JOURNAL DE LA SOCIÉTÉ STATISTIQUE DE PARIS

J.-F. BOULIER
H. SUEUR
E. TRUSSANT

Du mauvais usage des palmarès pour classer les gérants de fonds collectifs

Journal de la société statistique de Paris, tome 134, nº 4 (1993), p. 17-24
<http://www.numdam.org/item?id=JSFS_1993__134_4_17_0>

DU MAUVAIS USAGE DES PALMARÈS POUR CLASSER LES GÉRANTS DE FONDS COLLECTIFS

par J.-F. BOULIER
Directeur de la Recherche et de l'Innovation, CCF
par H. SUEUR
Ingénieur Financier, Direction de la Recherche et de l'Innovation, CCF
par E. TRUSSANT
Ingénieur Financier, Direction de la Recherche et de l'Innovation, CCF

Un statisticien averti en vaut deux ! Si vous cherchez à placer votre épargne dans un fonds d'investissement évitez de vous précipiter sur celui qui arrive en tête d'un des nombreux classements de la presse spécialisée tout particulièrement en fin d'année. Non pas qu'il soit a priori plus mauvais que tous les autres, mais plutôt parce que le fait que ce fonds soit premier relève par trop du hasard. Ce que vise à démontrer cet article à finalité tout à fait pratique, c'est que les « palmarès » de performance de fonds d'investissement apportent peu d'information, voire pas du tout.

Enjeux des mesures de performance

La gestion collective s'est considérablement développée ces dernières années. G. Gallais-Harmonno [1] trace dans son ouvrage un panorama des SICAV et Fonds Communs de Placement dont la croissance en nombre et encours a été spectaculaire. Cette « industrie » française, première en Europe, est particulièrement compétitive et s'adresse à tout public, depuis les personnes privées jusqu'aux institutions financières en passant par les entreprises. Qui n'a pas au moins une fois acheté un fonds monétaire ? Cette industrie vend bien sûr de la performance financière : une croissance régulière et sans heurt pour les fonds monétaires, une croissance en moyenne plus élevée mais également beaucoup plus erratique dans le cas d'un fonds investi en actions et donc soumis aux soubresauts de la Bourse. Car ce que nous enseigne l'expérience, c'est qu'il n'y a pas de rentabilité sans risque.

Cette loi d'équilibre a été formalisée par la théorie financière initiée il y a plus de trente ans par H. Markowitz [2] puis développée par W. Sharpe [3] et bien d'autres. Dans un monde où l'information est accessible à tous et où les agents en grand nombre peuvent librement échanger sur les marchés financiers, il ne peut y avoir sur une période donnée de placement de rentabilité r_p (en % annuel) plus élevée que celle r_0 d'un titre obligataire remboursé en fin de période, sans qu'un risque additionnel soit pris – c'est-à-dire qu'une incertitude soit attachée à cette rentabilité. En effet, si tel était le cas il y aurait « arbitrage », un agent emprunterait au taux r_0 et placerait à r_p ; En fin de période, il gagnerait après avoir remboursé sa dette $r_p - r_0$; Tout professionnel informé ne saurait laisser passer une telle opportunité. En d'autres termes, la théorie nous dit que les marchés financiers sont efficients. Par conséquent, si l'espérance de r_p est supérieure à r_0, l'écart type annualisé de r_p, que l'on dénomme volatilité v_p, n'est pas nul. En fait, $E(r_p) - r_0$ est une fonction croissante de v_p. Dans le modèle de Sharpe, où r_M est la rentabilité du « portefeuille de marché », cette fonction est une droite :

$$E(r_p) - r_0 = \beta\left(E(r_m) - r_0\right)$$

Dans les faits, la prime de risque $E(r_m) - r_0$ vaut quelques 3 % et le niveau de volatilité du portefeuille de marché avoisine 20 %. On gagne ainsi 0,15 % de plus par unité de risque pris, ce qui est relativement faible. Investir en bourse, c'est risqué, mais à la différence du casino la rentabilité est tout de même positive.

La mesure de la rentabilité des véhicules d'investissement est à l'évidence une question très importante. Mais la performance réalisée entre deux dates du passé ne donne qu'un aperçu des caractéristiques du fonds. Combien d'investisseurs en effet sont entrés et sortis à ces dates ? Force est de constater que les performances réelles vont dépendre de ces dates. À supposer que les caractéristiques du fonds soient stationnaires, et que l'on décale la période de mesure d'un semestre, l'écart type de la différence des rentabilités est $v_p/\sqrt{2}$, soit quelques 14 % dans le cas d'un fonds investi en actions ! Pour pouvoir comparer des fonds qui sont investis dans des actifs financiers de même nature ou dont les compositions en classe d'actifs sont en moyenne comparables, la Commission des Opérations de Bourse qui contrôle l'information et donne son agrément pour ces fonds vient de modifier ces catégories en retenant le principe d'une référence au risque (v_p) du fonds. Le lecteur intéressé par les méthodes de mesure de performance trouvera dans [4] un exposé plus complet et des illustrations de ces méthodes appliquées aux SICAV.

Sur le marché nord américain les mesures de performance sont réalisées par des sociétés indépendantes qui utilisent en général une forme de classement. Les fonds sont comparés dans un « univers » de véhicules aux caractéristiques communes, éventuellement référencées par un indice (action, obligation, international...). Les performances des fonds sont alors comparées sur plusieurs périodes, typiquement 3 mois, 1 an, 2 ans, 5 ans pour des fonds d'actions, par rapport aux quartiles de la distribution empirique des fonds concurrents. L'information diffusée est donc : dans quel quartile un fonds s'est-il situé ?

En France, l'usage des classements de 1 à n, nombre de fonds concurrents, ou encore de palmarès s'est répandu depuis le début des années 80. L'amplification médiatique est telle que les lauréats sont portés aux nues et les derniers voués à la réprobation des réseaux qui les distribuent.

Au risque d'enfoncer quelques portes ouvertes, nous allons dans les sections suivantes montrer que l'information contenue dans le palmarès est très faible. Après avoir rappelé quelques indicateurs de base pour mesurer la performance, nous considérerons un palmarès de deux fonds et nous calculerons la probabilité que le palmarès désigne par erreur ou malchance un « mauvais » gérant en compétition avec un « bon » gérant. Dans une deuxième étape, une compétition plus réaliste avec un plus grand nombre de gérants sera considérée – les calculs relatifs aux lois de probabilité seront exposés en annexe pour ne pas alourdir l'exposé.

Quelques outils fiables de mesure de performance

Nous avons vu en introduction l'importance, dans toute démarche d'évaluation de performance, de prendre en compte l'exposition au risque du portefeuille. Tout niveau de rentabilité atteint est à relativiser par le risque encouru. Le ratio de Sharpe, proposé par celui-ci en 1966, répond à ce souci. Défini par le rapport :

$$S_p = \frac{E\,[R_p] - R_f}{\sigma\,[R_p]}$$

il est égal à la prime de risque du portefeuille P (différence entre l'espérance ou moyenne sur la période d'étude de ses rentabilités logarithmiques et le taux sans risque) divisée par la volatilité de ces mêmes rentabilités. La performance est d'autant meilleure que ce ratio est élevé.

Lorsque l'on raisonne en comparaison à un portefeuille de référence ou benchmark (mesure de performance relative par opposition à une mesure absolue), on utilise alors son homologue le ratio d'information :

$$RI = \frac{E\,[R_p - R_b]}{\sigma\,[R_p - R_b]}$$

Au numérateur figure la sous- ou sur-performance obtenue par le gérant relativement à son benchmark. Au dénominateur apparaît l'écart type annualisé des différences de rentabilités $r_p - r_b$, encore appelé tracking-error. C'est l'équivalent de la volatilité dans un cadre relatif.

En termes d'ordre de grandeur, le ratio d'information est bien entendu avant tout à comparer à 0. Négatif, il condamne la mauvaise politique d'allocation suivie, les objectifs n'ayant pas été atteints. Positif, il est à comparer à 0,5. En effet, on considère classiquement pour une gestion de grande qualité que deux points de risque pris par rapport au benchmark doivent générer une plus-value d'un point. Compris entre 0 et 0,5, ce ratio est par conséquent le reflet d'une gestion sans doute quelque peu trop risquée étant donnée la sur-performance obtenue. En revanche, supérieur à 0,5, il consacre l'excellente politique du gestionnaire, caractérisée par une plus-value élevée associée à une limitation efficace du risque relatif.

Ces deux paramètres constituent des indicateurs fiables d'évaluation de performance, car ils s'intéressent non seulement à la rentabilité obtenue mais également au risque encouru. Ils ont de plus l'avantage de présenter une certaine stabilité aux variations soit de la périodicité des données utilisées (rentabilités quotidiennes, hebdomadaires ou mensuelles) soit de la période choisie pour leur estimation.

Premier cas : compétition entre deux gérants

Supposons que deux gérants aient les mêmes objectifs et contraintes. Ceux-ci sont résumés dans une structure identique de portefeuille de référence, le benchmark. C'est le cas par exemple pour deux portefeuilles indiciels, sensés répliquer le CAC40. Faisons également l'hypothèse que tous deux exposent leur fonds au même risque relatif par rapport à ce benchmark commun (même traking-error). En revanche, leurs sur-performances x (respectivement x') sont différentes. Le premier a ainsi un ratio d'information de + 0,5 tandis que le second a un ratio de – 0,5. C'est à dessein que nous avons choisi deux cas extrêmes afin de rendre la démonstration encore plus parlante.

Ces deux plus-values x et x' sont supposées suivre des lois normales de même écart-type et non corrélées :

$$x \sim N(R,\ \sigma) \quad \text{avec} \quad \frac{R}{\sigma} = +\,0{,}5$$

$$x' \sim N(R',\ \sigma) \quad \text{avec} \quad \frac{R'}{\sigma} = -\,0{,}5$$

Dans cette modélisation, la seule variable qui diffère d'un gérant à l'autre est l'espérance ou moyenne de leur rentabilité relative.

L'hypothèse de non corrélation peut s'appréhender intuitivement en considérant par exemple que l'un s'efforce de générer des plus-values par rapport au benchmark en vendant des Futures sur CAC40 alors que l'autre est un « stock picker » (il recherche les titres temporairement sous-évalués).

Quelle est alors la probabilité pour qu'un palmarès, fondé sur le seul critère d'un taux de rentabilité estimé entre deux dates arbitraires, se trompe et désigne le 2[e] gérant lauréat ? Nous cherchons donc à évaluer $P\,(x' > x)$.

Un simple calcul de probabilité montre que $P(x' > x) = 0{,}24$. C'est dire que l'on a une chance sur quatre de se tromper.

Dans un cadre multipériodique, le résultat va dépendre fortement de la période utilisée pour l'établissement du classement. En effet si T désigne le nombre d'années (ou de fractions d'années) sur lequel porte l'étude, la sur-performance suit une loi normale

$$x_T \sim N\left(RT,\ \sigma\sqrt{T}\right)$$

Le tableau ci-dessous répertorie, en fonction du temps, la probabilité d'erreur exprimée en % :

Période d'étude	1 jour	1 sem	1 mois	1 trim	1 an	2 ans	5 ans	10 ans
Probabilité	48,52	46,09	41,91	36,18	23,97	15,87	5,69	1,27

Dans un deuxième temps, l'hypothèse d'indépendance des deux gérants est relaxée. On suppose donc dorénavant que la corrélation entre leurs sur-performances est non nulle. Il semble bien difficile de présumer intuitivement de son signe. Alors que la non-corrélation s'explique aisément par des politiques de gestion différentes (market-timing, stock picking, utilisation de produits dérivés, ...), on se rend compte que deux éléments conduisent à des conclusions contradictoires : d'une part, les gérants de portefeuilles, ayant tous accès à la même information, ont des anticipations voisines sur les marchés et élaborent fréquemment des stratégies similaires ou du moins assez proches ; d'autre part, les deux fonds étudiés ayant des ratios d'information opposés, leurs plus-values par rapport au benchmark commun ont dû évoluer différemment au cours du temps.

Le tableau suivant, à double entrée, montre dans quelle mesure varient les probabilités de se tromper (en %) :

Corrélations	1 jour	1 sem	1 mois	1 trim	1 an	2 ans	5 ans	10 ans
– 0,3	48,71	46,57	42,90	37,82	26,76	19,02	8,28	2,49
0	48,52	46,09	41,91	36,18	23,97	15,87	5,69	1,27
+ 0,3	48,24	45,34	40,36	33,63	19,90	11,60	2,94	0,38

Il est intéressant de constater que plus la corrélation augmente entre les gérants plus les palmarès, établis à partir d'une seule rentabilité estimée entre deux dates arbitraires, réussissent à discerner le meilleur gérant du moins bon. En effet, le gérant n° 1 ayant en moyenne une rentabilité supérieure à celle du gérant n° 2, s'ils sont corrélés positivement, le nombre d'occurences d'un scénario tel que $x_1 > x_2$ augmente. Toutefois, on voit que les écarts restent faibles et qu'ils ne sont sensibles qu'à longue échéance.

Second cas : extension à *n* gérants

Lorsque l'on s'intéresse non plus à deux gérants mais à davantage (5, 10, 20 voire 100), on peut facilement imaginer que les conclusions seront accentuées et la démonstration encore plus convaincante. Examinons donc maintenant un éventail de n gérants parmi lesquels un seul possède un ration d'information non nul. Fixé à 0,5, il est révélateur d'une bonne gestion sur la période considérée. Tous les autres n'ont généré aucune sur-performance relativement au benchmark. Ils ont donc des ratios d'information nuls. Nous sommes par conséquent en présence de 1 bon et de $n - 1$ autres gérants. On peut supposer, sans nuire à la généralité du problème, que tous (le bon plus les autres) ont le même risque relatif ou tracking-error.

Supposant que les rentabilités suivent des lois normales indépendantes, la modélisation est ainsi la suivante :

$$x_1 \sim N(R,\ \sigma) \quad \text{avec} \quad \frac{R}{\sigma} = 0{,}5$$

$$x_i \sim N(0,\ \sigma) \quad \forall i = 2,\ 3,\ \ldots,\ n$$

Ce qui nous intéresse, ici aussi, est d'estimer la probabilité qu'un palmarès fondé sur la seule rentabilité déclare lauréat l'un des $n - 1$ gérants médiocres. On cherche donc à estimer $P\left(x_1 < \max\limits_{i = 2 \ldots n} (x_i)\right)$

Nous nous bornons dans cette section au cas d'indépendance entre les plus-values des différents portefeuilles. L'étude à deux gérants avait en effet conduit à la conclusion que les résultats n'étaient que peu altérés à court terme par l'existence d'une corrélation, qu'elle soit positive ou négative.

Nous avons calculé la probabilité exprimée en % pour qu'un palmarès décerne à tort la palme à l'un des $n - 1$ médiocres gérants et ce en fonction à la fois du nombre de gérants (n) et de la durée sur laquelle porte le plalmarès (T) :

n / T	Jour	Semaine	Mois	Trimestre	1 an	2 ans	5 ans	10 ans
5	79	78	76	74	67	60	47	32
10	90	89	88	86	80	75	62	46
20	94,8	94,3	93,5	92,2	88	84	74	59
50	97,8	97,7	97,3	96,6	94,4	91,9	85	72
100	98,9	98,8	98,6	98,2	96,8	95,2	90	80

Les gérants sont représentés par des variables aléatoires normales indépendantes et d'écart type annuel 5 %. Celle représentant le meilleur gérant a une espérance annuelle de 2,5 %, toutes les autres sont centrées.

Le tableau précédent montre de manière frappante l'influence déterminante du nombre de gérants, la probabilité d'erreur augmentant particulièrement rapidement. Dès que l'univers contient plus de 10 fonds, de tels classements ont moins d'une chance sur deux de déterminer avec succès le meilleur d'entre eux et ce quelque soit l'horizon choisi. Ne parlons pas du cas où l'on classe cent portefeuilles sur une journée, pour lequel la probabilité de réussite est de 1,1 % ! Le palmarès présente alors toutes les caractéristiques d'un tirage aléatoire et la probabilité de réussite se rapproche rapidement de $\frac{1}{n}$. Dans un groupe de cent gérants, même un palmarès établi sur dix ans se trompera quatre fois sur cinq. Notons, d'autre part, que la probabilité d'avoir raison la plus élevée n'est que de 68 %, valeur nettement inférieure au seuil de confiance généralement utilisé en statistique (95 %).

CONCLUSION

Les résultats obtenus sont éloquents. Comment faire confiance à des palmarès qui dans la majorité des cas conduisent à des résultats inexacts ? Ainsi, dans un scénario où deux gérants sont en compétition, à un horizon d'un trimestre (période d'étude rencontrée couramment dans la presse spécialisée), un classement « à l'aveugle » couronne le mauvais gérant dans plus d'un cas sur trois. Lorsque l'on augmente le nombre de gérants, la probabilité d'avoir raison se réduit très rapidement. Avec vingt gérants par exemple, la seule rentabilité calculée sur une semaine ne désigne effectivement le meilleur gérant que dans 5,7 % des cas. À un an, on a encore plus d'une chance sur huit de se tromper.

Lorsque l'on étudie le cas de non indépendance des politiques d'allocation des différents gestionnaires, la probabilité d'erreur est modifiée, passant de 27 à 20 % à horizon d'un an pour une corrélation variant de – 0,3 à + 0,3. Cependant, les résultats ne sont pas sensiblement différents surtout si l'on considère des échéances plus courtes.

Ajoutons que la démonstration est d'autant plus frappante que nous nous sommes placés délibérément dans un cadre d'étude extrême où un seul des gérants se distingue de manière caractéristique par sa bonne gestion. Évidemment, la réalité est plus subtile et des cas aussi tranchés ne se rencontrent que rarement. Imaginons à quels résultats nous aurions abouti si les différences entre les fonds étaient moins marquées ! Nous avons également supposé toutes les trasking-errors identiques. Relâcher une telle hypothèse revient à ajouter une incertitude sur les écarts types et donc à augmenter d'autant l'aléa associé à une mesure de la rentabilité telle que le ratio de Sharpe. L'influence du hasard serait de fait encore accentué.

Ce ne sont pas tant les performances publiées dans la presse que nous remettons en cause. Les rentabilités calculées sont tout ce qu'il y a de plus exact, si les protocoles de calcul sont bien respectés, mais elles sont bien plus le fruit du hasard que le reflet d'une bonne ou d'une mauvaise gestion. Elles sont par ailleurs particulièrement sensibles à de faibles variations des inputs et n'ont notamment d'intérêt que si l'investisseur entre et sort du fonds aux dates arbitrairement choisies. Une telle information, ni fiable ni robuste, ne peut et ne doit pas être utilisée pour sélectionner des OPCVM.

BIBLIOGRAPHIE

[1] GALLAIS-HAMONNO G. (1992) *SICAV et Fonds Communs de Placement, les OPCVM en France,* Que sais-je ? n° 2654, PUF, Paris, 128 pages.

[2] MARKOWITZ H. (1952) "Portfolio Selection", *Journal of Finance,* 77-91.

[3] SHARPE W. (1985) *Investments,* Englewood Cliffs, Prentice Hall, 746 pages.

[4] SUEUR H. et WALTER A. (1993) "Mesure et attribution de performance : Analyse des SICAV actions", *Quants,* n° 10, 31 pages.

ANNEXE

1. Premier cas : deux gérants

On suppose que le vecteur (x, x') est gaussien, d'espérance $(+RT, -RT)$ et de matrice de variance-covariance :

$$\begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \sigma^2 T$$

Le réel $x' - x$ suit alors la loi normale $N(-2RT, \Sigma^2 T)$ avec $\Sigma = \sigma\sqrt{2(1-\rho)}$. La probabilité d'erreur s'écrit alors :

$$P(x' > x) = P\left(y > 2R\sqrt{T}/\Sigma\right)$$

où $y = (x' - x + 2RT)/\Sigma\sqrt{T}$ est un vecteur gaussien centré réduit. D'où :

$$P(x' > x) = N\left(-2R\sqrt{T}/\Sigma\right)$$

en notant N la fonction cumulative normale. La probabilité d'erreur s'écrit finalement :

$$P(x' > x) = N\left(-\sqrt{\frac{2}{1-\rho}}\,\frac{R}{\sigma}\sqrt{T}\right)$$

2. Deuxième cas : extension à *n* gérants

On suppose que le vecteur $x = (x_1, x_2, ..., x_n)$ suit une loi normale d'espérance $(RT, 0, 0, ..., 0)$ et de matrice de variance-covariance $\sigma^2 T I_n$, où I_n est la matrice identité d'ordre n. En notant :

$$\xi = \{x_1 > x_i\,;\ 2 \le i \le n\}$$

La probabilité que l'événement ξ se réalise s'écrit :

$$P(\xi) = \int_D \exp\left(-\frac{(x_1 - RT)^2}{2\sigma^2 T}\right) \exp\left(-\frac{1}{2\sigma^2 T}\sum_{i=2}^{n} x_i^2\right) \frac{\mathrm{d}x}{\left(\sigma\sqrt{2\pi T}\right)^n}$$

L'intégration est effectuée sur le domaine $D = \{x\,;\ x_1\ x_i,\ 2\ \le i \le n\}$. Après intégration des $n - 1$ dernières variables, on se ramène à une intégrale réelle :

$$P(\xi) = \int_{-\infty}^{+\infty} \exp\left(-\frac{(x_1 - RT)^2}{2\sigma^2 T}\right) N\left(\frac{x}{\sigma\sqrt{T}}\right)^{n-1} \frac{\mathrm{d}x}{\sigma\sqrt{2\pi T}}$$

Le calcul s'effectue en utilisant une méthode du trapèze améliorée qui donne une précision de 10^{-4}. La probabilité d'erreur est alors le complémentaire à l'unité de $P(\xi)$.