

Implémentation d'un papier de recherche

False Discoveries in Mutual Fund Performance: Measuring Luck in Estimated Alphas
BARRAS, SCAILLET & WERMERS

BHAUGEERUTTY, KIRCH, MORIN

Université Paris Dauphine - PSL - Outils pour la Gestion Quantitative

Année académique 2023-2024

Table des matières

- 1 Introduction
- 2 Méthodologie
 - Performance des fonds
 - False Discovery Rate
- 3 Présentation des données
- 4 Résultats
 - Calcul de la performance des fonds
- 5 Critiques
 - Biais de l'estimateur FDR
 - Nuances
- 6 Extension - données de 1980 à 2023
 - Résultats
- 7 Conclusion

Introduction

Revue de littérature

Barras et al., 2010, *Journal of Finance* est un article novateur pour son utilisation du *False Discovery Rate* (FDR) en finance.

Il se situe dans une littérature très fournie sur l'évaluation de la performance des Fonds Commun de Placement (FCP) :

- Jensen, 1968, régression OLS des rendements sur un ensemble de facteurs de tarification pour dégager la performance excédentaire du fonds
- Blake et al., 2014 soulignent l'importance de prendre en compte l'information inter-fonds (Dépendance Croisée Sectionnelle (CSD)) : observation des données comme un panel
- Critique principale faite de l'article : le biais induit par le FDR, qui classe de manière incorrecte les FCP avec un réel α positif/négatif comme des FCP affichant un α nul (Andrikogiannopoulou and Papakonstantinou, 2019)
- Nouvelles recherches prennent en compte le biais, procédures avec plus de puissance statistique (Ferson and Chen, 2018, Harvey and Liu, 2020)

Méthodologie

Modèle à 4 facteurs

$$r_{i,t} = \alpha_i + b_i * r_{m,t} + s_i * r_{smb,t} + h_i * r_{hml,t} + m_i * r_{mom,t} + \epsilon_{i,t} \quad (1)$$

avec :

- r_i^t : le rendement du fond i à la période t ,
- r_x^t avec $x \in \{m, smb, hml, mom\}$: nos quatre variables exogènes
- α_i : la constante du modèle, la performance du fond i ,
- b_i, s_i, h_i, m_i : les autres coefficients du modèle,
- $\epsilon_{i,t}$: le résidu de la régression.

False Discovery Rate

$$\text{Overall FDR} = \frac{\pi_0 * \gamma}{\sum_{n=1}^N I_{(p\text{-val}_n < \gamma)} / N} \quad (2)$$

avec :

- π_0 : la proportion de fond zéro-alpha,
- γ : le seuil de significativité,
- $\sum_{n=1}^N I_{(p\text{-val}_n < \gamma)} / N$: la proportion de fonds tels que $p\text{-value} < \gamma$,

$$\text{FDR}^- = \frac{\pi_0 * \gamma / 2}{\sum_{n=1}^{N^-} I_{(p\text{-val}_n < \gamma \& \alpha_n < 0)} / N^-} \quad (3)$$

$$\text{FDR}^+ = \frac{\pi_0 * \gamma / 2}{\sum_{n=1}^{N^+} I_{(p\text{-val}_n < \gamma \& \alpha_n > 0)} / N^+} \quad (4)$$

avec :

- $\frac{\sum_{n=1}^{N^-} I_{(p\text{-val}_n < \gamma \& \alpha_n < 0)}}{N^-}$: la proportion de fonds tels que $(p\text{-val} < \gamma)$ et $(\alpha < 0)$
- $\frac{\sum_{n=1}^{N^+} I_{(p\text{-val}_n < \gamma \& \alpha_n > 0)}}{N^+}$: la proportion de fonds tels que $(p\text{-val} < \gamma)$ et $(\alpha > 0)$

Proportions

Proportion de fonds chanceux ou malchanceux (fonds à alpha-zero tout en ayant une p-value significative) :

$$\mathbf{E}(\mathbf{F}_{\gamma}) = \pi_0 \cdot \gamma/2 \quad (5)$$

$$\mathbf{E}(\mathbf{F}_{\gamma}^{+}) = \pi_A^{+} - E(F_{\gamma}) \quad (6)$$

$$\mathbf{E}(\mathbf{F}_{\gamma}^{-}) = \pi_A^{-} - E(F_{\gamma}) \quad (7)$$

Proportion de fonds performants & sous performants :

$$\mathbf{E}(\mathbf{T}_{\gamma}^{+}) = E(S_{\gamma}^{+}) - E(F_{\gamma}^{+}) \quad (8)$$

$$\mathbf{E}(\mathbf{T}_{\gamma}^{-}) = E(S_{\gamma}^{-}) - E(F_{\gamma}^{-}) \quad (9)$$

avec :

- F : fonds chanceux et malchanceux
- S^{+} (resp. S^{-}) : taux de fausses découvertes, ou FDR^{+} (resp. FDR^{-})
- T^{+} (resp. T^{-}) : fonds performants (resp. sous-performants)

Présentation des données

Jeux de données

- Données trimestrielles allant de 1980 à 2006.
- Variables explicatives :
 - ▶ Rendement excédentaire du portefeuille de marché NYSE/AMEX/NASDAQ contre le taux sans risque
 - ▶ Rendements des portefeuilles zéro-investissement des facteurs *Size*, *Book-to-Market* et *Momentum*
- Fonds Communs de Placement : fonds américains avec au minimum 20 observations trimestrielles.

Résultats

Calcul de la performance des fonds

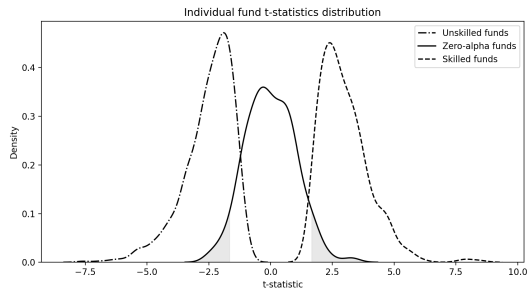


Figure: Distribution des t-statistiques des fonds individuels

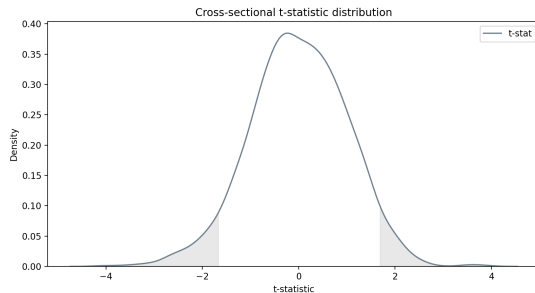


Figure: Distribution des t-statistiques en coupe transversale

Performance des fonds à long terme

	Zero alpha (π_0)				Unskilled (π_A^-)		Skilled (π_A^+)		
Proportion	18.1				52.8		29.1		
Number	271				791		436		

	Left Tail				Right Tail				
Signif. Level (γ)	0.05	0.10	0.15	0.20	0.20	0.15	0.10	0.05	Signif. Level (γ)
Signif. S_V^- (%)	7.5	8.3	9.6	10.2	21.3	21.7	20.5	28	Signif. S_V^+ (%)
Unlucky F_V^- (%)	3.8	6.6	7.4	8.9	0	0	0	0	Lucky F_V^+ (%)
Unskilled T_V^- (%)	3.7	1.7	2.2	1.3	21.3	21.7	20.5	28.0	Skilled T_V^+ (%)

Figure: Résultats obtenus sur la période 1980 - 2006

Critiques

Biais de l'estimateur FDR

Une des critiques principales : biais induit par l'estimateur FDR.

Par exemple : p-value élevée pour un fonds non-zéro alpha (vrai alpha faible, d'un bruit d'estimation élevé, ou les deux). Dans ce cas, $\hat{\pi}_0(\lambda) > \pi_0(\lambda)$ et inversement $\hat{\pi}_A(\lambda) < \pi_A(\lambda)$

Méthode pour estimer le biais :

$$\delta(\lambda) = \frac{F_{nc}(I(\lambda); \sigma_e, T)}{1 - \lambda}$$

avec :

- λ le seuil utilisé pour calculer $\hat{\pi}_0$
- T le nombre d'observations
- σ_e la volatilité résiduelle

Résultats : sur la période de 1980-2006, **biais = 69 %**. \Leftrightarrow probabilité de 69 % de mal classifier un fonds alpha négatif ou positif en fonds zéro alpha. Biais similaire sur le full-sample (1980 - 2023)

Contributions et nuances

- FDR facilement implémentable, sans simulation ni puissance de calcul trop importante. Méthodologie simple nécessitant uniquement la p-value et seuil γ
- Mais... résultats pas forcément transférables à une autre base de données (CSRP) qui peut poser une interrogation sur la généralisation de cette étude.
- Néanmoins, résultats sur le marché australien (Kim et al., 2014) et sur le marché du Royaume-Uni (Bredin et al., 2014) avec le FDR sont prometteurs

Extension - données de 1980 à 2023

Performance de Long Terme

	Zero alpha (π_0)				Unskilled (π_A^-)				Skilled (π_A^+)			
Proportion	29.7				18.9				51.4			
Number	2590				1652				4484			

	Left Tail				Right Tail				
Signif. Level (γ)	0.05	0.10	0.15	0.20	0.20	0.15	0.10	0.05	Signif. Level (γ)
Signif. S_Y^- (%)	11.1	12.5	13.9	13.8	12.1	12.7	13.5	13.5	Signif. S_Y^+ (%)
Unlucky F_Y^- (%)	0	0	0	0	7.4	4.9	2.7	1.3	Lucky F_Y^+ (%)
Unskilled T_Y^- (%)	11.1	12.5	13.9	13.9	4.7	7.8	10.8	12.1	Skilled T_Y^+ (%)

Figure: Résultats obtenus sur la période 1980 - 2023

	Zero alpha (π_0)				Unskilled (π_A^-)				Skilled (π_A^+)			
Proportion	18.1				52.8				29.1			
Number	271				791				436			

	Left Tail				Right Tail				
Signif. Level (γ)	0.05	0.10	0.15	0.20	0.20	0.15	0.10	0.05	Signif. Level (γ)
Signif. S_Y^- (%)	7.5	8.3	9.6	10.2	21.3	21.7	20.5	28	Signif. S_Y^+ (%)
Unlucky F_Y^- (%)	3.8	6.6	7.4	8.9	0	0	0	0	Lucky F_Y^+ (%)
Unskilled T_Y^- (%)	3.7	1.7	2.2	1.3	21.3	21.7	20.5	28.0	Skilled T_Y^+ (%)

Figure: Rappel : résultats obtenus sur la période 1980 - 2006

Conclusion

Conclusion

- La majorité des gérants de FCP n'ont pas les capacités de sélection pour obtenir une performance positive.
- Réflexion quant à la pertinence de la gestion active.
- Ouverture :
 - ▶ Méthode simple permettant une généralisation à d'autres zones géographiques et périodes temporelles.
 - ▶ Par exemple : Royaume-Uni (Bredin, 2014) et Australie (Kim, 2014).

Merci pour votre écoute !

A vos questions !