y Rates*. $\lambda_\star^{(0)}(\mathrm{MaxVar}(1,6,9))$ correspond to the optimal case (Maximum-Variance (1)) and $\lambda^{(0)}$ (BM-ERW$(6,9)$) to the benchmark case (top-bottom 20% (4)). We see that the optimal case enters easier in resonance like in April to August 2016 where the three factors were excited. Bottom: measure of the first three eigenvalues $\lambda_\star^{(1)}(\mathrm{MaxVar}(1,6,9))$ and λ^{Emp}. λ^{Emp} corresponds to the sample but noisy eigenvalue without any constraint. $\lambda_\star^{(1)}(\mathrm{MaxVar}(1,6,9))$ corresponds to the constrained eigenvalues using the Maximum-Variance optimisation (MaxVar(1,6,9)) . It appears that $\lambda_\star^{(1)}(\mathrm{MaxVar}(1,6,9))$ looks to be less noisy and be a good proxy of de-noised λ^{Emp}. We see how brutally the first eigenvalues increased in February 2018. We also see hat the excitation of the second eigenvalue in April to August 2016 corresponds to the excitation of *Momentum*, *Beta* and *10Y Rates* of the Top graphs. We also see a spike in July 2017 in the third eigenvalue with an interaction with the second one.

while *STR*, *Cash*, *Capitalisation* and *Book* factors all have positive skewness $(0.79, 0.18, 0.12, 0.05$ respectively). By the Central Limit Theorem argument the skewness is expected to decrease as (time scale)$^{-2}$.

the negative skewness of factor returns may justify theoretically and empirically the presence of alternative risk premia (see, for example, [60]). In short, investors prefer to combine occasional strong gains with frequent small losses. This translates into a positive skewness of the returns distribution. Once the skewness becomes negative, the same investors would like to have an alternative risk premia to compensate for the unattractive risk profile.

The Leverage Effect (LE), that is negative correlation between returns and volatility variation, is a well-known phenomena in stock market. That generates high negative skewness for a large variety of time scales [61, 62, 63]. We believe that for any portfolio this is more natural to study LE by analysing the variation of portfolio's FCL rather than its volatility (or variance) variation. The FCL is the $\boldsymbol{\Gamma}$-normalised variance of portfolio's returns. As an example, for the Maximum-Variance market portfolio the FCL variation accounts for the variation of the average correlation between single stocks rather than the variation of the average single stock volatility. Moreover, in this case, the FCL variation is a good proxy of the variation of the correlation matrix first eigenvalue, see Section 4.3.2.

To analyse LE for different factors, we start with the $\boldsymbol{\Gamma}$-normalised Maximum-Variance portfolio of a given factor. We then regress the monthly variations of its FCL against its monthly returns. The higher the coefficient of determination, R^2, the stronger the evidence for LE.

We found that the market mode is the only factor that exhibits a significant LE with $R^2 = 0.38$. All other

Factor	$\boldsymbol{\omega}_\star^{(1)}(\text{MaxVar}(1,6,9))_1$	Factor	$\boldsymbol{\omega}_\star^{(1)}(\text{MaxVar}(1,6,9))_2$
Market Mode	-9.17	Utilities	2.63
Capitalisation	0.10	10Y Rates	-1.65
Beta	0.07	Beta	-1.49
Energy	-0.07	Reits	1.37
Liquidity	-0.07	Energy	-0.92
Momentum	0.07	Dividend	0.73

Table 12: A composition measured from 2013 to 2018 in risk of first and second conditional constrained eigenvector $\boldsymbol{\omega}_\star^{(1)}(\text{MaxVar}(1,6,9))_1$, $\boldsymbol{\omega}_\star^{(1)}(\text{MaxVar}(1,6,9))_2$ obtained through the Maximum-Variance optimisation. We present only the 6 highest risk contributions. The constrained first eigenvector of the average matrix is exposed to the market mode risk and the capitalisation whereas the second eigenvector is exposed to the utilities, rates and beta factors. But if we focus on the eigenvectors of the conditional weekly matrix, we would see that the second eigenvector is changing and that the beta factor arrived in the top position for the first and second eigenvectors.

market-neutral factors do not exhibits any LE as $R^2 < 0.08$, see Table 13. Without any LE, the skewness is expected to converge quickly to zero at larger time-scales and that could challenge the theory of alternative risk premia.

5 New open problems

Here we summarize some results and open problems relevant for the realistic correlation matrix modelling given relatively precise measurements realized by means of our economics constraints filter:

- The first eigenvalue of the correlation matrix of $\boldsymbol{\gamma}^{-\frac{1}{2}}\boldsymbol{h}\boldsymbol{\gamma}^{-\frac{1}{2}}$ is weakly volatile (see the yellow line on Figure 11) and its dynamics seems to be governed by a systematic factor and the market mode FCL, $\lambda^{\mathrm{m}}{}_\star^{(0)}$, appears to be a good candidate. This first eigenvalue is relatively stable for different time periods and may be interpreted as the average correlation between fundamental factors as if the position overlaps were completely suppressed. The factor overlap (blue line on Figure 11) moves only moderately with time and seems to be correlated with the momentum performance.
- The time dependence of $\ln(\lambda^{(0)})(t)$ can be modelled by an Ornstein–Uhlenbeck process with a relaxation period of 60 days. It is tempting to model the correlation matrix diffusion by the FCLs diffusion, while keeping constant the correlation matrix of $\boldsymbol{\gamma}^{-\frac{1}{2}}\boldsymbol{h}\boldsymbol{\gamma}^{-\frac{1}{2}}$, and then to compare the patterns with those of the classical Wishart process [64].
- It will be interesting to use the autocorrelation model of [65, 66] to reproduce the time scale dependency of $\lambda^{(0)}$ and $\lambda^{(1)}$ (see Figures 10 and 11). This autocorrelation model introduces a drift following an Ornstein–Uhlenbeck process that might be justified by the lack of liquidity and the herding effect. As a consequence, the moving average of the factor returns may serve a good proxy for the *conditional* expected returns, while the *unconditional* ones remain zeros. The model captures inefficiency in the stock market, which is yet to be documented. It is different from the *Epps effect* [67, 56] that identifies stocks lags at the intra-day time scale. The autocorrelation of [65, 66] could be more robust than the classical anomalies describing the discrepancy between the measured unconditional expected factor

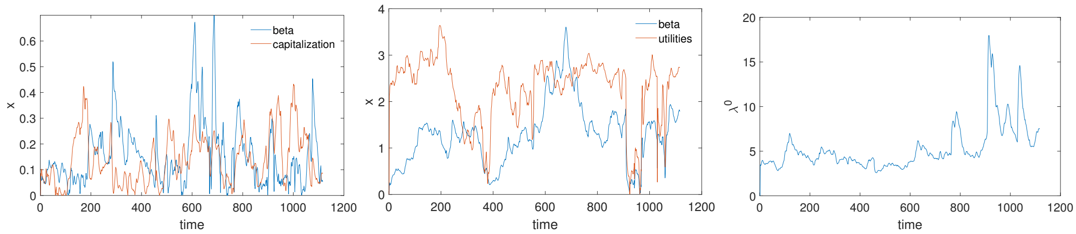

Figure 9: Left and Centre: A composition in risk from 2013 to 2018 of the first constrained eigenvector ($\boldsymbol{\omega}_\star^{(1)}(\mathrm{MaxVar}(1,6,9))_1$) and second conditional constrained eigenvector ($\boldsymbol{\omega}_\star^{(1)}(\mathrm{MaxVar}(1,6,9))_2$). We see that the interaction between the market mode, *Capitalisation* and *Beta* factors makes the first constrained eigenvector oscillate around the market mode, while the second eigenvector is exposed to the *Beta* factor despite its relatively low FCL. In July 2017, a new risk factor, the finance factor, replaced *Beta* and *Utilities*. The increases of the FCL of the financial factor appears at the same period (Right).

returns and the theoretical CAPM. Recently it has been argued in [23] that certain stock market anomalies become weaker after a study describing it has being published. In the same spirit most of the known anomalies were claimed to be fallacious and rather explained by over-fitting or selection bias [24].

6 Conclusion

We introduced the Maximum-Variance optimisation to build Maximum-Variance portfolios that capture as purely as possible the different signals used for extracting risk premia. We introduced the factor correlation level that the Maximum-Variance portfolio is optimising at any time scale. The Sharpe ratio under certain assumptions is also optimised and Maximum-Variance portfolios capture as best as possible the de-noised eigenvalues of the correlation matrix. An empirical test confirms the improvement from 5 minutes to 100 days time scales. The Maximum-Variance optimisation could therefore be used to reduce the dimension and to model and filter in a proper way the correlation matrix. The Maximum-Variance optimisation opens new problems to solve in the model of the correlation matrix around the dynamics and time scale dependency of eigenvalues and eigenvectors.

7 Acknowledgements

A PCA and Linear Regression

General idea

In this appendix we elaborate the connection between the principal component analysis (PCA) and the (ordinary, weighted or generalized) least squares approaches in a linear regression model. The results presented here appear rather scattered in the literature.

Let $\boldsymbol{Z}$ be an $m \times n$ matrix for $m \geqslant n$. We present it in a form

$$\boldsymbol{Z} = \boldsymbol{x}\boldsymbol{y}^{\mathrm{T}} + \boldsymbol{E}\,, \tag{58}$$

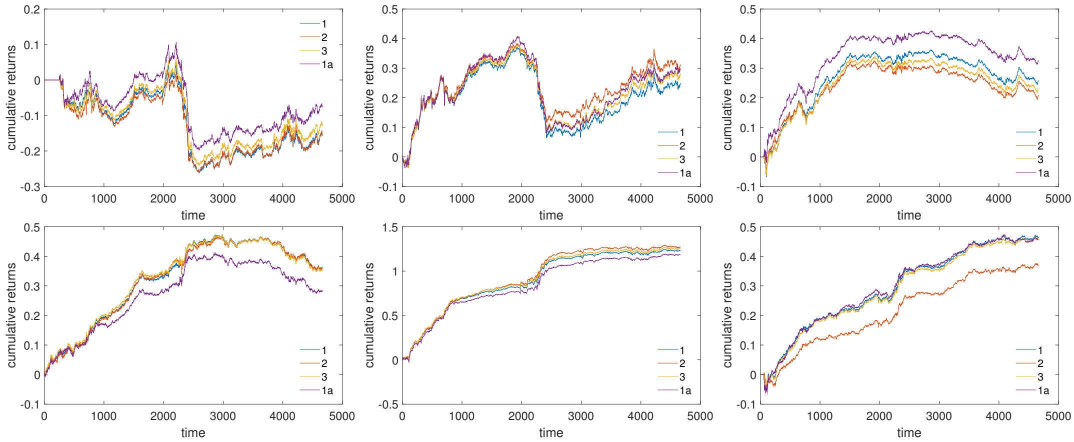

Figure 10: Cumulated gains at the same volatility for the *Momentum*, *Beta* and *Book* factors (top), and for the *Liquidity*, *STR* and *Cash* factors (bottom) from 2000 to May 2018. The three graphs on the top cover the most popular risk premia: *Beta* for quality, *Momentum* for trend-and-growth and *Book* for value. These factors were shown to be significantly profitable, but on a longer period (usually since 1960). The 2008 crisis generated exceptional losses to the beta and momentum factors that are not representative of a longer historical period. Bottom: the most significant factors on the period from 2000 to 2018: *Liquidity*, *STR* and *Cash*. We tested the Maximum-Variance optimisation with MaxVar(1,6,9), MaxVar(0,6,9), MaxVar(2,6,9) and BM-ERW(6,9). We see that the Maximum-Variance MaxVar(0,6,9) overperforms slightly for *Beta* factor but the Benchmark BM-ERW(6,9) overperforms slightly for *Momentum* and *Book* factors whereas the Maximum-Variance is theoretically expected to be the optimal solution for the Sharpe ratio. Our interpretation is that the backtest is noisy and not significant enough as the period is too short based on the weakness of different risk premia. Nevertheless the Maximum-Variance optimisation is confirmed empirically when risk premia is strong enough (Cash, STR and Liquidity factor) and if investors are convinced that a risk premium could actually exist, they should use the FCL (empirical or theoretical) to determine the best way to capture it.

where $\boldsymbol{x}$ and $\boldsymbol{y}$ are m-by-1 and n-by-1 vectors respectively, and the error matrix $\boldsymbol{E}$ has the same dimensions as $\boldsymbol{Z}$. There are two ways to minimise $\boldsymbol{E}$. Either one finds $\boldsymbol{y} = \boldsymbol{y}_{\min}$ that minimises $\operatorname{Tr}\left(\boldsymbol{E}^{\mathrm{T}}\boldsymbol{M}_x\boldsymbol{E}\right)$ for given $\boldsymbol{x}$ and a symmetric positive-definite $n \times n$ matrix $\boldsymbol{M}_x$, or $\boldsymbol{x} = \boldsymbol{x}_{\min}$ that does the same job for $\operatorname{Tr}\left(\boldsymbol{E}\boldsymbol{M}_y\boldsymbol{E}^{\mathrm{T}}\right)$ but this time for a fixed $\boldsymbol{y}$ and a different symmetric positive-definite $m \times m$ matrix $\boldsymbol{M}_y$. If the $\boldsymbol{M}$'s are unit matrices, the two minimised quantities are identical and in both cases we have the ordinary least squares (OLS), while for diagonal and general $\boldsymbol{M}$'s we have weighted (WLS) and generalised (GLS) least squares respectively. Starting (say) with $\boldsymbol{x}_{(\mathbf{0})}$ we may determine $\boldsymbol{y} = \boldsymbol{y}_{(\mathbf{0})}$ that minimises the square of $\boldsymbol{E} = \boldsymbol{Z} - \boldsymbol{x}_{(\mathbf{0})}\boldsymbol{y}^{\mathrm{T}}$, and then $\boldsymbol{x} = \boldsymbol{x}_{(\mathbf{1})}$ for $\boldsymbol{E} = \boldsymbol{Z} - \boldsymbol{x}\boldsymbol{y}_{(\mathbf{0})}^{\mathrm{T}}$, *etc.* Proceeding this way we will obtain the sequence

$$\boldsymbol{x}_{(\mathbf{0})} \to \boldsymbol{y}_{(\mathbf{0})} \to \boldsymbol{x}_{(\mathbf{1})} \to \boldsymbol{y}_{(\mathbf{1})} \to \boldsymbol{x}_{(2)} \to \boldsymbol{y}_{(2)} \to \cdots \tag{59}$$

with the following recursive identities:

$$\boldsymbol{y}_{(i)} = \frac{\boldsymbol{Z}^{\mathrm{T}}\boldsymbol{M}_x\boldsymbol{x}_{(i)}}{\boldsymbol{x}_{(i)}^{\mathrm{T}}\boldsymbol{M}_x\boldsymbol{x}_{(i)}} \qquad \text{and} \qquad \boldsymbol{x}_{(i+1)} = \frac{\boldsymbol{Z}\boldsymbol{M}_y\boldsymbol{y}_{(i)}}{\boldsymbol{y}_{(i)}^{\mathrm{T}}\boldsymbol{M}_y\boldsymbol{y}_{(i)}}\,. \tag{60}$$

Factor	$\lambda_\star^{(0)}$ MaxVar (1,6,9)	ρ_H	S MaxVar (1,6,9)	$S(\cdots) - S$(BM-ERW(6,9)) MaxVar (1,6,9)	MaxVar (0,6,9)	MaxVar (2,6,9)	$S(\cdots) - S$(MaxVar(1,6,9)) MaxVar (1,6,9, *Beta*)	MaxVar (1,6,9, *Book*)	MaxVar (1,6,9, *Size*)	R^2
10Y Rates	8.10	0.95	-0.20	-0.02	-0.03	-0.04	-0.02	-0.05	-0.24	0.01
Beta	7.63	0.96	0.34	-0.07	0.02	-0.02	-0.00	-0.24	-0.46	0.05
Momentum	6.10	0.97	-0.23	-0.12	-0.12	-0.08	-0.02	0.03	-0.04	0.01
Capitalisation	5.14	0.96	1.44	-0.15	-0.33	-0.23	-0.11	-0.58	0.00	0.01
Dividend	4.28	0.95	-0.22	-0.05	0.04	-0.04	-0.00	-0.00	1.01	0.00
Euro	3.85	0.95	0.40	0.05	0.12	0.08	0.00	-0.06	0.16	0.00
Liquidity	3.70	0.94	0.77	0.17	0.16	0.15	-0.14	-0.39	0.01	0.08
Book	3.53	0.93	0.51	-0.13	-0.24	-0.19	-0.05	0.00	-1.34	0.01
STR	3.44	0.96	1.89	0.07	0.13	0.10	0.01	0.05	-0.31	0.05
Sales	3.35	0.94	-0.07	0.08	0.26	0.13	0.07	0.50	0.98	0.07
Leverage	2.05	0.89	0.46	0.05	0.04	0.08	0.00	0.05	-0.73	0.03
Earning	2.04	0.91	0.52	-0.04	0.03	0.04	0.00	0.01	0.61	0.00
Cash	1.79	0.89	1.13	0.01	-0.22	0.00	0.00	-0.35	-0.17	0.01
Growth	1.24	0.91	0.27	-0.04	-0.02	0.01	0.00	0.00	0.11	0.00

Table 13: Difference in the Sharpe ratios between the Maximum-Variance and the benchmark portfolios (see Table 2) from 2000 to April 2018. ρ_H is the correlation between returns that are very high. Maximum-Variance and the 20% top-bottom are therefore highly correlated. The Sharpe ratio is statistically not significant for most factors except for *STR* (without cost), *Cash* and *Liquidity. Capitalisation*'s Sharpe ratio is overestimated as it suffers from the survival bias of our data. According to the literature (see Table 14 for a partial list of references) shows that for a much longer period (around 50 years) *Book*, *Momentum* and *Capitalisation* are the main significant risk premia even if there is no consensus on these anomalies that appear to be very sensitive to the normalisation method (see the discussion in the very end of Section 4.2). The differences between the empirical Sharpe ratios are statistically insignificant (below one sigma) as the period is too short and as risk premia are too weak. The last column presents the R^2-coefficient of the linear regression between the monthly variations of $\lambda_\star^{(0)}(\text{MaxVar}(1,6,9))$ and the monthly factor returns. The connection between these coefficients and the leverage effect for all these factors was covered in Section 4.4.2.

Eliminating $\boldsymbol{x}$ we arrive at the relation between $\boldsymbol{y}_{(i+1)}$ and $\boldsymbol{y}_{(i)}$:

$$\boldsymbol{y}_{(i+1)} = \kappa_{(i)} \cdot \boldsymbol{Z}^{\mathrm{T}} \boldsymbol{M}_x \boldsymbol{Z} \boldsymbol{M}_y \boldsymbol{y}_{(i)} \quad , \qquad \text{where} \qquad \kappa_{(i)} \equiv \frac{\boldsymbol{y}_{(i)}^{\mathrm{T}} \boldsymbol{M}_y \boldsymbol{y}_{(i)}}{\boldsymbol{y}_{(i)}^{\mathrm{T}} \boldsymbol{M}_y \boldsymbol{Z}^{\mathrm{T}} \boldsymbol{M}_x \boldsymbol{Z} \boldsymbol{M}_y \boldsymbol{y}_{(i)}} \,. \tag{61}$$

With a little algebra it can be shown that the sequence (59) converges to:

$$\boldsymbol{x}_\star = \frac{\boldsymbol{Z} \boldsymbol{M}_y^{\frac{1}{2}} \boldsymbol{v}_i}{\boldsymbol{v}_i^{\mathrm{T}} \boldsymbol{v}_i} \qquad \boldsymbol{y}_\star = \boldsymbol{M}_y^{-\frac{1}{2}} \boldsymbol{v}_i \,, \tag{62}$$

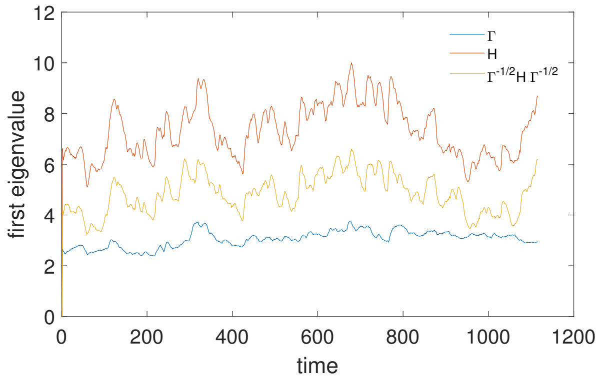

Figure 11: Measure of the first eigenvalue of the correlation matrices of the covariance matrix h, γ and $\boldsymbol{\gamma}^{-\frac{1}{2}}\boldsymbol{h}\boldsymbol{\gamma}^{-\frac{1}{2}}$. A random selection of factor with $\mu = 1.4$ generates an first eigenvalues of 7.52 for the correlation matrix of h and 3.15 for the correlation matrix of γ in agreement with the graph. The first eigenvalue of the average correlation matrix h and γ are only slightly different (6.13 and 2.6). We can guess that the dynamics of the first eigenvalue of the correlation matrix of $\boldsymbol{\gamma}^{-\frac{1}{2}}\boldsymbol{h}\boldsymbol{\gamma}^{-\frac{1}{2}}$ appears to be linked to the dynamics of the first eigenvalue of the correlation matrix of the returns of the signal stocks. When the stock market gets stress, the volatility increases, the first eigenvalue of the correlation matrix of single stock returns and the first eigenvalue of the correlation matrix of $\boldsymbol{\gamma}^{-\frac{1}{2}}\boldsymbol{h}\boldsymbol{\gamma}^{-\frac{1}{2}}$ increase meaning that the fundamentals factors tend to become more volatile and more correlated when market is stressed.

where $\boldsymbol{v}_i$ is the ith eigenvector of the matrix $\boldsymbol{M}_y^{\frac{1}{2}}\boldsymbol{Z}^{\mathrm{T}}\boldsymbol{M}_x\boldsymbol{Z}\boldsymbol{M}_y^{\frac{1}{2}}$, and its norm is fixed by the choice of $\boldsymbol{x}_{(\mathbf{0})}$.[17] It is very important to notice that (62) implies

$$\boldsymbol{E}\boldsymbol{M}_y\boldsymbol{y}_\star = 0 \qquad \text{and} \qquad \boldsymbol{x}_\star^{\mathrm{T}}\boldsymbol{M}_x\boldsymbol{E} = 0\,. \tag{63}$$

The relation to PCA becomes explicit if we replace $\boldsymbol{M}$'s by the identity matrices: $\boldsymbol{x}\boldsymbol{y}^{\mathrm{T}}$ is just the ith term of the PCA expansion of $\boldsymbol{Z}$, $\left(\lambda_i\boldsymbol{v}_i\boldsymbol{v}_i^{\mathrm{T}}\right)/\left(\boldsymbol{v}_i^{\mathrm{T}}\boldsymbol{v}_i\right)$, where λ_i is the ith largest eigenvalue of $\boldsymbol{Z}^{\mathrm{T}}\boldsymbol{Z}$ (and therefore also of $\boldsymbol{Z}\boldsymbol{Z}^{\mathrm{T}}$) and $\boldsymbol{v}_i$ is the corresponding eigenvector of $\boldsymbol{Z}^{\mathrm{T}}\boldsymbol{Z}$.

The Market Mode

To reproduce (3) and (4) one has to set $(m,n) = (T,N)$, $(\boldsymbol{Z})_{ti} = r_i(t)$ as well as $\boldsymbol{M}_x = T^{-1}\boldsymbol{I}_T$ and $\boldsymbol{M}_y = \boldsymbol{\Gamma}^{-1}$. Then plugging $\boldsymbol{x}_{(0)} = r^{\mathbf{m}}(t)$ and $i = 0$ into (60) we arrive at $\boldsymbol{y}_{(0)} = \boldsymbol{\beta}$ and then $\boldsymbol{x}_{(1)} = r_\star^{\mathbf{m}}(t)$ exactly as in the two formulae. The only differences are the conditional expectation in (3) replaced here by the regular mean, and the fact we used time-dependent betas in (4) rather then constant ones like here.

[17] Notice that the square root of $\boldsymbol{M}_y$ is well-defined since this matrix is positive-definite.

The Two-Factor Model Loadings

To find b_i in the Two-Factor model (11) using the linear regression we need the first equation of (60) for $i = 0$ with $(\boldsymbol{Z})_{ti} = r_i(t) - \beta_i r_\star^{\mathbf{m}}(t)$, $\boldsymbol{M}_x \propto \boldsymbol{I}_T$, $\boldsymbol{x_0} = r^{\mathbf{f}}(t)$ and $\boldsymbol{y_0} = \boldsymbol{b}$. This leads to

$$b_i = \frac{\sum_{t=1}^{T} \left(r_i(t) - \beta_i r_\star^{\mathbf{m}}(t)\right) r^{\mathbf{f}}(t)}{\sum_{t=1}^{T} \left(r^{\mathbf{f}}(t)\right)^2} = \left(\Sigma^{\mathbf{f}}\right)^{-2} \left\langle \left(r_i - \beta_i r_\star^{\mathbf{m}}\right) r^{\mathbf{f}} \right\rangle . \tag{64}$$

The market-neutrality $\sum_{i=1}^{N} \beta_i b_i = 0$ follows directly from (3) and (4) .

B Styles

In Table 14 we provide a brief description of all the financial styles used in the paper. The sectoral factors are presented in the next appendix.

C Maxima, minima and saddle points

Let $\boldsymbol{M}$ be an $N \times N$ positive matrix, $(\ell_i, \boldsymbol{v}_i)$ its sorted $(\ell_1 \geqslant \ell_2 \geqslant \cdots \geqslant \ell_N)$ eigenvalue/eigenvector pairs and $\boldsymbol{\mathcal{H}}_i$ be the Hessian matrix of the Lagrangian

$$\mathcal{L}(\boldsymbol{v}) = \frac{\boldsymbol{v}^{\mathrm{T}} \boldsymbol{M} \boldsymbol{v}}{\boldsymbol{v}^{\mathrm{T}} \boldsymbol{v}} \tag{65}$$

computed at a local optimal point $\boldsymbol{v} = \boldsymbol{v}_i$. We search for the signature of $\boldsymbol{\mathcal{H}}_i$. Substituting $\boldsymbol{v} + \boldsymbol{v}_i \delta \boldsymbol{v}$ into the Lagrangian one can easily see that

$$\mathcal{L}(\boldsymbol{v}) = \ell_i + \frac{1}{\boldsymbol{v}_i^{\mathrm{T}} \boldsymbol{v}_i} \cdot \delta \boldsymbol{v}^{\mathrm{T}} \left(\boldsymbol{M} - \ell_i \cdot \boldsymbol{I}_N\right) \delta \boldsymbol{v} + \mathcal{O}\left(\delta \boldsymbol{v}^3\right) , \tag{66}$$

where as expected the linear term vanishes upon $\boldsymbol{M}\boldsymbol{v}_i = \ell_i \boldsymbol{v}_i$. It immediately follows that the signature is:

$$\Bigg(\underbrace{-1, \cdots, -1}_{i-1}, 0, \underbrace{1, \cdots, 1}_{N-i}\Bigg) . \tag{67}$$

Here the flat direction corresponds to $\delta \boldsymbol{v} \propto \boldsymbol{v}_i$. We conclude that $\boldsymbol{v} = \boldsymbol{v}_i$ is a maximum (minimum) only for $i = 1$ $(i = N)$. For any other $1 < i < N$ the solution $\boldsymbol{v} = \boldsymbol{v}_i$ is a saddle point.

The derivation generalises trivially for the constrained eigensystems of Appendix D. Instead of $\delta \boldsymbol{v}$ one has to consider $\boldsymbol{P^c} \delta \boldsymbol{u}$, and replace $\boldsymbol{M}$ by $\boldsymbol{P^c M P^c}$.

D Constrained eigensystems

Let us search for an eigensystem (eigenvectors and eigenvalues pairs) of an $n \times n$ symmetric non-singular matrix $\boldsymbol{M}$, that is $\boldsymbol{v}^{\mathrm{T}} \boldsymbol{M} \boldsymbol{v}$ is optimised, under an additional constraint $\boldsymbol{v}^{\mathrm{T}} \boldsymbol{c} = 0$ for a unit vector $\boldsymbol{c}$, $\boldsymbol{c}^{\mathrm{T}} \boldsymbol{c} = 1$. It was shown in [42] that the $(n-1)$ *constrained* eigenvalues of $\boldsymbol{M}$ will coincide with the non-zero eigenvalues

Style	Definition	Long/Short high/low	Literature
Dividend Yield	Annual dividend income per share divided by the current share price	$\nearrow$	[68]
Capitalisation	Total market value of a company's shares	$\searrow$	[1, 17]
Liquidity	Volume of transaction in value divided by Capitalisation	$\searrow$	[69, 70, 71, 72, 73, 74, 75]
Short-term Reversion (STR)	Short-term reversal based on a 20 days moving average of returns	$\nearrow$	[76]
Momentum	Based on the last *12 months-1 month* moving average	$\nearrow$	[21]
Beta	Based on the 90 days regression on daily returns using the SP500	$\searrow$	[13, 77, 78, 19, 79, 80, 81, 82, 83]
Leverage	Debt to Equity	$\nearrow$	[84, 85]
Book	Book to Price ratio	$\nearrow$	[1]
Cash	Cash to Price ratio	$\nearrow$	[86, 87, 88]
Earning	Price to Earning ratio	$\nearrow$	[18, 89]
Growth	One year change of Earning divided by Equity	$\nearrow$	[90]
Euro	Price sensitivity to the weekly change in Euro/dollar based on the last 200 days	$\nearrow$	[91]
Rates	Price sensitivity to the weekly change in 10 years US Bond yield	$\searrow$	[92]
Sales	Sales to Price ratio	$\searrow$	[89, 93, 94, 95]

Table 14: summary of the basic information about the styles (non-sectoral factors) used in this paper. Apart from the definition and the relevant references we also present the long/short strategy. Our notations are as follows: $\nearrow$ stands for long high & short low and $\searrow$ for long low & short high.

of the matrix $\boldsymbol{P}^{\boldsymbol{c}}\boldsymbol{M}\boldsymbol{P}^{\boldsymbol{c}}$, where $\boldsymbol{P}^{\boldsymbol{c}} = \boldsymbol{I}_n - \boldsymbol{c}\boldsymbol{c}^{\mathrm{T}}$ is the projection matrix into the subspace of vectors orthogonal to $\boldsymbol{c}$. Moreover, if $\boldsymbol{u}$ is an eigenvector of $\boldsymbol{P}^{\boldsymbol{c}}\boldsymbol{M}\boldsymbol{P}^{\boldsymbol{c}}$ with a non-zero eigenvalue ℓ, then $\boldsymbol{P}^{\boldsymbol{c}}\boldsymbol{u}$ is necessarily an eigenvector of $\boldsymbol{P}^{\boldsymbol{c}}\boldsymbol{M}$ with exactly the same eigenvalue ℓ.

To prove this statement one starts with a Lagrangian (see (1.4) of [42]):

$$\mathcal{L}(\boldsymbol{v}, \Lambda_1, \Lambda_2) \equiv \boldsymbol{v}^{\mathrm{T}} \boldsymbol{M} \boldsymbol{v} - \Lambda_1 \left(\boldsymbol{v}^{\mathrm{T}} \boldsymbol{v} - 1\right) + 2\Lambda_2 \boldsymbol{v}^{\mathrm{T}} \boldsymbol{c}\,, \tag{68}$$

where Λ_1 and Λ_2 are the Lagrange multipliers ensuring the normalisation of $\boldsymbol{v}$ and the orthogonality of $\boldsymbol{c}$ respectively. From the equations of motion with respect to the three variables one finds then that

$$\boldsymbol{P}^{\boldsymbol{c}} \boldsymbol{M} \boldsymbol{v} = \Lambda_1 \boldsymbol{v}\,, \tag{69}$$

where $\boldsymbol{P}^{\boldsymbol{c}}$ is the projection operator satisfying $(\boldsymbol{P}^{\boldsymbol{c}})^2 = \boldsymbol{P}^{\boldsymbol{c}}$. This guarantees that the eigenvalues of $\boldsymbol{P}^{\boldsymbol{c}}\boldsymbol{M}$ coincides with those of $\boldsymbol{P}^{\boldsymbol{c}}\boldsymbol{M}\boldsymbol{P}^{\boldsymbol{c}}$ (recall that for any two square matrices $\boldsymbol{A}$ and $\boldsymbol{B}$, the eigenvalues of $\boldsymbol{AB}$ coincide with those of $\boldsymbol{BA}$.) We see, therefore, that Λ_1 is a constrained eigenvalue of $\boldsymbol{M}$. Moreover, $\boldsymbol{v}$ in the last equation might be written as $\boldsymbol{P}^{\boldsymbol{c}}\boldsymbol{u}$, where $\boldsymbol{u}$ is a standard eigenvector of $\boldsymbol{P}^{\boldsymbol{c}}\boldsymbol{M}\boldsymbol{P}^{\boldsymbol{c}}$.

In Appendix A we discussed linear regressions with the cost functions $\mathrm{Tr}\left(\boldsymbol{E}^{\mathrm{T}}\boldsymbol{M}_x\boldsymbol{E}\right)$ and $\mathrm{Tr}\left(\boldsymbol{E}\boldsymbol{M}_y\boldsymbol{E}^{\mathrm{T}}\right)$. It can be easily extended to the cost function minimisation under a given constraint, that is to say to the *constrained* WLS: the eigenvector $\boldsymbol{v}_i$ in (62) will be simply replaced by a constrained eigenvector of the matrix $\boldsymbol{M}_y^{\frac{1}{2}}\boldsymbol{Z}^{\mathrm{T}}\boldsymbol{M}_x\boldsymbol{Z}\boldsymbol{M}_y^{\frac{1}{2}}$. This is directly related to the Maximum-Variance portfolio of Section 2.5, where we had $\boldsymbol{M}_x = \mathcal{I}$, $\boldsymbol{M}_y = \boldsymbol{\Gamma}^{-1}$ and $\boldsymbol{Z}^{\mathrm{T}}\boldsymbol{Z} = \boldsymbol{H}$.

E The Maximum-Variance Market-Mode portfolio

One way to exploit the approach of Section 2.3 is to treat the market mode as a factor. To this end we may replace b_i and $r^{\mathbf{f}}$ in (11) by β_i and $r_\star^{\mathbf{m}}$ respectively and drop the $\beta_i r_\star^{\mathbf{m}}$ term on the left-hand side. This way we "discover" the Capital Asset Pricing Model (CAPM): $r_i(t) = \beta r_\star^{\mathbf{m}}(t) + \epsilon(t)_i$. As we have discussed above, it is related to the PCA analysis of the correlation matrix.

Proceeding as above we arrive at the *Maximum-Variance market-mode portfolio*:

$$\left(\omega_\star^{\mathbf{m}(0)}\right)_i \sim \Sigma_i^{-2}\beta_i\,. \tag{70}$$

As we have already mentioned below (4), the return of $\boldsymbol{\omega}_\star^{\mathbf{m}(0)}$ is equal to $r_\star^{\mathbf{m}}(t)$. The fastest way to obtain (70) is to replace b_i by β_i in (20) and to omit the irrelevant part of the market-neutrality projection.

It is worth to compare $\boldsymbol{\omega}_\star^{\mathbf{m}(0)}$ with other market portfolios proposed in the literature. In [10] the 1-factor model was used to introduce *Minimum Variance* and *Maximum Diversification* portfolios. The former is the optimal Sharpe ratio portfolio under the assumption that the expected stock returns are identical and positive. Maximum Diversification, on the other hand, is Sharpe-optimal if we assume that the expected returns are proportional to their volatilities. The Minimum Variance and Maximum Diversification portfolios may contain only long positions, although this condition is weak as all betas are close to one (see below). We will show later that the Maximum-Variance portfolio has the optimal Sharpe ratio if the expected returns are proportional to their betas, which is akin to the efficient market hypothesis, and if the residual return volatilities (Σ_i) and the stock volatilities are proportional. The latter is a meaningful and robust approximation that avoids problematic concentration due to fallacious correlation [40]. Table 15 brings together the three portfolios.

F Sectors

In Table 16 we report all *Sectors* and Industries of the Global Industry Classification Standard (GICS) used to construct the sectoral factors and the clusters as outlined in Subsection 3.4.

Maximum Diversification	Minimum Variance	Maximum-Variance
$\begin{cases} \dfrac{\Sigma_i}{\sigma_i^2}\left(1-\dfrac{\beta_i \Sigma^{\mathbf{m}}}{\Sigma_i \rho_L}\right) & \text{if} \quad \dfrac{\beta_i \Sigma^{\mathbf{m}}}{\rho_L \Sigma_i} < 1 \\ 0 & \text{otherwise} \end{cases}$	$\begin{cases} \dfrac{1}{\sigma_i^2}\left(1-\dfrac{\beta_i}{\beta_L}\right) & \text{if} \quad \dfrac{\beta_i}{\beta_L} < 1 \\ 0 & \text{otherwise} \end{cases}$	$\left(\omega_\star{}^{\mathbf{m}(0)}\right)_i = \dfrac{\beta_i}{{\Sigma_i}^2}$

Table 15: Summary of the weights of the two portfolios proposed in [10] (the first two columns) and the one defined by (70). Notice that the $\left(\omega_\star{}^{\mathbf{m}(0)}\right)_i \geqslant 0$ restriction does not affect many stocks. The parameters ρ_L and β_L are fixed from the Sharpe ratio optimisation, see [10]. The Minimum and the Maximum-Variance portfolios are invested in the low-beta and the high-beta stocks respectively. In these two portfolios the weights are inversely proportional to the square of volatilities, though for Minimum Variance these are the volatilities of the residual returns. Thus the Minimum and the Maximum-Variance portfolios can be seen as complementary.

Sector		Industry		GICS	Cluster
1	Energy	1	Energy Equipment & Services	101010	**1**
		2	Oil Gas & Consumable Fuels	101020	**1**
-	Materials	3	Chemicals	151010	**1**
			Construction Materials	151020	**1**
			Containers & Packaging	151030	**1**
			Paper & Forest Products	151050	**1**
		4	Metals & Mining	151040	**1**
2	Industrials	5	Aerospace & Defence	201010	**2**
			Building Products	201020	**2**
			Electrical Equipment	201040	**2**
			Trading Companies & Distributors	201070	**2**
		6	Machinery	201060	**2**
		7	Commercial Services & Supplies	202010	**3**
			Professional Services	202020	**3**
		8	Air Freight & Logistics	203010	**3**
			Airlines	203020	**3**
			Marine	203030	**3**
			Road & Rail	203040	**3**
		9	Transportation Infrastructure	203050	**3**

To be continued

Sector		Industry		GICS	Cluster
3	Consumer Discretionary	10	Auto Components	251010	**3**
			Automobiles	251020	**3**
		11	Household Durables	252010	**3**
			Leisure Products	252020	**3**
			Textiles, Apparel & Luxury Goods	252030	**3**
		12	Hotels, Restaurants & Leisure	253010	**3**
			Diversified Consumer Services	253020	**3**
		13	Media	254010	**3**
		14	Distributors	255010	**3**
			Internet & Direct Marketing Retail	255020	**3**
			Multiline Retail	255030	**3**
			Specialty Retail	255040	**3**
4	Consumer Staples	15	Food & Staples Retailing	301010	**4**
		16	Beverages	302010	**4**
			Food Products	302020	**4**
			Tobacco	302030	**4**
		17	Household Products	303010	**4**
			Personal Products	303020	**4**
5	Health Care	18	Health Care Equipment & Supplies	351010	**4**
			Health Care Providers & Services	351020	**4**
			Health Care Technology	351030	**4**
		19	Biotechnology	352010	**4**
			Pharmaceuticals	352020	**4**
			Life Sciences Tools & Service	352030	**4**
6	Financials	20	Banks	401010	**5**
			Thrifts & Mortgage Finance	401020	**5**
		21	Diversified Financial Services	402010	**5**
			Consumer Finance	402020	**5**
			Capital Markets	402030	**5**
		22	Mortgage Real Estate Investment Trusts (REITs)	402040	**5**
		23	Insurance	403010	**5**
7	Information Technology	24	Internet Software & Services	451010	**6**
			IT Service	451020	**6**
		25	Software	451030	**6**
		26	Communications Equipment	452010	**6**
			Technology Hardware Storage & Peripherals	452020	**6**
			Electronic Equipment Instruments & Components	452030	**6**
		27	Semiconductors & Semiconductor Equipment	453010	**6**

To be continued

Sector		Industry		GICS	Cluster
-	Telecommunication Services	28	Diversified Telecommunication Services	501010	**4**
			Wireless Telecommunication Services	501020	**4**
8	Utilities	29	Electric Utilities	551010	**4**
			Multi-Utilities	551030	**4**
			Water Utilities	551040	**4**
			Independent Power and Renewable Electricity Producers	551050	**4**
9	Real Estate	30	Equity Real Estate Investment Trusts (REITs)	601010	**5**
			Real Estate Management & Development	601020	**5**

Table 16: The first two columns on the left describe the nine GICS sectors used for all but the MaxVar(1,6,30) method, see Table 2. The next two columns contain the 30 industries used to build the sectoral factors of MaxVar(1,6,30). The last column marks the six clusters employed in all but the MaxVar(1,1,9) method, which has no clustering.

G Proof of optimal capture of eigenvalues

Let us denote by $\lambda_i(\boldsymbol{M})$ the i-th largest eigenvalue of a $n \times n$ Hermitian matrix $\boldsymbol{M}$. In other words,

$$\lambda_1(\boldsymbol{M}) \geqslant \lambda_2(\boldsymbol{M}) \geqslant \cdots \geqslant \lambda_n(\boldsymbol{M}) .$$

In this appendix we demonstrate that:

$$\lambda_k\left(\boldsymbol{\Gamma}^{-\frac{1}{2}}\boldsymbol{D}\boldsymbol{h}\boldsymbol{D}\boldsymbol{\Gamma}^{-\frac{1}{2}}\right) \leqslant \lambda_k\left(\boldsymbol{\Gamma}^{-\frac{1}{2}}\boldsymbol{h}\boldsymbol{\Gamma}^{-\frac{1}{2}}\right) \quad \text{for any} \quad k = 1, \ldots, K . \tag{71}$$

To remind the reader, $\boldsymbol{h}$ is a covariance matrix (and so by definition positive definite), $\boldsymbol{\Gamma}$ is a correlation matrix, and finally $\boldsymbol{D}$ is a diagonal matrix whose (real) entries belong to the range $(0, 1]$, or $0 < \lambda_k(\boldsymbol{D}) \leqslant 1$, but as we will see below this might be replaced by a weaker condition $0 < |\lambda_k(\boldsymbol{D})| \leqslant 1$, while D does not have to be diagonal, but rather only symmetric.

According to Lidskii [96] the following holds for two arbitrary $n \times n$ positive semidefinite Hermitian matrices $\boldsymbol{U}$ and $\boldsymbol{V}$:

$$\prod_{s=1}^{l} \lambda_{i_s}(\boldsymbol{U}\boldsymbol{V}) \leqslant \prod_{s=1}^{l} \lambda_{i_s}(\boldsymbol{U})\,\lambda_s(\boldsymbol{V}) , \tag{72}$$

with any

$$1 \leqslant i_1 \leqslant \cdots \leqslant i_l \leqslant K \quad \text{and} \quad l = 1, \ldots, K . \tag{73}$$

In particular, for $l = 1$ this theorem implies that

$$\lambda_i(\boldsymbol{U}\boldsymbol{V}) \leqslant \lambda_i(\boldsymbol{U})\,\lambda_1(\boldsymbol{V}) \quad \text{for any} \quad i = 1, \ldots, n . \tag{74}$$

We now apply this inequality for $\boldsymbol{U} = \boldsymbol{\Gamma}^{-\frac{1}{2}}\boldsymbol{h}\boldsymbol{\Gamma}^{-\frac{1}{2}}$ and $\boldsymbol{V} = \boldsymbol{\Gamma}^{\frac{1}{2}}\boldsymbol{D}\boldsymbol{\Gamma}^{-1}\boldsymbol{D}\boldsymbol{\Gamma}^{\frac{1}{2}}$:

$$\begin{aligned}\lambda_i\left(\boldsymbol{\Gamma}^{-\frac{1}{2}}\boldsymbol{D}\boldsymbol{h}\boldsymbol{D}\boldsymbol{\Gamma}^{-\frac{1}{2}}\right) =& \lambda_i\left(\boldsymbol{h}\boldsymbol{D}\boldsymbol{\Gamma}^{-1}\boldsymbol{D}\right) = \\ =& \lambda_i\left(\boldsymbol{\Gamma}^{-\frac{1}{2}}\boldsymbol{h}\boldsymbol{\Gamma}^{-\frac{1}{2}} \cdot \boldsymbol{\Gamma}^{\frac{1}{2}}\boldsymbol{D}\boldsymbol{\Gamma}^{-1}\boldsymbol{D}\boldsymbol{\Gamma}^{\frac{1}{2}}\right) \leqslant \lambda_i\left(\boldsymbol{\Gamma}^{-\frac{1}{2}}\boldsymbol{h}\boldsymbol{\Gamma}^{-\frac{1}{2}}\right)\lambda_1\left(\boldsymbol{\Gamma}^{\frac{1}{2}}\boldsymbol{D}\boldsymbol{\Gamma}^{-1}\boldsymbol{D}\boldsymbol{\Gamma}^{\frac{1}{2}}\right) .\end{aligned} \tag{75}$$

	$K=2$ ITPCSR	2-Step $K=1$ TPCSR (OLS)	2-Step $K=1$ TPCSR (WLS)
$\omega^{\mathbf{f}(0)}_{\star}$	$b_i - \frac{\sum_{k=1}^{N} b_k\beta_k}{\sum_{k=1}^{N}\beta_k^2}\beta_i$	$b_i - \frac{\sum_{k=1}^{N} b_k\beta_k}{\sum_{k=1}^{N}\beta_k^2}\beta_i$	$\Sigma_i^{-2}\left(b_i - \frac{\sum_{j=1}^{N} b_j\beta_j\Sigma_j^{-2}}{\sum_{k=1}^{N}\beta_k^2\Sigma_k^{-2}}\beta_i\right)$
$\omega^{\mathbf{m}(0)}_{\star}$	$\beta_i - \frac{\sum_{k=1}^{N}\beta_k b_k}{\sum_{k=1}^{N} b_k^2} b_i$	β_i	$\Sigma_i^{-2}\beta_i$

Table 17: The weights of the market (second line) and all other styles/factors (first line) in different approaches.

Here we used the fact that for two square non-singular matrices $\boldsymbol{A}$ and $\boldsymbol{B}$, the eigenvalue spectra of $\boldsymbol{AB}$ and $\boldsymbol{BA}$ are always identical. To complete the proof we notice that:

$$\lambda_1\left(\boldsymbol{\Gamma}^{\frac{1}{2}}\boldsymbol{D}\boldsymbol{\Gamma}^{-1}\boldsymbol{D}\boldsymbol{\Gamma}^{\frac{1}{2}}\right) \leqslant \left(\lambda_1\left(\boldsymbol{\Gamma}^{\frac{1}{2}}\boldsymbol{D}\boldsymbol{\Gamma}^{-\frac{1}{2}}\right)\right)^2 = \left(\lambda_1\left(\boldsymbol{D}\right)\right)^2 \leqslant 1\,, \tag{76}$$

where in the first inequality we used (74) once more.

H Comparison with Other Cross-Sectional Regressions

Two-pass cross-sectional regression (TPCSR) was frequently used to estimate the optimal factor loadings and weights from a general K-factor model. Recall that in our treatment we have K different two-factor models (11). Here we would like to outline the similarities and the differences between TPCSR and our model.

A good starting point is to notice that the model (11) might be re-written as $\boldsymbol{r} = \boldsymbol{B}^{\mathrm{T}}\boldsymbol{R}^{\mathbf{f}} + \boldsymbol{\epsilon}$, where $\boldsymbol{B}^{\mathrm{T}} = (\boldsymbol{\beta}, \boldsymbol{b})$ is the $(N,2)$-matrix of style loadings (market and a single style from our set of styles), $\boldsymbol{R}^{\mathbf{f}} = \left(r_i^{\mathbf{m}}, r_i^{\mathbf{f}}\right)$ is the $(2,T)$-matrix of factor returns, and $\boldsymbol{\epsilon}$ are the idiosyncratic returns. These matrices can be trivially generalised to capture $K > 2$ different factor/styles. The matrices $\boldsymbol{B}$ and $\boldsymbol{R}^{\mathbf{f}}$ will then be of size (K,T) and (K,N) respectively. With these conventions, TPCSR consists of recurrent time- and cross-sectional regressions. Starting from a given set of factor returns (for example, with $r_i^{\mathbf{b.m.}}(t)$) one fixes the factor loadings with the OLS linear regression of the time series, $r_i(t)$ against $\boldsymbol{R}^{\mathbf{f}}$. At the second step the $\boldsymbol{B}$-matrix is used to determine a new matrix of factor returns, $\boldsymbol{R}^{\mathbf{f}'}$, by means of the cross-sectional (cross-factorial to be more precise) OLS linear regression. One can repeat these two-step procedure indefinitely, in which case the model goes under the name ITPCSR. As was pointed out in [97], the iteration has a fixed point, where the factor returns converge to the K first eigenvectors (the eigenvectors corresponding to the K largest eigenvalues) of the $N \times N$ covariance matrix. This is just a $K > 1$ generalisation of the PCA decomposition discussed in Appendix A.

The loadings $\mathfrak{B}_{\star}^{(0)}$ of our model defined in Section 2.5 might be seen as a two-stage implementation of the $K = 1$ version of ITPCSR construction with the market-neutrality constraint.

- For a short time scale of 90 days, we use an exponential moving average of the $K = 1$ version of ITPCSR to determine the betas. In practise we do not repeat the iteration in order to get the beta with respect to the stock index return, $r^{\mathbf{m}}(t)$, rather than with respect to the Maximum-Variance market mode portfolio return, $r^{\mathbf{m}}_{\star}(t)$. It enables to take into account the variation of $\beta(t)$, which is an important feature of financial markets [98]. It was shown in [87] that time-varying beta makes the low volatility anomaly disappear, thus improving the empirical validation of the CAPM. Our methodology supports a time-varying beta which is not the case of multi-factorial ITPCSR that needs more data than 90 days in order to estimate as precisely as possible the eigenvectors.

- At the second stage, for a long time scale of several years, we applied $K = 1$ version of ITPCSR with market-neutrality constraint. This second ITPCSR when using a WLS instead of OLS and aggregation of similar stocks into $Q = 10$ different portfolios lead to the first constrained eigenvector of matrix $\widetilde{\boldsymbol{\gamma}}^{-\frac{1}{2}}\widetilde{\boldsymbol{h}}\widetilde{\boldsymbol{\gamma}}^{-\frac{1}{2}}$ (demonstrated in Appendix D). We will thus refer to our approach as the two-stage $K = 1$ version of ITPCSR.

Let us address the following crucial points:

- We used the WLS instead OLS, since it fits better with the heterogeneity of the stock volatilities and correct the effect of the heteroscedasticity. This can be seen in the Σ^2 factors in (20).
- The fixed point portfolio corresponds to the Maximum-Variance portfolio if the signal is strong enough ($\lambda^{(0)}\left(\boldsymbol{\omega}^{\mathbf{b.m.}}\right)$ should be high enough).
- In Table 17 we summarize the weights obtained using the standard $K = 2$ ITPCSR applied to the market and an additional factor (first column), and the weights derived using our method with the OLS linear regression (second columns) and the WLS (last column). Different versions give similar weights except that WLS is inversely proportional the variance and that in the 2 step $K = 1$ case the weights for the market mode is not interfering with the signal. By keeping $K = 1$ at every step we avoided mixing between the market return and the given style return. The $K = 2$ TPCSR with the same factor and the market mode will loose all the information about the original signal of this style after the first few iterations. In our approach, however, the factor return is kept mark-neutral at every iteration with equivalent weights to every stock in the same quantile, and thus the fixed point return and the sensitivity will preserve connection to the initial signal input. Even more importantly, the market weights, the betas, will not be affected at all at the second step.
- We obtain factors which are market-neutral with respect to the value-weighted stock index, while the factors of [97] for $K = 2$ (the market and an additional factor) become, after sufficiently many interactions, neutral with the respect to the first component of the covariance matrix.
- For $K \gg 1$ (including $K = 24$) the iterative procedure of [97] fails to incorporate the grouping of stocks into quantiles. For instance, $Q = 10$ necessitates 10^K different portfolios. This is in contrast with our approach, since we group stocks separately for each factor and so there are overall $K \times Q$ portfolios. As a result, the iteration can converge only to the noisy eigenvalues of the $N \times N$ correlation matrix, while quickly losing connection to the initial financial information.

To conclude, the input of the initial benchmark portfolio is washed away in the iterative process of linear regressions that yield to a solution that is finally not optimal. As a consequence the only available solution is the benchmark portfolio. Even if it is not optimized, $\omega^{\mathbf{b.m.}}$ remains the reference method to capture financial signals. As an example, in the mainstream Fama-French framework $\omega^{\mathbf{b.m.}}$ is implemented in its equal-weighted and neutral in nominal: $\sum_i \omega_i = 0$ instead of $\sum_i \beta_i \omega_i = 0$.

I Conditional and unconditional estimates based on the EMA

Estimating conditional volatilities and beta based on daily returns

To estimate the matrices $\boldsymbol{\gamma}(t)$, $\boldsymbol{h}(t)$ and all other matrices we have first to find the volatilities $\Sigma_i(t)$ and the betas $\beta_i(t)$. As we discussed in the main text, we aim to estimate the conditional values referred to in Section 2.1 as $\mathbb{E}_{t-1}(\cdots)$. The best way to do this is to average the time-dependent variables over the period $[t, t-\Delta t]$. The better practice is, however, to use the exponential moving average with $\alpha = \Delta^{-1}$ instead of the moving-window. Although it cannot account for all the delays, this approach still provides a good proxy for the conditional values.

We determine the variances with $\alpha_{\Sigma^2}^{-1} = 40$ days, and then use the conditional volatility of the stock index to estimate the betas with $\alpha_\beta^{-1} = 90$ days:

$$\begin{aligned} \Sigma_i^2(t+\delta t) &= (1-\alpha_{\Sigma^2}) \cdot \Sigma_i^2(t) + \alpha_{\Sigma^2} \cdot r_i^2(t) \\ \beta_i(t+\delta t) &= (1-\alpha_\beta) \cdot \beta_i(t) + \alpha_\beta \cdot \frac{r_i(t)\, r^{\mathbf{m}}(t)}{\Sigma^{\mathbf{m}2}(t)} \,. \end{aligned} \tag{77}$$

Here $\delta t = 1$ day and respectively all the returns are daily. We neglected the daily mean of the returns and used the same moving average to find the volatility index as for the single stocks:

$$(\Sigma^{\mathbf{m}})_i^2(t+\delta t) = (1-\alpha_{\Sigma^2}) \cdot (\Sigma^{\mathbf{m}})_i^2(t) + \alpha_{\Sigma^2} \cdot (r^{\mathbf{m}})_i^2(t) \,. \tag{78}$$

Estimating $\boldsymbol{h}$ and $\boldsymbol{\gamma}$ for different time scales based on daily returns

In this appendix we outline the algorithm to approximate the unconditional values of the matrices $\boldsymbol{h}(\tau)$ and $\boldsymbol{\gamma}(\tau)$ at different time scales τ from 1 day to 100 days. We will write down a single generalised formula that covers all the instances of these matrices discussed in this paper. The main differences between the instances are the portfolio weights, for example $v_i^{\mathfrak{p}}$ of Section 2.5 or $\boldsymbol{\omega}_\star^{(1)}$ of Section 2.4. Below we denote the portfolio weights simply by $\omega_{Ai}(t)$. As before the i index stands for the single stocks, whereas the capital letters indices should be replaced depending on the case at hand. For example, for the calculation of Section 2.4 one needs the factor indices, meaning $A, B, \cdots \to a, b, \cdots$, and the matrices $\boldsymbol{h}(\tau)$ and $\boldsymbol{\gamma}(\tau)$ are accordingly $K \times K$.

First, recall that the Maximum-Variance (20) as well as the benchmark weights (9) are well-defined only up to an overall time-dependent normalisation. This issue was already mentioned at the end of Section 4.2. For the purposes of this section we keep constant the following quantity:

$$\sum_{i=1}^{N} \omega_{Ai}^2(t) \Sigma_i^2(t) \,. \tag{79}$$

Notice that the normalisation constant is the same for all t and A.

Second, we introduce two sample means, the mean return of portfolio A and the mean (market-neutral) return of *all* single stocks:

$$\hat{r}_i(t) = r_i(t) - \eta \cdot \beta_i(t) r^{\mathbf{m}}(t) \qquad \mu_A \equiv \frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{N} \omega_{Ai}(t)\, \hat{r}_i(t) \qquad \bar{r} = \frac{1}{TN}\sum_{t=1}^{T}\sum_{i=1}^{N} r_i(t) \,. \tag{80}$$

In all cases but one we set $\eta = 0$. With these formulae we define the following two auxiliary functions:

$$\begin{aligned} \mathcal{E}^{\boldsymbol{h}}_{AB}(t,\tau) &\equiv \sum_{t'=t-\tau}^{t} \left(\sum_{i=1}^{N} \omega_{Ai}(t')\,\hat{r}_i(t') - \mu_A \right) \sum_{t'=t-\tau}^{t} \left(\sum_{i=1}^{N} \omega_{Bi}(t')\,\hat{r}_i(t') - \mu_B \right) \\ \mathcal{E}^{\boldsymbol{\gamma}}_{AB}(t,\tau) &\equiv \sum_{i=1}^{N} \left(\sum_{t'=t-\tau}^{t} \omega_{Ai}(t')\left(r_i(t') - \overline{r}\right) \right) \left(\sum_{t'=t-\tau}^{t} \omega_{Bi}(t')\left(r_i(t') - \overline{r}\right) \right) . \end{aligned} \tag{81}$$

Here τ is the *time-scale* used to estimate the matrices, see, for example, Table 10. In other words, the summation $\sum_{t'=t-\tau}^{t}$ gives the *accumulated* portfolio return for the period $[t-\tau, t]$. Notice also that all the parentheses in (81) contain only zero-mean quantities.

Next, we use the exponential moving average with $\alpha_{h,\gamma}^{-1} = 3\tau^{-1}$ days to "smooth" these matrices:

$$\begin{aligned} \widetilde{\mathcal{E}}^{\boldsymbol{h}}_{AB}(t+\delta t,\tau) &\equiv \left(1 - \frac{\alpha_{h,\gamma}}{\tau}\right) \cdot \widetilde{\mathcal{E}}^{\boldsymbol{h}}_{AB}(t,\tau) + \frac{\alpha_{h,\gamma}}{\tau} \cdot \mathcal{E}^{\boldsymbol{h}}_{AB}(t,\tau) \\ \widetilde{\mathcal{E}}^{\boldsymbol{\gamma}}_{AB}(t+\delta t,\tau) &\equiv \left(1 - \frac{\alpha_{h,\gamma}}{\tau}\right) \cdot \widetilde{\mathcal{E}}^{\boldsymbol{\gamma}}_{AB}(t,\tau) + \frac{\alpha_{h,\gamma}}{\tau} \cdot \mathcal{E}^{\boldsymbol{\gamma}}_{AB}(t,\tau) , \end{aligned} \tag{82}$$

where, again, $\delta t = 1$ day. We are finally in a position to write down the expressions for $\boldsymbol{h}$ and $\boldsymbol{\gamma}$:

$$\begin{aligned} h_{AB}(\tau) &= \frac{1}{T} \sum_{t=1}^{T} \frac{\widetilde{\mathcal{E}}^{\boldsymbol{h}}_{AB}(t,\tau)}{\sqrt{\widetilde{\mathcal{E}}^{\boldsymbol{\gamma}}_{AA}(t,\tau)\,\widetilde{\mathcal{E}}^{\boldsymbol{\gamma}}_{BB}(t,\tau)}} \\ \gamma_{AB}(\tau) &= \frac{1}{T} \sum_{t=1}^{T} \frac{\widetilde{\mathcal{E}}^{\boldsymbol{\gamma}}_{AB}(t,\tau)}{\sqrt{\widetilde{\mathcal{E}}^{\boldsymbol{\gamma}}_{AA}(t,\tau)\,\widetilde{\mathcal{E}}^{\boldsymbol{\gamma}}_{BB}(t,\tau)}} \end{aligned} \tag{83}$$

Let us explain these formulae. The time-dependent matrices $\mathcal{E}^{\boldsymbol{h}}$ and $\mathcal{E}^{\boldsymbol{\gamma}}$ in (81) follow directly from the definition of $\boldsymbol{h}$ and $\boldsymbol{\gamma}$, and the EMA is the standard procedure to reduce the impact of extremely volatile one-day returns. For $\mathcal{E}^{\boldsymbol{\gamma}}$ the measurement noise is further weakened for large N, because, in contrast to $\mathcal{E}^{\boldsymbol{h}}$, it has a single summation of the stock index. Regardless of the noise, the volatility is a rough stochastic process close to the fractional Brownian motion [99]. This is precisely where the normalisation ((79))becomes important. For larger $\Sigma_i(t)$'s, it keeps the portfolio weights lower, reducing the crisis impact on $\mathcal{E}^{\boldsymbol{\gamma}}$. This way the final results for $\boldsymbol{h}$ and $\boldsymbol{\gamma}$ are not overweighted by the contribution of large volatility periods. The normalisation comes with a cost though. The conditional volatilities defined in (77) have a $\alpha_{\Sigma^2}^{-1} = 40$ days delay that brings in an "artificial" volatility in both $\mathcal{E}^{\boldsymbol{h}}$ and $\mathcal{E}^{\boldsymbol{\gamma}}$. To treat this problem, we can adjust the normalisation. In (79) the weights were normalised by the diagonal matrix $\Gamma_{ij} = \delta_{ij}\Sigma_i^2$. It therefore makes sense to use $\widetilde{\mathcal{E}}^{\boldsymbol{\gamma}}$ for the last step normalisation, and this is precisely what the square roots in (83) were introduced for. The procedure is similar to what is usually done to reduce the heteroscedasticity of a stochastic volatility process: one divides the stochastic function by its standard deviation. The formulae (83) can also be seen as weighted means over the entire time period. Since the product $\boldsymbol{\gamma}^{-1/2}\boldsymbol{h}\boldsymbol{\gamma}^{-1/2}$ is invariant under a simultaneous rescaling of the matrices $\boldsymbol{h}$ and $\boldsymbol{\gamma}$, we do not have to divide by the sum of the weights, $\sum_t \left(\widetilde{\mathcal{E}}^{\boldsymbol{\gamma}}_{AA}\widetilde{\mathcal{E}}^{\boldsymbol{\gamma}}_{BB}\right)^{-1/2}$. Finally, let us add a comment on $\alpha_{h,\gamma}$. The EMA of $\mathcal{E}^{\boldsymbol{h}}$ is strictly speaking unnecessary, as its impact becomes irrelevant after the t-averaging in (83). The EMA of $\mathcal{E}^{\boldsymbol{\gamma}}$, on the other hand, is important as it changes the normalisation. As we wrote above, for large N the function $\mathcal{E}^{\boldsymbol{\gamma}}(t,\tau)$ is not too noisy and so we keep $\alpha_{h,\gamma}$ small compared to other EMA parameters, $\alpha_{\Sigma^2}^{-1}$ and α_{β}^{-1}.

	$A, B, \cdots$	$\omega_{Ai}(t)$	τ	η
The first step of the $\mathfrak{B}_\star(\mathfrak{q})$ evaluation in Section 2.5	$\mathfrak{q}, \mathfrak{q}', \cdots$	(31)	one day	1
The second step of the $\mathfrak{B}_\star(\mathfrak{q})$ evaluation in Section 2.5	$\mathfrak{p}, \mathfrak{p}', \cdots$	Maximum-Variance portfolio $v^{(\mathfrak{p})}$ derived from (20) and (35) with $\mathfrak{B}(\mathfrak{q})$ found at the previous step	one day	0
Section 4	$a, b, \cdots$	Maximum-Variance portfolios (20) of Table 2 with the universal law (43)	$1, \cdots, 100$ days (as in Tables day 8, 9, 12 and 10)	0
	$a, b, \cdots$	Benchmark portfolios (9) of Table 2	$1, \cdots, 100$ days (as in Tables 8, 9, 12 and 10)	0

Table 18: Different instances of applications of (83) in the paper.

Estimating $\boldsymbol{h}(t)$ and $\boldsymbol{\gamma}(t)$ based on 5-minutes returns

In this appendix we describe the computation of $\boldsymbol{h}(t)$, $\boldsymbol{\gamma}(t)$ and the correlation matrix both *conditional* (time-dependent) and *unconditional* (time-independent).

We start with two successive EMA filters:

$$\begin{aligned} \mathcal{E}^{\boldsymbol{C}}_{ij}(t + 5\text{ min.}) &= \left(1 - \frac{\alpha_C}{72}\right) \cdot \mathcal{E}^{\boldsymbol{C}}_{ij}(t) + \alpha_C \cdot r_i(t) r_j(t) \\ \widetilde{\mathcal{E}}^{\boldsymbol{C}}_{ij}(t + 1\text{ day}) &= (1 - \alpha_C) \cdot \widetilde{\mathcal{E}}^{\boldsymbol{C}}_{ij}(t) + \alpha_C \mathcal{E}^{\boldsymbol{C}}_{ij}(t)\,, \end{aligned} \tag{84}$$

where $\alpha_C^{-1} = 5\,$days and we set the 72 factor because this is the average number of available 5-minutes returns for one day. The five days averaging period is important because of the daily [100] and the weekly [101] U-patterns of the return variation.

With $\mathcal{E}^{\boldsymbol{C}}_{ij}(t)$ at hand we calculate the conditional empiric correlation matrix as:

$$C^{\text{Emp}}_{ij}(t) = \frac{\widetilde{\mathcal{E}}^{\boldsymbol{C}}_{ij}(t)}{\sqrt{\widetilde{\mathcal{E}}^{\boldsymbol{C}}_{ii}(t)\widetilde{\mathcal{E}}^{\boldsymbol{C}}_{jj}(t)}}\,. \tag{85}$$

The time-dependent eigenvalues of this matrix, $\lambda^{\text{Emp}}_i(t)$, appear in Figure 8, and the time-independent eigenvalues, λ^{Emp}_i, of the average of $C^{\text{Emp}}_{ij}(t)$, that is of

$$\overline{C}^{\text{Emp}}_{ij} = \frac{1}{T}\sum_t C^{\text{Emp}}_{ij}(t) \tag{86}$$

are given in Table 5.

To arrive at the conditional constrained eigenvalues $\lambda^{(1)}(t)$ we first define

$$\begin{aligned}
\mathcal{E}^{\boldsymbol{h}}_{AB}(t) &\equiv \sum_{i,j=1}^{N} \omega_{Ai}(t)\widetilde{\mathcal{E}}^{\boldsymbol{C}}_{ij}(t)\omega_{Bj}(t) \\
\mathcal{E}^{\boldsymbol{\gamma}}_{AB}(t) &\equiv \sum_{i=1}^{N} \omega_{Ai}(t)\widetilde{\mathcal{E}}^{\boldsymbol{C}}_{ii}(t)\omega_{Bi}(t)\,.
\end{aligned} \tag{87}$$

Here $\omega_{Ai}(t)$ are the weights from the last two lines of Table 18. We assume that they evolve weakly compared to the returns. This simplifies the formulae greatly. We then proceed to the third and final EMA:

$$\begin{aligned}
\widetilde{\mathcal{E}}^{\boldsymbol{h}}_{AB}(t+1\,\text{day}) &= (1-\alpha_C)\cdot\mathcal{E}^{\boldsymbol{h}}_{AB}(t) + \alpha_C\cdot\widetilde{\mathcal{E}}^{\boldsymbol{h}}_{AB}(t) \\
\widetilde{\mathcal{E}}^{\boldsymbol{\gamma}}_{AB}(t+1\,\text{day}) &= (1-\alpha_C)\cdot\mathcal{E}^{\boldsymbol{\gamma}}_{AB}(t) + \alpha_C\cdot\widetilde{\mathcal{E}}^{\boldsymbol{\gamma}}_{AB}(t)\ ,
\end{aligned} \tag{88}$$

Finally, the matrices we are interested in, follow from $\widetilde{\mathcal{E}}^{\boldsymbol{h}}_{AB}$ and $\widetilde{\mathcal{E}}^{\boldsymbol{\gamma}}_{AB}$ similarly to (83):

$$\begin{aligned}
h_{AB}(t) &= \frac{\widetilde{\mathcal{E}}^{\boldsymbol{h}}_{AB}(t)}{\sqrt{\widetilde{\mathcal{E}}^{\boldsymbol{\gamma}}_{AA}(t)\widetilde{\mathcal{E}}^{\boldsymbol{\gamma}}_{BB}(t)}} \\
\gamma_{AB}(t) &= \frac{\widetilde{\mathcal{E}}^{\boldsymbol{\gamma}}_{AB}(t)}{\sqrt{\widetilde{\mathcal{E}}^{\boldsymbol{\gamma}}_{AA}(t)\widetilde{\mathcal{E}}^{\boldsymbol{\gamma}}_{BB}(t)}}
\end{aligned} \tag{89}$$

The dynamics of the eigenvalues of $\boldsymbol{\gamma}^{-\frac{1}{2}}(t)\boldsymbol{h}(t)\boldsymbol{\gamma}^{-\frac{1}{2}}(t)$ is shown on Figure 8 and the time-independent eigenvalues of

$$\frac{1}{T}\sum_{t}\gamma_{ik}^{-\frac{1}{2}}(t)h_{kl}(t)\gamma_{lj}^{-\frac{1}{2}}(t) \tag{90}$$

are listed in Table 5.

References

[1] E. Fama and K. French, "The cross-section of expected stock returns," *Journal of Finance* **47** no. 2, (1992) 427–65.

[2] E. Fama and K. French, "Common risk factors in the returns on stocks and bonds," *Journal of Financial Economics* **33** no. 1, (1993) 3–56.

[3] E. F. Fama and K. R. French, "A five-factor asset pricing model," *Journal of Financial Economics* **116** no. 1, (2015) 1–22.

[4] R. Thaler and S. Benartzi, "Naive diversification strategies in defined contribution saving plans," *American Economic Review* **91** no. 1, (2001) 79–98.

[5] V. DeMiguel, L. Garlappi, F. J. Nogales, and R. Uppal, "A generalized approach to portfolio optimization: Improving performance by constraining portfolio norms," *Management Science* **55** no. 5, (2009) 798–812.

[6] D. B. Chaves, J. C. Hsu, F. Li, and O. Shakernia, "Risk parity portfolio vs. other asset allocation heuristic portfolios," *Journal of Investing* **20** no. 1, (2011) 108–118.

[7] J. Treynor, "Why market-valuation-indifferent indexing works," *Financial Analysts Journal* **61** no. 5, (2005) 65–69.

[8] H. Markowitz, "Portfolio selection," *Journal of Finance* **7** no. 1, (1952) 77–91.

[9] R. G. Clarke, H. de Silva, and S. Thorley, "Minimum-variance portfolio in the us equity market," *Journal of Portfolio Management* **33** no. 10, (2006) 33–10.

[10] R. G. Clarke, H. de Silva, and S. Thorley, "Risk parity, maximum diversification, and minimum variance: An analytic perspective," *Journal of Portfolio Management* **39** no. 3, (2012) 39–53.

[11] Y. Choueifaty and Y. Coignard, "Toward maximum diversification," *The Journal of Portfolio Management* **35** no. 1, (2008) .

[12] R. G. Clarke, H. de Silva, and S. Thorley, "Minimum-variance portfolio composition," *The Journal of Portfolio Management* **37** no. 2, (2012) 31–45.

[13] A. Ang, R. Hodrick, Y. Xing, and X. Zhang, "The cross-section of volatility and expected returns,".

[14] C. R. Harvey and Y. Liu, "Lucky factors," *Working paper* (2014) .

[15] D. A. Conway and M. R. Reinganum, "Stable factors in security returns: Identification using cross-validation," *Journal of Business & Economic Statistics* **6** no. 1, (1988) 1–15.

[16] C. S. Asness, T. J. Moskowitz, and L. Pedersen, "Value and momentum everywhere," *Journal of Finance* **68** no. 3, (2013) 929–985.

[17] R. W. Banz, "The relationship between return and market value of common stocks," *Journal of Financial Economics* **9** no. 1, (1981) 3–18.

[18] S. Basu, "Investment performance of common stocks in relation to their price-earnings ratios: A test of the efficient market hypothesis," *Journal of Finance* **32** no. 3, (1977) 663–82.

[19] F. Black, M. C. Jensen, and M. Scholes, "The capital asset pricing model: Some empirical tests," *Studies in the Theory of Capital Markets* (1972) 79–121.

[20] M. Carhart, "On persistence in mutual fund performance," *Journal of Finance* **52** no. 1, (1997) 57–82.

[21] N. Jegadeesh and S. Titman, "Returns to buying winners and selling losers: Implications for stock market efficiency," *The Journal of Finance* **48** no. 1, (1993) 65–91.

[22] J. Green, J. R. M. Hand, and X. F. Zhang, "The supraview of return predictive signals," *Review of Accounting Studies* **18** no. 3, (2013) 692–730.

[23] R. D. McLean and J. Pontiff, "Does academic research destroy stock return predictability?," *Journal of Finance* **71** no. 1, (2016) 5–32.

[24] C. Harvey, Y. Liu, and H. Zhu, "... and the cross-section of expected returns," *Review of Financial Studies* **29** no. 1, (2016) 5–68.

[25] G. Connor, "The three types of factor models: A comparison of their explanatory power," *Financial Analysts Journal* **51** no. 3, (1995) 42–46.

[26] S. Ross, "The arbitrage theory of capital asset pricing," *Journal of Economic Theory* **13** no. 3, (1964) 341–360.

[27] W. Sharpe, "Capital asset prices: A theory of market equilibrium under conditions of risk," *Journal of Finance* **19** no. 3, (1964) 425–442.

[28] R. Roll and R. Stephen, "An emprirical investigation of the arbitrage pricing theory," *Journal of Finance* **35** no. 5, (1980) 1073–1103.

[29] N. Chen, "Some empirical tests of the theory of arbitrage pricing," *Journal of Finance* **38** no. 2, (1983) 1393–1414.

[30] P. Dhrymes, I. Friend, and B. Gultekin, "A critical examination of the empirical evidence on the arbitrage pricing theory," *Journal of Finance* **39** no. 2, (1984) 323–346.

[31] G. Chamberlain and M. Rothschild, "Arbitrage and mean-variance analysis on large asset markets," *Econometrica* **51** no. 1, (1983) 1281–1301.

[32] C. Trzcinka, "On the number of factors in the arbitrage pricing model," *Journal of Finance* **41** no. 2, (1986) 347–368.

[33] S. J. Brown and M. I. Weinstein, "A new approach to testing asset pricing models: The bilinear paradigm," *Journal of Finance* **38** no. 2, (1983) 711–743.

[34] D. A. Conway and M. R. Reinganum, "Stable factors in security returns. identification through cross validation," *Journal of business and economic statistics* **6** no. 1, (1988) 1–15.

[35] S. J. Brown, "The number of factors in security returns," *Journal of Finance* **44** no. 5, (1989) 1247–1262.

[36] G. Connor and R. Korajczyk, "A test for the number of factors in an approximate factor model," *Journal of Finance* **48** no. 4, (1993) 1263–91.

[37] L. Laloux, P. Cizeau, J.-P. Bouchaud, and M. Potters, "Noise dressing of financial correlation matrices," *Physical Review Letters* **83** no. 1, (1998) 1467.

[38] P. Gopikrishnan, B. Rosenow, V. Plerou, and H. E. Stanley, "Quantifying and interpreting collective behavior in financial markets," *Physical Review E* **64** no. 1, (2001) 35106.

[39] V. Plerou, P. Gopikrishnan, B. Rosenow, L. A. N. Amaral, T. Guhr, and H. E. Stanley, "Random matrix approach to cross correlations in financial data markets," *Physical Review E* **65** no. 1, (2002) 66126.

[40] R. Michaud, "The Markowitz optimization enigma: Is optimized optimal?," *Financial Analysts Journal* **45** no. 1, (1989) 31–42.

[41] R. Allez and J.-P. Bouchaud, "Eigenvector dynamics: general theory and some applications," *Phys. Rev. E* **86** no. 4, (2012) 046201.

[42] G. H. Golub, "Some modified matrix eigenvalue problems," *SIAM Review* **2** (1973) 318–334.

[43] L. Laloux, P. Cizeau, J.-P. Bouchaud, and M. Potters, "Random matrix theory," *Risk Magazine* **12** no. 3, (1999) 69.

[44] J.-P. Bouchaud and M. Potters, "Financial applications of random matrix theory: a short review," *The Oxford handbook of Random Matrix Theory* (2011) 824–850.

[45] M. Potters, J.-P. Bouchaud, and L. Laloux, "Financial applications of random matrix theory: Old laces and new pieces," *Acta Physica Polonica B* **36** no. 9, (2005) 2767–2784.

[46] J. Bun, J.-P. Bouchaud, and M. Potters, "Cleaning large correlation matrices: tools from random matrix theory," *Physics Reports* **666** no. 1, (2017) 1–109.

[47] E. Fama and J. D. MacBeth, "Risk, return, and equilibrium: Empirical tests," *Journal of Political Economy* **81** no. 3, (1973) 607–36.

[48] M. Lambert and G. Hübner, "Size matters, book value does not! The Fama-French empirical CAPM revisited," (2014) .

[49] M. Lambert and G. Hübner, "Comoment risk and stock returns," *Journal of Empirical Finance* **23** (2013) 191–205.

[50] R. F.Kleina and V. K.Chow, "Orthogonalized factors and systematic risk decomposition," *The Quarterly Review of Economics and Finance* **53** no. 2, (2013) 175–187.

[51] P.-O. Löwdin, "On the nonorthogonality problem," *Advances in Quantum Chemistry* **5** (1970) 185–199.

[52] R. Benichou, Y. Lempérière, E. Sérié, J. Kockelkoren, P. Seager, J.-P. Bouchaud, and M. Potters, "Agnostic risk parity: Taming known and unknown-unknowns," *Journal of Investment Strategies* **6** no. 3, (2017) 1–12.

[53] D. Blitz, J. Huij, and M. Martens, "Residual momentum," *Journal of Empirical Finance* **18** no. 3, (2011) 506–521.

[54] R. Novy-Marx, "Operating leverage," *Review of Finance* **15** no. 1, (2011) 103–134.

[55] T. J. Moskowitz and M. Grinblatt, "Do industries explain momentum?," *The Journal of Finance* **54** no. 4, (1999) 1249–1290.

[56] S. Valeyre, D. Grebenkov, and S.Aboura, "Emergence of correlation of securities at short time scales," *Physica A* (submitted) .

[57] V. A. Marčenko and L. A. Pastur, "Distribution of eigenvalues for some sets of random matrices," *Math. USSR-Sbornik* **1** no. 4, (1967) 457–483.

[58] A. Moreira and T. Muir, "Volatility-managed portfolios," *Journal of Finance* **122** no. 4, (2017) 1611–1644.

[59] M. Pelger, *Essays in Financial Econometrics, Asset Pricing and Corporate Finance*. 2015.

[60] Y. Lemperiere, C. Deremble, T.-T. Nguyen, P. Seager, M. Potters, and J.-P. Bouchaud, "Risk premia: asymmetric tail risks and excess returns," *Quantitative Finance* **17** no. 1, (2017) 1–14.

[61] J.-P. Bouchaud, A. Matacz, and M. Potters, "The leverage effect in financial markets: retarded volatility and market panic," *Physical Review Letters* **87** no. 22, (2011) 228701.

[62] S. Valeyre, D. Grebenkov, S. Aboura, and Q. Liu, "The reactive volatility model," *Quanitiative Finance* **13** no. 11, (2013) 1697–1706.

[63] S. Valeyre, D. Grebenkov, and S. Aboura, "The reactive beta model," *Journal of Financial Research* (in revision) .

[64] C. Gourieroux, "Continuous time wishart process for stochastic risk," *Econometric Reviews* **25** no. 2-3, (2006) 177–217.

[65] D. S. Grebenkov and J. Serror, "Following a trend with an exponential moving average: Analytical results for a gaussian model," *Physica A: Statistical Mechanics and its Applications* **394** (2014) 288–30.

[66] D. Grebenkov and J. Serror, "Optimal allocation of trend following strategies," *Physica A: Statistical Mechanics and its Applications* **433** (2015) 107–125.

[67] T. W. Epps, "Comovements in Stock Prices in the Very Short Run," *Journal of the American Statistical Association* **74** (1979) .

[68] R. H. Litzenberger and K. Ramaswamy, "The effect of personal taxes and dividends on capital asset prices: Theory and empirical evidence," *Journal of Financial Economics* **7** no. 2, (1979) 163–195.

[69] V. T. Datar, N. Y. Naik, and R. Radcliffe, "Liquidity and stock returns: An alternative test," *Journal of Financial Markets* **1** no. 2, (1998) 203–219.

[70] R. Korajczyk and R. Sadka, "Pricing the commonality across alternative measures of liquidity," *Journal of Financial Economics* **87** no. 1, (2008) 45–72.

[71] W. Liu, "A liquidity-augmented capital asset pricing model," *Journal of Financial Economics* **82** no. 3, (2006) 631–671.

[72] M. Brennan and A. Subrahmanyam, "Market microstructure and asset pricing: On the compensation for illiquidity in stock returns," *Journal of Financial Economics* **41** no. 3, (1996) 441–464.

[73] Y. Amihud and H. Mendelson, "Asset pricing and the bid-ask spread," *Journal of Financial Economics* **17** no. 2, (1986) 223–249.

[74] V. Acharya and L. Pedersen, "Asset pricing with liquidity risk," *Journal of Financial Economics* **77** no. 2, (2005) 375–410.

[75] A. Frazzini and L. Pedersen, "Betting against beta," *Journal of Financial Economics* **111** no. 1, (2014) 1–25.

[76] N. Jegadeesh, "Evidence of predictable behavior of securities returns," *Journal of Finance* **45** no. 3, (1990) 881–898.

[77] G. W. Douglas, *Risk in the Equity Markets: An Empirical Appraisal of Market Efficiency.* University Microfilms, Inc., Ann Arbor, Michigan, 1968.

[78] A. Ali, L.-S. Hwang, and M. A. Trombley, "Arbitrage risk and the book-to-market anomaly," *Journal of Financial Economics* **69** no. 2, (2003) 355–373.

[79] F. Black, "Capital market equilibrium with restricted borrowing," *The Journal of Business* **45** no. 3, (1972) 444–55.

[80] R. A. Haugen and A. J. Heins, "Risk and the rate of return on financial assets: Some old wine in new bottles," *Journal of Financial and Quantitative Analysis* **10** no. 05, (1975) 775–784.

[81] R. A. Haugen and N. L. Baker, "The efficient market inefficiency of capitalization–weighted stock portfolios," *The Journal of Portfolio Management* **17** no. 3, (1991) 35–40.

[82] H. Hong and D. Sraer, "Speculative betas," *Journal of Finance* **71** no. 5, (2016) 2095–2144.

[83] M. P. Baker, B. Bradley, and R. Taliaferro, "The low beta anomaly: A decomposition into micro and macro effects," *Financial Analysts Journal* (forthcoming) .

[84] N.-F. Chen, R. Roll, and S. Ross, "Economic forces and the stock market," *The Journal of Business* **59** no. 3, (1986) 383–403.

[85] L. C. Bhandari, "Debt/Equity ratio and expected common stock returns: Empirical evidence," *Journal of Finance* **43** no. 2, (1988) 507–28.

[86] H. Desai, S. Rajgopal, and M. Venkatachalam, "Value-glamour and accruals mispricing: One anomaly or two?," *The Accounting Review* **79** no. 2, (2004) 355–385.

[87] T. G. Bali, R. F. Engle, and Y. Tang, "Dynamic conditional beta is alive and well in the cross section of daily stock returns," *Management Science* **63** no. 11, (2017) 3760–3779.

[88] Z. Da and M. C. Warachka, "Cashflow risk, systematic earnings revisions, and the cross-section of stock returns," *Journal of Financial Economics* **94** no. 3, (2009) 448–468.

[89] E. F. Fama and K. French, "Profitability, investment and average returns," *Journal of Financial Economics* **82** no. 3, (2006) 491–518.

[90] P. S. Mohanram, "Separating winners from losers among lowbook-to-market stocks using financial statement analysis," *Review of Accounting Studies* **10** no. 2–3, (2005) 133–170.

[91] M. Adler and B. Dumas, "International portfolio choice and corporation finance: A synthesis," *Journal of Finance* **38** no. 3, (1983) 925–84.

[92] K. C. Chan, N.-f. Chen, and D. A. Hsieh, "An exploratory investigation of the firm size effect," *Journal of Financial Economics* **14** no. 3, (1985) 451–471.

[93] William C. Barbee, Jr. and Sandip Mukherji and Gary A. Raines, "Do sales-price and debt-equity explain stock returns better than book-market and firm size?," *Financial Analysts Journal* **52** no. 2, (1996) 56–60.

[94] R. Balvers and D. Huang, "Productivity-based asset pricing: Theory and evidence," *Journal of Financial Economics* **86** no. 2, (2007) 405–445.

[95] R. Novy-Marx, "The other side of value: The gross profitability premium," *Journal of Financial Economics* **108** no. 1, (2013) 1–28.

[96] V. Lidskii, "The proper values of the sum and product of symmetric matrices (in Russian)," *Dokl. Akad. Nauk SSSR* **74** (1950) 769–772.

[97] G. Connor, R. Korajczyk, and R. T. Uhlaner, "A synthesis of two factor estimation methods," *Journal of Financial and Quantitative Analysis* **50** no. 04, (2015) 825–842.

[98] R. F. Engle, "Dynamic conditional beta," *Journal of Financial Econometrics* **14** no. 4, (2016) 643–667.

[99] J. Gatheral, T. Jaisson, and M. Rosenbaum, "Volatility is rough," *Quantitative Finance* **18** no. 6, (2018) 933–949.

[100] R. Allez and J.-P. Bouchaud, "Individual and collective stock dynamics: intra-day seasonalities,".

[101] H. Berument and H. Kiymaz, “The day of the week effect on stock market volatility,” *Journal of Economics and Finance* **25** no. 2, (2001) 181–193.

3. Time Scale Effect on Correlation between Securities at Long Time Horizon

Time Scale Effect on Correlation between Securities at Long Time Horizon

February 26, 2019

Abstract

The dependence of correlations on time scales larger than a week is difficult to measure due to insufficient data and strong noises. Here, such measurements are made possible by reducing the size of the correlation matrix to 24 risk factors. We observe that correlations continue to grow significantly. We propose a model for autocorrelation of increments of various risk factors that reproduces the scaling effect. While this model presents some inefficiencies that are more subtle than the alternative risk premia, it seems to be more robust.

1 Introduction

Valeyre, Grebenkov and Aboura (2018) measured and modeled a simple lead-lag effect on correlations at time scales from 1 minute to 1 day. However, such a measurement became too noisy above the 1 day time scale. This problem is well explained by the random matrix theory (Laloux *et al.* (1998)). Valeyre *et al.* (2018) proposed a practical solution of this problem by reducing the size of the correlation matrix to 24 major risk factors that reproduce the largest eigenvalues and their dynamics. This solution allows one to reduce noises and confirms the Epps effect up to the horizon of one year. The Epps effect (Epps (1979)) was observed initially only at the intraday time scales as noises were too problematic at daily scales. In our study, we model this effect differently by introducing an autocorrelation term into the returns of risk factors. This term of positive autocorrelations can be explained by a herd behavior of investors (Guedj and Bouchaud (2005); Michard and Bouchaud (2005); Cont and Bouchaud (2000); Wyart and Bouchaud (2007); Lux and Marchesi (1999)) and by lack of market liquidity implying that a move of market takes some time as an investor needs time for his transaction being executed. We rely on the model developed by Grebenkov and Serror (2014) that describes autocorrelations between different stock indices and explains the performances of trend following strategies and CTA funds over two centuries (Lemperiere *et al.* (2014)). This model of autocorrelation represents an inefficiency that is more subtle but also more robust than than the alternative risk premia. Harvey and Liu (2015) listed 316 potential factors from 313 articles published since 1967. The majority of these factors overlap with each other. Fama and French (2015) proposed a 5-factor model (size, book, cash, momentum et accrual). In order to justify such alternative risk premia (lack of liquidity, asymmetry), the financial theories are revised because these anomalies tend to disappear after their publication. McLean and Ponti (2015) propose multiple explainations: the bias in sample, along with problems in optimization and adaptation to markets. To our knowledge, no prior works attempted to reveal autocorrelations of risk factors that may present an inefficiency that is more subtle but also more robust in financial markets. Strategies based on factor timing are however well documented but do not use only trend indicators and there is not consensus about their robustness and profitability (Asness (2016); Lee (2017); Bender *et al.* (2018); Bass, Gladstone et Ang (2017); DeMiguel *et al.* (2017); Hodges *et al.* (2017); Dichtl *et al.* (2018); Brandt, Santa-Clara et Valkanov (2009)). Our autocorrelation model could support the profitability of strategies using a factor timing based on trend following signals.

2 Model and measures

We project the correlation matrix between single stocks onto the optimized subspace of 24 maximum variance portfolios introduced in Valeyre *et al.* (2018). The eigenvalues of the full correlation matrix, called "unconstrained eigenvalues", are then close to the "constrained eigenvalues" of the projected matrix $C(\tau) = \gamma_0^{-1/2} h_0(\tau) \gamma_0^{-1/2}$. Here h_0 and γ_0 are the 24×24 covariance and overlap matrices of 24 factors that were also introduced in Valeyre *et al.* (2018). τ is the time scale. We suppose that only the correlations between stocks depend on the time scale, whereas the matrix γ_0 and stocks' volatilities do not depend on τ. We also suppose that the correlation between factors ($Corrcov(h_0(\tau))$) does not depend on τ (see the plot on Fig. 1(bottom, right) that partly validates this hypothesis).

We suppose that there is an inertia in the returns of the Maximum variance portfolios that represent the main elements of the systematic trading. This inertia can be explained either via the herding effect, or by lack of liquidity. We start from the autocorrelation model developed in Grebenkov and Serror (2014) but apply it directly to the returns of the Maximum variance portfolios which are normalized in such a way that the diagonal elements of the matrix $\gamma(t)$ are equal to 1 in every time moment. Denoting r_{kt} the return of the k-th factor at day t, we get

$$r_{kt} = \epsilon_{kt} + \kappa_k \sum_{i=-\infty}^{t} (1-\chi)^{t-1-i} \xi_{ki} \tag{1}$$

where ϵ and ξ are Gaussian random variables without autocorrelation. One can show that

$$C(\tau) = \gamma_0^{-1/2} V_{\tau\infty}^{-1/2} h_0(1) V_{\tau\infty}^{-1/2} \gamma_0^{-1/2} \tag{2}$$

where $V_{\tau\infty}$ is a diagonal matrix where every element in the diagonal is derived from Eq. (9) of Grebenkov and Serror (2014):

$$V_k(\tau) = 1 + \frac{2\,(1-\chi)\,[\kappa_k^0]^2}{\chi\,(1+[\kappa_k^0]^2)} \left(1 - \frac{1-(1-\chi)^\tau}{\tau\chi}\right) \tag{3}$$

with

$$\kappa_k = \kappa_k^0 \sqrt{\chi\,(2-\chi)} \tag{4}$$

One can show that the eigenvalues at the time scale τ correspond to the eigenvalues at time scale of 1 day multiplied by the same coefficient $V_{\tau\infty}$ if all the factors have the same autocorrelation $\kappa_k = \kappa$.

The problem is getting more complicated when some factors are more correlated than the others. It seems to be the case as the sectorial factors are less cross-correlated than the style factors. This changes the eigenvectors as well:

- The measures (see Fig. 1) rely on the methods introduced in Valeyre *et al.* (2018) to estimate the returns of the maximum variance factors, $\gamma_0(\tau)$ and $h_0(\tau)$, then $C(\tau)$. One can use the method introduced in Grebenkov and Serror (2014) to estimate the variograms. However, the variagrams estimated from 20 years of historical daily data, are very noisy. The differences in measurement among factors are in agreement with the noise. On average, they show that the factor returns are positively correlated. The FCLs, introduced in Valeyre *et al.* (2018), increase for style factors but remain stable for sectorial ones. Finally, the eigenvalues increase with the time scale. The ranking of style factors according to their FCL is very sensitive to the time scale (see Fig. 3). Note that the Book and Capitalization factors appear in the 4th and 5th places after the Beta STR and Momentum factors, whereas they were ranked to be less relevant at the time scale of one hour.

- The simulation of the model (Fig. 2) is based on the matrices γ_0 et h_0 generated randomly only once by the method described in section 2.8 of Valeyre *et al.* (2018) from the empirical eigenvalues of the averaged correlation matrix of daily returns. The style and sectorial factors are generated randomly only once, with $\kappa^0 = 0.08$ for 13 style factors and $\kappa^0 = 0$ for 9 sectorial factors. One fixes $\chi = 0.011$ to reproduce the measurements correctly. It turns out that the parameters correspond also to the model of autocorrelation of the DowJones index that could be estimated accurately due to its 120 years of data, see Grebenkov and Serror (2014).

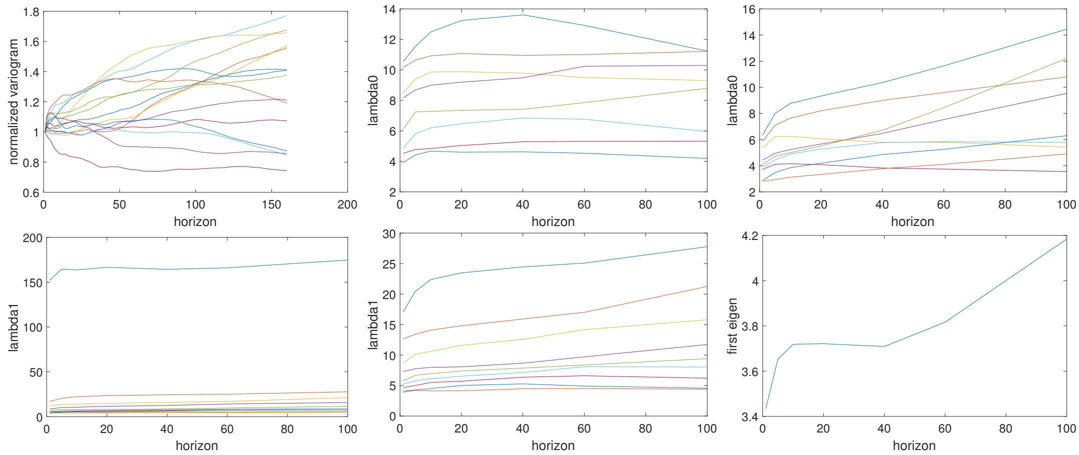

Figure 1: Top left: measurement of the variograms of the style factors; Top Middle: FCLs of sectorial factors that do not increase (i.e., these factors are just weakly autocorrelated); Top Right: FCLs of style factors that grow, particularly, the book factor by Fama and French that arrives at the 5-th place at long time scales (strong herding); Bottom left: constrained eigenvalues; Bottom Middle: constrained eigenvalues without the largest one. One can see that correlation keep strongly growing with the time scale; Bottom right: the first eigenvalue of $Corrcov(C)$ that is equivalent to $1 + 24\rho(\tau)^2$ with $\rho(\tau)$ being the average correlation between factors.

3 Conclusion

The dependence of correlations on time scales larger than a week is difficult to measure due to insufficient data and strong noises. Here, such measurements are made possible by reducing the size of the correlation matrix to 24 risk factors. We observe that correlations continue to grow significantly. We propose a model for autocorrelation of increments of various risk factors that reproduces the scaling effect. While this model presents some inefficiencies that are more subtle than the alternative risk premia, it seems to be more robust.

References

Asness, C. S., “The Siren Song of Factor Timing aka Smart Beta Timing aka Style Timing.”, Journal of Portfolio Management, (2016).

Bass, R., S. Gladstone, and A. Ang, “Total Portfolio Factor, not just Asset,Allocation.”, Journal of Portfolio Management, 43(5), 38. (2017).

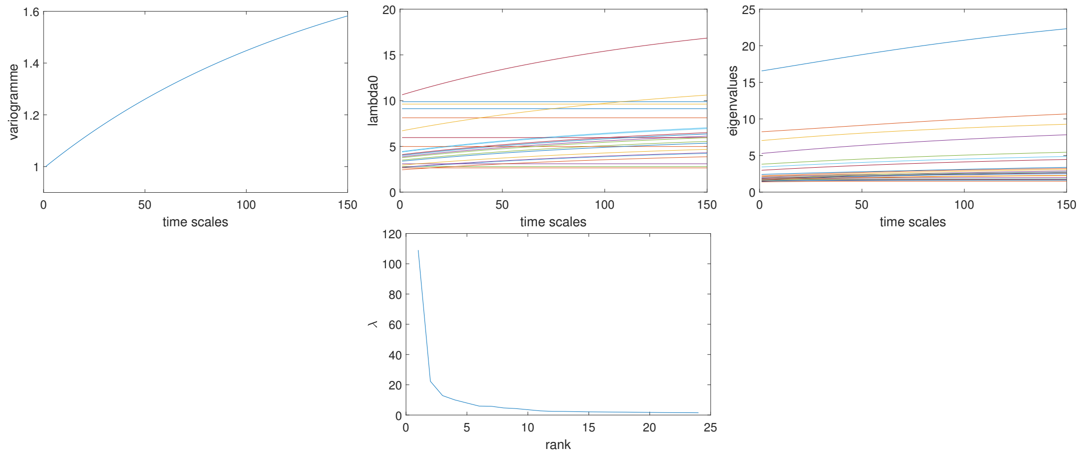

Figure 2: Top left: simulation of variograms of style factors with $\kappa^0 = 0.08$ and $\chi = 0.011$; Top middle: simulation of FCL (note that for sectorial factors $\kappa^0 = 0$); Top right: simulation of the eigenvalues; Bottom: empirical eigenvalues that were used an input for Monte Carlo simulation to generate γ_0 and $h_0(1)$; one can see that the model accurately reproduces the measurements.

Bender, J., X. Sun, R. Thomas, and V. Zdorovtsov, "The Promises and Pitfalls of Factor Timing.", Journal of Portfolio Management, 44(4), 79.(2018).

Brandt, M. W., P. Santa-Clara, and R. Valkanov, "Parametric Portfolio Policies: Exploiting Characteristics in the Cross-Section of Equity Returns.", Review of Financial Studies, 22(9), 3411 (2009).

Cont, R., and Bouchaud, J.-P., "Herd Behavior and Aggregate Fluctuations in Financial Markets.", Macroeconomic Dynamics Get access Volume 4, Issue 2 June (2000) , pp. 170-196.

Dichtl, H, W. Drobetz, H. Lohre, C. Rother, and P. Vosskamp, "Optimal Timing and Tilting of Equity Factors", Working Paper (2018).

Epps,T. W., "Comovements in Stock Prices in the Very Short Run.", J. Am. Stat. Asso. 74, 291-298 (1979).

Fama, E. F., and K. R. French, "A Five Factor Model.", Journal of Financial Economics, 116 (2015), pp. 1-22.

Grebenkov, D., and J. Serror, "Following a Trend with an Exponential Moving Average: Analytical Results for a Gaussian Model", Physica A. (2014).

factor	1 day	5 days	10 days	20 days	40 days	80 days	100 days
Beta	6.35	7.99	8.77	9.34	10.37	11.64	14.45
STR	4.10	4.73	5.00	5.51	6.72	8.45	12.21
Momentum	5.87	7.09	7.64	8.20	9.00	9.60	10.80
Capitalization	4.43	4.95	5.24	5.66	6.48	7.53	9.54
Book	2.49	2.94	3.40	3.99	4.89	5.50	6.40
Sales	2.88	3.47	3.85	4.19	4.85	5.26	6.30
Dividend	3.86	4.44	4.90	5.28	5.77	5.83	5.79
5Y Rates	5.38	6.24	6.26	6.06	5.80	5.80	5.42
Liquidity	2.82	2.93	3.12	3.34	3.77	4.11	4.91
Euro	3.72	4.12	4.15	4.08	3.82	3.74	3.55
Leverage	2.08	2.30	2.54	2.72	2.93	3.10	3.41
Earning	1.89	2.09	2.20	2.24	2.40	2.59	3.17
Cash	1.52	1.66	1.72	1.73	1.85	2.04	2.46
Growth	1.47	1.67	1.76	1.78	1.92	2.00	2.00

Figure 3: FCL at different time scales.

Guedj, O. , and Bouchaud, J.-P., “Experts Earning Forecasts: Bias, Herding and Gossamer Information”, International Journal of Theoretical and Applied Finance Vol. 08, No. 07, pp. 933-946 (2005).

Harvey, C.R., and Liu, Y., “Lucky factors.”, Working paper, University of Duke.(2015).

Hodges, P., K. Hogan, J. R. Peterson, and A. Ang, “Factor Timing with Cross-Sectional and Time-Series Predictors.”, Journal of Portfolio Management, 44(1), (2017).

Laloux L., Cizeau P., Bouchaud J.-P., and Potters M., “Noise Dressing of Financial Correlation matrices.”, Physical Review Letters 83, 1467. (1998).

Lee, W. (2017), “Factors Timing Factors”, Journal of Portfolio Management, 43(5), (2017).

Lemperiere,Y., C. Deremble, P. Seager, M. Potters, and J.-P. Bouchaud, “Two Centuries of Trend Following"", Journal of Investment Strategies, Volume3, number 3 June (2014).

Lux, T., and Marchesi, M., “Scaling and Criticality in a Stochastic Multi-Agent Model of a Financial Market.”, Nature volume 397, pages 498âĂŞ500 11 February (1999).

McLean, R.D., and Ponti, J. , “Does Academic Research Destroy Stock Return Predictability?.”, Forthcoming, Journal of Finance (2015).

Michard, Q., and Bouchaud, J.-P., “Theory of Collective Opinion Shifts: from Smooth Trends to Abrupt Swings.”, The European Physical Journal B - Condensed Matter and Complex Systems, September (2005), Volume 47, Issue 1, pp 151âĂŞ159.

DeMiguel,V., Martin-Utrera, F.J. Nogales, and R. Uppal, “A Portfolio Perspective on the Multitude of Firm Characteristics.”, LBS working paper (2017). Best Paper Award at the XXIV Finance Forum, presented at 2017 EFA Annual Meeting and 2018 AFA Annual Meeting (2017).

Valeyre, S., S. Kuperstein, D. S. Grebenkov, and S. Aboura, “The Market Neutral Fundamental Maximum Variance Portfolios” Working paper (2018).

Valeyre, S., D. S. Grebenkov, and S. Aboura, “Emergence of Correlation between Securities at Short Time Scales” Submited to Physica A (2018).

Wyart, M., and Bouchaud, J.-P., “Self-Referential Behaviour, Overreaction and Conventions in Financial Markets.”, Journal of Economic Behavior & Organization Volume 63, Issue 1, May (2007), Pages 1-24.

4. The Reactive Beta Model

The Reactive Beta Model

February 14, 2019

Abstract

We present a reactive beta model that accounts for the leverage effect and beta elasticity. For this purpose, we derive a correlation metric for the leverage effect to identify the relation between the market beta and volatility changes. An empirical test based on the most popular market neutral strategies is run from 2000 to 2015 with exhaustive data sets, including 600 US stocks and 600 European stocks. Our findings confirm the ability of the reactive beta model to withdraw an important part of the bias from the beta estimation and from most popular market neutral strategies. To examine the robustness of the reactive beta measurement, we conduct Monte Carlo simulations over seven market scenarios against five alternative methods. The results confirm that the reactive model significantly reduces the bias overall when financial markets are stressed.

Keywords: Beta, Correlation, Volatility, Leverage effect, Market Neutral Strategies.
JEL classification: C5, G01, G11, G12, G32.

1 Introduction

Finding an appropriate measurement of market betas is of paramount importance for many financial applications, including market neutral hedge fund managers who target a near-zero beta. Contrary to common belief, perfect beta neutral strategies are difficult to achieve in practice, as the mortgage crisis in 2008 exemplified, when most market neutral funds remained correlated with stock markets and experienced considerable unexpected losses. This exposure to the stock index (Banz, 1981; Fama and French, 1992, 1993; Carhart, 1997; Ang *et al.*, 2006) is even stronger during down market conditions (Mitchell and Pulvino, 2001; Agarwal and Naik, 2004; Bussière *et al.*, 2015). In such a period of market stress, hedge funds may even add no value (Asness *et al.*, 2001).

In this paper, we derive a stock market beta measure that we implement to test the quality of hedging for four popular strategies in the hedge funds industry. The first and most important strategy captures the low beta anomaly (Black, 1972; Black *et al.*, 1972; Haugen and Heins, 1975; Haugen and Baker, 1991; Ang *et al.*, 2006; Baker *et al.*, 2013; Frazzini and Pedersen, 2014; Hong and Sraer, 2016) that defies conventional wisdom on the risk and reward trade-off predicted by the CAPM (Sharpe, 1964). According to this anomaly, high beta stocks underperform low beta stocks. Similarly, stocks with high idiosyncratic volatility earn lower returns than stocks with low idiosyncratic volatility (Malkiel and Xu, 1997; Goyal and Santa-Clara, 2003; Ang *et al.*, 2006, 2009). The related strategy consists of shorting high beta stocks and buying low beta stocks. The second important strategy captures the size effect (Banz, 1981; Reinganum, 1981; Fama and French, 1992), in which stocks of small firms tend to earn higher returns, on average, than stocks of larger firms. The related strategy consists of buying stocks with small market capitalization and shorting those with high market capitalization. The third strategy captures the momentum effect (Jegadeesh and Titman, 1993; Carhart, 1997; Grinblatt and Moskowitz, 2004; Fama and French, 2012), where past winners tend to continue to show high performance. This strategy consists of buying the past year's winning stocks and shorting the past year's losing ones. The fourth strategy captures the short-term reversal effect (Jegadeesh, 1990), where past winners in the last month tend to show low performance. This strategy consists of buying the past month's losing stocks and shorting the past month's winning stocks, which would be highly profitable if there were no transaction cost and no market impact. Testing the quality of the hedge of the strategies is equivalent to assessing the quality of the beta measurements, which is difficult to realize directly as the true beta is not known.

The implementation of all these strategies requires a reliable estimation of the betas to maintain the hedge. Ordinary least squares (OLS) estimation remains the most frequently employed method, even though it is impaired in the presence of outliers, especially from small companies (Fama and French, 2008), illiquid companies (Amihud, 2002; Acharyaa and Pedersen, 2005; Ang *et al.*, 2013), and business cycles (Ferson and Harvey, 1999). In these circumstances, the OLS beta estimator might be inconsistent. To overcome these limitations, our approach consists of renormalizing the returns to make them closer to Gaussian and thus to make the OLS estimator more consistent. In addition, many papers report that betas are time varying (Blume, 1971; Fabozzi and Francis, 1978; Jagannathan and Wang, 1996; Fama and French, 1997; Bollerslev *et al.*, 1988; Lettau and Ludvigson, 2001; Lewellen and Nagel, 2006; Ang and Chen, 2007; Engle, 2016). This can lead to measurement errors that could create serious bias in the cross-sectional asset pricing test (Shanken, 1992; Chan and Lakonishok, 1992; Meng *et al.*, 2011; Bali *et al.*, 2017) . In fact, firms' stock betas do change over time for several reasons. The firm's assets tend to vary over time via acquiring or replacing new businesses, which makes them more diversified. The betas also change for firms that change in dimension to be safer or riskier. For instance, financial leverage may increase when firms become larger, as they can issue more debt. Moreover, firms with higher leverage are exposed to a more unstable beta (Galai and Masulis, 1976; DeJong and Collins, 1985). One way to account for the time dependence of betas is to consider regime changes when the return history used in the beta estimation is long enough. Surprisingly, only one paper (Chen *et al.*, 2005) suggests a solution to capture the time dependence and discusses regime changes for the beta using a multiple structural change methodology. The study shows that the risk related to beta regime changes is rewarded by higher returns. Another approach is to examine the correlation dynamics. Francis (1979) finds that "the correlation with the market is the primary cause of changing betas... the standard deviations of individual assets are fairly stable". This finding calls for special attention to the correlation dynamics addressed in our paper but that are apparently insufficiently investigated in other works.

Despite the extensive literature on this issue, little attention has been paid to the link between the leverage effect[1] and the beta. The leverage effect is defined as the negative correlation between the securities' returns and their volatility changes. This correlation induces residual correlations between the stock overperformances and beta changes. In fact, earlier studies have heavily focused on the role of the leverage effect on volatility (Black, 1976; Christie, 1982; Campbell and Hentchel, 1992; Bekaert and Wu, 2000; Bouchaud *et al.*, 2001; Valeyre *et al.*, 2013). Surprisingly, despite its theoretical and empirical underpinnings, the leverage effect has not been considered so far in beta modeling, while it is a measure of risk. We aim to close this gap. Our paper starts by investigating the role of the leverage effect in the correlation measure by extending the reactive volatility model (Valeyre *et al.*, 2013),

[1]Note that we are not dealing with the restricted definition of the "leveraged beta" that comes from the degree of leverage in the firm's capital structure. Notice that the market beta may be non-linearly related to the market return, which could lead to spurious inference in beta measurement (DeBondt and Thaler, 1987) while leverage effect could possibly be a major explanation of such non-linearity (e.g. Garlappi and Yan (2011) relate leverage to default probability; Daniel, Jagannathan and Kim (2012) relate the financial leverage to the operating leverage; Choi (2013) relates leverage to economic conditions; Moreira and Muir (2017) relate leverage and volatility managed portfolios; Liu, Stambaugh and Yuan (2018) relate the leverage to the beta-idiosyncratic volatility relation). In this context, the time-variation effect in conditional beta adds on this bias (Boguth *et al.*, 2011).

which efficiently tracks the implied volatility by capturing both the retarded effect induced by the specific risk and the panic effect, which occurs whenever the systematic risk becomes the dominant factor. This allows us to set up a reactive beta model incorporating three independent components, all of which contribute to a reduction in the bias of the hedging. First, we take into account the leverage effect on beta, where the beta of underperforming stocks tends to increase. Second, we consider a leverage effect on correlation, in which a stock index decline induces an increase in correlations. Third, we model the relation between the relative volatility (defined as the ratio of the stock's volatility to the index's volatility) and the beta. When the relative volatility increases, the beta increases as well. All three independent components contribute to a reduction in the biases in the naive regression estimation of the beta and therefore considerably improve hedging strategies.

The main contribution of this paper is the formulation of a *reactive beta model*. The economic intuition behind the reactive beta model is the derivation of a suitable beta measure allowing market beta estimation with reduced bias and a smaller standard deviation. The model is coined "reactive" because the beta measurement is adjusted as soon as prices move. An empirical test is performed based on an exhaustive dataset that includes the 600 largest American stocks and the 600 largest European stocks over the period from 2000 to 2015, which includes several business cycles. This test validates the superiority of the reactive beta model over conventional methods.

We further examine the robustness of the reactive beta measurement using Monte Carlo simulation against five alternative methods (ordinary least squares, minimum absolute deviation, trimean quantile regression, and dynamic conditional correlation with or without asymmetry) over seven scenarios that reflect various market conditions from calm (Gaussian universe) to stressed (Non-Gaussian universe). The results confirm that the reactive beta presents a lower bias when stressed market conditions are included.

The article is organized as follows. Section 2 outlines the methodology employed for the reactive beta model. Section 3 describes the data and empirical findings. Section 4 provides several robustness checks to assess the quality of the reactive beta model against alternative methods. Section 5 expands the discussion beyond the field of portfolio management, while Section 6 concludes.

2 The reactive beta model

In this section, we present the reactive beta model with three independent components. First, we take into account the specific leverage effect on the beta. Second, we consider the systematic leverage effect on the correlation. Third, we model the relation between the relative volatility and the beta via nonlinear beta elasticity.

2.1 The leverage effect on beta

We first account for relations among returns, volatilities, and the beta, which are characterized by the so-called leverage effect. This component takes into account the phenomenon where a beta increases as soon as a stock underperforms the index. Such a phenomenon can be fairly well described by the leverage effect captured in the reactive volatility model. We call the *specific leverage effect* the negative relation between specific returns and the risk (here, the beta), where the specific return is the nonsystematic part of the returns (a stock's overperformance). The specific leverage effect on the beta follows the same dynamics as the specific leverage effect introduced in the reactive volatility model.

2.1.1 The reactive volatility model

This section aims to capture the dependence of betas on stock overperformance (when a stock is overperforming, its beta tends to decrease). For this purpose, we rely on the methodology of the reactive volatility model (Valeyre *et al.*, 2013) to derive a stable measure of the beta by using the renormalization factor that depends on the stock's overperformance. The model describes the systematic and specific leverage effects. Systematic leverage, which is due to the panic effect, and specific leverage, which is due to a retarded effect, have very different relaxation times and intensities. These two different effects were investigated by Bouchaud *et al.* (2001), who introduced the measurement of the returns' volatility correlation function at different time scales τ. They defined this measurement as $\mathscr{L}(\tau) = E\left(r^2(t+\tau)r(t)\right)/E^2\left(r^2(t)\right)$, where $r(t)$ is the daily return at day t, and they showed that it exhibits an exponential decay curve depending on τ with 2 parameters: the relaxation time and the initial amplitude, which describes the intensity of the leverage. The intensity measured is 9 times higher for the stock index than for the single stocks, and the relaxation time is 6 times smaller for the stock index. The higher intensity and the shorter relaxation times applied to the stock index were explained by the panic effect that occurs as soon as all single stocks decrease at the same time. The low intensity and the longer relaxation time applied to single stocks were explained by the retarded effect: On short time scales, the standard deviation of differences in price is the criteria used by traders to assess the risk, whereas on longer time scales, the standard deviation of returns is used. The retarded effect works as if traders need time to take into account a change in price in the analysis of the risk. The reactive volatility model reproduces very well the measurement of $\mathscr{L}(\tau)$ for the stock index and for the single stocks.

We start by recalling the construction of the reactive volatility model, which explicitly accounts for the leverage effect on volatility. Let $I(t)$ be a stock index at day t. It is well known that arithmetic returns, $r_I(t) = \delta I(t)/I(t-1)$, are heteroscedastic, partly due to price-volatility correlations. Throughout the text, δ refers to the difference between successive values, e.g., $\delta I(t) = I(t) - I(t-1)$. The reactive volatility model aims to construct an appropriate "level" of the stock index, $L(t)$, to replace the original returns $\delta I(t)/I(t-1)$

with less-heteroscedastic returns $\delta I(t)/L(t-1)$.

For this purpose, we first introduce two "levels" of the stock index as exponential moving averages (EMAs) with two time scales: a slow level $L_s(t)$ and a fast level $L_f(t)$. In addition, we denote by $L_{is}(t)$ the EMA (with the slow time scale) of the price $S_i(t)$ of the stock i at time t. These EMAs can be computed using standard linear relations:

$$L_s(t) = (1-\lambda_s)L_s(t-1) + \lambda_s I(t), \tag{1}$$
$$L_f(t) = (1-\lambda_f)L_f(t-1) + \lambda_f I(t), \tag{2}$$
$$L_{is}(t) = (1-\lambda_s)L_{is}(t-1) + \lambda_s S_i(t), \tag{3}$$

where λ_s and λ_f are the weighting parameters of the EMAs that we set to $\lambda_s = 0.0241$ and $\lambda_f = 0.1484$, relying on the estimates by Bouchaud *et al.* (2001). The slow parameter corresponds to the relaxation time of the retarded effect for specific risk, whereas the fast one corresponds to the relaxation time of the panic effect for systematic risk. These two relaxation times are found to be rather universal, as they are stable over years and do not change among different mature stock markets. The appropriate levels, $L(t)$ and $L_i(t)$, accounting for the leverage effect on the volatility to correctly normalize the difference in price, were introduced for the stock index and individual stocks, respectively. [2]

$$L(t) = I(t)\left(1 + \frac{L_s(t) - I(t)}{I(t)}\right)\left(1 + \ell\,\frac{L_f(t) - I(t)}{L_f(t)}\right), \tag{4}$$
$$L_i(t) = S_i(t)\underbrace{\left(1 + \frac{L_{is}(t) - S_i(t)}{S_i(t)}\right)}_{\text{specific risk}}\underbrace{\left(1 + \ell_i\,\frac{L_f(t) - I(t)}{L_f(t)}\right)}_{\text{systematic risk}}, \tag{5}$$

with the parameters ℓ and ℓ_i quantifying the leverage. The parameter ℓ was introduced by Valeyre *et al.* (2013) to reproduce the exponential fit of the returns' volatility correlation function $\mathscr{L}(\tau)$ at different time scales τ. The initial parameters of the exponential fit were estimated on 7 major stock indexes so that ℓ was deduced to be approximately 8. If $\ell = \ell_i$, the correlation between the stock index and the individual stock i is not impacted by the leverage effect. In turn, if $\ell > \ell_i$, the correlation increases when the stock index decreases. Although ℓ_i can generally be specific to the considered i-th stock, we ignore its possible dependence on i and set $\ell_i = \ell'$. Using the levels $L(t)$ and $L_i(t)$, we introduce the normalized returns:

$$\tilde{r}_I = \tilde{r}_I(t) = \frac{\delta I(t)}{L(t-1)}, \qquad \tilde{r}_i = \tilde{r}_i(t) = \frac{\delta S_i(t)}{L_i(t-1)} \tag{6}$$

[2] In practice, a filtering function is introduced to attenuate the contribution from eventual outliers (extreme events or wrong data). The filter was applied to $z = \frac{L_s(t)-I(t)}{I(t)}$ and $z = \frac{L_{is}(t)-S_i(t)}{S_i(t)}$ in Eqs. (4, 5) and was defined as $F_\phi(z) = \tanh(\phi z)/\phi$ with $\phi = 3.3$ (in the limit $\phi = 0$, there is no filter: $F_0(z) = z$).

and compute the renormalized variances $\tilde{\sigma}_I^2$ and $\tilde{\sigma}_i^2$ through the EMAs as:

$$\begin{aligned} \tilde{\sigma}_I^2(t) &= (1-\lambda_\sigma)\tilde{\sigma}_I^2(t-1) + \lambda_\sigma \tilde{r}_I^2(t), \quad (7) \\ \tilde{\sigma}_i^2(t) &= (1-\lambda_\sigma)\tilde{\sigma}_i^2(t-1) + \lambda_\sigma \tilde{r}_i^2(t), \quad (8) \end{aligned}$$

where λ_σ is a weighting parameter that has to be chosen as a compromise between the accuracy of the estimated renormalized volatility and the reactivity of that estimation. Indeed, the renormalized returns are constructed to be homoscedastic only at short times because the renormalization based on the leverage effect with short relaxation times (λ_s, λ_f) cannot account for long periods of changing volatility related to economic cycles. Since economic uncertainty does not change significantly over a period of two months (40 trading days), we set λ_σ to $1/40 = 0.025$. This sample length leads to a statistical uncertainty of approximately $\sqrt{1/40} \approx 16\%$. Finally, these renormalized variances can be converted into the reactive volatility $\sigma_I(t)$ of the stock index quantifying the systematic risk governed by the panic effect, and the reactive volatility $\sigma_i(t)$ of each individual stock quantifying the specific risk governed by the leverage effect:

$$\begin{aligned} \sigma_I(t) &= \tilde{\sigma}_I(t)\frac{L(t)}{I(t)}, \quad (9) \\ \sigma_i(t) &= \tilde{\sigma}_i(t)\frac{L_i(t)}{S_i(t)}. \quad (10) \end{aligned}$$

This reactive volatility captures a large part of the heteroscedasticity, i.e., a large part of the volatility variation is completely explained by the leverage effect. That was the main result of Valeyre *et al.* (2013): For instance, if the stock index loses 1%, $\frac{L(t)}{I(t)}$ increases by $\ell \times 1\% = 8\%$, and the stock index volatility increases by 8%. That effect is enough to capture a large part of the VIX variation, with $R^2 = 0.46$. In turn, if the stock underperforms the stock index by 1%, $\frac{L_i(t)}{S_i(t)}$ increases by 1%, and the single stock volatility increases by 1%.

2.1.2 The specific leverage effect in the reactive beta model

The volatility estimation procedure naturally impacts the estimation of the beta. Many financial instruments rely on the estimated beta, β_i, which corresponds to the slope of a linear regression of the stocks' arithmetic returns r_i on the index arithmetic return r_I:

$$r_i = \beta_i r_I + \epsilon_i, \qquad \text{with} \quad r_i = \frac{\delta S_i(t)}{S_i(t-1)}, \quad r_I = \frac{\delta I(t)}{I(t-1)}, \quad (11)$$

where ϵ_i is the residual random component specific to stock i. We consider another beta estimate, $\tilde{\beta}_i$, based on the reactive volatility model, in which the renormalized stock returns $\tilde{r}_i$ are regressed on the renormalized stock index returns $\tilde{r}_I$:

$$\tilde{r}_i = \tilde{\beta}_i \tilde{r}_I + \tilde{\epsilon}_i, \qquad \text{with} \quad \tilde{r}_i = \frac{\delta S_i(t)}{L_i(t-1)}, \quad \tilde{r}_I = \frac{\delta I(t)}{L(t-1)}. \quad (12)$$

We then obtain a reactive beta measure:

$$\beta_i(t) = \tilde{\beta}_i(t)\frac{\sigma_i(t)\,\tilde{\sigma}_I(t)}{\sigma_I(t)\,\tilde{\sigma}_i(t)} = \tilde{\beta}_i\frac{L_{is}(t)I(t)}{L_s(t)S_i(t)}, \tag{13}$$

which includes two improvements:

- $\tilde{\beta}_i$, which becomes less sensitive to price changes by accounting for the specific leverage effect;
- $\sigma_i\tilde{\sigma}_I/(\sigma_I\tilde{\sigma}_i)$, which changes instantaneously with price changes.

When taking into account the short-term leverage effect in correlations, the reactive term is reduced to $\frac{L_{is}(t)I(t)}{L_s(t)S_i(t)}$. This term has a significant impact, as the beta of underperforming stocks should increase.

2.2 The systematic leverage effect on correlation

2.2.1 The empirical estimation of ℓ' for single stocks

We use the term *systematic leverage effect* to denote the negative relation between systematic returns and the risk (here, the correlation), where the systematic returns are the nonspecific part of the returns (stock index performance). The systematic leverage effect on the correlation follows the same dynamics as the systematic leverage effect introduced in the reactive volatility model (the phenomenon's duration is approximately 7 days for $\lambda_f = 0.1484$). All correlations are impacted together in the same way by the systematic leverage effect, and single stocks and their stock indexes should also shift in the same direction. This explains why the stock's beta will not change with respect to the index. The implication is that betas are not very sensitive to the systematic leverage effect, in contrast to the specific leverage effect. We consider the impact of the short-term systematic leverage effect on correlation. Assuming that the correlation between each individual stock and the stock index is the same for all stocks, one can define the implied correlation as: [3]

$$\rho(t) = \frac{\sigma_I^2(t) - \sum_i w_i^2\sigma_i^2(t)}{\sum_{i\neq j} w_i w_j \sigma_i(t)\sigma_j(t)}, \tag{14}$$

where w_i represents the weight of stock i in the index. Denoting

$$e_I(t) = \frac{\hat{L}_s(t)}{I(t)} - 1, \qquad e_i(t) = \frac{\hat{L}_{is}(t)}{S_i(t)} - 1, \tag{15}$$

[3] See http://www.cboe.com/micro/impliedcorrelation/impliedcorrelationindicator.pdf

we use Eqs. (9, 10) to obtain:

$$\rho = \frac{\tilde{\sigma}_I^2 (1+e_I)^2 \left(1+\ell \frac{L_f - I}{L_f}\right)^2 - \left(1+\ell' \frac{L_f - I}{L_f}\right)^2 \sum_i w_i^2 (1+e_i)^2 \sigma_i^2}{\left(1+\ell' \frac{L_f - I}{L_f}\right)^2 \sum_{i \neq j} w_i w_j \tilde{\sigma}_i \tilde{\sigma}_j (1+e_i)(1+e_j)}. \tag{16}$$

If the weights w_i are small, we can ignore the second term; in addition, if e_i are small, then

$$\sum_{i \neq j} w_i w_j \tilde{\sigma}_i \tilde{\sigma}_j (1+e_i)(1+e_j) \approx (1+e_I)^2 \tilde{\sigma}_0^2,$$

where $\tilde{\sigma}_0^2$ is an average of $\tilde{\sigma}_i^2$. Keeping only the leading terms of the expansion in terms of the small parameter $(L_f - I)/L_f$, one thus obtains

$$\rho \approx \frac{\tilde{\sigma_I}^2}{\tilde{\sigma_0}^2} \left(1 + 2(\ell - \ell') \frac{L_f - I}{L_f}\right). \tag{17}$$

This relation shows the dynamics of the implied correlation ρ induced by the leverage effect (accounted for through the factor $(L_f - I)/L_f$). We assume that the same dynamics are applicable to correlations between individual stocks, i.e.,

$$\rho_{i,j} = \tilde{\rho}_{i,j} \left(1 + 2(\ell - \ell') \frac{L_f - I}{L_f}\right), \tag{18}$$

where $\tilde{\rho}_{i,j}$ are the parameters specific to each pair of stocks i and j. From this relation, we derive a measure of correlation accounting for the leverage effect between the single stock i and the stock index:

$$\rho_i = \tilde{\rho}_i \left(1 + (\ell - \ell') \frac{L_f - I}{L_f}\right), \tag{19}$$

where $\tilde{\rho}_i$ are the parameters specific to each stock i. Note that there is no factor 2 in front of $(\ell - \ell')$ in Eq. (19) because we have a one-factor model here. We use Eq. (19) in the reactive beta model (see Eqs. (34, 36) below) to take into account the varying nature of the correlation in the regression. We rescale the measurement by the normalization factor $(1 + (\ell - \ell')(L_f - I)/L_f)$ and then recover the variation of the correlation through the denormalization factor $1/(1 + (\ell - \ell')(L_f - I)/L_f)$. We emphasize that the parameter ℓ in Eq. (4) that quantifies the systematic leverage for the stock index is slightly different from the parameter ℓ' in Eq. (5) that quantifies the systematic leverage for single stocks. According to Eq. (18), when the market decreases, correlations between stocks increase as $\ell > \ell'$, and therefore, the stock index volatility increases more than the single stock's volatility: $\delta(\sigma_i/\sigma_I) < 0$. Once again, the beta is, in contrast to the correlation, weakly impacted by the systematic leverage effect, as all correlations increase in the same way. More precisely, this

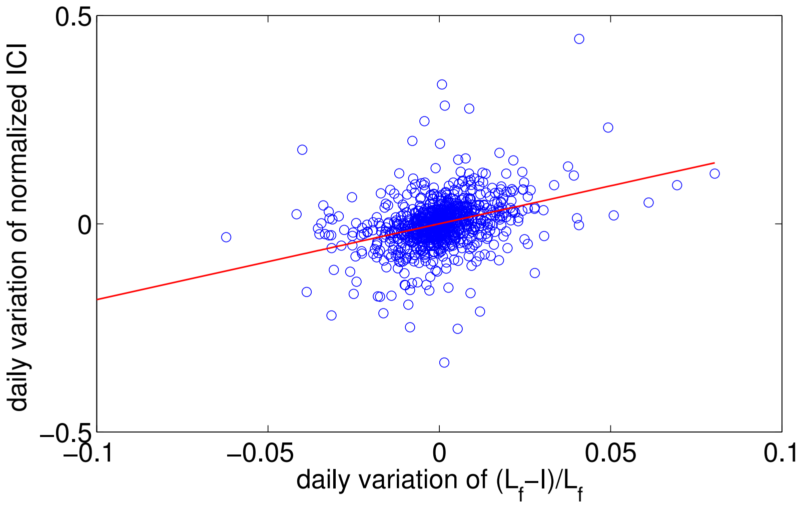

Figure 1: Daily variations of the CBOE S&P 500 Implied Correlation Indices (ICI) since their inception, divided by their mean, versus daily variations of the leverage factor $(L_f - I)/L_f$. A linear regression (solid line) yields the coefficient 1.82 ± 0.16 (i.e., $2(\ell - \ell') = 1.82$), with $R^2 = 0.13$ and t-statistics of 11.4. Period: 2007-2015.

means that the impact of the increase in correlation in the beta measurement is compensated by a decrease in the relative volatility: $\delta(\sigma_i/\sigma_I) < 0$, i.e., the single stock volatility increase is lower than that of the stock index volatility. For this reason, the reactive beta model in Eqs. (34, 36) is not very sensitive to the choice of ℓ'. Nevertheless, we explain in this section how ℓ' is calibrated using the implied volatility index. We measure the level of the systematic leverage effect ℓ' for a single stock by regressing Eq. (17) with data from the market-implied correlation S&P 500 index. Figure 1 illustrates the slope of this regression. By regressing $\frac{L_f - I}{L_f}$ against $\frac{\rho}{\tilde{\rho}_0}$, where $\tilde{\rho}_0$ is the average of ρ, we deduce that empirically, we can set:

$$\ell - \ell' = 0.91 \pm 0.08, \tag{20}$$

with a t-statistic of 11.4. Since $\ell - \ell' \ll \ell (= 8)$, we deduce an important result, namely, that the systematic leverage impact on the correlation is more than 8 times smaller than the systematic leverage impact on the volatility. The main consequence is that although it is statistically significant, the leverage effect is not a major component of the correlation.

2.2.2 The systematic leverage effect component in the reactive model

As discussed above, the correlation increases when the stock index price decreases. This effect could generate a bias in the beta measurement, as stock index prices could fluctuate in a sample used to measure the slope. Our solution is to adjust the beta between renormalized returns through the correction factor $\mathcal{L}(t)$, defined as

$$\mathcal{L}(t) = 1 + (\ell - \ell') \left(\frac{L_f(t-1) - I(t-1)}{L_f(t-1)} \right), \tag{21}$$

The correction factor $\mathcal{L}(t)$ should be used to estimate the slope between the stock index and single stock returns and should then be used to denormalize the slope to obtain the reactive beta that depends directly on $\mathcal{L}(t)$.

2.3 The relation between the relative volatility and beta

2.3.1 The empirical estimation of beta elasticity

In this part, we identify correlations between the relative volatility and beta changes. We choose the relative volatility, defined as the ratio $\tilde{\sigma}_i/\tilde{\sigma}_I$, as an explanatory variable of $\tilde{\beta}_i$ because $\tilde{\beta}_i$ is expected to be constant if the ratio $\tilde{\sigma}_i/\tilde{\sigma}_I$ is constant. However, empirically, the ratio $\tilde{\sigma}_i/\tilde{\sigma}_I$ can change dramatically between periods of high dispersion (i.e., when stocks are, on average, weakly correlated) and low systematic risk (i.e., when stock indexes are not stressed) and periods of low dispersion and high systematic risk. Figure 2 illustrates, for both European and US markets, that the dispersion among stocks decreases, on average, when markets become volatile. A linear regression of rescaled daily variations of $\tilde{\sigma}_i$ yields:

$$\frac{\delta\tilde{\sigma}_i(t)}{\tilde{\sigma}_i(t-1)} \approx 0.4\, \frac{\delta\tilde{\sigma}_I(t)}{\tilde{\sigma}_I(t-1)} + \epsilon_i, \tag{22}$$

where ϵ_i is the residual (specific) noise. Using the standard rules for infinitesimal increments, we find from this regression the following:

$$\delta\left(\frac{\tilde{\sigma}_i}{\tilde{\sigma}_I}\right) \simeq \frac{\delta\tilde{\sigma}_i}{\tilde{\sigma}_I} - \frac{\tilde{\sigma}_i\,\delta\tilde{\sigma}_I}{\tilde{\sigma}_I^2} = \frac{\tilde{\sigma}_i}{\tilde{\sigma}_I}\left(\frac{\delta\tilde{\sigma}_i}{\tilde{\sigma}_i} - \frac{\delta\tilde{\sigma}_I}{\tilde{\sigma}_I}\right) \simeq -0.6\frac{\tilde{\sigma}_i}{\tilde{\sigma}_I}\,\frac{\delta\tilde{\sigma}_I}{\tilde{\sigma}_I}, \tag{23}$$

i.e., the relative volatility $\tilde{\sigma}_i/\tilde{\sigma}_I$ is relatively stable, but its small variations can still impact the beta estimation. This empirical relation shows that when there is a volatility shock in the market, the stock index volatility increases much faster than the average single stock volatility.

Because we want to take into account the impact of the relative volatility change on the beta measurement, we introduce the beta elasticity as the derivative of the beta with respect to the logarithm of the squared relative volatility:

$$f(\tilde{\beta}_i) = \frac{d\tilde{\beta}_i}{d\ln(\tilde{\sigma}_i^2/\tilde{\sigma}_I^2)} = \frac{d\tilde{\beta}_i}{d(\tilde{\sigma}_i/\tilde{\sigma}_I)}\,\frac{\tilde{\sigma}_i}{2\tilde{\sigma}_I}. \tag{24}$$

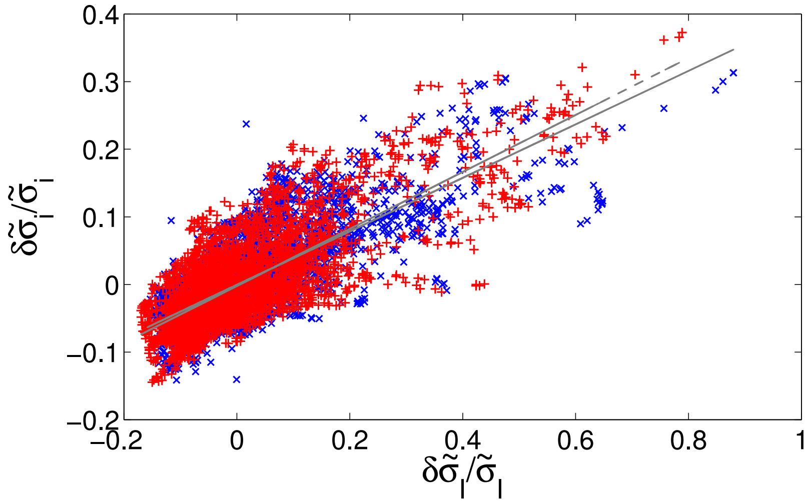

Figure 2: Normalized daily variations of $\tilde{\sigma}_i$, $\delta\tilde{\sigma}_i/\tilde{\sigma}_i = \frac{\tilde{\sigma}_i(t)-\tilde{\sigma}_i(t-1)}{\tilde{\sigma}_i(t-1)}$ versus normalized daily variations of $\tilde{\sigma}_I$, $\delta\tilde{\sigma}_I/\tilde{\sigma}_I = \frac{\tilde{\sigma}_I(t)-\tilde{\sigma}_I(t-1)}{\tilde{\sigma}_I(t-1)}$ for the European market (blue crosses) and the US market (red pluses). The two gray lines show the linear regressions of the two datasets, with regression coefficients of 0.40 ($R^2 = 0.60$) and 0.42 (with $R^2 = 0.59$) for the European and US markets, respectively. The time frame includes observations from the technology bubble burst, U.S. subprime, and Euro debt crises. Period: 1998-2015.

We expect that $f(\tilde{\beta}_i)$ is positive and increasing with $\tilde{\beta}_i$. Indeed, we expect that a stock with a low beta should have a stable beta (less sensitive to its relative volatility increase), as the increase in this case is most likely due to a specific risk increase. In such a case, the sensitivity of the beta to the relative volatility is weak. In the opposite case of a high beta, a stock that is highly sensitive to the stock index will face a beta decline as soon as its relative volatility decreases. Consequently, when there is a volatility shock in the market, $\delta(\frac{\tilde{\sigma}_i}{\tilde{\sigma}_I})$ is negative, and therefore, the beta of stocks with a high beta and a high f is reduced. In turn, the stocks with a low beta are less impacted because f is smaller and $\delta(\tilde{\sigma}_i/\tilde{\sigma}_I)$ is expected to be less negative.

When the correlation of the stock with the stock index is constant, we can use a linear model: $f(\tilde{\beta}_i) = \tilde{\beta}_i/2$. In fact, using the relation $\tilde{\beta}_i = \tilde{\rho}_i \frac{\tilde{\sigma}_i}{\tilde{\sigma}_I}$ and the assumption that $\tilde{\rho}_i$ is constant (i.e., it does not depend on $\frac{\tilde{\sigma}_i}{\tilde{\sigma}_I}$), one obtains from Eq. (24) $f = \tilde{\rho}_i \frac{\tilde{\sigma}_i}{2\tilde{\sigma}_I} = \tilde{\beta}_i/2$. In general, however, the correlation can depend on the relative volatility, and thus, the function f may be more complicated. To estimate f, one needs the renormalized beta and the relative volatility. For a better estimation, we aim to reduce the heteroscedasticity even further by using an exponential moving regression of the returns $\tilde{r}_i$ and $\tilde{r}_I$ that are renormalized by the estimated normalized index volatility $\tilde{\sigma}_I$. We denote these renormalized returns as:

$$\hat{r}_i(t) = \frac{\tilde{r}_i(t)}{\tilde{\sigma}_I(t-1)}, \qquad \hat{r}_I(t) = \frac{\tilde{r}_I(t)}{\tilde{\sigma}_I(t-1)}. \tag{25}$$

Computing the EMAs,

$$\hat{\phi}_i(t) = (1-\lambda_\beta)\hat{\phi}_i(t-1) + \lambda_\beta \hat{r}_i(t)\,\hat{r}_I(t), \tag{26}$$

$$\hat{\sigma}_I^2(t) = (1-\lambda_\beta)\hat{\sigma}_I^2(t-1) + \lambda_\beta \left[\hat{r}_I(t)\right]^2, \tag{27}$$

with $\lambda_\beta = 1/90$, we estimate the beta as:

$$\hat{\beta}_i(t) = \frac{\hat{\phi}_i(t)}{\hat{\sigma}_I^2(t)}. \tag{28}$$

Here, $\hat{\phi}_i$ is an estimation of the covariance between stock index returns and single stock returns that includes two normalizations: the levels L_i and L from the reactive volatility model and $\tilde{\sigma}_I$ to further reduce heteroscedasticity. We write $\hat{\beta}_i$ instead of $\tilde{\beta}_i$ to stress this particular way of estimating the beta. Similarly, the hat symbol in Eq. (27) is used to distinguish $\hat{\sigma}_I(t)$, computed with renormalized index returns, from $\tilde{\sigma}_I(t)$. In principle, the above estimate $\tilde{\beta}$ could be directly regressed to the ratio of earlier estimates of $\tilde{\sigma}_i$ and $\tilde{\sigma}_I$ from Eqs. (7). However, to use the normalization by $\tilde{\sigma}_I$ consistently, we consider the ratio of these volatilities obtained in the renormalized form, i.e., $\hat{\sigma}_i(t)/\hat{\sigma}_I(t)$, where $\hat{\sigma}_I(t)$ is given in Eq. (27), and

$$\hat{\sigma}_i^2(t) = (1-\lambda_\beta)\hat{\sigma}_i^2(t-1) + \lambda_\beta \big[\hat{r}_i(t)\big]^2. \tag{29}$$

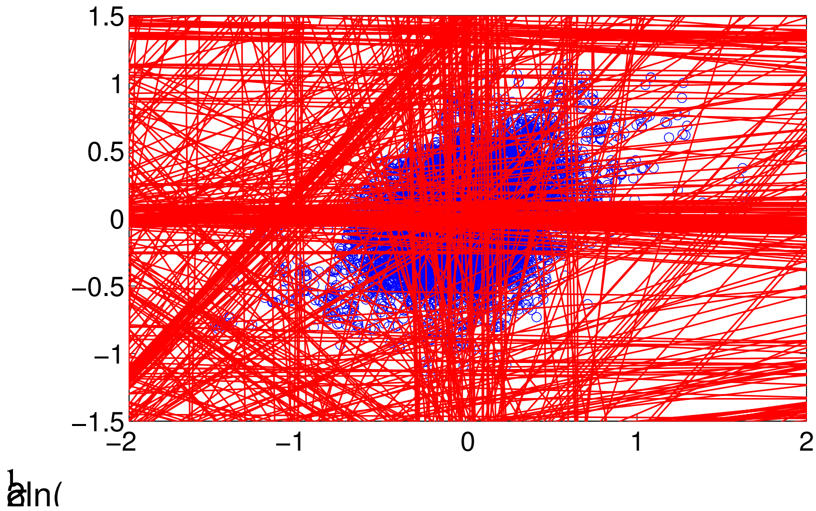

Figure 3: Relation between the beta $\hat{\beta}_i$ and the doubled logarithm of the relative volatility $\ln(\hat{\sigma}_i/\hat{\sigma}_I)$, from which the mean values $\langle\hat{\beta}_i\rangle$ and $\ln(\langle\hat{\sigma}_i/\hat{\sigma}_I\rangle)$ were subtracted (the mean is obtained by averaging over time for each i). A linear regression is shown by the solid line: $\hat{\beta}_i - \langle\hat{\beta}_i\rangle = 0.76\big[\ln(\hat{\sigma}_i/\hat{\sigma}_I) - \ln(\langle\hat{\sigma}_i/\hat{\sigma}_I\rangle)\big]$, with $R^2 = 0.14$. For better visualization, only 10,000 randomly selected points are shown (by circles) among 271,958 points from the European dataset. Period: 2014-2015.

Figure 3 illustrates the sensitivity of the beta to relative volatilities by plotting $\hat{\beta}_i(t)$ from Eq. (28) versus $\ln(\hat{\sigma}_i(t)/\hat{\sigma}_I(t))$ for all stocks i and times t from 2000 to 2015, although we only display the time frame of 2014-2015 for clarity of illustration. On both axes, we subtract the mean values $\langle\hat{\beta}_i\rangle$ and $\ln(\langle\hat{\sigma}_i/\hat{\sigma}_I\rangle)$ averaged over all times in the whole sample. This plot enables us to measure the average of the $f(\tilde{\beta}_i)$ in Eq. (24), which is close to $0.76/2 = 0.38$.

To obtain the dependence of f on the beta, we estimate the slope between $\hat{\beta}_i(t) - \langle\hat{\beta}_i\rangle$ from Eq. (28) and $2\ln(\hat{\sigma}_i(t)/\hat{\sigma}_I(t)) - 2\ln(\langle\hat{\sigma}_i/\hat{\sigma}_I\rangle)$ *locally* around each value of $\hat{\beta}_i$. For this purpose, we sort all collected values of $\hat{\beta}_i$ and group them into successive subsets, each with 10,000 points. In each subset, we estimate the slope between $\hat{\beta}_i(t) - \langle\hat{\beta}_i\rangle$ from Eq. (28) and $2\ln(\hat{\sigma}_i(t)/\hat{\sigma}_I(t)) - 2\ln(\langle\hat{\sigma}_i/\hat{\sigma}_I\rangle)$ by a standard linear regression over 10,000 points. This regression yields the value of f of that subset that corresponds to some average value of $\hat{\beta}_i$. Repeating this procedure over all subsets, we obtain the dependence of f on $\hat{\beta}_i$, which is plotted in Figure 4. We show that f increases with the beta. For both European and US markets, we propose the following approximation of the function f with three different

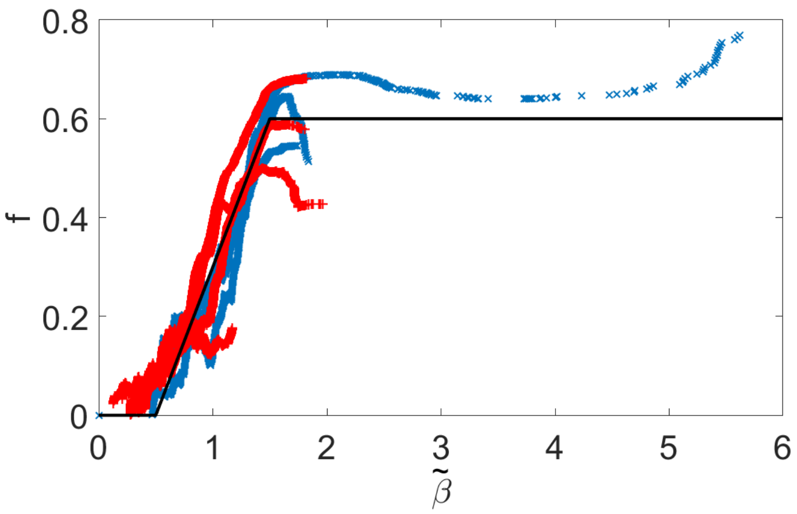

Figure 4: The function f from Eq. (24) versus the beta for the European market (blue crosses) and the US market (red pluses). This function is estimated locally for 4 different time periods. The solid black line shows the approximation (30). Period: 2000-2015.

regimes:

$$f(\tilde{\beta}_i) = \begin{cases} 0, & \tilde{\beta}_i < 0.5, \\ 0.6(\tilde{\beta}_i - 0.5), & 0.5 < \tilde{\beta}_i < 1.6, \\ 0.6 & \tilde{\beta}_i > 1.6. \end{cases} \tag{30}$$

In the first regime, for low beta stocks (mostly quality and value stocks), the beta elasticity is zero, which is equivalent to the constant beta case. For the intermediate regime, the elasticity increases linearly with $\tilde{\beta}_i$ and is close to the constant correlation case with $f(\tilde{\beta}_i) = \tilde{\beta}_i/2$. In the third regime for high beta stocks (mostly speculative and growth stocks), the elasticity is constant. The shape of the beta elasticity is similar for the European market and the US market.

2.3.2 The nonlinear beta elasticity component in the reactive model

According to Eq. (30), the sensitivity of the normalized beta to changes in the relative volatility is nonlinear. This elasticity could generate bias in the beta estimation if the relative volatility changes in a sample used to measure the slope. Our solution is to adjust the beta between normalized returns through the correction factor $\mathcal{F}(t)$, defined as:

$$\mathcal{F}(t) = 1 + \frac{2f(\tilde{\beta}_i(t))}{\tilde{\beta}_i(t)} \Delta\left(\frac{\tilde{\sigma}_i}{\tilde{\sigma}_I}\right). \tag{31}$$

The function f is approximated by Eq. (30), $\ell - \ell'$ is given by Eq. (20), and

$$\Delta\left(\frac{\tilde{\sigma}_i}{\tilde{\sigma}_I}\right) = \frac{\tilde{\sigma}_i(t-1)/\tilde{\sigma}_I(t-1) - \sqrt{\kappa_i(t-1)}}{\sqrt{\kappa_i(t-1)}} \tag{32}$$

with

$$\kappa_i(t) = (1-\lambda_\beta)\kappa_i(t-1) + \lambda_\beta\left(\frac{\tilde{\sigma}_i(t)}{\tilde{\sigma}_I(t)}\right)^2 \tag{33}$$

being the EMA of the squared relative volatility $(\tilde{\sigma}_i/\tilde{\sigma}_I)^2$. The $\Delta(\tilde{\sigma}_i/\tilde{\sigma}_I)$ quantifies deviations of the relative volatility from its average over the sample that will be used to estimate the beta.

The correction factor $\mathcal{F}(t)$ should be used to estimate the slope between stock index and single stock returns and should then be used to denormalize the slope to obtain the reactive beta that depends directly on $\mathcal{F}(t)$.

2.4 Summary of the reactive beta model

In this section, we recapitulate the reactive beta model that combines the three independent components that we described in the previous sections: the specific leverage effect on the beta, the systematic leverage effect on correlation, and the relation between the relative volatility and the beta. Starting with the time series $I(t)$ and $S_i(t)$ for the stock index and individual stocks, one computes the levels $L_f(t)$, $L(t)$, and $L_i(t)$ from Eqs. (2, 4, 5), the normalized stock index and individual stock returns $\tilde{r}_I(t)$ and $\tilde{r}_i(t)$ from Eqs. (6), the normalized stock index volatility $\tilde{\sigma}_I(t)$ from Eq. (7), the renormalized stock index and individual stock returns $\hat{r}_I(t)$ and $\hat{r}_i(t)$ from Eq. (25), the associated volatilities $\hat{\sigma}_I(t)$ and $\hat{\sigma}_i(t)$ from Eqs. (27, 29), and the renormalized beta $\hat{\beta}_i(t)$ from Eq. (28). From these quantities, one re-evaluates the covariance between $\hat{r}_i$ and $\hat{r}_I$ by accounting for the leverage effects and excluding the other effects. In fact, we compute $\hat{\Phi}_i(t)$ as an EMA of the normalized covariance of the normalized daily returns:

$$\hat{\Phi}_i(t) = (1-\lambda_\beta)\hat{\Phi}_i(t-1) + \lambda_\beta \frac{\hat{r}_i(t)\,\hat{r}_I(t)}{\mathcal{L}(t)\,\mathcal{F}(t)}, \tag{34}$$

where $\mathcal{L}(t)$ and $\mathcal{F}(t)$ are two correction factors defined in Eq. (21) and Eq. (31) that are used to withdraw bias from the systematic leverage and the beta elasticity. The parameter λ_β describes the look-back used to estimate the slope and is set to 1/90, as 90 days of look-back appears to us as a good compromise. In fact, for a longer look-back, variations in beta, correlation and volatilities are expected to happen due to changes in market stress and business cycles and are not taken into account properly by our reactive renormalization. In turn, for a shorter look-back, the statistical noise of the slope would be too high.

Finally, the stable estimate of the normalized beta is

$$\tilde{\beta}_i(t) = \frac{\hat{\Phi}_i(t)}{\hat{\sigma}_I^2(t)}, \tag{35}$$

with $\hat{\sigma}_I^2(t)$ defined in Eq. (27) from which the estimated reactive beta of stock i is deduced as

$$\beta_i(t) = \tilde{\beta}_i(t) \left(\frac{L_i(t)\, I(t)}{S_i(t)\, L(t)} \right) \mathcal{L}(t)\, \mathcal{F}(t). \tag{36}$$

The estimation of the normalized stable beta $\tilde{\beta}_i(t)$ is close to the slope estimated by an OLS[4] but with exponentially decaying weights to accentuate recent returns and with normalized returns to withdraw different biases. Then, the normalized stable beta $\tilde{\beta}_i(t)$ is "denormalized" by the factor that combines the three main components: the specific leverage effect on beta, $(L_i/S_i)(I/L)$, the systematic leverage effect, $\mathcal{L}(t)$, and nonlinear beta elasticity, $\mathcal{F}(t)$. The final beta estimation $\beta_i(t)$ is a reactive dynamic conditional estimation that is adjusted as soon as prices moves through the instantaneous variations of the 3 correction factors.

Every term impacts the hedging of a certain strategy:

- the term with $\mathcal{L}(t)$ does not have a significant impact on the beta, as it is compensated in L_i/L, which models the short-term systematic leverage effect on the correlation in Eqs. (34, 36) (introduced in Sec. 2.2), whereas the levels L_i and L were introduced in the reactive volatility model. However, it could impact the correlation by +10% if the market decreases by 10%.

- the term with $L_i I/(L S_i)$ that models the specific leverage effect on volatilities (introduced in Sec. 2.1.2) could impact the beta by 10% if the stocks underperform by 10%. This term impacts the hedging of the short-term reversal strategy.

- the term with $\mathcal{F}(t)$ that models the nonlinear beta elasticity, which is the sensitivity of the beta to the relative volatility (introduced in Sec. 2.3), could impact the beta by 10% if the relative volatility increases by 10%. This term impacts the hedging of the low volatility strategy.

The reduced version of the reactive beta model, when only the leverage effect is introduced without beta elasticity and stochastic normalized volatilities, defines an interesting class of stochastic processes that appears to be a mean reverting with a standard deviation linked to $\tilde{\sigma}_i\sqrt{1/\lambda_s}$ and a relaxation time linked to $1/\lambda_s$.

The reactive beta model is based on the fit of several well-identified effects. Implied parameters work universally for all stock markets ($\ell - \ell'$ is the only one that was fitted only

[4]We assumed that the average of daily returns was zero. That assumption makes sense as at a daily time scale, and the average of returns can be completely neglected compared to the standard deviation.

on the US market, as the implied correlations for other countries are not traded). Here, we summarize the different parameters used in the reactive beta model:

- $\lambda_f = 0.1484$, which describes the relaxation time of 7 days for the panic effect;
- $\lambda_s = 0.0241$, which describes the relaxation time of 40 days for the retarded effect;
- $l = 8$, which describes the leverage intensity of the panic effect;
- $\ell - \ell' = 0.91$ based on implied correlations on the US stock market;
- the different thresholds in the function $f(\tilde{\beta}_i)$ from Eq. (30) that describes the nonlinear beta elasticity.

3 Empirical findings

3.1 Data description

We used only daily returns. For the empirical calibration of $\ell - \ell'$, we chose the CBOE S&P 500 Implied Correlation Index (ICI), which is the first widely disseminated market-based estimate of implied average correlation of the stocks that comprise the S&P 500 Index (SPX). This index begins in July 2009, with historical data going back to 2007. We take the front-month correlation index data from 2007 and roll it to the next contract until the previous one expires. We also use the daily S&P 500 stock index. For the empirical calibration of the other parameters of the reactive beta model, we use the daily S&P 500 stock index and the 600 largest US stocks from January 1, 2000, to May 31, 2015. For the European market, we consider the EuroStoxx50 index and the 600 largest European stocks over the same period. The same data are used for both the calibration of parameters and empirical tests.

For consistency, we kept the parameters of the reactive volatility model that describe the intensity of the panic effect (l), the relaxation time of the panic effect (λ_f) and the relaxation time of the retarded effect(λ_s) identical to those that were calibrated in a period prior to 2000 by Bouchaud *et al.* (2001), since they are seen as being universal.

3.2 Empirical results

In this section, we show that exposure to common risk factors can sometimes lead to a high exposure of market neutral funds to the stock market index if the betas are not correctly assessed. Indeed, although market neutral funds should be orthogonal to traditional asset classes, this is not always the case during extreme moves (Fung and Hsieh, 1997). For instance, Patton (2009) tests the zero correlation against the nonzero correlation and finds that approximately 25% of the market neutral funds exhibit significant non-neutrality, concluding that "many market neutral hedge funds are in fact not market neutral, but overall they are, at least, more market neutral than other categories of hedge funds." The reactive beta model can help hedge funds be more market neutral than others. To demonstrate this, we empirically test the efficiency of hedging using the most popular market neutral strategies (low volatility, short term reversal, momentum and size):

- low volatility (beta) strategy: buying the stocks with the highest 30% beta and shorting those with the lowest 30% beta (estimated by the OLS);
- short term reversal: shorting the stocks with the highest 15% one-month returns and buying those with the lowest 15% one-month returns;
- momentum strategy: buying the stocks with the highest 15% two-year returns and shorting those with the lowest 15% two-year returns;
- size strategy: buying the stocks with the highest 30% capitalization and shorting those with the lowest 30% capitalization.

The construction of the four most popular strategies that target beta neutrality is explained in Appendix B. The different portfolios are dynamic. The efficiency of the hedge depends on the accuracy of the beta estimation. For each strategy, we compare two different methods to estimate the beta that use only the past information to avoid look-ahead bias: ordinary least squares (OLS) (which corresponds to a specific case of our model with $L_i = S_i$, $L = I$, $\ell = \ell' = 0$, and $f = 0$, with the same exponential weighting scheme) and our reactive method. We analyze two statistics:

- Statistic 1: the CorSTD, which is defined as the standard deviation of the 90-day correlation of the strategy with the stock index returns, describes the lack of robustness of the hedge and, consequently, the inefficiency of the beta measurement. The more robust the strategy is, the lower the CorSTD statistics are. If the strategy was well hedged, the correlation would fluctuate around 0, within the theoretical 10% standard deviation, and CorSTD would be 10% (a CorSTD of 10% is obtained with two independent Gaussian variables for 90-day correlations).
- Statistic 2: the Bias, which is defined as the correlation of the strategy with the stock index returns on the whole period, describes the bias in the hedge of the strategy and, therefore, the bias of the beta measurement.

These statistics present a proxy for assessing the quality of the beta measurement, which is very difficult to realize directly, as true betas are not known.

Table 1 summarizes the statistics of the four strategies for the US and Europe markets. We see the highest bias for the low volatility strategy when hedged with the standard approach (-25.5% for USA and -22.4% for Europe). The CorSTD is approximately 20%, i.e., twice as high as expected if the volatility was stable, which means that the efficiency of the hedge is time-varying. This could represent an important risk for the funds of funds managers, where hidden risk could accumulate and arise especially when the market is stressed. Indeed, the bias seems to have been higher by approximately -60% for both the USA and Europe when the market was stressed in 2008. The use of the reactive beta model reduces the bias in the low volatility factor, and the residual bias comes from the selection bias (see Appendix A). When using the OLS, the possible loss in 2008 would have been -9.6% ($= -60\% \times 40\% \times 8\%/20\%$) for a 40% stock decline with a fund invested entirely in a low volatility anomaly with a bias of -60% and a target annualized volatility of 8% for the fund and 20% for the index.

We also see a significant bias for the short-term reversal strategy when hedged with the standard approach (approximately 13.1% in the USA and in Europe). The CorSTD is approximately 18%. The efficiency of the hedge depends on the recent past performance of the strategy. As soon as the strategy starts to lose, the efficiency will decline and risk will arise, as in 2009. Again, we see that the reactive beta model reduces the bias in the short-term reversal factor. The biases and CorSTD are lower for the momentum strategy (-6.3% in the USA, with a CorSTD of 18.3%) and are of the same magnitude for the size strategy (-7.6% in the USA with a CorSTD of 17.0%). The reactive beta model further reduces the bias and the CorSTD. This is also valid for the European market.

We conclude that the reactive beta model reduces the bias of the low volatility factor when it is stressed by the market. The remaining residual is most likely explained by the selection bias (see Appendix A for a formal proof). The improvement is more significant for the momentum factor and the size factor in the U.S. only.

We also illustrate these findings by presenting the correlation between the stock index and the low volatility strategy (Figure 5) and the short-term reversal strategy (Figure 6), which are the strategies with the highest bias. A period surrounding the financial crisis was chosen (2007-2010). One can see that the beta computed by the OLS is highly positively exposed to the stock index in 2008. In turn, the exposure is reduced within the reactive model. The improvement becomes even more impressive in extreme cases when the strategies are stressed by the market. We see that in some extreme cases (a stress period with extreme strategies), the common approach could generate high biases (-50% for the short-term reversal strategies in 2008-2009 and -71% for the beta strategy in 2008). In each case, our methodology allows one to significantly reduce the bias.

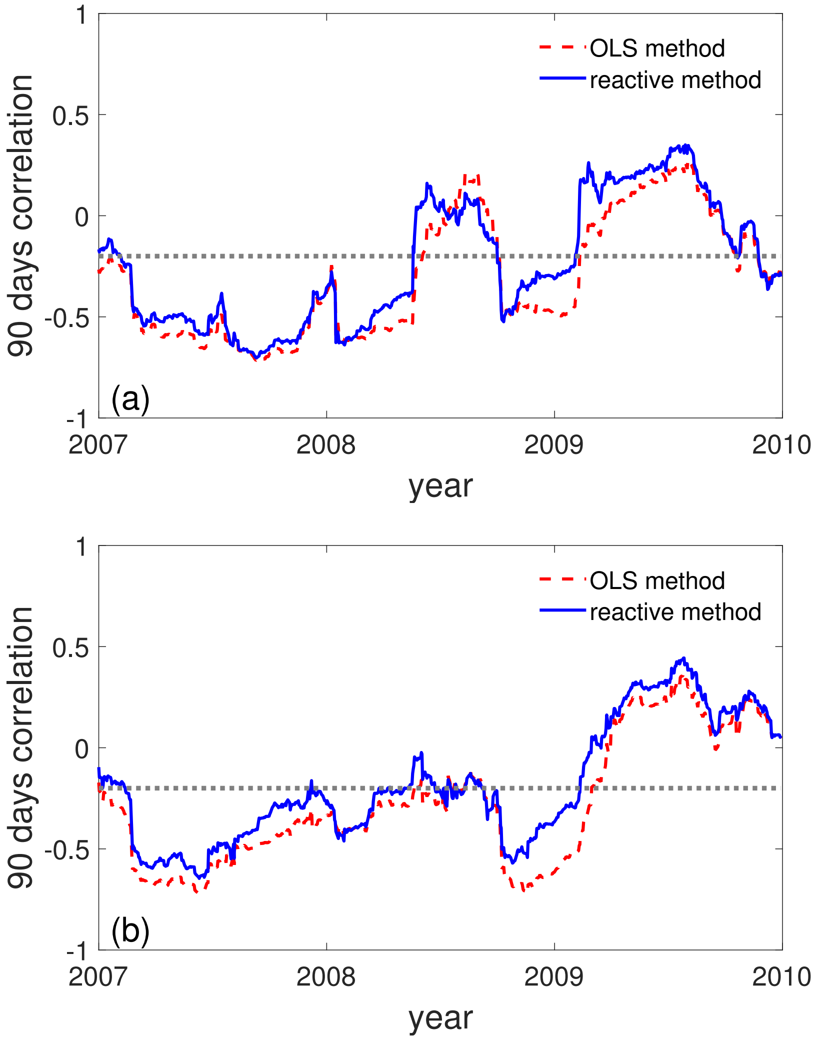

Figure 5: Ninety-day correlation of the low volatility factor with the stock index in the European market **(a)** and in the USA market **(b)**. Solid and dashed lines present the proposed reactive beta model and the OLS methodology, respectively. The dotted horizontal line shows the selection bias of -19.10%, as shown in Appendix A. A time frame surrounding the financial crisis is chosen. Period: 2007-2010.

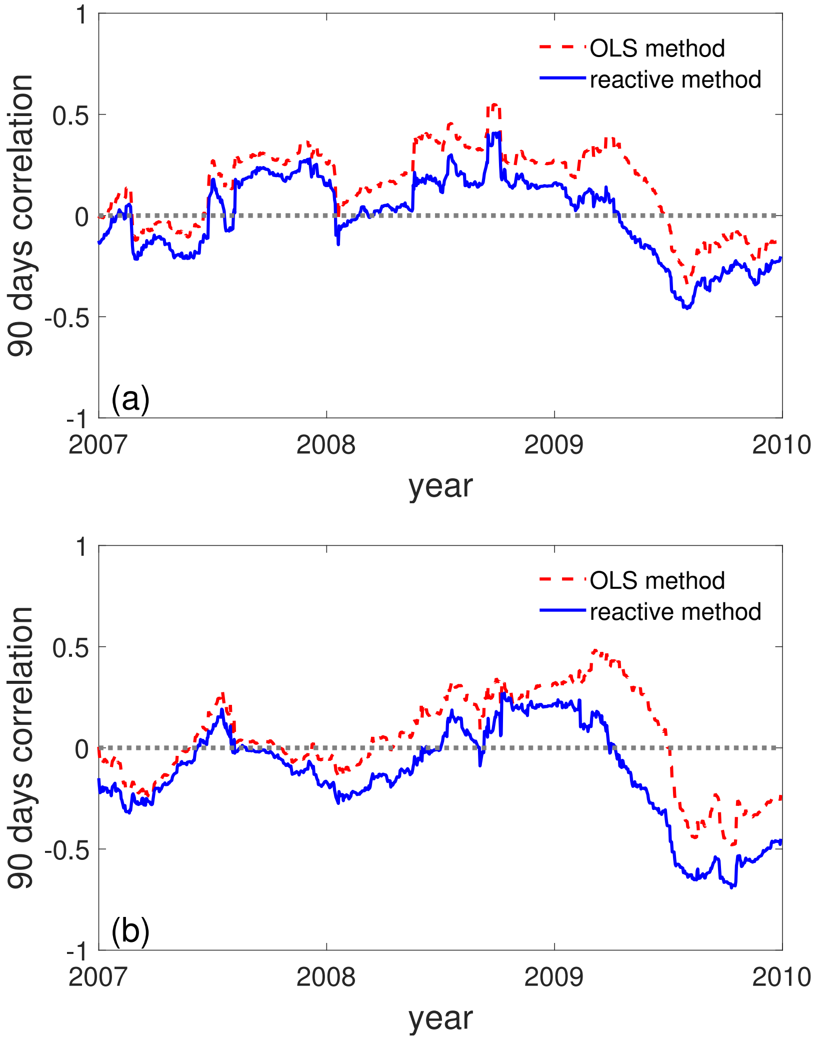

Figure 6: Ninety-day correlation of the short-term reversal factor with the stock index in the European market **(a)** and in the US market **(b)**. Solid and dashed lines present the proposed Reactive beta model and the OLS methodology, respectively. A time frame surrounding the financial crisis is chosen. Period: 2007-2010.

	Strategy \ Method	OLS		Reactive	
	Statistics :	Bias	CorSTD	Bias	CorSTD
US	low volatility	-25.54%	21.73%	-16.79%	21.43%
	short-term reversal	13.09%	18.96%	-6.06%	18.50%
	momentum	-6.27%	18.28%	-2.95%	16.54%
	size	-7.56%	17.00%	-1.84%	17.26%
Europe	low volatility	-22.39%	19.97%	-14.68%	20.94%
	short-term reversal	13.05%	17.51%	0.64%	14.52%
	momentum	-4.42%	18.03%	-1.55%	17.23%
	size	3.12%	17.15%	3.79%	15.63%

Table 1: Bias is defined as the correlation over the whole sample between the stock index and each of the OLS and Reactive strategies for the US and Europe markets. CorSTD is defined as the standard deviation of the 90-day correlation over the whole sample between the stock index and each of the OLS and Reactive strategies for the US and Europe markets. The residual bias for the low volatility strategy in the reactive method can be explained by the selection bias, as demonstrated in Appendix A. Period: 2000-2015.

4 Robustness Checks

This section presents a robustness check analysis by comparing the quality of several methods for beta measurements against the reactive beta model. We build the comparative analysis based on two important articles. Chan and Lakonishok (1992) enables the assessment of the robustness statistics of some alternative methods to the classical ordinary least squares (OLS) method when assuming implicitly that betas are static and returns are homoscedastic. This section extends their work by including alternative dynamics beta estimators to be consistent with our reactive model and with the work by Engle (2016) that demonstrates that the betas are significantly time-varying using dynamic conditional betas. The models and the methods are presented in detail in Appendix C.

4.1 Monte Carlo simulations

In financial research, one often resorts to simulated data to estimate the error of measurements. For instance, Chan and Lakonishok (1992) built their main results on numerical simulation while applying real data for simple comparison between betas estimated with OLS and quantile regression (QR).

The comparative analysis is based on a two-step procedure. The first step simulates returns using different models that capture some market patterns, and the second step estimates the beta from simulated returns by using our reactive method and alternative methods. We tested the same estimators as used by Chan and Lakonishok (1992), including OLS, the minimum absolute deviation (ABSD), and the Trimean quantile regression (TRM). We also added two variations of the dynamical conditional correlation (DCC), which has become a mainstream model to measure the conditional beta when the beta is stochastic (Bollerslev *et al.*, 1988; Bollerslev, 1990; Engle, 2002; Cappiello *et al.*, 2006). We analyze the error of measurements, which we defined as the difference between the measured beta and the true beta of the simulated data.

4.1.1 The first step: simulation

The first step simulates 30,000 paths of T=1,000 consecutive returns for both the stock index (r_I) and the single stock (r_i). It also allows one to generate 1,000 conditional "true" expected betas per path (Fig. 7). To that end, following Chan and Lakonishok (1992), normally distributed residuals and Student-t distributed residuals are considered to take into consideration the robustness of different methods to outliers.

In our setting, we implemented seven Monte Carlo simulations for the returns r_i and r_I. In the simulations, we target the realistic case of an unconditional single stock annualized volatility of 40%, an unconditional stock index volatility of 15% and an unconditional beta of 1. We also target the realistic case of a correlation between the index and the stock of 0.4, since the relative precision of the beta measurement is inversely proportional to the square root of the number of returns when the correlation is close to zero. First, we consider the naive version of the market model, based on Eq. (11), which we call "the basic market model". For the case of a constant beta, as in the paper by Chan and Lakonishok (1992), the simulated data are based on the hypothesis of a null intercept, and the beta is set equal to 1 to characterize the ideal case with a Gaussian (MC1) or a Student-t distribution (MC2) for residuals. In the most simple reactive version of the market model, which we call "the reduced reactive market model" (MC3 and MC4), normalized returns $\tilde{r}_i$ and $\tilde{r}_I$ are first generated randomly through Eq. (12) with a normalized beta set to 1. Then, based on the levels L_s, L_{is}, which are respectively the slow moving averages of the stock index and the stock prices defined in Eq. (1), we generate δI and δS defined in Eq. (6) and then r_i and r_I. Finally, we update L_s and L_{is}. That model is sufficient to capture the leverage effect on beta with increasing beta as soon as a single stock underperforms the stock index. Even if the normalized beta is set to unity, the denormalized beta in Eq. (13) becomes time dependent (Fig. 7). As previously, MC3 and MC4 differ by the distribution of residuals, Gaussian (MC3) versus Student-t (MC4).

The full reactive market model (MC5) includes all the components described in Sec. 2, i.e., the leverage effect and the nonlinear beta elasticity. For the full version, we generated stochastic $\tilde{\sigma}_i$ and $\tilde{\sigma}_I$, which generate $\tilde{r}_i$ and $\tilde{r}_I$ from Eq. (12), using the normalized beta fixed to $\mathcal{F}(t)\mathcal{L}(t)$ (see definitions in Eqs. (31) and (21)). This allows the generation of returns that capture the leverage effect pattern and the empirical nonlinear beta elasticity (Fig. 3

and Fig. 4).

We also used another way to generate random returns that captures a time-varying beta through the implementation of the dynamic conditional correlation (DCC) model (Engle, 2002, 2016), which generalizes the GARCH(1,1) process to two dimensions. This is a mainstream model that has two variations: symmetric and asymmetric, the latter capturing the leverage effect. The symmetric and asymmetric versions of the DCC model are denoted as MC6 and MC7, respectively.

To summarize, the seven Monte Carlo simulations are the following:

- MC 1: The basic market model in Eq. (12), where residuals (ϵ_i) are normally distributed, and the constant beta is set to 1.
- MC 2: The basic market model in Eq. (12), where residuals (ϵ_i) follow a Student-t distribution (with three degrees of freedom), and the constant beta is set to 1.
- MC 3: The reduced reactive market model in Eq. (12), where residuals ($\tilde{\epsilon}_i$) are normally distributed with constant volatilities ($\tilde{\sigma}_i$, $\tilde{\sigma}_I$) and constant renormalized beta ($\tilde{\beta}$) is set to 1, but the denormalized beta is now time-dependent (Fig 7). The conditional beta (β) is a mean reversion process with a relaxation time $1/\lambda_s = 40$ days. MC3 includes only the leverage effect and ignores the nonlinear beta elasticity.
- MC 4: The reduced reactive market model in Eq. (12), where residuals ($\tilde{\epsilon}_i$) follow a Student-t distribution (with three degrees of freedom) with constant relative volatility and a constant renormalized beta set to 1, as in MC3.
- MC 5: The full reactive market model in Eq. (12), where residuals ($\tilde{\epsilon}_i$) follow a Student-t distribution (with three degrees of freedom) whose standard deviation (s_i) is stochastic and where the normalized stock index return ($\tilde{r}_I$) is a Gaussian whose standard deviation (s_I) is also stochastic. We suppose that $\log(s_I)$ and $\log(s_i) - \log(s_I)$ follow two independent Ornstein-Uhlenbeck processes (with a relaxation time of 100 days and a volatility of volatility of 0.04). In that way, the stock index annualized volatility could jump up to 40%. The normalized beta that was set to 1 in MC4 is now set to $\mathcal{F}(t)\mathcal{L}(t)$ to take into account the nonlinear beta elasticity (see definitions in Eqs. (31) and (21)). Both the leverage effect and stochastic normalized volatilities make the volatilities and the beta defined in Eq. (36) time-dependent (Fig. 7).
- MC 6: The symmetric DCC model in two dimensions, which generates volatilities of volatilities and a correlation of similar amplitude as MC5 (Fig. 7).
- MC 7: The ADCC model in two dimensions, which generates volatilities of volatilities and a correlation of similar amplitude as MC5 (Fig. 7).

In Fig. 7, we plot a Monte Carlo path generated for a true beta for MC 3 to 7 (MC1 and MC2 are excluded, as they generate a true beta of 1). We also plot the conditional correlation and volatilities that are highly volatile and thus make the estimation of the conditional beta complicated.

4.1.2 The second step: measurements

The second step is devoted to the analysis of the error measurement of the beta estimations, defined as the difference between the measured beta and the true beta of the simulated data. In our setting, we test 5 alternative beta estimations that should replicate the true beta as closely as possible. Note that in all five configurations, we use an exponentially weighted scheme to give more weight to recent observations, to be in line with the reactive market model ($1/\lambda_\beta = 90$). Consequently, in a path of T=1,000 generated returns, only the 90 last returns truly matter (note that Chan and Lakonishok (1992) is based on the statistics from 35 returns with an equal weight scheme). The first alternative method is the ordinary least squares (OLS) of the returns, which was also implemented in the empirical test based on real data. Note that the OLS[5] would give the same measurement as our reactive method if the parameters were set differently ($\lambda_s = 1$, $\lambda_f = 1$, $l = l' = 0$, $f = 0$). The square errors in the OLS are weighted by $(1 - \lambda_\beta)^{T-t}$. The second method estimates the beta by using the minimum absolute deviation (MAD), which is supposed to be less sensitive to outliers because absolute errors are minimized instead of square errors. The absolute errors are weighted by $(1-\lambda_\beta)^{T-t}$. The third alternative is the beta computed from the Trimean quantile regression (TRM), which is reputed to be more robust to outliers according to Chan and Lakonishok (1992). The absolute errors are also weighted by $(1 - \lambda_\beta)^{T-t}$. The fourth and fifth methods are the conditional beta computed from the DCC model. The DCC method was calibrated using the same exponential $(1 - \lambda_\beta)^{T-t}$ weights introduced in the log-likelihood function to extract the optimal unconditional volatilities and correlations, while other parameters such as the relaxation time and volatilities of volatilities and volatilities of correlations were set to the values that were used for the Monte Carlo simulation.

We summarize the reactive method and the five alternative methods that were implemented to estimate the beta as follows:

- β_{OLS}: beta estimated by the ordinary least squares method;
- β_{MAD}: beta estimated by the minimum absolute deviation method;
- β_{TRM}: beta estimated by the trimean quantile regression;
- β_{DCC}: T^{th} conditional beta estimated by using the DCC model;

[5] We assumed that the average of daily returns was zero. That assumption makes sense because at a daily time scale, the average of returns can be completely neglected compared to the standard deviation.

- β_{ADCC}: T^{th} conditional beta estimated by using the ADCC model;
- β_R: beta estimated by the reactive method in Eq. (36).

4.1.3 The statistics

We analyze for every path the error of measurement, defined as the difference between the measured beta based on different methods applied to T returns and the true beta value at time T.

To assess the quality of different methods, we use three statistics following Chan and Lakonishok (1992). The first statistic is the bias, which gives the average error of measurement. Obtaining the bias is more informative than simply obtaining an estimated average estimation of the beta, because in our case, the true beta is not always 1 but fluctuates around 1 for time-varying models from MC3 to MC7. Because we focus on capturing the leverage effect in the beta measurement, we also define winner (loser) stocks, which are stocks that have outperformed (underperformed) the stock index during the last month. Due to the leverage effect, the OLS method is expected to underestimate the beta for loser stocks and to overestimate the beta for winner stocks. It would be interesting to see how robust the improvement of the reactive beta estimation is. We therefore measure the average error among the loser stocks and among the winner stocks. The loser and winner biases are related to the bias in hedging of the short-term reversal strategy measured on real data, and they can confirm the robustness of the empirical measurements. We also define the low (high) beta stocks, which are the stocks whose conditional true beta is lower (higher) than 1. We measure the average error among low and high beta stocks that are related to the bias in hedging of the low beta strategy measured from real data and can confirm the robustness of the beta measurement when adding the component describing the nonlinear beta elasticity.

The second statistic is the ABSolute Deviation (ABSD) of a measurement. It reflects the average absolute errors such that the positive and negative sign errors cannot be mutually compensated. It is a measurement of the robustness.

The third statistic, which is equivalent to ABSD, is the inverse of the variance of the errors of measurement ($\frac{V_{OLS}}{V_m}$) to characterize the relative robustness of the alternative beta estimation. The alternative beta method (with subscript m) that brings the highest improvement is the one with the highest ratio.

The three statistics that were implemented are summarized as follows:

- Statistic 1: the bias, the winner bias and the loser bias, the low beta bias and the high beta bias;
- Statistic 2: the absolute deviation of measurement (ABSD);
- Statistic 3: the relative variance statistics $\frac{V_{OLS}}{V_m}$.

4.2 Empirical tests

We summarize the statistics in Table 2. We see that all methods are unbiased on average in most Monte Carlo simulations. However, this is misleading, as biases from one group of stocks can be significant and can offset others.

4.2.1 Winner and loser bias

The estimated β_{DCC} and β_{ADD} appear to be biased as soon as fat tails are included (MC2).

The reactive beta is the only one to be unbiased for winner and loser stocks when the leverage effect is introduced in Monte Carlo simulations (MC 3, 4, 5). The biases for winner stocks and loser stocks are significant for all methods except for the reactive beta. The biases are amplified when a fat tail of residual distribution is introduced (MC 4). Winner/loser biases can reach 14%. This is in line with the empirical test implemented on real data, where we see that the reactive method reduces the bias of hedging of the short-term reversal strategy (Tab. 1).

When all components that deviate from the Gaussian market model are mixed in MC5 (fat tails, nonlinear beta elasticity, stochastic volatilities, leverage effect), we see a kind of cocktail effect, as bias is generated for most methods on average and not only in some groups of stocks. The reactive method provides the best results and is the only method that has no bias. β_{MAD} and β_{TRM}, which were supposed to be robust, appear to perform very poorly, with high bias (average, loser or winner) as soon as the stochastic volatility is added, which is confirmed with MC6 and MC7.

We also see that the reactive model looks to be incompatible with the DCC and ADCC models. Indeed, MC5 generates high bias for β_{DCC} and β_{ADD} in the winner and loser stocks even if the leverage effect and the dynamic beta are implemented in the ADCC. In the same way, MC 6 generates bias for the reactive method that is even amplified when leverage effect is generated through MC7. We can wonder which model is the most realistic. Both ADCC and the reactive model capture the volatility clustering and leverage effect patterns, but their dynamics are very different. For example, in the reactive model, volatility increases as soon as the price decreases, and it decreases as soon as the price increases. In contrast, the volatility in ADCC increases only if the return is more negative than the unconditional standard deviation ($\gamma\left(\sigma_i^2[\xi_i^-(t)]^2 - \tilde{\sigma}_i^2\right) > 0$, see Eqs. (67, 69)). The reactive beta model has three components that were tailored to three well-identified effects (the specific leverage through the retarded effect, the systematic leverage through the panic effect and the nonlinear beta elasticity) whose main parameters appear to be stable and universal for all markets. Bouchaud *et al.* (2001) measured most of the parameters for seven main stock indexes. The relaxation time is approximately 1 week for the panic effect ($\lambda_s = 0.1484$), the relaxation time is 40 days for the retarded effect ($\lambda_s = 0.0241$), and the leverage parameter for the panic effect is $l = 8$. The systematic leverage parameter on correlation $\ell - \ell' = 0.91$

was the only one that has been measured through the implied correlation only from the US market. The parameters of the beta elasticity were measured for both the European and the US markets. Whereas the reactive model was tailored to the market, the DCC and ADCC models are more generic constrained models whose parameters were obtained when optimizing the log-likelihood of returns, which in reality are a result of the complex mixture of the three effects, without focusing on each element. DCC and ADCC do not take into account that, in fact, a large part of the heteroscedasticity comes in reality from the complex leverage effect and not from simple autoregressive conditional heteroskedasticity. The parameters that we used describing the relaxation times, the volatilities of volatility, the volatility of correlation and the asymmetries of the DCC and ADCC models were based on the work by Sheppard (2017) and Cappiello *et al.* (2006). Relaxation times of 10 days and 13 days were estimated for the US market and are different from those used in the reactive volatility model. In Table 5 of Cappiello *et al.* (2006), we see that the relaxation time would be 2.5 days for Belgian stocks (the decay factor is $\beta = 0.6184$ for the univariate GARCH), 4 days for French stocks (the decay factor is $\beta = 0.7497$) and 14 days for Spanish stocks (the decay factor is $\beta = 0.9360$). It is not surprising to see this variation if simple autoregressive conditional heteroskedasticity cannot capture the complexity of the leverage effect. We think it would be better to apply autoregressive conditional heteroskedasticity to model the residual part of the heteroskedasticity of returns once the part due to the leverage effect is withdrawn through normalized returns from the reactive model. The relaxation time in that case is expected to be a couple of months.

4.2.2 High and low beta bias

The reactive beta is the only one that reduces the bias for low and high beta stocks when stochastic volatility is introduced and when the empirical nonlinear beta elasticity is implemented (MC 5). This is in line with the empirical test applied to real data, where we see that the reactive method reduces the bias of hedging of the low volatility strategy (Tab. 1).

4.2.3 ABSD and V_{OLS}/V_m

The β_{OLS}, which is the theoretical optimal estimation for Monte Carlo simulated returns with the Gaussian market model (MC1), gives similar statistics to that of the reactive beta for the MC3. In this case (MC3), the reactive method outperforms the other considered methods. The ABSD of 0.17 is entirely explained by irreducible statistical noise that is intrinsic to any regression based on approximately 90 points with a weak correlation.

When a fat tail is incorporated to the residual (MC4), the ABSD of the reactive beta is increased and becomes intermediate between the ABSD of β_{OLS}, β_{MAD} and β_{TRM}. β_{MAD} and β_{TRM} are more robust in the presence of fat tails. The reactive beta is expected to be as sensitive as the OLS would be due to the outliers. The reactive method could be still improved if a TRM regression were implemented instead of the classical OLS to measure the normalized beta between normalized returns. When stochastic volatility and correlation are introduced (MC5, MC6 and MC7), the reactive beta becomes as robust as β_{MAD} and β_{TRM} based on ABSD.

Method	Bias	Winner Bias	Loser Bias	Low Bias	High Bias	ABSD	Vols/Vm
MC1 Gaussian basic market model							
β_{OLS}	-0.00	-0.00	-0.00			0.16	1.00
β Reactive	0.00	-0.05*	0.05*			0.18	0.79
β_{DCC}	0.04*	0.05*	0.03*			0.23	0.51
β_{ADCC}	0.09*	0.01	0.17*			0.25	0.44
β_{MAD}	-0.00	0.00	-0.01			0.20	0.65
β_{TRM}	-0.00	0.00	-0.01			0.20	0.68
MC2 t-Student basic market model							
β_{OLS}	-0.00	0.01	-0.01			0.28	1.00
β Reactive	0.01	-0.06*	0.08*			0.31	0.82
β_{DCC}	0.13*	0.14*	0.12*			0.39	0.67
β_{ADCC}	0.25*	0.15*	0.35*			0.46	0.57
β_{MAD}	-0.00	-0.00	-0.00			0.22	2.18
β_{TRM}	-0.00	-0.00	-0.00			0.22	2.24
MC3 Gaussian reduced reactive market model							
β_{OLS}	-0.00	0.07*	-0.07*	0.07*	-0.07*	0.19	1.00
β Reactive	-0.00	0.02*	-0.02*	0.02*	-0.02*	0.17	1.27
β_{DCC}	0.04*	0.10*	-0.02	0.11*	-0.02	0.24	0.62
β_{ADCC}	0.09*	0.06*	0.12*	0.07*	0.11*	0.24	0.66
β_{MAD}	-0.01	0.06*	-0.08*	0.06*	-0.08*	0.22	0.73
β_{TRM}	-0.01	0.06*	-0.08*	0.06*	-0.08*	0.22	0.75
MC4 t-Student reduced reactive market model							
β_{OLS}	0.01	0.13*	-0.11*	0.12*	-0.10*	0.35	1.00
β Reactive	-0.01	0.02	-0.04*	0.03	-0.05*	0.31	1.30
β_{DCC}	0.12*	0.22*	0.02	0.27*	-0.01	0.47	0.84
β_{ADCC}	0.26*	0.24*	0.28*	0.30*	0.21*	0.52	0.83
β_{MAD}	-0.03*	0.09*	-0.14*	0.10*	-0.14*	0.27	2.68
β_{TRM}	-0.03*	0.09*	-0.14*	0.10*	-0.14*	0.27	2.76
MC5 t-Student full reactive market model							
β_{OLS}	-0.01	0.13*	-0.14*	0.14*	-0.22*	0.50	1.00
β Reactive	-0.04*	-0.00	-0.07*	0.05*	-0.17*	0.41	1.42
β_{DCC}	-0.01	0.10*	-0.12*	0.20*	-0.32*	0.52	1.31
β_{ADCC}	0.10*	0.10*	0.11*	0.29*	-0.17*	0.54	1.32
β_{MAD}	-0.09*	0.04*	-0.22*	0.09*	-0.37*	0.38	2.43
β_{TRM}	-0.09*	0.04*	-0.22*	0.09*	-0.36*	0.37	2.46
MC6 Gaussian symmetric DCC model							
β_{OLS}	-0.11*	-0.10*	-0.11*	0.06*	-0.27*	0.32	1.00
β Reactive	-0.07*	-0.11*	-0.02	0.09*	-0.23*	0.33	0.93
β_{DCC}	-0.01	-0.00	-0.02*	-0.01	-0.01	0.16	4.09
β_{ADCC}	0.02*	-0.08*	0.12*	0.05*	-0.01	0.22	2.06
β_{MAD}	-0.14*	-0.13*	-0.15*	0.04*	-0.32*	0.34	0.89
β_{TRM}	-0.14*	-0.13*	-0.15*	0.04*	-0.32*	0.34	0.90
MC7 Gaussian asymmetric DCC model							
β_{OLS}	-0.09*	0.03	-0.24*	0.09*	-0.25*	0.30	1.00
β Reactive	-0.07*	0.02	-0.17*	0.10*	-0.21*	0.27	1.21
β_{DCC}	-0.04*	0.04*	-0.15*	-0.00	-0.08*	0.21	2.08
β_{ADCC}	-0.01	-0.01	-0.01	-0.00	-0.01	0.15	3.74
β_{MAD}	-0.13*	-0.02	-0.28*	0.06*	-0.29*	0.32	0.92
β_{TRM}	-0.13*	-0.01	-0.28*	0.06*	-0.29*	0.32	0.92

Table 2: Monte Carlo robustness tests. Statistics are provided for seven Monte Carlo simulations and six different methods to estimate the beta. We estimated statistics such as bias, which is the average error of beta measurements. Winner/loser biases are the biases among winner/loser stocks. Low/High biases are the biases among low/high beta stocks. ABSD is the average of the error in absolute value. Vols/Vm is the variance of the error in the OLS case divided by the variance of the error. * indicates a bias greater than 3 standard deviations.

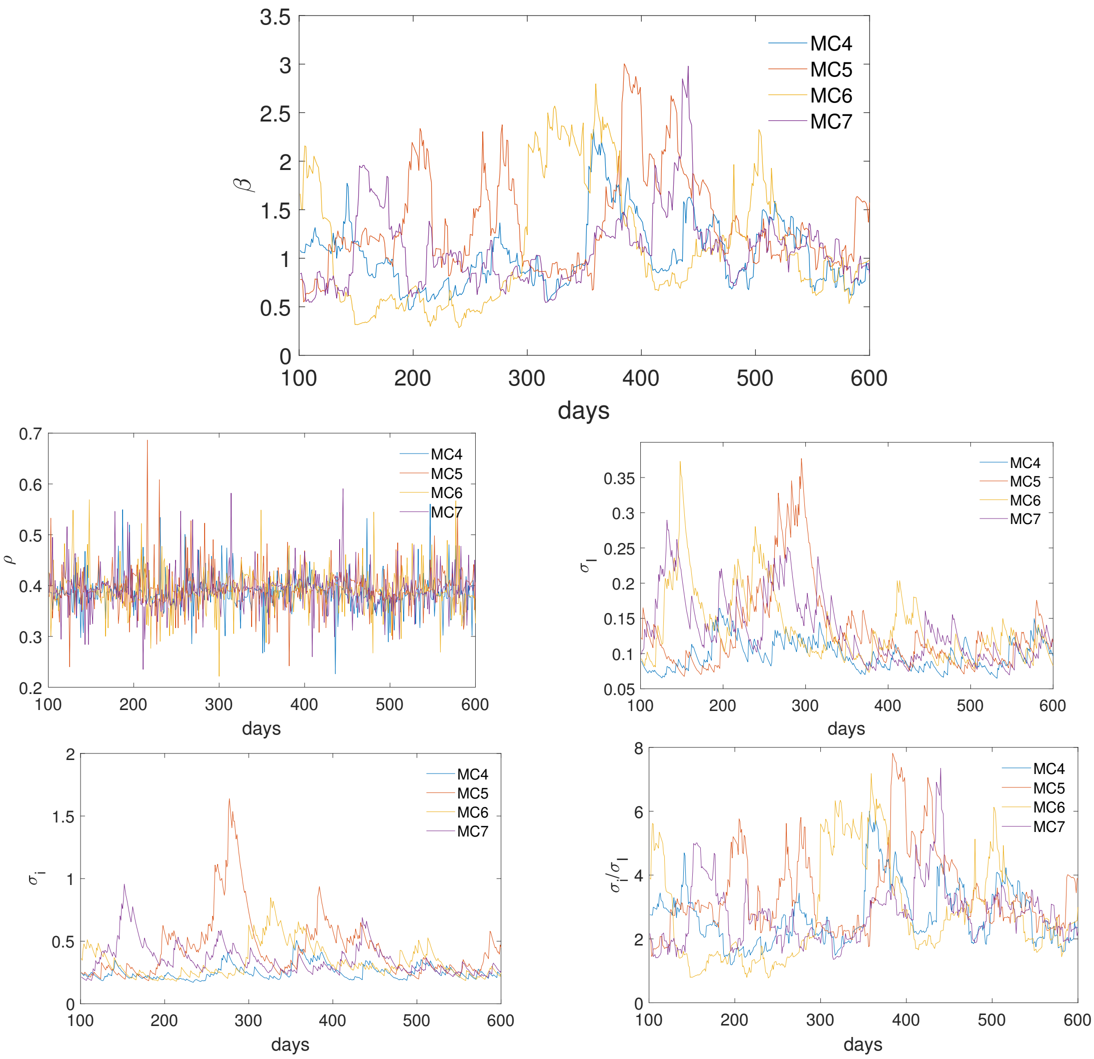

Figure 7: Simulated paths for models MC4 – MC7. The true conditional beta (top), true conditional correlation (middle left), true conditional stock index volatility (middle right), true conditional single stock volatility (bottom left), and true conditional relative volatility (bottom right) are plotted. Paths limited to 500 days, which are independent from model to model, capture the same order of magnitude of variation in volatilities, beta and correlation.

5 Open problems in other fields

The estimated beta is used in a wide range of financial applications, including security valuation, asset pricing, portfolio management and risk management. This extends also to corporate finance in many applications, such as financing decisions to quantify risk associated with debt, equity and assets and firm valuation when discounting cash-flows using the weighted average cost of capital. The most likely reason is that the beta describes systematic risk that could not be diversified and that should be remunerated. However, as explained, the OLS estimator of the beta is subject to measurement errors, which include the presence of outliers, time dependence, the leverage effect, and the departure from normality.

5.1 Asset Pricing

Bali *et al.* (2017) apply the DCC model by Engle (2016) to assess the cross-sectional variation in expected stock returns. They estimate the conditional beta for the S&P 500 using daily data from 1963 to 2009. They test whether the betas have predictive power for the cross-section of individual stock returns over the next one to five days. They show that there is no link between the unconditional beta and the cross-section of expected returns. Most remarkably, they also show that the time-varying conditional beta is priced in the cross-section of daily returns. At the portfolio level, they indicate that a long-short trading strategy of buying the highest conditional beta stocks and selling the lowest conditional beta stocks yields average returns of 8% per year. Thus, conditional CAPM is empirically valid, whereas unconditional CAPM is not empirically valid. Moreover, they show that improvements in beta measurement from unconditional to conditional betas would not have significant pricing impacts on major anomalies (size, book, momentum...). Thus, one can see that DCC greatly changes the pricing of the low volatility anomaly that disappears and improves the empirical validation of the CAPM but does not change the pricing of other major anomalies. We expect that the reactive method can bring further improvements. Indeed, as revealed by our robustness tests in Sec. 4, the leverage effect and the nonlinear beta elasticity are likely to generate bias in the DCC estimation. Because our reactive method was designed to correct for these biases, its use can help reveal pricing effects of the dynamic beta on major anomalies. This point is an interesting perspective for future research.

5.2 Corporate Finance

To determine a fair discount rate for valuing cash-flows, the firm's manager must select the appropriate beta of the project given that the discount rate remains constant over time, while the project may exhibit significant variation over time and the leverage effect due to the debt-to-equity ratio. As such, Ang and Liu (2004) discuss how to discount cash-flows with time-varying expected returns in a traditional set-up. For instance, the traditional dividend discount model assumes that the expected return along with the expected rate of cash-flow growth are set as constant while they are time-varying and correlated. In practice, in the first step, the manager computes the expected future cash-flows from financial forecasts. In the second step, the manager uses a constant discount rate, usually relying on the CAPM for the discounting factor. In contrast, Ang and Liu (2004) derive a valuation formula that incorporates the correlation among stochastic cash-flows, betas and risk premia. They show that the greater the magnitude of the difference between the true discount rate and the constant discount rate, the greater the project's misvaluation. They even show that when computing perpetuity values from the discounting model, the potential mispricing can even become worse. They conclude that accounting for time-varying expected returns can lead to different prices from using a constant discount rate from the traditional unconditional CAPM. The impact of the leverage effect and of the nonlinear elasticity of the beta on potential mispricing deserves to be investigated. Indeed, our results seem to indicate that the mispricing might be higher for low and high beta stocks over a long period. This could be an interesting topic for future work.

5.3 Alternative asset classes

Notice that in this paper, the reactive beta model is tailored for stocks. However, it could help to withdraw some bias in a context involving assets other than stocks such like hedge funds or mutual funds. Indeed the simple market neutral strategies (Short term reversal, momentum, size) can be extended to simple directional strategies (contrarian, trend following) to model the behaviors of funds managers. Some identified bias in beta measurement described in Sec 3.2 captured by the market neutral strategies are also most likely to occur in the directional ones.

An application of the reactive beta model on hedge funds would raise interesting concerns about a better estimation of non-linearity features that would stem from option-like strategies or higher moments as documented by the literature: Fung and Hsieh (2001) warn that hedge funds employ dynamic trading strategies that have option-like returns even if the manager does not trade in derivatives markets. This means that asset pricing models of investment styles are not designed to capture non-linear returns that commonly characterized hedge fund industry. Agarwal and Naik (2004) observe that hedge funds report large losses during crisis episodes, which suggests that they may be bearing significant left-tail risk particularly during large market downturn. They find that the non-linear option like pay-offs from a wide range of equity-oriented hedge funds resembles to a strategy of writing a put option on the equity index. Recall that hedge funds generally employ long-short dynamic strategies to capture non-standard risk premia, in constrast to mutual funds that employ overall long position on buy-and-hold strategies to capture standard risk premia like equity/bond risk premia. Agarwal, Arisoy and Naik (2004) build on an augmented version of the Fung and Hsieh (2004) a seven-factor model to find that hedge funds with greater leverage, longer time in existence and larger assets under management have more negative uncertainty betas. This echoes the findings of Bali, Brown and Tang (2017) for stocks that provide evidence of significant non-linearity in uncertainty premium. Agarwal, Green, and and Ren (2018) measure risk-adjusted hedge fund performance using a range of single and multi-factor models to find that, surprisingly, hedge fund flows are being better explained by CAPM alpha than by more sophisticated models. This echoes the findings of Berk and van Binsbergen (2016) who first use capital flows of mutual funds for asset pricing models to finally conclude that the CAPM better explains risk than no model at all.

An application on mutual funds would also raise interesting concerns about the estimation errors in the individual beta estimate because beta is exposed to estimation errors for individual stocks (see e.g. Chordia, Goyal and Shanken (2015)). But since mutual funds are themselves diversified portfolios, it should alleviate the estimation error in the beta estimate. This is important because it addresses the controversy in the literature as to whether some expected return variations associated with factor loadings (betas) are due to economic risk, or are due to mispricing effects linked to this measurement error. At the same time, using portfolios could also hide some precious information that exists at the individual stock level as documented by the literature (Black *et al.*, 1972; Fama and MacBeth, 1973). Such an investigation on alternative asset classes is left open towards a future research.

6 Conclusion

We propose a reactive beta model with three components that account for the specific leverage effect (when a stock underperforms, its beta increases), the systematic leverage effect (when a stock index declines, correlations increase), and beta elasticity (when relative volatility increases, the beta increases). The three components were fitted and incorporated through elaborate statistical measurements. An empirical test is run from 2000 to 2015 with exhaustive data sets including both American and European securities. We compute the bias for hedging the most popular market neutral strategies (low volatility, short-term reversal, momentum and capitalization) using the standard approach of the beta measurement and the reactive beta model. Our main findings emphasize the ability of the reactive beta model to significantly reduce the biases of these strategies, particularly during stress periods. We further extend the research design to include robustness checks based on simulated data to compare the reactive method with five alternative methods (ordinary least squares, minimum absolute deviation, trimean quantile regression, and dynamic conditional correlation with or without asymmetry) over seven Monte Carlo scenarios reflecting different market conditions from calm (Gaussian residuals, no leverage effect, constant beta) to stress (non-Gaussian residuals, leverage effect, nonlinear beta elasticity, stochastic volatility, nonconstant volatility of volatility and volatility of correlation). We find that the overall results confirm that the reactive beta presents the lower bias when stressed market conditions are included. Further, the reactive model can be useful in other empirical applications such as asset pricing and corporate finance and alternative asset classes such as hedge funds and mutual funds. This provides a good starting point for future research.

References

Acharyaa, V., and L. H., Pedersen. "Asset Pricing with Liquidity Risk." Journal of Financial Economics, 77 (2005), pp. 375-410.

Agarwal, V., and N. Y. Naik. "Risks and Portfolio Decisions Involving Hedge Funds." Review of Financial Studies, 17 (2004), pp. 63-98.

Agarwal, V., Y.E., Arisoy, and N.Y., Naik. "Volatility of Aggregate Volatility and the Cross-Section of Hedge Fund Returns." Journal of Financial Economics, 125 (2017), pp. 491-510.

Agarwal, V., T.C., Green, and H., Ren. "Alpha or Beta in the Eye of the Beholder: What Drives Hedge Fund Flows?" Journal of Financial Economics, 127 (2018), pp. 417-434.

Amihud, Y. "Illiquidity and Stock Returns: Cross-section and Time-series Effects." Journal of Financial Markets, 5 (2002), pp. 31-56.

Ang, A., and G. Bekaert. "International Asset Allocation with Time-Varying Correlations" Review of Financial Studies, 15 (2000), pp. 1137-1187.

Ang, A., and J., Chen. "Asymmetric Correlations of Equity Portfolios." Journal of Financial Economics, 63 (2002), pp. 443-494.

Ang, A., and J. Liu. "How to Discount Cashflows with Time-Varying Expected Returns." Journal of Finance, 59, 6 (2004), pp. 2745-2783.

Ang, A., and J. Chen. "CAPM over the Long Run: 1926-2001." Journal of Empirical Finance, 14 (2007), pp. 1-40.

Ang, A., R. Hodrick, Y. Xing, and X. Zhang. "The Cross-Section of Volatility and Expected Returns." Journal of Finance, 61 (2006), pp. 259-299.

Ang, A., R., Hodrick, Y. Xing, and X. Zhang. "High Idiosyncratic Volatility and Low Returns: International and Further U.S. Evidence." Journal of Financial Economics, 91 (2009), pp. 1-23.

Ang, A., A. Shtauber, and P. C. Tetlock. "Asset Pricing in the Dark: The Cross-Section of OTC Stocks." Review of Financial Studies, 26 (2013), pp. 2985-3028.

Asness, C., R. Krail, and J. Liew. "Do Hedge Funds Hedge?" Journal of Portfolio Management, 28 (2001), pp. 6-19.

Baker, M., B. Bradley, and R. Taliaferro. "The Low Beta Anomaly: A Decomposition into Micro and Macro Effects." Working paper, Harvard Business School, 2013.

Bali, T. G. and R. F. Engle, and Y. Tang. "Dynamic Conditional Beta Is Alive and Well in the Cross Section of Daily Stock Returns." Management Science, 63, 11 (2017), pp. 3760-3779.

Bali, T.G., S.J., Brown, and Y. Tang. "Is Economic Uncertainty Priced in the Cross-Section of Stock Returns?" Journal of Financial Economics, 126 (2017), pp. 471-489.

Banz, R. W. “The Relationship Between Return and Market Value of Common Stocks.” Journal of Financial Economics, 9 (1981), pp. 3-18.

Bekaert, G., and G. Wu. “Asymmetric Volatility and Risk in Equity Markets.” Review of Financial Studies, 13 (2000), pp. 1-42.

Berk, J.B., and J.H., van Binsbergen. “Assessing Asset Pricing Models using Revealed Preference.” Journal of Financial Economics, 119 (2016), pp. 1-23.

Black, F. “Capital Market Equilibrium with Restricted Borrowing.” Journal of Business, 4 (1972), pp. 444-455.

Black, F., M. Jensen, and M. Scholes. “The Capital Asset Pricing Model: Some Empirical Tests.” In Studies in the Theory of Capital Markets, edited by Michael Jensen. New York: Praeger, 1972.

Black, F. “Studies in Stock Price Volatility Changes.” American Statistical Association, Proceedings of the Business and Economic Statistics Section, 177-181, 1976.

Blume, M. E. “On the Assessment of Risk.” Journal of Finance, 26 (1971), pp. 1-10.

Boguth, O., Carlson, M., Fisher, A., and M., Simutin. “Conditional Risk and Performance Evaluation: Volatility Timing, Overconditioning, and New Estimates of Momentum Alphas.” Journal of Financial Economics, 102 (2011), pp. 363-389.

Bollerslev, T., R. Engle, and J. Wooldridge. “A Capital Asset Pricing Model with Time-Varying Covariances.” Journal of Political Economy, 96 (1988), pp. 116-131.

Bollerslev, T. “Modelling the coherence in short-run nominal exchange rates: A multivariate generalized ARCH model.” Review of Economics and Statistics, 72 (1990), pp. 498-505.

Bouchaud, J.-P., A. Matacz, and M. Potters. “Leverage Effect in Financial Markets: The Retarded Volatility Model.” Physical Review Letters, 87 (2001), pp. 1-4.

Bussière, M., M. Hoerova, and B. Klaus. “Commonality in Hedge Fund Returns: Driving Factors and Implications.” Journal of Banking & Finance, 54 (2015), pp. 266-280.

Campbell, J. Y., and L. Hentchel, “No News is Good News: An Asymmetric Model of Changing Volatility in Stock Returns.” Journal of Financial Economics, 31 (1992), pp. 281-318.

Cappiello, L., R. Engle, and K. Sheppard. “Asymmetric Dynamics in the Correlations of Global Equity and Bond Returns.” Journal of Financial Econometrics, 4 (2006), pp. 537-572.

Carhart, M. “On Persistence in Mutual Fund Performance.” Journal of Finance, 52 (1997), pp. 57-82.

Chan, L., and J. Lakonishok. “Robust Measurement of Beta Risk.” Journal of Financial and Quantitative Analysis, 27 (1992), pp. 265-282.

Chen, A-S., T.-W. Zhang, and H. Wu. “The Beta Regime Change Risk Premium.” Working paper, National Chung Cheng University, 2005.

Choi, J. “What Drives the Value Premium? The Role of Asset Risk and Leverage.” Review of Financial Studies, 26 (2013), pp. 2845-2875.

Chordia, T., Goyal, A., and J., Shanken. “Cross-Sectional Asset Pricing with Individual Stocks: Betas versus Characteristics.” Working paper, Emory University, (2015).

Christie, A. “The Stochastic Behavior of Common Stock Variances - Value, Leverage, and Interest Rate Effects.” Journal of Financial Economic Theory, 10 (1982), pp. 407-432.

Daniel, K., R., Jagannathan, and S., Kim. “Tail Risk in Momentum Strategy Returns.” Working paper, National Bureau of Economic Research, (2012).

DeBondt, W.M., and R.H., Thaler. “Further Evidence on Investor Overreaction and Stock Market Seasonality.” Journal of Finance, 42 (1987), pp. 557-581.

DeJong, D., and D. W. Collins. “Explanations for the Instability of Equity Beta: Risk-Free Rate Changes and Leverage Effects.” Journal of Financial and Quantitative Analysis, 20 (1985), 73-94.

Fabozzi, J. F., and C. J. Francis. “Beta Random Coefficient.” Journal of Financial and Quantitative Analysis, 13 (1978), pp. 101-116.

Engle, R. F. ”Dynamic Conditional Correlation: A Simple Class of Multivariate Generalized Autoregressive Conditional Heteroskedasticity Models.” Journal of Business and Economic Statistics, 20 (2002), pp. 339-350.

Engle, R. F. “Dynamic Conditional Beta.” Journal of Financial Econometrics, 14, Issue 4 (2016), pp. 643-667.

Fama, E.F., and J.D., MacBeth. “Risk, Return and Equilibrium: Empirical Tests.” Journal of Political Economy, 81 (1973), pp. 607-636.

Fama, E. F., and K. R. French. “The cross-section of expected returns.” Journal of Finance, 47 (1992), pp. 427-465.

Fama, E. F., and K. R. French. “Common risk factors in the returns on stocks and bonds.” Journal of Financial Economics, 33 (1993), pp. 3-56.

Fama, E. F., and K. R. French. “Industry Costs of Equity.”, Journal of Financial Economics, 43 (1997), pp. 153-193.

Fama, E. F., and K. R. French. “Dissecting Anomalies.” Journal of Finance, 63 (2008), pp. 1653-1678.

Fama, E. F., and K. R. French. “Size, Value, and Momentum in International Stock Returns.” Journal of Financial Economics, 105 (2012), pp. 457-472.

Ferson, W. E., and C. R. Harvey. “Conditioning Variables and the Cross-Section of Stock Returns.” Journal of Finance, 54 (1999), pp. 1325-1360.

Forbes, K., and R., Rigobon. “No Contagion, only Interdependence: Measuring Stock Market Comovements.” Journal of Finance, 57 (2002), pp. 2223-2261.

Francis, J. C. “Statistical Analysis of Risk Surrogates for NYSE Stocks.” Journal of Financial and Quantitative Analysis, 14 (1979), pp. 981-997.

Frazzini, A., and L. Pedersen. “Betting Against Beta.” Journal of Financial Economics, 111 (2014), pp. 1-25.

Fung, W., and D. A. Hsieh. “Empirical Characteristics of Dynamic Trading Strategies: The Case of Hedge Funds.” The Review of Financial Studies, 10 (1997), pp. 275-302.

Fung, W., and D.A., Hsieh. “The Risk in Hedge Fund Strategies: Theory and Evidence from Trend Followers.” Review of Financial Studies, 14 (2001), pp. 313-341.

Fung, W., and D.A., Hsieh. “Hedge Fund Benchmarks: A Risk-Based Approach.” Financial Analyst Journal, 60, (2004), pp. 65-81.

Galai, D. and R. W. Masulis. “The Option Pricing Model and the Risk Factor of Stock.” Journal of Financial Economics, 3 (1976), 53-81.

Garlappi, L., and H., Yan. “Financial Distress and the Cross-Section of Equity Returns.” Journal of Finance, 66 (2011), pp. 789-822.

Glosten, L. R., R. Jagannathan, and D. E. Runkle. “On The Relation between The Expected Value and The Volatility of Nominal Excess Return on stocks.” Journal of Finance, 48 (1993), pp. 1779-1801.

Goyal, A., and P. Santa-Clara, “Idiosyncratic Risk Matters!” Journal of Finance, 58 (2003), pp. 975-1007.

Grinblatt, M., and T., Moskowitz. “Predicting Stock Price Movements from Past Returns: The Role of Consistency and Tax-Loss Selling.” Journal of Financial Economics, 71 (2004), pp. 541-579.

Haugen, R. A., and A. Heins. “Risk and the Rate of Return on Financial Assets: Some Old Wine in New Bottles.” Journal of Financial and Quantitative Analysis, 10 (1975), pp. 775-784.

Haugen, R. A., and N. L. Baker. “The Efficient Market Inefficiency of Capitalization-Weighted Stock Portfolios.” Journal of Portfolio Management, 17 (1991), pp. 35-40.

Hong, Y., Tu, J., and G., Zhou. “Asymmetries in Stock Returns: Statistical Tests and Economic Evaluation.” Review of Financial Studies, 20 (2006), pp. 1547-1581.

Hong, H., and D. Sraer. “Speculative Betas.” Forthcoming, Journal of Finance, (2016).

Jagannathan, R., and Z. Wang, “The Conditional CAPM and the Cross-Section of Expected Returns.” Journal of Finance, 51 (1996), pp. 3-53.

Jegadeesh, N. “Evidence of Predictable Behavior of Securities Returns.” Journal of Finance, 45 (1990), pp. 881-898.

Jegadeesh, N., and S. Titman. “Returns to buying winners and selling losers: Implications for stock market efficiency.” Journal of Finance, 48 (1993), pp. 65-91.

Jiang, L., K., Wu, and G., Zhou. “Asymmetry in Stock Comovements: An Entropy Approach.” Journal of Financial and Quantitative Analysis, 53 (2018), pp. 1479-1507.

Koenker, R., and G. Bassett. “Regression Quantiles.” Econometrica, 46 (1978), pp. 33-50.

Lettau, M., and S. Ludvigson. “Resurrecting the (C)CAPM: a Cross-Sectional Test when Risk Premia are Time-Varying.” Journal of Political Economy, 109 (2001), pp. 1238-1287.

Lewellen, J., and S. Nagel. “The Conditional CAPM does not Explain Asset Pricing Anomalies.” Journal of Financial Economics, 82 (2006), pp. 289-314.

Liu, J., R.F., Stambaugh, and Y., Yuan. “Absolving Beta of Volatility’s Effects.” Journal of Financial Economics, 128 (2018), pp. 1-15.

Longin, F., and B., Solnik. “Extreme Correlation of International Equity Markets.” Journal of Finance, 56 (2001), pp. 649-676.

Malkiel, B. G., and Y. Xu. “Risk and Return Revisited.” Journal of Portfolio Management, 23 (1997), pp. 9-14.

Meng, J. G., G. Hu, and J. A. Bai. “A Simple Method for Estimating Betas When Factors Are Measured with Error.” Journal of Financial Research, 34 (2011), pp. 27-60.

Mitchell, M., and T. Pulvino. “Characteristics of Risk and Return in Risk Arbitrage.” Journal of Finance, 56 (2001), pp. 2135-2175.

Moreira, A., and T., Muir. “Volatility-Managed Portfolios.” Journal of Finance, 72 (2017), pp. 1611-1644.

Patton, A. J. “Are ‘Market Neutral’ Hedge Funds Really Market Neutral?” Review of Financial Studies, 22 (2009), pp. 2295-2330.

Reinganum, R. “Misspecification of Capital Asset Pricing: Empirical Anomalies based on Earnings Yields and Market Values.” Journal of Financial Economics, 9 (1981), pp. 19-46.

Roll, R. “A Critique of the Asset Pricing Theory’s Tests Part I: On Past and Potential Testability of the Theory.” Journal of Financial Economics, 4 (1977), pp. 129-176.

Sharpe, W. F. “Capital Asset Prices: A Theory of Market Equilibrium under Risk.” Journal of Finance, 19 (1964), pp. 425-442.

Shanken, J. “On the Estimation of Beta-Pricing Models.” Review of Financial Studies, 5 (1992), 1-33.

Sheppard, K. “Univariate volatility modeling.” Notes, Chapter 7, University of Oxford (2017).

Valeyre, S., D. S. Grebenkov, S. Aboura, and Q. Liu. “The Reactive Volatility Model.” Quantitative Finance, 13, (2013), pp. 1697-1706.

A Selection bias

Here, we provide some evidence that the bias in beta of the low volatility factor comes from the selection bias: selection of the bottom beta stocks yields the stocks whose beta is underestimated.

The measured beta β_{im} of stock i is obtained by a standard linear regression of the i-th stock returns, r_i, to the stock index returns, r_I,

$$r_i = \beta_{im} r_I + \epsilon_i, \tag{37}$$

where ϵ_i is the residual return. We suppose that the measured beta of the stock i, β_{im}, is affected by noise,

$$\beta_{im} = \beta_{iT} + \eta_i, \tag{38}$$

where β_{iT} is the true beta (which is unknown), and η_i is the error of the measurement inherent to the linear regression. The standard deviation of η_i, σ_{η_i}, depends on the average correlation between the single stock i and the stock index and on the number n of independent points used for the regression (which we set at $n = \frac{1}{\lambda_\beta} = 90$):

$$\sigma_{\eta_i} = \frac{\sigma_{\epsilon i}}{\sigma_I} \frac{1}{\sqrt{n}}, \tag{39}$$

where $\sigma_{\epsilon i}$ is the standard deviation of the residual returns ϵ_i. Averaging the above relation over all stocks, we obtain

$$\sigma_\eta = \frac{\langle \sigma_{\epsilon i} \rangle}{\sigma_I} \sqrt{\lambda_\beta}, \tag{40}$$

where $\langle \sigma_{\epsilon i} \rangle$ denotes the average. According to Eq. (37), the standard deviation of the stock returns, σ_i, is

$$\sigma_i = \sqrt{\beta_{im}^2 \sigma_I^2 + \sigma_{\epsilon i}^2} \approx \sigma_{\epsilon i}, \tag{41}$$

because $(\beta_{im}\sigma_I/\sigma_i)^2 \ll 1$ (stocks are much more volatile than the index). We thus obtain

$$\sigma_\eta \approx \frac{\langle \sigma_i \rangle}{\sigma_I} \sqrt{\lambda_\beta}. \tag{42}$$

The low volatility factor is 50% long of the 30% top β_{im} stocks and 50% short of the 30% bottom β_{im} stocks (here, we consider only one sector for simplicity). We adjust the most volatile leg to target a beta neutral factor if we suppose that η_i are null. In reality, when taking into account the difference between the measured and the true beta, we obtain the beta of the low volatility factor as:

$$\beta_{\text{low factor}} = -50\% \langle \beta_{iT} | i \in \text{Bottom} \rangle + 50\% \frac{\langle \beta_{im} | i \in \text{Bottom} \rangle}{\langle \beta_{im} | i \in \text{Top} \rangle} \langle \beta_{iT} | i \in \text{Top} \rangle. \tag{43}$$

This is essentially the beta neutral condition that we impose when constructing the factor (see Appendix B). Here, $\langle \beta_{im} | i \in \text{Bottom} \rangle$ is the average of the measured beta over the stocks i in the 30% bottom in the *measured* beta values β_{im} (similar for other averages).

Defining $\Delta\beta_B$ and $\Delta\beta_T$ as

$$\langle \beta_{iT} | i \in \text{Bottom} \rangle = \langle \beta_{im} | i \in \text{Bottom} \rangle + \Delta\beta_B, \tag{44}$$

$$\langle \beta_{iT} | i \in \text{Top} \rangle = \langle \beta_{im} | i \in \text{Top} \rangle + \Delta\beta_T, \tag{45}$$

we rewrite Eq. (43) as

$$\begin{aligned}\beta_{\text{low factor}} &= -50\% \left(\langle\beta_{im}|i \in \text{Bottom}\rangle + \Delta\beta_B\right) + 50\% \frac{\langle\beta_{im}|i \in \text{Bottom}\rangle}{\langle\beta_{im}|i \in \text{Top}\rangle} \left(\langle\beta_{im}|i \in \text{Top}\rangle + \Delta\beta_T\right) \\ &= -50\%\Delta\beta_B + 50\% \frac{\langle\beta_{im}|i \in \text{Bottom}\rangle}{\langle\beta_{im}|i \in \text{Top}\rangle} \Delta\beta_T. \end{aligned} \tag{46}$$

Given that $\langle\beta_{im}|i \in \text{Bottom}\rangle \ll \langle\beta_{im}|i \in \text{Top}\rangle$ (as the β_{im} in the top quantile are higher than the β_{im} in the bottom quantile), we obtain the following approximation

$$\beta_{\text{low factor}} \approx -50\%\Delta\beta_B. \tag{47}$$

If one knew the true β_{iT} values and used them for constructing the low volatility factor, the excess $\Delta\beta_B$ would be zero. However, the true values are unknown, and one uses the measured beta β_{im} that creates a selection bias and the nonzero $\Delta\beta_B$, as shown below.

To estimate $\Delta\beta_B$, we consider the true beta β_{iT} and the measurement error η_i as independent random variables and replace the average over stocks by the following conditional expectation

$$\Delta\beta_B = \langle\beta_{iT} - \beta_{im}|i \in \text{Bottom}\rangle \approx \mathbb{E}\{\beta_{iT} - \beta_{im}|i \in \text{Bottom}\} = B. \tag{48}$$

We have, then,

$$\begin{aligned} -B &= \mathbb{E}\{\eta_i|i \in \text{Bottom}\} = \int_{-\infty}^{\infty} \eta\, \mathbb{P}\{\eta_i \in (\eta, \eta + d\eta)|i \in \text{Bottom}\} \\ &= \int_{-\infty}^{\infty} \eta \frac{\mathbb{P}\{\eta_i \in (\eta, \eta + d\eta),\ i \in \text{Bottom}\}}{\mathbb{P}\{i \in \text{Bottom}\}}, \end{aligned} \tag{49}$$

where we wrote explicitly the conditional probability. The denominator is precisely the threshold determining the bottom quantile, $\mathbb{P}\{i \in \text{Bottom}\} = p$, which we set to 30%. We thus obtain

$$-B = \frac{1}{p} \int_{-\infty}^{\infty} \eta\, \mathbb{P}\{\eta_i \in (\eta, \eta + d\eta),\ \beta_{im} - \beta_0 < Q\}, \tag{50}$$

where the event $i \in \text{Bottom}$ is equivalently written as $\beta_{im} < \beta_0 + Q$, where Q is the value of the measured beta that corresponds to the quantile p, and β_0 is the mean of β_{im}. Using Eq. (38) and the assumption that β_{iT} and η_i are independent, one obtains

$$\begin{aligned} -B &= \frac{1}{p} \int_{-\infty}^{\infty} \eta\, \mathbb{P}\{\eta_i \in (\eta, \eta + d\eta),\ \beta_{iT} - \beta_0 < Q - \eta\} \\ &= \frac{1}{p} \int_{-\infty}^{\infty} \eta\, \mathbb{P}\{\eta_i \in (\eta, \eta + d\eta)\}\, \mathbb{P}\{\beta_{iT} - \beta_0 < Q - \eta\}. \end{aligned} \tag{51}$$

To obtain some quantitative estimates, we make a strong assumption that both β_{iT} and η_i are Gaussian variables, with means β_0 and 0 and standard deviations σ_β and σ_η, respectively. We then obtain

$$-B = \frac{1}{p} \int_{-\infty}^{\infty} d\eta\, \eta\, \frac{\exp(-\eta^2/(2\sigma_\eta^2))}{\sqrt{2\pi}\,\sigma_\eta} \Phi\big((Q - \eta)/\sigma_\beta\big), \tag{52}$$

where

$$\Phi(x) = \int_{-\infty}^{x} dy\, \frac{e^{-y^2/2}}{\sqrt{2\pi}} \tag{53}$$

is the cumulative Gaussian distribution. Changing the integration variable, one obtains

$$-B = \frac{\sqrt{2}\sigma_\eta}{p\sqrt{\pi}} \int_{-\infty}^{\infty} dx\, x\, \exp(-x^2)\Phi\big((Q - x\sqrt{2}\sigma_\eta)/\sigma_\beta\big). \tag{54}$$

Integrating by parts and omitting technical computations, we obtain

$$B = \frac{\sqrt{2}\sigma_\eta}{p\sqrt{\pi}} \frac{\sigma_\eta}{2\sigma_\beta\sqrt{1+b^2}} \exp\biggl(-\frac{a^2}{1+b^2}\biggr), \tag{55}$$

where $a = Q/(\sqrt{2}\sigma_\beta)$ and $b = \sigma_\eta/\sigma_\beta$. Setting

$$Q = \sigma_\beta\sqrt{2}\, q, \qquad q = \mathrm{erf}^{-1}(2p-1), \tag{56}$$

we obtain

$$B = \frac{\sigma_\eta}{p\sqrt{2\pi}} \frac{1}{\sqrt{1+(\sigma_\beta/\sigma_\eta)^2}} \exp\biggl(-\frac{q^2}{1+(\sigma_\eta/\sigma_\beta)^2}\biggr), \tag{57}$$

from which

$$\beta_{\text{low factor}} \approx -50\% \frac{\sigma_\eta}{p\sqrt{2\pi}} \frac{1}{\sqrt{1+(\sigma_\beta/\sigma_\eta)^2}} \exp\biggl(-\frac{q^2}{1+(\sigma_\eta/\sigma_\beta)^2}\biggr). \tag{58}$$

From the data for the USA, we estimate the standard deviation of the measured beta ($\sigma_\beta = 0.43$), the volatility of the stock index ($\sigma_I = 19.77\%$), the volatility of the low volatility factor (3.46%), and $\langle\sigma_i\rangle/\sigma_I = 1.53$. Setting $\lambda_\beta = 1/90$, we obtain from Eq. (42) $\sigma_\eta = 1.53\sqrt{1/90} = 0.1613$. For $p = 0.3$ (bottom 30%), we obtain $q = -0.3708$ and, thus, $\beta_{\text{low factor}} \approx 0.0334$ from Eq. (58). Finally, we conclude that $\rho_{\text{low factor}} = 3.34\% \frac{19.77\%}{3.46\%} = 19.1\%$.

B Construction of the beta-neutral factors

We implement the four most popular strategies as four beta-neutral factors that are constructed as follows. First, we split all stocks into six clusters of sectors of similar sizes to minimize sectorial correlations. For each trading day, the stocks of the chosen cluster are sorted according to the indicator (e.g., the capitalization) available the day before (we use the publication date and not the valuation date). The related indicator-based factor is formed by buying the first pN stocks in the sorted list and shorting the last pN stocks, where N is the number of stocks in the considered cluster and p is a chosen quantile level. As described in Sec. 3.2, we use $p = 0.15$ for short-term reversal and long-term momentum factors and $p = 0.30$ for the capitalization and low volatility factors. The other stocks (with intermediate indicator values) are not included (weighted by 0). To reduce the specific risk, the weights of the selected stocks are set inversely proportional to the stock's volatility σ_i, whereas the weights of the remaining stocks are 0. Moreover, the inverse stock volatility is limited to reduce the impact of extreme specific risk. For each trading day, we recompute the weight w_i as follows

$$w_i = \begin{cases} +\mu_+ \min\{1, \sigma_{\text{mean}}/\sigma_i\}, & \text{if } i \text{ belongs to the first } pN \text{ stocks in the sorted list,} \\ -\mu_- \min\{1, \sigma_{\text{mean}}/\sigma_i\}, & \text{if } i \text{ belongs to the last } pN \text{ stocks in the sorted list,} \\ 0, & \text{otherwise,} \end{cases} \tag{59}$$

where $\sigma_{\text{mean}} = \frac{1}{N}(\sigma_1 + \ldots + \sigma_N)$ is the mean estimated volatility over the cluster of sectors. In this manner, the weights of low-volatility stocks are reduced to avoid strongly unbalanced portfolios concentrated in such stocks. The two common multipliers, $\mu_\pm$, are used to ensure the beta market neutral condition:

$$\sum_{i=1}^{N} \beta_i w_i = 0, \tag{60}$$

where β_i is the sensitivity of stock i to the market obtained either by an OLS or by our reactive method. In every case, the method to estimate beta uses the rolling daily returns and only past information to avoid the look-ahead bias. If the aggregated sensitivity of the long part of the portfolio to the market is higher than that of the short part of the portfolio, its weight is reduced by the common multiplier $\mu_+ < \frac{1}{2pN}$, which is obtained from Eq. (60) by setting $\mu_- = \frac{1}{2pN}$ (which implies that the sum of absolute weights $|w_i|$ does not exceed 1). In the opposite situation (when the short part of the portfolio has a higher aggregated beta), one sets $\mu_+ = \frac{1}{2pN}$ and determines the reducing multiplier $\mu_- < \frac{1}{2pN}$ from Eq. (60). The resulting factor is obtained by aggregating the weights constructed for each supersector. We emphasize that the factors are constructed on a daily basis, i.e., the weights are re-evaluated daily based on updated indicators. However, most indicators do not change frequently, so the transaction costs related to changing the factors are not significant.

C Description of alternative methods

C.1 Unconditional beta

The theory. Chan and Lakonishok (1992) produce an empirical analysis that describes various robust methods for estimating constant beta, as they provide an alternative to ordinary least squares (OLS). Their method is built on the work of Koenker (1978), which provides robust alternatives to the sample mean using a more complex linear combination of order statistics in order to face the case of non-Gaussian errors, which are the source of outliers. Instead of minimizing the sum of squared residuals, they consider an estimator that is based on minimizing the criterion, including a penalty function ϱ on the residuals ϵ:

$$\sum_{t=1}^{T} \varrho_\theta(\epsilon_t) \tag{61}$$

for $\varrho_\theta(\epsilon_t) = \theta\,|\epsilon_t|$ if $\epsilon_t \geq 0$, or $(1-\theta)\,|\epsilon_t|$ if $\epsilon_t < 0$, where $0 < \theta < 1$.

Chan and Lakonishok (1992) minimize the sum of absolute deviations of the residuals ϵ_{it} from the market model instead of the sum of squared deviations. The resulting minimum absolute deviations (MAD) estimator of the regression parameters corresponds to the special case of $\theta = 1/2$, where half of the observations lie above the line, while half lie below. More generally, large or small values of the weight θ attach a penalty to observations with large positive or negative residuals. Varying θ between 0 and 1 yields a set of regression quantile estimates $\hat{\beta(\theta)}$ that is analogous to the quantiles of any sample of data. However, they recognize that MAD does not prove itself to be a clearly superior method, and they suggest that it may be improved via linear combinations of sample quantiles such as trimmed means.

For that reason, Chan and Lakonishok (1992) test different combinations of regressions quantiles serving as the basis for the robust estimators. They discuss the general case of the trimmed regression quantile (TRQ) given as a weighted average of the regression quantile statistics:

$$\hat{\beta}_\alpha = (1-2\alpha)^{-1} \int_\alpha^{1-\alpha} \hat{\beta}(\theta) d\theta \tag{62}$$

where $0 < \alpha < 1/2$ and $0 < \theta < 1$.

More specifically, Chan and Lakonishok (1992) suggest a more straightforward and equivalent method that considers estimators that are finite linear combinations of regression quantiles (QR) and are computationally simpler:

$$\beta_\omega = \sum_{i=1}^{N} \omega_i \hat{\beta}(\theta_i) \tag{63}$$

where weights $0 < \omega_i < 1$, $i = 1, ..., N$ and $\sum_{i=1}^{N} \omega_i = 1$. The specific case of the weighted average is given by Tukey's trimean (TRM) estimator:

$$\hat{\beta}_{TRM} = 0.25\hat{\beta}(1/4) + 0.5\hat{\beta}(1/2) + 0.25\hat{\beta}(3/4) \tag{64}$$

The application. Their analysis is based mainly on simulated return data, although they add some tests with actual return data. The main advantages of a simulation are that the true values of the underlying parameters are known and that the extent of departures from normality can be controlled. They begin with a baseline simulation with 25,000 replications using data generated from a normal distribution. They also consider the case where the residual term is drawn from a Student-distribution with three degrees of freedom in order to explain the observed leptokurtosis in daily return data. We follow the same methodology to assess the quality of the OLS, the MAD and the TRM beta estimators using Gaussian and t-Student residuals in the seven types of Monte Carlo simulations (MC1,...,MC7).

To replicate the exponential weight scheme of the reactive model ($\lambda_\beta = 1/90$), Eq. (61) is replaced by

$$\sum_{t=1}^{T} (1 - \lambda_\beta)^{T-t} \varrho_\theta(\epsilon_t) \tag{65}$$

C.2 Conditional Beta

The theory. The first application of time-varying beta was proposed in Bollerslev *et al.* (1988), since the beta was computed as the ratio of the conditional covariance to the conditional variance. Engle (2002) generalizes Bollerslev (1990)'s constant correlation model by making the conditional correlation matrix time-dependent with the Dynamic Conditional Correlation (DCC) model, which constrains the time-varying conditional correlation matrix to be positive definite and the number of parameters to grow linearly by a two-step procedure. The first step requires the GARCH variances to be estimated univariately. Their parameter estimates remain constant for the next step. The second stage is estimated conditioned on the parameters estimated in the first stage.

Hereafter, we extend the modeling of the DCC beta for the inclusion of an asymmetric term in the conditional variance equation. In the case of asymmetry in the conditional variance, we select the GJR-GARCH(1,1) specification by Glosten *et al.* (1993), which assumes a specific parametric form with leverage effect in the conditional variance (DCC-GJR beta). The basic idea is that negative shocks at period $(t-1)$ have a stronger impact on the conditional variance at period t than positive shocks. Note that even though the conditional distribution is Gaussian, the corresponding unconditional distribution can still present excess kurtosis.

We select the ADCC model by Cappiello *et al.* (2006) to incorporate asymmetry in correlation.[6] The case mixing asymmetry located in the variance equation (GJR-GARCH) and in the correlation equation

[6]There is a rich literature documenting the existence of asymmetry in correlation overall during bear markets. To cite a few examples, Ang and Bekaert (2000) find evidence of the presence of a high volatility and high correlation regime that tend to coincide with a bear market. Longin and Solnik (2001) find that

(ADCC) is examined (ADCC-GJR GARCH). In our paper, the symmetric GARCH DCC will be called simply DCC, and the asymmetric ADCC-GJR will be called simply ADCC.

Let us consider r_i and r_I as the returns of a single stock and the stock index, respectively. We assume that their respective conditional variances follow a (GJR-)GARCH(1,1) specification. The stock return r_i is defined by its conditional volatility, σ_i, and a zero-mean white noise $\xi_i(t)$:

$$r_i(t) = \sigma_i(t-1)\xi_i(t) \tag{66}$$

The conditional variation specification of the stock return is the following:

$$\sigma_i^2(t) = (1-a-b-\gamma/2)\tilde{\sigma}_i^2 + a\sigma_i^2(t-1)[\xi_i(t)]^2 + b\sigma_i^2(t-1) + \gamma\sigma_i^2[\xi_i^-(t)]^2 \tag{67}$$

where $\tilde{\sigma}_i$ is the unconditional volatility, and a, b, and γ are parameters reflecting respectively the ARCH, GARCH and asymmetry effects. When $\gamma = 0$, the specification collapses to a GARCH model; otherwise, it stands for the GJR-GARCH model, where the asymmetric term is defined as $\xi_i^-(t) = \xi_i(t)$ if $\xi_i(t) > 0$, or $\xi_i^-(t) = 0$ otherwise.

The stock index return r_I is defined by its conditional volatility, σ_I, and a zero-mean white noise $\xi_I(t)$ that is correlated to $\xi_i(t)$:

$$r_I(t) = \sigma_I(t-1)\xi_I(t) \tag{68}$$

The conditional variance specification of the stock index return is the following:

$$\sigma_I^2(t) = (1-a-b-\gamma/2)\tilde{\sigma}_I^2 + a\sigma_I^2(t-1)[\xi_I(t)]^2 + b\sigma_I^2(t-1) + \gamma\sigma_I^2[\xi_I^-(t)]^2 \tag{69}$$

We define the normalized conditional variance diagonal terms as follows:

$$q_{ii}(t) = (1-a_\rho-b_\rho-\gamma_\rho/2) + a_\rho\xi_i(t-1)\xi_i(t-1) + b_\rho q_{ii}(t-1) + \gamma_\rho\xi_i^-(t-1)\xi_i^-(t-1) \tag{70}$$

$$q_{II}(t) = (1-a_\rho-b_\rho-\gamma_\rho/2) + a_\rho\xi_I(t-1)\xi_I(t-1) + b_\rho q_{II}(t-1) + \gamma_\rho\xi_I^-(t-1)\xi_I^-(t-1) \tag{71}$$

The normalized conditional covariance term $q_{iI}(t)$ is given by:

$$q_{iI}(t) = (1-a_\rho-b_\rho-\gamma_\rho/4)\tilde{\rho} + a_\rho\xi_i(t-1)\xi_I(t-1) + b_\rho q_{iI}(t-1) + \gamma_\rho\xi_i^-(t-1)\xi_I^-(t-1) \tag{72}$$

---

correlation among large negative returns is much larger than the correlation among large positive returns. Forbes and Rigobon (2002) warn that the correlation can increase only because the volatility increases even if the beta remains constant. To that end, there has been a controversy in the literature on the statistical significance of such an asymmetry. For this purpose, Ang and Chen (2002) develop a summary statistic that quantifies the degree of asymmetry in correlations across downside and upside markets relative to a particular model. They find that stocks from either small firms, value firms, or low past returns firms, tend to exhibit greater asymmetric correlations. Hong, Tu and Zhou (2006) extend the Ang and Chen (2002) analysis to a model-free approach so that if symmetry is rejected, then the data cannot be modeled by any symmetrical distributions. They find that the betas can be asymmetric even if there is no asymmetry in the correlation. They also find strong evidence of asymmetries for both size and momentum portfolios, but no evidence for book-to-market portfolios. Jiang, Wu and Zhou (2018) extend the Hong, Tu and Zhou (2006)'s correlation-based test approach to finally find that asymmetry is much more pervasive than previously thought. Indeed, they address asymmetry beyond the second moment as the correlation coefficient is a measure of linear dependence, captured by the market beta, between individual stock returns and the market portfolio return. In contrast to Hong, Tu and Zhou (2006), they finally find evidence of asymmetry in some portfolios sorted by the book-to-market ratio.

When $\gamma_\rho = 0$, the specification collapses to a DCC model; otherwise, it stands for the ADCC model, where the asymmetric term is defined as $\xi_i^-(t) = \xi_i(t)$ if $\xi_i(t) > 0$, or $\xi_i^-(t) = 0$ otherwise.

The conditional correlation between $\xi_I(t+1)$ and $\xi_i(t+1)$ is then updated by:

$$\rho_{iI}(t) = q_{iI}(t)/\sqrt{q_{II}(t)q_{ii}(t)} \tag{73}$$

The beta DCC and beta ADCC estimation are defined in the same way:

$$\beta_{DCC}(t) = \rho_{iI}(t)\sigma_i(t)/\sigma_I(t) \tag{74}$$

The log-likelihood function is optimized to calibrate the parameters $\tilde{\rho}$, $\tilde{\sigma}_I$ and $\tilde{\sigma}_i$ for estimation:

$$L_{DCC} = \frac{1}{2}\sum_t^T (L_V(t) + L_C(t)) \tag{75}$$

$$L_V(t) = -2\log(2\pi) - \xi_I(t)^2 - \xi_i(t)^2 - 2\log(\sigma_I(t)) - 2\log(\sigma_i(t)) \tag{76}$$

$$L_C(t) = -\log(det(R(t))) - U'(t)R(t)^{-1}U(t) - U'(t)U(t) \tag{77}$$

with det as the determinant of a matrix, and

$$R(t) = \begin{bmatrix} 1 & \rho_{iI}(t) \\ \rho_{iI}(t) & 1 \end{bmatrix}, \qquad U(t) = \begin{bmatrix} \xi_i(t) \\ \xi_I(t) \end{bmatrix} \tag{78}$$

The application. For Monte Carlo simulation purposes:

- $\xi_i(t)$ is either generated randomly in MC6 and MC7 according to a standard Gaussian or measured through returns $r_i(t)$ and $\sigma_i(t-1)$ for beta DCC estimation.
- $\gamma = 0$ for MC6 and beta DCC estimation but $\gamma > 0$ for MC7 and beta ADCC, which captures the asymmetry term of the GJR-GARCH.
- $\xi_I(t)$ is either generated randomly in MC6 and MC7 according to a standard Gaussian random variable that is correlated with the random variable $\xi_i(t)$ (the correlation between $\xi_i(t)$ and $\xi_I(t)$ is $\rho_{iI}(t-1)$) or is measured through returns $r_I(t)$ and $\sigma_I(t-1)$ for beta DCC estimation.
- $\gamma_\rho = 0$ for MC6 and beta DCC but $\gamma_\rho > 0$ for MC7 and beta ADCC, which captures the asymmetry term of the ADCC.

The fixed parameters that are supposed to be known when testing the beta DCC are set to US market estimates from Sheppard (2017):

- fixed parameters for the univariate symmetric GARCH(1,1) process (MC6, i.e., DCC):

  $b = 0.89$, b is the decay coefficient, and $1/(1-b)$ is related to the number of days the process needs to mean revert;

  $a = 0.099$ describes the level of the volatility of the volatility.

- fixed parameters for the univariate asymmetric GJR-GARCH(1,1,1) process (MC7, i.e., ADCC):

  $b = 0.901$, b is the decay coefficient, and $1/(1-b)$ is related to the number of days the process needs to mean revert;

  $a = 0.0$, $a + \gamma/2$ describe the level of the volatility of the volatility;

  $\gamma = 0.171$, γdescribe the asymmetry.

The fixed parameters that are supposed to be known when testing the DCC and ADCC betas are set to US market estimates from Cappiello *et al.* (2006):

- fixed parameters for the symmetric cross-term process (MC6, i.e., DCC):
  $b_\rho = 0.9261$, b_ρ is the decay coefficient and is linked to the relaxation time;
  $a_\rho = 0.0079$ describes the level of the volatility.
- fixed parameters for the asymmetric cross-term process (MC7, i.e., ADCC):
  $b_\rho = 0.9512$, b_ρ is the decay coefficient and is linked to the relaxation time;
  $a_\rho = 0.0020$, $a_\rho + \gamma_\rho/4$ describes the level of the volatility of the correlation;
  $\gamma_\rho = 0.0040$, γ_ρ describes the asymmetry.

The fixed parameters that are not known when testing the DCC beta and are estimated through the optimization of log-likelihood are set by MC simulation to:

- $\tilde{\rho} = 0.15/0.4$, unconditional correlation;
- $\tilde{\sigma}_I = 0.15/\sqrt{255}$, $\tilde{\sigma}_i = 0.4/\sqrt{255}$ unconditional stock index volatility;
- $\tilde{\sigma}_i = 0.4/\sqrt{255}$ unconditional single stock volatility.

To replicate the exponential weight scheme in the reactive model ($\lambda_\beta = 1/90$), Eq. (75) is replaced by

$$L_{DCC} = \frac{1}{2} \sum_{t}^{T} (1 - \lambda_\beta)^{T-t} \left(L_V(t) + L_C(t)\right) \tag{79}$$

5. The Model of Diffusion of Correlations between Securities

[1]

1. The results of this chapter were obtained in collaboration with Stanislav Kuperstein.

The model of diffusion of the correlations between securities

June 7, 2019

Abstract

The measurement of diffusion of the correlation matrix between securities is almost impossible as diffusion of the correlation of population is hidden by measurement noises. The use of five minutes returns and the reduction of the dimension of the matrix from 500 single stocks to the 23 main market neutral risk factors by taking into account financial information allow us to measure some diffusion patterns. The empirical distribution of the eigenvalues of the increments of the correlation matrix is estimated. The deformation of that distribution with time scale is also studied. We introduce an alternative model that is based on a stochastic equation governing the volatilities of the risk factors. The non-orthogonality of the factors enables to generate an interesting behavior and enables to fit the empirical diffusion patterns of the eigenvectors: The eigenvectors of the matrix tend to be invested at time t on the risk factors that are the most volatile at time t and therefore diffuse as well due to the endless rotation of the most volatile factors.

1 Introduction

The measurement of diffusion of the correlation matrix between securities is almost impossible as the diffusion is hidden by in measurement noises. Allez and Bouchaud (2012) show that the empirical main eigenvector is oscillating around the population one and that the angle of oscillation is related to the ratio between the second eigenvalue and the first eigenvalue and to the ratio between the size of the matrix and the number of independent returns per instrument. We could extend this rule for other eigenvectors and we guess that they are measured with a lot of noises as eigenvalues become closer to their neighbors. As a result the eigenvectors with corresponding small eigenvalues are not well defined or are chaotic. Valeyre *et al.* (2018) introduces a powerful method that takes advantage of the financial information to filter the measurement noises by reducing the dimension of the correlation matrix of single stock returns to 24 main risk factors, among which 23 are market neutral. This method reproduces the main eigenvectors orthogonal to the market mode and their dynamics. We exclude the market mode from the analysis as the first eigenvector appears to be less noisy and less random with a high eigenvalue. So we focus the study on the market neutral sub space. When the method introduced by Valeyre *et al.* (2018) is applied with 5 minutes returns to estimate the correlation matrix, we can measure properly the weekly variation of the correlation. That enables to test for the first time how well the different mainstream and theoretical models of the literature, that describe how population covariance matrix can change and be stochastic, can reproduce the empirical statistics of the variations.

The Wishart process is a mainstream tool to model the dynamics of the covariance matrix. This process can be interpreted as the "square" of a matrix of Brownian motions or in its stationary version as the square of a matrix of Ornstein-Uhlenbeck processes. Bru (1991) derived the stochastic equation that describes the dynamics of the eigenvalues that repulse each other. Their distribution follows the Marčenko-Pastur law for high-dimensional matrices.

Motivated by price multi-asset option or default intensities, Cuchiero *et al.* (2011) analyzed the foundation of the stochastic continuous affine process on the universe of covariance matrices. The Wishart process extends in fact the Feller diffusion from one dimension to several dimensions. Gourieroux (2007) introduced a mean reversion term and extended the process of Cox, Ingersoll and Ross (1985) from one to several dimensions. The process of Cox, Ingersoll and Ross (1985) is very popular in finance to model the dynamics of the interest rate or the volatility of single stocks. In that way, Gourieroux (2007) could model properly the risk of a portfolio taking into account the risk that correlation could change. In the same way Fonseca, Grasselli and Tebaldi (2008) extended to several dimensions to price basket options the model of Heston (1993) where the volatility of the Brownian process, that describes prices, is stochastic and modeled by a CIR process.

Other stochastic matrices are very well documented. Ahdida and Alfonsi (2013) worked on a mean reversion process of correlation matrix through the Wright-Fisher diffusion. Plenty of algorithms were also documented to generate random walk among the ensemble of the rotation matrices, that can be used to describe directly the diffusion of the eigenvectors of the correlation matrix. As an example the Walk by Kac (1959) is a very efficient algorithm that generates random paths but there is no mean reversion component so that after a while the matrix loses the connection with the initial matrix. Gaussian matrices were also very well studied. The distribution of the eigenvalues of the symmetric Gaussian matrix is the well-known Wigner semi-circle law. This is an important point if we consider that the increment of a covariance or correlation matrix could be well modeled by a Gaussian matrix, we can guess that the distribution of the eigenvalues of the increment should be close to the semi-circle law.

In this paper we first define in Section 2 the empirical diffusion patterns we want to reproduce. We then introduce our diffusion model in Section 3. We then present how well the model captures the empirical patterns in Section 4. We finally compare with the results obtained with the mainstream models from the literature in Section 5.

2 The description of the empirical diffusion patterns

2.1 The introduction of the proxy of the correlation matrix between single stocks

Valeyre *et al.* (2018) proposed a practical solution to filter noises of the correlation matrix between single stocks, noted as $\boldsymbol{C(t)}$, by reducing the size of the correlation matrix from 500 or more single stocks to 24 major risk factors that reproduce the largest eigenvalues and their dynamics. These factors were named 'Fundamental Maximum Variance Market neutral' portfolios as their construction was optimized to capture as best as possible the empirical eigenvalues.

We use the same data and factors as in Valeyre *et al.* (2018): we selected the 500 most liquid stocks from the US stock market, from 2013 to 2018 and the $K = 23$ most popular market neutral factors according to the literature (dividend yield, capitalization, volume/capitalization, STR, momentum, beta, leverage, sales to price, book to price, cash to price, price to earning, growth of earning, sensitivity to Euro dollar, sensitivity to 10 years rates, energy, finance, IT, utilities, consumer, industry, pharmacy, consumer discretionary vs. staple, REITs). So we exclude from the analysis the market mode and select only market neutral factors.

Instead of analyzing diffusion of the large and noisy matrix $\boldsymbol{C(t)}$, we analyze diffusion of its proxy which is the reduced matrix $\boldsymbol{C}_p(t)$, Eq. (1), introduced in Eq. (25) of Valeyre *et al.* (2018). $\boldsymbol{C}_p(t)$ depends on the reduced $\boldsymbol{h}(t)$ and $\boldsymbol{\gamma}(t)$ matrices, that can be interpreted as the overlap between the K factors positions and as the covariance between the K factors returns. They can be estimated accurately using 5 minutes returns based on Eq. (89) of Valeyre *et al.* (2018).

$$\boldsymbol{C}_p(t) = \boldsymbol{\gamma}^{-\frac{1}{2}}(t)\boldsymbol{h}(t)\boldsymbol{\gamma}^{-\frac{1}{2}}(t) \tag{1}$$

Valeyre *et al.* (2018) argue that the main eigenvalues from $\boldsymbol{C}_p(t)$ measured from 5 minutes returns with a lookback of 1 week are very close to the main empirical eigenvalues of the correlation matrix of single stocks and that the dynamics of eigenvalues are also well reproduced. Therefore $\boldsymbol{C}_p(t)$ can be used as a good proxy of $\boldsymbol{C(t)}$.

2.2 The introduction of diffusion patterns

In this subsection, we introduce some diffusion patterns that we believe are important to reproduce. As diffusion of the eigenvalues of the correlation matrix is already well documented in the literature, we focus on diffusion of the eigenvectors. We only know from the literature that Allez and Bouchaud (2012) model the stability of the subspace generated by successive eigenvectors based on the overlap between the new subspace and the old one. But they assumed that the population correlation matrix was constant and that only measurement noise could explain the diffusion of the empirical eigenvectors.

We define the matrix $\boldsymbol{O}(t)$ of the eigenvectors that diagonalize $\boldsymbol{C}_p(t)$. These eigenvectors could be interpreted as the constrained and filtered eigenvectors of the large correlation matrix between single stocks. We have $\boldsymbol{O}^{\mathrm{T}}(t)\boldsymbol{C}_p(t)\boldsymbol{O}(t) = \Omega(t)$, where the diagonal matrix $\Omega(t)$ contains the eigenvalues of $\boldsymbol{C}_p(t)$ but could be interpreted as the constrained eigenvalues of the large correlation matrix between single stocks.

We define $\boldsymbol{S}^0(t,\tau)$ in Eq. (2) as the increment of the correlation matrix corresponding to the time increment τ. We changed the basis and set $\boldsymbol{S}^0(t,\tau)$ into the basis generated by the initial eigenvectors $\boldsymbol{O}(t)$. We define $\boldsymbol{S}^1(t,\tau)$ in Eq. (3) as a tilted version to minor the impact of change in eigenvalues $\Omega(t+\tau)-\Omega(t)$ after the time increment τ and get measurement only sensitive to diffusion of eigenvectors of single stocks but not to the diffusion of eigenvalues. We define, in Eq. (4), $\boldsymbol{S}^2(t,\tau)$ as a tilted version to minor the impact of the eigenvalues $\Omega(t)$ in the weighting and get measurement that is based on the same weight for all eigenvectors (major or minor ones).

$$\boldsymbol{S}^0(t,\tau) = \boldsymbol{O}^{\mathrm{T}}(t)\left(\boldsymbol{C}_p(t+\tau) - \boldsymbol{C}_p(t)\right)\boldsymbol{O}(t)\,, \quad (2)$$

$$\boldsymbol{S}^1(t,\tau) = \left(\Omega(t)/\Omega(t+\tau)\right)^{\frac{1}{2}}\boldsymbol{O}^{\mathrm{T}}(t)\boldsymbol{C}_p(t+\tau)\boldsymbol{O}(t)\left(\Omega(t)/\Omega(t+\tau)\right)^{\frac{1}{2}} - \boldsymbol{O}^{\mathrm{T}}(t)\boldsymbol{C}_p(t)\boldsymbol{O}(t)\,, \quad (3)$$

$$\boldsymbol{S}^2(t,\tau) = \mathrm{CorrCov}\left(\boldsymbol{O}^{\mathrm{T}}(t)\boldsymbol{C}_p(t+\tau)\boldsymbol{O}(t)\right) - Id\,, \quad (4)$$

Matrices $\boldsymbol{S}^0(t,\tau)$, $\boldsymbol{S}^1(t,\tau)$ and $\boldsymbol{S}^2(t,\tau)$ quantify whether "old" eigenvectors $\boldsymbol{O}(t)$ are still close to be the "new" eigenvectors $\boldsymbol{O}(t+\tau)$. $\boldsymbol{S}(t,\tau)$ measures the way the portfolios that were initially fixed as eigenvectors start to be correlated to each other as time τ elapses. We can interpret the eigenvalues of $\boldsymbol{S}(t,\tau)$ as a measure of the way how the eigenvectors of the proxy of the large correlation matrix between single stocks diffuse. The direct way that would have consisted in measuring the distance between old and new eigenvectors would have not made any sense as eigenvectors with corresponding small eigenvalues are close to be chaotic and change dramatically as τ changes a little, while the eigenvalues of $\boldsymbol{S}$ remain empirically continuous with τ.

To be more precise as the eigenvalues of $\boldsymbol{S}^0(t,\tau)$ are also simply the eigenvalues of $\boldsymbol{C}_p(t+\tau)-\boldsymbol{C}_p(t)$, the eigenvalues of $\boldsymbol{S}^0(t,\tau)$ could be interpreted as the eigenvalues of the increments of the large but cleaned correlation matrix between single stocks.

The eigenvalues of $\boldsymbol{S}^2(t,\tau)$ are also simply the eigenvalues of the change in correlation between main risk factors. $\boldsymbol{S}^2(t,\tau)$ is particularly interesting and could reveal peculiar properties as $\boldsymbol{S}^2(t,\tau)$ is rather homogeneous and random so it could be close to symmetric Gaussian matrix and we can expect that the distribution of the eigenvalues looks like to a deformed version of the Wigner semi-circle law, without any tails and singularities.

We define two diffusion patterns to measure from real data that we want to reproduce with stochastic models for both $\boldsymbol{S}^0$, $\boldsymbol{S^1}$ and $\boldsymbol{S^2}$:

- $|\boldsymbol{\lambda}|_{\max}(\tau)$ is the largest eigenvalues in absolute value of $\boldsymbol{S}^0(t,\tau)$, $\boldsymbol{S}^1(t,\tau)$ or $\boldsymbol{S}^2(t,\tau)$ the depending on the time scale of the increment τ and averaged on the different t. The increase of $|\boldsymbol{\lambda}|_{\max}(\tau)$ with τ describes the way the portfolios that were initially set as eigenvectors starts to be more and more correlated. $|\boldsymbol{\lambda}|_{\max}(\tau)$ will converge toward an asymptotic value when τ tends to ∞ with a certain relaxation time;
- $\boldsymbol{\rho_\tau}(\lambda)$ is the distribution of the eigenvalues of of $\boldsymbol{S}^0(t,\tau)$, $\boldsymbol{S}^1(t,\tau)$ or $\boldsymbol{S}^2(t,\tau)$ for all different t. The shape of the distribution depends on τ, the time scale of the increment. We cannot reduce diffusion of the largest eigenvalue $|\boldsymbol{\lambda}|_{\max}(\tau)$ and it is important to worry about the distribution of all eigenvalues. Indeed that is important to understand if the risk of correlations change can be extreme and localized by concentrating on only one factor, i.e that is important to determine whether the distribution $\boldsymbol{\rho_\tau}(\lambda)$ could have tails or not if the tails are the results of the distribution of the eigenvalues of the correlation matrix or the results of a more sophisticated phenomenon. That explains why it is interesting to measure the shape of $\boldsymbol{\rho_\tau}(\lambda)$ for both $\boldsymbol{S}$, $\boldsymbol{S^1}$ and $\boldsymbol{S^2}$.

3 The diffusion model governed by the diffusion of the FCL

$$\lambda_i^0(t) = \frac{\boldsymbol{h}_{ii}(t)}{\boldsymbol{\gamma}_{ii}(t)} \tag{5}$$

$\lambda_1^0(t), ..., \lambda_K^0(t)$ are the time dependent Factor correlation levels, "FCL", of the K factors, which were introduced in Valeyre *et al.* (2018). The interpretation of Eq. (5), that defines the "FCL", is that it corresponds to the conditional variance of the returns of the corresponding normalized risk factor. The "FCL" has very appealing properties as for example it corresponds to the weighted average of the eigenvalues of the correlation matrix of the single stocks returns, with the weights given by the squares of the eigenvectors projections of the risk factor. Each market neutral factor was determined with the Maximum Variance formula (Eq.43 combined with Eq.54 of with Valeyre *et al.* (2018) $\nu = 1$), that enables to optimize the "FCL" and to reproduce very well the empirical eigenvalues and their dynamics.

$x_1(t), ..., x_K(t)$ are the logarithm of the ratio between $\lambda_i^0(t)$ and λ_{0i}^0 the unconditional "FCL" estimation (Eq.6) and according to Valeyre *et al.* (2018) they could be well modeled by independent Ornstein-Uhlenbeck processes (Eq.7). σ describes the volatility and α describes the inverse of the relaxation time.

$$x_i(t) = \ln\left(\frac{\lambda_i^0(t)}{\lambda_{0i}^0}\right) \tag{6}$$

$$dx_i(t) = -\alpha x_i(t)dt + \sigma dB_i(t)\,. \tag{7}$$

The parameter α could be fitted from a normalized variogram introduced in Grebenkov and Serror (2014) through Eq. (11). The left graph of Fig. 1 exhibits the normalized variogram of the daily variation of $x_1(t), .., x_K(t)$ for the K risk factors. The normalized variograms are fitted for most factors by an Ornstein-Uhlenbeck process with a relaxation time of 60 days. Some factors deviate but it could be explained by the noise of the measurement of the method. The right graph of Fig. 1 shows the empirical eigenvalues of the unconditional correlation matrix $\langle \boldsymbol{C(t)} \rangle$.

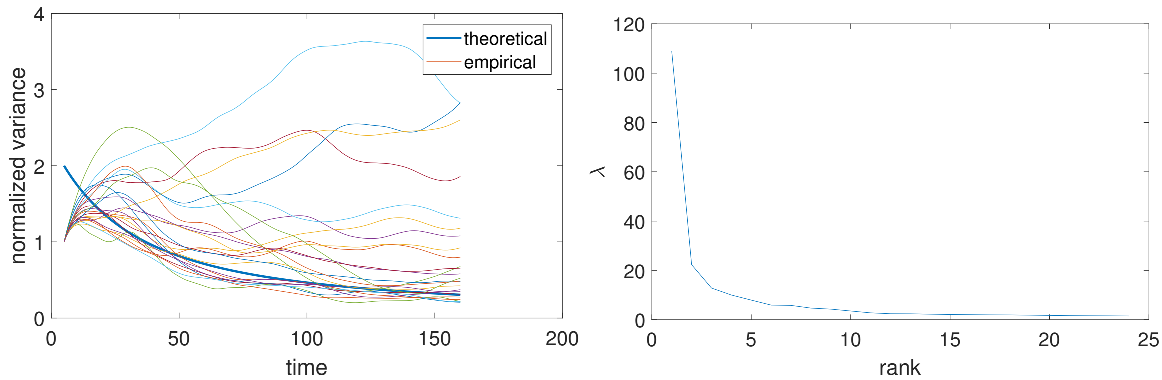

Figure 1: Left: Normalized variogram of the daily variation of the logarithm of the FCL $\lambda_1^0, .., \lambda_K^0$ using five minutes returns from 2013 to 2018. Each curve corresponds to a risk factor. Most of the empirical measurement are fitted by theoretical variogram obtained for the Ornstein-Uhlenbeck process with $1/\alpha = 60$. Right: Eigenvalues of the empirical unconditional correlation matrix, we exclude from the analysis the first one that corresponds to the market mode.

We define $\boldsymbol{\gamma}_0$ and $\boldsymbol{h}_0$ as the unconditional covariance and the unconditional overlap matrices whereas $\boldsymbol{h}(t)$ et $\boldsymbol{\gamma}(t)$ are the conditional matrices depending on time.

We introduce in Eq. (8) the model of diffusion of the correlation matrix that is governed by diffusion of the "FCL". $\boldsymbol{C}_p^{\text{sim}}(t)$ could be generated completely in a random way, to replicate empirical patterns of $\boldsymbol{C}_p(t)$. Eq. (44) of Valeyre *et al.* (2018) describes how to generate the random unconditional matrices ($\boldsymbol{h}_0^{\text{sim}}$, $\boldsymbol{\gamma}_0^{\text{sim}}$ and $\boldsymbol{C}_{p0}^{\text{sim}} = \boldsymbol{\gamma}_0^{\text{sim}-\frac{1}{2}} \boldsymbol{h}_0^{\text{sim}} \boldsymbol{\gamma}_0^{\text{sim}-\frac{1}{2}}$)

through the random selection of the K factors. They are generated randomly once based on the eigenvalues of the empirical unconditional correlation matrix between single stocks. A summary is described in Appendix B. The unconditional empirical eigenvalues Ω_0 are based on empirical eigenvalues using the 5 minutes data from 2013 to 2018 and are reported in Table 1. The first eigenvalue was excluded. Eq. (8) helps to simulate the heteroscedsacticity of the correlation matrix through change in "FCL" with the stochastic processes $x_1(t)$,...$x_K(t)$ that can be updated step by step randomly from $t = 1$ to T using the parameters of Eq. (7). That enables to generate stochastic "FCL" that we could interpret as volatilities for the K risk factors. For numerical simulation we set $T = 1071$.

	λ^{Emp}
1	109.02
2	22.32
3	12.84
4	10.01
5	7.94
6	5.92
7	5.79
8	4.70
9	4.29
10	3.53
11	2.77
12	2.42
13	2.38
14	2.23
15	2.10
16	2.03
17	1.96
18	1.92
19	1.84
20	1.73
21	1.64
22	1.60
23	1.56
24	1.50

Table 1: λ^{Emp} are the sample unconstrained eigenvalues of the correlation matrix between single stocks obtained from 2013-2018 with 5 minutes returns. Ω_0 is set as λ_2^{Emp},...,$\lambda_{24}^{\text{Emp}}$.

$$\boldsymbol{C}_p^{\text{sim}}(t) = \begin{pmatrix} e^{x_1(t)} & ... & 0 \\ ... & & \\ 0 & ... & e^{x_K(t)} \end{pmatrix}^{-\frac{1}{2}} \boldsymbol{C}_{p0}^{\text{sim}} \begin{pmatrix} e^{x_1(t)} & ... & 0 \\ ... & & \\ 0 & ... & e^{x_K(t)} \end{pmatrix}^{-\frac{1}{2}} \tag{8}$$

4 Simulations against measurements

Fig. 2 displays the time scale dependency of $|\boldsymbol{\lambda}|_{\max}(\tau)$ for empirical $\boldsymbol{S}^0$ and simulated $\boldsymbol{S}^{0\text{sim}}$. We see that simulation captures well the measurement. The fit could have been much better if the parameter σ was optimized. The curve looks like a square root law that is a characteristic of diffusion but in fact it seems to converge to a value between 8 and 12 with an exponential decay and a relaxation time close to 60 days $(1/\alpha)$. We can interpret that the asymptote between 8 and 12 as close to the first eigenvalue of $\boldsymbol{C}_p^{\text{sim}}$ multiplied by $\sigma\sqrt{1/\alpha}$. We include $|\boldsymbol{\lambda}|_{\max}(\tau)$ obtained for $\boldsymbol{S}^1$ when withdrawing the impact of diffusion of the eigenvalues $\Omega(t)$ of the correlation matrix on the measurement. We see a small difference between $|\boldsymbol{\lambda}|_{\max}(\tau)$ obtained for $\boldsymbol{S}^0$ and $|\boldsymbol{\lambda}|_{\max}(\tau)$ obtained for $\boldsymbol{S}^1$ and that diffusion of eigenvalues $\Omega(t)$ has a smaller impact than diffusion of the eigenvectors into diffusion of the correlation matrix of single stocks. The very likely scenario that explains diffusion of the correlation matrix is therefore a permanent rotation between factors that matters for risk. In that scenario the distribution of the eigenvalues of the correlation matrix is maintained rather stable.

Fig.3 exhibits $|\boldsymbol{\lambda}|_{\max}(\tau)$ for empirical $\boldsymbol{S}^2$ and simulated $\boldsymbol{S}^{2\text{sim}}$. At $\tau = 60$ days $|\boldsymbol{\lambda}|_{\max}(\tau)$ is close to 1.5 but we have to add another 0.5 to achieve the likely asymptote and another 1 to estimate the first eigenvalue of the correlation matrix between factors that were set initially as eigenvectors of the empirical conditional correlation matrix. Indeed that first eigenvalue corresponds to the first eigenvalue of $Id + \boldsymbol{S}^{2\text{sim}}(t, \tau = \infty)$. So 3 should be compared to 4.01 that corresponds to the value of first eigenvalue of the correlation matrix derived from the empirical unconditional $\boldsymbol{\gamma}_0^{-\frac{1}{2}} \boldsymbol{h}_0 \boldsymbol{\gamma}_0^{-\frac{1}{2}}$ or to 4.14 that corresponds to the average of the first eigenvalue of the correlation matrix derived from the random matrix $\boldsymbol{\gamma}_0^{\text{sim}-\frac{1}{2}} \boldsymbol{h}_0^{\text{sim}} \boldsymbol{\gamma}_0^{\text{sim}-\frac{1}{2}}$ based on a random selection of the K factors. It is therefore almost not worth to try to orthogonalize the random factors as correlation change will destroy a large part of the effect of the orthogonalization after 60 days or more.

Fig. 4 displays $\boldsymbol{\rho_\tau}(\lambda)$, the empirical histogram of the eigenvalues of $\boldsymbol{S}^0$ and the histogram of the simulated $\boldsymbol{S}^{0\text{sim}}$ for different time scales τ from 1 to 30 days. We see that the empirical and simulated histograms are very close. They have tails and are very different from the Wigner semi-circle law, we could have expected if the variations of the correlation matrix were Gaussian. The model appears to be realistic. That is confirmed by Fig. 5 obtained for $\boldsymbol{S}^1$ that is very similar to Fig. 4 obtained for $\boldsymbol{S}^0$. Fig. 6 obtained for $\boldsymbol{S}^2$, that minors the impact of the eigenvalues $\Omega(t)$ in the weighting and get the measurement of diffusion that is based on the same weight for all eigenvectors (major or minor ones), still exhibits a

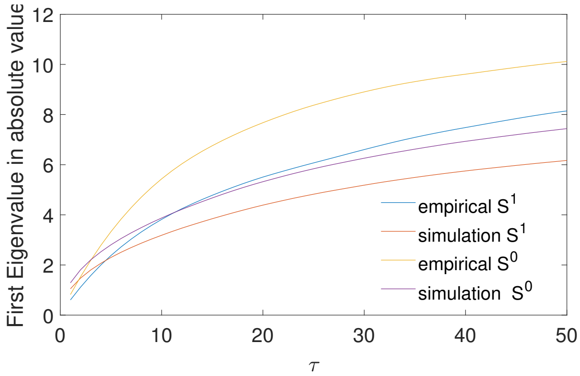

Figure 2: $|\boldsymbol{\lambda}|_{\max}(\tau)$ based on $\boldsymbol{S}^0$, $\boldsymbol{S}^{0\text{sim}}$, $\boldsymbol{S}^1$ and $\boldsymbol{S}^{1\text{sim}}$. It corresponds to the largest eigenvalues in absolute value of the increment of the proxy of the correlation matrix between single stocks depending on τ, the time horizon of the increment. "Empirical S^0" and "Empirical S^1" are the empirical measurements. "Empirical S^1" is the case where the impact of diffusion of the eigenvalues $\Omega(t)$ is artificially withdrawn. "simulation S^0" and "simulation S^1" were obtained with $\sigma = 0.0545$ and $1/\alpha = 60$ to make the "FCL" stochastic. The simulation is generated by our model that captures well the empirical measurements based on the 5 minutes returns on the period 2013-2018.

distribution with tails and remains very different from the Wigner semi-circle law or from the pointed hat shape distribution for both empirical and simulated cases.

5 Comparison with the mainstream models from the literature

To determine how well our model is adapted to the reality, we compare it to other models, that were selected among mainstream models from the literature. We will check if the standard models from the literature also could generate histogram $\boldsymbol{\rho_\tau}(\lambda)$ corresponding to S^0, S^1 and S^2 that are close to the empirical ones.

5.1 Feller Diffusion with the Wishart process

Feller Diffusion with the Wishart process is used to model diffusion of the covariance between portfolios that were initially fixed as eigenvectors.

$$C_{p,t}^{\text{sim}}(\tau) = \Omega_0^{\frac{1}{2}} \frac{Id + \sigma \boldsymbol{B}_t^{\text{T}}(\tau)\boldsymbol{B}_t(\tau)}{1 + \sigma^2 \tau L} \Omega_0^{\frac{1}{2}}$$

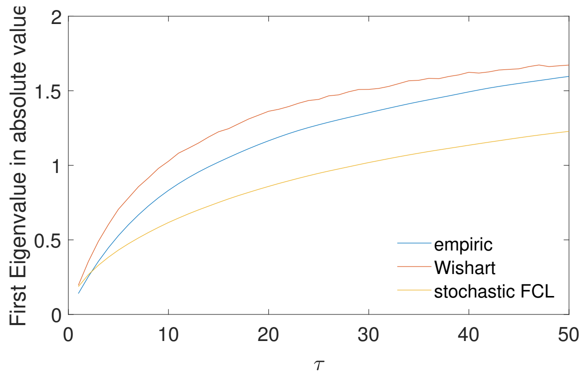

Figure 3: $|\boldsymbol{\lambda}|_{\max}(\tau)$ based on $\boldsymbol{S}^2$, $\boldsymbol{S}^{2\text{sim}}$ for both the Wishart model (derived from Eq. 9) and the model governed by the stochastic "FCL". It corresponds to the largest eigenvalues in absolute value of the normalized increment of the proxy of the correlation matrix between single stocks depending on τ, the time horizon of the increment.

with $\boldsymbol{B}_t(\tau)$ set as a Brownian matrix of size $K \times L$ with the path t, as if diffusion was coming only from statistical error of measurement of the correlation as in Allez and Bouchaud (2012). The diffusion model of the population eigenvectors by the error of measurement of the empirical eigenvectors sounds weird. We simulated $|\boldsymbol{\lambda}|_{\max}(\tau)$ and $\boldsymbol{\rho_\tau}(\lambda)$ for $\boldsymbol{S}^{0\text{sim}}(t,\tau)$ defined by the Eq. (9)

$$\boldsymbol{S}^{0\text{sim}}(t,\tau) = \Omega_0^{\frac{1}{2}} \frac{Id + \sigma^2 \boldsymbol{B}_t^{\mathrm{T}}(\tau)\boldsymbol{B}_t(\tau)}{1+\sigma^2\tau L}\Omega_0^{\frac{1}{2}} - \Omega_0 \tag{9}$$

Here we chose to have independent paths t. We generate plenty of paths $t = 1, .., t = T = 1071$. Ω_0 is set to empirical unconditional eigenvalues of the large correlation matrix between single stocks, the measurements are reported in the Tab.1. The first empirical eigenvalue that corresponds to the market mode was excluded. L and σ were fitted to replicate approximatively, without any optimization, the empirical measurement of $|\boldsymbol{\lambda}|_{\max}(\tau)$.

Fig. 7 shows that the quality of the fit whereas Fig. 8 exhibits how the model reproduces well the empirical histogram of the eigenvalues of $\boldsymbol{S}^0$.

When we simulate

$$\boldsymbol{S}^{2\text{sim}}(t,\tau) = \text{CorrCov}\left(\Omega_0^{\frac{1}{2}} \frac{Id + \sigma^2 \boldsymbol{B}_t^{\mathrm{T}}(\tau)\boldsymbol{B}_t(\tau)}{1+\sigma^2\tau L}\Omega_0^{\frac{1}{2}}\right) - Id,$$

we see in Fig. 9 that the model could not manage to replicate the tails of $\boldsymbol{\rho_\tau}(\lambda)$, the histogram of the increments and generate only distribution without any tails close to the deformed semi-circle law of Wigner.

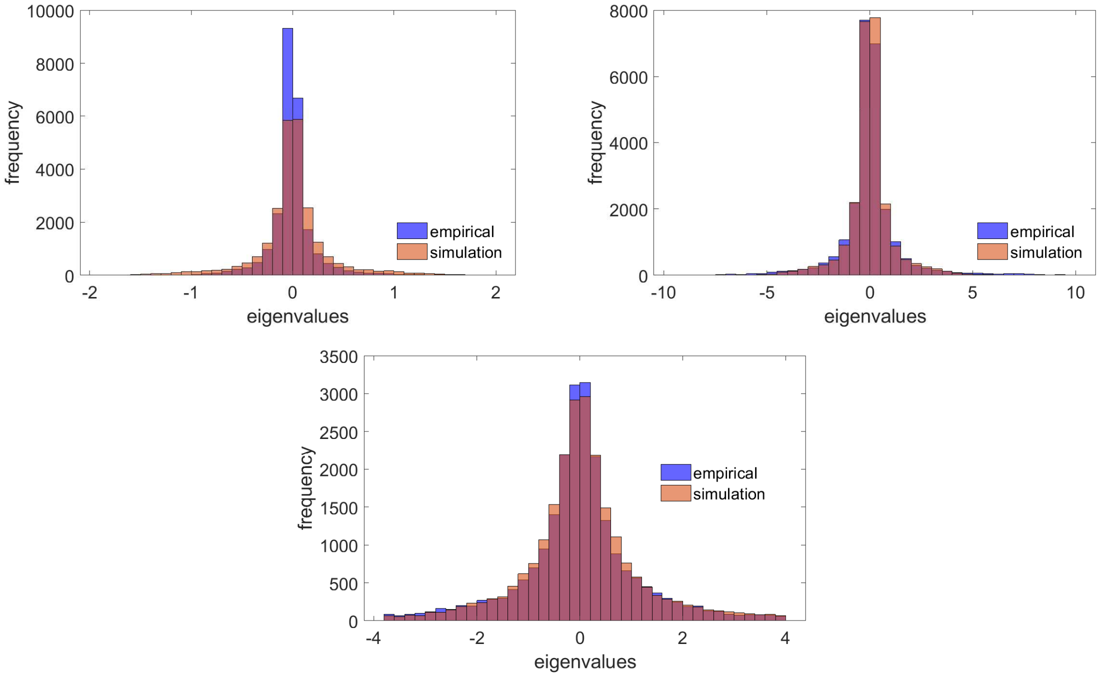

Figure 4: $\boldsymbol{\rho_\tau}(\lambda)$ based on $\boldsymbol{S}^0$ and $\boldsymbol{S}^{0\text{sim}}$ with $\tau = 1$, 10 or 30 days. The histograms correspond to the empirical distribution of the eigenvalues of the increments of the correlation matrix. $\tau = 1$, 10 or 30 days correspond to the time horizon of the increment. The simulation, that was obtained with $\sigma = 0.0545$ and $1/\alpha = 60$ to make the "FCL" stochastic, captures well the measurement based on the 5 minutes returns on the period 2013-2018 at any time scale.

In conclusion the Wishart model could not reproduce the empirical normalized change in correlation. We could have tested with the large matrix $\boldsymbol{C}^{\text{sim}}$ of dimension $N = 500$ instead of $\boldsymbol{C}_p^{\text{sim}}$ of dimension K but it would have generate the same disappointing results with the incapacity to get for $\boldsymbol{\rho_\tau}(\lambda)$ a distribution whose shape is very different from the Wigner semi-circle law.

5.2 Wright-Fisher diffusion with mean reversion term

This diffusion is used to model diffusion of the correlation between portfolios that were initially fixed as eigenvectors. We simulated directly the stochastic correlation matrix, $A(t, \tau)$ generated by the stochastic process introduced by Ahdida and Alfonsi (2013). The different parameters were fitted to replicate approximatively, without any optimization, the empirical measurement of $|\boldsymbol{\lambda}|_{\max}(\tau)$ (Fig. 10). We generated plenty of path $t = 1, ..., t = T = 1071$ and

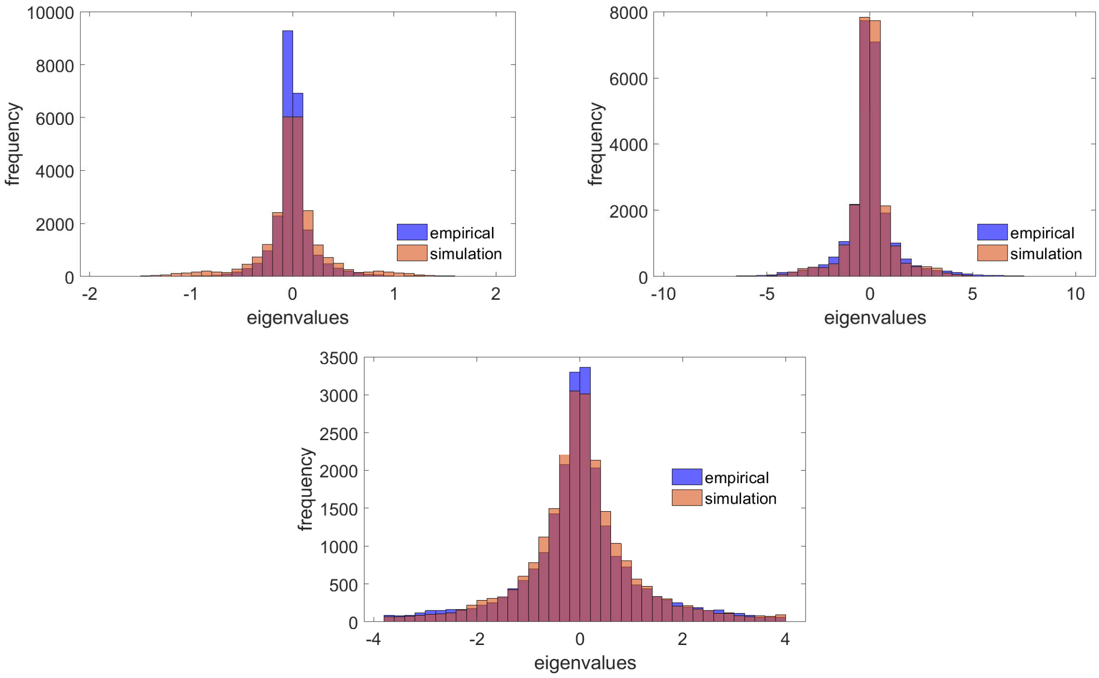

Figure 5: $\boldsymbol{\rho_\tau}(\lambda)$ based on $\boldsymbol{S}^1$ and $\boldsymbol{S}^{1\text{sim}}$ with $\tau = 1$, 10 or 30 days. The histograms correspond to the empirical distribution of the eigenvalues of the increments of the correlation matrix readjusted by the eigenvalues increments. $\tau = 1$, 10 or 30 days correspond to the time horizon of the increment. The simulation, that was obtained with $\sigma = 0.0545$ and $1/\alpha = 60$ to make the "FCL" stochastic, captures well the measurement based on the 5 minutes returns on the period 2013-2018 at any time scale.

$A(t, 0)$ was initialized at the identity matrix for every path. The "mean matrix" of the mean reversion term was also set to the identity. We also plot the histogram of the eigenvalues of $\boldsymbol{S}^{0\text{sim}}(t, \tau)$ defined by Eq. (10) (Fig. 11) that looks realistic.

$$\boldsymbol{S}^{0\text{sim}}(t, \tau) = \Omega_0^{\frac{1}{2}} A(t, \tau) \Omega_0^{\frac{1}{2}} - \Omega_0 \tag{10}$$

When we simulate $\boldsymbol{S}^{2\text{sim}}(t, \tau) = A(t, \tau) - Id$, the Wright-Fisher diffusion could not help to avoid for getting the distribution $\boldsymbol{\rho_\tau}(\lambda)$ that looks like a Wigner semi-circle law (Fig.12). In conclusion the diffusion of Wright-Fisher with mean reversion term could not reproduce the empirical normalized change in correlation;

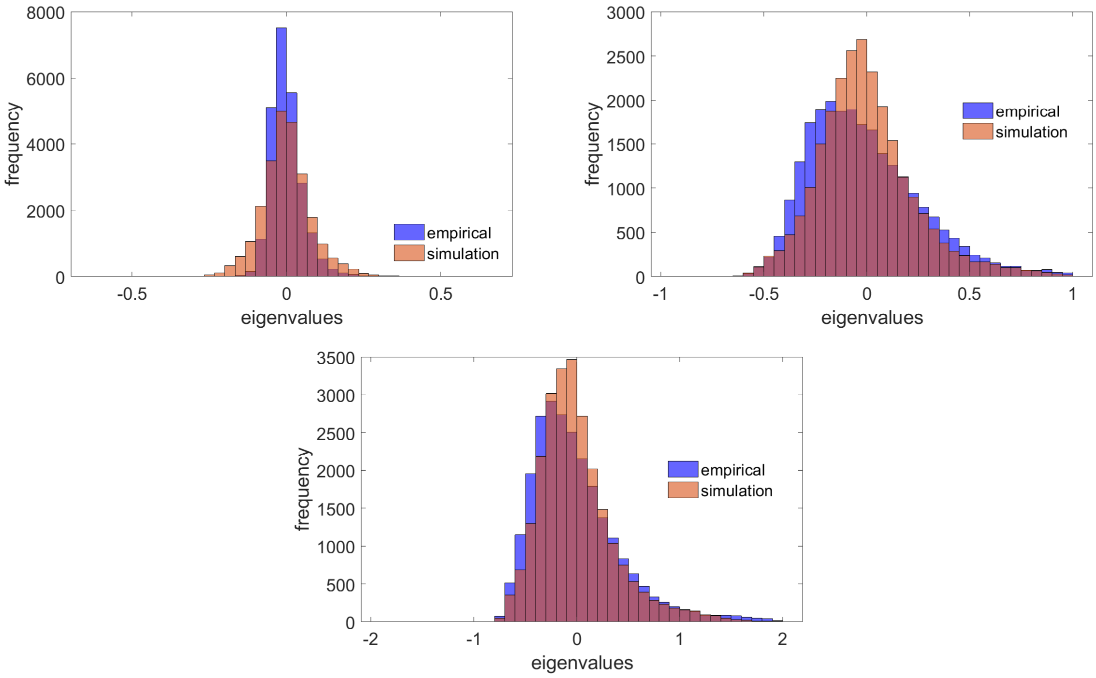

Figure 6: $\boldsymbol{\rho_\tau}\left(\lambda\right)$ based on $\boldsymbol{S}^2$ and $\boldsymbol{S}^{2\text{sim}}$ with $\tau = 1$, 10 or 30 days. The histograms correspond to the empirical distribution of the eigenvalues of the increments of the correlation matrix readjusted by the eigenvalues increments. $\tau = 1$, 10 or 30 days correspond to the time horizon of the increment. The simulation, that was obtained with $\sigma = 0.0545$ and $1/\alpha = 60$ to make the "FCL" stochastic, captures well the measurement based on the 5 minutes returns on the period 2013-2018.

5.3 Mean reversion random walk on the ensemble of the rotation matrices

Mean reversion random walk on the ensemble of the rotation matrices describes directly diffusion of the eigenvectors of C_p^{sim}. We simulate directly $\boldsymbol{O}(t,\tau)$ initialized to the identity matrix. We simulated $|\boldsymbol{\lambda}|_{\max}(\tau)$ and $\boldsymbol{\rho_\tau}\left(\lambda\right)$ for $\boldsymbol{S}^{0\text{sim}}(t,\tau) = \boldsymbol{O}^{\text{T}}(t,\tau)\Omega_0\boldsymbol{O}(t,\tau) - \Omega_0$ for different path $t = 1, .., t = T = 1071$. We also test $\boldsymbol{S}^{2\text{sim}}(t,\tau) = \text{CorrCov}\left(\boldsymbol{O}^{\text{T}}(t,\tau)\Omega_0\boldsymbol{O}(t,\tau)\right) - Id$. Parameters were set to reproduce approximatively, without any optimization, the empirical measurement of $|\boldsymbol{\lambda}|_{\max}(\tau)$ for the three different following methods:

- Gram-Schmidt algorithm that is described in Appendix C.1. We fit the parameters through Fig.13. But Fig.14 exhibits two abnormal bumps on the right and the left and Fig.15 exhibits an abnormal pointed hat shape;

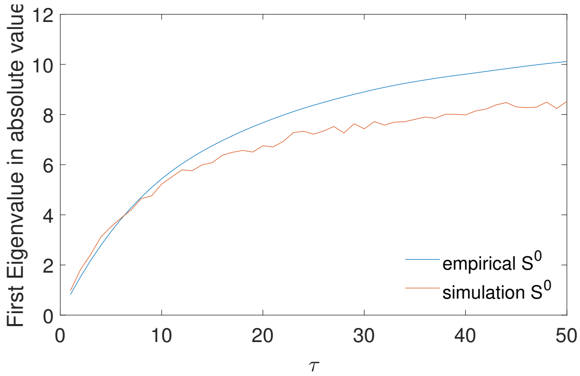

Figure 7: $|\boldsymbol{\lambda}|_{\max}(\tau)$ based on $\boldsymbol{S}^0$, $\boldsymbol{S}^{0\text{sim}}$, $\boldsymbol{S}^1$ and $\boldsymbol{S}^{1\text{sim}}$. It corresponds to the largest eigenvalues in absolute value of the increment of the proxy of the correlation matrix between single stocks depending on τ, the time horizon of the increment. "Empirical S^0" is the empirical measurements. "simulation S^0" is derived from a Wishart process in Eq. (9) derived from a Gaussian matrix of dimension $L \times K$ with $L = 30$ and $K = 23$. The last parameter is $\sigma = 0.06$.

- Algorithms based on the Walk by Kac (1959) that was tuned to include a mean reversion term that is described in Appendix C.1;
- A new stochastic differential equation that is described in Appendix C.1.

In the two cases, we fit the parameters through Fig.13 (Fig.16). But Fig.14 (Fig.17) exhibits two abnormal bumps on the right and the left and Fig.15 (Fig.21)exhibits an abnormal pointed hat shape.

5.4 The classical model

Classical model where the eigenvalues of the matrix, instead of the "FCL", are stochastic. In that model the eigenvectors are stable and do not diffuse.

6 Conclusion

The measurement of diffusion of the correlation matrix is almost impossible as diffusion is hidden by measurement noises. The use of five minutes returns and the reduction of the dimension of the matrix from 500 single stocks to 24 main risk factors allow us to measure some diffusion patterns. The distribution of the eigenvalues of the variation of the matrix is measured. The deformation of the distribution with time scale is studied. The empirical

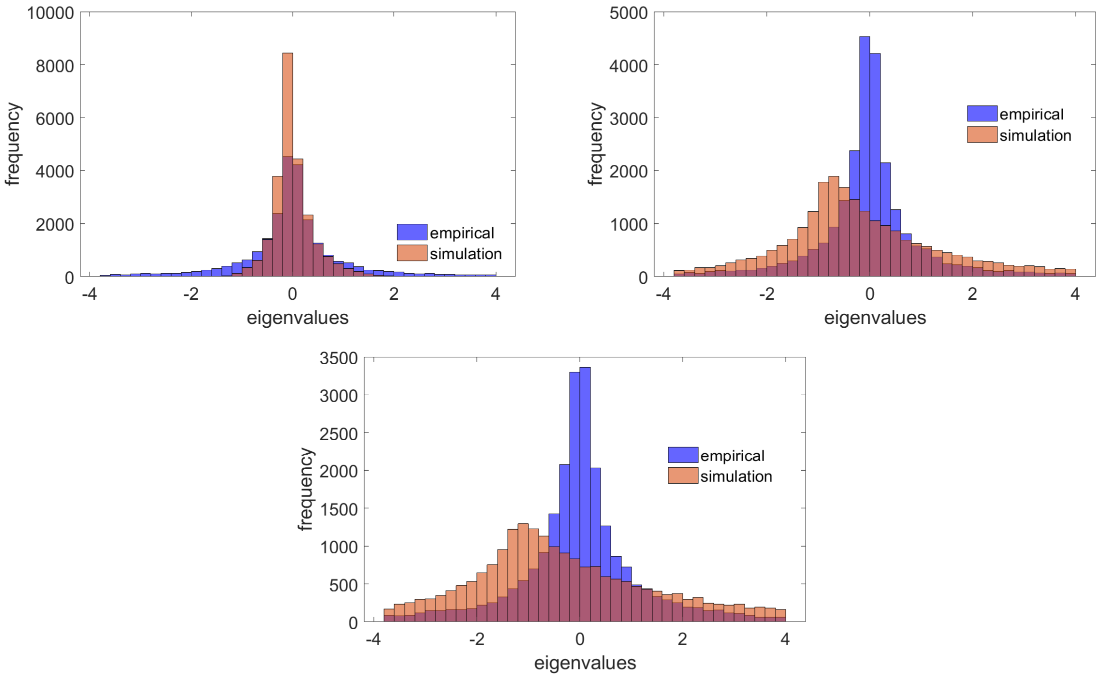

Figure 8: $\boldsymbol{\rho_\tau}(\lambda)$ based on $\boldsymbol{S}^0$ and $\boldsymbol{S}^{0\text{sim}}$ using the Wishart model (Eq. 9) with $\tau = 1$, 10 or 30 days. The histograms correspond to the empirical distribution of the eigenvalues of the increments of the correlation matrix readjusted by the eigenvalues increments. $\tau = 1$, 10 or 30 days correspond to the time horizon of the increment. The simulation was obtained with $L = 30$ and $\sigma = 0.06$.

patterns are not well reproduced by the standard stochastic models derived from the Wishart process or from standard random walk on rotation matrices. We introduce a new alternative model that is based on a stochastic equation for the volatilities of the risk factors that fit the empirical patterns. The eigenvectors of the matrix tend to be invested on the risk factors that are the most volatile and therefore diffuse. Mainstream models could not capture extreme changes of correlation localized in one direction. Our alternative model appears more robust and realistic.

A Variogram

$$\mathbf{V}_i(\tau) = \frac{\text{Var}_{u=T_1}^{T_2}\left(l_i(u) - l_i(u-\tau)\right)}{\text{Var}_{u=T_1}^{T_2}\left(l_i(u) - l_i(u-1)\right) + ...\text{Var}_{u=T_1}^{T_2}\left(l_i(u-\tau) - l_i(u-\tau-1)\right)} \tag{11}$$

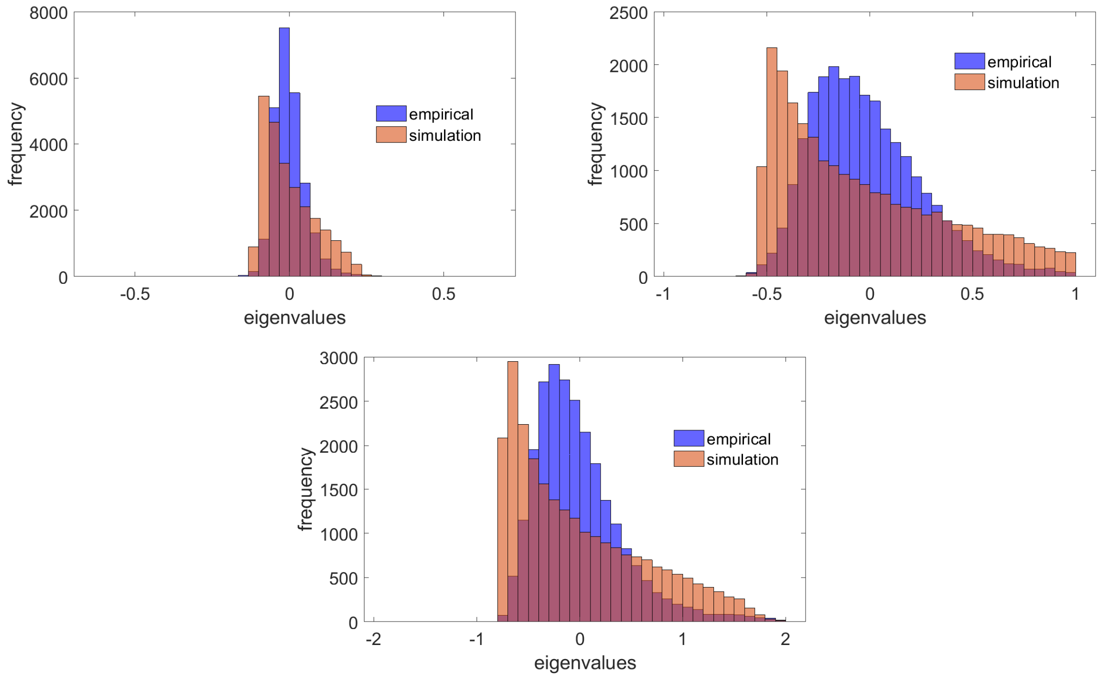

Figure 9: $\boldsymbol{\rho_\tau}(\lambda)$ based on $\boldsymbol{S}^2$ and $\boldsymbol{S}^{2\text{sim}}$ using the Wishart model (normalized version of Eq. 9) with $\tau = 1$, 10 or 30 days. The histograms correspond to the empirical distribution of the eigenvalues of the increments of the correlation matrix readjusted by the eigenvalues increments. $\tau = 1$, 10 or 30 days correspond to the time horizon of the increment. The simulation was obtained with $L = 30$ and $\sigma = 0.06$.

where $\text{Var}_{u=T_1}^{T_2}$ is the empirical variance based on the sample from T_1 to T_2.

B Generating $\boldsymbol{C}_{\text{sim}}(t)$ governed by stochastic FCL

In the Paper we we modelled the returns by

$$r_i(t) = \sum_{j=1}^{N} \sqrt{\ell_j^{\mathbf{r}.}} \left(\boldsymbol{e}_j^{\mathbf{r}.}\right)_i \varepsilon_j(t) \, , \tag{12}$$

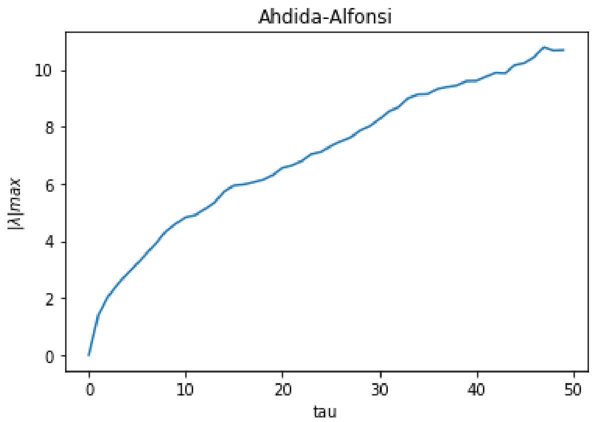

Figure 10: $|\boldsymbol{\lambda}|_{\max}(\tau)$ based on $\boldsymbol{S}^{0\mathrm{sim}}$. It corresponds to the largest eigenvalue in absolute value of the increment of the proxy of the correlation matrix between single stocks depending on τ, the time horizon of the increment. $\boldsymbol{S}^{0\mathrm{sim}}$ is derived from Ahdida and Alfonsi (2013) (Eq. 10) with $|\lambda|_{\max}(\tau)$ for $(k, a) = (1., 0.9)$ and $c = \mathcal{I}$. This choice of k and a provides the best match.

where $\varepsilon_j(t)$ are TN standard normal random variables, $\boldsymbol{e}^{\mathrm{r}\cdot}$ is a random orthonormal basis (meaning that $\boldsymbol{e}^{\mathrm{r}\cdot}_j$ are entries of an $SO(N)$ matrix) and

$$\ell^{\mathrm{r}\cdot}_i = \begin{cases} \lambda^{\mathrm{Emp}}_i & \text{for} \quad i \leqslant K \\ 1 - (N-K)^{-1} \sum\limits_{a=1}^{K} \lambda^{\mathrm{Emp}}_a & \text{for} \quad i > K \end{cases} . \tag{13}$$

For large N the covariance matrix of $r_i(t)$, $\boldsymbol{H}^{\mathrm{r}\cdot}$, is close to the correlation obtained from the same returns. In other words, $H^{\mathrm{r}\cdot}_{ii} \approx 1$. Next, we modelled the Maximum-Variance market neutral portfolios as

$$\left(\boldsymbol{\omega}^{(0)}_{\star\ a}\right)_i = \sum_{j=1}^{N} \left(\ell^{\mathrm{r}\cdot}_j\right)^{\mu/2} \left(\boldsymbol{e}^{\mathrm{r}\cdot}_j\right)_i \mathcal{E}_{ja}\,, \tag{14}$$

where $\boldsymbol{\mathcal{E}}$ is a $N \times K$ matrix of standard normal random variables simulating our factor loadings, and μ is a free parameter fixed to $\mu \approx 1.4$ in order to match the observations. With these conventions the unconditional matrices become:

$$h^{\mathrm{r}\cdot}_{0\ ab} = \boldsymbol{\omega}^{(0)\ \mathrm{T}}_{\star\ a} \boldsymbol{H}^{\mathrm{r}\cdot} \boldsymbol{\omega}^{(0)}_{\star\ b} \qquad \text{and} \qquad \boldsymbol{\gamma}^{\mathrm{r}\cdot}_{0\ ab} = \sum_{i=1}^{N} \left(\omega^{(0)}_{\star\ a}\right)_i H^{\mathrm{r}\cdot}_{ii} \left(\omega^{(0)}_{\star\ b}\right)_i . \tag{15}$$

These matrices are used to define

$$\boldsymbol{C}^{\mathrm{r}\cdot}_0 = \boldsymbol{\gamma}^{\mathrm{r}\cdot\,-1/2}_0 \boldsymbol{h}^{\mathrm{r}\cdot}_0 \boldsymbol{\gamma}^{\mathrm{r}\cdot\,-1/2}_0 \tag{16}$$

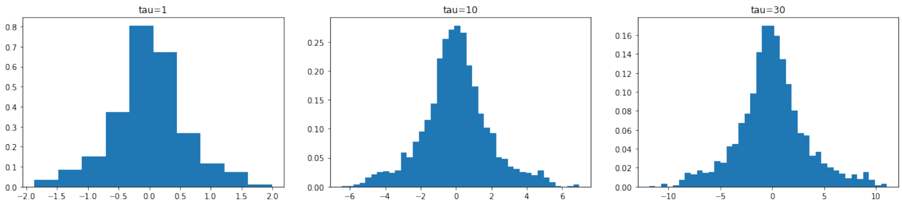

Figure 11: $\boldsymbol{\rho_\tau}(\lambda)$ based on $\boldsymbol{S}^{0\text{sim}}$ using Ahdida and Alfonsi (2013) (Eq. 10) with $\tau = 1$, 10 or 30 days. The histograms corresponds to the simulated distribution of the eigenvalues of the increments of the correlation matrix. $\tau = 1$, 10 or 30 days corresponds to the time horizon of the increment. The simulation, that was obtained with $(k, a) = (1., 0.9)$ and with 100 iterations (paths).

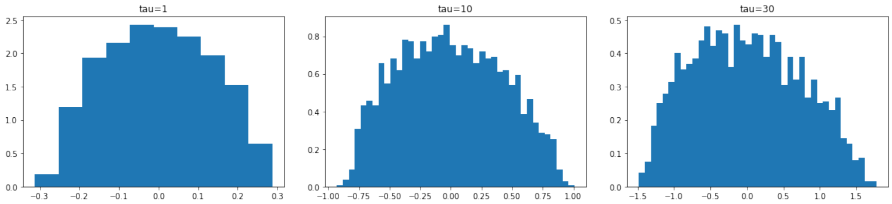

Figure 12: $\boldsymbol{\rho_\tau}(\lambda)$ based on $\boldsymbol{S}^{2\text{sim}}$ using Ahdida and Alfonsi (2013) (normalized version of Eq. 10) with $\tau = 1$, 10 or 30 days. The histograms correspond to the simulated distribution of the eigenvalues of the increments of the correlation matrix. $\tau = 1$, 10 or 30 days correspond to the time horizon of the increment. The simulation was obtained with $(k, a) = (1., 0.9)$ and with 100 iterations (paths).

Finally we model the FCL dynamics by the following OrnsteinâĂŞUhlenbeck process:

$$dx_a(t) = -\alpha x_a(t)dt + \sigma dB_a(t)\,. \tag{17}$$

Here $B_a(t)$ are independent Wiener processes and the parameters α and σ are determined from the best match to the FCL variograms. The time variation is then mimicked by

$$\boldsymbol{C}_{\text{sim}}(t) = \begin{pmatrix} e^{x_1(t)} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & e^{x_{K-1}(t)} \end{pmatrix} \boldsymbol{C}_0^{\mathbf{r.}} \begin{pmatrix} e^{x_1(t)} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & e^{x_{K-1}(t)} \end{pmatrix} \tag{18}$$

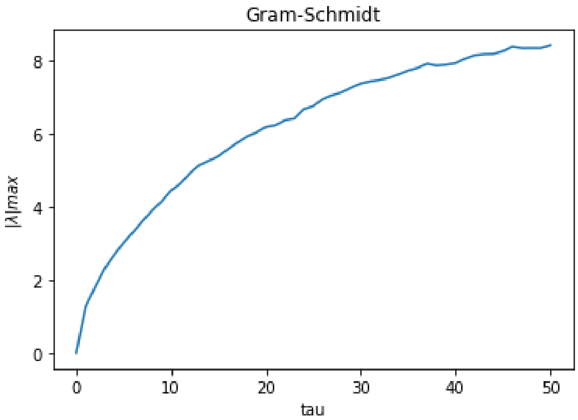

Figure 13: $|\boldsymbol{\lambda}|_{\max}(\tau)$ based on $\boldsymbol{S}^{0\mathrm{sim}}$. It corresponds to the largest eigenvalue in absolute value of the increment of the proxy of the correlation matrix between single stocks depending on τ, the time horizon of the increment. $\boldsymbol{S}^{0\mathrm{sim}}$ is derived by a Gram-Schmidt algorithm for $\epsilon = 0.015$

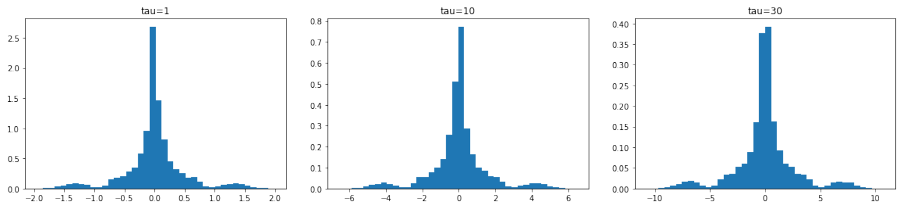

Figure 14: $\boldsymbol{\rho_\tau}(\lambda)$ based on $\boldsymbol{S}^{0\mathrm{sim}}$ using a Gram-Schmidt algorithm with $\tau = 1$, 10 or 30 days. The histograms correspond to the simulated distribution of the eigenvalues of the increments of the correlation matrix. $\tau = 1$, 10 or 30 days correspond to the time horizon of the increment. The simulation was obtained for $\epsilon = 0.015$

C Random Matrices from the literature

To measure the “distance" between two matrices we will use the normalized version of the *Hilbert-Schmidt operator* also known as the *Frobenius inner product*:

$$\mathbf{d}(A, B) = \frac{1}{2N}\langle A, B\rangle_{\mathrm{HS}} \equiv \frac{1}{2N}\mathrm{Tr}\left(A^\dagger B\right) . \tag{19}$$

If A is real and orthonormal ($AA^T = A^T A = \mathbb{1}_{N\times N}$), while $B = \mathbb{1}_{N\times N}$, then the above reduces to:

$$\mathbf{d}(A, \mathbb{1}_{N\times N}) = \left(1 - \frac{\mathrm{Tr}A}{N}\right) . \tag{20}$$

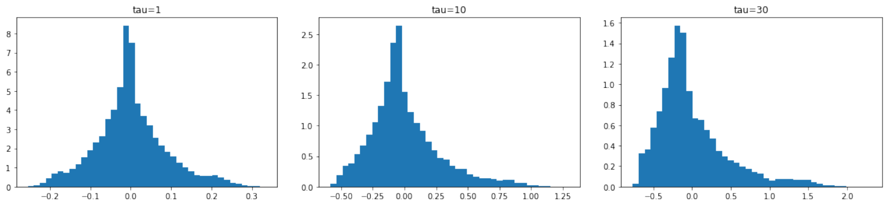

Figure 15: $\boldsymbol{\rho_\tau}(\lambda)$ based on $\boldsymbol{S}^{2\text{sim}}$ using a Gram-Schmidt algorithm with $\tau = 1$, 10 or 30 days. The histograms correspond to the simulated distribution of the eigenvalues of the increments of the correlation matrix. $\tau = 1$, 10 or 30 days correspond to the time horizon of the increment. The simulation was obtained for $\epsilon = 0.015$

The distance is, therefore, equal to 0 if and only if $A = \nVdash_{N\times N}$. Similarly the maximal possible distance $\mathbf{d}(A, \nVdash_{N\times N} = 2)$ requires $A = -\nVdash_{N\times N}$ (notice that it is possible only for even N).

C.1 (Ornstein-Uhlenbeck) random walk on $SO(N)$

Below we list various options to generate a random walk on $SO(N)$ that will not depart "too far" from the identity matrix (a random work around an arbitrary orthogonal matrix is a trivial generalisation).

C.1.1 Gram-Schmidt based algorithm

Let us denote by $GS(V)$ the Gram-Schmidt orthogonalisation (and normalisation) procedure that acts on the rows of a (square) matrix V. We require that the algorithm doesn't mix different rows and columns, ans so $GS(V + \delta V)$ is close to V if V is orthonormal and δV is sufficiently small.

We can then generate the aforementioned walk with:

$$O(t+1) = GS\left(O(t) + \mu \cdot \mathbf{I} + \epsilon \cdot W\right), \tag{21}$$

where μ is the drift parameter, W is a random $N \times N$ matrix and ϵ is the parameter controlling the random walk around the identity matrix $\mathbf{I}$. Figure 22 presents the simulation output for $N = 50$ and parameters listed in the caption.

The result seems to be satisfactory, but it comes with a performance cost. A "hand-written" Python code runs 0.2 seconds for a single (!) 500×500 matrix and is therefore impractical for $T = 70000$. At the same time the already existing `numpy` routine for the GS algorithm is ten times faster but unfortunately it mixes the rows of the matrix and so cannot be used for the numerical simulation.

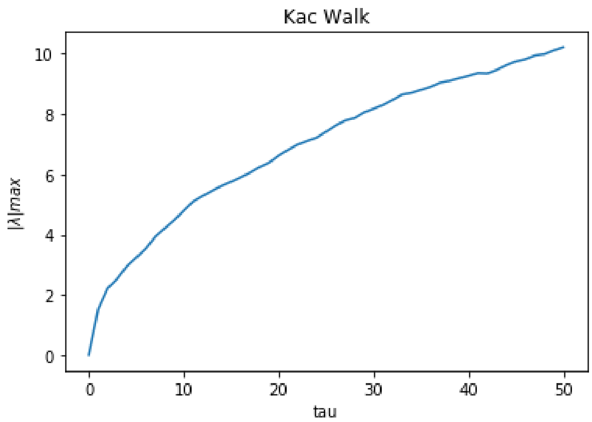

Figure 16: $|\boldsymbol{\lambda}|_{\max}(\tau)$ based on $\boldsymbol{S}^{0\text{sim}}$. It corresponds to the largest eigenvalue in absolute value of the increment of the proxy of the correlation matrix between single stocks depending on τ, the time horizon of the increment. $\boldsymbol{S}^{0\text{sim}}$ is derived the Kac Walk for $\theta = 0.07$ and $n = 15$ random 2×2 rotations per $\delta t = 1$.

C.2 Kac walk based approaches

To reduce the running time we have to avoid using costly matrix operations (like, for instance, matrix products) since they have N^2 complexity, and instead operating on selected rows/columns of $\mathbf{O}(t)$. For example, the diffusion can be modelled by a sequence of rotations

$$\begin{pmatrix} O(t+1)_i \\ O(t+1)_j \end{pmatrix} = \begin{pmatrix} c & -s \\ s & c \end{pmatrix} \begin{pmatrix} O(t)_i \\ O(t)_j \end{pmatrix}, \tag{22}$$

where $c, s = \cos(\delta), \sin(\delta)$ for a small angle δ, and the pair of rows i, j is selected randomly. In terms of the $\mathbf{O}(t)$'s columns this is just the Kac walk often used to model diffusion on a sphere.

Repeating these so-called Givens rotations leads to a random walk on $SO(N)$ though definitely with no "mean-reversion". To introduce the drift we have somehow to rotate $\mathbf{O}(t)$ back each time in order to bring it closer to the identity matrix.

We tried two different approaches:

1. Identify a pair of indices i, j for which the value $|O(t)_{i,j} - O(t)_{j,i}|$ is maximal. This corresponds to a plane where $\mathbf{O}(t)$ deviates the most from the identity matrix. Rotate then in this plane by a *fixed portion* of the angle needed to bring this part of $\mathbf{O}(t)$ maximally close to the identity matrix:

$$\boldsymbol{\gamma} \cdot \arcsin\left(\frac{O_{j,i} - O_{i,j}}{\sqrt{(O_{j,i} - O_{i,j})^2 + (O_{i,i} + O_{j,j})^2}}\right) \tag{23}$$

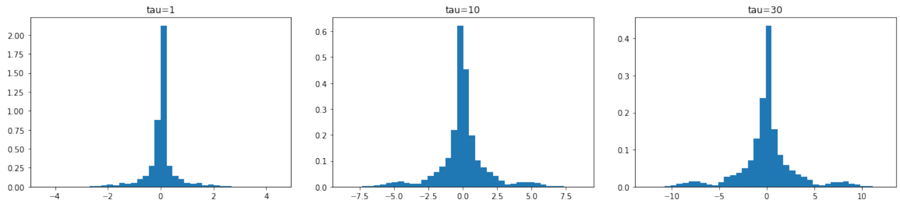

Figure 17: $\boldsymbol{\rho_\tau}(\lambda)$ based on $\boldsymbol{S}^{0\text{sim}}$ using Kac walk with $\tau = 1$, 10 or 30 days. The histograms correspond to the simulated distribution of the eigenvalues of the increments of the correlation matrix. $\tau = 1$, 10 or 30 days correspond to the time horizon of the increment. The simulation was obtained for with $\theta = 0.07$, $n = 15$ random 2×2 rotations per $\delta t = 1$ and with 1000 iterations (paths) with Kac walk

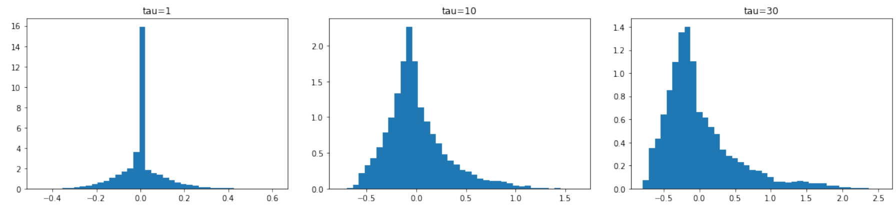

Figure 18: $\boldsymbol{\rho_\tau}(\lambda)$ based on $\boldsymbol{S}^{2\text{sim}}$ using Kac walk with $\tau = 1$, 10 or 30 days. The histograms correspond to the simulated distribution of the eigenvalues of the increments of the correlation matrix. $\tau = 1$, 10 or 30 days correspond to the time horizon of the increment. The simulation was obtained for with $\theta = 0.07$, $n = 15$ random 2×2 rotations per $\delta t = 1$ and with 1000 iterations (paths) with Kac walk

2. Proceed the same as above but select i for which the value $|O(t)_{i,i} - 1|$ is maximal and take j with the largest $|O(t)_{i,j} - O(t)_{j,i}|$.

The two algorithms have overall four different parameters:

- The number of random consecutive Givens rotations, n_{RW}.
- The constant angle used for these rotations, δ.
- The number of the consecutive "reversions" applied after Givens rotations, n_{Rev}.
- The parameter $\boldsymbol{\gamma}$ needed to control the reversion/drift.

The first approach performs slightly better, but the search for the pair (i, j) has an N^2 cost, while the latter has only N-order complexity. Figure 23 demonstrates the $N = 50$ implementation of the second algorithm with parameters described in the caption.

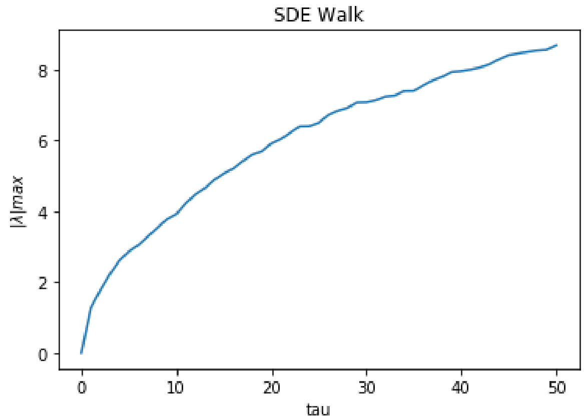

Figure 19: $|\boldsymbol{\lambda}|_{\max}(\tau)$ based on $\boldsymbol{S}^{0\text{sim}}$. It corresponds to the largest eigenvalue in absolute value of the increment of the proxy of the correlation matrix between single stocks depending on τ, the time horizon of the increment. $\boldsymbol{S}^{0\text{sim}}$ is derived the new Stochastic Differential Equation for $\mu = 1.$ and $\sigma = 0.2$

References

Ahdida,A., A. Alfonsi. "A mean-reverting SDE on correlation matrices." Stochastic process and their application 2013

R. Allez and J.-P. Bouchaud, "Eigenvector dynamics: General theory and some applications",Phys. Rev. E 86, 046202 (2012).

Bru, M.F.. "Wishart processes". J. Theoret. Probab., 4(4):725âĂŞ751, 1991.

Cox,J. , J. Ingersoll, and S. Ross. "A theory of the term structure of interest rates". Econometrica, 53:385âĂŞ407, 1985

Cuchiero,C., D. Filipovi , E. Mayerhofer, and J. Teichmann. "Affine processes on positive semidefinite matrices". Ann. Appl. Probab., 21(2):397âĂŞ463, 2011.

Fonseca,J. Da, M. Grasselli, and C. Tebaldi. "Option pricing when correlations are stochastic: an analytical framework". Springer, 2008.

Grebenkov, D., J. Serror "Following a trend with an exponential moving average: Analytical results for a Gaussian model" Physica A 2014

Gourieroux, C., "Continuous time Wishart process for stochastic risk". Econometric Reviews, 25:2:177 âĂŞ 217, 2007.

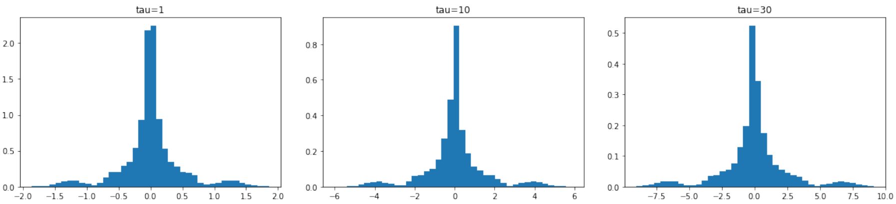

Figure 20: $\boldsymbol{\rho_\tau}(\lambda)$ based on $\boldsymbol{S}^{0\mathrm{sim}}$ using the new Stochastic Differential Equation with $\tau = 1$, 10 or 30 days. The histograms correspond to the simulated distribution of the eigenvalues of the increments of the correlation matrix. $\tau = 1$, 10 or 30 days correspond to the time horizon of the increment. The simulation was obtained for $\mu = 1$, $\sigma = 0.2$, and with 1000 iterations (paths) with Stochastic Differential Equation.

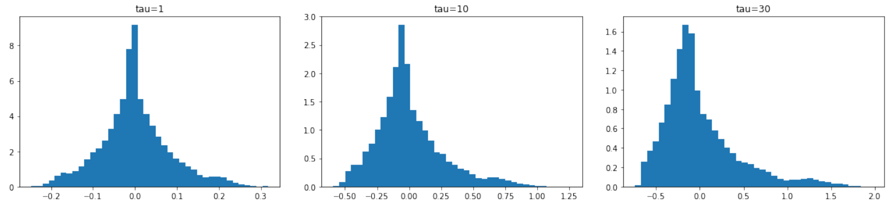

Figure 21: $\boldsymbol{\rho_\tau}(\lambda)$ based on $\boldsymbol{S}^{2\mathrm{sim}}$ using the new Stochastic Differential Equation with $\tau = 1$, 10 or 30 days. The histograms correspond to the simulated distribution of the eigenvalues of the increments of the correlation matrix. $\tau = 1$, 10 or 30 days correspond to the time horizon of the increment. The simulation was obtained for $\mu = 1$, $\sigma = 0.2$, and with 1000 iterations (paths) with Stochastic Differential Equation.

Heston, Steven L. (1993). “A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options”. The Review of Financial Studies. 6 (2): 327âĂŞ343. doi:10.1093/rfs/6.2.327. JSTOR 2962057.

Kac, M. Probability and related topics in physical sciences, Book American Mathematical Soc. 1959

Valeyre, S., S. Kuperstein, D. S. Grebenkov, S. Aboura “The market neutral fundamental maximum variance portfolios” Working Paper 2018.

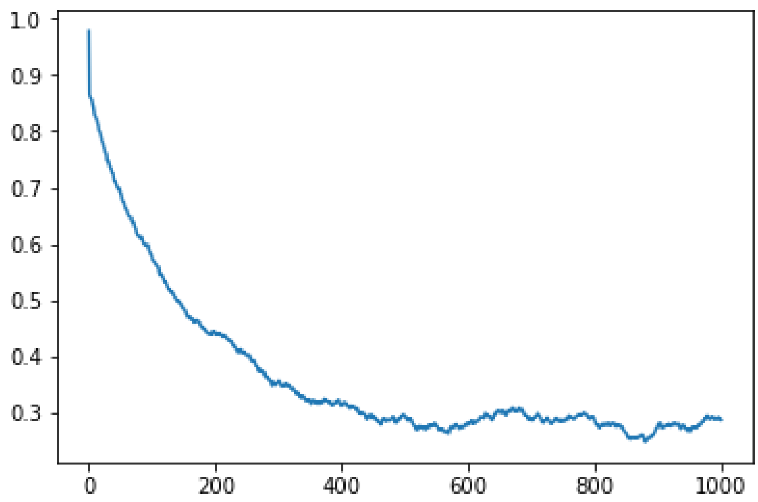

Figure 22: The OU walk defined by (21) for $N = 50$ and $(\mu, \epsilon) = (0.01, 0.015)$. The horizontal axis is t and the vertical axis corresponds to the distance between $\mathbf{O}(t)$ and the identity matrix as defined by (20). The matrix $\mathbf{O}(0)$ is a random orthogonal matrix, so the distance is close to 1. The process starts to oscillate around $\mathbf{I}_d$ at $t = 400$. Notice that $\mathbf{d} = 0.3$ is in fact a very small distance for $N = 50$.

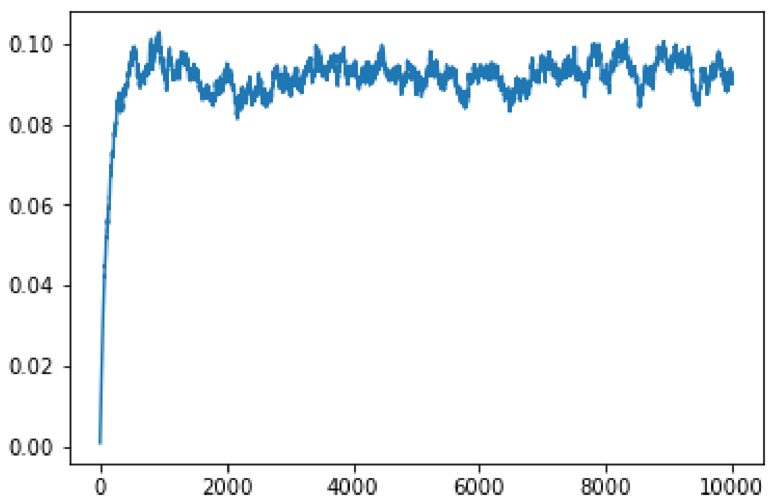

Figure 23: The horizontal axis is t and the vertical axis corresponds to the distance between $\mathbf{O}(t)$ and the identity matrix as defined by (20). The parameters are $(n_{\mathrm{RW}}, \delta, n_{\mathrm{Rev}}, \boldsymbol{\gamma}) = \left(10, 0.05 \cdot \frac{\pi}{2}, 10, 0.2\right)$. In this simulation $\mathbf{O}(0) = \mathbf{I}$.

6. Should Employers Pay Better their Employees? An Asset Pricing Approach

Should employers pay their employees better? An asset pricing approach

ABSTRACT

We uncover a new anomaly in asset pricing that is linked to the remuneration: the more a company spends on salaries and benefits per employee, the better its stock performs, on average. Moreover, the companies adopting similar remuneration policies share a common risk, which is comparable to that of the value premium. For this purpose, we set up an original methodology that uses firm financial characteristics to build factors that are less correlated than in the standard asset pricing methodology. We quantify the importance of these factors from an asset pricing perspective by introducing the factor correlation level as a directly accessible proxy of eigenvalues of the correlation matrix. A rational explanation of the remuneration anomaly involves the positive correlation between pay and employee performance.

JEL classification: G12, G32, J30, C4
Keywords: Anomalies, Asset Pricing, Remuneration, Performance, Factor Correlation.

"The wages of labour are the encouragement of industry, which like every other human quality, improves in proportion to the encouragement it receives. Where wages are high, accordingly, we shall always find the workmen more active, diligent, and expeditious, than where they are low." *Adam Smith (1776).*

I. Introduction

Should employers pay their employees better? Although this question might appear provoking because lowering production costs remains a cornerstone of the contemporary economy, we present the first attempt to report the real effects of employee remuneration on asset pricing. Remuneration – defined as the annual salaries and benefits expenses (e.g., wages, bonuses, pension expenses, health insurance payment, etc.) per employee – is the basis of any employment contract. For instance, pay was shown to explain, on average, 65% of the variance in evaluations of overall job attractiveness (Rynes *et al.*, 1983). Classical theory states that profit-maximizing firms choose the level of labor pay by setting the marginal cost of labor (i.e., the wage rate) equal to the marginal revenue product of labor (i.e., the marginal benefit). Beyond this paradigm, we provide strong evidence that firms that pay their employees better tend to over-perform on the stock market.

Our objective is to examine whether remuneration is an anomaly that can be priced in asset pricing models. Schwert (2003) defines anomalies as "empirical results that seem to be inconsistent with maintained theories of asset-pricing behavior (the CAPM). They indicate either market inefficiency (profit opportunities) or inadequacies in the asset-pricing model. After they are documented and analyzed in the academic literature, anomalies often seem to disappear, reverse, or attenuate." Anomalies are typically identified either by regressing a cross-section of average returns (e.g., the seminal Fama and MacBeth (1973) approach uses the capitalization and book-to-market values), or by using a panel regression of the cross-section of returns with different factor returns through the F-Statistic (Gibbons *et al.*, 1989), or by using a portfolio-based approach that segregates individual stocks with similar capitalization and book-to-market values into different style portfolios (Fama and French, 1993). In the latter case (which we refer to as the "FF approach"), the factors formed on small minus big market capitalization portfolios (SMB) and high minus low book-to-market portfolios (HML) explain an important part of the identified anomalies (Fama and French, 1996). Over recent decades, the growing number of discovered anomalies suggests that the standard asset pricing models fail to explain much of the cross-sectional variation in average stock returns. Meanwhile, the effect of remuneration on company performance has surprisingly never been tested, despite the fact that employers pay particular attention to labor costs in attempting to maximize profits.

This research contributes empirically to the asset pricing literature by introducing an observable firm characteristic, namely the remuneration, as a candidate anomaly. More precisely, we focus on remuneration as a priced factor. Indeed, it remains unclear how far remuneration can explain the cross-section of returns despite a sizeable literature on labor economics that relates labor to asset pricing. This branch of literature has intensively investigated the impact of labor decisions on the firm's value, notably through the operating leverage, which affects the equity returns riskiness. However, to our best knowledge, there are no asset pricing studies that incorporate employee's wages as a pricing factor. Besides, based on the impressive list of anomalies analyzed by Harvey *et al.* (2016), we find only one paper that highlights income as a potential factor. Indeed, Gomez *et al.* (2015) analyze the relation between U.S. census division-level labor income and the cross-section of returns using the standard Fama and French (1993) approach. More specifically, these authors use per capita personal income (from the Bureau of Economic Analysis) as a new candidate factor and conclude that the cross-section of stock returns depends on the census district in which the headquarters of the firm are located. Unfortunately, as Harvey *et al.* (2016) has noted, "most of the division level labor income have a non-significant t-statistic. *We do not count their factors*". Moreover, we use remuneration at the company level to generate results that are more realistic from an asset pricing perspective, which contrasts with Gomez *et al.* (2015), whose scope is limited to income per state and per division.

This research contributes also theoretically to the asset pricing literature by introducing a new methodology to build factors that is conceptually close to principal component analysis (PCA) but goes beyond its noise-induced limitations. This methodology presents many advantages compared with the conventional multi-factor approach developed by Fama and French (1992, 1993). We propose a new measure of "explanatory power" of factors where the relevance of the factor does not depend on the number of considered factors, in contrast to the R-squared argument of the FF setting. Hence, we introduce the Factor Correlation Level (FCL) as a metrics of common risks that measures the ability of stocks within the factor to fluctuate in a common way. Importantly, it allows ordering the factors according to their capacity of taking into account the variability of stocks, and therefore to their importance from an asset pricing perspective. In this respect, our ranking by the FCL indicator resembles principal component analysis. At the same time, this indicator is also linked to the R-squared value of the factor in the asset pricing model: higher FCLs correspond to higher R-squared values in the asset pricing model with one factor. The empirical validation of the FCL methodology is founded on an exhaustive testing protocol. First, we use ten factors that summarize most of the existing factors: dividend, capitalization, liquidity, momentum, low-volatility, debt-to-book, sales-to-market, book-to-market, cash and, of course, the remuneration factor; those which are not present in this list remain correlated with some of these factors; we check that performance associated with the remuneration factor is not explained by other major factors such as low-volatility, capitalization, book-to-market, or momentum. Second, we consider six "supersectors" that are used to split stocks into comparable groups since remuneration varies strongly from one sector to another. Third, we employ a large data set of 3612 daily single stock close prices from January 2001 to July 2015 for the 569 biggest companies in Europe. For comparison, we also treat the same number of randomly selected companies in the U.S.A. whose capitalization exceeds 1 billion of dollars. Although we do not access the remuneration data for these companies, the analysis of other factors allows us to validate the FCL methodology on the U.S. market (often considered as a benchmark) and to compare our predictions to whose of the FF approach. Fourth, we perform several robustness checks to examine if the results change with the tested variations; for instance, we perform a separate analysis with the 258 biggest companies from U.K. to check for potential domestic biases; we also run the methodology on monthly data to check the role of time scale; in the spirit of comparability, we evaluate the factor performances with seven incremental transitions from the standard FF approach to our methodology. Finally, we compare our results with the basic PCA and illustrate its limitations. Our main result indicates that a market neutral investment strategy based on the remuneration anomaly would likely deliver positive annual returns of 2.42% above the market.

The remainder of the paper is organized as follows. Section II offers a literature review that covers several fields of research. Section III describes the novel methodology. Section IV presents the data, whereas Section V presents the empirical results. Section VI discusses the advantages and limitations of our methodology and compares it with the FF approach. Section VII summarizes the main findings and concludes.

II. Literature review

A. The asset pricing

This article is mainly related to the asset pricing literature in which previous studies have shown that the average returns of common stocks are related to firm characteristics such as capitalization, price-earnings ratio, cash flow, book-to-market, past sales growth and past returns. For example, stocks with lower market capitalization tend to have higher average returns (Banz, 1981). Another important anomaly is the value premium: value stocks have higher returns than growth stocks, which is likely because the market undervalues distressed stocks (Fama and French, 1998). More precisely, small stocks and value stocks have higher average returns than their betas can explain (Campbell and Vuolteenaho, 2004). Profitability and investment also add to the description of average returns (Fama and French, 2015). The low volatility anomaly was revealed for medium and big stocks in addition to growth stocks (Jordan and Riley, 2013). Those stocks that are expected to have high idiosyncratic risk earn high returns in the cross-section (Fu, 2009). This result contradicts previous findings made by Ang *et al.* (2006), who posit that stocks with high idiosyncratic volatility have low average returns. Macroeconomic risk has also been connected with the cross-section of returns. For instance, the growth rate of industrial production is seen as a priced risk factor in standard asset pricing tests (Chen *et al.*, 1986; Liu and Zhang, 2008). There is a size effect in bank stock returns that differs from the market capitalization effects documented in non-financial stock returns (Gandhi and Lustig, 2015). The most popular anomaly is momentum: stocks with low past returns tend to have low future returns while stocks with high past returns tend to have high future returns (Jegadeesh and Titman, 1993). Hence, the momentum strategy that buys past winners and sells past losers should earn abnormal returns in upcoming years. Return momentum has also been observed when spreads in average momentum returns decrease from smaller to bigger stocks (Fama and French, 2012). However, momentum strategies seem to produce losses specifically in January (Jegadeesh and Titman, 1993), probably based on taxation effects (Grinblatt and Moskowitz, 2004). Similarly, changes in book equity appear to be more informative about expected stock returns than price returns (Bali *et al.*, 2013). Notably, certain stock market anomalies may appear and then disappear after publication in academic journals (McLean and Pontiff, 2015). In spite of the abundant literature, the work by Gomez *et al.* (2015) seems to be the sole article that considers income as a candidate anomaly although it is still not an income per employee but rather per state and per division. Several models have been developed to provide economic interpretations of numerous stylized anomalies and to improve the performance of the CAPM.[1] Simultaneously, the anomaly-based evidence against the CAPM has been questioned because anomalies have primarily been confined to small stocks (Cederburg *et al.*, 2015).[2]

[1] Campbell and Vuolteenaho (2004) introduced a two-beta model to explain the capitalization and book-to-market value anomalies in stock returns by splitting the CAPM into a cash-flow beta with a higher price of risk than a discount-rate beta. Fama and French (1993) proposed a three-factor model to capture the patterns in U.S. average returns associated with capitalization and value-versus-growth. Even after a theoretical rationale for the three-factor model was provided by Ferguson and Shockley (2003), many anomalies remain unexplained by the three-factor model (Fama and French, 2015). Although a four-factor model has been derived (Carhart, 1997), it has also failed to absorb all the momentum in U.S. average stock returns (Avramov and Chordia, 2006). Recently, a five-factor model was introduced to capture capitalization, value, profitability, and investment patterns in average stock returns and is reputed to perform better than the three-factor model (Fama and French, 2015).

[2] In line with this criticism, doubt was cast on the set of anomalies to consider in a multi-factorial setup, given that Harvey *et al.* (2016) have summarized 316 potential factors by reviewing 313 papers published

B. Corporate finance

This article is also related to the extensive literature on corporate finance, which has also continued to investigate the relation between remuneration and performance, although it has usually focused on managerial pay as opposed to the broader category of employees that we consider in the present study. This branch of literature typically examines the wage as a managerial incentive likely to reduce agency costs by designing an optimal job contract. In that sense, we may consider that solving the incentive problem leads to shareholder value creation affecting stock returns. Indeed, managers face both discipline and opportunities provided by the free market economy that leads to the notion that there is no need for explicit contracts to resolve incentive problems (Fama, 1980). Nevertheless, market forces cannot act as a complete substitute for contracts (Holmstrom, 1999) because career concerns must be considered to design optimal contracts and to arrive at strong incentives (Gibbons and Murphy, 1992). The effects of incentives depend on how they are designed (Gneezy *et al.*, 2011), given that managers have considerable power to shape their own pay arrangements – and perhaps to even hurting shareholder interest (Bebchuk *et al.*, 2002). Indeed, public company disclosures do not provide a comprehensive measure of managerial incentive to increase shareholder value (O'Byrne and Young, 2010). Many explanations were brought forward to justify top managers' remuneration. Firms with abundant investment opportunities pay their executives better (Gaver and Gaver, 1995). The increase in the level of stock-option compensation can be explained by the inability of boards to evaluate its real costs (Hall and Murphy, 2003; Jensen *et al.*, 2004). The capitalization of large firms explains many patterns in top manager pay across firms, over time, and between countries (Gabaix and Landier, 2008). Manager fixed effects, interpreted as unobserved managerial attributes and understood as a proxy for latent managerial ability, are important in explaining the level of executive remuneration (Graham *et al.*, 2012). Overall, remuneration matters because it may affect a corporation's level of risk as bonus-driven remuneration might encourage excessive risk-taking. However, pay and risk are correlated not because mis-aligned pay drives risk-taking, but rather because principal agent theory predicts that riskier but more profitable firms must pay more remuneration than less risky firms to provide a risk-averse manager the same incentives (Cheng *et al.*, 2015).

since 1967. In the same vein, 38 out of 80 potential firm-level anomalies were shown to be insignificant in the broad cross-section of average stock returns (Hou *et al.*, 2015). In addition, mistakes can easily be made in this field due to multiple testing or data mining methods. As noted by Harvey and Liu (2015), many discovered factors are likely to be false if their t-statistics do not exceed 3. Finally, these papers suggest that many claims in the anomalies literature are likely to be exaggerated regarding the associated t-statistics.

C. Labor economics

The labor economics literature treats this question through the "efficiency wage theory" by relating it to unemployment. Yellen (1984) and Akerlof and Yellen (1990) did a remarkable work with an analysis that is built – unlike most economic models – mainly on sociology and psychology with experimentation that delivers salient stylized facts on human behavior in a working context. Efficiency wage theory maintains that rising wages is the best way to increase output per employee because it links pecuniary incentives to employee performance. In particular, the use of performance pay packages by employers has been shown to increase employee productivity (Lazear, 2000) and job satisfaction (Green and Heywood, 2008). There are several interesting studies that relate labor market to asset pricing. All these empirical results emphasize the significant impact of labor decisions, in which wage plays a prominent role, onto firm's value. Santos and Veronesi (2006) show that labor income to consumption ratio is a strong predictor of long horizon returns. Danthine and Donaldson (2002) explain that operating leverage is more significant for the riskiness of equity returns than financial leverage. In other words, attention should be paid to wages, particularly because the priority nature of wages enhances the risk of dividends. In this spirit, Kuehn *et al.* (2013) note that a high value of unemployment makes wages inelastic, which gives rise to operating leverage. The impact of inelastic wages is even stronger in bad times as it amplifies the equity risk premium. Gourio (2007) argues that because wages are smooth, revenues are more cyclic than costs, making the profits more volatile. In particular, firms with high book-to-market or with low productivity, i.e. value firms, have more pro-cyclic earnings. Ochoa (2013) finds a positive and statistically significant relation between the reliance on skilled labor and expected returns. In times of high volatility, firms with a high share of skilled workers earn an annual return of 2.7% above those with a high share of unskilled workers notably because their labor is more costly to adjust. Labor decisions made by workers can affect firm risk (Donangelo, 2014) while hiring decisions can also be the determinants of firm risk (Carlson *et al.*, 2004; Belo *et al.*, 2014). Indeed, Donangelo (2014) discusses the idea that mobile workers carry some of the firm's capital productivity when they leave an industry. He finds that portfolios that hold long positions in stocks of high-mobility industries (general workers) and short positions in stocks of low-mobility industries (industry-specific workers) earn an annual return spread of over 5%. Like Monika and Yashiv (2007) who explain that labor should matter since firms' market value embodies the value of hiring, Belo *et al.* (2014) argue that the market value of a firm reflects the value of its labor force because the firm can extract rents as compensation for the costs associated with adjusting its labor force. They find that long positions in stocks of low-hiring firms and short positions in high-hiring firms earn an average annual excess stock return of 5.6%. Favilukis and Xiaoji (2016) introduce infrequent renegotiation in standard wages model showing that it leads to smooth average wages. Due to this wage rigidity, they find that wage growth forecasts long-horizon excess equity returns.

D. Social sciences

This article is also broadly related to several streams of research in various social sciences, including sociology, psychology and human resources. In these fields, wage acts like a motivator since it typically reflects a social preference for rewards likely to affect the employee's performance. Sociological studies have developed a theory of social exchange in which there are equivalent rewards on both sides (Blau, 1955), which is consistent with the preference for reciprocity that is viewed as a social preference, as it depends on the behavior of the reference person (Fehr and Falk, 2002). Reciprocity induces agents to cooperate voluntarily with the principal when the principal treats them correctly; the evidence for reciprocity is based on a so-called gift exchange experiment.

Psychological studies highlight the exchange in working situations in which the perceived value of labor equals the perceived value of remuneration, based on the theory of equity (Adams, 1963). When there is no mismatch between effort and wages, employees may change their perceived effort and even their perceived level of remuneration by redefining the non-pecuniary component.

Human resources studies generally offer evidence that money is an important motivator for most people (Rynes *et al.*, 2004), as pay can help climbing on the Maslow's motivational hierarchy of needs, including social esteem and self-actualization. Nevertheless, tangible rewards might also produce secondary negative effects on motivation (Baker, 1992) by forestalling self-regulation (Deci *et al.*, 1999).

III. Methodology

In this section, we introduce a new methodology to build factors that combines advantages of the PCA and the Fama and French (1993) approach. As would be the case with the PCA, our factors are built to be uncorrelated with the market index and with sectorial factors. For each factor, **we introduce and estimate the Factor Correlation Level (FCL) that allows us to order the factors based on their importance and to select the most important ones in asset pricing models.**

A. Conventional diagonalization of the covariance and correlation matrices

Identifying common risks of multiple assets is necessary to diversify investments and can help to profit from style's arbitrage opportunities. Conventional approaches, such as PCA, attempt to diagonalize the empirical covariance (or correlation) matrix of the traded universe, i.e., to decorrelate assets by constructing independent linear combinations (portfolios) of assets. Each eigenvector of the covariance matrix represents the coefficients of one such combination while the corresponding eigenvalue gives its variance. If the covariance matrix does not contain negative elements (i.e., if there are no negatively correlated assets), the eigenvector corresponding to the largest eigenvalue has positive elements that can be interpreted as relative weights of stocks in the market mode. The classical long portfolio, following the market, can be constructed by investing in proportion to these weights. In turn, market neutral portfolios should be orthogonal to the market mode and therefore have both long and short positions (the latter corresponding to negative weights). The other eigenvectors capture different common risks of the traded universe, and the most common include sectorial risks (e.g., banking sector, commodities, energy, etc.).

In mathematical terms, if the covariance matrix Ω of stocks was known precisely, it might be diagonalized to identify uncorrelated linear combinations of stocks and their variances to assess the related risks. For a traded universe with n stocks, let $r_1, \ldots, r_n$ denote the daily returns of these stocks at a given time. The covariance matrix has n eigenvalues $\lambda_1, \ldots, \lambda_n$ and n eigenvectors $V_1, \ldots, V_n$ satisfying $\Omega V_\alpha = \lambda_\alpha V_\alpha$ (for each $\alpha = 1, \ldots, n$). Each eigenvector V_α determines one linear combination of stocks, $(V_\alpha)_1 r_1 + \ldots + (V_\alpha)_n r_n$, which is decorrelated from the others, while the eigenvalue λ_α is its variance (under the condition that V_α is appropriately normalized).

The above eigenbasis can be interpreted as follows. For any linear combination of stocks with weights w_i, $r_\pi = w_1 r_1 + \ldots + w_n r_n = (w \cdot r)$ (written as a scalar product), the variance of such a portfolio π can be expressed as

$$\langle {r_\pi}^2 \rangle = \langle \left(\sum_{i=1}^{n} w_i r_i \right)^2 \rangle = \sum_{i,j=1}^{n} w_i w_j \Omega_{i,j} = \sum_{i,j=1}^{n} w_i w_j \sum_{\alpha=1}^{n} \lambda_\alpha (V_\alpha)_i (V_\alpha)_j = \sum_{\alpha=1}^{n} \lambda_\alpha (w \cdot V_\alpha)^2, \quad (1)$$

where $\langle \ldots \rangle$ denotes the expectation, and the returns r_k were assumed to be centered. In other words, the variance is decomposed into a sum of variances λ_α of independent linear combinations proportional to the projection of the weights w_i onto the corresponding eigenvector V_α. If the weights w_i are chosen in proportion to the elements of one eigenvector, i.e.,

$w_i = c(V_\alpha)_i$ for some α and c, then the orthogonality of V_α to other eigenvectors yields

$$\langle {r_\pi}^2 \rangle = \lambda_\alpha c^2 (V_\alpha \cdot V_\alpha)^2 = \lambda_\alpha \; (w \cdot w), \tag{2}$$

where we used the L^2-normalization of the eigenvectors: $(V_\alpha \cdot V_\alpha) = 1$. As expected, the variance of such a linear combination is fully determined by the corresponding eigenvalue λ_α. Notably, the above relation can also be written as

$$\lambda_\alpha = \frac{\langle {r_\pi}^2 \rangle}{\sum_{i=1}^n w_i^2} \tag{3}$$

to estimate the variance of the linear combination whose weights are constructed close to an eigenvector.

As different stocks exhibit quite distinct volatilities, it is convenient to rescale the stock's return r_i by its realized volatility σ_i: $\tilde{r}_i = r_i/\sigma_i$. This rescaling is also known to reduce heterogeneity of volatilities among stocks and heteroskedasticity (Andersen *et al.*, 2000; Bouchaud *et al.*, 2001; Valeyre *et al.*, 2013). In other words, one can write

$$\langle {r_\pi}^2 \rangle = \langle \left(\sum_{i=1}^n w_i \sigma_i \tilde{r}_i \right)^2 \rangle = \sum_{i,j=1}^n \tilde{w}_i \tilde{w}_j C_{i,j}, \tag{4}$$

where $\tilde{w}_i = w_i \sigma_i$ and $C = \langle \tilde{r}_i \tilde{r}_j \rangle$ is the covariance matrix of the renormalized returns $\tilde{r}_i$ or, equivalently, the correlation matrix of returns r_i: $\Omega_{i,j} = \sigma_i \sigma_j C_{i,j}$. To proceed, the eigenvalues and eigenvectors of Ω can be replaced by the eigenvalues $\tilde{\lambda}_\alpha$ and eigenvectors $\tilde{V}_\alpha$ of the correlation matrix C, $C\tilde{V}_\alpha = \tilde{\lambda}_\alpha \tilde{V}_\alpha$, i.e.,

$$\langle {r_\pi}^2 \rangle = \sum_{i,j=1}^n \tilde{w}_i \tilde{w}_j \sum_{\alpha=1}^n \tilde{\lambda}_\alpha (\tilde{V}_\alpha)_i (\tilde{V}_\alpha)_j = \sum_{\alpha=1}^n \tilde{\lambda}_\alpha (\tilde{w} \cdot \tilde{V}_\alpha)^2. \tag{5}$$

If the volatility-normalized weights $\tilde{w}_i$ are chosen to be proportional to the elements of an eigenvector, $\tilde{w}_i = c(\tilde{V}_\alpha)_i$, one obtains $\langle {r_\pi}^2 \rangle = \tilde{\lambda}_\alpha c^2 (\tilde{V}_\alpha \cdot \tilde{V}_\alpha) = \tilde{\lambda}_\alpha c^2 = \tilde{\lambda}_\alpha (\tilde{w}, \tilde{w})$, from which

$$\tilde{\lambda}_\alpha = \frac{\langle {r_\pi}^2 \rangle}{\sum_{i=1}^n w_i^2 \sigma_i^2}, \tag{6}$$

where the L^2-normalization of $\tilde{V}_\alpha$ was used: $(\tilde{V}_\alpha \cdot \tilde{V}_\alpha) = 1$. As previously discussed, $\tilde{\lambda}_\alpha$ is the rescaled variance of the linear combination of the volatility-normalized returns $\tilde{r}_i$ (given by the eigenvector $\tilde{V}_\alpha$), each of which is decorrelated from other such combinations. By construction, the variance $\tilde{\lambda}_\alpha$ is normalized, which facilitates the comparison of different factors and different markets. We emphasize that diagonalizations of the covariance and correlation matrices are generally not equivalent; in particular, the eigenvalues λ_α, $\tilde{\lambda}_\alpha$ and the eigenvectors V_α, $\tilde{V}_\alpha$ are different (though in our case, their interpretations should be close). We choose the second option (i.e., Eq. (6)), which inherently reduces stock heterogeneity and heteroskedasticity due to rescaling.

Unfortunately, a straightforward diagonalization of the empirical covariance or correlation matrix estimated from stock price series is known to be very sensitive to noise (Laloux *et al.*, 1999; Plerou *et al.*, 1999, 2002; Potters *et al.*, 2005; Wang *et al.*, 2011; Allez and Bouchaud, 2012). In particular, only a few eigenvectors corresponding to the largest eigenvalues can be estimated, as illustrated and further discussed in Sec. V.D. As a consequence, conventional diagonalization does not appear suitable for building various representative factors.

B. *Our methodology: Indicator-based factors*

We propose a different approach to building factors. We begin from the available economic and financial indicators regarding the traded companies, such as their capitalization, sales-to-market, dividend yields, etc. We expect that companies with comparable indicators – at least those with comparable indicators in the extreme quantiles of the indicator distribution – will exhibit correlations in their stock performance. This hypothesis allows us to construct and then test indicator-based factors beyond sectors. To minimize sectorial correlations, we split the stocks into six supersectors of similar sizes, as detailed in Appendix A. The following construction is performed separately for each supersector and then the data are aggregated (see below).

We consider ten indicator-based factors:

1. The dividend factor, which is based on the dividend yield.
2. The capitalization (or size) factor, which is based on capitalization.
3. The liquidity factor, which is based on the ratio of the weekly exponential moving average to the total number of shares (i.e., capitalization/close price).
4. The momentum factor, which is based on the 3-year exponential moving average of past daily returns.
5. The low-volatility (or beta) factor, which is based on the sensitivity to the stock index.
6. The leverage factor, which is based on the debt-to-book value ratio.
7. The sales-to-market factor, which is based on the ratio of sales to market value at the end of the fiscal period.
8. The book-to-market factor, which is based on the ratio of the book value to the market value at the end of the fiscal period.

9. The remuneration factor, which is based on salaries and benefits expense per employee.
10. The cash factor, which is based on the ratio between the free cash flow and the latest market value.

We believe that considering these 10 factors is sufficient and including additional factors will not significantly change our results. In particular we might have included the investment and profitability factors following Fama and French (2015), but we expect that our 10 factors already capture the common risk from these two factors. Indeed, sales and cash should be correlated with profitability, whereas the dividend yield and leverage ratio should be correlated with investment.

For each trading day, the stocks of the chosen supersector are sorted according to the indicator (e.g., remuneration) available the day before (we use the publication date and not the valuation date). The related indicator-based factor is formed by buying the first qn_s stocks in the sorted list and shorting the last qn_s stocks, where n_s is the number of stocks in the considered supersector, and $0 < q < \frac{1}{2}$ is a chosen quantile level. The other stocks (with intermediate indicator values) are not included (weighted by 0). In the simplest setting, one can choose equal weights:

$$w_i = \begin{cases} +1, & \text{if } i \text{ belongs to the first } qn_s \text{ stocks in the sorted list,} \\ -1, & \text{if } i \text{ belongs to the last } qn_s \text{ stocks in the sorted list,} \\ 0, & \text{otherwise.} \end{cases} \tag{7}$$

In attempting to reduce the specific risk, the common practice suggests to invest inversely proportional to the stock's volatility σ_i, i.e., to set $w_i = \pm 1/\sigma_i$ or 0. Moreover, the inverse stock volatility should also be bounded to reduce the impact of extreme specific risk. Each trading day, we recompute the weight w_i as follows

$$w_i = \begin{cases} +\mu_+ \min\{1, \sigma_{\text{mean}}/\sigma_i\}, & \text{if } i \text{ belongs to the first } qn_s \text{ stocks in the sorted list,} \\ -\mu_- \min\{1, \sigma_{\text{mean}}/\sigma_i\}, & \text{if } i \text{ belongs to the last } qn_s \text{ stocks in the sorted list,} \\ 0, & \text{otherwise,} \end{cases} \tag{8}$$

where $\sigma_{\text{mean}} = \frac{1}{n_s}(\sigma_1 + \ldots + \sigma_{n_s})$ is the mean estimated volatility over the supersector. In this manner, the weights of low-volatility stocks are reduced to avoid strongly unbalanced portfolios concentrated in such stocks. The two common multipliers, $\mu_\pm$, are used to ensure the beta market neutral condition:

$$\sum_{i=1}^{n_s} \beta_i w_i = 0, \tag{9}$$

where β_i is the sensitivity of stock i to the market (obtained by a linear regression of the normalized stock and index returns based on the reactive volatility model (Valeyre *et al.*, 2013); note that the use of standard daily returns leads to similar results, see Appendix B). If the aggregated sensitivity of the long part of the portfolio to the market is higher than that of the short part of the portfolio, its weight is reduced by the common multiplier $\mu_+ < \frac{1}{2qn_s}$, which is obtained from Eq. (9) by setting $\mu_- = \frac{1}{2qn_s}$ (which implies that the sum of absolute weights $|w_i|$ does not exceed 1). In the opposite situation (when the short part of the portfolio has a higher aggregated beta), one sets $\mu_+ = \frac{1}{2qn_s}$ and determines the reducing multiplier $\mu_- < \frac{1}{2qn_s}$ from Eq. (9). This method of ensuring the market neutral condition is better than leaving the residual beta (as in the FF approach) or withdrawing it by subtracting an appropriate constant from all weights. Indeed, under our approach, the factor is maintained to be invested only in stocks that are sensitive to this factor. In turn, subtracting a constant would affect all stocks, even those that were "excluded" and whose weights were set to 0 in Eq. (8). We also emphasize the difference with the conventional FF approach: **our factors are built to be market-neutral under Eq. (9), whereas the FF portfolio is built to be delta-neutral** (i.e., to have zero net investment):

$$\sum_{i=1}^{n_s} w_i = 0. \tag{10}$$

The resulting factor is obtained by aggregating the weights constructed for each supersector. This construction is repeated for each of the ten factors listed above. We emphasize that the factors are constructed on a daily basis, i.e., the weights are re-evaluated daily based on updated indicators. However, most indicators do not change frequently so that the transaction costs related to changing the factors are not significant.

The above procedure can be extended to construct factors from other quantiles, in addition to the first and the last. In this manner, we will consider three portfolios for each factor:

- Q1: long positions for stocks whose indicator belongs to the first 15% quantile and short positions for stocks in the last 15% quantile, as discussed above (for $q = 0.15$).
- Q2: long positions for stocks in the second 15% quantile and short positions for stocks in the next-to-last 15% quantile (i.e., positive weights are assigned to stocks ranging between $0.15n_s$ and $0.30n_s$ in the list, and negative weights are assigned to stocks ranging between $0.70n_s$ and $0.85n_s$).
- Q3: long positions for stocks in the third 15% quantile ($0.30n_s - 0.45n_s$) and short positions for stocks in the third-to-last 15% quantile ($0.55n_s - 0.70n_s$).

To evaluate common risk with each factor, we introduce the *factor correlation level* (FCL) as the square root of the ratio between the empirical variance of the indicator-based factor and the total empirical variance of the constituent stocks:

$$\mathrm{FCL}(t) = \left(\frac{\mathrm{EMA}\left\{ {r_\pi}^2(t) \right\}}{\mathrm{EMA}\left\{ \sum_{i=1}^{n} w_i^2(t)\sigma_i^2(t) \right\}} \right)^{1/2}, \tag{11}$$

where $r_\pi(t)$ is the daily return of the factor,

$$r_\pi(t) = \sum_{i=1}^{n} w_i(t) r_i(t), \tag{12}$$

where $w_i(t)$ is the weight of the stock i in the factor, and $\sigma_i(t)$ is the volatility of the stock i estimated using the reactive volatility model (Valeyre *et al.*, 2013). The exponential moving average (EMA) is used with a long averaging period of 200 days to reduce noise by smoothing measurements. We emphasize that the above sum aggregates stocks from all supersectors. We also considered the standard volatility estimator based on a 40-days exponential moving average and obtained similar results (see Appendix B). The square root in Eq. (11) is taken to operate with volatilities instead of variances. The estimator (11) is built analogously to Eq. (6) for the eigenvalues $\tilde{\lambda}_\alpha$ of the correlation matrix. This analogy relies on the assumption that the indicator-based weights w_i are close to an eigenvector of the correlation matrix. Since the true correlation matrix is unavailable, it is impossible to directly validate this strong assumption. We will therefore resort to indirect validations based on empirical correlations of the constructed factors and on the profitability of trading strategies derived from such factors. Note also that the weights w_i depend on the choice of the quantile q, such that we expect to have slightly different results for different quantiles (see Fig. 4 below). Simultaneously, the analogy to eigenvalues of the correlation matrix allows various factors to be classified according to their "importance": larger values of FCL mean stronger volatility of the factor and therefore higher common risks. For example, when the correlation of small capitalization firms increases while the volatility of individuals stocks remains stable, the FCL of the capitalization factor will increase, and the volatility of the factor will increase. In general, the risk of a factor is proportional to the average individual volatility multiplied by the FCL. For this reason, **FCL can be interpreted as an average correlation measure between stocks within the factor that is also directly linked to the common risk level underpinning the factor.** It must also be emphasized that the FCL estimator is dynamic, i.e., it can capture changes in the correlation structure of the market over time.

IV. Data

In this study, we use only liquid stocks (most with capitalization greater than 800 million euros), thus excluding microcap firms that are typically the main focus of the labor studies we have cited. Thanks to the European accounting regulations, the remuneration must be provided by European companies on a regular basis and can thus be accessed through commercial databases such as FACTSET. Lacking such information for the U.S. market, we mainly focus on the European companies. To reveal possible nation-specific features, the analysis is performed for two trading universes: (i) the 569 biggest companies in Europe (London Stock Exchange, Euronext, Eurex, Sixt), and (ii) the 258 biggest companies on the London Stock Exchange only. Although the twice-as-large European universe is expected to increase the statistical significance of the results, the consideration of the U.K.-bounded universe allows us to eliminate country biases and additional fluctuations (e.g., due to currency exchange rate variations). We will show that the major conclusions are similar for both universes. In addition, we will validate our indicator-based methodology on the U.S. universe that includes the 569 randomly selected companies whose capitalization is above 1 billions of dollar. Note that the universe of the 1229 biggest firms in the U.S. studied by Fama and French (2008) is comparable to our European universe in terms of capitalization and liquidity.

All the companies that we include in the European and U.K. universes belong to the small (below 1 billion euros), mid (between 1 and 5 billion euros), large (between 5 and 20 billion euros), or big (above 20 billion euros) capitalization categories. The data set consists of 3612 daily single stock close prices from January 2001 to July 2015. **Note that most Fama and French data begin from 1963, which leads to greater t-statistics.** We rely on daily prices (instead of the monthly prices that are commonly used in the literature) to have more precision in the temporal granularity of our FCL estimation. In addition, several economic and financial indicators are extracted from the FACTSET database: book-to-market, capitalization, sales-to-market, dividend yield, debt-to-book, free cash flow, salaries and benefit expenses, and the number of employees on an annual basis (see Table I). For the European universe, we partly offset geographical biases in each indicator by renormalizing it to its median in the country. For instance, remuneration is divided by its median by country, whereas the median by country is subtracted from the moving average of returns in the case of momentum.

	Capitalization	Number of employees	Remuneration
Europe	(13 ± 25) B€	(41 ± 78) thousand	(0.13 ± 0.99) M€
U.K.	(11 ± 21) B£	(38 ± 87) thousand	(0.08 ± 0.08) M£

Table I Basic statistics (mean and standard deviation) regarding capitalization (in billions of euro/pounds), number of employees (in thousands), and remuneration (in millions of euro/pounds) from the FACTSET database. Since minimal capitalization is approximately 800 million euros, the distribution is truncated at small capitalizations.

V. Empirical results

In this section, we present the main results of our methodology applied to the European, the U.K. and the U.S. universes. We mainly focus on the remuneration indicator, which has largely been ignored so far. We will show that remuneration yields a non-negligible common risk and represents a small anomaly. The possibility of revealing the role of the remuneration factor relies on the proposed FCL methodology.

A. Correlation between remuneration and capitalization

First, we inspect the empirical joint distribution of remuneration and capitalization. This inspection is important because a positive size-wage effect has already been well documented in the economic literature for microcapitalization firms (Lallemand *et al.*, 2007). The wage gap due to firm size is approximately 35% (Oi and Idson, 1999) because large firms (but remaining in the microcapitalization category) demand a higher quality of labor and set a higher performance standard that must be supported by a compensating wage difference. Note that the magnitude and determinants of the employer-size wage premium vary across industrialized countries. Indeed, individual effects explain approximately 90% of inter-industry and firm-size wage differences in France (Abowd *et al.*, 1999), while almost 50% of the firm-size wage differentials in Switzerland derive from a firm-size effect (Winter-Ebmer *et al.*, 1999). In the U.K., larger firms pay better because of internal labor markets that reward effort and firm-specific capital (Belfield *et al.*, 2004). A visual inspection of Figure 1 (top) suggests that there is almost no correlation between remuneration and capitalization within the class of liquid stocks (that excludes microcapitalization firms) and, in any case, residual correlation is not significant. As a consequence, a larger firm from our sample does not necessarily pay its employees more. This result is consistent with the literature.

To confirm that the remuneration anomaly exists for different capitalizations, we split our sample in two groups: the above-median group of stocks whose capitalization exceeds the median size of our sample, and the below-median group with the remaining stocks

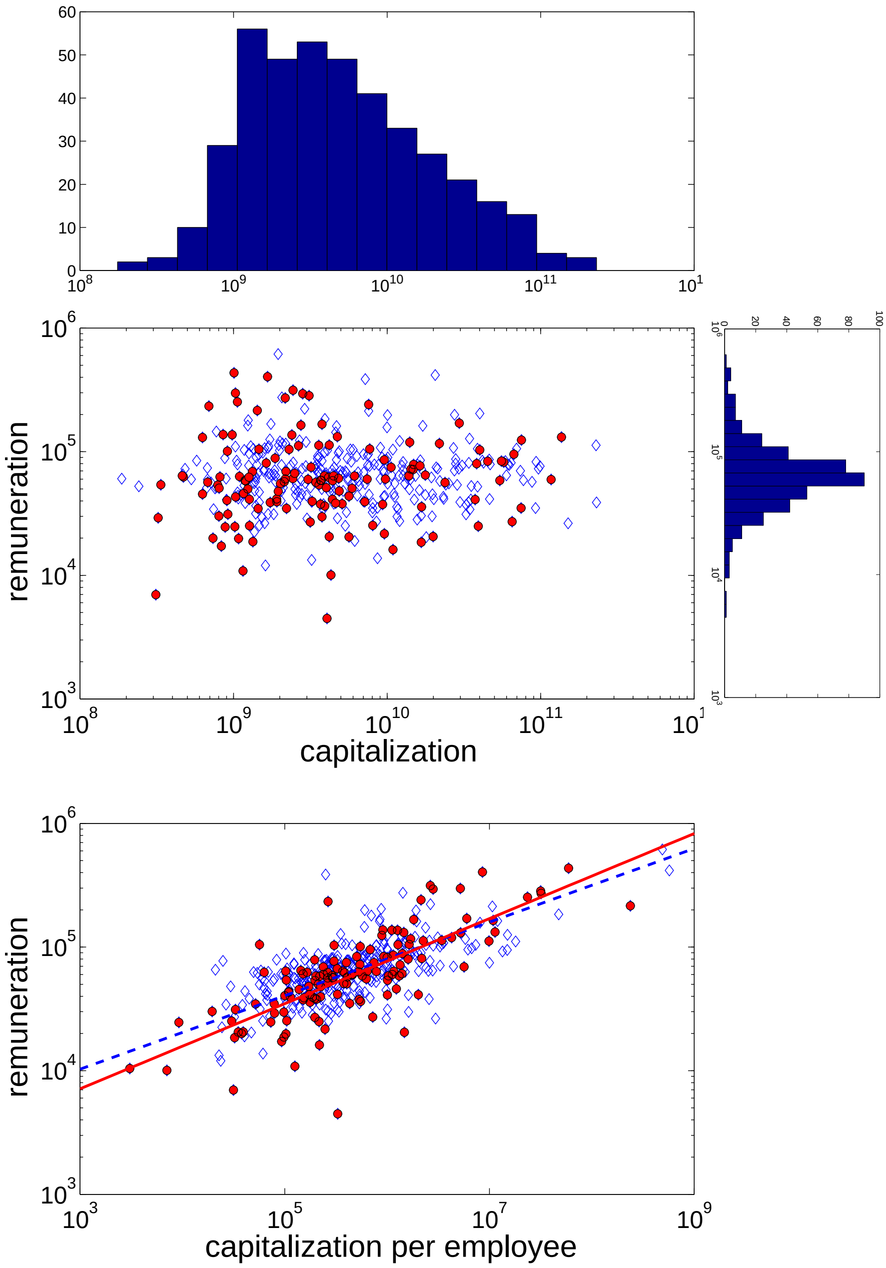

Figure 1. Remuneration versus capitalization (top) and remuneration versus capitalization per employee (bottom). Full circles and empty diamonds present large U.K. and European companies, respectively. Both quantities are shown in local currency and plotted on a logarithmic scale to account for significant dispersion in capitalization and remuneration. Solid and dashed lines indicate the linear regression between the logarithms of these quantities for the U.K. and European universes, respectively (the respective slopes are 0.34 and 0.30, and R^2 goodness of fit are 0.48 and 0.58, respectively). Since the records on remuneration and capitalization of each company in the FACTSET database are updated at different moments of the year, data were averaged over the period from 15/12/2014 to 30/07/2015. Similar results were obtained by taking the latest record for each company (not shown). Two subplots show the empirical distributions of capitalization (top) and remuneration (right) among the biggest European companies.

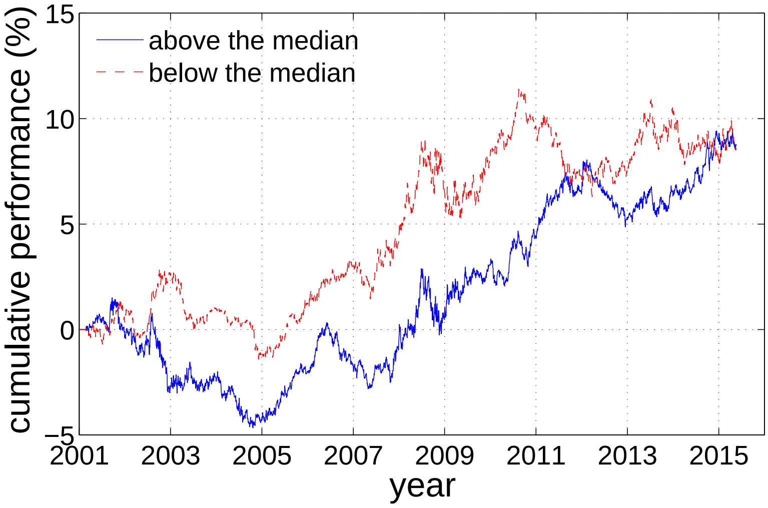

Figure 2. Similar cumulative performance anomalies of two remuneration factors for quantile Q1: one is constructed from stocks whose capitalization exceeds the median size of our sample, and the other is constructed from the remaining stocks. The cumulative performance of both factors after 15 years is approximately 9%, yielding an annualized performance of 0.6% (compared with 0.68% in Table IV). These curves are obtained for the European universe (the results for the U.K. universe are similar and thus not shown). The annualized performance for the remuneration factor is thus biased and cannot be fully explained by an unbiased random walk.

(we recall that both groups exclude microcapitalization firms). For each group, we build its own remuneration factor. Figure 2 shows that the cumulative performances of both remuneration factors are statistically different from 0 and behave similarly. An apparent slight outperformance of the factor constructed for the below-median group is not significant and can be attributed to statistical fluctuations.

Further investigations on the size-wage effect compel us to explore this relation per employee. Figure 1 (bottom) reveals that **remuneration is positively correlated to capitalization per employee**, i.e., remuneration increases with the amount of capitalization per employee. One plausible explanation for this phenomenon might be that reducing the number of employees (in particular, underperforming employees) increases marginal remuneration. In summary, there is no correlation between capitalization and remuneration for both universes of firms with capitalization over 800 million euros. Simultaneously, remuneration increases with the amount of capitalization per employee – as if the cake had to be shared fewer times.

B. Remuneration as a common risk

The motivation for building indicator-based factors relies on the hypothesis that the stocks with close indicator values behave similarly and thus share common risks. To verify this hypothesis, we compare three realizations of the remuneration factor built on different quantiles (Q1, Q2, and Q3), as described in Section III.B. Figure 3 shows weak but highly significant correlation between the daily returns of the remuneration factors from quantiles Q1 and Q2 (top) and Q1 and Q3 (bottom), notwithstanding that these factors have no stocks in common, which is the indirect proof that **the companies adopting similar remuneration policies (e.g., paying their employees well) share a common risk.** The weak correlation can be explained by a rapid decrease of the stock sensitivity to the remuneration factor with the quantile: the correlation level of (Q1,Q3) is measured to be half that of (Q1,Q2). The common risk is of the same order of magnitude as the residual risk, even for Q1. In summary, only the stocks in the extreme quantiles are the most sensitive to the remuneration factor. This observation is also confirmed by **the anomalies that are more important for extreme quantiles**, as shown in Figure 4.

C. Factor correlation level as a proxy of the eigenvalues

Ordering the factors based on their importance is central for the asset pricing analysis. As discussed in Sec. III.B, the relevance of indicator-based factors can be characterized using the factor correlation level (FCL) defined by Eq. (11). If the factor weights were approximately proportional to the elements of an eigenvector of the correlation matrix, the FCL would be an estimator of the volatility of this factor. The factors with larger FCL would most likely have greater impact on the portfolio returns for the same exposure. In general, the risk of a factor is proportional to the average individual volatility multiplied by the FCL. Thus, **FCL can be interpreted as an average correlation measurement between stocks within the factor.**

Using the daily returns of each factor and estimating the realized volatility of each stock, we compute the FCL for each factor based on Eq. (11). Figure 5 shows the time evolution of the FCLs for ten indicator-based factors defined in Sec. III.B. For comparison, we plot the FCLs for the European and the U.S. universes (the FCLs for the U.K. universe behave similarly and are thus not shown). First, the FCLs exhibit strong variations over time. In particular, the FCLs of two factors can cross each other, i.e., the ordering of the factors based on their "importance" can evolve over time. For both universes, the low-volatility factor appears as the most important, followed by capitalization and momentum factors. Other factors are smaller but statistically significant. Averaging the FCL over 15 years al-

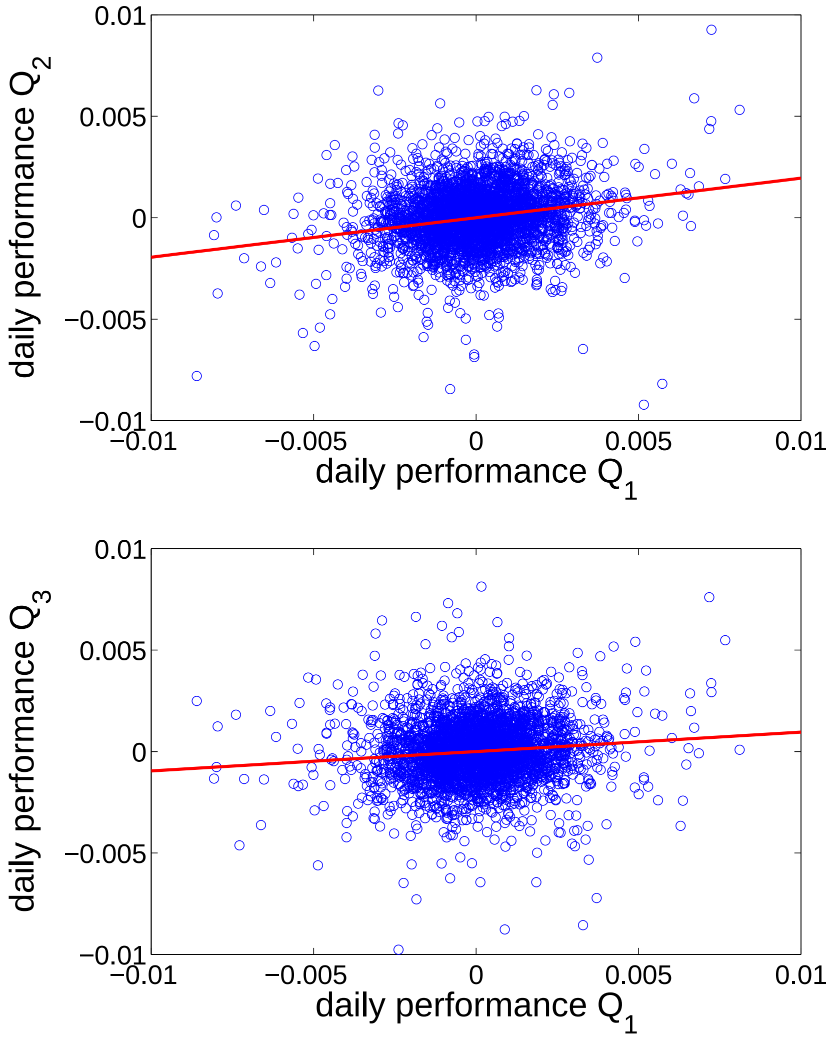

Figure 3. **(Top)** Correlation between the daily returns of the two remuneration factors constructed on quantiles Q1 (0%–15% and 85%–100%) and Q2 (15%–30% and 70%–85%), which have no stocks in common. The daily returns of these factors are weakly correlated but correlation is significant: the slope and its 95%-confidence interval is 0.19 ± 0.03. **(Bottom)** For comparison, the correlation between the daily returns of the remuneration factors Q1 and Q3 (30%–45% and 55%–70%) is shown, with the slope and its 95%-confidence interval 0.10 ± 0.03. Both graphs were obtained for the European universe. Similar graphs for the U.K. universe yield the slopes 0.23 ± 0.03 and 0.02 ± 0.03 for Q1-Q2 and Q1-Q3 correlations, respectively (graphs are not shown but are available upon request).

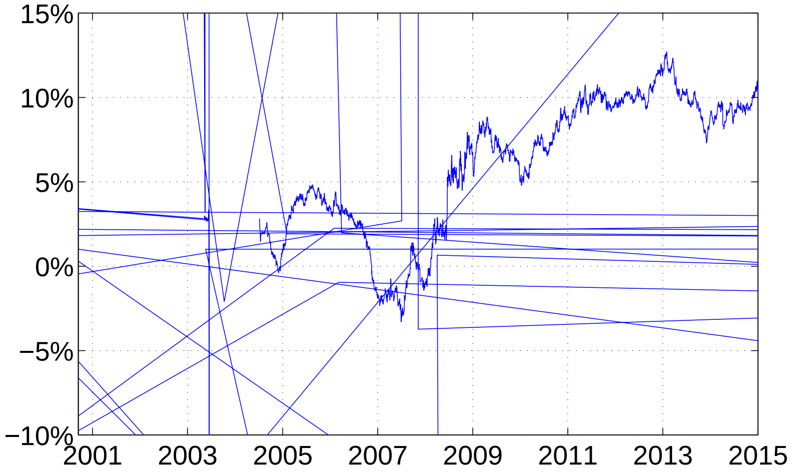

Figure 4. The cumulative performance of the remuneration factor for the three quantiles (Q1, Q2 and Q3) for the European universe (the graph for the U.K. universe is similar and is available upon request). Biases are more pronounced for Q1 than for Q2 or Q3, which might be explained by the possibility that stocks belonging to the extreme quantile are the most sensitive to the remuneration anomaly.

lows us to order the factors according to their importance. Table II suggests **the following order for the European universe: low-volatility (1.73), capitalization (1.72), momentum (1.41), sales-to-market (1.22), liquidity (1.19), book-to-market (1.13), dividend (1.09), leverage (1.07), remuneration (0.99), and cash (0.92).** All these FCLs are higher than the noise level of 0.78 that we estimated by building a "noise factor" according to an arbitrary non-financial indicator, such as an alphabetic order. Even though the remuneration factor is relatively small, its magnitude remains statistically relevant in comparison with other well-known factors. For example, the FCLs of the book-to-market, dividend, leverage and cash factors are close to that of the remuneration factor. Their low values mean that these factors are not particularly volatile and that the related common risks are low. Conversely, the low-volatility factor (excluded from the FF approach) has the highest FCL and is thus identified as the first potential source of risk in a portfolio, after market index and sectorial risks. Notably, the low-volatility factor is comparable to the capitalization factor and greatly exceeds the book-to-market factor, the two "major" factors identified in the Fama and French (1993) model.

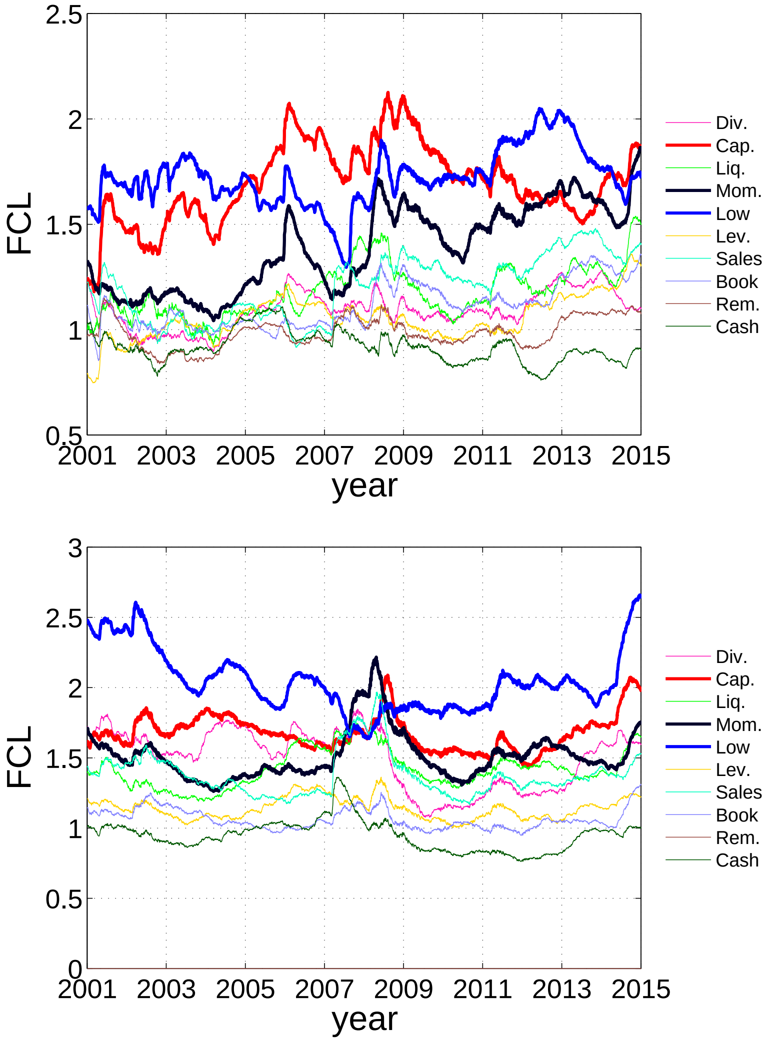

Figure 5. Evolution of the factor correlation level (FCL) for ten factors (quantile Q1): the European (top) and USA (bottom) universes (the behavior for the U.K. universe is similar and available upon request). In our interpretation, FCL is a measure of "importance" of factors in asset pricing models. Thick lines highlight the three major factors: low-volatility, capitalization, and momentum. The mean FCLs averaged over 14 years are summarized in Table II. All FCLs are highly volatile, but this volatility is not linked to stock market volatility. In addition, we can see the jump- and cross-over of FCLs. During the 2007–2008 financial crisis, several FCLs collapse for the U.S. universe. Note that we could not construct the remuneration factor for the U.S. universe because of lack of systematic remuneration data for U.S. companies.

FCL	Div.	Cap.	Liq.	Mom.	Low	Lev.	Sales	Book	Rem.	Cash	Market
Europe	1.09	1.72	1.19	1.41	1.73	1.07	1.22	1.13	0.99	0.92	10.41
U.K.	0.97	1.45	0.92	1.15	1.38	0.96	1.03	0.96	0.93	0.83	6.73
U.S.	1.49	1.73	1.49	1.62	2.10	1.15	1.41	1.12	–	0.95	12.35

Table II The mean value of the FCL for ten factors (quantile Q1) averaged over the period from 10/08/2001 to 31/07/2015, for the European, U.K., and U.S. universes. According to these values, the main factors for asset pricing are the low-volatility factor (excluded from the FF approach), followed by the capitalization, and momentum factors. We see that the book-to-market and remuneration factors are of the same order of magnitude such that the remuneration factor should have the same importance in asset pricing models as the book-to-market factor. We also estimated the FCL of the market (last column). The FCL of a noise factor was estimated to be around 0.8 for three universes implying that all presented factors exceed noise. Note that we could not construct the remuneration factor for the U.S. universe because of lack of systematic remuneration data for U.S. companies.

D. Comparison with the principal component analysis

The principal component analysis (PCA), which is applied to decorrelate time series, consists in forming the empirical correlation matrix from daily stock returns and then finding its eigenvalues and eigenvectors. In practice, the number of stocks in a traded universe (typically 500 - 1000) is often comparable to the number of available historic returns per stock (for instance, 3612 daily returns in our dataset), that makes this general method strongly sensible to noise, as discussed in (Laloux *et al.*, 1999; Plerou *et al.*, 1999, 2002; Potters *et al.*, 2005; Wang *et al.*, 2011; Allez and Bouchaud, 2012).

In order to illustrate this limitation, we apply the PCA to the European universe and compute numerically 569 eigenvalues. Figure 6 shows the histogram of square roots of the obtained eigenvalues, i.e., how many eigenvalues are contained in successive bins of size 0.0626. The largest value, $\lambda_{\text{market}}^{1/2} \approx 12.62$, corresponding to the market mode, was excluded from the plot for a better visualization of other values. One can identify approximately ten well-separated single eigenvalues that are typically attributed to market sectors. In turn, the remaining part of (smaller) eigenvalues lying close to each other and thus almost indistinguishable, can be rationalized by using the random matrix theory (Laloux *et al.*, 1999). If the daily stock returns were distributed as independent Gaussian variables (with mean zero and variance one), the eigenvalues of the underlying empirical correlation matrix would asymptotically be distributed according to the Marcenko-Pastur density

$$\rho(\lambda) = \frac{\sqrt{4q\lambda - (\lambda + q - 1)^2}}{2\pi q\lambda}, \tag{13}$$

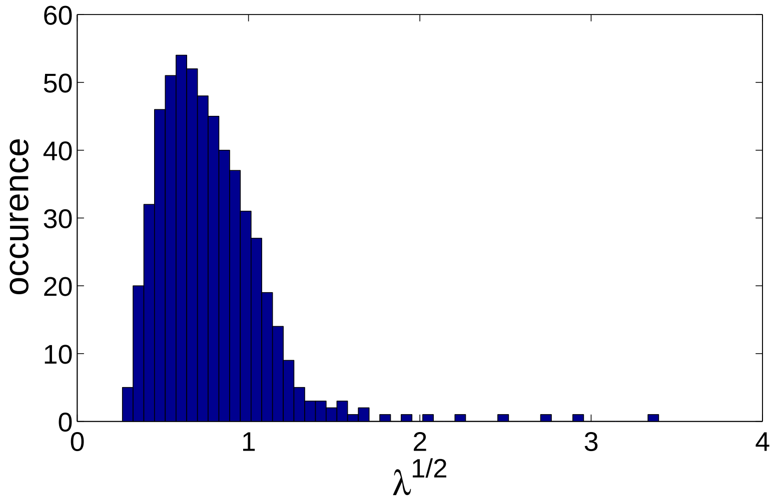

Figure 6. Histogram of square roots of eigenvalues, $\lambda^{1/2}$, of the empirical correlation matrix obtained from daily returns of 569 stocks in the Europe universe over the period from 10/08/2001 to 31/07/2015. The largest value, $\lambda_{\mathrm{market}}^{1/2} \approx 12.62$, corresponding to the market mode, was excluded from the plot for a better visualization of other values.

where $q = N/T$ is the ratio between the number of stocks, N, and the number of daily returns per stock, T. These eigenvalues lie between two critical values, $\lambda_{\min} = (1 - \sqrt{q})^2$ and $\lambda_{\max} = (1 + \sqrt{q})^2$. As a consequence, the eigenvalues obtained by diagonalizing the empirical correlation matrix and lying below $\lambda_{\max}$ can be understood as statistical uncertainty of the PCA. In other words, the PCA cannot reliably identify the factors with $\lambda < \lambda_{\max}$. For our European universe, $q = 569/3612$ so that $\sqrt{\lambda_{\max}} \approx 1.4$ determines a theoretical threshold between larger, significant eigenvalues, and smaller, noisy ones.

Comparing large values in Fig. 6 to the FCL from Table II, we conclude that PCA might identify three major factors: low-volatility (1.73), capitalization (1.72), and momentum (1.41). In turn, the other factors whose the FCL is smaller than the PCA threshold 1.4, would thus be understood as statistical uncertainty in the PCA method. **The crucial advantage of our method, in which factors are built from firm-based indicators while market and sectorial correlations are eliminated by construction, is the possibility to go beyond this PCA limit and to identify the factors with smaller FCLs.** Moreover, this identification can be performed over time.

E. Net investment as a proxy of the exposure to the low-volatility factor

Building market-neutral portfolios requires nonzero net investment when the portfolio is exposed to the low volatility anomaly. This anomaly is governed by the low-volatility factor, which is the most influential factor (after market and sectors) according to our FCL measurement (Table II), and unfortunately a residual exposure to the low-volatility factor cannot be easily reduced. As a result, most factors can still be correlated to the low-volatility factor. Thus, when the average beta of long stocks in a factor is significantly different from the average beta of short stocks, the factor is also exposed to the low-volatility factor with a nonzero net investment. The net investment is defined as the difference between long ($\omega_i > 0$) and short ($\omega_i < 0$) investments normalized by total investment, i.e.,

$$\Delta = \frac{\sum_{i=1}^{n} w_i}{\sum_{i=1}^{n} |w_i|}. \tag{14}$$

By construction, Δ can vary between -1 and 1 or, equivalently, between -100% and 100%.

Replacing the individual sensitivities β_i in the market neutral relation (9) by the averages $\langle \beta_L \rangle$ and $\langle \beta_S \rangle$ for long and short stocks, the net investment Δ from Eq. (14) can also be expressed as

$$\Delta = \frac{\langle \beta_S \rangle - \langle \beta_L \rangle}{\langle \beta_S \rangle + \langle \beta_L \rangle}. \tag{15}$$

When the average sensitivities for long and short stocks are similar, net investment is close to 0. In turn, a net bias in Δ occurs when the average beta is different for long and short stocks. Δ is a proxy of the exposure to the low-volatility factor that is more reactive and more precise than the estimation obtained through the usual regression of returns.

The bias in the long and short betas in Eq. (15) may also be related to the sensitivity to the market (i.e., to the stock index) of a factor built with the FF approach (i.e., neutral in nominal but not in beta):

$$\beta_{FF} = \langle \beta_L \rangle - \langle \beta_S \rangle = -2\langle \beta \rangle \Delta, \tag{16}$$

where $\langle \beta \rangle = \frac{1}{2}(\langle \beta_S \rangle + \langle \beta_L \rangle)$ is the average beta of the universe that we estimated as $\langle \beta \rangle \approx 0.65$ for the period from 2001 to 2015. The net investment Δ can also be related to the sensitivity of any beta neutral portfolio or factor (both in the FF approach and in our methodology) to the low-volatility factor (the most influential factor, according to the FCL).

Figure 7 shows that the low-volatility factor has the most important short investment (negative values of Δ ranging between -80% and -60%), although its sensitivity to the market was maintained at 0. Other factors also have a bias in Δ, including the capitalization

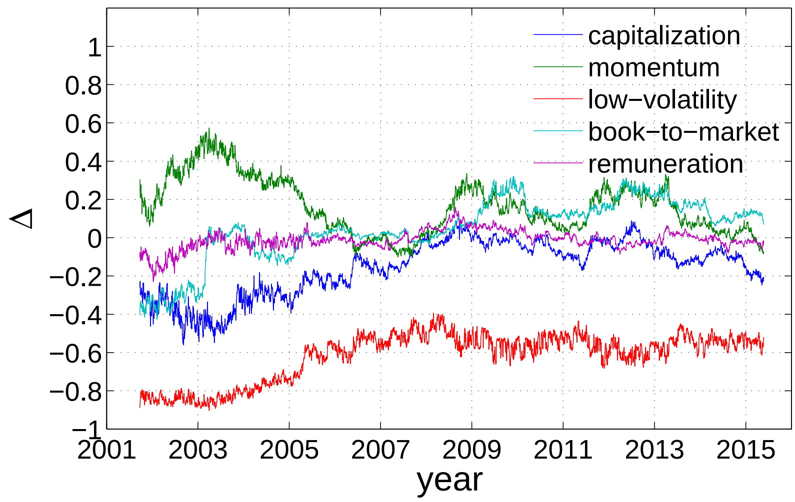

Figure 7. Evolution of the net investment Δ for five indicator-based factors for the European universe: capitalization, momentum, low-volatility, book-to-market, and remuneration (the results for the U.K. universe are not shown but are available upon request). We recall that Δ is a proxy of the exposure to the low-volatility factor. The remuneration Δ is around zero, and the factor therefore has no correlation with the low-volatility factor. Other factors seem to be more exposed to the low-volatility factor.

and the momentum factors, in particular. In the FF approach, these factors would therefore also have a significant sensitivity to the market. In particular, the low-volatility factor built with the FF approach would be strongly correlated to the market. Moreover, Δ indicates that most factors have a residual correlation with the low-volatility factor that remains uncorrected by our method. Since 2003, the Δ of the book-to-market factor (one of the major anomalies investigated by Fama and French) has shrunk, and **the related book-to-market anomaly has almost disappeared** (see Table IV). Finally, the remuneration factor shows nearly zero net investment, i.e., it remains uncorrelated with the low-volatility factor.

F. Other inter-factor correlations

Correlations between factors matter as long as one needs uncorrelated portfolios for asset pricing purposes. The indicator-based factors were introduced to build as many uncorrelated portfolios as possible. At the same time, such an explicit construction does not guarantee to yield truly uncorrelated combinations, such as the eigenvectors of the covariance (or correlation) matrix. Moreover, some indicators may capture the same economic or financial features of the company and may thus be correlated; in other words, different factors may approximate the same eigenvector and thus be highly correlated. In particular, adding new indicator-based factors does not necessarily help to capture new features and may thus be redundant. The choice of the ten indicator-based factors studied in this paper is judged as sufficient with respect to the trade-off between capturing information and remaining uncorrelated. Table III presents the correlation coefficients between ten indicator-based factors estimated from their volatility-normalized daily returns. Clearly, many indicator-based factors remain correlated. If the same estimation was applied to ten independent Gaussian vectors of the same length ($m = 3612$ elements), the standard deviation of the estimated correlation coefficients would be $1/\sqrt{m} \approx 0.0166$. In other words, the presented correlations between the indicator-based factors are highly significant.

The remuneration factor exhibits correlations with some other factors, and the most significant of these include the following: the sales-to-market (-0.38), dividend (-0.23), and momentum (0.20) factors. These correlations can be explained as follows. First, **the companies with low sales-to-market ratios have a high margin and thus the ability to pay their employees well** (strong negative correlation -0.38). The direct link between a firm's margin and wage is well documented in the labor economics literature. More precisely, there is a relation between margin and labor cost. For instance, a study by the European Central Bank (ECB) and the Organization for Economic Co-operation and Development (OECD) reveals that larger firms make more extensive use of margin for labor cost-cutting strategies, i.e., firms choose to reduce benefits as a cost-cutting strategy (Babecky *et al.*, 2012). In addition, the positive relation between firm size and the use of cost-cutting strategies that is monotonically increasing and highly significant, is uncovered. Second, **the companies that pay high dividends to shareholders tend to remunerate their employees less**, yielding a negative correlation of -0.23, which is a direct representation of profit-sharing within firms. Indeed, dividend payments are charged on the profits of the business after all salaries and benefits expenses are paid out. Although this result appears intuitive, it remains important as it reveals the level of correlation between both quantities. The labor economics literature and the corporate finance literature are not very well documented on this particular issue. Finally, **companies that perform well and show strong momentum can offer higher remuneration to their employees or, alternatively, the higher remuneration stimulates employees to work better and to imbue the company with momentum** (positive correlation 0.20). This is a central and very important result of our research because it highlights the positive relation between pay and performance. The rationale behind this result is discussed in Section VI.

It is worth emphasizing that these correlations between factors are not static (as presented in Table III by averaging over 15 years) but evolve over time. For example, Fig.

	Div.	Cap.	Liq.	Mom.	Low	Lev.	Sales.	Book.	Rem.	Cash
Div.		0.10	0.14	**-0.33**	0.02	0.29	0.26	0.18	-0.23	0.14
Cap.	0.10		0.08	0.10	0.13	0.21	-0.20	-0.06	0.05	-0.01
Liq.	0.14	0.08		-0.21	0.20	0.09	0.05	0.05	-0.06	0.05
Mom.	**-0.33**	0.10	-0.21		-0.18	-0.24	-0.25	**-0.36**	0.20	-0.04
Low	0.02	0.13	0.20	-0.18		0.02	0.01	0.07	-0.03	0.05
Lev.	0.29	0.21	0.09	-0.24	0.02		0.23	0.11	-0.17	-0.02
Sales.	0.26	-0.20	0.05	-0.25	0.01	0.23		**0.31**	**-0.38**	0.23
Book.	0.18	-0.06	0.05	**-0.36**	0.07	0.11	**0.31**		-0.13	0.05
Rem.	-0.23	0.05	-0.06	0.20	-0.03	-0.17	**-0.38**	-0.13		-0.11
Cash	0.14	-0.01	0.05	-0.04	0.05	-0.02	0.23	0.05	-0.11	

Table III Correlation coefficients between 10 indicator-based factors for the U.K. companies: Dividend (1), capitalization (2), liquidity (3), momentum (4), low-volatility (5), leverage (6), sales-to-market (7), book-to-market (8), remuneration (9), and cash (10). These coefficients were estimated from daily returns of these factors over the period from 23/02/2001 to 27/07/2015. Daily returns of each factor were normalized by their volatility averaged over 20 days to reduce the effects of heteroskedasticity. Similar correlation coefficients were obtained for the European companies (available upon request).

8 shows the evolution of two correlation coefficients between volatility-normalized daily returns of remuneration, low-volatility, and sales-to-market factors. The correlation between the remuneration and low-volatility factors remains close to zero, with eventual deviations beyond the Gaussian significance range (e.g., during the subprime and financial crises in 2007-2009). These two factors can be considered uncorrelated. In turn, the negative correlation between the remuneration and sales-to-market factors always remains beyond the Gaussian significance range.

G. The anomaly of the remuneration factor and its interpretation

Table IV compares the remuneration anomaly with other factors in terms of the annualized bias (the annualized cumulative return between the last and the first observation days), the Sharpe ratio (the annualized bias normalized by annualized volatility), and t-statistics (the Sharpe ratio multiplied by the square root of the total duration in years). In particular, the t-statistic allows one to reject the null hypothesis of no bias at the 90% confidence level.

The bias reveals the level of overperformance due to a particular factor. We observe a significant bias for the dominant capitalization and low-volatility factors, which have been previously documented. The anomaly of the book-to-market factor seems to have disappeared (see Table IV). In fact, the Sharpe ratio that we estimated to be 0.49 for the period

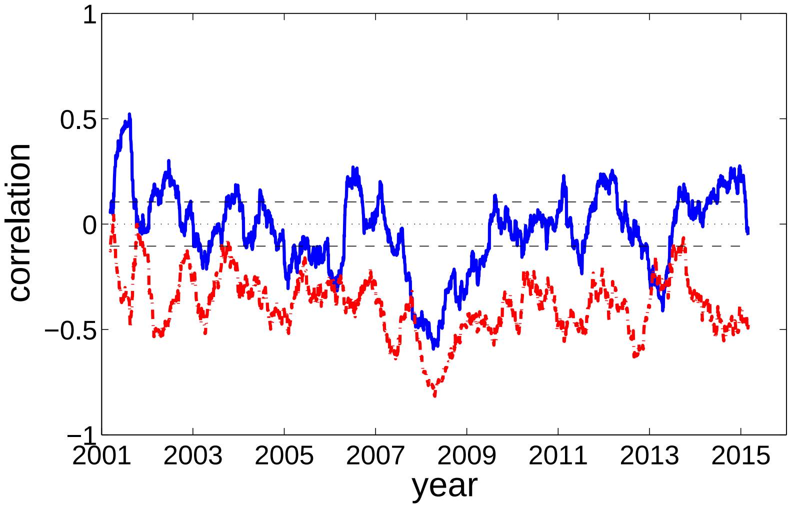

Figure 8. Correlation coefficients between daily returns of the remuneration factor and of the low-volatility factor (solid line) or the sales-to-market factor (dashed line) for the largest U.K. companies. The coefficients were computed over a sliding window of 90 days. Prior to computation, the daily returns were renormalized by their average volatility over the previous 20 days. The mean values over 15 years are -0.03 and -0.38 (see Table III), respectively. Horizontal dashed lines show the standard deviation, 0.105, of the same estimator applied to two independent Gaussian samples. Similar results were obtained for the European universe (available upon request).

from 1926 to 2008 in the U.S.[3], became much smaller in recent years (and even changed the sign for the European universe, becoming −0.08). We suspect that this result can be explained by the change in its exposition to the low-volatility factor. The momentum factor has also changed direction.

The remuneration factor appears as the sixth most important anomaly in the U.K. market, and the eighth most important anomaly in the European market. **A bias of 1.21% means that companies that pay better should overperform their less paying competitors by $2 \times 1.21\%$.** The prefactor 2 appears if we assume that 50% is invested in high remuneration and 50% in low remuneration (i.e., there is no exposure to the low-volatility factor and volatility is nearly homogeneously distributed). This is one of the most important results in this paper, as it shows that **a market neutral investment style arbitrage strategy based on the remuneration anomaly is likely to deliver positive returns.** Next, assuming that the bias in the remuneration factor consists of an intrinsic bias and contributions from biases of other factors due to inter-factor correlations, the relative impacts of these biases can be estimated by multiplying them by the correlation coefficients in the 9^{th} line of Table III. These relative impacts are summarized in the last line of Table IV. Since most contributions from other factors are negative, it might be surmised that the intrinsic remuneration bias is even higher than 1.21% (estimated to be around 2.85%) but that its value is reduced due to correlations with other factors. If we were able to build a remuneration factor fully decorrelated from other factors, we would have obtained most likely a t-statistic above 3 (around 3.29, see Table IV) that fulfills the requirements formulated by Harvey *et al.* (2016). Note also that there is no selection bias in our study (we have not analyzed all the different possibilities to finally retain the remuneration factor), such that the condition requiring a t-statistic greater than 3 when taking into account the number of possible anomaly candidates is not applicable. In any event, the observed bias of 1.21% cannot simply be explained by the biases of other factors. The Sharpe ratio of 0.37 indicates that a horizon of $1/0.37 \approx 2.7$ years is required for the anomaly to be captured and to have a positive return with a likelihood of 84%. From an asset management point of view, it suggests the recommended time horizon to take profits based on this market anomaly.

H. The rationale behind the remuneration anomaly

Our analysis clearly reveals correlations between remuneration policies of a company and performances of its stock. Do higher wages imply better performances, or better performances lead to higher wages? More generally, is the relation between remuneration and

[3] Based on the publicly available data from Fama and French, `http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html`

		Div.	Cap.	Liq.	Mom.	Low	Lev.	Sales.	Book.	Rem.	Cash
Europe	Bias, %	2.39	-5.72	-0.95	-1.60	-4.15	-1.95	0.08	-0.23	0.68	1.66
	Sharpe	0.80	-1.69	-0.41	-0.42	-1.46	-0.74	0.03	-0.08	0.25	0.65
	t-stat	3.04	-6.42	-1.57	-1.59	-5.57	-2.81	0.11	-0.30	0.97	2.46
U.K.	Bias, %	2.12	-4.29	-0.11	-2.81	-3.81	-1.01	0.92	0.34	1.21	2.60
	Sharpe	0.65	-1.38	-0.05	-0.71	-1.25	-0.35	0.31	0.11	0.37	0.92
	t-stat	2.48	-5.24	-0.18	-2.71	-4.77	-1.34	1.16	0.40	1.40	3.51
	Impact, %	-0.49	-0.21	0.01	-0.56	0.11	0.17	-0.35	-0.04	2.85	-0.29

Table IV The annualized bias (the annualized cumulative return between the last and the first observation days, as a percentage), the Sharpe ratio (annualized bias normalized by annualized volatility), and the t-statistic (the Sharpe ratio multiplied by the square root of the total duration in years, i.e., by $\sqrt{14.5} \simeq 3.81$) for the following 10 indicator-based factors (quantile Q1): dividend (1), capitalization (2), liquidity (3), momentum (4), low-volatility (5), leverage (6), sales-to-market (7), book-to-market (8), remuneration (9), and cash (10). These quantities are estimated for the period from January 2001 to July 2015, for the largest European companies (top lines) and for the largest U.K. companies (bottom lines). The last line shows the relative impacts of the biases of various factors on the remuneration bias (1.21) for the U.K. companies. These impacts are obtained by multiplying the biases in the fourth line by the correlation coefficients from the 9^{th} line of Table III. The annualized bias for the remuneration factor in the U.K. universe is 1.21% with a t-statistic of 1.40. Moreover, if we subtract all the impacts from remuneration's annualized bias, we obtain an intrinsic remuneration bias of 2.85%. **Therefore, we would have a t-statistic of approximately 2.85 $\times$ 1.40/1.21 = 3.29 that would fulfill the requirements formulated by Harvey** ***et al.*** **(2016)**.

performance causal? In analogy to the chicken or the egg causality dilemma, wages and performances are likely to be entangled, while the causality direction can change from firm to firm or even over time. Although these challenging questions are difficult to answer in a quantitative way, we provide two arguments in favor of a causal relation between wages and performances. First, the remuneration factor is built on the published balance sheets that reflect wages of the past year. As a consequence, there is a significant one or even two-year delay between earlier remunerations and current performances. In this way, we capture the impact of wages on performances. Second, the top and bottom remuneration quantiles are not static but change in time (half of the companies in each quantile are replaced in approximately 5 years). One can speculate that the management of a company competes with others by offering higher remuneration to attract the best employees who will make the performance of the company stronger. As football team managers, companies could buy success by investing in human resources (Simmons and Forrest, 2004).

In a survey paper, Yellen (1984) poses the question of why firms do not cut wages in an economy characterized by involuntary unemployment? Indeed, unemployed workers would prefer to work at the real wage rather than being unemployed, but firms will not hire them at a lower wage simply because any reduction in wage would lower employee productivity. This is Yellen's most-cited paper, and it stipulates that the amount of effort that employees put into their job depends on the difference between the wage they are getting paid and what they perceive as a "fair wage". The bigger the difference, the less hard they tend to work, which highlights the idea that paying employees more than the market clearing wage may boost productivity and ends up being worthwhile for the employer. Paradoxically, cutting wages may end up raising labor costs since it will negatively affect productivity (Stiglitz, 1981). Hence, productivity is the main argument, which is confirmed by other theoretical papers that consider employees to be more productive in larger firms and thus explain why they demand higher wages (Idson and Oi, 1999). The other arguments are as follows. Given job contract incompleteness, not all duties of an employee can be specified in advance. For this reason, monitoring is a central instrument to control production costs (Alchian and Demsetz, 1972). Unfortunately, monitoring is too costly and sometimes inaccurate due to measurement error. Instead of having costly and imperfect monitoring, firms can offer higher wages to their employees to create an incentive for the employee not to lose their high wage by being fired (Shapiro and Stiglitz, 1984). In this context, paying a wage in excess of the market clearing wage can be seen as an efficient way to prevent employees from shirking. The attractiveness of wages to skillful workers also contributes to reduce their turnover. Moreover, raising wages partly eliminates job demands from less performing candidates who would fear competing with overperforming candidates. This adverse selection is a subtle support for the fair wage hypothesis because paying fair wages will attract only the more skillful workers and deter lemons and will thus help avoid costly monitoring devices in the recruitment processes. In summary, the motivation for the fair wage-effort hypothesis is a simple observation of human nature arguing that employees who receive less than what they perceive to be a fair wage will not work as hard as a consequence. In the very same vein, Akerlof and Yellen (1990) set up a model of unemployment in which “people work less hard if they are paid less than they deserve, but not harder if they receive more than they deserve”. The model puts in equation the fair wage-effort hypothesis to represent the idea that a poorly paid employee may be keen on taking its revenge on its employer.

VI. Discussion

A. Fama and French approach

Fama and French (1993, 2015) use time series of 25 portfolios, each portfolio built with similar capitalization and book-to-market stocks. They regress the monthly performance $R_i(t)$ of each portfolio i on the returns $f_j(t)$ of different factors j:

$$R_i(t) = a_i + \sum_j b_{i,j} f_j(t) + \varepsilon_i(t),$$

where a_i and $\varepsilon_i(t)$ are portfolio-specific intercept and noise, and $b_{i,j}$ is the estimated sensitivity of the i-th portfolio to the j-th factor.

If the remuneration factor had to be investigated using the FF approach, how could one proceed? Five different portfolios might be built with stocks sorted according to remuneration and then at least three major factors might be used: the market index, capitalization, and book-to-market factors (the factor returns, $f_j(t)$, would be estimated through the performance of the long-short portfolio, e.g., buying the high capitalization and shorting the low capitalization, or buying the high book-to-market and shorting the low book-to-market). The intercept, a_i, for the 5 different portfolios might be measured with their t-statistics to assess whether the remuneration is an anomaly. One might also measure the $a_{\text{high}} - a_{\text{low}}$ and its t-statistics, as in Table 2 by Fama and French (2008). Finally, the remuneration factor might be added to the regression panel and the R^2 for every portfolio might be measured to quantify how well the data fit the statistical model and how well the common factors explain the price returns.

Instead, we simply measure the average returns of the HML portfolio (see Table IV) built to be beta-neutral without any regression, as we construct our remuneration factor

Sectors	Median book-to-market	Median remuneration (in euros)
Consumer discretionary	0.31443	22 859.96
Consumer staple	0.24681	39 416.51
Energy	0.81440	137 625.91
Financial	0.87972	126 498.10
Health	0.24442	51 452.06
Industrial	0.32765	58 626.27
IT	0.19867	77 854.94
Material	0.55733	32 516.14
Telecom	0.39122	66 283.21
Utilities	0.32572	47 014.69

Table V Sectorial variations of the median of the book-to-market and of the remuneration (in euros) for the U.K. universe in 2014. Both book-to-market value and remuneration vary substantially across different sectors.

as uncorrelated to the main factors. That should be close to the $\frac{1}{2}(a_{\text{high}} - a_{\text{low}})$ of the FF approach, or close to the average return of the HML portfolio built to be delta-neutral (see Table I from Fama and French (2015)). This is due to the fact that the remuneration factor is not exposed to the market index, low-volatility and book-to-market factors. However, the FF approach would not account for the fact that remuneration depends on sectors (see Table V). Using the volatility of the portfolio, we can also measure the t-statistics to learn whether the anomaly is statistically significant, and we measure the FCL to quantify how well the common factors explain the price returns.

In Appendix B we compare the FF approach to our methodology. In particular, we show that sectorial constraint and beta-neutral property were the two key advantages of our factors construction: without them, **the FF approach applied to the same period, would give insignificant results for the remuneration factor** (we recall that most Fama and French data begin from 1963, which leads to greater t-statistics).

B. Advantages and limitations of the methodology

Our methodology has several advantages over the FF approach:

1. The estimated FCL quantifying the relevance of the factor does not depend on the number of considered factors, in contrast to the R^2 argument of the FF approach (e.g., see Table 6 in Fama and French (1993)). Thus, one can select the most important factors (e.g., stock index, low-volatility, capitalization, liquidity, and momentum factors) in asset pricing models.
2. The sensitivities of the different common risk factors to the market (i.e., to the stock

index) are maintained at zero even for the low-volatility factor, which is an important feature because the market mode may have a hundred times greater impact on portfolio returns than other factors.

3. The factors are constructed to be sector neutral, which allows one to better identify their impacts on price variations, which is important because intra-sector correlations are typically more important than within-factor correlations. Notably, the book-to-market factor of FF approach also captures sectorial risk, as the firms are not priced in the same way from one sector to another (see Table V). In particular, the remuneration is very different from one sector to another.
4. Weights (w_i) of the stocks that are close in capitalization (or in book-to-market, or in remuneration, etc., depending on the factor) are of the same order of magnitude that reduces the specific risk of the factor.
5. Maintaining factors beta-neutral at any time reduces the noise of factors, even those that are not supposed to be correlated to the stock index. In fact, we will show in Appendix B that in the case of factors uncorrelated to the stock index, the beta-neutral constraint reduces the volatility of the factor by 1.2% on an annualized basis.
6. Our method enables the inclusion of the low-volatility factor into the cross-section of average returns (in contrast to the FF approach) without any multiregression model. The low-volatility and capitalization factors were found to provide the largest anomaly (see Table IV). In addition, the low-volatility factor was also identified as the major contribution to risk, according to our measurement (see Fig. 5). Surprisingly, the capitalization factor, which had previously been considered as the most important, now occupies the second position. Moreover, the book-to-market factor identified by Fama and French (1993) as important, has eventually become a minor factor (and is just slightly more important than the remuneration factor) after having eliminated the sectoral and market modes.

The main limitations to our methodology are related to the methodology itself. Indeed, although introducing indicator-based factors and their relevance assessments through the FCL were inspired by eigenbasis, this construction does not pretend to yield true eigenvectors and eigenvalues of the covariance (or correlation) matrix. In particular, correlations observed between several factors (e.g., the remuneration and sales-to-market factors) indicate that the decorrelation performed is not perfect. Although the construction of factors can be further refined to make them less correlated (e.g., by splitting the stocks into smaller groups than supersectors), it is difficult to quantitatively assess the quality of such improvements.

VII. Conclusion

We identify a new anomaly in asset pricing that is statistically significant and economically relevant. It is linked to remuneration: the more a company pays for salaries and benefits expenses per employee, the better its stock performs. We show that remuneration is a common risk factor although its magnitude appears relatively small compared with dominant factors such as low-volatility or capitalization. It also appears that only the companies that belong to extreme quantiles are sensitive to the remuneration factor. To validate the abnormal performance associated with the remuneration factor, we check that performance is not explained by other major factors such as low-volatility, capitalization, book-to-market, or momentum. This finding is an empirical contribution to the asset pricing because employee's remuneration has not been accounted for in so far, while it is a determinant element in social sciences including labor economics, sociology or management. These various strands of literature show that strong attention should be paid to wages and more generally to labor decisions that are likely to affect firms' value. The economic interpretation of our key finding is mainly based on a rational explanation of the remuneration anomaly: wages and employee performance are positively correlated. This argument is overall supported by the efficiency wage theory, which claims that rising wages is the best way to increase output per employee because it links pecuniary incentives to employee performance. But it is also supported by several studies highlighting the prominent role of operating leverage as a main source of riskiness of equity returns that is comparable in magnitude to financial leverage.

For this purpose, we introduce an original methodology, coined "Factor Correlation Level" (FCL), to build indicator-based factors. The FCL describes the ability of stocks within the factor to move in a common way and thus reflects the common risk level underpinning each factor. The FCL methodology is a theoretical contribution to the asset pricing literature. Indeed, it allows ordering the factors according to their capacity of taking into account the variability of stocks. This ranking can help fund managers to select the most important factors to set up an asset pricing model and well balanced portfolios. The FCL approach is an alternative to the common practice in asset pricing studies where factor selection depends on several statistical criteria that do not necessarily convey the same information.

Implications of this work are important, numerous and go far beyond asset pricing literature. A first investment style implication of our finding is that the companies that pay better should overperform their competitors by 2.42% per year. In other words, a market neutral investment style arbitrage strategy based on the remuneration anomaly would likely deliver positive returns. A second economics implication is that a company might operate better if it could attract the best human resources while maintaining the company as competitive as possible by keeping only those employees who are productive. While we find that a company that pays too much its shareholders, pays less to its employees according to the negative correlation between remuneration and dividend factors, attention should be brought by top managers to this trade-off between equity capital and labor remuneration. A third research implication is that our new methodology suggests the following ranking for the European stocks according to their respective FCLs: low-volatility (1.73), capitalization (1.72), momentum (1.41), sales-to-market (1.22), liquidity (1.19), book-to-market (1.13), dividend (1.09), leverage (1.07), remuneration (0.99), and cash (0.92). In particular, the low-volatility factor, which is excluded from the FF approach, is the next most important component following the market factor (i.e., the stock index). The remuneration factor is comparable to the book-to-market factor and thus not negligible. We conclude that a five factor model should encapsulate the first five anomalies ordered by their FCL.

Appendix A. Supersectors

Following the Global Industry Classification Standard (GICS), we constructed six supersectors as summarized in Table VI. This redistribution has been performed manually and has aimed at minimizing intrasector correlations and at obtaining an almost equal number of stocks in each supersector. We emphasize that final portfolios include the stocks from all supersectors, i.e., this redistribution is only an intermediate technical step to improve the factors.

Appendix B. Comparison with FF approach

In order to highlight the advantages of our methodology as compared to the standard FF approach, it is instructive to consider *incremental* transformations from one method to the other. In this way, one can analyze the respective roles of several proposed improvements. For this purpose, we implement the standard FF approach and its progressive modifications.

- A0 (the standard FF approach): According to Table I from Fama and French (2015), stocks are subdivided two groups of small (below median) and large (above median) capitalization. Within each of two groups, assets are ordered according to the chosen indicator (e.g., remuneration) and then split into three subgroups (top, medium and bottom 33%). The related portfolio is constructed by buying the top 33% and selling the bottom 33% assets from the sorted list with equal weights. Such prepared two portfolios (for small and large capitalization groups) are then merged into a single

1	Food & Staples Retailing Food, Beverage & Tobacco Health Care Equipment & Services Household & Personal Products Pharmaceuticals, Biotechnology & Life Sciences
2	Banks Diversified Financials Insurance
3	Consumer Durables & Apparel Consumer Services Media Retailing
4	Materials Real Estate
5	Energy Transportation Utilities
6	Automobiles & Components Capital Goods Commercial & Professional Services Software & Services Technology Hardware & Equipment Telecommunication Services

Table VI Six supersectors that we used to split stocks and to construct the indicator-based factors (from the FACTSET database). Note that we mixed very different industries to have 6 supersectors with approximately the same number of stocks. Even if different industries were grouped randomly into six supersectors, we show in Appendix B that our methodolody would reduce significantly the sectorial risk of different factors.

FF portfolio. To be comparable with our methodology, the portfolio is rebalanced on daily basis (note that the original FF approach stipulated monthly rebalancing). The constucted portfolio is delta-neutral.

- A1: The same rules as A0 except for buying top 15% and selling bottom 15% assets (as in our methodology);
- A2: The same rules as A1 except that the splitting into small and large capitalization groups is withdrown;
- A3: The same rules as A2 except that we add sectorial and geographical constraints as in our methodology. In other words, assets are split into 6 supersectors (see Appendix A), the portfolio construction is performed individually for each supersector and then the obtained portfolios are merged. In addition, we normalize the chosen indicator (e.g., remuneration) by the median per country to correct for geographical biases;
- A4: The same rules as A3 except that equal weights are replaced by volatility-based weights as in our methodology;
- A5: The same rules as A4 except that the volatility-based weights are rescaled by factors $\mu_\pm$ to get beta-neutral portfolios (beta's are estimated throuh a standard methodology);
- A6 (our methodology): The same rules as A5 except that a standard volatility and beta estimations (by exponential moving averages) are replaced by the reactive volatility model.

Each of these seven approaches (A0, ..., A6) has been applied to both U.K. and European universes. We computed the mean return and volatility of ten factor-based portfolios introduced in this paper. To be closer to the standard Fama and French framework, we present results on *monthly* basis, in contrast to the main text, in which daily basis was used. Table VII recapitulates the main findings for the European universe (similar results were obtained for the U.K. universe, available upon request).

As expected, the change of quantiles (passage from the standard A0 approach to A1) almost does not affect the results. Similarly, a standard volatility/beta estimator and the reactive volatility/beta model lead to similar results (passage from A5 to A6). The most significant changes are observed when passing from A2 to A3 and from A4 to A5.

• In the former case, adding the sectorial constraints (see Appendix A) reduces sectorial biases and allows one to better capture the indicator-based factors. To illustrate this point, let us suppose that remuneration is very high in the energy industry and is low (at approximately the same level) in all other industries. If there was no sectorial constraint, the remuneration factor would be long on the energy industry and short in all other industries. In other words, it would be 100% invested in energy, with eventual high risks. In turn, the

		Div.	Cap.	Low	Mom.	Liq.	Lev.	Sales.	Book.	Rem.	Cash.
A0	Mean	0.35%	-0.93%	0.30%	-0.64%	0.68%	0.02%	-0.46%	-0.33%	-0.03%	0.35%
	Std	3.20%	0.56%	4.90%	5.72%	3.92%	2.87%	3.16%	3.16%	2.04%	1.99%
	t-stat	1.46	-22.40	0.83	-1.49	2.32	0.10	-1.97	-1.40	-0.18	2.37
A1	Mean	0.37%	-0.92%	0.34%	-0.79%	0.75%	0.07%	-0.53%	-0.42%	-0.02%	0.32%
	Std	3.15%	0.44%	4.81%	5.52%	3.77%	2.80%	2.98%	3.04%	1.98%	1.97%
	t-stat	1.58	-27.98	0.95	-1.91	2.66	0.32	-2.40	-1.87	-0.12	2.17
A2	Mean	0.37%	-1.12%	-0.21%	-0.49%	0.31%	-0.23%	-0.49%	-0.27%	-0.07%	0.38%
	Std	3.41%	1.39%	4.80%	6.01%	3.85%	2.54%	3.18%	3.47%	2.00%	1.96%
	t-stat	1.45	-10.82	-0.59	-1.09	1.07	-1.20	-2.09	-1.05	-0.44	2.62
A3	Mean	0.41%	-0.96%	-0.19%	-0.61%	0.31%	-0.21%	-0.40%	-0.39%	0.00%	0.39%
	Std	2.65%	1.17%	3.85%	4.99%	3.35%	2.31%	3.05%	2.60%	1.91%	1.69%
	t-stat	2.06	-10.97	-0.68	-1.63	1.22	-1.23	-1.77	-2.03	0.02	3.11
A4	Mean	0.41%	-0.96%	-0.19%	-0.60%	0.30%	-0.21%	-0.41%	-0.40%	0.00%	0.40%
	Std	2.65%	1.17%	3.85%	4.98%	3.34%	2.31%	3.05%	2.59%	1.91%	1.68%
	t-stat	2.06	-10.97	-0.68	-1.62	1.22	-1.19	-1.79	-2.06	0.03	3.17
A5	Mean	0.41%	-1.16%	-0.86%	-0.11%	-0.34%	-0.46%	0.02%	-0.08%	0.22%	0.25%
	Std	2.09%	1.97%	1.90%	3.34%	2.37%	1.58%	1.95%	1.94%	1.53%	1.61%
	t-stat	2.61	-7.88	-6.04	-0.43	-1.94	-3.94	0.13	-0.58	1.92	2.06
A6	Mean	0.45%	-1.17%	-0.82%	-0.16%	-0.36%	-0.40%	-0.03%	-0.10%	0.19%	0.24%
	Std	2.05%	1.91%	1.94%	3.33%	2.44%	1.59%	2.06%	2.00%	1.50%	1.60%
	t-stat	2.94	-8.19	-5.63	-0.66	-2.00	-3.36	-0.22	-0.66	1.73	1.98

Table VII Progressive evaluation of factor performances with incremental transition from the FF approach (A0, top) to our methodology (A6, bottom). For each factor, we present mean monthly return (Mean) and volatility (Std), as well as their ratio (t-stat).

sectorial constraint reduces this risk by approximately 1/6 because the strong concentration on energy only remains in the 5th supersector while investments in other industries are necessarily imposed for other supersectors. For instance, if the annualized sectorial volatility is 12%, such an enforced diversification would reduce it to 2% on an annualized basis.

• In the latter case, we switch from the delta-neutral to beta-neutral portfolios, i.e., we (partly) remove correlations with the stock market index. We evoque two possible origins to rationalize the significant decrease of volatility when passing from A4 to A5. First, if we suppose that stock beta's follow a distribution with standard deviation s_β, the average aggregated beta of a random delta-neutral factor built with $2 \times 15\% \times 500 = 150$ stocks would be 0, while its standard deviation would be $2s_\beta/\sqrt{150} \approx 16\% s_\beta \approx 6\%$, where we estimated $s_\beta \approx 0.37$ from our data. As a consequence, the volatility added by the random exposure to the market index is around $6\% \times \sigma_m \approx 1.2\%$ on an annualized basis, where $\sigma_m \approx 21\%$ is the annualized volatility of the market index. Second, our construction of beta-neutral portfolio reduces their leverage to ensure Eq. (9). Consequently, smaller investments lead to smaller volatility, as compared to the Fama and French construction with a constant investment.

One also observes that volatilities of factors progressively diminish when passing from A0 to A6. This observation indicates that our modifications better withdraw other common risks and manage to concentrate on the risk of interest.

Looking more specifically to the remuneration factor, one can observe a significant increase of t-stat, from -0.18 (insignificant) to 1.73 (significant), when passing from the standard FF approach (A0) to our methodology (A6). In other words, **implementing the above improvements allowed us to level up the remuneration factor from noise to a small but significant anomaly.**

We complete this Appendix by the following general remark. The variability of results presented in Table VII indicates their dependence on a chosen data analysis method and its parameters. The methodology plays therefore the crucial role, especially when dealing with small anomalies such as remuneration. This highlights the advantage of our method that enabled to detect and quantify such small features in the market behavior. At the same time, our methodology remains robust against some changes in construction of factors, such as replacing conventional volatility estimator by reactive volatility model, using volatility renormalized weights, or changing daily to monthly returns.

REFERENCES

Abowd, J.M., Kramarz, F., and Margolis, D.M. 1999. High wage workers and high wage firms. *Econometrica* 67, 251-333.

Adams, J.S. 1963. Toward an understanding of inequity. *Journal of Abnormal and Social Psychology* 67, 422-436.

Akerlof, G.A. 1982. Labour contracts as a partial gift exchange. *Quarterly Journal of Economics* 97, 543-69.

Akerlof, G.A. and Yellen, J.L. 1990. The fair wage-effort hypothesis and unemployment. *Quarterly Journal of Economics* 105, 255-283.

Alchian, A. and Demsetz, H. 1972. Production, information costs, and economic organization. *American Economic Review* 62, 777-795.

Allez R. and Bouchaud J.-P., 2012. Eigenvector dynamics: General theory and some applications. *Physical Reviews E* 86, 046202.

Andersen T. G., Bollerslev T., Diebold F. X., and Labys P., 2000. Exchange rate returns standardized by realized volatility are (nearly) gaussian. *Multinational Finance Journal* 4, 159-179.

Ang, A., Hodrick, R., Xing, Y., and Zhang., X., 2006. The cross-section of volatility and expected returns. *Journal of Finance* 61, 259-299.

Avramov, D., Chordia, T. 2006. Asset pricing models and financial market anomalies. *Review of Financial Studies* 19, 1001-1040.

Babecky, J., Du Caju, P., Kosma, T., Lawless, M., Messina, J., and Room, T. 2012. How do European firms adjust their labour costs when nominal wages are rigid? *Labour Economics* 19, 732-801.

Baker, G. 1992. Incentive contracts and performance measurement. *Journal of Political Economy* 100, 598-614.

Bali, T.G., Cakici, N., and Fabozzi, F.J. 2013. Book-to-market and the cross-section of expected returns in international stock markets. *Journal of Portfolio Management* 39, 101-115.

Banz, R.W. 1981. The relationship between return and market value of common stocks. *Journal of Financial Economics* 9, 3-18.

Bebchuk, L., Fried, J.M., and Walker, D.I. 2002. Managerial power and rent extraction in the design of executive compensation. *University of Chicago Law Review* 69, 751-846.

Belfield, C.R., and Wei, X. 2004. Employer size-wage effects : evidence from matched employer-employee survey data in the UK. *Applied Economics* 36, 185-193.

Belo, F., Lin, X., and Bazdresch, S. 2014. Labor hiring, investment and stock return predictability in the cross section. *Journal of Political Economy* 122, 129-177.

Black, F., Jensen, M.C., and Scholes, M. 1972. The capital asset pricing model: Some empirical tests. *Studies in the Theory of Capital Markets*, New York: Praeger.

Blau, P.M. 1955. The Dynamics of bureaucracy: A study of interpersonal relations in two government agencies. (Chicago: Chicago University Press).

Bouchaud J.-P., Matacz A., and Potters M., 2001. Leverage effect in financial markets: The retarded volatility model. *Physical Review Letters* 87, 1-4.

Campbell, J.Y., and Vuolteenaho, T. 2004. Bad beta, good beta. *American Economic Review* 94, 1249-1275.

Carhart, M. 1997. On persistence in mutual fund performance. *Journal of Finance* 52, 57-82.

Carlson, M., Fisher, A., and Giammarino, R. 2004. Corporate investment and asset price dynamics: Implications for the cross-section of returns. *Journal of Finance* 59, 2577-2603.

Cederburg, S., Davies, P., O'Doherty, M. 2015. Asset-pricing anomalies at the firm level. *Journal of Econometrics* 186, 113-128.

Chen, N.F., Roll, R., and Ross, S.A. 1986. Economic forces and the stock market. *Journal of Business* 59, 383-403.

Cheng, I.-H., Hong, H. and Scheinkman, J.E.A. 2015. Yesterday's heroes: Compensation and risk at financial firms. *Journal of Finance* 70, 839-879.

Cochrane, J. 2005. Asset pricing. Revised edition. *Princeton University Press.*

Danthine, J., and Donaldson, J. 2002. Labor relations and asset pricing. *Review of Economic Studies* 69, 41-64.

Deci, E.L., Koestner, R., and Ryan, R. 1999. A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin* 125, 627-668.

Donangelo, A. 2014. Labor mobility: Implications for asset pricing. *Journal of Finance* 3, 1321-1346.

Fama, E.F., MacBeth, J.D. 1973. Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy* 81, 607-636.

Fama, E. 1980. Agency problems and the theory of the firm. *Journal of Political Economy* 88, 288-307.

Fama, E., and French, K.R. 1992. The cross-section of expected stock returns. *Journal of Finance* 47, 427-465.

Fama, E.F., and French, K.R. 1993. Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics* 33, 3-56.

Fama, E.F., and French, K.R. 1996. Multifactor explanations of asset pricing anomalies. *Journal of Finance* 51, 55-84.

Fama, E.F., and French, K.R. 1998. Value vs. Growth: The international evidence. *Journal of Finance* 53, 1975-1999.

Fama, E.F., and French, K.R. 2008. Dissecting anomalies. *Journal of Finance* 63, 1653-1678.

Fama, E.F., and French, K.R. 2012. Size, value, and momentum in international stock returns. *Journal of Financial Economics* 105, 457-472.

Fama, E.F., and French, K.R. 2015. A five factor model. *Journal of Financial Economics* 116, 1-22.

Fama, E.F., and French, K.R. 2015. Dissecting Anomalies with a Five-Factor Model. Working paper.

Favilukis, J., and Xiaoji L. 2016. Wage rigidity: A quantitative solution to several asset pricing puzzles. *Review of Financial Studies* 29, 148-192.

Fehr, E. and Falk, A. 2002. Psychological foundations of incentives. *European Economic Review* 46, 687-724.

Ferguson, M.F., and Shockley, R.L. 2003. Equilibrium anomalies. *Journal of Finance* 58, 2549-2580.

Fu, F., 2009. Idiosyncratic risk and the cross-section of expected stock returns. *Journal of Financial Economics* 91, 24-37.

Gabaix, X., and Landier, A. 2008. Why has CEO pay increased so much? *Quarterly Journal of Economics* 123, 49-100.

Gandhi, P. and Lustig, H. 2015. Size anomalies in U.S. bank stock returns. *Journal of Finance* 70, 733-768.

Gaver, J., and Gaver, K.M. 1995. Compensation policy and the investment opportunity set. *Financial Management* 24, 19-32.

Gibbons, M.R., Ross, S.A., and Shanken, J. 1989. A test of the efficiency of a given portfolio. *Econometrica* 57, 1121-1152.

Gibbons, M.R., and Murphy, K.J. 1992. Optimal incentive contracts in the presence of career concerns: Theory and evidence. *Journal of Political Economy* 100, 468-505.

Gneezy, U., Meier, S. and Rey-Biel, P. 2011. When and why incentives (don't) work to modify behavior. *Journal of Economic Perspectives* 25, 191-209.

Gomez, J-P., Priestley, R., and Zapatero, F. 2015. Labor income, relative wealth concerns, and the cross-section of stock returns. Forthcoming, *Journal of Financial and Quantitative Analysis.*

Gourio, F. 2007. Labor leverage, firms heterogeneous sensitivities to the business cycle, and the cross-section of returns. Working paper, Boston University

Graham, J.R., Li, S. and Qiu, J. 2012. Managerial attributes and executive compensation. *Review of Financial Studies* 25, 144-186.

Green, C., and Heywood, J.S. 2008. Does performance pay increase job satisfaction? *Economica* 75, 710-728.

Grinblatt, M., and Moskowitz, T. 2004. Predicting stock price movements from past returns: The role of consistency and tax-loss selling. *Journal of Financial Economics* 71, 541-579.

Hall, B.J., and Murphy, K.J. 2003. The trouble with stock options. *Journal of Economic Perspectives* 17, 49-70.

Harvey, C.R., Liu, Y., and Zhu, H. 2016. ... and the cross-section of expected returns. *Review of Financial Studies*, 29, 5-68.

Harvey, C.R., and Liu, Y. 2015. Lucky factors. Working paper, University of Duke.

Holmstrom, B. 1999. Managerial incentive problems: A dynamic perspective. *Review of Economic Studies* 66, 169-182.

Hou, K., Xue, C., and Zhang, L. 2015. Digesting anomalies: An investment approach. *Review of Financial Studies* 28, 650-705.

Idson, T.L., and Oi, W.Y. 1999. Workers are more productive in large firms. *American Economic Review* 89, 104-108.

Jegadeesh, N., and Titman, S. 1993. Returns to buying winners and selling losers: Implications for stock market efficiency. *Journal of Finance* 48, 65-91.

Jensen, M., Murphy, K.J. and Wruck, E. 2004. Remuneration: Where we've been, how we got to here, what are the problems, and how to fix them. Mimeo, Harvard University.

Jordan, B.D. and Riley, T.B. 2013. Dissecting the low volatility anomaly. Working paper, University of Kentucky.

Kuehn, L-A., Petrosky-Nadeau, N., and Zhang, L. 2013. An equilibrium asset pricing model with labor market search. Working paper, NBER.

Thierry Lallemand, T., Plasman, R., and Rycx, F. 2007. The establishment-size wage premium: Evidence from European countries. *Empirica* 34, 427-451.

Laloux L., Cizeau P., Bouchaud J.-P., and Potters M. 1999. Noise dressing of financial correlation matrices. *Physical Review Letters* 83, 1467.

Lazear, E.P. 2000. Performance pay and productivity. *American Economic Review* 90, 1346-1361.

Liu, X., and L. Zhang, 2008, Momentum profits, factor pricing, and macroeconomic risk. *Review of Financial Studies* 21, 2417-2448.

McLean, R.D. and Pontiff, J. 2015, Does academic research destroy stock return predictability? *Journal of Finance* 71, 5-32.

Monika, M., and Yashiv, E. 2007. Labor and the market value of the firm. *American Economic Review* 97, 1419-1431.

O'Byrne, S.F. and Young, S.D. 2010. What investors need to know about executive pay. *Journal of Investing* 19, 36-44.

Ochoa, M. 2013. Volatility, labor heterogeneity and asset prices. Working paper, Federal Reserve Board of Washington D.C.

Oi, W.Y. and Idson, T.L. 1999. Firm size and wages. *Handbook of labor economics* 3, 2165-2214.

Plerou V., Gopikrishnan P., Rosenow B., Amaral L. A. N., and Stanley H.E., 1999. Universal and nonuniversal properties of cross correlations in financial time series. *Physical Review Letters* 83, 1471.

Plerou V., Gopikrishnan P., Rosenow B., Amaral L. A. N., Guhr T., and Stanley H. E., 2002. Random matrix approach to cross correlations in financial data. *Physical Reviews E* 65, 066126.

Potters M., Bouchaud J.-P., and Laloux L., 2005. Financial applications of random matrix theory: Old laces and new pieces. *Acta Physica Polonica B* 36, 2767-2784.

Rynes, S., Schwab, D., and Heneman, H. 1983. The role of pay and market pay variability in job application decisions. *Organizational Behavior and Human Performance* 31, 353-364.

Rynes, S., Gerhart, B., and Minette, K. 2004. The importance of pay in employee motivation: Discrepancies between what people say and what they do. *Human Resource Management* 43, 381-394.

Santos, T., and Veronesi, P. 2006. Labor income and predictable stock returns. *Review of Financial Studies* 19, 1-44.

Schwert, G.W. 2003. Anomalies and market efficiency. Chapter 15 in Handbook of the economics of finance. eds. G. Constantinides, M. Harris, and R. M. Stulz, North-Holland, 937-972.

Shapiro, C., and Stiglitz, J.E. 1984. Equilibrium unemployment as a worker discipline device. *American Economic Review* 74, 433-444.

Simmons, R., and Forrest, D. 2004. Buying success: Relationships between team performance and wage bills in the US and European sports leagues. In *International sports economics comparisons*, ed. R. Fort and J. Fizel, 123-40. Westport: Praeger.

Smith, A. 1776. An Inquiry into the Nature and Causes of the Wealth of Nations. London: Methuen and Co., Ltd., ed. Edwin Cannan, 1904. Fifth edition.

Stiglitz, J.E. 1981. Alternative theories of wage determination and unemployment: The efficiency wage model. Working paper, University of Princeton.

Valeyre S., Grebenkov D. S., Aboura S., and Liu Q., 2013. The reactive volatility model. *Quantitative Finance* 13, 1697-1706.

Wang D., Podobnik B., Horvatic D., and Stanley H. E., 2011. Quantifying and modeling long-range cross correlations in multiple time series with applications to world stock indices. *Physical Reviews E* 83, 046121.

Winter-Ebmer, R., and Zweimuller, J. 1999. Firm-size wage differentials in Switzerland: Evidence from job-changers. *American Economic Review* 89, 89-93.

Yellen, J.L. 1984. Efficiency wage models of unemployment. *American Economic Review* 74, 200-205.

III Conclusion Générale

Les propriétés empiriques de la matrice de corrélation des rendements des actions ne sont pas bien documentées dans la littérature car elles sont noyées dans le bruit de mesure. L'originalité de la méthode que j'ai introduite permet de débruiter la matrice de corrélation en profitant de données disponibles en plus des rendements pour contraindre les vecteurs propres. Elle ouvre donc de nouvelles portes. La méthode est particulièrement adaptée aux matrices de corrélation des actions car la première valeur propre est bien plus grande que les autres et de nombreuses données financières sont disponibles ("Book", "Capitalization", "Cash Flow", etc.). Le débruitage de la matrice a permis de mettre en évidence de nouvelles propriétés importantes de la matrice de corrélation des actions:

- l'instabilité des valeurs propres et des vecteurs propres. Ces derniers sont investis en priorité sur les facteurs de risque les plus importants. L'importance d'un facteur est mesurée à travers le "FCL", notion que j'ai introduite. Le "FCL" est la variance normalisée d'un facteur de risque et correspond aussi à la moyenne pondérée des valeurs propres par les projections au carré du facteur sur les différents vecteurs propres ;
- la diffusion du logarithme des "FCL" modélisée par de simples processus d'Orstein-Uhlenbeck semble suffire pour expliquer une grande partie de la diffusion de la matrice de corrélation. Cela permet de retrouver une distribution des valeurs propres des incréments de la matrice de corrélation;
- les poids des facteurs de risque qui optimisent les "FCL" sont repartis de manière uniforme ce qui n'est pas compatible avec une distribution aléatoires des vecteurs propres. En effet on aurait à priori attendu une

distribution gaussienne des poids qui aurait été naturelle si les vecteurs propres étaient complètement aléatoires. Cela a beaucoup d'applications notamment dans la construction des portefeuilles "risk premia" qui sont devenus importants dans l'industrie de la gestion d'actifs. Ces portefeuilles, qui capturent un style donné sont construits, selon la méthode de Fama et French, avec une fonction "double Heavyside" c'est-à-dire investis à l'achat sur les top 20% et à la vente sur le bottom 20% par rapport à un critère donné ("Book", "Capitalization", "Momentum",etc.). Ces portefeuilles peuvent être optimisés avec une règle linéaire compatible avec la distribution uniforme au lieu de la "double Heavyside" de Fama et French. J'ai nommé ces portefeuilles optimaux "Fundamental Market Neutral Maximum Variance Portfolios" car ces portefeuilles capturent de manière optimale un style donné en minimisant le risque spécifique. Ils ont théoriquement un Sharpe et un "FCL" optimaux ;

- l'effet d'échelle sur les corrélations avec deux régimes:

  - aux petites échelles de temps entre quelques secondes et quelques minutes, un effet de retard de l'ensemble des actions avec un temps de relaxation de quelques minutes explique les petites autocorrélations et l'augmentation des valeurs propres avec l'échelle de temps. J'ai développé un modèle de retard et j'ai dérivé une formule simple qui décrit cette augmentation qui intègre curieusement une loi en puissance. Le modèle reproduit précisément les mesures. Aussi, on peut interpréter les corrélations entre actions comme la conséquence des interactions entre les actions par l'intermédiaire des traders ;

  - aux grandes échelles de temps entre 1 jour et plusieurs mois une faible autocorrélation est initiée par un manque de liquidité et un comportement moutonnier des acteurs. De la même façon un modèle d'autocorrélation qui inclut des tendances qui suivent un processus d'Ornstein-Ulhenbeck permet de reproduire les augmentations des valeurs propres sur des échelles de temps longues.
- l'effet de levier qui est caractérisé par l'augmentation des corrélations et de la première valeur propre avec la baisse du marché, ne se généralise pas aux autres facteurs de risque. Lorsqu'un facteur chute, son "FCL" et les valeurs propres n'augmentent pas. Cela est théoriquement intéressant économiquement dans la mesure où les facteurs de risque alternatifs ne peuvent pas avoir de risque asymétrique sur un horizon de temps long à cause de la loi des grands nombres, s'il n'y a pas d'effet de levier et ne peuvent pas justifier une prime de risque positive. En effet c'est l'effet de levier principalement avec ou sans les queues épaisses des distributions des rendements qui rend la convergence vers la distribution gaussienne très lente en maintenant l'asymétrie. Sans effet de levier les rendements des primes de risque doivent converger plus rapidement vers la distribution gaussienne.

Par ailleurs j'ai aussi étudié finement la dynamique des beta qui est la sensibilité d'une action par rapport aux variations de l'indice, qui est directement liée à la composition du premier vecteur propre de la matrice de corrélations et qui constitue le paramètre clef de risque. J'ai proposé un modèle réactif avec 3 composants intégrant l'effet de levier spécifique (lorsqu'une action sous performe, son beta augmente), l'effet de levier systématique (lorsque l'indice baisse les corrélations augmentent), l'élasticité des beta (quand la volatilité relative augmente, les beta augmentent). Les trois composants ont été calibrés et testés. J'ai testé le biais du modèle à partir de 4 stratégies "market neutre" de base et j'ai montré la supériote du modèle par rapport à une simple régression linéaire. J'ai aussi procédé à un test Monte-Carlo qui confirme la supériorité du modèle par rapport aux méthodes alternatives ("Minimum Absolute Deviation", "Trimean Quantile Regression" et "Dynamic Conditional Correlation" avec ou sans asymetrie).

Enfin j'ai présenté une application très pratique qui présente des implications concrètes pour la gestion d'entreprise en montrant empiriquement que les entreprises qui rémunèrent bien leurs employés partagent une partie significative de leur risque et ont tendance à surperformer. La finesse de methode de mesure permet d'identifier cette anomalie de marché et met en lumière les limitations de la méthode classique de Fama et French. Cette anomalie qui reste néanmoins relativement faiblement significative semble intuitive et évidente aux professionnels.

0. Bibliographie

Acharya, Viral and Lasse Pedersen (2005), "Asset pricing with liquidity risk." *Journal of Financial Economics*, 77, 375–410.

Agarwal, Vikas (2004), "Risks and portfolio decisions involving hedge funds." *Review of Financial Studies*, 17, 63–98.

Ahdida, Abdelkoddousse and Aurélien Alfonsi (2013), "A mean-reverting sde on correlation matrices." *Stochastic Processes and their Applications*, 123, 1472–1520.

Allez, Romain and Jean-Philippe Bouchaud (2012), "Eigenvector dynamics: General theory and some applications." *Phys. Rev. E*, 86, 046202.

Amihud, Yakov (2002), "Illiquidity and stock returns: cross-section and time-series effects." *Journal of Financial Markets*, 5, 31–56.

Ang, Andrew and Joseph Chen (2007), "Capm over the long run: 1926-2001." *Journal of Empirical Finance*, 14, 1–40.

Ang, Andrew, Robert Hodrick, Yuhang Xing, and Xiaoyan Zhang (2006), "The cross-section of volatility and expected returns." *Journal of Finance*, 61, 259–299.

Ang, Andrew, Robert Hodrick, Yuhang Xing, and Xiaoyan Zhang (2009), "High idiosyncratic volatility and low returns: International and further u.s. evidence." *Journal of Financial Economics*, 91, 1–23.

Ang, Andrew, Assaf A. Shtauber, and Paul C. Tetlock (2013), "Asset pricing in the dark: The cross-section of otc stocks." *Review of Financial Studies*, 26, 2985–3028.

Asness, Clifford S, Robert J Krail, and John M Liew (2001), "Do hedge funds hedge?" *The Journal of Portfolio Management*, 28, 6–19.

Asness, Clifford S., Tobias J. Moskowitz, and Lasse Pedersen (2013), "Value and momentum everywhere." *Journal of Finance*, 68, 929–985.

Baker, Malcolm, Brendan Bradley, and Ryan Taliaferro (2014), "The low-risk anomaly: A decomposition into micro and macro effects." *Financial Analysts Journal*, 70, 43–58.

Bali, Turan G., Robert F. Engle, and Yi Tang (2017), "Dynamic conditional beta is alive and well in the cross section of daily stock returns." *Management Science*, 63, 3760–3779.

Banz, Rolf W. (1981), "The relationship between return and market value of common stocks." *Journal of Financial Economics*, 9, 3–18.

Barra (1998), *Risk Model Handbook, United States Equity, Version 3 (E3).*

Bass, Robert, Scott Gladstone, and Andrew Ang (2017), "Total portfolio factor, not just asset, allocation." *The Journal of Portfolio Management*, 43, 38–53.

Bekaert, Geert and Guojun Wu (2000), "Asymmetric volatility and risk in equity markets." *Review of Financial Studies*, 13, 1–42.

Bender, Jennifer, Xiaole Sun, Ric Thomas, and Volodymyr Zdorovtsov

(2018), “The promises and pitfalls of factor timing.” *The Journal of Portfolio Management*, 44, 79–92.

Benichou, Raphael, Yves Lempérière, Emmanuel Sérié, Julien Kockelkoren, Philip Seager, Jean-Philippe Bouchaud, and Marc Potters (2017), “Agnostic risk parity: Taming known and unknown-unknowns.” *Journal of Investment Strategies*, 6, 1–12.

Benzaquen, Michael, Iacopo Mastromatteo, Zlotan Eisler, and Jean-Philippe Bouchaud (2017), “Dissecting cross-impact on stock markets: an empirical analysis.” *Journal of Statistical Mechanics: Theory and Experiment*, 2017, 023406.

Black, Fischer (1972), “Capital market equilibrium with restricted borrowing.” *The Journal of Business*, 45, 444–55.

Black, Fischer (1976), “Studies of stock price volatility changes.” *Proceedings of the 1976 Meetings of the American Statistical Association, Business and Economics Statistics Section*, 177–181.

Black, Fischer, Michael C. Jensen, and Myron Scholes (1972), “The capital asset pricing model: Some empirical tests.” *Studies in the Theory of Capital Markets, Praeger Publishers Inc.*

Blume, Marshall E (1971), “On the assessment of risk.” *Journal of Finance*, 26, 1–10.

Bollerslev, Tim, Robert Engle, and Jeffrey Wooldridge (1988), “A capital asset pricing model with time-varying covariances.” *Journal of Political Economy*, 96, 116–31.

Bouchaud, Jean-Philippe, Andrew Matacz, and Marc Potters (2001), “Leverage effect in financial markets: The retarded volatility model.” *Phys.*

Rev. Lett., 87, 228701.

Bouchaud, Jean-Phillipe and Marc Potters (2018), *The Oxford Handbook of Random Matrix Theory*, chapter Financial applications of random matrix theory: a short review. Oxford University Press.

Brandt, Michael W., Pedro Santa-Clara, and Rossen Valkanov (2009), "Parametric portfolio policies: Exploiting characteristics in the cross-section of equity returns." *Review of Financial Studies*, 22, 3411–3447.

Brown, Stephen J. (1989), "The number of factors in security returns." *The Journal of Finance*, 44, 1247–1262.

Bru, Marie-France (1991), "Wishart processes." *Journal of Theoretical Probability*, 4, 725–751.

Bun, Joel, Jean-Philippe Bouchaud, and Marc Potters (2016), "Cleaning correlation matrices." *Risk magazine*.

Bussiere, Matthieu, Marie Hoerova, and Benjamin Klaus (2015), "Commonality in hedge fund returns: Driving factors and implications." *Journal of Banking & Finance*, 54, 266–280.

Campbell, John and Ludger Hentschel (1992), "No news is good news *1: An asymmetric model of changing volatility in stock returns." *Journal of Financial Economics*, 31, 281–318.

Carhart, Mark (1997), "On persistence in mutual fund performance." *Journal of Finance*, 52, 57–82.

Cederburg, Scott and Michael S. O'Doherty (2015), "Asset-pricing anomalies at the firm level." *Journal of Econometrics*, 186, 113–128.

Chan, Louis K. C. and Josef Lakonishok (1992), "Robust measurement of beta risk." *Journal of Financial and Quantitative Analysis*, 27, 265–282.

Chen, An-Sing and Tai-Wei Zhang (2005), "The beta regime change risk premium." *Available at SSRN*.

Choueifaty, Yves and Yves Coignard (2008), "Toward maximum diversification." *The Journal of Portfolio Management*, 35, 40–51.

Christie, Andrew A. (1982), "The stochastic behavior of common stock variances: Value, leverage and interest rate effects." *Journal of Financial Economics*, 10, 407 – 432.

Ciliberti, Stefano, Emmanuel Sérié, Guillaume Simon, Yves Lemperiere, and Jean-Philippe Bouchaud (2017), "The 'size premium' in equity markets: Where is the risk?" *Available at SSRN*.

Clarke, Roger, Harindra de Silva, and Steven Thorley (2013), "Risk parity, maximum diversification,and minimum variance: An analytic perspective." *The Journal of Portfolio Management*, 39, 39–53.

Connor, Gregory (1995), "The three types of factor models: A comparison of their explanatory power." *Financial Analysts Journal*, 51, 42–46.

Connor, Gregory and Robert Korajczyk (1993), "A test for the number of factors in an approximate factor model." *Journal of Finance*, 48, 1263–91.

Cont, Rama and Jean-Philipe Bouchaud (2000), "Herd behavior and aggregate fluctuations in financial markets." *Macroeconomic Dynamics*, 4, 170–196.

Conway, Delores A and Marc R Reinganum (1988), "Stable factors in security returns: Identification using cross-validation." *Journal of Business & Economic Statistics*, 6, 1–15.

Cox, John C, Jonathan Ingersoll, and Stephen Ross (1985), "A theory of the term structure of interest rates." *Econometrica*, 53, 385–407.

Cuchiero, Christa, Damir Filipović, Eberhard Mayerhofer, and Josef Teichmann (2011), "Affine processes on positive semidefinite matrices." *Ann. Appl. Probab.*, 21, 397–463.

Da Fonseca, José, Martino Grasselli, and Claudio Tebaldi (2007), "Option pricing when correlations are stochastic: an analytical framework." *Review of Derivatives Research*, 10, 151–180.

DeJong, Douglas V. and Daniel W. Collins (1985), "Explanations for the instability of equity beta: Risk-free rate changes and leverage effects." *Journal of Financial and Quantitative Analysis*, 20, 73–94.

DeMiguel, Victor, Martin-Utrera Alberto, Francisco J. Nogales, and Raman Uppal (2017), "A portfolio perspective on the multitude of firm characteristics." *CEPR Discussion Papers.*

Dhrymes, Phoebus J, Irwin Friend, and N Bulent Gultekin (1984), "A critical reexamination of the empirical evidence on the arbitrage pricing theory." *Journal of Finance*, 39, 323–46.

Dichtl, Hubert, Wolfgang Drobetz, Harald Lohre, Carsten Rother, and Patrick Vosskamp (2018), "Optimal timing and tilting of equity factors." *Available at SSRN.*

Duan, Jin-Chuan (1995), "The garch option pricing model." *Mathematical Finance*, 5, 13–32.

Engle, Robert (1982), "Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation." *Econometrica*, 50, 987–1007.

Engle, Robert (2002), "Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models." *Journal of Business & Economic Statistics*, 20, 339–50.

Engle, Robert (2016), "Dynamic conditional beta." *Journal of Financial Econometrics*, 14, 643–667.

Epps, Thomas W. (1979), "Comovements in stock prices in the very short run." *Journal of the American Statistical Association*, 74, 291–298.

Eynard, Bertrand, Taro Kimura, and Sylvain Ribault (2015), "Random matrices." *ARXIV:1510.04430*.

Fabozzi, Frank J. and Jack Clark Francis (1978), "Beta as a random coefficient." *The Journal of Financial and Quantitative Analysis*, 13, 101–116.

Fama, Eugene F. (1965), "Random walks in stock market prices." *Financial Analysts Journal*, 21, 55–59.

Fama, Eugene F. (1968), "Risk, return and equilibrium: Some clarifying comments." *Journal of Finance*, 23, 29–40.

Fama, Eugene F. and Kenneth R. French (1992), "The cross-section of expected stock returns." *The Journal of Finance*, 47, 427–465.

Fama, Eugene F. and Kenneth R. French (1993), "Common risk factors in the returns on stocks and bonds." *Journal of Financial Economics*, 33, 3–56.

Fama, Eugene F. and Kenneth R. French (1997), "Industry costs of equity." *Journal of Financial Economics*, 43, 153 – 193.

Fama, Eugene F. and Kenneth R. French (1998), "Value versus growth: The international evidence." *The Journal of Finance*, 53, 1975–1999.

Fama, Eugene F. and Kenneth R. French (2008), "Dissecting anomalies." *Journal of Finance*, 63, 1653–1678.

Fama, Eugene F. and Kenneth R. French (2012), "Size, value, and momentum in international stock returns." *Journal of Financial Economics*, 105, 457–472.

Fama, Eugene F. and Kenneth R. French (2015), "A five-factor asset pricing model." *Journal of Financial Economics*, 116, 1 – 22.

Fama, Eugene F. and James D. MacBeth (1973), "Risk, return, and equilibrium: Empirical tests." *Journal of Political Economy*, 81, 607–36.

Ferson, Wayne E. and Stephen R. Foerster (1994), "Finite sample properties of the generalized method of moments in tests of conditional asset pricing models." *Journal of Financial Economics*, 36, 29–55.

Ferson, Wayne E. and Campbell R. Harvey (1999), "Conditioning variables and the cross section of stock returns." *The Journal of Finance*, 54, 1325–1360.

Francis, Jack Clark (1979), "Statistical analysis of risk surrogates for nyse stocks." *Journal of Financial and Quantitative Analysis*, 14, 981–997.

Frazzini, Andrea and Lasse Pedersen (2014), "Betting against beta." *Journal of Financial Economics*, 111, 1–25.

Fu, Fangjian (2009), "Idiosyncratic risk and the cross-section of expected stock returns." *Journal of Financial Economics*, 91, 24 – 37.

Fung, William and David A Hsieh (1997), "Empirical characteristics of dynamic trading strategies: The case of hedge funds." *Review of Financial Studies*, 10, 275–302.

Galai, Dan and Ronald W. Masulis (1976), "The option pricing model and the risk factor of stock." *Journal of Financial Economics*, 3, 53 – 81.

Golub, Gene H. (1973), "Some modified matrix eigenvalue problems." *SIAM Review*, 15, 318–334.

Gourieroux, Christian (2006), “Continuous time wishart process for stochastic risk.” *Econometric Reviews*, 25, 177–217.

Goyal, Amit and Pedro Santa-Clara (2003), “Idiosyncratic risk matters!” *Journal of Finance*, 58, 975–1008.

Graham, B. and D.L.F. Dodd (1934), *Security Analysis*. McGraw-Hill Education.

Grinblatt, Mark and Tobias J. Moskowitz (2004), “Predicting stock price movements from past returns: the role of consistency and tax-loss selling.” *Journal of Financial Economics*, 71, 541–579.

Guedj, Olivier and Jean-Philippe Bouchaud (2005), “Experts’ earning forecasts: Bias, herding and gossamer information.” *International Journal of Theoretical and Applied Finance*, 08, 933–946.

Harvey, Campbell R. and Yan Liu (2018), “Lucky factors.” *Available at SSRN*.

Haugen, Robert A. and Nardin L. Baker (1991), “The efficient market inefficiency of capitalization–weighted stock portfolios.” *The Journal of Portfolio Management*, 17, 35–40.

Haugen, Robert A. and A. James Heins (1975), “Risk and the rate of return on financial assets: Some old wine in new bottles.” *The Journal of Financial and Quantitative Analysis*, 10, 775–784.

Heston, Steven L (1993), “A closed-form solution for options with stochastic volatility with applications to bond and currency options.” *Review of Financial Studies*, 6, 327–43.

Hodges, Philip, Ked Hogan, Justin R. Peterson, and Andrew Ang (2017), "Factor timing with cross-sectional and time-series predictors." *The Journal of Portfolio Management*, 44, 30–43.

Hong, Harrison and David Sraer (2016), "Speculative betas." *Journal of Finance*, 71, 2095–2144.

Hong, Harrison and Jeremy Stein (1999), "A unified theory of underreaction, momentum trading, and overreaction in asset markets." *Journal of Finance*, 54, 2143–2184.

Jagannathan, Ravi and Zhenyu Wang (1996), "The conditional capm and the cross-section of expected returns." *The Journal of Finance*, 51, 3–53.

Jegadeesh, Narasimhan and Sheridan Titman (1993), "Returns to buying winners and selling losers: Implications for stock market efficiency." *Journal of Finance*, 48, 65–91.

Jobson, J. D. and Bob Korkie (1980), "Estimation for markowitz efficient portfolios." *Journal of the American Statistical Association*, 75, 544–554.

Jordan, Bradford D. and Timothy Brandon Riley (2013), "Dissecting the low volatility anomaly." *Available at SSRN*.

Kac, Marc (1959), *Probability and Related Topics in Physical Sciences*. Lectures in applied mathematics (American Mathematical Society) ; 1.A, American Mathematical Society.

Khandani, Amir E. and Andrew Lo (2011), "What happened to the quants in august 2007? evidence from factors and transactions data." *Journal of Financial Markets*, 14, 1–46.

Kyle, Albert (1985), "Continuous auctions and insider trading." *Econometrica*, 53, 1315–35.

Lakonishok, Josef, Andrei Sheifer, and Robert W. Vishny (1994), "Contrarian investment, extrapolation, and risk." *The Journal of Finance*, 49, 1541–1578.

Laloux, Laurent, Pierre Cizeau, Jean-Philippe Bouchaud, and Marc Potters (1999), "Noise dressing of financial correlation matrices." *Phys. Rev. Lett.*, 83, 1467–1470.

Ledoit, Olivier and Sandrine Péché (2011), "Eigenvectors of some large sample covariance matrix ensembles." *Probability Theory and Related Fields*, 151, 233–264.

Ledoit, Olivier and Michael Wolf (2003), "Improved estimation of the covariance matrix of stock returns with an application to portfolio selection." *Journal of Empirical Finance*, 10, 603–621.

Ledoit, Olivier and Michael Wolf (2004), "Honey, i shrunk the sample covariance matrix." *The Journal of Portfolio Management*, 30, 110–119.

Ledoit, Olivier and Michael Wolf (2012), "Nonlinear shrinkage estimation of large-dimensional covariance matrices." *The Annals of Statistics*, 40, 1024–1060.

Lee, Wai (2017), "Factors timing factors." *The Journal of Portfolio Management*, 43, 66–71.

Lettau, Martin and Sydney Ludvigson (2001), "Resurrecting the (c)capm: A cross-sectional test when risk premia are time-varying." *Journal of Political Economy*, 109, 1238–1287.

Lewellen, Jonathan and Stefan Nagel (2006), "The conditional capm does not explain asset-pricing anomalies." *Journal of Financial Economics*, 82, 289–314.

Luedecke, B. P. (1984), "An empirical investigation of arbitrage and approximate k-factor structure on large asset markets." *Dissertation, Department of Economics, University of Wisconsin.*

Lux, Thomas and Michele Marchesi (1984), "Scaling and criticality in a stochastic multi-agent model of a financial market." *Nature*, 397, 498.

Maillard, Sébastien, Thierry Roncalli, and Jérôme Teïletche (2010), "The properties of equally weighted risk contribution portfolios." *The Journal of Portfolio Management*, 36, 60–70.

Malkiel, Burton G. and Yexiao Xu (1997), "Risk and return revisited." *The Journal of Portfolio Management*, 23, 9–14.

Markowitz, Harry (1952), "Portfolio selection." *Journal of Finance*, 7, 77–91.

McCulloch, Robert and Peter Rossi (1991), "A bayesian approach to testing the arbitrage pricing theory." *Journal of Econometrics*, 49, 141–168.

McLean, R. David and Jeffrey Pontiff (2016), "Does academic research destroy stock return predictability?" *Journal of Finance*, 71, 5–32.

Meng, J. Ginger, Gang Hu, and Jushan Bai (2011), "Olive: A simple method for estimating betas when factors are measured with error." *Journal of Financial Research*, 34, 27–60.

Michard, Quentin and Jean-Philippe Bouchaud (2005), "Theory of collective opinion shifts: from smooth trends to abrupt swings." *Eur. Phys. J. B*, 47, 151–159.

Michaud, Richard O. (1989), "The markowitz optimization enigma: Is 'optimized' optimal?" *Financial Analysts Journal*, 45, 31–42.

Mitchell, Mark (2001), "Characteristics of risk and return in risk arbitrage." *Journal of Finance*, 56, 2135–2175.

Patton, Andrew (2009), "Are "market neutral" hedge funds really market neutral?" *Review of Financial Studies*, 22, 2295–2330.

Plerou, Vasiliki, Parameswaran Gopikrishnan, Bernd Rosenow, Luís A. Nunes Amaral, Thomas Guhr, and H. Eugene Stanley (2002), "Random matrix approach to cross correlations in financial data." *Phys. Rev. E*, 65, 066126.

Plerou, Vasiliki, Parameswaran Gopikrishnan, Bernd Rosenow, Luís A. Nunes Amaral, and H. Eugene Stanley (1999), "Universal and nonuniversal properties of cross correlations in financial time series." *Phys. Rev. Lett.*, 83, 1471–1474.

Possamai, Dylan and Pierre Gauthier (2011), "Prices expansion in the wishart model." *The IUP Journal of Computational Mathematics*, 4, 44 – 71.

Potters, Marc, Jean-Philippe Bouchaud, and Laurent Laloux (2005), "Financial applications of random matrix theory: Old laces and new pieces." *ACTA PHYSICA POLONICA B*, 36, 2767–2782.

Reinganum, Marc R. (1981), "Misspecification of capital asset pricing: Empirical anomalies based on earnings' yields and market values." *Journal of Financial Economics*, 9, 19–46.

Roll, Richard (1977), "A critique of the asset pricing theory's tests part i: On past and potential testability of the theory." *Journal of Financial Economics*, 4, 129–176.

Roll, Richard and Stephen Ross (1984), "A critical reexamination of the empirical evidence on the arbitrage pricing theory: A reply." *Journal of Finance*, 39, 347–50.

Roll, Richard and Stephen A Ross (1980), "An empirical investigation of the arbitrage pricing theory." *The Journal of Finance*, 35, 1073–1103.

Ross, Stephen (1976), "The arbitrage theory of capital asset pricing." *Journal of Economic Theory*, 13, 341–360.

Shanken, Jay (2015), "On the Estimation of Beta-Pricing Models." *The Review of Financial Studies*, 5, 1–33.

Sharpe, William (1963), "A simplified model for portfolio analysis." *Management Science*, 9, 277–293.

Sharpe, William (1964), "Capital asset prices: A theory of market equilibrium under conditions of risk." *Journal of Finance*, 19, 425–442.

Stein, Jeremy (2009), "Presidential address: Sophisticated investors and market efficiency." *Journal of Finance*, 64, 1517–1548.

Tofallis, Chris (2008), "Investment volatility: A critique of standard beta estimation and a simple way forward." *European Journal of Operational Research*, 187, 1358 – 1367.

Trzcinka, Charles (1986), "On the number of factors in the arbitrage pricing model." *The Journal of Finance*, 41, 347–368.

Valeyre, Sebastien (2012), "Herding behavior also applies to fundamentals." *The Hedge Fund Journal*, 91.

Valeyre, Sebastien (2018), "Time scale effect on correlation at larger horizon."

Valeyre, Sebastien, Denis Grebenkov, and Sofiane Aboura (2019a), "Emergence of correlations between securities at short time scales." *Physica A: Statistical Mechanics and its Applications*, 526, 121026.

Valeyre, Sebastien, Denis Grebenkov, and Sofiane Aboura (2019b), "The reactive beta model." *Journal of Financial Research*, 42, 71–113.

Valeyre, Sebastien, Denis Grebenkov, Sofiane Aboura, and Qian Liu (2013), "The reactive volatility model." *Quantitative Finance*, 13, 1697–1706.

Valeyre, Sebastien, Denis Grebenkov, Qian Liu, Sofiane Aboura, and Francois Bonnin (2016), "Should employers pay better their employees? an asset pricing approach." *Available at SSRN*.

Valeyre, Sebastien and Stanislav Kuperstein (2018), "Model of diffusion of the correlation between securities."

Valeyre, Sebastien, Stanislav Kuperstein, Denis Grebenkov, and Sofiane Aboura (2018), "The market neutral fundamental maximum variance portfolios."

Wang, Duan, Boris Podobnik, Davor Horvatić, and H. Eugene Stanley (2011), "Quantifying and modeling long-range cross correlations in multiple time series with applications to world stock indices." *Phys. Rev. E*, 83, 046121.

Wyart, Matthieu and Jean-Philippe Bouchaud (2007), "Self-referential behaviour, overreaction and conventions in financial markets." *Journal of Economic Behavior & Organization*, 63, 1–24.

Titre: Modélisation fine de la matrice de covariance/corrélation des actions

Résumé: Une nouvelle méthode a été mise en place pour débruiter la matrice de corrélation des rendements des actions en se basant sur une analyse par composante principale sous contrainte en exploitant les données financières. Des portefeuilles, nommés "Fundamental Maximum variance portfolios", sont construits pour capturer de manière optimale un style de risque défini par un critère financier ("Book", "Capitalization",etc.). Les vecteurs propres sous contraintes de la matrice de corrélation, qui sont des combinaisons linéaires de ces portefeuilles, sont alors étudiés. Grâce à cette méthode, plusieurs faits stylisés de la matrice ont été mis en évidence dont: i) l'augmentation des premières valeurs propres avec l'échelle de temps de 1 minute à plusieurs mois semble suivre la même loi pour toutes les valeurs propres significatives avec deux régimes; ii) une loi "universelle" semble gouverner la composition de tous les portefeuilles "Maximum variance". Ainsi selon cette loi, les poids optimaux seraient directement proportionnels au classement selon le critère financier étudié; iii) la volatilité de la volatilité des portefeuilles "Maximum Variance", qui ne sont pas orthogonaux, suffirait à expliquer une grande partie de la diffusion de la matrice de corrélation; iv) l'effet de levier (augmentation de la première valeur propre avec la baisse du marché) n'existe que pour le premier mode et ne se généralise pas aux autres facteurs de risque. L'effet de levier sur les beta, sensibilité des actions avec le "market mode", rend les poids du premier vecteur propre variables.

Mots clefs: corrélation, filtre, diagonalisation sous contrainte, modèle multifactoriel, portefeuilles optimaux, gestion d'actifs, diffusion

Discipline: Sciences Economique/ Gestion de portefeuille

Title: Refined model of the covariance/correlation matrix between securities

Summary: A new methodology has been introduced to clean the correlation matrix of single stocks returns based on a constrained principal component analysis using financial data. Portfolios were introduced, namely "Fundamental Maximum Variance Portfolios", to capture in an optimal way the risks defined by financial criteria ("Book", "Capitalization", etc.). The constrained eigenvectors of the correlation matrix, which are the linear combination of these portfolios, are then analyzed. Thanks to this methodology, several stylized patterns of the matrix were identified: i) the increase of the first eigenvalue with a time scale from 1 minute to several months seems to follow the same law for all the significant eigenvalues with 2 regimes; ii) a universal law seems to govern the weights of all the "Maximum variance" portfolios, so according to that law, the optimal weights should be proportional to the ranking based on the financial studied criteria; iii) the volatility of the volatility of the "Maximum Variance" portfolios, which are not orthogonal, could be enough to explain a large part of the diffusion of the correlation matrix; iv) the leverage effect (increase of the first eigenvalue with the decline of the stock market) occurs only for the first mode and cannot be generalized for other factors of risk. The leverage effect on the beta, which is the sensitivity of stocks with the market mode, makes variable the weights of the first eigenvector.

Key words: correlation, filter, constrained diagonalization, multi factorial model, optimal portfolios, portfolio management, diffusion

Discipline: Economics/ Portfolio Management

Centre d'Économie de l'Université Paris Nord
U.F.R Sciences Economiques et Gestion
École Doctorale ERASME
Université Paris 13 – Campus Villetaneuse
99 avenue Jean-baptiste Clément
93430 Villetaneuse