

Société de Calcul Mathématique SA

Outils d'aide à la décision

depuis 1995



Analyse critique de la réplication d'indices

en Ingénierie Financière

II : Analyse critique d'articles

Rapport rédigé par la

Société de Calcul Mathématique SA

Mai 2018

rédaction : Bernard Beauzamy, Adrien Schmitt

Dans cette seconde partie, nous faisons une analyse détaillée des principaux articles publiés dans la littérature scientifique à propos de la réplification d'indices. Tous comportent des erreurs mathématiques majeures.

I. Analyse détaillée du document Lo and Hasanhodzic

CAN HEDGE-FUND RETURNS BE REPLICATED? THE LINEAR CASE

Jasmina Hasanhodzic and Andrew W. Lo,

JOURNAL OF INVESTMENT MANAGEMENT, Vol. 5, No. 2, (2007), pp. 5–45

1. Résumé opérationnel de notre analyse

L'article vise à répondre à la question suivante : peut-on "cloner" des fonds, en les remplaçant par un ajustement linéaire (encore appelé "régression") obtenu à partir d'un petit nombre (par exemple 5 ou 6) de "facteurs" (par exemple le S&P 500 et quelques autres) ? Les avantages d'une telle substitution seraient évidents en terme de simplicité. L'article passe en revue 1610 fonds sur une longue période (février 1986 à septembre 2005) et analyse la performance des "clones", par comparaison avec celle des fonds d'origine ; la conclusion est positive, en ce sens que, selon l'article, ce "clonage" est prometteur.

Malheureusement, cet article est entièrement dépourvu de valeur scientifique ; il est également dépourvu de toute valeur prédictive pour un gestionnaire de fonds ou pour un investisseur.

Mathématiquement parlant, il utilise des techniques très simples et très anciennes : celles de l'ajustement linéaire. Mais les conclusions mathématiques déduites de cet ajustement sont loin d'être ce que croit l'article : dans bien des cas, cet ajustement est fondamentalement mauvais, tout simplement parce que la fonction initiale est loin d'être linéaire. De plus, on peut avoir le sentiment d'un très bon ajustement sans que cela corresponde à une réalité, précisément parce que le phénomène de départ n'est pas linéaire.

Donnons une illustration simple pour faire comprendre ceci : on peut admettre que, dans une ville en France, la peinture moyenne pour les hommes soit 41 et dans une ville au Portugal elle soit 40. Soit ; ceci est un fait dont l'énoncé est correct. Pourtant, si vous alimentez les magasins avec uniquement ces peintures dans chaque cas, personne ne sera content.

L'article commet une erreur mathématique très grave en essayant de normaliser des coefficients de signe quelconque (que feraient-ils si $\beta_1 = 1, \beta_2 = -1$, coefficients opposés, somme nulle ?). On s'étonne qu'un tel article ait pu être publié dans une revue à comité de lecture.

Sur le plan de la méthodologie, les méthodes "fixed weight", qui utilisent l'ensemble de l'historique, sont évidemment à proscrire : si vous voulez faire une prévision au 1^{er} janvier 2000, vous ne disposez pas des données postérieures à cette date. C'est tricher que de les incorporer. Seules les méthodes utilisant l'historique avant la date de prévision sont acceptables. Chose bizarre, l'article analyse les "fenêtres glissantes" ("rolling windows"), par exemple sur 24 mois, sans se rendre compte que l'on a le droit d'utiliser l'ensemble de l'historique antérieur.

Les erreurs les plus importantes, pour un investisseur, sont des erreurs "de conversion". Nous avons des concepts mathématiques simples : un ajustement linéaire, qui fournit des coefficients bien définis : α, β_i , etc. Vouloir les convertir en éléments techniques factuels est impropre : convertir α en "manager related" et les β_i en évaluation du risque n'est pas correct : c'est une sorte de mysticisme, qui consiste à croire que l'on a décrit la Nature à partir de deux ou trois équations linéaires simplifiées.

Quoi qu'en dise l'article, les conclusions ne sont nullement favorables. Si on se limite à l'approche "rolling window", les performances des clones sont médiocres ; on ne sait jamais quels clones auront des performances satisfaisantes, ni quand, ni pourquoi : cette approche, pour un investisseur, est entièrement dépourvu de valeur prédictive.

2. Analyse d'un exemple

Ce document prend l'exemple (hypothétique) d'un fonds appelé "*Capital Multiplication Partners*" (§2.2, page 10), qui repose sur une "dynamic asset-allocation strategy between the S&P 500 and 1-month US Treasury Bills" ; en langage clair, le gestionnaire choisit à chaque instant le meilleur des deux (c'est un "max" au sens mathématique du mot). L'article dit que ceci est bien hypothétique, parce que personne ne peut suivre le marché aussi précisément (tout dépend en vérité de l'intervalle de temps retenu). Les auteurs vont cependant essayer d'ajuster une droite au seul S&P 500, puis au seul US Treasury Bills ; le résultat est pitoyable :

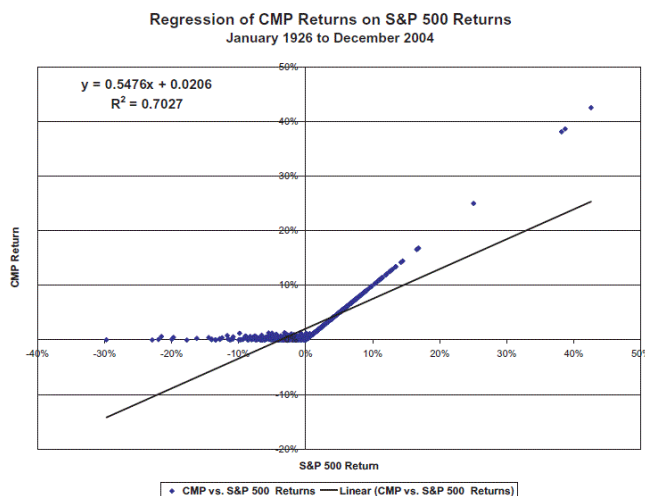


Figure 1 Scatter plot of simulated monthly returns of a perfect market-timing strategy between the S&P 500 and 1-month US Treasury bills, against monthly returns of the S&P 500, from January 1926 to December 2004.

Les auteurs concluent que "visually, it is obvious that the linear regression does not capture the essence of this inherently nonlinear strategy". Ils observent cependant que l'approximation est bonne : "the formal measure of how well the linear regression fits the data, the "R2," is 70.3% in this case, which suggests a very strong linear relationship indeed".

Il y a ici plusieurs erreurs de méthode. Tout d'abord, la régression a été faite sur des pourcentages (on ne sait pas de quoi), ce qui est formellement incorrect : l'ajustement devrait toujours être fait sur des valeurs réelles (la mesure elle-même) et non sur des pourcentages, qui vont dépendre de la valeur choisie comme référence.

On ne sait pas très bien de quels pourcentages il s'agit. Ce sont probablement des performances sur une périodicité donnée :
$$\frac{\text{prix}(t+1) - \text{prix}(t)}{\text{prix}(t)}$$
.

Ensuite, il est évident que le R^2 (souvent utilisé en finance) n'est pas une bonne mesure de la qualité de l'ajustement ; on peut avoir des R^2 proche de 1 et un très mauvais ajustement : voir Première Partie.

Une variable obtenue comme maximum de deux autres (ici le S&P 500 and 1-month US Treasury Bills) n'en dépend pas linéairement (à la différence, par exemple, de la somme).

Sans avoir compris les erreurs mathématiques commises, les auteurs concluent cependant avec bon sens : "Despite the high R2 achieved by the linear regression of CMP's returns on the market index, the actual performance of the linear clone falls far short of the strategy because a linear model will never be able to capture the option-like payoff structure of the perfect market-timer".

3. Biais de survie

Les auteurs s'inquiètent du fait que, dans la base qui sert à suivre les fonds et à construire les ajustements, certains disparaissent. A priori, ce sont les plus mauvais, mais, selon les auteurs, il arrive que des fonds qui ont du succès disparaissent aussi ("many successful funds leave the sample as well as the poor performers"). Le taux de disparition est estimé à 3% par an pour les uns, 2.4% pour les autres, ce qui est considérable sur une longue période. De surcroît, aucune analyse n'est faite sur la pondération que pouvaient avoir les fonds disparus : s'ils jouaient un rôle important dans les calculs, toute la suite en est faussée. L'affirmation "any survivorship bias should impact both funds and clones identically" n'est donc pas correcte.

L'analyse porte sur la période février 1986 à septembre 2005, ce qui est suffisamment long pour que l'on puisse utiliser les résultats. Après diverses éliminations, 1 610 fonds sont conservés et 5 catégories sont établies, représentant 78% des fonds retenus : " Long/Short Equity Hedge (520), Fund of Funds (355), Event Driven (169), Managed Futures (114), and Emerging Markets (102)."

4. Méthode suivie

Les auteurs font un ajustement linéaire ("time-series regression") de chacun des 1 610 fonds retenus, au moyen de six "facteurs" (on dirait plutôt "variables explicatives", en statistiques). Ces variables explicatives sont :

- (1) USD: the US Dollar Index return;
- (2) BOND: the return on the Lehman Corporate AA Intermediate Bond Index;
- (3) CREDIT: the spread between the Lehman BAA Corporate Bond Index and the Lehman Treasury Index;
- (4) SP500: the S&P 500 total return;
- (5) CMDTY: the Goldman Sachs Commodity Index (GSCI) total return;
- (6) DVIX: the first-difference of the end-of-month value of the CBOE Volatility Index (VIX).

Des formules mathématiques très simples sont établies :

$$E[R_{it}] = \alpha_i + \beta_{i1}E[\text{RiskFactor}_{1t}] + \dots + \beta_{iK}E[\text{RiskFactor}_{Kt}] \quad (2)$$

$$\text{Var}[R_{it}] = \beta_{i1}^2 \text{Var}[\text{RiskFactor}_{1t}] + \dots + \beta_{iK}^2 \text{Var}[\text{RiskFactor}_{Kt}] + \text{Covariances} + \text{Var}[\epsilon_{it}] \quad (3)$$

Elles sont absolument standard, et définissent les coefficients α_i comme l'ordonnée à l'origine de l'ajustement linéaire et les coefficients β_i comme la pente de cet ajustement. Il faut noter aussi que "covariances" s'entend ici au sens du coefficient de corrélation linéaire, ce qui n'est pas correct. Comme expliqué dans la Première Partie, deux variables non corrélées ne sont pas nécessairement indépendantes.

Outre cette inexactitude à caractère mathématique, qui reflète le fait que les auteurs ont une vision "linéaire" des statistiques, se pose la question de l'identification des coefficients. On peut toujours faire un ajustement linéaire : ceci est parfaitement légitime, même si les résultats sont peu probants, mais associer le coefficient α à la qualité du gestionnaire ("manager-specific alpha") est peu convaincant.

Il est complètement évident, en effet, que si l'on fait le même ajustement à partir de données légèrement différentes (par exemple en tenant compte des incertitudes, ce que l'article ne fait pas), tout sera modifié : pas seulement les coefficients de pente β_i , mais aussi les ordonnées à l'origine α . Suivant l'horizon de temps considéré (un an, deux ans, etc.), l'ajustement ne sera pas le même, donc le α ne sera pas le même, bien que le manager soit resté le même.

Associer l'un au manager et les autres au "risque factor" est donc illégitime. Nous avons ici un exemple d'identification (passage d'un élément purement mathématique à une explication "vie courante"), commune à une profession, mais complètement illégitime. Cela résulte d'un mysticisme simplificateur : nous avons écrit trois équations et elles décrivent la Nature dans toute sa complexité.

L'article se sert de cette présentation mathématique pour illustrer les diverses sources de risque : "From Eq. (3), we see that a hedge fund's variance has three distinct sources: the variances of the risk factors multiplied by the squared beta coefficients, the variance of the residual ϵ_{it} (which may be related to the specific economic sources of α_i), and the weighted covariances among the factors." Mais, là encore, c'est une explication factice, déduite d'une représentation mathématique simpliste et impropre.

L'article recense les statistiques réalisées pour l'ensemble des fonds en réalisant l'ajustement linéaire à partir des six variables explicatives ("factors") retenus. Il conclut que l'exposition au risque (représentée par les coefficients β) varie considérablement selon les catégories "Convertible Arbitrage funds have three main exposures (long credit, long bonds, and long volatility), whereas Emerging Markets funds have four somewhat different exposures (long stocks, short USD, long credit, and long commodities). The category with the smallest overall risk

exposures is Equity Market Neutral, and not surprisingly, this category exhibits the second lowest average mean return, 8.09%."

L'article mesure la qualité de l'ajustement au moyen du coefficient R^2 ; cet ajustement est toujours mauvais : "the R2 statistic of the regression (1). The mean R2s range from a low of 10.4% for EquityMarket Neutral (as expected, given this category's small average factor exposures to all six factors) to a high of 40.4% for Dedicated Short Bias (which is also expected given this category's large negative exposure to the S&P 500)."

5. Application à la répllication des fonds

Il s'agit de savoir (c'est la question posée à l'origine) si un fonds donné peut être remplacé par un ajustement linéaire obtenu à partir des six variables explicatives. L'article ne répond pas clairement à cette question ; il se borne à indiquer, dans chaque cas, la qualité de l'ajustement linéaire, ce qui n'est pas la même chose. Mais il indique cependant que "For the entire sample of 1 610 funds, 61% of the average total return is attributable to manager-specific alpha", ce qui est un aveu d'impuissance !

6. Choix de l'horizon de temps

Selon l'article, on peut avoir deux approches : ou bien chercher à réaliser l'ajustement en considérant l'ensemble des données disponibles (comme indiqué, période février 1986 à septembre 2005), ou bien se limiter à un horizon plus restreint, par exemple un an. La première version suppose qu'on connaît le futur : par exemple si on veut faire une estimation en l'an 2000, on utilise les données jusqu'en 2005, ce qui est à l'évidence incorrect (l'article appelle cela "look ahead bias").

Le sixième facteur explicatif est abandonné ("we drop the DVIX factor because its returns are not as easily realized with liquid instruments"), pour des raisons non-mathématiques mais de validité pratique. L'ordonnée à l'origine est abandonnée et les coefficients β_i sont renormalisés pour que leur somme fasse 1. Une discussion tente de justifier ces choix : "We omit the intercept because our objective is to estimate a weighted average of the factors that best replicates the fund's returns, and omitting the constant term forces the least-squares algorithm to use the factor means to fit the mean of the fund, an important feature of replicating hedge-fund expected returns with factor risk premia. And we constrain the beta coefficients to sum to one to yield a portfolio interpretation for the weights."

Ces arguments n'ont de sens que par référence aux vœux des auteurs : évaluer les risques, mais ils sont entièrement dépourvus de validité mathématique. De surcroît, les pentes sont (par essence même !) de signe quelconque ; elles peuvent fort bien être négatives. Normaliser des nombres de signe quelconque en divisant par la somme n'a aucun sens : cela ne se fait que pour des nombres positifs (calculer des pourcentages).

Une approche par "fenêtre glissante" ("rolling window") est réalisée, au moyen d'un horizon de temps de 24 mois ; l'ajustement est refait chaque mois en utilisant les données connues sur les 24 mois précédents. L'article mentionne ce défaut évident : on disposera de moins de données que si l'on utilise l'historique tout entier.

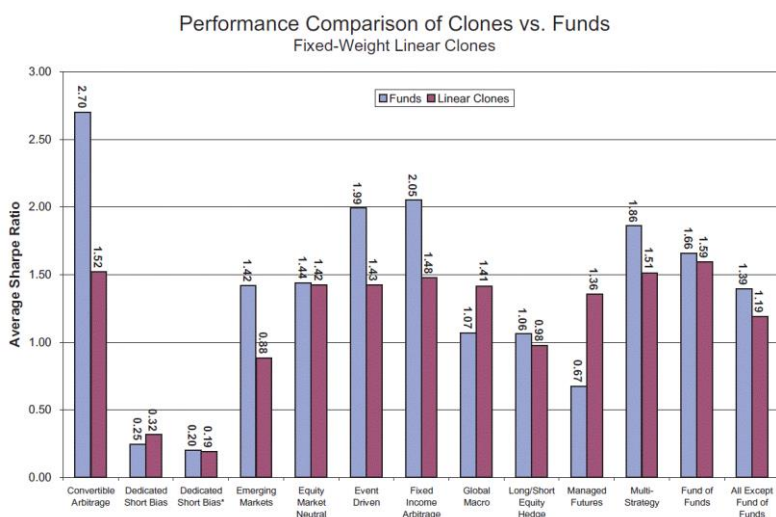
L'article ne mentionne pas l'approche qui, pourtant, aurait le meilleur sens scientifique : utiliser à chaque instant toutes les données du passé. On n'utilise pas l'ensemble de l'historique, puisque cela signifie "tricher", mais, par exemple, si l'on veut faire un pronostic au 1^{er} janvier 2000, on utilise toutes les données de février 1986 à décembre 1999, et pas seulement les 24 mois qui précèdent.

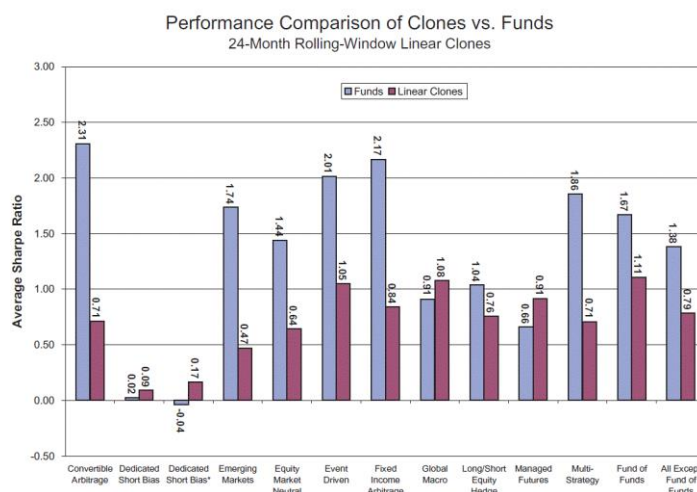
7. Analyse de performance

L'article conclut "The results are striking—for several categories, the average mean return of the clones is only slightly lower than that of their fund counterparts, and in some categories, the clones do better". Cela est très possible, mais l'approche "fixed weight", qui utilise l'ensemble de l'historique, est incorrecte. Même ainsi, certaines estimations sont mauvaises, ce qui est surprenant !

Pour l'approche "rolling window", les performances sont moins bonnes "the average performance of rolling window clones is typically lower than that of their fixed-weight counterparts". L'article estime cependant qu'elles restent satisfaisantes dans certains cas : "the performance of the rolling-window clones is all the more remarkable in categories such as Dedicated Short Bias (6.83% average mean return vs. 2.58% average mean return for the corresponding sample of funds), Equity Market Neutral (4.43% clones vs. 5.71% funds), Global Macro (12.97% clones vs. 9.01% funds), Long/Short Equity Hedge (9.08% clones vs. 11.90% funds), and Managed Futures (19.24% clones vs. 11.84% funds). Dans d'autres cas, les performances sont insatisfaisantes : "rollingwindow clones also fall short substantially in the categories of Emerging Markets (5.17% clones vs. 21.12% funds), Event Driven (6.96% clones vs. 11.65% funds), and Fixed Income Arbitrage (4.47% vs. 7.80%)".

Les résultats sont résumés dans les deux figures qui suivent :





Le résultat global, vu sur ce graphique, est loin d'être satisfaisant :

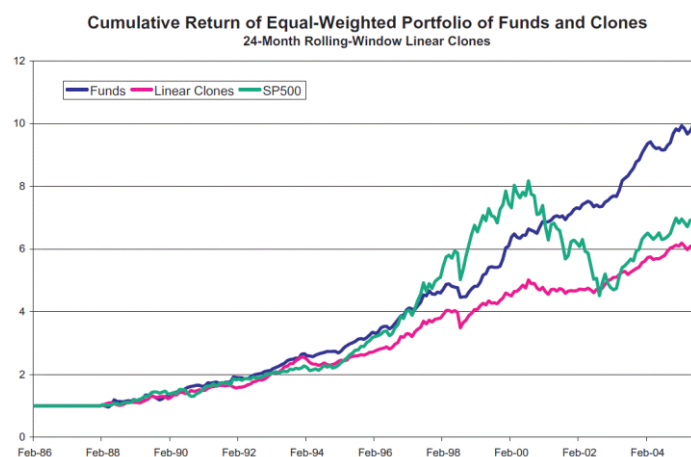


Figure 5 Cumulative returns of equal-weighted portfolios of funds and fixed-weight and 24-month rolling-window linear clones, and the S&P 500 index, from February 1986 to September 2005.

8. Liquidité

L'article introduit le coefficient de corrélation ρ_1 entre la performance sur un mois et celle sur le mois précédent (ce qui a un sens mathématique précis) et lui attribue un sens financier : "Lo (2001, 2002) and Getmansky et al. (2004) observe that positive values for ρ_1 in hedge-fund returns is a proxy for illiquidity risk." Comme nous l'avons vu précédemment, cette tentative de conversion à la vie courante est très sujette à caution. Ce coefficient de corrélation est calculé de manière linéaire, ce qui lui ôte toute valeur réelle, comme nous l'avons déjà dit. L'assertion "the clones have much lower average autocorrelations than their fund counterparts" est entièrement dépourvue de fondement pratique.

9. Conclusion

Les auteurs, pour conclure, assurent : "Our empirical findings suggest that the possibility of cloning hedge-fund returns is real". Mais les méthodes utilisées sont entièrement incorrectes et les résultats n'ont aucune valeur prédictive : on ne sait pas par avance quels seront les "clones" qui donneront des résultats satisfaisants, ni pourquoi.

II. Analyse détaillée du document Factor Modelling and Benchmarking of Hedge Funds:

Can passive investments in hedge fund strategies deliver?

Lars Jaeger and Christian Wagner, 2005

1. Résumé opérationnel

Il ne s'agit pas réellement d'une investigation scientifique ; l'article ne comporte qu'une seule formule, celle de la régression linéaire (voir article précédent). Il se veut plus "philosophique". Il commence par reconnaître (ceci est évident pour nous) qu'il n'y a pas consensus sur la signification des concepts "alpha" et "beta", mais, pour ajouter à la confusion, il introduit un troisième concept, qu'il appelle "alternative beta", et qui n'est pas mieux défini que les deux autres.

L'article fait référence à des modèles d'équilibre général, qui ne sont absolument pas appropriés dans un tel contexte (voir Première Partie), mais la référence s'arrête là : aucun modèle d'équilibre n'est utilisé dans l'article.

Les exemples choisis sont de mauvaise qualité ; l'article plaide pour la construction de "répliques", mais les résultats obtenus au moyen des régressions ne sont absolument pas convaincants : la méthode est mauvaise en soi, mais de plus elle ne produit pas de résultats !

A lire cet article, on a davantage l'impression d'une "propagande commerciale" pour un ensemble d'idées que d'une analyse scientifique objective et honnête.

2. Présentation générale

L'article annonce "investable benchmarks based on risk factor analysis and replication offers a valid, theoretically more sound, and cheaper alternative to the currently offered hedge fund index products".

3. Introduction à la discussion "alpha versus beta"

"The proponents on the one side claim that the essential part of hedge fund returns come from the funds' exposure to systematic risks, i.e. comes from their betas. Conversely, the "alpha protagonists" argue that hedge fund returns depend mostly on the specific skill of the hedge fund managers"

L'article admet (et ceci est important ici) qu'il n'y a pas de consensus de définition : "There is no consensus definition of 'alpha'". Il donne sa propre définition : "We define alpha as the part of the return that cannot be explained by the exposure to systematic risk factors in the global capital markets and is thus the return part that stems from the unique ability and skill set of the hedge fund manager". Nous observons que cette définition est extrêmement vague : même si on parvient à automatiser une gestion, il y a une part de compétence dans la définition même des algorithmes.

L'article soumet une idée intéressante en soi : "we argue that the better part of the confusion around hedge fund returns arises from the inability of conventional risk measures and theories to properly measure the diverse risk factors of hedge funds".

Mais, très vite, il ajoute à la confusion : "we are starting to realize that hedge fund beta is different from traditional beta ; hedge fund beta is more complex than traditional beta".

Et la définition donnée est très compliquée : "If the specific return is available only to a handful investors and the scheme of extracting it cannot be simply specified by a systematic process, then it is most likely real alpha. If it can be specified in a systematic way, but it involves nonconventional techniques such as short selling, leverage and the use of derivatives (techniques which are often used to specifically characterize hedge funds), then it is possibly beta, however in an alternative form, which we will refer to as "alternative beta". In the hedge fund industry "alternative beta" is often sold as alpha, but is not real alpha".

On est bien dans du marketing, de la vente, donc le but est de montrer (au moins d'affirmer) que la bonne performance provient d'un talent qui ne se trouve pas ailleurs dans le marché (ce sont les bêtas). Pour en profiter, il faut souscrire au produit proposé.

L'article voudrait donc donner des méthodes pour distinguer correctement entre alpha, "alternative beta" et "traditional beta". Il voudrait obtenir ceci sous la forme d'un "general equilibrium model". Il y a là une approche très incorrecte, mathématiquement parlant. Un modèle qui recherche l'équilibre est soumis (comme en physique) à des règles strictes, qui ne sont pas du tout celles des marchés financiers. De manière générale, rappelons (voir la Première Partie) que la Nature ne recherche jamais un équilibre (idée fautive à la base de la plupart des modèles économiques), mais procède par grandes oscillations. Ceci est particulièrement évident pour les marchés financiers : si une situation particulière apparaît (par exemple une restriction sur le pétrole, ou bien, à l'inverse, le développement de nouvelles énergies), elle donne lieu à un engouement immédiat, qui se traduit par une "bulle" ; au bout de quelques années, cette bulle explose et une autre bulle apparaît.

L'article voudrait reconnaître "that there is likely more beta than alpha in hedge funds".

4. Analyse des méthodes

Le modèle de Sharpe, 1992, repose sur une combinaison linéaire d'indices ; il est considéré comme satisfaisant. Il a été étendu aux hedge funds par Fung and Hsieh en 1997. L'équation de base est la même que précédemment :

Hedge fund excess return = Manager's alpha + $\sum (\beta_i * \text{Factor}_i)$ + random fluctuations

Fung and Hsieh ont considéré 8 classes d'indices et ont identifié 5 facteurs de risques. L'article fait observer que le résultat dépend du choix des facteurs de risque, ce qui est juste. Si on en oublie, cela se retrouvera à tort dans le "alpha".

L'article passe en revue les biais de méthode, notamment le biais de disparition de certains fonds (voir analyse précédente), et la reconstitution des valeurs ("backfilling"), mais estime que leur effet ne dépasse pas 3% à 4% : ceci par dire d'expert, sans aucune justification.

L'article se plaint de l'insuffisance de la recherche académique sur ces questions, mais nous dirons à l'inverse que la recherche académique existante va dans le mauvais sens et ne fait qu'ajouter à la confusion : elle adopte par principe des axiomes de base qui sont incorrects et que ladite recherche académique n'est pas capable de remettre en cause.

L'article annonce qu'une combinaison à poids égal de trois stratégies simples, à savoir :

1. A simple trend following model on 25 liquid futures markets summarized on what is known as the "sgfi index" (Bloomberg ticker "SGFII <Index>");
 2. The BXM index - an index defined by the Chicago Board of Trade for a simple "buy write" strategy on the S&P 500 (Bloomberg ticker "BXM <Index>");
 3. The Credit Suisse High Yield Bond Index (Bloomberg ticker "CSHY <Index>").
- donne d'excellents résultats sur une période de 11 ans : 1996 à 2005. Ces résultats sont meilleurs que ceux des indices des hedge funds "on a risk-adjusted basis". Mais, sur la figure qui suit, ce n'est guère évident :

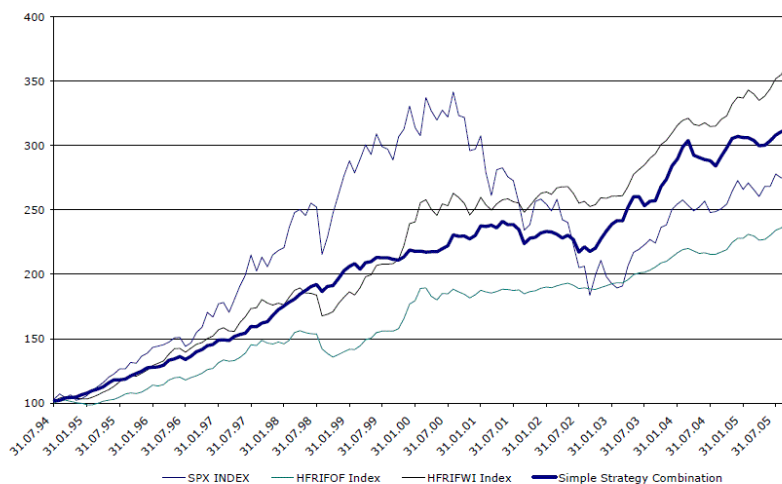


Fig. 2: Performance of an equally weighed combination of three strategies: the sgfi trend following index, the BXM covered call writing index, and long the Credit Suisse High Yield Bond Index (annualised return: 10.3%, annualised volatility: 5.6%). For comparison, we show the performance of the HFR Composite (annualised return: 11.7%, annualised volatility: 7.2%), the HFR Fund of Funds Index (annualised return: 7.9%, annualised volatility: 5.8%) and the S&P 500.

Les résultats comparent les ratios performance annuelle et volatilité annuelle. Plus ce ratio est élevé, mieux le risque (la volatilité) est rémunéré. Cette normalisation a posteriori est très utilisée pour comparer des fonds sur long terme. Mais, comme le risque est mal défini en amont, cela ne veut pas dire grand-chose. Quand on en parle, c'est en général que les performances sont médiocres...

De toute façon, à l'évidence, une telle assertion est vide de sens, mathématiquement parlant : on a choisi la combinaison a posteriori, ainsi que la période. Pour que cela soit effectif, il faudrait annoncer d'avance le choix de l'outil et vérifier ensuite les résultats.

Les régressions sont dans l'ensemble de mauvaise qualité, et les résultats sont différents selon les secteurs, comme dans l'article précédent. Cependant, l'article conclut "We estimate that up to 80% of the returns from hedge funds originate as the result of beta exposure (i.e. exposure to systematic risk factors) with the balance accounting for manager skill based alpha (or not yet identified risk factors)", sur la base des discussions précédentes. Pourtant, les résultats ne sont guère probants :

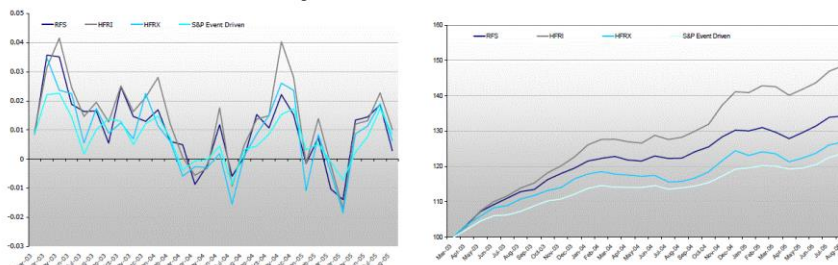


Fig. 10: Returns (monthly and cumulated) of the non-investable HFRI Event Driven Index and the investable HFRX Event Driven Index (in light color) versus the RFS cumulative return (in dark color) based on the factor returns (see text for details).

La conclusion "The key to the hedge fund 'black box' is the understanding that hedge funds generate returns primarily through risk premia and only secondarily by exploiting inefficiencies in imperfect markets" est très étrange, parce que les marchés sont par définition imparfaits (l'information n'est pas disponible, ou bien elle est biaisée) et que c'est précisément là que se situe l'analyse du risque et la performance du manager : on est amené, dans tous les cas, à prendre une décision dans un contexte d'information imparfaite. La distinction qui est faite entre "analyse du risque" et "marchés imparfaits" est donc fondamentalement incorrecte.

III. Analyse détaillée du document Hedge Fund Benchmarks: A Risk Based Approach

By William Fung & David A. Hsieh, March 2004

1. Résumé opérationnel

L'intérêt principal de cet article est qu'il tente une validation du modèle : reconstituer des données d'une année (2003), à partir des informations antérieures et comparer aux résultats réels. En moyenne annuelle, la prédiction est satisfaisante, mais cela n'est pas suffisant pour qualifier la méthode.

2. Analyse de l'article

Cet article date de 2004 et son argument principal "les données n'existent pas" n'est plus d'actualité. La discussion porte, comme dans les articles précédents, sur la part respective des "alpha" et "beta" dans des formules de régression, dont la validité n'est pas mise en cause.

L'article prétend faire usage d'outils sophistiqués, comme le "filtre de Kalman" ; la justification donnée est que "the information content of the return data is likely to decline with the age of the return observations". Il semble au contraire que la qualité de l'information s'améliore

avec le temps et, de toute façon, nous ne voyons pas le lien avec le filtre de Kalman, dont l'utilisation n'est pas détaillée.

Les analyses ont toujours un caractère rétroactif : "This led us to identify March 2000, and September 1998 as triggering market events—they are, respectively the end of the Internet bubble, and the LTCM debacle. Accordingly, we divided the sample into two sub periods—January 1994 to September 1998, and April 2000 to Dec 2002". En procédant de la sorte, on peut effectivement démontrer n'importe quoi.

L'article observe que le alpha peut varier d'une période à l'autre : "Perhaps the most troublesome change is in the intercept, or alpha, term of the regressions. For the first sub period, both the magnitude and significance level dropped dramatically for the intercept term, indicating little to no added value from the average fund-of-hedge fund manager beyond systematic bets". Ceci devrait conduire à remettre en cause le dogme "alpha = performance propre au manager", puisque le manager est en principe le même. Mais cette analyse n'est pas faite.

La plupart des managers sont très irréguliers en pratique dans la création de surperformance. Il suffit de regarder la variabilité des classements de leur fonds pour s'en apercevoir.

Nous constatons une erreur sur la signification du coefficient de corrélation : "the HFR Composite Index ("HFRI") and CSFB/Tremont Composite Index ("CTI") are not highly correlated—the correlation coefficient is only 0.76". Mais, comme nous l'avons déjà dit, il s'agit du coefficient de corrélation linéaire. Deux variables peuvent être absolument corrélées sans avoir un bon coefficient de corrélation linéaire.

3. Tentative de validation du modèle

L'article tente d'appliquer ses méthodes pour prédire les résultats de 2003 à partir des données des années antérieures, ce qui est une approche satisfaisante. Les résultats ne sont pas mauvais à l'échelle de l'année :

Table 5
Conditional Forecasts of Various Hedge Fund Indices
Annual Return In 2003

	HFR	CTI	MSCI	SPHF
Actual	18.10%	14.48%	14.05%	10.58%
Predicted	18.17%	12.13%	14.84%	10.72%

HFRI: HFR equally-weighted composite index.

CTI: CSFB/Tremont asset-weighted composite index.

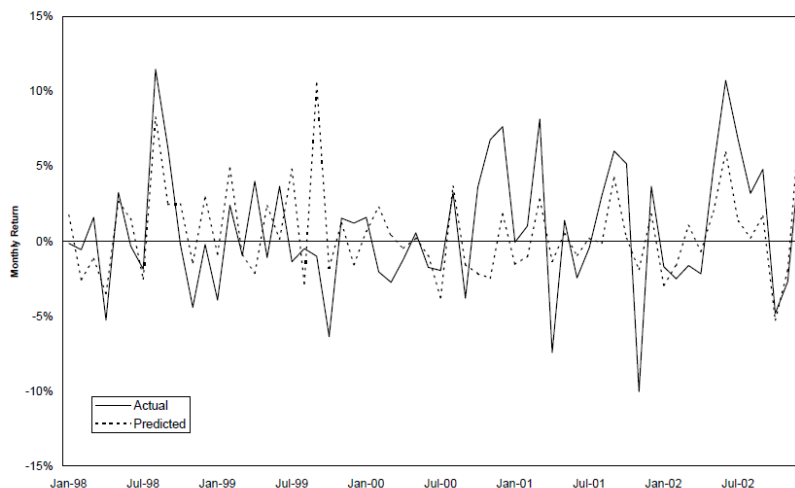
MSCI: MSCI equally-weighted composite index.

SPHF: S&P hedge fund composite index.

mais il faudrait réitérer l'expérience sur d'autres périodes. Une période calme se prête facilement à la prospective.

Une reconstruction à l'échelle du mois ne donne pas de bons résultats :

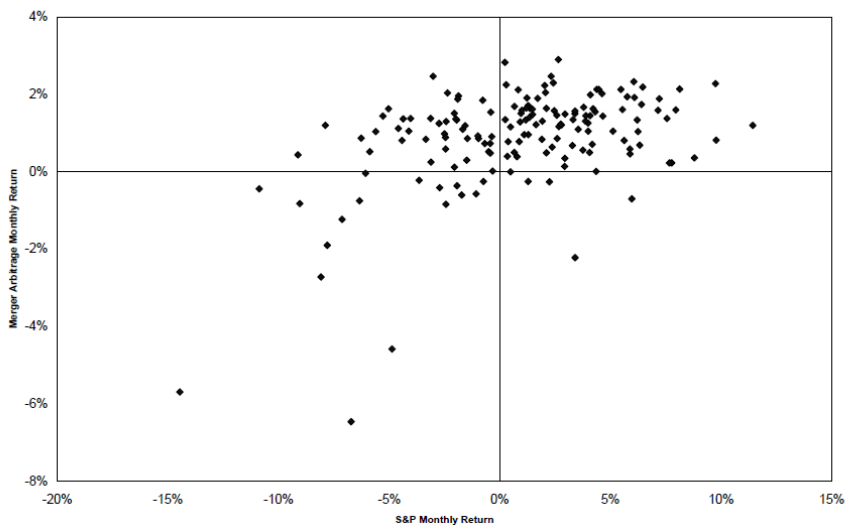
Figure 1: Trend Followers' Average Return (Actual Vs Predicted)



L'article s'interroge sur la nécessité de prendre en compte davantage de "risk factors" ; mais cela remet en cause le principe de simplicité qui est la base même de l'approche de réplcation. Il apparaît que passer de 5 à 10 ne change rien à la validité des résultats.

Un exemple de nuage de points d'où l'article déduit des corrélations :

Figure 2: Merger Arbitrage Vs S&P (Jan 1990-Dec 2002)



IV. Analyse du document How do hedge fund clones manage the real world?

Erik Wallerstein, Nils S. Tuchschild, Sassan Zaker, 2009

L'article analyse les performances de 20 produits de réplcation, obtenus par trois techniques différentes, offerts par 17 institutions.

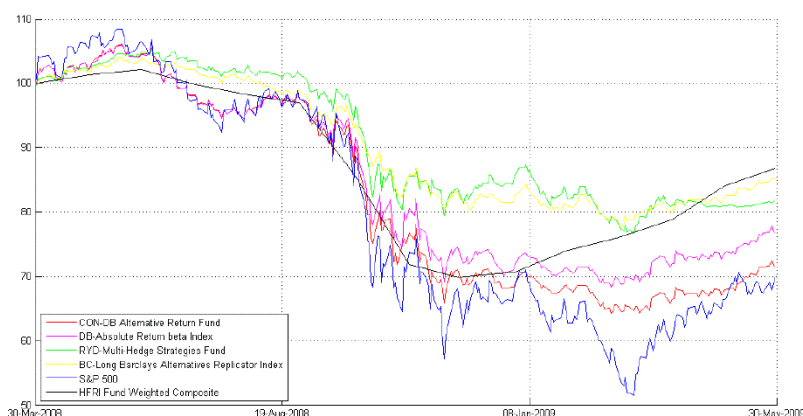
- Première approche : "linear regression analysis of hedge fund returns" ;

- Seconde approche : "dynamic trading techniques to replicate a pay-off, with similar dependence structures as between a hedge fund and an appropriate choice of an investor's portfolio" ;
- Troisième approche : reverse engineering.

L'article conclut que la principale difficulté pour la première approche est l'insuffisance de données. Nous ne partageons pas cette conclusion : les données sont nombreuses ; c'est la théorie qui est défailante.

La période d'analyse de la performance est mars 2008 à mai 2009 (période étonnante, car elle se situe en pleine crise financière des subprimes). L'article conclut qu'il n'existe pas de raison de préférer l'une des trois techniques. Les produits proposés par les grandes institutions ne se révèlent pas meilleurs que ceux proposés par les petites.

Le résultat ne se révèle guère probant : "All of them fall short of delivering absolute returns, but neither did hedge funds over the sample period". L'explication donnée est : "It is first of all relatively short and more importantly it only covers a recessionary economic cycle. If replication product is a conservative asset in general it is little surprise that they have outperformed hedge funds over this sample period".



Retours cumulés des indices répliqués, par comparaison avec S&P

V. Analyse de l'article **Performance of Passive Hedge Fund Replication Strategies, 2009**,

by Noël Amenc, Lionel Martellini, Jean-Christophe Meyfredi, Volker Ziemann

Le résumé est clair : utiliser des modèles non-linéaires n'améliore pas la performance ; les facteurs explicatifs doivent être choisis sur critères économiques ; la performance des "clones" est systématiquement inférieure à celle des modèles d'origine.

L'article ne remet pas en cause la pertinence d'un ajustement linéaire. Il reprend les méthodes de Hasanhodzic and Lo, y compris la normalisation par la somme de nombres non nécessairement positifs (voir plus haut).

Dans le paragraphe Markov Regime-Switching Model, les auteurs font l'hypothèse que les résidus ont une forme gaussienne, hypothèse purement académique. Il en va de même pour l'introduction, au paragraphe suivant, des filtres de Kalman. On est ici en présence de modèles purement académiques, sans aucune validité pratique.

Une analyse de performance est réalisée ; la période de calibration est janvier 1997 à décembre 1998. On utilise une fenêtre glissante de 24 mois avec une réévaluation mensuelle. L'article conclut : "These results clearly draw a less optimistic picture of the replication quality of economic factor models. Indeed, with the exception of Emerging Market clones, most other clones perform much less well than the full factor model clones".

VI. Analyse de l'article WHO NEEDS HEDGE FUNDS?

A Copula-Based Approach to Hedge Fund Return Replication,

Harry M. Kat, Helder P. Palaro, 2005.

1. Résumé opérationnel

Cet article est totalement sans intérêt. Il est purement académique. Il part du principe que l'on cherche à cloner les propriétés statistiques d'un fonds (ce qui est évidemment absurde) et construit des théories mathématiques reposant sur des hypothèses purement académiques (lois normales, etc.).

2. Analyse des méthodes

L'article dit : "We do not necessarily have to replicate a fund's month-to month returns. For most applications it will be enough if we can generate returns with the same statistical properties as the returns generated by the fund". Non, certainement pas : l'investisseur se moque bien des propriétés statistiques ; ce qu'il veut, c'est gagner de l'argent : voir l'introduction du présent rapport.

Il s'agit, dit l'article de répliquer le profil de risque d'un fonds, sans considérer explicitement le retour attendu : "The replication procedure concentrates on replicating a fund's risk profile without explicitly considering the fund's expected return". Nous ne sommes certainement pas d'accord avec cette approche.

Dans la procédure de réplification, trois lois marginales sont prises en compte : normale, Student-t and Johnson SU. Mais ceci est une approche purement académique. Dans la réalité, aucune de ces lois n'est appropriée.

VII. Analyse de l'article : Quels indices pour la gestion alternative?

Document Edhec, Noël Amenc et Mathieu Vaissié, 2003

Cet article voudrait se donner une apparence mathématique :

"Théorème 1 : Un indice des indices est toujours plus représentatif que l'ensemble des indices qui le composent"

qui est appuyée par des formules :

$$I_p = \left(\bigcap_{i=1, \dots, n} I_i \right) \cup \left(\overline{\bigcap_{i=1, \dots, n} I_i} \right) = \bigcup_{i=1, \dots, n} I_i$$

Soit $\text{Card}(I_i)$ le nombre de fonds contenus dans la base de données de l'indice I_i

$$\text{Card}(I_p) = \text{Card}\left(\bigcap_{i=1, \dots, n} I_i\right) + \text{Card}\left(\overline{\bigcap_{i=1, \dots, n} I_i}\right) = \text{Card}\left(\bigcup_{i=1, \dots, n} I_i\right) \geq \max_{i=1, \dots, n} (\text{card}(I_i))$$

entièrement dépourvues de contenu. Ce sont des formules de logique formelle, qui n'ont aucun sens ici. Cela revient à dire : connaître la peinture moyenne d'un ensemble d'individus est plus représentatif que connaître la population pour chaque peinture. Allez dire cela au gérant d'un magasin de chaussures !

Le théorème 2 n'améliore pas les choses :

Théorème 2 : Un indice des indices est toujours moins biaisé que la moyenne des indices qui le composent

$$\sigma_{\varepsilon_p}^2 \leq \sum_{i=1}^n w_i \sigma_{\varepsilon_i}^2$$

En effet, **puisque les biais des indices concurrents sont supposés indépendants** :

$$\sigma_{\varepsilon_p}^2 = w' \cdot \sum_{\varepsilon} \cdot w = \sum_{i=1}^n w_i^2 \sigma_{\varepsilon_i}^2$$

Où w est le vecteur des poids du portefeuille, i.e. $w = (w_1, \dots, w_n)$. Nous obtenons alors que :

$$\sum_{i=1}^n w_i^2 \sigma_{\varepsilon_i}^2 \leq \sum_{i=1}^n w_i \sigma_{\varepsilon_i}^2$$

On lit bien en rouge une hypothèse d'indépendance qui n'est absolument pas réalisée en pratique.

Les auteurs utilisent une Analyse en Composante Principale (ACP), ce qui est commun, mais recèle un biais méthodologique : c'est une projection linéaire. Aucun résultat financier n'est présenté, qui illustrerait la validité de l'approche.