



GRENOBLE INP - ENSIMAG

École nationale supérieure d'informatique et de mathématiques appliquées

MICROSTRUCTURE ET MÉCANISME DES MARCHÉS FINANCIERS

3^{ÈME} ANNÉE - INGÉNIERIE POUR LA FINANCE

Projet *Asymétrie informationnelle*

07 avril 2018

Equipe

AMRAOUI Samy | samy.amraoui@grenoble-inp.org
BERNARD Julien | julien.bernard@grenoble-inp.org
CAMUS Julien | julien.camus@grenoble-inp.org
MOINE Yoann | yoann.moine@grenoble-inp.org

Table des matières

1	Introduction	2
2	Aspects théoriques	2
2.1	Modèle de Grossman et Stiglitz (1980)	2
2.2	Généralisation dans un cadre multi-actifs	3
2.2.1	Hypothèses	3
2.2.2	Calcul du prix d'équilibre \tilde{P}^0	3
2.3	Étude de $Cov(\tilde{P}^1 - \tilde{P}^0, \tilde{P}^0)$	8
2.3.1	$Cov(\tilde{P}^1 - \tilde{P}^0, \tilde{P}^0)$ fonction de N	8
2.3.2	$Cov(\tilde{P}^1 - \tilde{P}^0, \tilde{P}^0)$ selon le degré d'asymétrie informationnelle	9
2.3.3	Application numérique	10
3	Aspects empiriques	11
3.1	Corrélations entre différentes mesures d'asymétrie informationnelle	11
3.2	Analyse de la mesure AIM par régression linéaire	12
3.2.1	Explications	12
3.2.2	Régression sur bêta	13
3.2.3	Régression sur bêta et AIM	14
3.2.4	Régression sur bêta, AIM et PIN	14
3.2.5	Régression sur bêta et AIM quand on connaît PIN	14
3.2.6	Régression sur bêta, AIM et la taille	15
3.2.7	Régression sur bêta, AIM et BTM	15
3.2.8	Régression sur bêta, AIM et FSRV	15
3.2.9	Régression sur bêta, AIM et illiq_amihud	16
3.2.10	Régression sur bêta, AIM et turnover	16
3.2.11	Régression sur bêta, AIM, taille et FSRV	16
3.2.12	Régression sur bêta, AIM, taille, FSRV et turnover	17
3.2.13	Régression sur bêta, AIM, PIN, taille, BTM, FSRV, turnover et illiq_amihud	17
3.3	Stratégies actives de gestion de portefeuilles	17
4	Conclusion	18
5	Annexes	19
5.1	Annexe 1	19
5.2	Annexe 2	19

1 Introduction

Ce projet a pour objectif d'étudier une mesure alternative du degré d'asymétrie informationnelle d'un titre (AIM) qui est fondée sur les modèles d'équilibre à anticipations rationnelles (EAR). 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 :

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

Cette mesure est approchée théoriquement par la covariance entre rentabilité et prix :

$$Cov(\tilde{P}^1 - \tilde{P}^0; \tilde{P}^0)$$

Dans cette étude, nous allons dans un premier temps justifier l'utilisation de cette mesure sur un plan théorique pour, dans un second temps, procéder de même sur un plan pratique.

2 Aspects théoriques

A l'aide du modèle de Grossman et Stiglitz (1980) dans un cadre mono-actif que nous allons généraliser dans un cadre multi-actifs, nous allons montrer que $Cov(\tilde{P}^1 - \tilde{P}^0; \tilde{P}^0)$ est négative et décroît avec le degré d'asymétrie informationnelle.

2.1 Modèle de Grossman et Stiglitz (1980)

Grossman et Stiglitz (1980) ont proposé un modèle mono-actif dans lequel une proportion λ d'investisseurs **informés** connaît la distribution \tilde{P}^1 des revenus futurs de l'unique actif risqué présent sur le marché qui s'écrit alors :

$$\tilde{P}^1 = \tilde{\theta} + \tilde{\epsilon}$$

où $\tilde{\theta}$ et $\tilde{\epsilon}$ suivent des lois normales indépendantes.

Ainsi, en notant :

- \tilde{z} la variable aléatoire représentant l'offre ;
- a le coefficient d'aversion au risque des investisseurs (supposé identique pour chaque investisseur) ;
- $\sigma_\epsilon^2 = Var(\tilde{\epsilon})$ la variance du bruit $\tilde{\epsilon}$;
- $\sigma_\theta^2 = Var(\tilde{\theta})$ la variance de la variable aléatoire $\tilde{\theta}$ corrélée avec le revenu futur ;
- $\sigma_z^2 = Var(\tilde{z})$ la variance de l'offre \tilde{z} ;
- $R = 1 + r_f$ le revenu de l'actif sans risque en $t = 1$;

il existe une unique solution analytique pour la distribution initiale \tilde{P}^0 de l'unique actif risqué dans la classe des fonctions linéaires en $\tilde{\theta}$ et \tilde{z} donnée par :

$$\tilde{P}^0(\tilde{\theta}, \tilde{z}) = A_0 + A_1 \left(\tilde{\theta} - \frac{a}{\lambda} \sigma_\epsilon^2 \tilde{z} \right) = A_0 + A_1 \tilde{w}(\tilde{\theta}, \tilde{z})$$

où A_0 et A_1 sont des réels constants et :

- \tilde{w} est une combinaison linéaire de variables aléatoires indépendantes suivant une loi normale, donc suit elle-même une loi normale ;

- $A_1 = \frac{1}{R}(\sigma_\theta^2 + \sigma_\epsilon^2 - V_m)(\sigma_\theta^2)^{-1}$;
- $A_0 = \frac{1}{R} \left[(1 - RA_1)\mathbb{E}(\tilde{\theta}) + \left(\frac{Ra}{\lambda} A_1 \sigma_\epsilon^2 - aV_m \right) \mathbb{E}(\tilde{z}) \right]$;

avec :

- $V_m = [\lambda(\sigma_\epsilon^2)^{-1} + (1 - \lambda)(N + \sigma_\epsilon^2)^{-1}]^{-1}$;
- $N = Var(\tilde{\theta}|P^0) = \sigma_\theta^2 - \sigma_\theta^4(Var^{-1}(\tilde{w}))$;
- $Var(\tilde{w}) = \sigma_\theta^2 + \frac{a^2}{\lambda^2} \sigma_\epsilon^4 \sigma_z^2$

2.2 Généralisation dans un cadre multi-actifs

Dans cette partie, nous allons étendre la réflexion de Grossman et Stiglitz à un modèle multi-actifs dans un contexte d'information asymétrique.

2.2.1 Hypothèses

Les hypothèses du modèle sont les suivantes :

- Il y a une seule période. Les investisseurs constituent leur portefeuille en $t = 0$ et le consomment en $t = 1$;
- Le marché comprend n actifs risqués dont les prix futurs (en $t = 1$) sont notés $(\tilde{P}_i^1)_{i \in \{1..n\}}$ ainsi qu'un actif sans risque dont le revenu futur (en $t = 1$) est égal à $R = 1 + r_f$;
- Les revenus futurs des actifs risqués sont représentés par un modèle de la forme $\tilde{P}_i^1 = \tilde{\theta}_i + \tilde{\epsilon}_i$ où $\tilde{\theta}_i$ représente une variable aléatoire suivant une loi normale spécifique au titre i , et $\tilde{\epsilon}_i$ est un bruit correspondant à un terme d'erreur.
En notant \tilde{P}^1 le vecteur des revenus futurs des n actifs risqués, $\tilde{\theta}$ le vecteur aléatoire spécifique aux titres et $\tilde{\epsilon}$ le vecteur aléatoire spécifique au bruit, nous pouvons réécrire le prix futur sous forme vectorielle : $\tilde{P}^1 = \tilde{\theta} + \tilde{\epsilon}$;
- Les variables aléatoires $\tilde{\theta}$ et $\tilde{\epsilon}$ suivent des lois normales indépendantes respectivement d'espérances $\mathbb{E}[\tilde{\theta}]$ et 0, et de matrices de variance-covariance $Var(\tilde{\theta}) = T$ et $Var(\tilde{\epsilon}) = E$;
- Nous notons $\tilde{z} = (\tilde{z}_i)_{i \in \{1..n\}}$ le vecteur aléatoire représentant l'offre pour les n actifs risqués. \tilde{z} est indépendant de $\tilde{\theta}$ et $\tilde{\epsilon}$, et a pour espérance $\mathbb{E}[\tilde{z}]$ et variance $Var(\tilde{z}) = Z$;
- Nous considérons une proportion λ d'investisseurs informés, c'est-à-dire la proportion d'investisseurs connaissant la réalisation future θ de la variable aléatoire $\tilde{\theta}$;
- Tous les investisseurs possèdent la même aversion au risque a supposée constante.
- Chaque investisseur j possède une richesse finale (en $t = 1$) notée w_j^1 et évolue selon une fonction d'utilité exponentielle négative : $U(w_j^1) = -e^{-aw_j^1}$.
- Il existe une unique solution analytique pour \tilde{P}^0 dans la classe des fonctions linéaires en $\tilde{\theta}$ et \tilde{z} donnée par : $\tilde{P}^0(\tilde{\theta}, \tilde{z}) = A_0 + A_1 \left(\tilde{\theta} - \frac{a}{\lambda} E \tilde{z} \right) = A_0 + A_1 \tilde{w}(\tilde{\theta}, \tilde{z})$ où A_0 et A_1 sont des matrices avec A_1 régulière.

2.2.2 Calcul du prix d'équilibre \tilde{P}^0

L'objectif de cette partie est de déterminer l'expression finale du vecteur \tilde{P}^0 représentant le prix d'équilibre en $t = 0$ des n actifs risqués.

Remarquons tout d'abord que le prix d'équilibre \tilde{P}^0 est tel que l'offre (\tilde{z}) égalise la demande. Or, la demande de chaque investisseur est telle qu'elle maximise l'espérance d'utilité de sa richesse. Pour calculer cette espérance, nous avons besoin de considérer l'information connue par chaque investisseur. Ainsi, nous introduisons, pour chaque investisseur j , l'information I_j qui lui est associée. Cette information diffère selon le statut de l'investisseur considéré (informé ou non informé).

Considérons un investisseur j . Notons $X_j = \begin{pmatrix} x_1^j \\ \vdots \\ x_n^j \end{pmatrix}$ le vecteur représentant la composition

de son portefeuille en actifs risqués. En notant w_j^0 la richesse initiale de l'investisseur, nous avons l'égalité suivante :

$$w_j^1 = w_j^0 R + X_j'(\tilde{P}^1 - R\tilde{P}^0) \quad (1)$$

Remarquons que si la richesse finale w_j^1 suit une loi normale, les fonctions exponentielles négatives de paramètre a représentent les mêmes préférences qu'une fonction de Markowitz linéaire de paramètre $\frac{a}{2}$ ¹.

Ainsi, en supposant que la distribution de w_j^1 est normale conditionnellement à I_j et en utilisant une variante de la propriété précédente (en rajoutant un facteur $-a$), nous pouvons écrire l'espérance d'utilité de la richesse finale conditionnellement à l'information détenue par l'investisseur j :

$$\mathbb{E}[U(w_j^1|I_j)] = U \left[\mathbb{E}(w_j^1|I_j) - \frac{a}{2} \text{Var}(w_j^1|I_j) \right] \quad (2)$$

La maximisation de cette fonction d'utilité s'écrit alors :

$$\begin{aligned} \max_{X_j} \{ \mathbb{E}[U(w_j^1|I_j)] \} &= \max_{X_j} \left\{ U \left[\mathbb{E}[w_j^1|I_j] - \frac{a}{2} \text{Var}(w_j^1|I_j) \right] \right\} \\ &= \max_{X_j} \left\{ \mathbb{E}[w_j^1|I_j] - \frac{a}{2} \text{Var}(w_j^1|I_j) \right\} \\ &= \max_{X_j} \left\{ \mathbb{E}[w_j^0 R + X_j'(\tilde{P}^1 - R\tilde{P}^0)|I_j] - \frac{a}{2} \text{Var}(w_j^0 R + X_j'(\tilde{P}^1 - R\tilde{P}^0)|I_j) \right\} \\ &= \max_{X_j} \left\{ \mathbb{E}[X_j'(\tilde{P}^1 - R\tilde{P}^0)|I_j] - \frac{a}{2} \text{Var}(X_j'(\tilde{P}^1 - R\tilde{P}^0)|I_j) \right\} \end{aligned} \quad (3)$$

Or, nous pouvons réécrire les espérance et variance considérées ici telles que :

$$\begin{aligned} \mathbb{E}[X_j'(\tilde{P}^1 - R\tilde{P}^0)|I_j] &= X_j' \mathbb{E}[\tilde{P}^1 - R\tilde{P}^0|I_j] \\ \text{et } \text{Var}(X_j'(\tilde{P}^1 - R\tilde{P}^0)|I_j) &= X_j' \text{Var}(\tilde{P}^1 - R\tilde{P}^0|I_j) X_j \end{aligned}$$

Donc il vient :

$$\max_{X_j} \{ \mathbb{E}[U(w_j^1|I_j)] \} = \max_{X_j} \left\{ X_j' \mathbb{E}[\tilde{P}^1 - R\tilde{P}^0|I_j] - \frac{a}{2} X_j' \text{Var}(\tilde{P}^1 - R\tilde{P}^0|I_j) X_j \right\} \quad (4)$$

Et après dérivation de l'expression à maximiser, on trouve que la solution s'écrit :

$$X_j = \frac{1}{a} \text{Var}^{-1}(\tilde{P}^1 - R\tilde{P}^0|I_j) \mathbb{E}[\tilde{P}^1 - R\tilde{P}^0|I_j] \quad (5)$$

Intéressons-nous maintenant à l'information I_j . Si l'investisseur j est informé, alors il connaît la réalisation θ de la variable aléatoire $\tilde{\theta}$ et donc $I_{j,I} = \{\theta, P^0\}$, sinon $I_{j,NI} = \{P^0\}$ ².

On peut alors réécrire le vecteur X_j selon le statut de l'investisseur j :

1. Voir le raisonnement en [Annexe 1](#).

2. $I_{j,I}$ et $I_{j,NI}$ représentent respectivement l'information d'un investisseur j informé et l'information d'un investisseur j non informé.

— Pour un investisseur j informé :

$$\begin{aligned} X_{j,I} &= \frac{1}{a} Var^{-1}(\tilde{P}^1 - R\tilde{P}^0 | \theta, \tilde{P}^0 = P^0) \mathbb{E}[\tilde{P}^1 - R\tilde{P}^0 | \theta, \tilde{P}^0 = P^0] \\ &= \frac{1}{a} Var^{-1}(\tilde{\epsilon} | \theta, \tilde{P}^0 = P^0) (\theta - RP^0) \\ &= \frac{1}{a} E^{-1} (\theta - RP^0) \end{aligned} \quad (6)$$

— Pour un investisseur j non informé :

$$\begin{aligned} X_{j,NI} &= \frac{1}{a} Var^{-1}(\tilde{P}^1 - R\tilde{P}^0 | \tilde{P}^0 = P^0) \mathbb{E}[\tilde{P}^1 - R\tilde{P}^0 | \tilde{P}^0 = P^0] \\ &= \frac{1}{a} Var^{-1}(\tilde{\theta} + \tilde{\epsilon} | \tilde{P}^0 = P^0) \left(\mathbb{E}[\tilde{\theta} + \tilde{\epsilon} | \tilde{P}^0 = P^0] - RP^0 \right) \\ &= \frac{1}{a} (N + E)^{-1} \left(\mathbb{E}[\tilde{\theta} | \tilde{P}^0 = P^0] - RP^0 \right) \end{aligned} \quad (7)$$

où $N = Var(\tilde{\theta} | \tilde{P}^0 = P^0)$.

A l'équilibre, l'offre égalise la demande, ce que l'on traduit par :

$$\lambda X_{j,I} + (1 - \lambda) X_{j,NI} = \tilde{z} \quad (8)$$

En utilisant les expressions de $X_{j,I}$ et $X_{j,NI}$, il vient :

$$\Leftrightarrow \lambda E^{-1} (\theta - RP^0) + (1 - \lambda) (N + E)^{-1} \left(\mathbb{E}[\tilde{\theta} | \tilde{P}^0 = P^0] - RP^0 \right) = a\tilde{z} \quad (9)$$

Nous nous plaçons juste avant l'observation des investisseurs donc l'égalité est également vérifiée avec les variables aléatoires, soit :

$$\begin{aligned} &\lambda E^{-1} (\tilde{\theta} - R\tilde{P}^0) + (1 - \lambda) (N + E)^{-1} \left(\mathbb{E}[\tilde{\theta} | \tilde{P}^0 = P^0] - R\tilde{P}^0 \right) = a\tilde{z} \\ \Leftrightarrow &-R [\lambda E^{-1} + (1 - \lambda) (N + E)^{-1}] \tilde{P}^0 + \lambda E^{-1} \tilde{\theta} + (1 - \lambda) (N + E)^{-1} \mathbb{E}[\tilde{\theta} | \tilde{P}^0] - a\tilde{z} = 0 \end{aligned} \quad (10)$$

Le terme $\lambda E^{-1} + (1 - \lambda) (N + E)^{-1}$ représente ici une moyenne harmonique de la précision des informés et des non informés sur le marché.

Posons $V_m := [\lambda E^{-1} + (1 - \lambda) (N + E)^{-1}]^{-1}$, V_m représente alors ici un **risque**.

L'équation (10) se réécrit alors :

$$\tilde{P}^0 = \frac{1}{R} V_m \left[\lambda E^{-1} \tilde{\theta} + (1 - \lambda) (N + E)^{-1} \mathbb{E}[\tilde{\theta} | \tilde{P}^0] - a\tilde{z} \right] \quad (11)$$

$\tilde{\theta}$ et \tilde{P}^0 ayant une distribution normale conjointe, on peut réécrire :

$$\mathbb{E}[\tilde{\theta} | \tilde{P}^0] = \mathbb{E}[\tilde{\theta}] + Cov(\tilde{\theta}, \tilde{P}^0) Var^{-1}(\tilde{P}^0) (\tilde{P}^0 - \mathbb{E}[\tilde{P}^0]) \quad (12)$$

A partir de l'équation (11), déterminons l'expression des coefficients matriciels A_0 et A_1 tels que $\tilde{P}^0(\tilde{\theta}, \tilde{z}) = A_0 + A_1 \left(\tilde{\theta} - \frac{a}{\lambda} E\tilde{z} \right) = A_0 + A_1 \tilde{w}(\tilde{\theta}, \tilde{z})$.

Pour cela, commençons par remplacer \tilde{P}^0 par l'expression ci-dessus dans l'équation (12) :

$$\begin{aligned}
\mathbb{E}[\tilde{\theta}|\tilde{P}^0 = P^0] &= \mathbb{E}[\tilde{\theta}] + T A_1' \left[A_1 \left(T + \frac{a^2}{\lambda^2} E Z E \right) A_1' \right]^{-1} A_1 \left[\tilde{\theta} - \mathbb{E}[\tilde{\theta}] - \frac{a}{\lambda} E (\tilde{z} - \mathbb{E}[\tilde{z}]) \right] \\
&= \mathbb{E}[\tilde{\theta}] + T A_1' A_1'^{-1} \left(T + \frac{a^2}{\lambda^2} E Z E \right)^{-1} A_1^{-1} A_1 \left[\tilde{\theta} - \mathbb{E}[\tilde{\theta}] - \frac{a}{\lambda} E (\tilde{z} - \mathbb{E}[\tilde{z}]) \right] \\
&= \mathbb{E}[\tilde{\theta}] + T \underbrace{(A_1^{-1} A_1)'}_{I_n} \left(T + \frac{a^2}{\lambda^2} E Z E \right)^{-1} \left[\tilde{\theta} - \mathbb{E}[\tilde{\theta}] - \frac{a}{\lambda} E (\tilde{z} - \mathbb{E}[\tilde{z}]) \right] \\
&= \mathbb{E}[\tilde{\theta}] + T \left(T + \frac{a^2}{\lambda^2} E Z E \right)^{-1} \left[\frac{a}{\lambda} E \mathbb{E}[\tilde{z}] - \mathbb{E}[\tilde{\theta}] \right] \\
&\quad + \left[T \left(T + \frac{a^2}{\lambda^2} E Z E \right)^{-1} \right] \tilde{\theta} \\
&\quad - \frac{a}{\lambda} E \left[T \left(T + \frac{a^2}{\lambda^2} E Z E \right)^{-1} \right] \tilde{z}
\end{aligned} \tag{13}$$

En remarquant que $N := \text{Var}(\tilde{\theta}|\tilde{P}^0) = T - T (\text{Var}^{-1}(\tilde{w})) T$ avec $\text{Var}(\tilde{w}) = T + \frac{a^2}{\lambda^2} E Z E$, on peut réécrire :

$$\begin{aligned}
\mathbb{E}[\tilde{\theta}|\tilde{P}^0] &= \mathbb{E}[\tilde{\theta}] + (T - N) T^{-1} \left[\frac{a}{\lambda} E \mathbb{E}[\tilde{z}] - \mathbb{E}[\tilde{\theta}] \right] \\
&\quad + [(T - N) T^{-1}] \tilde{\theta} \\
&\quad - \frac{a}{\lambda} E [(T - N) T^{-1}] \tilde{z}
\end{aligned} \tag{14}$$

Injectons cette expression dans l'équation (11) :

$$\begin{aligned}
\tilde{P}^0 &= \frac{1}{R} V_m \left[\lambda E^{-1} \tilde{\theta} + (1 - \lambda) (N + E)^{-1} \mathbb{E}[\tilde{\theta}|\tilde{P}^0] - a \tilde{z} \right] \\
&= \frac{1}{R} V_m \left[(1 - \lambda) (N + E)^{-1} \left(\mathbb{E}[\tilde{\theta}] + (T - N) T^{-1} \left(\frac{a}{\lambda} E \mathbb{E}[\tilde{z}] - \mathbb{E}[\tilde{\theta}] \right) \right) \right] \\
&\quad + \frac{1}{R} V_m \left[\lambda E^{-1} + (1 - \lambda) (N + E)^{-1} (T - N) T^{-1} \right] \tilde{\theta} \\
&\quad + \frac{1}{R} V_m \left[-\frac{a}{\lambda} E (1 - \lambda) (N + E)^{-1} (T - N) T^{-1} - a \right] \tilde{z} \\
&= \frac{1}{R} V_m \left[(1 - \lambda) (N + E)^{-1} \left(\mathbb{E}[\tilde{\theta}] + (T - N) T^{-1} \left(\frac{a}{\lambda} E \mathbb{E}[\tilde{z}] - \mathbb{E}[\tilde{\theta}] \right) \right) \right] \\
&\quad + \frac{1}{R} V_m \left[\lambda E^{-1} + (1 - \lambda) (N + E)^{-1} (T - N) T^{-1} \right] \left(\tilde{\theta} - \frac{a}{\lambda} E \tilde{z} \right)
\end{aligned} \tag{15}$$

Ainsi, par identification des uniques constantes A_0 et A_1 , il vient :

$$\begin{aligned}
& \begin{cases} A_0 = \frac{1}{R} V_m \left[(1 - \lambda)(N + E)^{-1} \left(\mathbb{E}[\tilde{\theta}] + (T - N)T^{-1} \left(\frac{a}{\lambda} E \mathbb{E}[\tilde{z}] - \mathbb{E}[\tilde{\theta}] \right) \right) \right] \\ A_1 = \frac{1}{R} V_m [\lambda E^{-1} + (1 - \lambda)(N + E)^{-1}(T - N)T^{-1}] \end{cases} \\
& \Leftrightarrow \begin{cases} A_0 = \frac{1}{R} V_m \left[(1 - \lambda)(N + E)^{-1} \left((I_n - (T - N)T^{-1}) \mathbb{E}[\tilde{\theta}] + \frac{a}{\lambda} (T - N)T^{-1} E \mathbb{E}[\tilde{z}] \right) \right] \\ A_1 = \frac{1}{R} V_m [\lambda E^{-1}T + (1 - \lambda)(N + E)^{-1}(T - N)] T^{-1} \end{cases} \\
& \Leftrightarrow \begin{cases} A_0 = \frac{1}{R} V_m \left[\left((1 - \lambda)(N + E)^{-1} - (1 - \lambda)(N + E)^{-1}(T - N)T^{-1} \right) \mathbb{E}[\tilde{\theta}] \right. \\ \quad \left. + \left[\frac{a}{\lambda} \underbrace{(1 - \lambda)(N + E)^{-1}(T - N)T^{-1} E}_{RV_m^{-1}A_1 - \lambda E^{-1}} \right] \mathbb{E}[\tilde{z}] \right] \\ A_1 = \frac{1}{R} V_m \left[\underbrace{(\lambda E^{-1} + (1 - \lambda)(N + E)^{-1})}_{V_m^{-1}} T - (1 - \lambda)(N + E)^{-1}N \right] T^{-1} \end{cases} \\
& \Leftrightarrow \begin{cases} A_0 = \frac{1}{R} V_m \left(\left[(1 - \lambda)(N + E)^{-1} - \underbrace{(1 - \lambda)(N + E)^{-1}(T - N)T^{-1} - \lambda E^{-1}}_{-RV_m^{-1}A_1} + \lambda E^{-1} \right] \mathbb{E}[\tilde{\theta}] \right. \\ \quad \left. + \left[\frac{a}{\lambda} RV_m^{-1}A_1 - a \right] \mathbb{E}[\tilde{z}] \right) \\ A_1 = \frac{1}{R} V_m \left[V_m^{-1}T - \underbrace{(1 - \lambda)(N + E)^{-1}N - \lambda E^{-1}N}_{-V_m^{-1}N} + \lambda E^{-1}N \right] T^{-1} \end{cases} \quad (16)
\end{aligned}$$

En remarquant que $V_m = (N + E)(I_n + \lambda E^{-1}N)^{-1}$, on a $\lambda E^{-1}N = V_m^{-1}(N + E) - I_n$, et donc :

$$\begin{aligned}
& \begin{cases} A_0 = \frac{1}{R} V_m \left(\left[-RV_m^{-1}A_1 + \underbrace{\lambda E^{-1} + (1 - \lambda)(N + E)^{-1}}_{V_m^{-1}} \right] \mathbb{E}[\tilde{\theta}] \right. \\ \quad \left. + \left[\frac{a}{\lambda} RV_m^{-1}A_1 - a \right] \mathbb{E}[\tilde{z}] \right) \\ A_1 = \frac{1}{R} V_m [V_m^{-1}T - V_m^{-1}N + V_m^{-1}(N + E) - I_n] T^{-1} \end{cases} \\
& \Leftrightarrow \begin{cases} A_0 = \frac{1}{R} \left[(I_n - RA_1) \mathbb{E}[\tilde{\theta}] + \left(\frac{Ra}{\lambda} A_1 E - a V_m \right) \mathbb{E}[\tilde{z}] \right] \\ A_1 = \frac{1}{R} [T + E - V_m] T^{-1} \end{cases} \quad (17)
\end{aligned}$$

Nous remarquons que, par combinaison linéaire, A_1 est bien une matrice régulière donc cette hypothèse est vérifiée.

Par ailleurs, l'équation du prix d'équilibre prend toujours la forme d'une espérance de revenu futur à laquelle on soustrait une prime de risque. Cependant, en asymétrie informationnelle, nous pouvons faire face à de nouvelles situations où par exemple le prix d'équilibre diminue avec une anticipation de revenu futur à la hausse.

2.3 Étude de $Cov(\tilde{P}^1 - \tilde{P}^0; \tilde{P}^0)$

En considérant que \tilde{P}^0 peut s'écrire comme fonction linéaire en $\tilde{\theta}$ et \tilde{z} telle que :

$$\tilde{P}^0(\tilde{\theta}, \tilde{z}) = A_0 + A_1 \left(\tilde{\theta} - \frac{a}{\lambda} E \tilde{z} \right) = A_0 + A_1 \tilde{w}(\tilde{\theta}, \tilde{z})$$

où A_0 et A_1 sont des matrices avec A_1 régulière et :

$$\begin{aligned} A_1 &= \frac{1}{R}(T + E - V_m)T^{-1} \\ A_0 &= \frac{1}{R} \left[(I_n - RA_1)\mathbb{E}[\tilde{\theta}] + \left(\frac{Ra}{\lambda} A_1 E - aV_m \right) \mathbb{E}[\tilde{z}] \right] \end{aligned}$$

nous avons alors :

$$\begin{aligned} \tilde{P}^1 - \tilde{P}^0 &= \tilde{\theta} + \tilde{\epsilon} - \left[A_0 + A_1 \left(\tilde{\theta} - \frac{a}{\lambda} E \tilde{z} \right) \right] \\ &= -A_0 + (I_n - A_1)\tilde{\theta} + \frac{a}{\lambda} A_1 E \tilde{z} + \tilde{\epsilon} \end{aligned} \quad (18)$$

et donc :

$$\begin{aligned} Cov(\tilde{P}^1 - \tilde{P}^0, \tilde{P}^0) &= (I_n - A_1)T(A_1)' + \left(\frac{a}{\lambda} A_1 E \right) Z \left(\frac{a}{\lambda} A_1 E \right)' \\ &= (I_n - A_1)T A_1' + \frac{a^2}{\lambda^2} (A_1 E) Z (A_1 E)' \end{aligned} \quad (19)$$

Démontrons maintenant que $Cov(\tilde{P}^1 - \tilde{P}^0, \tilde{P}^0)$ est une matrice définie négative et qu'elle décroît avec le degré d'asymétrie informationnelle. Pour cela, écrivons tout d'abord $Cov(\tilde{P}^1 - \tilde{P}^0, \tilde{P}^0)$ en fonction de $N = Var(\tilde{\theta}|P^0)$.

2.3.1 $Cov(\tilde{P}^1 - \tilde{P}^0, \tilde{P}^0)$ fonction de N

Dans cette partie, nous prenons pour les calculs $R = 1$.

À partir de l'équation (19) et en remplaçant A_1 par son expression, il vient :

$$\begin{aligned} Cov(\tilde{P}^1 - \tilde{P}^0, \tilde{P}^0) &= T A_1' - A_1 T A_1' + \frac{a^2}{\lambda^2} A_1 E Z E' A_1' \\ &= T A_1' - A_1 \left(T + \frac{a^2}{\lambda^2} E Z E' \right) A_1' \\ &= \left[T - A_1 \underbrace{\left(T + \frac{a^2}{\lambda^2} E Z E' \right)}_{Var(\tilde{w})} \right] A_1' \\ &= [T - (T + E - V_m)T^{-1}Var(\tilde{w})] T^{-1}(T + E - V_m) \\ &= (T + E - V_m) - (T + E - V_m) \underbrace{[T(Var^{-1}(\tilde{w}))T]^{-1}}_{(T-N)^{-1}} (T + E - V_m) \\ &= (T + E - V_m) - (T + E - V_m)(T - N)^{-1}(T + E - V_m) \\ &= (T + E - V_m) [I_n - (T - N)^{-1}(T + E - V_m)] \end{aligned} \quad (20)$$

En remarquant que V_m peut également s'écrire :

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

l'équation (20) se réécrit :

$$\begin{aligned} Cov(\tilde{P}^1 - \tilde{P}^0, \tilde{P}^0) &= (T + \underbrace{E + N}_{V_m(I_n + \lambda E^{-1}N)} - N - V_m) \left[I_n - (T - N)^{-1}(T + \underbrace{E + N}_{V_m(I_n + \lambda E^{-1}N)} - N - V_m) \right] \\ &= (T - N + \lambda V_m E^{-1}N) [I_n - (T - N)^{-1}(T - N + \lambda V_m E^{-1}N)] \\ &= (T - N + \lambda V_m E^{-1}N) [-\lambda (T - N)^{-1} V_m E^{-1}N] \\ &= -\lambda V_m E^{-1}N - \lambda^2 V_m E^{-1}N (T - N)^{-1} V_m E^{-1}N \end{aligned}$$

En posant $f(N) := V_m E^{-1}N$, nous obtenons que :

$$Cov(\tilde{P}^1 - \tilde{P}^0, \tilde{P}^0) = -\lambda f(N) - \lambda^2 f(N)(T - N)^{-1}f(N) \quad (21)$$

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

λ est une proportion donc est un réel positif, $(T - N)^{-1} = [T(Var^{-1}(\tilde{w}))T]^{-1}$ est une matrice définie positive, et par combinaison linéaire on remarque que f est une fonction positive. Ainsi, $Cov(\tilde{P}^1 - \tilde{P}^0, \tilde{P}^0)$ est bien une **matrice définie négative**.

Étudions maintenant sa monotonie selon la variable N .

2.3.2 $Cov(\tilde{P}^1 - \tilde{P}^0, \tilde{P}^0)$ selon le degré d'asymétrie informationnelle

Montrons dans cette partie qu'une augmentation de la valeur de $N = Var(\tilde{\theta}|P^0)$ conduit à une diminution de la matrice $Cov(\tilde{P}^1 - \tilde{P}^0, \tilde{P}^0)$.

Financièrement, une augmentation de la variance N correspond à une plus grande incertitude pour les investisseurs non informés, donc le résultat escompté semble cohérent.

Montrons que la fonction f est croissante. Pour cela, considérons N_1 et N_2 deux matrices symétriques définies positives telles que $N_1 \leq N_2$ et montrons que $f(N_1) \leq f(N_2)$.

En passant à l'inverse, il vient :

$$N_1^{-1} \geq N_2^{-1}$$

Avec E symétrique définie positive, il vient³ :

$$N_1^{-1}EN_1^{-1} \geq N_2^{-1}EN_2^{-1}$$

Par ailleurs, avec $\lambda \geq 0$, on a :

$$\lambda N_1^{-1} \geq \lambda N_2^{-1}$$

Par combinaison linéaire, on en déduit que :

$$(N_1^{-1}EN_1^{-1} + \lambda N_1^{-1})^{-1} \leq (N_2^{-1}EN_2^{-1} + \lambda N_2^{-1})^{-1} \quad (22)$$

Et que :

$$(N_1^{-1} + \lambda E^{-1}) \geq (N_2^{-1} + \lambda E^{-1})$$

3. Voir la propriété (p7) de la boîte à outils.

Donc que :

$$(N_1^{-1} + \lambda E^{-1})^{-1} \leq (N_2^{-1} + \lambda E^{-1})^{-1} \quad (23)$$

En combinant les équations (22) et (23), nous avons bien $f(N_1) \leq f(N_2)$, ce qui prouve que la fonction f est bien croissante en N .

Enfin, par combinaison linéaire (de facteurs $-\lambda$ et $-\lambda^2$), nous concluons bien que la matrice $Cov(\tilde{P}^1 - \tilde{P}^0, \tilde{P}^0)$ **décroît avec le degré d'asymétrie informationnelle**.

2.3.3 Application numérique

Considérons un marché constitué de deux actifs risqués et de l'actif sans risque et ayant les caractéristiques suivantes :

$$a = 0,01 \quad ; \quad R = 1 \quad ; \quad \lambda = 30,00\% \quad ; \quad E = \begin{pmatrix} 120 & 30 \\ 30 & 8 \end{pmatrix} \quad ; \quad Z = \begin{pmatrix} 50 & 0 \\ 0 & 5 \end{pmatrix}$$

Notre objectif est de montrer graphiquement qu'une augmentation de la matrice N entraîne une diminution de $Cov(\tilde{P}^1 - \tilde{P}^0, \tilde{P}^0)$.

Nous définissons ici la mesure d'une matrice par la [norme de Frobenius](#) qui est telle que, pour A une matrice à coefficients réels :

$$\|A\|_F := \sqrt{\text{tr}(AA')} = \sqrt{\text{tr}(A'A)} \quad (24)$$

La variance $N = \text{Var}(\tilde{\theta}|P^0)$ étant une matrice endogène⁴, pour la mesurer nous allons faire varier la matrice $T = \text{Var}(\tilde{\theta})$ qui possède la même monotonie. Ainsi, nous définissons :

$$T = \begin{pmatrix} 90 + x & 10 \\ 10 & 30 + x \end{pmatrix}, \quad \text{avec } x \in [0, 150]$$

En rappelant que :

$$N = T - T \left[T + \frac{a^2}{\lambda^2} EZE \right] T$$

nous pouvons construire le graphique suivant :

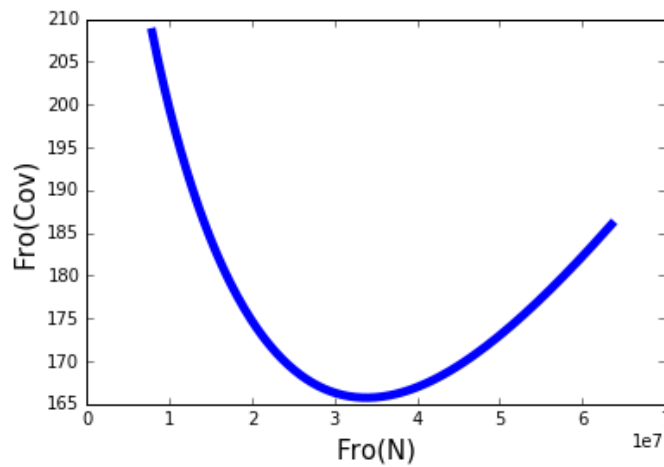


FIGURE 1 – Évolution de la norme de Frobenius de $Cov(\tilde{P}^1 - \tilde{P}^0, \tilde{P}^0)$ en fonction de celle de N .

4. Une variable endogène est une variable dépendant directement d'autres variables. Ici, N dépend de T , E et Z .

Nous remarquons tout d'abord qu'une augmentation de la variance N entraîne une diminution de $Cov(\tilde{P}^1 - \tilde{P}^0, \tilde{P}^0)$, puis se produit le phénomène contraire. On pourrait alors penser que ceci est dû à une divergence dans la mesure de Frobenius.

Pour vérifier cette conjecture, procédons à un changement de mesure et construisons alors notre graphe en fonction du terme d'indices $(0, 0)$ des matrices N et $Cov(\tilde{P}^1 - \tilde{P}^0, \tilde{P}^0)$. Avec les mêmes autres paramètres, nous avons le graphique suivant :

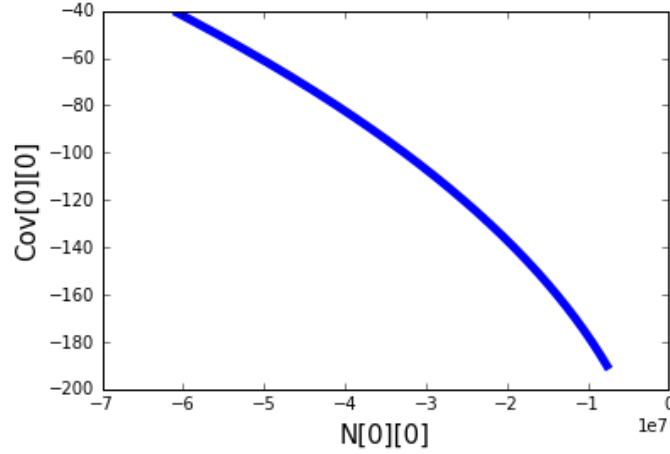


FIGURE 2 – Évolution du terme d'indices $(0, 0)$ de $Cov(\tilde{P}^1 - \tilde{P}^0, \tilde{P}^0)$ en fonction de celui de N .

Ainsi, nous observons qu'une augmentation de la variance N entraîne bien une diminution de $Cov(\tilde{P}^1 - \tilde{P}^0, \tilde{P}^0)$ et nous n'observons plus l'inversion de tendance de la courbe.

Le code Python utilisé pour tracer ces différents graphiques est disponible en [Annexe 2](#). Il est alors possible de modifier les valeurs des différents paramètres pour observer l'évolution de la courbe dès lors que l'on modifie telle ou telle valeur.

Nous laissons la possibilité au lecteur de considérer un troisième actif risqué (voire plus encore) pour étudier les modifications que cela peut entraîner sur la tendance de la courbe.

3 Aspects empiriques

En suivant la méthodologie proposée par Easley et al. (2002), nous allons mener une analyse⁵ montrant que la mesure AIM est positivement et significativement liée aux rentabilités des titres. Ainsi, nous pourrions justifier l'utilisation de cette mesure d'asymétrie informationnelle.

Nous pourrions ensuite proposer des stratégies actives de gestion de portefeuilles basées sur l'exploitation du différentiel informationnel entre investisseurs.

3.1 Corrélations entre différentes mesures d'asymétrie informationnelle

Afin de s'assurer de la pertinence des données étudiées et de leur cohérence, nous avons calculé les coefficients de corrélation entre plusieurs variables servant à expliquer la rentabilité des actifs. Les corrélations entre deux variables ne sont calculées que lorsque des données sont présentes pour ces deux variables.

5. Régression en coupe transversale des rentabilités des titres sur le bêta du marché, la mesure AIM ainsi que d'autres mesures utilisées usuellement comme mesures d'asymétrie informationnelle.

	return	beta	aim	log_market_cap	log_btm	pin	fsrv	turnover	illiq_amihud
return	1.0	-0.0236	0.02433	0.00386	0.03145	-0.00907	-0.01737	-0.04316	-0.00115
beta	-0.0236	1.0	-0.13555	0.03292	-0.06215	0.09105	-0.13098	0.16161	-0.00994
aim	0.02433	-0.13555	1.0	-0.32608	0.17278	0.09325	0.29502	-0.13092	0.04781
log_market_cap	0.00386	0.03292	-0.32608	1.0	-0.27852	-0.5998	-0.57324	0.0604	-0.10899
log_btm	0.03145	-0.06215	0.17278	-0.27852	1.0	0.20234	0.13158	-0.10723	0.04654
pin	-0.00907	0.09105	0.09325	-0.5998	0.20234	1.0	0.26401	-0.06188	0.11358
fsrv	-0.01737	-0.13098	0.29502	-0.57324	0.13158	0.26401	1.0	-0.07661	0.04793
turnover	-0.04316	0.16161	-0.13092	0.0604	-0.10723	-0.06188	-0.07661	1.0	-0.01137
illiq_amihud	-0.00115	-0.00994	0.04781	-0.10899	0.04654	0.11358	0.04793	-0.01137	1.0

TABLE 1 – Coefficients de corrélation des différentes variables étudiées

	return	beta	aim	log_market_cap	log_btm	pin	fsrv	turnover	illiq_amihud
return	0.0	0.0	0.0	0.14929	0.0	0.04969	0.0	0.0	0.67395
beta	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.8E-4
aim	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
log_market_cap	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
log_btm	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
pin	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
fsrv	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
turnover	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.0E-5
illiq_amihud	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

TABLE 2 – P-valeurs associées aux différentes variables étudiées

Les p-valeurs inférieures à 10^{-6} sont ici marquées comme nulles.

Au vu des p-valeurs calculées, seule la mesure d'illiquidité n'est pas liée à la rentabilité des actifs. Les autres mesures sont toutes corrélées de manière significative.

On peut remarquer ici que la mesure AIM est positivement corrélée avec la rentabilité d'un titre. Nous nous intéressons aux corrélations entre AIM et les mesures naïves d'asymétrie :

- *log_market_cap* : Les deux mesures sont négativement corrélées. Cela semble cohérent, car plus une entreprise a d'impact sur le marché, plus il est cohérent que l'information sur celle-ci soit connue.
- *log_btm* : Les deux mesures sont positivement corrélées. Un BTM fort indique que les investisseurs ne sont pas prêts à acheter le titre à une grande valeur. Cela semble cohérent avec un manque d'information, donc cette corrélation positive est cohérente.
- *pin* : Les deux mesures sont positivement corrélées. Les deux mesurant l'asymétrie informationnelle, il est très cohérent d'obtenir ce résultat.
- *illiq_amihud* : Les deux mesures sont positivement corrélées. Un titre peu liquide est peu échangé, donc il est cohérent que l'information manque sur ces titres.

3.2 Analyse de la mesure AIM par régression linéaire

3.2.1 Explications

Nous avons effectué plusieurs régressions linéaires du rendement d'un actif en fonction des autres paramètres du modèle à la date précédente. Nous avons dans un premier temps choisi plusieurs n -uplets de paramètres sur lesquels régresser (par exemple AIM et le bêta du marché). À chaque date, nous avons effectué une régression sur ces paramètres sur l'ensemble des actifs.

Nous avons enfin calculé la moyenne des coefficients de régression à chaque date pour avoir une valeur de coefficient de régression par paramètre sur l'ensemble des données analysées, et avons calculé la p-valeur du test de différence de cette moyenne à 0 en modélisant la loi suivie par cette moyenne par une loi de Student de nombre de degrés de liberté égal à (*nombre de dates* - 1).

Nous avons calculé le coefficient moyen et sa p-valeur associée de deux manières : dans un premier temps en calculant la moyenne de façon équipondérée, puis en pondérant les coefficients par $\frac{1}{\sigma\gamma}$, de sorte à augmenter l'impact des γ de faible variance, donc plus représentatifs.

Il nous semble cependant pertinent de noter que cette méthode de calcul, notamment de la p-valeur, est critiquable dans le sens où la seule mesure de significativité des coefficients eux-mêmes est la pondération par $\frac{1}{\sigma\gamma}$. Le test de p-valeur mathématiquement correct pour la moyenne des coefficients demanderait de déterminer la loi de cette moyenne, c'est-à-dire si on suit les conventions mathématiques usuelles, la loi de la moyenne des lois de Student de nombre de degrés de liberté égal au nombre d'observations moins le nombre de paramètres de régression moins un, pour cette date, et uniquement une fois ceci fait pourrions nous obtenir la probabilité que ce coefficient soit supérieur ou égal à celui observé en supposant que sa vraie espérance soit nulle.

L'inversion de la fonction de répartition de cette loi étant extrêmement complexe, nous avons utilisé la méthode consistant à simplement moyenner les coefficients obtenus, mais il nous semblait important de préciser que ceci est une approximation.

Dans la suite de ce rapport, on se référera aux résultats obtenus sans et avec pondération par respectivement (1) et (2). Par ailleurs, on utilisera comme mesure de la taille le logarithme de la capitalisation sur le marché, et comme mesure du BTM la valeur *logBtm* fournie.

3.2.2 Régression sur bêta

Avec *intercept* :

	Moyenne (1)	Moyenne (2)	P-valeur (1)	P-valeur (2)
Intercept	0,003587	0,003807	0,031682	0,016665
Bêta	-0,004406	-0,003920	0,012374	0,015137

Sans *intercept* :

	Moyenne (1)	Moyenne (2)	P-valeur (1)	P-valeur (2)
Bêta	-0,001194	-0,000281	0,544945	0,876555

Dans un premier temps nous régressons sur bêta pour tester la validité de nos calculs, et nous trouvons à notre grande surprise que le bêta est négatif, et de surcroît, significativement négatif, avec une p-valeur sur la moyenne pondérée $\simeq 1,5\%$.

Ceci pourrait s'expliquer par un échantillon trop faible si la significativité de la mesure n'était pas aussi grande, mais, cette dernière étant élevée, on ne peut pas justifier cette mesure ainsi.

L'intercept signifie qu'il semble exister une prime systématique supérieure à R_f , en revanche, la négativité du bêta est incompréhensible. Cependant, notons que nous ne faisons pas la régression sur $\beta_{t-1}(R_{m_t} - R_{f_t})$, mais seulement sur β_{t-1} . Ceci est un problème important car le CAPM modélise le marché par $R_x = R_f + \beta_x * (R_m - R_f)$. Ainsi, le coefficient mesuré en β à chaque régression

devrait théoriquement être $R_m - R_f$, et le coefficient moyen devrait quant à lui être la moyenne de ces coefficients ; le bon sens financier voulant donc que ce coefficient soit supérieur à 0. La seule explication rationnelle est que nous étions sur une période où $R_m - R_f < 0$, mais ceci est financièrement absurde.

Nous pourrions probablement obtenir des résultats plus sensés en faisant la régression sur $\beta_{t-1}(R_{m_t} - R_{f_t})$, cependant nous rencontrons des difficultés à mettre en place ces calculs notamment dû aux valeurs manquantes, mais notons qu'il s'agit d'une piste d'amélioration intéressante.

Devant la significativité de l'intercept qui pourtant devrait être nul selon les divers modèles étudiés (en supposant que les paramètres supplémentaires soient centrés en 0), il nous a semblé pertinent de ne pas forcer l'intercept à 0 dans les régressions suivantes.

3.2.3 Régression sur bêta et AIM

	Moyenne (1)	Moyenne (2)	P-valeur (1)	P-valeur (2)
Intercept	0,007036	0,006760	0,002657	0,002090
Bêta	-0,004192	-0,003786	0,012714	0,013876
AIM	0,001398	0,001227	0,038107	0,050016

On constate ici que AIM est significativement différent de 0, de façon positive, ce qui signifie qu'il existe bien une prime de risque – selon nos théories, informationnelle – pour laquelle les investisseurs demandent rémunération. Ici, si on oublie le bêta, tout va bien.

3.2.4 Régression sur bêta, AIM et PIN

	Moyenne (1)	Moyenne (2)	P-valeur (1)	P-valeur (2)
Intercept	0,008968	0,008508	0,095312	0,079508
Bêta	-0,005678	-0,004404	0,059534	0,094567
AIM	-0,000186	-0,000158	0,831359	0,843531
PIN	-0,015861	-0,007566	0,347641	0,609659

Quant à ici, tout va mal, l'ajout de PIN et de AIM a rendu le coefficient lié à AIM non significativement différent de 0, et PIN n'est pas significatif pour autant. On remarque également que la significativité de l'intercept et de bêta a également baissé. Ces mesures sont très probablement dues au nombre réduit de données d'études suite au filtrage de ces dernières pour disposer de PIN.

3.2.5 Régression sur bêta et AIM quand on connaît PIN

Pour vérifier cette supposition, nous faisons une régression sur AIM seulement sur la période durant laquelle nous disposons de la valeur de PIN.

	Moyenne (1)	Moyenne (2)	P-valeur (1)	P-valeur (2)
Intercept	0,006204	0,007433	0,090280	0,029467
Bêta	-0,005724	-0,004550	0,066418	0,093861
AIM	-0,000357	-0,000235	0,697824	0,779084

Et nous trouvons que la p-valeur associée au coefficient d'AIM est maintenant $\simeq 77,9\%$ quand cette dernière était $\simeq 5\%$ avec toutes les valeurs, ce qui nous conforte dans notre croyance qu'inclure PIN ruine la significativité par le manque de données qui découle de l'utilisation de PIN.

Nous allons maintenant observer quels paramètres affectent fortement négativement la significativité d'AIM lorsqu'ils sont inclus dans le modèle.

3.2.6 Régression sur bêta, AIM et la taille

	Moyenne (1)	Moyenne (2)	P-valeur (1)	P-valeur (2)
Intercept	-0,003838	-0,002203	0,495931	0,669492
Bêta	-0,004074	-0,003724	0,013455	0,013974
AIM	0,002793	0,002299	0,000170	0,000822
Taille (<i>log_market_cap</i>)	0,001111	0,000969	0,025685	0,037726

La significativité d'AIM ne baisse pas – au contraire, elle augmente quand on inclut une mesure de la taille dans le modèle, ce qui signifie qu'il existe une prime de risque différente de celle liée à la taille, et donc que le risque informationnel liée à AIM est différent de celui lié à la taille.

L'augmentation signifie même qu'elles sont complémentaires.

3.2.7 Régression sur bêta, AIM et BTM

	Moyenne (1)	Moyenne (2)	P-valeur (1)	P-valeur (2)
Intercept	0,004737	0,004614	0,040709	0,034604
Bêta	-0,003933	-0,003388	0,017903	0,025979
AIM	0,000477	0,000465	0,479369	0,455302
<i>log</i> (BTM)	0,001333	0,000948	0,074803	0,146720

L'inclusion du *Book To Market* dans le modèle, en revanche, affecte très négativement la significativité d'AIM, qui a désormais une p-valeur de seulement 45,53% (partant de 5%). Le BTM n'affecte pas non plus significativement le modèle. Comme le PIN précédemment, il s'avère que ces résultats sont probablement dûs à un manque de données.

Voici les résultats de la régression sur bêta et AIM en ne gardant que les données où le *log_btm_t1* est disponible :

	Moyenne (1)	Moyenne (2)	P-valeur (1)	P-valeur (2)
Intercept	0,005262	0,005083	0,026416	0,023131
Bêta	-0,004088	-0,003562	0,014684	0,020249
AIM	0,000735	0,000659	0,290231	0,307225

Cela suffit à rendre AIM peu significatif (p-valeur de 30,7%).

3.2.8 Régression sur bêta, AIM et FSRV

	Moyenne (1)	Moyenne (2)	P-valeur (1)	P-valeur (2)
Intercept	0,018539	0,021860	0,010001	0,000648
Bêta	-0,003508	-0,003320	0,029091	0,023205
AIM	0,002780	0,002341	0,000064	0,000242
FSRV	-0,009272	-0,012683	0,194862	0,040854

Similairement à la taille, l'inclusion de FSRV fait significativement augmenter la significativité d'AIM, et FSRV est de même significatif.

Les mêmes conclusions peuvent donc en être tirées – il ne s'agit pas du même risque informationnel et AIM apporte ici une information qui lui est propre.

3.2.9 Régression sur bêta, AIM et illiq_amihud

	Moyenne (1)	Moyenne (2)	P-valeur (1)	P-valeur (2)
Intercept	0,006414	0,006012	0,008735	0,009634
Bêta	-0,004477	-0,003875	0,012409	0,017355
AIM	0,001045	0,000776	0,118948	0,211414
Illiq_amihud	0,000004	-0,000084	0,989930	0,408806

L'*Amihud Illiquidity Measure* tombe dans la même situation, et par conséquent, les mêmes conclusions que PIN et BTM, comme démontré par la régression suivante sur bêta et AIM quand on connaît les données de l'*Amihud Illiquidity Measure* :

	Moyenne (1)	Moyenne (2)	P-valeur (1)	P-valeur (2)
Intercept	0,006399	0,005938	0,008732	0,009636
Bêta	-0,004606	-0,004027	0,010203	0,013285
AIM	0,001053	0,000781	0,121721	0,213782

Notons la p-valeur qui est passée de 5% à 21,38%.

On ne peut pas conclure sur PIN, BTM, et Illiq_amihud par manque de données.

3.2.10 Régression sur bêta, AIM et turnover

	Moyenne (1)	Moyenne (2)	P-valeur (1)	P-valeur (2)
Intercept	0,006404	0,006188	0,004343	0,003273
Bêta	-0,002754	-0,002412	0,069406	0,080258
AIM	0,000697	0,000612	0,257370	0,284549
Turnover	-0,044943	-0,039061	0,010484	0,000467

Après courtes recherches, le *Stock Turnover* est considéré par certains comme une mesure d'asymétrie informationnelle, entre autres, bien que négativement corrélé dans nos résultats avec les autres mesures.

Nous avons inclus notre régression quant au *Stock Turnover* car cette dernière a des résultats intéressants : ici la significativité d'AIM a baissé, cependant le turnover est très significatif, ce qui serait un état difficile à obtenir si la quantité de données était fortement diminuée par l'inclusion du turnover. Et pour cause, en limitant les données à celles où on connaît le turnover :

	Moyenne (1)	Moyenne (2)	P-valeur (1)	P-valeur (2)
Intercept	0,007036	0,006760	0,002657	0,002090
Bêta	-0,004192	-0,003786	0,012714	0,013876
AIM	0,001398	0,001227	0,038107	0,050016

La p-valeur de l'AIM reste à 5%, ce qui signifie que le turnover rend moins probable l'existence d'une prime de risque spécifiquement associée à AIM ; signifiant qu'AIM pourrait être une proxy bruitée du *Stock Turnover*.

3.2.11 Régression sur bêta, AIM, taille et FSRV

	Moyenne (1)	Moyenne (2)	P-valeur (1)	P-valeur (2)
Intercept	-0,005186	-0,003343	0,676506	0,757175
Bêta	-0,004021	-0,003592	0,007828	0,009215
AIM	0,003573	0,002976	0,000001	0,000012
Taille (<i>log_market_cap</i>)	0,001064	0,001039	0,057039	0,045418
FSRV	0,004292	0,000163	0,593788	0,980954

Pour cette régression, AIM est extrêmement significatif, et positif comme à chaque fois (p-valeur de 0,0012%!). En revanche, FSRV n'est plus considéré comme significatif, alors qu'il l'était quand la taille n'était pas incluse dans le modèle. FSRV n'est donc pas complémentaire à la taille.

3.2.12 Régression sur bêta, AIM, taille, FSRV et turnover

	Moyenne (1)	Moyenne (2)	P-valeur (1)	P-valeur (2)
Intercept	-0,000137	0,000257	0,990723	0,980030
Bêta	-0,002632	-0,002320	0,065138	0,073588
AIM	0,002914	0,002422	0,000025	0,000158
Taille (<i>log_market_cap</i>)	0,000931	0,000929	0,089310	0,067478
FSRV	-0,000169	-0,002627	0,981979	0,683239
Turnover	-0,036107	-0,034874	0,098886	0,001976

Contre nos attentes, en incluant le turnover cette fois, AIM reste (positif et) très significatif, et le turnover de même, ce qui voudrait dire que, plus tôt, nous avions juste des données non convaincantes de la réelle différence de prime de risque entre AIM et turnover.

L'existence avec forte significativité d'une prime de risque pour un AIM grand alors que la taille, le FSRV et le turnover sont pris en compte (les autres mesures ne pouvant pas être prises en compte faute de données suffisantes) et sont elles-mêmes d'importance significative sur le prix à l'exception de FSRV qui entre en conflit avec la taille nous convainc donc qu'AIM est une bonne mesure d'asymétrie informationnelle et non une proxy de celles étudiées jusqu'à présent.

3.2.13 Régression sur bêta, AIM, PIN, taille, BTM, FSRV, turnover et illiq_amihud

	Moyenne (1)	Moyenne (2)	P-valeur (1)	P-valeur (2)
Intercept	-0,016985	-0,019140	0,255057	0,147166
Bêta	-0,004974	-0,003969	0,056120	0,091741
AIM	0,000004	0,000205	0,996218	0,796765
PIN	0,008535	0,024477	0,529579	0,035957
Taille (<i>log_market_cap</i>)	0,001644	0,001707	0,029194	0,012910
<i>log</i> (BTM)	0,000733	-0,000194	0,661718	0,894349
FSRV	0,003900	0,002143	0,630235	0,754362
Turnover	-0,046566	-0,025553	0,041029	0,142488
Illiq_amihud	-0,000276	-0,000222	0,643016	0,337655

A titre indicatif, voici la régression lorsqu'on inclut tous les paramètres. On obtient une p-valeur très mauvaise pour AIM (79,67%). On obtient des p-valeurs à peu près faibles pour certains des paramètres, mais nos études précédentes nous laissent supposer que cela n'est dû qu'à un phénomène d'*overfitting*.

3.3 Stratégies actives de gestion de portefeuilles

Nous nous sommes demandé s'il était possible d'utiliser l'asymétrie informationnelle dans la gestion active de portefeuille. Pour répondre à cette question, nous avons choisi d'implémenter une stratégie active de gestion de portefeuille basée sur l'asymétrie informationnelle et plus particulièrement la mesure AIM, afin de mesurer l'effet de cette mesure sur la rentabilité d'un portefeuille.

Tous les mois, nous regardons quels ont été les titres avec la plus grande asymétrie informationnelle le mois passé. Nous sélectionnons les 10% de titres au plus grand AIM, et le portefeuille est composé de manière équirépartie de ces titres.

Ceci est réitéré tous les mois pendant 10 ans, entre janvier 1991 et décembre 2000.

Nous avons comparé notre portefeuille au portefeuille de marché, et nous nous sommes intéressés à l'excès de rentabilité associé : l'alpha de Jensen.

Sur notre stratégie, nous obtenons un alpha de 0,765% avec une p-valeur associée de 1,59%. Ces résultats montrent que la stratégie possède un excédent de rentabilité, avec une significativité importante.

Nous n'avons pas comparé notre stratégie avec une autre, mais la présence d'un alpha significativement positif est un résultat intéressant, puisque cela signifie que dans le modèle traditionnel du MEDAF, nous pourrions nous baser sur la mesure AIM pour obtenir un excédent de rentabilité par rapport au marché.

La valeur obtenue est surprenante, car il s'agit d'un alpha calculé sur des rentabilités mensuelles, et nous avons obtenu une grande valeur. Cet excès de rentabilité est très certainement lié au risque d'asymétrie informationnelle non pris en compte dans le modèle de régression utilisé. Il est fort probable que l'on obtienne un alpha nul si nous prenions en compte l'asymétrie d'information lors de ce calcul.

4 Conclusion

Si l'on ignore les résultats trouvés sur le bêta, les régressions nous montrent qu'AIM est une mesure d'asymétrie informationnelle qui apporte des informations distinctes de celles apportées par la taille, le FSRV et le *Stock Turnover*. Nous n'avons pas pu conclure sur le PIN, le BTM et la mesure d'illiquidité d'Amihud par manque d'information.

Les corrélations sont cohérentes avec ce résultat.

Il faudrait cependant refaire le travail avec assez de données pour pouvoir tester correctement la relation d'AIM à ces 3 mesures, et sur des données où le coefficient en bêta est cohérent avec le modèle théorique, avant de pouvoir se prononcer.

5 Annexes

5.1 Annexe 1

Si W suit une loi normale, alors $\mathbb{E}(e^{-aW}) = e^{-a\mathbb{E}(W) + \frac{a^2}{2}\text{Var}(W)}$. En effet :

$$\begin{aligned} X &\sim \mathcal{N}(\mathbb{E}(W), \text{Var}(W)) \\ \Rightarrow -aW &\sim \mathcal{N}(-a\mathbb{E}(W), a^2\text{Var}(W)) \\ \Rightarrow e^{-aW} &\sim \text{Log-}\mathcal{N}(-a\mathbb{E}(W), a^2\text{Var}(W)) \\ \Rightarrow \mathbb{E}(e^{-aW}) &= e^{-a\mathbb{E}(W) + \frac{a^2}{2}\text{Var}(W)} \end{aligned}$$

5.2 Annexe 2

```
# -*- coding: utf-8 -*-
"""
Created on Thu Mar 15 15:06:54 2018

@author: camusj
"""

# Imports
try:
    import numpy as np
    np.random.seed(0)
except ImportError:
    raise ImportError('numpy module needs to be imported.')
try:
    import matplotlib.pyplot as plt
except ImportError:
    raise ImportError('matplotlib.pyplot module needs to be imported.')

def f(lmbda, N, E):
    return np.linalg.inv(np.dot(np.linalg.inv(N), np.dot(E, np.linalg.inv(N)))
    + lmbda*np.linalg.inv(N)) + np.linalg.inv(np.linalg.inv(N) +
    lmbda*np.linalg.inv(E))

def frobenius(A):
    return np.linalg.norm(A, 'fro')

def plot_cov_curve(x_array, y_array):
    plt.plot(x_array, y_array, linewidth=5)

    plt.xticks(fontsize=10);
    plt.yticks(fontsize=10);
```

```

plt.xlabel('Fro(N)', fontsize=15);
plt.ylabel('Fro(Cov)', fontsize=15);

plt.legend(loc='best', fontsize=5);

# Main function
if __name__ == '__main__':

    # Initialization
    a = 0.01
    rf = 0.0    # R = 1
    lambda = 0.3
    E = np.array([[120, 30],[30, 8]])
    Z = np.array([[50, 0],[0, 5]])
    x_array = []
    y_array = []

    # Loop over x
    for x in range(150):
        T = np.array([[90 + x, 10],[10, 30 + x]])
        N = T - np.dot(T, np.dot(T + (a/lambda)**2 * np.dot(E, np.dot(Z, E)),
                                T))
        Cov = -lambda*f(lambda, N, E) - lambda**2*np.dot(f(lambda, N, E),
                                np.dot(np.linalg.inv(T - N), f(lambda, N, E)))

        # Append to arrays
        x_array.append(frobenius(N))
        y_array.append(frobenius(Cov))
        #x_array.append(N[0][0])
        #y_array.append(Cov[0][0])

    # Plot
    plot_cov_curve(x_array, y_array)

```