Projet Microstructure et mécanismes des marchés financiers

Yassine ARAHOU, Lionel FELIHO, Anass JMARI, Ilyass KETTIOUI

Professeur Enseignant : Sonia JIMENEZ GARCES

Mars 2019

Contents

1 Description du Projet

1.1 Asymetrie Informationnelle

Dans le cours de **'Microstructure et mécanisme des marchés financiers'**, nous avons étudié certains modèles théoriques et empiriques qui aspirent à modéliser les différentes structures de marché responsables de la formation des prix des titres et l'estimation des coûts de transaction. L'analyse approfondie de ces coûts exigent une réflexion poussée sur la composante de sélection adverse et, par extension, l'***asymetrie informationnelle*** des marchés.

L'asymétrie d'information décrit une situation dans laquelle tous les participants à un marché ne disposent pas de la même information. C'est une imperfection du marché qui peut aboutir à une sélection adverse ou à un aléa moral.

On comprend alors que l'importance de la quantification de cette asymetrie informationnelle. Plusieurs mesures ambitionnent d'estimer le degré d'asymetrie informationnelle auquel est soumis un titre.

1.2 Mesures issues de la Microstructure

On peut citer quelques mesures d'asymetrie informationnelle issues de la microstructure :

- Le bid-ask spread est une mesure de l'asymetrie d'information, mais une mesure biaisée, car elle tient compte d'autres facteurs qui n'ont rien à voir avec l'asymetrie informationnelle.
- Mesure "macro" des entreprises : le Market-to-Book

  $$\text{Market-to-Book} = \frac{\text{valeur du marché (valeur des flux futurs actualisées au bon taux)}}{\text{valeur comptable (réalisations passées)}}$$

  Plus une entreprise a des opportunités de croissance fortes, plus son ratio Market-to-Book est fort.

Cependant, ces mesures présentent deux inconvénients majeurs :

1. Faible fréquence de disponibilité de ces données comptables, qui entraine une mauvaise précision de calcul.
2. Sensibilité aux normes comptables, ces mesures peuvent être altérées si les données comptables sont modifiées.

1.3 Objectif du projet

Un des buts de ce projet est de fournir un conseil aux practiciens : "Vous n'avez pas le temps de vous engagez dans tous les types de marchés, il faut se focaliser sur un marché dans lequel vous pourrez bénéficier des mesures d'asymétrie d'information."

Ainsi, il faut leur proposer une mesure d'asymetrie d'information efficace pour orienter leurs activités sur le marché.

Dans ce contexte, il nous est proposé d'étudier une mesure alternative du degré d'asymétrie informationnelle d'un titre : **AIM (Asymetric Information Measure)** qui est fondée sur les modèles d'équilibre à anticipations rationnelles. C'est une mesure qui a été validée et publiée dans *Journal of financial economics.*

Cette mesure est obtenue par projection de la rentabilité d'un titre sur son prix ainsi que sur le prix de portefeuilles informationnellement pertinents pour cet actif.

$$r_{i,t} = \alpha_i + \beta_i^i P_i^{t-1} + \sum_{j \in J, j \neq i} \beta_i^j P_{j,t-1} + \epsilon_{i,t} \quad \Rightarrow \quad \text{AIM} = R^2$$

2 Étude théorique

Le but de cette première partie est de montrer à l'aide du modèle de Grossman et Stiglitz (dans un cadre multi-actif) que : $cov(\tilde{P}_1 - \tilde{P}_0; \tilde{P}_0)$, qui est une approximation de $AIM = R^2$, est négative et décroit avec le degré d'asymétrie informationnelle pour conclure avec un raisonnement théorique que AIM est une bonne mesure d'asymétrie informationnelle.

2.1 Les hypothèses

Il s'agit des hypothèses du modèle de Grossman et Stiglitz, A savoir:

- Modèle à deux dates 0 et 1.
- N actifs risqués et un actif sans risque.
- λ reprèsente la part des investisseurs informés et $(1 - \lambda)$ représente la part des investisseurs non informés.
- L'information dont dispose les informés pour chaque actif est représentée par un vecteur de taille N de loi normale $\tilde{\theta} \sim \mathcal{N}_N(\mu_\theta, \sigma_\theta)$.
- Le prix $\tilde{P}_1$ à la date 1 s'écrit sous la forme $\tilde{P}_1 = \tilde{\theta} + \tilde{\epsilon}$ avec $\tilde{\epsilon}$ un bruit qui suit la loi normale $\mathcal{N}_N(0, \sigma_\epsilon)$.
- La demande en actif risqués $\tilde{z}$ suit la loi normale $\mathcal{N}_N(\mu_z, \sigma_z)$.
- Les trois vecteurs aléatoires $\tilde{\theta}$, $\tilde{\epsilon}$ et $\tilde{z}$ sont indépendants.
- La fonction d'utilité des investisseurs est supposée exponentielle négative.

2.2 Analyse

On note M_i la somme investi par l'investisseur i et X_i le vecteur des proportions d'actif de l'investisseur i, ainsi les agents ont une dotation initiale :
À la date 0 :

$$\tilde{w}_{0i} = M_i + X_i \tilde{P}_0$$

Et à la date 1 :

$$\tilde{w}_{1i} = M_i R + X_i \tilde{P}_1 = \tilde{w}_{0i} R + X_i(\tilde{P}_1 - \tilde{P}_0)$$

Chaque investisseur cherche à maximiser en 0 l'espérance de l'utilité de son portefeuille en 1 sachant leur information, c'est à dire maximiser :

$$E[U(\tilde{w}_{1i})/I_i]$$

Ceci est équivalent à maximiser $U(E[\tilde{w}_{1i}/I_i]) - aVar[\tilde{w}_{1i}/I_i]$ avec a l'aversion au risque des investisseurs.

Ce dernier terme peut s'écrire sous la forme :

$$\tilde{w}_{0i}R + X_iE[\tilde{P}_1 - \tilde{P}_0/I_i] - \frac{a}{2}X_iVar[\tilde{P}_1 - R\tilde{P}_0/I_i]X_i^t$$

Maximiser ce terme revient à dire que le gradient par rapport à X_i est nul, on obtient alors :

$$X_i = \frac{1}{a}Var[\tilde{P}_1 - R\tilde{P}_0/I_i]^{-1}E[\tilde{P}_1 - \tilde{P}_0/I_i]$$

On distingue deux catégories d'investisseurs :

2.2.1 Les non-informés $I_i = \tilde{P}_0$

Pour cette catégorie l'information se limite à $\tilde{P}_0$, on a donc :

- $Var[\tilde{P}_1 - R\tilde{P}_0/I_i] = Var[\tilde{P}_1/\tilde{P}_0] = Var[\tilde{\theta}/\tilde{P}_0] + Var[\tilde{\epsilon}] = N + E$
  Où on pose $N = Var[\tilde{\theta}/\tilde{P}_0]$.
- $E[\tilde{P}_1/I_i] = E[\tilde{\theta}/\tilde{P}_0]$.

On conclut:

$$\boxed{X_{NI} = \frac{1}{a}(N + E)^{-1}(E[\tilde{\theta}/\tilde{P}_0] - R\tilde{P}_0)}$$

2.2.2 Les informés $I_i = (\tilde{P}_0, \tilde{\theta})$

Pour cette catégorie l'information est représentée par l'ensemble $(\tilde{P}_0, \tilde{\theta})$, dans ce cas :

- $Var[\tilde{P}_1 - R\tilde{P}_0/I_i] = Var[\tilde{P}_1/(\tilde{P}_0, \tilde{\theta})] = Var[\tilde{\theta}/(\tilde{P}_0, \tilde{\theta})] + Var[\tilde{\epsilon}/(\tilde{P}_0, \tilde{\theta})] = Var[\tilde{\epsilon}] = E$.
- $E[\tilde{P}_1/I_i] = E[\tilde{\theta}/\tilde{P}_0, \tilde{\theta}] + E[\tilde{\epsilon}/\tilde{P}_0, \tilde{\theta}] = \tilde{\theta}$.

On conclut:

$$\boxed{X_I = \frac{1}{a}E^{-1}(\tilde{\theta} - R\tilde{P}_0)}$$

2.3 Notations et Relations utiles

2.3.1 Notations

On fixe les notations suivantes dont quelques unes sont déjà rencontrées avant :

- $E = Var(\tilde{\epsilon})$
- $T = Var(\tilde{\theta})$
- $N = Var(\tilde{\theta}/\tilde{P}_0)$
- $Z = Var(\tilde{z})$
- $V_m = [\lambda E^{-1} + (1-\lambda)(N+E)^{-1}]^{-1}$
- $w(\tilde{\theta}, \tilde{z}) = \tilde{\theta} - \frac{a}{\lambda} E\tilde{z}$

On veut montrer que le prix à l'instant t = 0 peut s'écrire sous la forme $\tilde{P}_0 = A_0 + A_1 \tilde{w}(\tilde{\theta}, \tilde{z})$ et donc de determiner A_0 et A_1.

2.3.2 Relations utiles

$$\boxed{Var(\tilde{w}) = T + \frac{a^2}{\lambda^2} EZE} \tag{1}$$

Preuve :

$$\begin{aligned}
Var(\tilde{w}) &= Var(\tilde{\theta} - \frac{a}{\lambda} E\tilde{z}) \\
&= Var(\tilde{\theta}) + Var(-\frac{a}{\lambda} E\tilde{z}) \\
&= T + \frac{a^2}{\lambda^2} Var(E\tilde{z}) \\
&= T + \frac{a^2}{\lambda^2} EVar(\tilde{z})E
\end{aligned}$$

$$\boxed{N = T - T(Var^{-1}(\tilde{w}))T} \tag{2}$$

Preuve :

$$\begin{aligned}
N &= Var(\tilde{\theta}/\tilde{P}_0) \\
&= Var(\tilde{\theta}/\tilde{w}(\tilde{\theta}, \tilde{z})) \\
&= Var(\tilde{\theta}) - cov(\tilde{\theta}, \tilde{w}) Var(\tilde{w})^{-1} cov(\tilde{w}, \tilde{\theta}) \\
&= T - T(Var^{-1}(\tilde{w}))T
\end{aligned}$$

En effet :

$$\begin{aligned}
cov(\tilde{\theta}, \tilde{w}) &= cov(\tilde{\theta}, \tilde{\theta} - \frac{a}{\lambda} E\tilde{z}) \\
&= Var(\tilde{\theta}) - \frac{a}{\lambda} cov(\tilde{\theta}, E\tilde{z}) \\
&= Var(\tilde{\theta}) \qquad (independance)
\end{aligned}$$

De cette relation on déduit:

$$\boxed{T^{-1}N = I_n - (Var^{-1}(\tilde{w}))T \iff NT^{-1} = I_n - T(Var^{-1}(\tilde{w}))} \tag{3}$$

Et aussi :

$$\boxed{E(\tilde{\theta}/\tilde{P}_0) = E(\tilde{\theta}) - (I_n - NT^{-1})(\tilde{w} - E[\tilde{w}])} \tag{4}$$

Preuve :

$$\begin{aligned}
E(\tilde{\theta}/\tilde{P}_0) &= E(\tilde{\theta}/\tilde{w}(\tilde{\theta}, \tilde{z})) \\
&= E[\tilde{\theta}] + cov(\tilde{\theta}, \tilde{w})Var[\tilde{w}]^{-1}(\tilde{w} - E[\tilde{w}]) \\
&= E(\tilde{\theta}) - TVar(\tilde{w})^{-1}(\tilde{w} - E[\tilde{w}]) \\
&= E(\tilde{\theta}) - (I_n - NT^{-1})(\tilde{w} - E[\tilde{w}])
\end{aligned}$$

On a aussi cette relation importante :

$$\boxed{V_m = [\lambda E^{-1} + (1-\lambda)(N+E)^{-1}]^{-1} = (E+N)(I_n + \lambda E^{-1}N)^{-1}} \tag{5}$$

Preuve :

$$\begin{aligned}
(N+E)(N+E)^{-1} &= I_n \\
N(N+E)^{-1} &= I_n - E(N+E)^{-1} \\
(N+E)^{-1} + \lambda E^{-1}N(N+E)^{-1} &= \lambda E^{-1} + (1-\lambda)(N+E)^{-1} \\
(E+N)(I_n + \lambda E^{-1}N)^{-1} &= [\lambda E^{-1} + (1-\lambda)(N+E)^{-1}]^{-1}
\end{aligned}$$

$$\boxed{(1-\lambda)V_m(N+E)^{-1}E = E - \lambda V_m} \tag{6}$$

Preuve :

$$\begin{aligned}
V_m^{-1} &= \lambda E^{-1} + (1-\lambda)(N+E)^{-1} \\
I_n - \lambda V_m E^{-1} &= (1-\lambda)V_m(N+E)^{-1} \\
E - \lambda V_m &= (1-\lambda)V_m(N+E)^{-1}E
\end{aligned}$$

$$\boxed{(1-\lambda)V_m(N+E)^{-1}N = V_m - E} \tag{7}$$

Preuve :

$$\begin{aligned}
(N+E)^{-1}N &= I_n - (N+E)^{-1}E \\
(1-\lambda)V_m(N+E)^{-1}N &= (1-\lambda)V_m - (1-\lambda)V_m(N+E)^{-1}E \\
&= (1-\lambda)V_m - (E - \lambda V_m) \\
&= V_m - E
\end{aligned}$$

À partir des deux relations précedentes on en déduit

$$\boxed{(V_m - E)N^{-1}E = E - \lambda V_m} \tag{8}$$

2.4 Expression du prix en t=0

Le but de cette partie est d'écrire $\tilde{P}_0$ sous la forme :

$$\tilde{P}_0 = A_0 + A_1(\tilde{\theta} - \tfrac{a}{\lambda}E\tilde{z}) = A_0 + A_1\tilde{w}(\tilde{\theta}, \tilde{z})$$

L'équilibre s'écrit :

$$\lambda X_I + (1-\lambda)X_{NI} = \tilde{z}$$

En developpant avec les expressions de X_I et X_{NI}, et en utilisant la relation 4 on a :

$$\begin{aligned}
0 &= \lambda X_I + (1-\lambda)X_{NI} - \tilde{z} \\
&= \frac{\lambda}{a}E^{-1}(\tilde{\theta} - R\tilde{P}_0) + \frac{1-\lambda}{a}(N+E)^{-1}(E[\tilde{\theta}/\tilde{P}_0] - R\tilde{P}_0) - \tilde{z} \\
&= \frac{\lambda}{a}E^{-1}(\tilde{\theta} - R(A_0 + A_1\tilde{w})) + \frac{1-\lambda}{a}(N+E)^{-1}(E(\tilde{\theta}) - TVar(\tilde{w})^{-1}(\tilde{w} - E[\tilde{w}]) \\
&\quad - R(A_0 + A_1\tilde{w})) - \tilde{z}
\end{aligned}$$

En Remplaçant $\tilde{w}$ par son expression $\tilde{\theta} - \frac{a}{\lambda}E\tilde{z}$ et en factorisant selon les coefficients de $\tilde{\theta}$ et $\tilde{z}$ on obtient :

$$\begin{aligned}
0 = \tilde{\theta}[&\frac{\lambda}{a}(E^{-1} - RE^{-1}A_1) + \frac{(1-\lambda)}{a}(N+E)^{-1}(TVar(\tilde{w})^{-1} - RA_1)] \\
&+ \tilde{z}[RE^{-1}A_1E + \frac{(1-\lambda)}{\lambda}(N+E)^{-1}(RA_1E - TVar(\tilde{w})^{-1}E) - I_n] \\
&\quad + [-\frac{\lambda}{a}RE^{-1}A_0 + \frac{(1-\lambda)}{a}(N+E)^{-1}(E[\tilde{\theta}] - TVar(\tilde{w})^{-1}E[\tilde{w}] - RA_0)]
\end{aligned}$$

Par indépendance des variables aléatoires on obtient le système suivant :

$$\begin{cases}
0 &= \lambda(E^{-1} - RE^{-1}A_1) + (1-\lambda)(N+E)^{-1}(TVar(\tilde{w})^{-1} - RA_1) \\
0 &= \lambda RE^{-1}A_0 + (1-\lambda)(N+E)^{-1}(E[\tilde{\theta}] - TVar(\tilde{w})^{-1}E[\tilde{w}] - RA_0)
\end{cases}$$

2.4.1 Calcul de A_1

À partir de la première équation du système on a:

$$(\lambda E^{-1} + (1-\lambda)(N+E)^{-1})RA_1 = \lambda E^{-1} + (1-\lambda)(N+E)^{-1}TVar(\tilde{w})^{-1}$$

d'où

$$\begin{aligned}
A_1 &= \frac{1}{R}V_m(\lambda E^{-1} + (1-\lambda)(N+E)^{-1}TVar(\tilde{w})^{-1}) \\
&= \frac{1}{R}V_m(\lambda E^{-1} + (1-\lambda)(N+E)^{-1} - (1-\lambda)(N+E)^{-1}NT^{-1}) && \textit{(D'après 3)} \\
&= \frac{1}{R}(I_n - (1-\lambda)V_m(N+E)^{-1}NT^{-1}) && \textit{(Défintion de } V_m\textit{)} \\
&= \frac{1}{R}(T - (1-\lambda)V_m(N+E)^{-1}N)T^{-1} \\
&\boxed{= \frac{1}{R}(T + E - V_m)T^{-1}} && \textit{(D'après 7)}
\end{aligned}$$

2.4.2 Calcul de A_0

À partir de la deuxième équation du système on a:

$$(\lambda E^{-1} + (1-\lambda)(N+E)^{-1})RA_0 = (1-\lambda)(N+E)^{-1}(E[\tilde{\theta}] - TVar(\tilde{w})^{-1}E[\tilde{w}])$$

d'où :

$$\begin{aligned}
A_0 &= \frac{(1-\lambda)}{R}V_m(N+E)^{-1}(E[\tilde{\theta}](I_n - TVar(\tilde{w})^{-1}) + \frac{a}{\lambda}TVar(\tilde{w})^{-1}EE[\tilde{z}]) \\
&= \frac{1}{R}(V_m - E)N^{-1}(E[\tilde{\theta}]NT^{-1} + \frac{a}{\lambda}(I_n - NT^{-1})EE[\tilde{z}]) && \textit{(D'après 3 et 7)} \\
&= \frac{1}{R}((V_m - E)T^{-1}E[\tilde{\theta}] + \frac{a}{\lambda}((V_m - E)N^{-1} - (V_m - E)T^{-1})EE[\tilde{z}])
\end{aligned}$$

Or d'après l'expression de A_1 on :

$$(V_m - E)T^{-1} = I_n - RA_1$$

d'où :

$$\begin{aligned}
A_0 &= \frac{1}{R}((I_n - RA_1)E[\tilde{\theta}] + \frac{a}{\lambda}((V_m - E)N^{-1}E - (E - RA_1E))E[\tilde{z}]) \\
&= \frac{1}{R}([I_n - RA_1]E[\tilde{\theta}] + [\frac{Ra}{\lambda}A_1E + \frac{a}{\lambda}(V_m - E)N^{-1}E - E)]E[\tilde{z}]) \\
&\boxed{= \frac{1}{R}((I_n - RA_1)E(\tilde{\theta}) + (\frac{Ra}{\lambda}A_1E - aV_m)E(\tilde{z}))} && \textit{(D'après 8)}
\end{aligned}$$

2.5 Calcul de $cov(\tilde{P}_1 - \tilde{P}_0, \tilde{P}_0)$

2.5.1 Expression de la covariance

On commence par calculer la $Var(\tilde{P}_0)$:

$$\begin{aligned}
Var(\tilde{P}_0) &= Var(A_0 + A_1(\tilde{\theta} - \frac{a}{\lambda}E\tilde{z})) \\
&= A_1 Var(\tilde{\theta} - \frac{a}{\lambda}E\tilde{z})A_1^t \\
&= A_1 Var(\tilde{\theta})A_1^t + \frac{a^2}{\lambda^2}A_1 Var(E\tilde{z})A_1^t \\
&= A_1 T A_1^t + \frac{a^2}{\lambda^2}A_1 E\tilde{z}E^t A_1^t
\end{aligned}$$

D'autre part :

$$\begin{aligned}
cov(\tilde{P}_1, \tilde{P}_0) &= cov(\tilde{\theta} + \tilde{\epsilon}, A_0 + A_1(\tilde{\theta} - \frac{a}{\lambda}E\tilde{z})) \\
&= cov(\tilde{\theta} + \tilde{\epsilon}, \tilde{\theta} - \frac{a}{\lambda}E\tilde{z})A_1^t \\
&= (var(\tilde{\theta}) + cov(\tilde{\epsilon}, \tilde{\theta} - \frac{a}{\lambda}E\tilde{z}))A_1^t \\
&= var(\tilde{\theta})A_1^t = TA_1^t && \textit{(Par indépendances des variables aléatoires)}
\end{aligned}$$

On déduit que :

$$\begin{aligned}
cov(\tilde{P}_1 - \tilde{P}_0, \tilde{P}_0) &= cov(\tilde{P}_1, \tilde{P}_0) - Var(\tilde{P}_0) \\
&\boxed{= TA_1^t - A_1TA_1^t - \frac{a^2}{\lambda^2}(A_1E)Z(A_1E)^t} \\
&\boxed{= (I_n - A_1)TA_1^t - \frac{a^2}{\lambda^2}(A_1E)Z(A_1E)^t}
\end{aligned}$$

2.5.2 La covariance en fonction d'une fonction f

Montrons que :

$$cov(\tilde{P}_1 - \tilde{P}_0, \tilde{P}_0) = -\lambda f(N) - \lambda^2 f(N)(T-N)^{-1}f(N)$$

Avec :

$$f(N) = (N^{-1}EN^{-1} + \lambda N^{-1})^{-1} + (N^{-1} + \lambda E^{-1})^{-1}$$

On a :

$$\begin{aligned}
TA_1^t &= \frac{1}{R}T[(T + E - V_m)T^{-1}]^t \\
&= \frac{1}{R}T(T^{-1})^t(T + E - V_m)^t && \textit{(tous les termes sont symétriques (variances), et} \\
&= \frac{1}{R}(T + E - V_m) && \textit{combinaison et inverse de symétriques est symétrique)}
\end{aligned}$$

On prendra dorénavant $R = 1$. Par conséquence $TA_1^t = T + E - V_m$.

D'autre part, d'après la relation (1) on a :

$$\frac{a^2}{\lambda^2} EZE = Var(\tilde{w}) - T$$

Ainsi :

$$\begin{aligned}
\frac{a^2}{\lambda^2}(A_1E)\tilde{z}(A_1E)^t &= A_1(Var(\tilde{w}) - T)A_1^t \\
&= (T + E - V_m)T^{-1}Var(\tilde{w})T^{-1}(T + E - V_m) - A_1TA_1^t \\
&= TA_1^t(T - N)^{-1}TA_1^t - A_1TA_1^t && \textit{(D'après 2)}
\end{aligned}$$

On conclut que :

$$\boxed{cov(\tilde{P}_1 - \tilde{P}_0, \tilde{P}_0) = TA_1^t - TA_1^t(T - N)^{-1}TA_1^t} \qquad (*)$$

On a :

$$\begin{aligned}
V_mE^{-1}N &= (E + N)(I_n + \lambda E^{-1}N)^{-1}E^{-1}N && \textit{(D'après 5)} \\
&= E(I_n + \lambda E^{-1}N)^{-1}E^{-1}N + N(I_n + \lambda E^{-1}N)^{-1}E^{-1}N \\
&= (N^{-1} + \lambda E^{-1})^{-1} + (N^{-1}EN^{-1} + \lambda N^{-1})^{-1} \\
&= f(N)
\end{aligned}$$

Or,

$$\begin{aligned}
E - V_m &= (\lambda V_m - E)E^{-1}N && \textit{(D'après 8)} \\
&= \lambda V_mE^{-1}N - N \\
&= \lambda f(N) - N
\end{aligned}$$

Ainsi :

$$\boxed{TA_1^t = T + E - V_m = T - N + \lambda f(N)}$$

En injectant ce résultat dans (*) on trouve :

$$\begin{aligned}
cov(\tilde{P}_1 - \tilde{P}_0, \tilde{P}_0) &= T - N + \lambda f(N) - (T - N + \lambda f(N))(T - N)^{-1}(T - N + \lambda f(N)) \\
&= -\lambda f(N)(T - N)^{-1}(T - N + \lambda f(N)) \\
&= -\lambda f(N) - \lambda^2 f(N)(T - N)^{-1}f(N) && \text{CQFD}
\end{aligned}$$

2.5.3 Caractère négatif et Monotonie de la covariance

2.5.3.1 Définie négative

La matrice $f(N)$ est définie positive car elle est construite par des opérations qui conservent le caractère définie positive sur des matrices définies positives (la variance).
La matrice $(T-N)^{-1}$ est régulière alors $f(N)(T-N)^{-1}f(N)$ est définie positive.
Par conséquence la matrice $cov(\tilde{P}_1 - \tilde{P}_0, \tilde{P}_0) = -\lambda f(N) - \lambda^2 f(N)(T-N)^{-1}f(N)$ est définie négative.

2.5.3.2 Monotonie

soit N_1 et N_2 deux matrices définies positives tel que $N_1 > N_2$

$$\begin{aligned}
N_1 > N_2 &\Rightarrow N_1^{-1} < N_2^{-1} \\
&\Rightarrow \begin{cases} N_1^{-1}EN_1^{-1} < N_2^{-1}EN_2^{-1} \\ \lambda N_1^{-1} < \lambda N_2^{-1} \\ N_1^{-1} + \lambda E^{-1} < N_2^{-1} + \lambda E^{-1} \end{cases} \\
&\Rightarrow \begin{cases} (N_1^{-1}EN_1^{-1} + \lambda N_1^{-1})^{-1} > (N_2^{-1}EN_2^{-1} + \lambda N_2^{-1})^{-1} \\ (N_1^{-1} + \lambda E^{-1})^{-1} > (N_2^{-1} + \lambda E^{-1})^{-1} \end{cases} \\
&\Rightarrow f(N_1) > f(N_2) \\
&\Rightarrow |cov_1(\tilde{P}_1 - \tilde{P}_0, \tilde{P}_0)| > |cov_2(\tilde{P}_1 - \tilde{P}_0, \tilde{P}_0)|
\end{aligned}$$

2.6 Conclusion de l'étude théorique

$cov(\tilde{P}_1 - \tilde{P}_0, \tilde{P}_0)$ est une matrice définie négative et elle augmente en valeur absolue avec le degré d'asymétrie informationnelle. Ce que justifie donc que, théoriquement, AIM est une bonne mesure d'asymétrie informationnelle.

Reste à nous assurer empiriquement de nos résultats théoriques !

Étude Empirique AIM comme mesure d'IA

Yassine ARAHOU, Lionel FELIHO, Anass JMARI, Ilyass KETTIOUI

Mars 2019

Contents

Étude empirique

Pour valider l'usage de la mesure AIM pour quantifier l'asymetrie informationnelle, nous suivons la méthodologie proposée par Easley et al. (2002) dans l'article de validation de la mesure PIN comme mesure d'asymetrie informationnelle.

On essayera de démontrer que AIM est positivement et significativement corrélée aux rentabilités des titres pour justifier qu'il s'agit bien d'une bonne mesure d'asymetrie informationnelle.

Chargement des données du marché

Statistiques analysées

Dans notre étude nous allons étudier plusieurs paramètres en fonction de la rentabilité des titres moins le taux sans risque. Voici une brève description des paramètres :

beta

Le bêta, ou coefficient bêta, d'un titre financier est un coefficient de volatilité ou de sensibilité qui indique la relation existant entre les fluctuations de la valeur du titre et les fluctuations du marché.

Il s'obtient en régressant la rentabilité de ce titre sur la rentabilité de l'ensemble du marché. Il s'agit de modéliser la corrélation entre un titre et l'indice auquel il appartient.

Concrètement, le marché a un bêta de 1. Une action ayant un bêta supérieur à celui du marché, par exemple 1.5, revient à dire qu'elle réplique les variations du marché mais les amplifie; on dit qu'elle est plus sensible. Inversement, une action ayant un bêta de 0.5 répliquera son marché de référence mais aura tendance à amortir les variations de ce dernier.

AIM

AIM est la mesure alternative du degré d'asymétrie informationnelle que nous cherchons à étudier par le biais de ce projet. AIM représente l'information asymétrique, elle se produit lorsqu'une partie à une transaction économique possède une connaissance matérielle plus grande que l'autre partie. Cela se produit normalement lorsque le vendeur d'un bien ou d'un service a une meilleure connaissance que l'acheteur, bien que l'inverse soit possible. Presque toutes les transactions économiques impliquent des asymétries d'information.

Les asymétries d'information tendent à être plus marquées dans les domaines où l'information est complexe, difficile à obtenir, ou les deux. Par exemple, il est relativement difficile d'obtenir des informations exclusives lorsqu'on échange des cartes de baseball, mais c'est relativement facile dans des domaines comme le droit, la médecine, la technologie ou les finances.

AIM se calcule directement par le R^2 d'une régression linéaire des rentabilités sur les prix :

$$r_{i,t} = \alpha_i + \beta_i^i P_{i,t-1} + \sum_{j \in J, j \neq i} \beta_i^j P_{j,t-1} + \epsilon_{i,t} \quad \Rightarrow \quad \text{AIM} = R^2$$

log_market_cap

Il s'agit du logarithme népérien de la capitalisation boursière de l'entreprise. Plus une entreprise est grande plus elle possède une grande capitalisation. Or, Clarke et Shastri (2001) montrent que les résultats sont sensibles à la taille des entreprises. Ces auteurs émettent des doutes quant à la capacité des mesures basées sur les prévisions des analystes à déterminer la quantité d'information privée des titres.

Ainsi la market capitalization est une mesure d'asymétire informationelle.

log_btm

Il s'agit du logarithme néprien du book-to-market. Le book-to-market est défini de la façon suivante :

$$BooktoMarket = \frac{BookValue}{MarketCapitalization}$$

Les dirigeants d'entreprises en forte croissance ont une information plus précise sur les opportunités d'investissement (Smith et Watts, 1992). Le market-to-book est donc une mesure de l'asymétrie d'information.

fsrv

1. Risque total moins risque systématique.
2. FSRV : risque total moins le risque expliqué par les variations du portefeuille de marché et d'un portefeuille de titres appartenant au même secteur d'activité que le titre étudié.

$$r_{i,j,t} = \beta_{j,0} + \beta_{j,m} r_{m,t} + \beta_{j,i} r_{i,t} + \epsilon_{i,j,t}$$

$FSRV = 1 - R^2$ de cette régression.

turnover

Le taux de rotation est un indicateur important pour déterminer le caractère plus ou moins actif d'une gestion. Il mesure le volume de transactions effectuées par le gérant sur une année, en proportion de l'actif du fonds. Il y a plusieurs manières de le calculer, notamment parce que flux de souscriptions et rachats engendrent structurellement des transactions, sans qu'il y ait nécessairement de changement dans la composition du portefeuille. Or, le volume anormal d'ordres d'achat/vente génère de l'information privée.

pin

1. Probabilité pour le MM que l'échange sur un titre soit basé sur de l'information privée.
2. Cette mesure est basée sur des modèles de microstructure qui interprètent le volume anormal d'ordres d'achat/vente comme contenant de l'information privée.
3. PIN: fraction des ordres issus d'investisseurs informés

$$PIN = \frac{\alpha\mu}{\alpha\mu + \epsilon_b + \epsilon_s}$$

L'inconvénient du PIN est qu'il est souvent un proxy du volume de transaction et qu'il mesure plus la liquidité et le volume des transactions que la mesure de l'asymétrie d'information.

Statistique descriptive des mesures

```
##                  return_rf          beta           AIM log_market_cap
## nbr.val       1.394130e+05  1.394150e+05  1.392390e+05   1.394150e+05
## nbr.null      1.000000e+01  0.000000e+00  0.000000e+00   0.000000e+00
## nbr.na        2.000000e+00  0.000000e+00  1.760000e+02   0.000000e+00
## min          -3.948270e+00 -4.793150e+00 -7.736740e+00   5.024300e+00
## max           2.137420e+00  8.053500e+00  1.747150e+00   1.927560e+01
## range         6.085690e+00  1.284665e+01  9.483890e+00   1.425130e+01
## sum           7.773300e-01  1.191842e+05 -3.678742e+05   1.739736e+06
## median        1.620000e-03  7.659400e-01 -2.631040e+00   1.255050e+01
## mean          5.575735e-06  8.548881e-01 -2.642034e+00   1.247883e+01
## SE.mean       3.290062e-04  1.509545e-03  1.923623e-03   5.061551e-03
## CI.mean.0.95  6.448459e-04  2.958680e-03  3.770265e-03   9.920545e-03
## var           1.509077e-02  3.176886e-01  5.152297e-01   3.571715e+00
## std.dev       1.228445e-01  5.636388e-01  7.177950e-01   1.889898e+00
## coef.var      2.203198e+04  6.593129e-01 -2.716827e-01   1.514484e-01
##                    log_btm          fsrv      turnover          pin
## nbr.val       1.280080e+05 1.083400e+05  1.394150e+05 4.684100e+04
## nbr.null      0.000000e+00 0.000000e+00  2.224000e+03 0.000000e+00
## nbr.na        1.140700e+04 3.107500e+04  0.000000e+00 9.257400e+04
## min          -9.758060e+00 2.657500e-01 -3.000000e-04 1.930870e-02
## max           6.806020e+00 1.000000e+00  3.216780e+01 6.201992e-01
## range         1.656408e+01 7.342500e-01  3.216810e+01 6.008905e-01
## sum          -3.837368e+04 9.474821e+04  8.090623e+03 8.033490e+03
## median       -2.846200e-01 9.152750e-01  2.830000e-02 1.607668e-01
## mean         -2.997757e-01 8.745450e-01  5.803266e-02 1.715055e-01
## SE.mean       2.936698e-03 3.777278e-04  4.742479e-04 2.932832e-04
## CI.mean.0.95  5.755876e-03 7.403411e-04  9.295169e-04 5.748394e-04
## var           1.103966e+00 1.545777e-02  3.135598e-02 4.029031e-03
## std.dev       1.050698e+00 1.243293e-01  1.770762e-01 6.347465e-02
## coef.var     -3.504947e+00 1.421645e-01  3.051320e+00 3.701026e-01
```

Pour commencer nos analyses statistiques nous réalisons quelques calculs de statistiques descriptive pour éviter tous problèmes d'hétérogéinité des données.

Nombre de valeurs

Pour la première étude d'hétérogéinité, intérressons-nous aux nombres de valeurs présentes dans le fichier. Les valeurs les plus présentes dans le fichier sont return_rf, beta, le turnover, log_market_cap et enfin le nombre de coefficient AIM quasiment égal au nombre de coefficient précédents.

Les premières valeurs étant facilement calculables, il est normal que nous ayons toutes les données présentes. En revanche, le coefficient AIM, d'asymétrie d'information est plus complexe à calculer et est généralement présent seulement pour lesgrandes entreprises. Cela veut dire que notre échantillon de données comporte un grand nombre de grandes entreprises étant donné le nombre de coefficients. Cela va donc être plus difficile de se rendre compte que ce coefficient impact sur les prix Car pour de grandes entreprises qui sont connues, l'information est plus homogène entre les investisseurs et donc ce coefficient n'impacte que très peu le prix.

En revanche on constate que le PIN est une variable très peu représentée dans cet échantillon de données. Nous appliquons donc des calculs différents lorsque nous intégrons ce paramètre étant donnée qu'il réduit de deux tiers les données. Le paramètre fsrv possède également un déficit d'échantillons. Il perd environ un tiers des données.

Statistiques descriptives

En moyenne, l'espérance de rentabilités des titres soustrait à la rentabilité du placement sans risque est quasiment nul : 5.575735×10^{-6}, cela montre que le marché est efficient. Et que toute l'information des titres est contenue dans le prix. Cela va encore aller à l'encontre de l'estimation de significativité du AIM qui représente l'hétérogéinité de l'information des investisseurs. Cependant on peut tout de même remarquer que ce coefficient est en moyenne non nul et négatif -2.642034 avec un intervalle de confiance à 95% de 3.770265×10^{-3}. On peut donc dire que le coefficient directeur lors de la régression linéaire sera de valeur absolue petite, sinon auquel cas, ce coefficient aura une trop grande importance sur les rentabilités. Bien sûr, tout ceci sous l'hypothèse d'une p-value faible.

Corrélation entre les différentes mesures

Pour éviter qu'un partie trop importante des informations du marché ne soit tronquée par R en calculant la corrélation de tous les vecteurs en même temps (R ne garde que les lignes où toutes les colonnes ont bien des valeurs), nous avons opté pour un calcul des corrélations deux à deux, puis nous les regroupons dans une seule et même matrice de corrélation pour puvoir visualiser l'ensemble.

Afin de nous assurer de la significativité des résultats, nous opérons aussi un test de moment-produit de corrélation de Pearson à probabilité de rejet $\alpha = 5\%$.

```
##                  return_rf        beta         AIM log_market_cap
## return_rf       1.000000000 -0.02310218  0.02139276    0.006209259
## beta           -0.023102181  1.00000000 -0.13555456    0.032917841
## AIM             0.021392757 -0.13555456  1.00000000   -0.326080042
## log_market_cap  0.006209259  0.03291784 -0.32608004    1.000000000
## log_btm         0.028659371 -0.06215395  0.17277822   -0.278523728
## fsrv           -0.018463787 -0.13097743  0.29501695   -0.573238875
## turnover       -0.040818441  0.16160967 -0.13091794    0.060400787
## pin            -0.010510582  0.09104824  0.09324622   -0.599796578
##                    log_btm        fsrv    turnover         pin
## return_rf       0.02865937 -0.01846379 -0.04081844 -0.01051058
## beta           -0.06215395 -0.13097743  0.16160967  0.09104824
## AIM             0.17277822  0.29501695 -0.13091794  0.09324622
## log_market_cap -0.27852373 -0.57323887  0.06040079 -0.59979658
## log_btm         1.00000000  0.13158239 -0.10722750  0.20234392
## fsrv            0.13158239  1.00000000 -0.07661446  0.26401429
## turnover       -0.10722750 -0.07661446  1.00000000 -0.06188466
## pin             0.20234392  0.26401429 -0.06188466  1.00000000
```

```
##                  return_rf          beta          AIM log_market_cap
## return_rf      0.000000e+00  6.299434e-18 1.423142e-15   2.042662e-02
## beta           6.299434e-18  0.000000e+00 0.000000e+00   9.732800e-35
## AIM            1.423142e-15  0.000000e+00 0.000000e+00   0.000000e+00
## log_market_cap 2.042662e-02  9.732800e-35 0.000000e+00   0.000000e+00
## log_btm        1.115243e-24 9.254331e-110 0.000000e+00   0.000000e+00
## fsrv           1.218767e-09  0.000000e+00 0.000000e+00   0.000000e+00
## turnover       1.723007e-52  0.000000e+00 0.000000e+00  7.975800e-113
## pin            2.292032e-02  8.705482e-87 7.145681e-91   0.000000e+00
##                    log_btm          fsrv      turnover          pin
## return_rf       1.115243e-24  1.218767e-09  1.723007e-52 2.292032e-02
## beta           9.254331e-110  0.000000e+00  0.000000e+00 8.705482e-87
## AIM             0.000000e+00  0.000000e+00  0.000000e+00 7.145681e-91
## log_market_cap  0.000000e+00  0.000000e+00 7.975800e-113 0.000000e+00
## log_btm         0.000000e+00  0.000000e+00 9.881313e-324 0.000000e+00
```

```
## fsrv           0.000000e+00  0.000000e+00 1.010328e-140 0.000000e+00
## turnover      9.881313e-324 1.010328e-140  0.000000e+00 5.572616e-41
## pin            0.000000e+00  0.000000e+00  5.572616e-41 0.000000e+00
```

Interpretation

Significativité

On remarque que l'ensemble des p-valeurs des tests de corrélation sont inférieurs à 5%, on en déduit donc que les corrélations entre les différentes mesures sont bel et bien significatives.

Corrélations des rentabilités

avec les risques systématiques β

Étonnament, la corrélation est significativement négative. On aurait pu s'attendre à une corrélation positive, étant donnée que plus un titre est risqué, plus grande est la rentabilité exigée par les investisseurs.

avec la capitalisation du titre (size)

Un autre résultat inattendu, la corrélation ici est positive, alors qu'on s'attendait à l'inverse, pensant que moins le titre était connu, plus son asymétrie informationnelle serait élevée, et donc plus la rentabilité serait importante.

avec le reste des mesures

Les résultats concordent à nos prévisions. La corrélation est bien confirmée avec les rentabilités.

Régressions Linéaires

Dans cette partie, nous faisons des régressions linéaires des rentabilités en fonction des mesures d'asymétrie informationnelle afin de valider leurs corrélation et mesurer l'impact de chacune des mesures sur les rentabilités.

Nous opérons la régression à chaque date (sans doublon) de notre grille pour observer si il y a bien corrélation. Nous le faisons aussi pour tous les couples (mesure_x , AIM) pour détecter, dans l'éventualité où AIM serait en effet une mesure qui approcherait un risque pricé par le marché, si ce n'est qu'un "proxy" de mauvaise qualité de la mesure mesure_x.

Enfin, pour conclure une réponse définitive, nous faisons un test de Student pour tester la nullité des coefficients.

renta ~ β

```
##
##  One Sample t-test
##
## data:  betas
## t = -1.9195, df = 226, p-value = 0.05618
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  -0.0119500065  0.0001567284
## sample estimates:
```

```
##    mean of x
## -0.005896639
```

```
## $test
## [1] "One Sample Weighted T-Test"
##
## $coefficients
##     t.value          df     p.value
##  -1.6285113 226.0000000   0.1048097
##
## $additional
##   Difference         Mean  Alternative     Std. Err
## -0.004394852 -0.004394852  0.000000000  0.002698693
```

Interpretation

Dans un premier temps, nous souhaitons analyser les risques systematiques, s'il y a bien un lien positive avec la rentabilité pour répondre à l'équation imposée par l'équilibre des marchés financiers, le MEDAF traditionnel.

Les résultats non pondérés par les poids nous donne un coefficient négatif de -0.005897 pour une p-value de 0.05618 qui nous donne l'information que ce coefficient est significativement différent de zéro et que l'on peut interpréter. Donc notre portefeuille répond bien à l'équation du MEDAF traditionnel. Étant donné qu'il a une valeur négative, il s'agit ici d'une valeur "refuge". Cela veut donc dire que le portefeuille varie de façon opposé au marché.

Les résultats pondérés par les poids nous donne un coefficient négatif de -0.004394 pour une p-value de 0.1048 qui nous donne l'information que ce coefficient n'est pas totalement significativement différent de zéro, on ne peut donc pas l'interpétrer avec une très grande certitude. On peut donner les mêmes interprétations que que pour le coefficient non pondéré par les poids, on remarque que le coefficient a diminué en valeur absolue pour se rapprocher de zéro. Cela veut dire que les valeurs les plus significatives de beta ont tendance à diminuer la dépendance de beta sur les rentabilités.

Cependant, nos p-valeur reste tout de même très petites, cela veut donc dire que le MEDAF est bien respecté. Nous pouvons donc étudier d'autres facteurs liés au rentabilités. Nous porterons donc notre choix sur les facteurs de mesure d'asymétrie d'information.

renta ~ AIM

```
##
##  One Sample t-test
##
## data:  AIMs
## t = -0.6521, df = 226, p-value = 0.515
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  -0.002649552  0.001331958
## sample estimates:
##     mean of x
## -0.0006587974
```

```
## $test
## [1] "One Sample Weighted T-Test"
##
## $coefficients
##     t.value          df     p.value
```

```
##  -0.6444166 226.0000000   0.5199589
##
## $additional
##    Difference          Mean   Alternative      Std. Err
## -0.0005831366 -0.0005831366  0.0000000000  0.0009049062
```

Interpretation

AIM est notre paramètre coeur de notre analyse. Nous souhaitons donc vérifier la validité de paramètre en tant que mesure d'asymétrie informationelle.

Cependant, au vue de la valeur et de la p-valeur du test de Student avec ou sans poids, on ne peut pas valider la dépendance du prix en fonction de ce paramètre. En effet, le coefficient gamma n'est que de -0.000658 pour la version non pondérée par les poids et -0.000583 pour la version pondéré. La p-valeur pour les deux est d'environ un demi, nous avons donc une chance sur deux que l'hypothèse d'observation de nos valeurs sachant que la moyenne donnée est correcte soit vraie, ce qui est non significatif. Comme dit lors de la première partie, cela est surement dû à l'échantillon qui contient beaucoup de grandes entreprises.

AIM ne représente donc pas un risque pricé par le marché, les investisseurs n'exigent pas une prime pour ce risque. On ne peut conclure que c'est une bonne mesure d'asymétrie d'information.

renta ~ β + AIM

```
##
##  One Sample t-test
##
## data:  AIMs
## t = -1.0609, df = 226, p-value = 0.2899
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  -0.002960155  0.000888192
## sample estimates:
##    mean of x
## -0.001035981
```

```
##
##  One Sample t-test
##
## data:  betas
## t = -2.0022, df = 226, p-value = 0.04646
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  -1.214018e-02 -9.687768e-05
## sample estimates:
##    mean of x
## -0.006118527
```

Interpretation

Cette fois-ci nous décidons d'introduire bêta. Toujours pour tester la validité d'AIM. On remarque que la valeur absolue de bêta diminue quant à la valeur absolue d'AIM augmente :

L'interprétation purement mathématiques de ce phénomène est que Bêta est une fonction d'AIM et qu'il vient brouiller la régression linéaire. On peut alors tester avec d'autres facteurs pour voir si ce n'est pas juste un mauvais proxy d'une autre mesure.

Variable	Équipondéré	Pondération
AIM	-0.001035981	-0.0008879540
AIM p-value	0.2899	0.3105887
Bêta	-0.006118527	-0.004613529
Bêta p-value	0.04646	0.08703831

renta ~ size

Par le biais de cette régression nous souhaitons valider la conjecture de Clarke et Shastri. Ainsi, dans un premier temps nous vérifions si la taille de l'entreprise est intégrée dans le prix du titre et que les investisseurs sont prêts à payer une prime en fonction de ce risque.

```
##
##  One Sample t-test
##
## data:  Sizes
## t = 2.5584, df = 226, p-value = 0.01117
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  0.0004190833 0.0032284492
## sample estimates:
##   mean of x
## 0.001823766
```

```
## $test
## [1] "One Sample Weighted T-Test"
##
## $coefficients
##      t.value           df      p.value
##   2.47024255 226.00000000   0.01424201
##
## $additional
##   Difference         Mean  Alternative     Std. Err
## 0.0015954140 0.0015954140 0.0000000000 0.0006458532
```

Interpretation

Variable	Équipondéré	Pondération
Size	0.001823766	0.0015954140
Size p-value	0.01117	0.01424201

Or, au calcul du coefficient, nous observons une moyenne du coefficient de 0.001823766 avec une moyenne de 12.4 pour la size on peut dire qu'en moyenne elle un impact de $0,02232$ sur le prix. La p-valeur étant de 1% on peut alors valider l'hypothèse que ce coefficient est non nul. Il s'agit d'une mesure impactant le prix et dont les investisseurs exigent une prime de risque. Cela valide donc l'hypoyhèse émise par Clarke et Shastri.

renta ~ size + AIM

Vérifions si AIM n'est pas simplement un mauvais proxy de la taille.

```
##
##  One Sample t-test
##
```

```
## data:  AIMs
## t = 0.58005, df = 226, p-value = 0.5625
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  -0.001459902  0.002677947
## sample estimates:
##    mean of x
## 0.0006090225
```

```
## $test
## [1] "One Sample Weighted T-Test"
##
## $coefficients
##     t.value          df     p.value
##   0.5456760 226.0000000   0.5858268
##
## $additional
##   Difference         Mean  Alternative     Std. Err
## 0.0005065829 0.0005065829 0.0000000000 0.0009283583
```

```
##
##  One Sample t-test
##
## data:  Sizes
## t = 2.5409, df = 226, p-value = 0.01173
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  0.0004227331 0.0033433778
## sample estimates:
##   mean of x
## 0.001883055
```

```
## $test
## [1] "One Sample Weighted T-Test"
##
## $coefficients
##      t.value           df      p.value
##   2.44906798 226.00000000   0.01508366
##
## $additional
##   Difference         Mean  Alternative     Std. Err
## 0.0016313873 0.0016313873 0.0000000000 0.0006661258
```

Interpretation

Étant donné que les valeurs ne changent pas significativement lorsqu'elles sont couplées ou découplées, cela veut dire que la taille représente bien un facteur sur la rentabilité mais qu'il n'est une foncion déterministe d'AIM. Essayon d'inclure le book-to-market comme mesure d'asymétrie.

renta ~ BTM

D'après les hypothèses de Fama & French, les primes book-to-market des petites (SMB) et des grandes entreprises (HML) obtenues sont en moyenne positives. Le modèle empirique d'évaluation des actifs financiers de FF est:

$$R_{it} = \alpha_i + \beta_{im}(R_m - R_f) + \beta_{is}\text{SMB}_t + \beta_{ih}\text{HML}_t + \epsilon_{it}$$

Nous nous attendons donc à avoir un coefficient positif ou nul.

```
##
##  One Sample t-test
##
## data:  BTM
## t = -1.0635, df = 226, p-value = 0.2887
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  -0.005012041  0.001498409
## sample estimates:
##    mean of x
## -0.001756816
```

```
## $test
## [1] "One Sample Weighted T-Test"
##
## $coefficients
##     t.value          df     p.value
##  -1.3268507 226.0000000   0.1858971
##
## $additional
##   Difference         Mean  Alternative     Std. Err
## -0.002003460 -0.002003460  0.000000000  0.001509937
```

Interpretation

Pas de coefficients positifs et une valeur non significative puisque la p-value est bien trop grande. On peut seulement dire que sur cet échantillon le book-to-market n'est pas une mesure significative et on ne peut donc pas interpréter ce résultat.

BTM ~ AIM

Vérifions si AIM est un proxy du book to market

```
##
##  One Sample t-test
##
## data:  AIMs
## t = 15.398, df = 226, p-value < 2.2e-16
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  0.09368641 0.12118432
## sample estimates:
## mean of x
## 0.1074354
```

```
## $test
## [1] "One Sample Weighted T-Test"
##
## $coefficients
##   t.value        df   p.value
```

```
##  15.67081 226.00000   0.00000
##
## $additional
##  Difference        Mean Alternative    Std. Err
## 0.106775301 0.106775301 0.000000000 0.006813643
```

Interpretation

Dans un premier temps, on remarque qu'à une p-valeur inférieure à 2.2×10^{-16} on a un coefficient moyen de 0.106. Cela veut clairement dire que AIM est une fonction du book-to-market. Ainsi, on peut remplacer le book-to-market par $0.106 \times$ AIM. Ainsi, étant donné le test de Student sur AIM, il est normal que l'on trouve un book-to-market qui n'impacte pas le prix.

renta ~ PIN

Maintenant, nous souhaitons vérifier si le PIN est une variable dont les investisseurs exigent une prime de risque.

```
##
##  One Sample t-test
##
## data:  Pins
## t = -1.2023, df = 226, p-value = 0.2305
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  -0.05589810  0.01353367
## sample estimates:
##   mean of x
## -0.02118221
```

```
## $test
## [1] "One Sample Weighted T-Test"
##
## $coefficients
##     t.value          df     p.value
##  -0.5982073 226.0000000   0.5503007
##
## $additional
##   Difference         Mean  Alternative     Std. Err
## -0.009474665 -0.009474665  0.000000000  0.015838432
```

Interpretation

Pour une p-valeur de 23%, on ne peut pas dire que le coefficient est significativement différent de zéro. On ne peut donc dire que le PIN est un facteur de risque dont les investisseurs demandent une prime.

renta ~ PIN + AIM

Vérifions si AIM est un proxy du PIN

```
##
##  One Sample t-test
##
## data:  AIMs
```

```
## t = -0.20172, df = 226, p-value = 0.8403
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  -0.002152207  0.001752488
## sample estimates:
##     mean of x
## -0.0001998592
```

```
## $test
## [1] "One Sample Weighted T-Test"
##
## $coefficients
##     t.value          df     p.value
##  -0.4030484 226.0000000   0.6872937
##
## $additional
##    Difference          Mean   Alternative      Std. Err
## -0.0003557463 -0.0003557463  0.0000000000  0.0008826392
```

```
##
##  One Sample t-test
##
## data:  Pins
## t = -1.1174, df = 226, p-value = 0.265
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  -0.05483727  0.01514910
## sample estimates:
##   mean of x
## -0.01984408
```

```
## $test
## [1] "One Sample Weighted T-Test"
##
## $coefficients
##     t.value          df     p.value
##  -0.5130992 226.0000000   0.6083831
##
## $additional
##  Difference        Mean Alternative    Std. Err
## -0.00813942 -0.00813942  0.00000000  0.01586325
```

Interpretation

Les p-values sont trop élevées, toutes supérieures à 25%, dans les modèles pondérés encore plus, elles sont supérieures à 60%. On ne peut donc rien conclure.

renta ~ FSRV

Voyons voir maintenant si la mesure FSRV exprime un risque pricé par le marché :

```
##
##  One Sample t-test
##
## data:  FSRVs
```

```
## t = -1.1173, df = 226, p-value = 0.265
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  -0.026355673  0.007282177
## sample estimates:
##    mean of x
## -0.009536748
```

```
## $test
## [1] "One Sample Weighted T-Test"
##
## $coefficients
##    t.value         df    p.value
##  -1.313106 226.000000   0.190479
##
## $additional
##   Difference         Mean  Alternative     Std. Err
## -0.009355549 -0.009355549  0.000000000  0.007124749
```

Interpretation

Coefficients relativement importants (≈ -0.01), mais encore une fois, les p-values sont trop élevées pour pouvoir conclure.

renta ~ FSRV + AIM

Vérifions si AIM est un proxy du FSRV

```
##
##  One Sample t-test
##
## data:  AIMs
## t = -0.38181, df = 226, p-value = 0.703
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  -0.002352696  0.001588950
## sample estimates:
##     mean of x
## -0.0003818729
```

```
## $test
## [1] "One Sample Weighted T-Test"
##
## $coefficients
##     t.value          df     p.value
##  -0.1966401 226.0000000   0.8442859
##
## $additional
##    Difference          Mean   Alternative      Std. Err
## -0.0001743412 -0.0001743412  0.0000000000  0.0008866004
```

```
##
##  One Sample t-test
##
## data:  FSRVs
```

```
## t = -1.0337, df = 226, p-value = 0.3024
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  -0.025624402  0.007991108
## sample estimates:
##    mean of x
## -0.008816647
```

```
## $test
## [1] "One Sample Weighted T-Test"
##
## $coefficients
##     t.value          df     p.value
##  -1.2851886 226.0000000   0.2000419
##
## $additional
##   Difference         Mean  Alternative     Std. Err
## -0.009138641 -0.009138641  0.000000000  0.007110740
```

Interpretation

Les p-values sont trop élevées, toutes supérieures à 20%. On ne peut donc rien conclure. Néanmoins, les coefficients tendent à exprimer que AIM est effectivement un proxy de FSRV.

renta ~ Turnover

Voyons voir maintenant si la mesure Turnover exprime un risque pricé par le marché :

```
##
##  One Sample t-test
##
## data:  Turnovers
## t = -2.1807, df = 226, p-value = 0.03023
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  -0.096420421 -0.004882894
## sample estimates:
##   mean of x
## -0.05065166
```

```
## $test
## [1] "One Sample Weighted T-Test"
##
## $coefficients
##     t.value          df     p.value
##  -1.4315276 226.0000000   0.1536605
##
## $additional
##  Difference        Mean Alternative    Std. Err
## -0.02659243 -0.02659243  0.00000000  0.01857626
```

Interpretation

Dans le modèle de la moyenne équirépartie, la p-value est de 3.02%, ce qui exprime que le Turnover est bien significative, mais négativement corrélées à la rentabilité.

renta ~ Turnover + AIM

Vérifions si AIM est un proxy du Turnover

```
##
##  One Sample t-test
##
## data:  Turnovers
## t = -2.2996, df = 226, p-value = 0.02238
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  -0.099117943 -0.007638946
## sample estimates:
##   mean of x
## -0.05337844
```

```
## $test
## [1] "One Sample Weighted T-Test"
##
## $coefficients
##     t.value          df     p.value
##  -1.5551128 226.0000000   0.1213184
##
## $additional
##  Difference        Mean Alternative    Std. Err
## -0.02878352 -0.02878352  0.00000000  0.01850896
```

```
##
##  One Sample t-test
##
## data:  AIMs
## t = -1.7136, df = 226, p-value = 0.08797
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  -0.003542429  0.000247018
## sample estimates:
##    mean of x
## -0.001647705
```

```
## $test
## [1] "One Sample Weighted T-Test"
##
## $coefficients
##     t.value          df     p.value
##  -1.4231178 226.0000000   0.1560809
##
## $additional
##    Difference          Mean   Alternative      Std. Err
## -0.0012257732 -0.0012257732  0.0000000000  0.0008613294
```

Interpretation

Pour le modèle de moyenne équirépartie, les p-values sont bien inférieures au seuil de 5%, les coefficients des deux mesures de mêmes signes (toutes deux négatives) et $|\gamma_{AIM}| < |\gamma_{Turnover}|$. On en déduit que AIM est bien un mauvais proxy de Turnover.

Par contre, dans le modèle de moyenne pondéré, qui est intuitivement plus significatif que le précédent, les p-values sont certes faibles mais supérieures au seul de 5%. On ne peut donc pas faire de conclusion similaire.

renta ~ β + AIM + Size

Testons la significativité de la mesure AIM en comparaison avec β et Size

```
##
##  One Sample t-test
##
## data:  AIMs
## t = 0.22796, df = 226, p-value = 0.8199
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  -0.001722842  0.002173601
## sample estimates:
##    mean of x
## 0.0002253794
```

```
## $test
## [1] "One Sample Weighted T-Test"
##
## $coefficients
##     t.value          df     p.value
##   0.2451769 226.0000000   0.8065420
##
## $additional
##   Difference         Mean  Alternative     Std. Err
## 0.0002143930 0.0002143930 0.0000000000 0.0008744424
```

```
##
##  One Sample t-test
##
## data:  betas
## t = -2.0606, df = 226, p-value = 0.04048
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  -0.0121005386 -0.0002705585
## sample estimates:
##    mean of x
## -0.006185549
```

```
## $test
## [1] "One Sample Weighted T-Test"
##
## $coefficients
##      t.value           df      p.value
##  -1.74819862 226.00000000   0.08178755
##
## $additional
##   Difference         Mean  Alternative     Std. Err
## -0.004610503 -0.004610503  0.000000000  0.002637288
```

```
##
##  One Sample t-test
##
```

```
## data:  Sizes
## t = 2.5233, df = 226, p-value = 0.01231
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  0.0004084032 0.0033199229
## sample estimates:
##   mean of x
## 0.001864163
```

```
## $test
## [1] "One Sample Weighted T-Test"
##
## $coefficients
##     t.value          df     p.value
##   2.4540515 226.0000000   0.0148817
##
## $additional
##  Difference        Mean Alternative    Std. Err
## 0.001618513 0.001618513 0.000000000 0.000659527
```

renta ~ β + AIM + Size + Turnover

Testons la significativité de la mesure AIM en comparaison avec β, Size et Turnover

```
##
##  One Sample t-test
##
## data:  AIMs
## t = -0.26344, df = 226, p-value = 0.7924
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  -0.002132821  0.001629789
## sample estimates:
##     mean of x
## -0.0002515162
```

```
## $test
## [1] "One Sample Weighted T-Test"
##
## $coefficients
##     t.value          df     p.value
##  -0.0870964 226.0000000   0.9306720
##
## $additional
##    Difference          Mean   Alternative      Std. Err
## -7.356248e-05 -7.356248e-05  0.000000e+00  8.446099e-04
```

```
##
##  One Sample t-test
##
## data:  betas
## t = -1.8454, df = 226, p-value = 0.06628
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  -0.0106947974  0.0003505401
```

```
## sample estimates:
##    mean of x
## -0.005172129
```

```
## $test
## [1] "One Sample Weighted T-Test"
##
## $coefficients
##      t.value           df      p.value
##  -1.71429810 226.00000000   0.08784494
##
## $additional
##   Difference         Mean  Alternative     Std. Err
## -0.004273095 -0.004273095  0.000000000  0.002492621
```

```
##
##  One Sample t-test
##
## data:  Sizes
## t = 2.8717, df = 226, p-value = 0.004471
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  0.0006915143 0.0037157614
## sample estimates:
##   mean of x
## 0.002203638
```

```
## $test
## [1] "One Sample Weighted T-Test"
##
## $coefficients
##      t.value           df      p.value
## 2.723076e+00 2.260000e+02 6.972612e-03
##
## $additional
##   Difference         Mean  Alternative     Std. Err
## 0.0018500701 0.0018500701 0.0000000000 0.0006794046
```

```
##
##  One Sample t-test
##
## data:  Turnovers
## t = -1.9869, df = 226, p-value = 0.04814
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  -0.0913284663 -0.0003782455
## sample estimates:
##   mean of x
## -0.04585336
```

```
## $test
## [1] "One Sample Weighted T-Test"
##
## $coefficients
##     t.value          df     p.value
##  -1.2217622 226.0000000   0.2230704
```

```
##
## $additional
##  Difference        Mean Alternative    Std. Err
## -0.02223748 -0.02223748  0.00000000  0.01820115
```

Interpretation

Dans les deux tests, la p-value pour notre mesure étudiée AIM est supérieure à 80%, tandis que les autres p-values sont nettement inférieurs et donc leurs coefficients plus significatifs.

Conclusion de l'étude empirique

Dans l'ensemble, les données du marché qui nous ont été fournies semblent un peu ambigües et ne permettent pas d'établir un lien clair et net entre la mesure AIM et l'asymétrie informationnelle de ce marché. Néanmoins, les différentes études comparatives menées entre l'AIM et les autres mesures classiques confirment nos espérances à l'issue de l'étude théorique et valident que l'AIM est bel et bien un candidat sérieux et ambitieux pour mesurer l'asymétrie informationnelle.