

Université de Montréal

**Étude sur les modèles à facteurs dynamiques et les
rendements financiers**

par

Oscar Calderon
CALO23068000

Département de Sciences Économiques
Faculté des arts et sciences

Rapport de recherche présenté
en vue de l'obtention du grade de
Maîtrise en Sciences Économiques

Août 2011

Sommaire

Les facteurs dynamiques permettent de résumer l'information de plusieurs séries en quelques facteurs communs. Ici, j'applique les méthodes de seuillages dur et mou aux séries afin de diminuer la dimension des panels avant d'estimer les facteurs. Ces facteurs sont ensuite utilisés pour prédire les rendements excédentaires. Dans un cadre prédictif, les facteurs estimés à partir des panels ayant été soumis à un seuillage sont plus informatifs que les facteurs estimés avec toutes les données disponibles. En général, on peut aussi diminuer l'erreur quadratique moyenne des prévisions des rendements excédentaires faites à partir des facteurs en soumettant les séries à un seuillage. Pour les données trimestrielles, les facteurs estimés à partir d'un panel augmenté soumis à un seuillage mou expliquent jusqu'à 19% de la variation des rendements excédentaires et démontrent un pouvoir de prévision hors échantillon statistiquement significatif sur toutes les périodes considérées. Les facteurs sont ensuite utilisés comme instruments dans des tests du MÉDAF conditionnel pour expliquer la coupe transversale des 25 portefeuilles Fama-French. Dans ce cadre, en se limitant à deux instruments on réussit à obtenir un meilleur ajustement, tel que calculé par le \bar{R}^2 en coupe transversale, qu'avec le MÉDAF ou le modèle à trois facteurs de Fama-French et à réduire considérablement les moyennes des erreurs d'évaluation des actifs par rapport à ces deux modèles.

Table des matières

Table des matières	iii
Liste des Tableaux	iv
Graphiques	iv
Introduction	1
1 Revue de la littérature	2
2 Modèle théorique	4
3 Seuillage	6
3.1 Seuillage dur	7
3.2 Seuillage mou	8
3.2.1 Régression des moindres angles	9
3.2.2 Elastic Net	11
4 Données empiriques	12
5 Résultats empiriques	14
5.1 Régressions prédictives mensuelles	15
5.2 Régressions prédictives trimestrielles	17
6 Expérience hors-échantillon	19
7 Coupe Transversale des Rendements	22
Conclusion	24
Références	26

Liste des tableaux

1	Description des panels utilisés pour extraire les facteurs dynamiques	29
2	Nombre de facteurs communs estimés par IC_{p2} pour les divers panels	29
3	Régressions des rendements excédentaires mensuels sur les facteurs estimés	30
4	Régressions des rendements excédentaires mensuels sur les facteurs estimés et Z_t	31
5	Régressions des rendements excédentaires mensuels sur les facteurs estimés et Z_t , configurations optimales	32
6	Régressions des rendements excédentaires trimestriels sur les facteurs estimés et Z_t , configurations optimales	33
7	Résultat des expériences hors-échantillon Facteurs vs Constante pour les données mensuelles	34
8	Résultat des expériences hors-échantillon Facteurs vs Constante pour les données trimestrielles	35
9	Résultat des expériences hors-échantillon Facteurs vs Constante et cay_t pour les données trimestrielles	36
10	Tests Fama-MacBeth du MÉDAF conditionnel avec deux Instruments	37

Table des figures

1	Graphique des rendements estimés sur les rendements réalisés	38
---	--------------------------------------------------------------	----

Introduction

La prévision des rendements excédentaires est un problème sur lequel plusieurs économistes se sont penchés. La question principale étant de savoir si les rendements peuvent être prévus ou non. En général, on s'accorde pour dire qu'ils le peuvent. La plupart de recherches sur le sujet utilisent des variables qui devraient, dans un cadre théorique du moins, contenir de l'information sur les rendements financiers. Ici, je suppose que les performances de l'économie réelle et des marchés financiers sont reliées et que cette relation doit se refléter sur un panel de plusieurs séries temporelles. Le problème principal est que je ne peux pas faire une régression avec toute l'information disponible. J'estime alors une petite quantité de facteurs communs à partir des séries temporelles qui seront utilisés dans la prévision des rendements excédentaires. Les facteurs sont estimés par la méthode des composantes principales.

Le problème qui se pose désormais c'est à savoir si le fait qu'on puisse travailler avec un panel contenant un nombre substantiel de séries implique que c'est ce qui devrait être fait. Boivin et Ng (2006) ont démontré que si les séries sont non informatives le fait d'inclure celles-ci dans notre panel peut en fait être nocif dans un contexte de prévision. Je suis ici la méthodologie de Bai et Ng (2008), qui utilisent des méthodes de réduction et de sélection de panels pour améliorer les prévisions des facteurs sur l'inflation. Je permets aussi une relation non linéaire entre les facteurs dynamiques et les séries temporelles en incluant le carré des séries dans un panel augmenté. J'applique donc les mêmes méthodes de réduction sur deux panels : un contenant les séries temporelles et l'autre qui est augmenté du carré de celles-ci. Afin de voir si ceci aura un effet sur les prévisions, je teste ensuite le pouvoir explicatif des facteurs dynamiques, ainsi obtenus, avec une régression prédictive des rendements excédentaires sur ceux-ci. Les caractéristiques des facteurs estimés de panels soumis à un seuillage sont aussi évaluées dans le cadre d'une expérience hors échantillon. Les facteurs sont ensuite utilisés afin de mesurer leur pouvoir explicatif de la coupe transversale des rendements des 25 portefeuilles Fama-French.

Après une courte revue de la littérature dans la section 1, je décris dans la section 2 le modèle et les fondements théoriques des modèles à facteurs dynamiques. La section 3 décrit les méthodes de seuillages que j'applique à mon panel. Par la suite, je décris les données qui sont utilisées dans ce travail dans la section 4. La section 5 présente les résultats empiriques des régressions prédictives des rendements excédentaires sur les facteurs dynamiques. La section 6 offre les résultats des expériences hors échantillon. Le

pouvoir explicatif des facteurs dynamiques sur la coupe transversale des rendements des 25 portefeuilles Fama-French triés par taille et valeur est évalué à la section 7.

1 Revue de la littérature

Le premier aspect important de ce travail est que je dispose d'un panel duquel je veux estimer les facteurs dynamiques. La méthode que j'utilise pour faire ceci est la même que dans Stock et Watson (2002a,b) c'est-à-dire la méthode des composantes principales. Dans Stock et Watson (2002a), les auteurs utilisent 215 séries temporelles afin de faire des prévisions en expérience hors échantillons de huit séries macroéconomiques, 4 séries de variables réelles et 4 séries de variables d'inflation, sur des horizons de 6 mois, 1 an et 2 ans. Ils comparent les prévisions de ces séries avec d'autres méthodes en utilisant l'erreur quadratique moyenne relative. Ils trouvent que les facteurs dynamiques sont plus performants dans les prévisions des séries réelles que les autres méthodes. De plus, dans Stock et Watson (2002b) les auteurs montrent que les facteurs latents peuvent être estimés de façon consistante par la méthode des composantes principales en présence d'une faible autocorrélation des erreurs ainsi que d'une faible corrélation des erreurs en coupe transversale, à condition d'avoir un nombre assez grand de séries et d'observations. Ces facteurs peuvent par la suite être utilisés pour faire des prévisions de la variable dépendante.

Boivin et Ng (2006) étudient les effets du nombre de séries temporelles qui sont utilisées pour faire l'analyse des facteurs. Ils se penchent sur les effets du nombre de séries utilisées pour estimer les facteurs. Ils montrent ainsi que le pouvoir de prévision des facteurs dynamiques peut être inférieur lorsqu'en présence d'un nombre plus important de séries, surtout si celles-ci ont des erreurs corrélées. Ils concluent aussi que certains facteurs plus importants pour faire des prévisions de la variable en question et estimés à partir de panels moins grands peuvent être dominés dans des panels avec un nombre plus important de séries.

Dans ce travail, je suis les méthodes utilisées dans Bai et Ng (2008) pour la sélection des séries temporelles. Dans cet article, les auteurs utilisent des méthodes de seuillages, dont une basée sur l'algorithme de régression du moindre angle¹ de Efron et al. (2004), appliqué à l'estimateur "elastic net"

1. Least Angle Regression

afin de choisir les séries qui sont utilisées pour l'analyse de facteurs. Ils appliquent la méthode des composantes principales à ces séries, pour estimer les facteurs et faire des prévisions sur l'inflation à différents horizons. Ils concluent que de soumettre les séries à une forme de seuillage permet de diminuer l'erreur quadratique moyenne des prévisions à tous les horizons. Ils considèrent aussi une méthode des composantes principales carrée qui permet la relation entre les facteurs et les séries d'être non linéaire, et trouve qu'en général il y a des gains additionnels à faire ceci.

Finalement, le dernier aspect de mon travail c'est la prévision des rendements. Les modèles à facteurs sont déjà utilisés dans des recherches sur la prédictibilité des rendements. Par exemple, Ludvigson et Ng (2007) considèrent des facteurs extraits de deux panels, un contenant des séries macroéconomiques et l'autre contenant des séries financières, afin de prévoir les rendements excédentaires trimestriels du marché et sa volatilité. Ils concluent que même en conditionnant sur d'autres variables communément utilisées, les facteurs ont de l'information supplémentaire lorsqu'estimés à partir des séries financières et les facteurs estimés à partir des données macroéconomiques sont importants pour les modèles de prévisions de la volatilité. Mönch (2006) quant à lui utilise des facteurs dynamiques afin d'étudier les modèles conditionnels d'évaluation des actifs financiers. Il conclut que la coupe transversale des rendements excédentaires des 25 portefeuilles de Fama-French est mieux expliquée avec des facteurs dynamiques comme instruments que par les autres variables utilisées.

D'ailleurs, je conclus ce travail en faisant une analyse du pouvoir explicatif des facteurs sur la coupe transversale des rendements des 25 portefeuilles de Fama-French. Fama et MacBeth (1973) développent la méthode que je vais utiliser. Ils étudient la coupe transversale des rendements excédentaires dans le cadre du MÉDAF². Fama et French (1993) utilisent comme facteurs le rendement excédentaire du marché, ainsi que SMB "small minus big" et HML "high minus low", deux facteurs qu'ils développent dans cet article, pour expliquer la coupe transversale des rendements excédentaires dans le cadre d'un MÉDAF inconditionnel. Ce modèle donne un meilleur ajustement que celui du MÉDAF. Lettau et Ludvigson (2001b) utilisent comme instrument le ratio consommation-richeesse "cay" pour étudier la coupe transversale des rendements excédentaires dans le cadre de modèles d'évaluation des actifs financiers conditionnels. Ils obtiennent ainsi un ajustement comparable à celui du modèle à trois facteurs de Fama-French.

2. Modèle d'évaluation des actifs financiers

2 Modèle théorique

L'intérêt est de savoir si nous pouvons faire des prévisions sur les rendements excédentaires du marché, $r_{m,t+1}^e$. Pour ce faire, je veux estimer la régression prédictive suivante :

$$r_{m,t+1}^e = \beta' Z_t + \epsilon_{t+1}. \quad (1)$$

Pour estimer le modèle de l'équation (1), j'ai accès à un panel X de $T \times N$ avec les éléments x_{it} , $i = 1, \dots, N$ et $t = 1, \dots, T$, où N est grand. Dans le cadre d'une régression multivariée ordinaire telle que celle de l'équation (1), il n'est pas possible d'utiliser la totalité du panel et il faut donc trouver un moyen de choisir les séries temporelles qui formeront Z_t et ainsi estimer le modèle.

Une façon de garder le modèle parcimonieux tout en gardant l'information pertinente contenue dans le panel X est de supposer que les données admettent une représentation de structure à facteur de la forme :

$$x_{it} = \lambda_i' f_t + e_{it} \quad (2)$$

où f_t est un vecteur $r \times 1$ de facteurs communs ; λ_i est un vecteur correspondant $r \times 1$ de sensibilités et e_{it} est l'erreur idiosyncratique. Il est important de noter qu'ici $r \ll N$, d'où l'attrait pour cette méthode.

La relation entre les rendements excédentaires du marché et les facteurs communs est représentée par l'équation suivante :

$$r_{m,t+1}^e = \alpha' F_t + \beta' Z_t + \epsilon_{t+1} \quad (3)$$

où ϵ_{t+1} est l'erreur de prévision et $F_t \subset f_t$, donc on fait la distinction entre les facteurs communs qui expliquent la covariance des séries temporelles et ceux qui sont importants dans la régression de l'équation (3), F_t peut inclure des fonctions non linéaires des facteurs f_t . Si les facteurs communs f_t étaient observés on pourrait simplement inclure ceux qui sont importants dans la régression de l'équation (3), mais ce n'est pas le cas et il faut donc les estimer. Stock et Watson (2002b) montrent que lorsque $N \rightarrow \infty$ et $T \rightarrow \infty$, on peut estimer de façon consistante l'espace engendré par les facteurs communs par la méthode des composantes principales. Ces estimations des facteurs peuvent par la suite être utilisées comme si c'était les facteurs qu'on observait dans la régression prédictive de l'équation (3).

Puisque le nombre r de facteurs sous-jacents est inconnu il faut estimer ce nombre. Dans Stock et Watson (1998, 2002a) les auteurs utilisent

un critère d'information Bayésien appliqué à la variable dépendante y_{t+1} . Cependant, Bai et Ng (2002) notent que ceci n'est pas la même chose que d'estimer le nombre de facteurs sous-jacents à un panel de séries temporelles. Ils développent donc des critères d'information pour estimer de façon consistante le nombre de facteurs sous-jacents à un panel, en considérant la somme des erreurs du panel au carré au lieu de la somme de l'erreur au carré de la régression de la variable dépendante sur les facteurs. De plus, les pénalités sont fonction des deux dimensions soit N et T . Ce sont ces critères qui seront utilisés pour choisir le nombre de facteurs à estimer pour chaque modèle considéré, principalement le critère IC_{P2} :

$$IC_{P2} = \ln(V(k, \hat{f}^k)) + k \left(\frac{N+T}{NT} \right) \ln C_{NT}^2 \quad (4)$$

où k est le nombre de facteurs, $C_{NT}^2 = \min(N, T)$ et

$$V(k, \hat{f}^k) = \frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T \left(x_{it} - \lambda_i^{k'} \hat{f}_t^k \right)^2 \quad (5)$$

Par la méthode des composantes principales, j'estime k facteurs \hat{f}_t , pour k choisi par le critère IC_{P2} . Par la suite, je forme \hat{F}_t en choisissant différentes combinaisons de \hat{f}_t , en pratique je limite à 10 le nombre de facteurs pouvant être estimés dans chaque cas sauf pour les panels 1 et 2 de données mensuelles, dans quel cas $kmax = 20$. Pour chaque configuration de variable, je régresse $r_{m,t+1}^e$ sur \hat{F}_t et Z_t et j'évalue le critère d'information Bayésien (BIC) et le R^2 , choisissant la combinaison qui diminue le BIC. Par la suite, diverses fonctions des facteurs qui ont contribué à diminuer le BIC sont considérées comme le carré de ceux-ci et leurs produits croisés. Je ne considère donc pas des fonctions des facteurs qui n'étaient pas significatifs, ceci dans le but de limiter la recherche de données.

L'équation (2) présuppose une relation linéaire entre les facteurs communs et les séries temporelles. Comme tels, les facteurs estimés sont des combinaisons linéaires des séries, qui minimisent les sommes sur les deux dimensions du carré des résidus. Une relation plus flexible entre les facteurs et les séries peut être décrite par l'équation suivante :

$$g(X_{it}) = \phi_i J_t + e_{it} \quad (6)$$

où $g(\cdot)$ est une fonction non linéaire, J_t sont les facteurs communs et ϕ_i représente les sensibilités. Par exemple, la fonction objective peut impliquer

une combinaison entre les séries et leur carré, on aurait donc le panel augmenté $X^* = [X_1 X_2 \dots X_N X_1^2 X_2^2 \dots X_N^2]$. Par conséquent, on obtiendrait comme relation entre les facteurs communs et nos données l'équation suivante :

$$X_t^* = \phi J t + e_t \quad (7)$$

La méthode des composantes principales peut être utilisée pour estimer les facteurs ainsi que les sensibilités de l'équation (7). Les facteurs estimés seront maintenant une combinaison linéaire des séries incluses dans X_t^* permettant donc une non-linéarité avec les séries X_t .

Afin de permettre une non-linéarité entre les facteurs estimés et les séries, j'ajoute au panel le carré des données soit x_{it}^2 , comme le font Bai et Ng (2008). Ils nomment cette méthode, la méthode des composantes principales carrée. Ils proposent aussi un modèle où les produits croisés des séries temporelles sont ajoutés et appellent ce modèle la méthode des composantes principales quadratique. Cependant, les auteurs affirment que les gains à utiliser un panel quadratique ne sont pas assez importants pour les coûts computationnels que ceux-ci entraînent. Par conséquent, dans ce rapport je n'utilise qu'un panel carré comme celui de l'équation (7). En considérant ce type de panel, on se retrouve maintenant avec $2N$ séries. Dépendamment de N , ceci peut nous donner un panel avec un nombre important de séries. Il est possible que certaines séries dans ce type de panel ne contiennent pas beaucoup d'information sur les rendements excédentaires et qu'elles entraînent donc plus de bruit, ce qui aurait comme effet de baisser la qualité des facteurs pour l'utilisation qu'on veut en faire. La prochaine section répond à la question de comment faire pour choisir les séries utilisées pour en extraire les facteurs dynamiques, de manière à ce que ceux-ci contiennent le plus d'information pertinente sur les rendements excédentaires.

3 Seuillage

Le seuillage³ est une méthode utilisée dans Bai et Ng (2008) afin de choisir les séries à inclure dans le panel de données, avant d'estimer les facteurs. Ils utilisent deux types de seuillage : le seuillage dur qui ne prend pas en compte les séries déjà incluses et choisit donc si la série est à ajouter si elle est statistiquement significative dans une régression univariée⁴; et le

3. Thresholding

4. En pratique ce n'est pas une régression univariée, car les autres variables pertinentes dans le modèle sont ajoutées dans la régression.

seuillage mou qui choisit les séries en fonction de celles déjà incluses. La justification pour soumettre les panels à un seuillage avant d'estimer les facteurs est que ceux-ci peuvent inclure des séries individuelles qui ne contiennent pas beaucoup d'information sur la variable dépendante. Boivin et Ng (2006) ont conclu que le fait d'avoir plus de données lorsque celles-ci n'ont pas beaucoup d'information sur les facteurs n'améliore pas les prévisions et dans certains cas peut considérablement augmenter les erreurs de prévisions. On peut donc diminuer le nombre de séries temporelles afin d'utiliser celles qui contiennent le plus d'information sur les rendements excédentaires du marché. Dans ce qui suit, je décris les deux méthodes de seuillage utilisées dans ce rapport.

3.1 Seuillage dur

La méthode de seuillage dur telle qu'utilisée par Bai et Ng (2008) consiste à faire un simple test avec la statistique de Student (t) d'une régression de y_{t+h} sur x_{it} en incluant les autres variables pertinentes w_t qui seront utilisées dans le modèle. Par la suite, on applique une règle de seuillage pour décider quelles séries on garde dans le panel.

L'algorithme utilisé par Bai et Ng (2008) est comme suit :

1. Pour chaque série temporelle $i = 1, \dots, N$ on fait la régression de y_{t+h} sur W_t et x_{it} .
2. Par la suite, on prend la valeur absolue des statistiques-t et on les met en ordre descendant soit $|t_1|, |t_2|, \dots, |t_N|$.
3. Pour un seuil α particulier, on garde les séries qui ont une valeur absolue de la statistique-t supérieure à la valeur de la statistique-t pour le seuil choisi.
4. Pour un nombre k_α^* qui est le nombre de séries dont la valeur de la statistique-t est supérieure au seuil choisi, on forme le panel de donnée $x_t(\alpha) = (x_{1t}, x_{2t}, \dots, x_{k_\alpha^* t})$.
5. On estime les facteurs f_t avec $x_t(\alpha)$.

Bai et Ng (2008) utilisent les valeurs de statistique-t de 1.28, 1.65 et 2.58, qui correspondent aux seuils pour des tests bilatéraux de 20%, 10% et 1%, respectivement. Dans le cas des rendements excédentaires une valeur de la statistique-t trop élevée telle que 2.58, fait en sorte que la dimension du panel est trop petite afin d'en estimer les facteurs. Par conséquent, le seuil que j'ai choisi pour appliquer cette méthode prend en considération le fait que je ne veux pas me retrouver avec un nombre de séries trop petit, ce qui entraînerait une mauvaise estimation des facteurs. Bien que l'objectif soit de diminuer la dimension du panel, il faut garder en tête que la méthode des

composantes principales requiert en fait un N assez grand afin d'estimer de façon consistante les facteurs.

Un problème peut se poser lorsqu'on considère une série à la fois afin de prendre la décision de l'inclure dans le panel. Les séries choisies peuvent être très semblables et, par conséquent, contenir très peu d'information marginale supplémentaire. De plus, en choisissant des séries trop semblables on augmente la possibilité que celles-ci aient des erreurs corrélées. Dans Boivin et Ng (2006), une des raisons pour considérer un panel avec moins de séries était justement le fait que d'avoir trop de séries pouvait entraîner des erreurs corrélées en coupe transversale. Dans le cadre de la méthode des composantes principales appliquée à une structure de facteur dynamique approximative, on fait l'hypothèse que les erreurs sont faiblement corrélées en coupe transversale. Donc en choisissant des séries très semblables on peut en fait affecter l'estimation des facteurs qui seront ensuite utilisés pour la régression décrite à l'équation (3). Ceci sert de justification à chercher d'autres méthodes afin de faire la sélection des séries, surtout une méthode qui prend en compte l'information qui est déjà contenue dans les autres séries. Dans la prochaine section, je décris la méthode de seuillage mou qui sera utilisée dans ce rapport et qui permet de corriger ce problème.

3.2 Seuillage mou

Une méthode permettant de prendre en considération les autres variables qui forment le panel est de faire la sélection des séries temporelles par une régression pénalisée. Un type de régression pénalisée particulière est le Lasso⁵ de Tibshirani(1996), qui choisit les coefficients comme la solution au problème :

$$\min_{\beta} RSS + \lambda \sum_j |\beta_j| \quad (8)$$

où RSS est la somme des résidus au carré de la régression de la variable dépendante sur toutes les variables. L'avantage de cet estimateur est que le Lasso va en fait mettre des coefficients à zéro, contrairement à l'estimateur ridge qui a comme pénalité $\sum_j \beta_j^2$. L'estimateur ridge va baisser les estimations des coefficients vers zéro, mais il ne les mettra pas égaux à zéro, il va plutôt faire baisser la valeur absolue de tous les coefficients. Dans un cadre orthogonal, le Lasso va estimer les coefficients à zéro si leur valeur absolue est trop petite, les estimations seront donc :

$$\tilde{\beta}_i = \text{sign}(\hat{\beta}_i)(|\hat{\beta}_i| - \gamma)^+ \quad (9)$$

5. "Least absolute shrinkage selection operator"

où γ est déterminé par la condition $\sum |\beta_j| = t$. D'ailleurs, dans Tibshirani (1996) l'auteur note que cet estimateur a la même forme que l'estimateur "seuillage mou"⁶ de Donoho et Johnstone. Par conséquent, le Lasso est plus adapté à l'usage qu'on veut en faire, qui est de faire la sélection des variables et diminuer le panel.

3.2.1 Régression des moindres angles

L'estimation des coefficients de Lasso peut être faite par l'algorithme LARS⁷ proposé par Efron et al. (2004). L'algorithme tel que décrit dans l'article pour implémenter LARS est expliqué ci-après. Supposons que nous sommes à l'étape k , nous avons un sous-ensemble des variables K qui est l'ensemble des indices actifs et $\hat{\mu}_K$ est l'estimation courante. On trouve les corrélations courantes $\hat{c} = X'(y - \hat{\mu}_K)$. On définit l'ensemble des indices actifs K comme les indices des corrélations absolues les plus grandes :

$$\hat{C} = \max_j \{|\hat{c}_j|\} \quad K = \{j : |\hat{c}_j| = \hat{C}\} \quad (10)$$

On définit la matrice active associée à K comme :

$$X_K = (\dots s_j x_j \dots)_{j \in K} \quad (11)$$

où s_j est le signe de \hat{c}_j . On définit aussi par la suite :

$$G_K = X_K' X_K, \quad A_K = (1_K' G_K^{-1} 1_K)^{-\frac{1}{2}}, \quad (12)$$

où 1_K est un vecteur de 1's dont la longueur est égale à la taille de K , $|K|$. On définit ensuite le vecteur équiangulaire unitaire, qui forme des angles égaux avec les colonnes de la matrice de l'ensemble actif X_K :

$$u_K = X_K w_K \quad \text{où} \quad w_K = A_K G_K^{-1} 1_K \quad (13)$$

de manière à ce que $X_K' u_K = A_K 1_K$ et $\|u_K\|^2 = 1$. On trouve le produit intérieur $a_K = X_K' u_K$.

Finalement, la nouvelle estimation $\hat{\mu}_{K+}$ est :

$$\hat{\mu}_{K+} = \hat{\mu}_K + \hat{\gamma} u_K \quad (14)$$

où

6. soft threshold
7. Least Angle Regression

$$\hat{\gamma} = \min_{j \in K^c} + \left(\frac{\hat{C} - \hat{c}_j}{A_K - a_j}, \frac{\hat{C} + \hat{c}_j}{A_K + a_j} \right) \quad (15)$$

Le minimum dans l'équation (15) n'est pris que sur les valeurs positives de l'ensemble. La valeur de $\hat{\gamma}$ de l'équation (15) est la valeur la plus petite de γ dans $\mu(\gamma) = \hat{\mu}_K + \gamma u_K$, afin qu'une autre variable entre dans l'ensemble actif. Pour l'indice \hat{j} , qui est l'indice correspondant à la valeur qui minimise (15), le nouvel ensemble actif K_+ est $K \cup \hat{j}$. À la première étape, l'estimation courante est $\hat{\mu}_0 = 0$. Avec cet algorithme, après K étapes nous allons avoir K variables dans l'ensemble actif et les estimations des coefficients des autres variables seront donc égales à zéro. L'intérêt pour mon cas ne porte pas sur les valeurs estimées pour les β , mais sur l'ordre des variables que cet algorithme va me donner.

Afin d'avoir une certaine intuition de ce que fait cet algorithme, commençons au début. À la première étape, on trouve les corrélations entre la variable dépendante et les variables indépendantes. On commence donc en ajoutant la série dont la corrélation absolue avec y est la plus forte, supposons x_1 . L'estimation $\hat{\mu}$ est donc renouvelée pour inclure l'information de la nouvelle variable. Cependant, au lieu d'augmenter l'estimation dans la direction de x_1 au nouveau maximum qu'on pourrait aller, qui correspond à la projection de y sur x_1 , l'algorithme va augmenter $\hat{\mu}$ de manière à ce que la corrélation entre les résidus $(y - \hat{\mu})$ et la variable x_1 baisse et soit égale à la prochaine variable ayant la corrélation la plus élevée, x_2 , et on ajoute celle-ci à l'ensemble de variables actives. L'estimation $\hat{\mu}$ est donc augmentée à nouveau dans une direction qui est équiangulaire entre les deux variables, faisant donc baisser la corrélation de ces deux variables et les résidus du même niveau jusqu'à ce que celles-ci soient égales à la corrélation d'une troisième variable, qui sera donc ajoutée. L'algorithme poursuit ainsi son estimation équiangulairement entre x_1, x_2 et x_3 , soit dans la direction des moindres angles, jusqu'à ce qu'une quatrième variable entre dans l'ensemble des plus corrélés et ainsi de suite. L'algorithme arrête lorsque K variables sont incluses dans l'ensemble des variables actives, pour un K choisi au préalable. Si nous poursuivons l'algorithme jusqu'à $K = N$, où N est le nombre de variables incluses dans le panel, nous obtiendrions les coefficients correspondants à ceux des MCO.

Comme l'ont démontré Efron et al. (2004), cet algorithme nous donne des résultats semblables à ceux du Lasso, et avec une petite altération on obtient exactement les mêmes résultats qu'avec le Lasso. Pour implémenter

l'algorithme je définis $K = 30$ comme dans Bai et Ng (2008) pour les données mensuelles et $K = 45$ pour les données trimestrielles. Je choisis donc, 30 et 45 séries parmi celles disponibles pour estimer les facteurs pour les deux types de panels soient le standard et le panel augmenté. J'ai fait des tests avec des nombres plus importants de séries, mais ceci a eu pour effet de diminuer la significativité des facteurs dans les régressions prédictives, ainsi que de diminuer les coefficients de détermination des régressions, pour un nombre de facteurs estimés fixes. Dans la pratique, afin d'implémenter l'algorithme LARS la variable dépendante a été centrée et les séries temporelles du panel ont été centrées et réduites.

3.2.2 Elastic Net

Bien que la caractéristique du Lasso qui fait qu'il met les coefficients égaux à zéro soit une caractéristique intéressante dans le cadre d'un seuillage, il reste que cet estimateur a certaines faiblesses. La plus importante pour notre cas est que s'il y a un groupe de variables qui sont fortement corrélées l'estimateur Lasso va en choisir une seule parmi le groupe et estimer les coefficients des autres variables du groupe à zéro, ou dans notre cas une seule de ces variables va entrer dans l'ensemble actif. Pour corriger cet effet, Zou et Hastie (2005) proposent un estimateur appelé "elastic net". La fonction objective de cette régression pénalisée est :

$$\min_{\beta} RSS + \lambda_1 \sum_{j=1}^N |\beta_j| + \lambda_2 \sum_{j=1}^N \beta_j^2 \quad (16)$$

où RSS est la somme des résidus au carré de la régression de la variable dépendante sur toutes les variables. La pénalité de cet estimateur est donc une combinaison convexe de la pénalité du Lasso et de la pénalité de l'estimateur ridge. Ceci a donc pour effet de diminuer les faiblesses qui affectent les deux estimateurs.

Une caractéristique intéressante de cet estimateur est qu'on peut redéfinir les variables ainsi :

$$X^+ = (1 + \lambda_2)^{-1/2} \begin{pmatrix} X \\ \sqrt{\lambda_2} I_N \end{pmatrix} \text{ et } y^+ = \begin{pmatrix} y \\ 0_N \end{pmatrix} \quad (17)$$

On peut donc reformuler la fonction objective de cet estimateur comme suit :

$$\min_{\beta} \sum (y^+ - X^+ \beta^*)^2 + \gamma \sum_j |\beta_j^*| \text{ où } \gamma = \frac{\lambda_1}{\sqrt{1 + \lambda_2}} \quad (18)$$

On peut voir que sous cette nouvelle formulation le problème devient un problème de Lasso. Ceci est important, car il nous permet d'utiliser les mêmes algorithmes qui sont utilisés pour estimer la régression Lasso, tel que l'algorithme LARS décrit plus haut.

Lorsqu'on applique le LARS-EN, le choix de λ_1 devient en fait un choix sur K , le nombre de variables dans l'ensemble actif. Pour l'autre variable de choix λ_2 Bai et Ng (2008) choisissent des valeurs de 0.25, 0.5, et 1.5. J'ai trouvé que pour ces valeurs de λ_2 les résultats étaient exactement les mêmes, ce qui me fait conclure que, pour mon cas, c'est λ_1 qui domine par le choix très parcimonieux de K et que par conséquent pour ces valeurs le choix des variables est le même que pour le Lasso. Pour voir les effets des variables de choix, le problème d'optimisation peut être réécrit comme suit :

$$\hat{\beta} = \arg \min_{\beta} \sum (y - X\beta)^2 \text{ sujet à } (1 - \alpha) \sum_j |\beta_j| + \alpha \sum_j \beta_j^2 \leq t \quad (19)$$

où $\alpha = \lambda_2 / (\lambda_1 + \lambda_2)$. Ainsi on note que lorsque α se rapproche de 1, nous obtenons comme résultat l'estimateur ridge. On obtient à l'inverse l'estimateur Lasso lorsque $\alpha = 0$. Donc, pour obtenir une combinaison convexe des deux je choisis une valeur de λ_2 plus élevée que celles utilisées dans Bai et Ng (2008), dans les expériences que j'ai conduites je trouve que le choix optimal était aux alentours de $\lambda_2 = 100$ pour le cadre mensuel et $\lambda_2 = 500$ pour le cadre trimestriel, ceci permet de choisir un panel qui contrairement au Lasso permet à certaines séries corrélées d'entrer dans l'ensemble actif.

4 Données empiriques

Les séries macroéconomiques pour le panel mensuel que j'utilise sont celles qu'on retrouve dans Stock et Watson (2005) et qui sont disponibles sur le site internet de Mark Watson⁸. Le panel contient 132 séries temporelles qui couvrent plusieurs catégories économiques telles que : la production réelle, le salaire réel, le taux d'emploi et les heures de travail, les ventes au détail, les ventes de manufacture et de commerce, la consommation, les ventes de maisons et les débuts de chantiers de construction de maisons, les inventaires réels, les commandes, les prix d'actions, les taux de change, les taux d'intérêt, les agrégats de monnaie et de crédit, les indices de prix, le salaire horaire moyen et autres. Puisque le panel doit être I(0) afin de pouvoir estimer l'espace généré par les facteurs communs, les séries sont transformées

8. <http://www.princeton.edu/~mwatson/>

afin qu'elles soient stationnaires. Les transformations sont les mêmes que dans Stock et Watson (2005), on peut ainsi trouver la liste des séries ainsi que les transformations qui sont appliquées dans cet article. Avant d'estimer les facteurs ou d'appliquer les méthodes de seuillage, les séries sont normalisées. L'échantillon utilisé s'échelonne de 1959 :11 à 2002 :11, je perds les deux premières observations pour faire les transformations. Au final, j'ai 515 observations mensuelles s'échelonnant de 1960 :1 à 2002 :11.

Pour le panel de données trimestrielles, j'utilise celui qu'on retrouve dans Stock et Watson (2009).⁹ Celui-ci consiste en 144 séries macroéconomiques, dont les transformations qui ont été appliquées sont les mêmes que celles de l'article. L'échantillon utilisé correspond à 1959 :III à 2002 :III. Encore une fois, on perd les deux premières observations pour les transformations donc il reste un panel s'échelonnant de 1960 :I à 2002 :III pour un total de 171 observations trimestrielles. Contrairement à Stock et Watson qui utilisent seulement 110 séries pour estimer les facteurs, j'utilise 143 des séries. Je laisse tomber une série qui correspond à la variable TRM_t décrite plus bas.

Je teste le pouvoir prédictif des facteurs sur le log rendement excédentaire du marché. Ceux-ci sont calculés ainsi :

$$r_{m,t+1}^e = r_{m,t+1} - r_{f,t+1} \quad (20)$$

où $r_{m,t+1}$ c'est le rendement du marché qui est remplacé par une variable proxy : le log rendement à valeur pondérée du NYSE/AMEX du CRSP¹⁰ ; r_{t+1}^f c'est le taux sans risque, pour les données mensuelles c'est le rendement du T-Bill de 30 jours et pour les données trimestrielles c'est le rendement du T-Bill de 3 mois.¹¹ J'utilise aussi comme autres variables pour Z_t : le ratio dividende prix $d_t - p_t$; DEF_t l'écart entre le rendement d'une obligation corporative BAA et le rendement d'une obligation corporative AAA ; TRM_t l'écart entre le rendement d'un bon du trésor américain de 10 ans et celui de 3 mois ; et $RREL_t$ qui est l'écart entre le rendement sans risque et sa moyenne sur la dernière année, donc pour les données mensuelles c'est le rendement du T-bill de 30 jours moins sa moyenne sur les 12 derniers mois et pour les données trimestrielles c'est le rendement du T-bill de 3 mois moins sa moyenne sur les 4 derniers trimestres. Cette dernière variable a été construite à partir des données de Yogo Motohiro. J'utilise ces variables pour les régressions prédictives mensuelles et trimestrielles. En plus, pour

9. Disponible sur le même site.

10. Center for Research in Security Prices

11. J'ai trouvé les valeurs pour ces variables sur le site internet de Yogo Motohiro - <http://www.nber.org/myogo/> les données sont utilisées dans Campbell et Yogo (2006)

les régressions avec des données trimestrielles j'utilise la variable cay_t de Lettau et Ludvigson (2001a,b), disponibles sur le site internet du professeur Ludvigson.¹² J'utilise les données révisées pour cette variable.

Pour l'analyse en coupe transversale, j'utilise les rendements sur les 25 portefeuilles triés par taille et valeur de Fama-French. Ces rendements sont disponibles sur le site internet du professeur Kenneth French.¹³ En plus de ces portefeuilles j'utilise aussi les facteurs de Fama et French (1993) qui sont disponibles sur le même site.

5 Résultats empiriques

Pour évaluer s'il y a des gains à utiliser les méthodes de seuillage ou la méthode des composantes principales au carré, j'estime des facteurs sur 6 panels différents pour les observations mensuelles ainsi que pour les observations trimestrielles. La description des panels est donnée au Tableau 1. Pour former le panel d'observations mensuelles à être soumis au seuillage dur j'utilise comme valeur critique $t^c = 1.645$, qui correspond à la valeur de la statistique-t pour un seuil d'un test bilatéral de 10%. Ceci me donne $N = 38$ pour le cas où le seuillage est appliqué aux 132 séries temporelles. Pour les données trimestrielles j'utilise une valeur critique de $t^c = 1.28$, qui correspond au seuil pour un test bilatéral de 20%, ceci me donne $N = 43$. Ceci a été fait car une valeur critique plus élevée réduisait trop la dimension N . J'ai utilisé comme règle, un peu arbitraire peut-être, de ne pas avoir un panel avec moins de 30 séries. Les mêmes valeurs critiques ont été utilisées pour les panels qui incluent le carré des séries, ainsi on ne fait qu'ajouter au panel 3 les nouvelles séries qui sont significatives pour obtenir le nouvel ensemble.

Pour appliquer le seuillage mou aux séries, on doit faire le choix du nombre de séries à inclure dans le panel, K , ainsi que de λ_2 puisque nous utilisons l'algorithme LARS-EN. Pour les données mensuelles K est fixé à 30, et diverses valeurs de λ_2 ont été essayées pour finalement fixer $\lambda_2 = 100$. Le choix de cette valeur repose sur la performance des facteurs estimés dans les régressions prédictives. De même, pour les données trimestrielles, je fixe $N = 45$ et je fais varier λ_2 . La valeur choisie dans ce cas est $\lambda_2 = 500$, encore une fois cette valeur a été choisie comme celle dont les facteurs estimés performant le mieux dans les régressions prédictives.

12. <http://www.econ.nyu.edu/user/ludvigsons/>

13. <http://www.mba.tuck.dartmouth.edu/pages/faculty/ken.french/>

Le choix du nombre de facteurs est fait avec le critère d'information IC_{p2} de Bai et Ng (2002). L'estimation du nombre de facteurs pour chaque cas, par ce critère, est donnée au tableau 2. Pour le cas particulier des panels formés par seuillage dur pour les données mensuelles les critères d'informations de Bai et Ng (2002) ne permettent pas de conclure puisqu'on obtient toujours le maximum. Par conséquent, le choix est effectué par un autre critère qu'ils avaient considéré, soit le BIC_3 qui dans les simulations avaient eu de bonnes propriétés en présence de corrélation entre les séries. Dans le cas du panel 3 j'estime donc $\hat{k} = 5$ facteurs et $\hat{k} = 6$ facteurs pour le panel 4, pour \hat{k} estimé par BIC_3 .

5.1 Régressions prédictives mensuelles

Les tableaux 3 à 5 présentent les résultats pour les régressions prédictives de l'équation (3). Le tableau 3 présente les résultats des régressions prédictives des rendements excédentaires sur les facteurs estimés à partir de chacun des panels. Les résultats présentés sont ceux pour les configurations optimales de facteurs, c'est-à-dire celles qui ont contribué le plus à faire diminuer le BIC, tel que suggéré dans Stock et Watson (2002b). Le coefficient de détermination ajusté (\bar{R}^2) est donné pour chaque cas. Les statistiques-t, corrigées pour l'hétéroscédasticité et l'autocorrélation de Newey-West (1987), sont celles pour un test de l'hypothèse nulle que les coefficients sont nuls, elles sont fournies entre parenthèses sous les coefficients. La valeur calculée pour le critère d'information Bayésien (BIC) est aussi fournie.

D'abord, notons que tous les coefficients des facteurs sont statistiquement significatifs au seuil de 1% pour un test unilatéral que le coefficient est nul. Le nombre de variables qui contribuent le plus à faire baisser les erreurs tel que calculé par le BIC est supérieur ou égal lorsque les facteurs sont soumis à une forme de seuillage. Ces variables sont soit les facteurs estimés ou bien des fonctions formées à partir des facteurs estimés. Le seul cas où des fonctions des variables ont eu un pouvoir explicatif suffisamment élevé pour être inclus dans la configuration rapportée est pour le cas des facteurs estimés à partir des panels soumis à un seuillage mou. Soit le modèle 5, qui inclut les variables $\hat{F}_{1t} \times \hat{F}_{5t}$ et le modèle 6 qui inclut les variables \hat{F}_{2t}^2 et $\hat{F}_{3t} \times \hat{F}_{4t}$. Je rappelle que les fonctions des facteurs n'ont été considérées que pour les facteurs qui avaient le plus d'information sur les rendements excédentaires tels que choisis par le BIC. Par exemple, pour les facteurs estimés à partir des 132 séries, seulement \hat{F}_{2t} contenait suffisamment de pouvoir explicatif pour être un candidat, donc la seule fonction qui a été considérée était \hat{F}_{2t}^2 .

Le fait que le nombre de facteurs importants dans la régression prédictive

des rendements excédentaires soit plus important pour les séries soumises à un seuillage est évocateur. Ceci nous indique que les facteurs estimés contiennent plus d'information sur la variable dépendante lorsque nous choisissons la composition du panel. De ce fait, même si dans chaque cas un nombre inférieur de facteurs est estimé à partir des séries ils sont plus importants comme variables explicatives dans le cadre de l'équation (3). Il y a deux explications possibles pour cet effet comme le notent Boivin et Ng (2006). La première est que dans les panels plus grands les facteurs estimés dominent d'autres facteurs moins importants pour expliquer les co-mouvements dans le panel, mais qui seraient plus informatifs sur les rendements excédentaires. La deuxième explication serait que le fait d'inclure plus de séries entraînerait plus de corrélation entre les erreurs des séries et donc des facteurs moins efficaces.

En comparant le BIC des configurations optimales de facteurs pour chaque panel, on note qu'en général les modèles estimés à partir du panel augmenté sont préférés par le critère d'information. Le seul cas où ceci n'est pas vrai c'est pour le cas où l'on compare les résultats pour les facteurs estimés à partir des 132 séries et le cas où l'on compare les résultats qui incluent les facteurs estimés à partir du panel augmenté. Ceci peut être dû à la situation 2 décrite ci-haut. Les méthodes de seuillage dur et mou semblent aussi améliorer la performance du modèle. Ainsi on voit une certaine amélioration dans les coefficients de détermination des modèles avec seuillage dur par rapport aux modèles avec les panels complets, et une amélioration encore plus marquée pour les cas avec seuillage mou. D'ailleurs, le coefficient de détermination est le plus élevé pour le cas avec seuillage mou sur le panel augmenté. C'est aussi ce cas que le critère d'information Bayésien privilégie.

Afin de voir si les résultats sont conséquents en présence d'autres variables je fais des régressions qui incluent $d_t - p_t$, REL_t , TRM_t et DEF_t . Les résultats de ces régressions sont présentés au tableau 4. D'abord, notons que pour l'échantillon considéré le log du ratio dividende prix est significatif pour tous les modèles. Les variables $RREL_t$ et DEF_t ne sont quant à elles significatives pour aucun des modèles considérés. Notons qu'il y a des variables financières dans le panel et que par conséquent l'information marginale que ces deux variables pourraient amener au modèle s'en voit diminuée. Pour un seuil de 5%, d'un test bilatéral la variable TRM_t est significative pour les quatre premiers modèles, mais elle ne l'est pas pour le cas avec des modèles ayant subi un seuillage mou.

Les coefficients des facteurs restent tous statistiquement significatifs en présence de ces autres variables. Encore une fois le critère d'information pri-

vilégie les modèles ayant été soumis à un seuillage, principalement le seuillage mou. Dans cette configuration, le modèle qui performe le mieux demeure celui avec le seuillage mou appliqué au panel augmenté avec un coefficient de détermination ajusté de 0.107.

Le tableau 5 présente les configurations optimales pour les six modèles. Dans ce cas, les variables non significatives soient $RREL_t$ et DEF_t sont retirées. Le log du ratio dividende prix est inclus pour tous les modèles, ainsi que TRM_t pour les modèles 1 à 4. Les résultats et conclusions demeurent les mêmes pour ces cas.

En comparant les deux cas où le seuillage dur a été appliqué, nous notons que bien que le BIC soit plus bas ou égal pour les régressions réalisées avec les facteurs du modèle 4, le coefficient de détermination est plus élevé pour les diverses configurations du panel 3. Notons que le panel 4 inclut un total de 80 séries temporelles et parmi celles-ci il y a les 38 qui constituent le panel 3. Les facteurs estimés sont donc moins informatifs sur les rendements excédentaires et seulement les deux premiers sont significatifs dans les régressions en présence des autres régresseurs. On peut donc voir ici un des cas où l'effet d'avoir plus de séries ne conduit pas nécessairement à un meilleur résultat. Par la méthode appliquée pour obtenir ces panels, il y a la possibilité que les nouvelles séries ajoutées au panel 3 pour obtenir le panel 4 induisent plus de corrélations dans les erreurs. De plus, le fait d'augmenter le critère de seuillage pour former le panel pourrait exacerber cet effet.

5.2 Régressions prédictives trimestrielles

Afin de voir si les résultats sont consistants en présence de la variable cay_t de Lettau et Ludvigson 2001(a,b), je considère aussi des régressions prédictives à un intervalle trimestriel. Cette variable n'est disponible que sur une base trimestrielle et son intérêt repose sur le fait que contrairement aux variables de contrôle disponibles sur une base mensuelle celle-ci est une variable macroéconomique. C'est une variable proxy du ratio consommation-richeesse. Dans Ludvigson et Ng (2007), les facteurs estimés à partir du panel macroéconomique, qui avaient un pouvoir explicatif marginal lorsque seulement les facteurs étaient considérés, n'étaient plus significatifs lorsque cay_t était incluse. Donc l'information que contenaient les facteurs était bien résumée par cette variable.

Le tableau 6 rapporte les résultats des régressions prédictives pour les données trimestrielles. Seuls les résultats pour les configurations optimales

de chaque modèle sont rapportés. Des fonctions non linéaires des facteurs démontrant un fort pouvoir explicatif ont été considérées comme variables, mais elles n'ont pas eu d'effet notable, elles ne sont donc pas incluses dans les résultats présentés à ce tableau. Encore une fois la variable $RREL_t$ n'est pas significative en présence des facteurs et des autres variables, elle n'est donc pas incluse dans aucune des configurations rapportées au tableau 6. La variable TRM_t n'est pas incluse non plus dans aucun des résultats rapportés.

La variable DEF_t est informative en présence des autres régresseurs seulement pour les modèles où le seuillage dur a été appliqué. Le log du ratio dividende prix n'est significatif que pour le modèle incluant les facteurs estimés sur la base de données obtenue par seuillage mou sur le panel augmenté. En accord avec les travaux de Lettau et Ludvigson (2002a) et de Ludvigson et Ng (2007), la variable cay_t a un fort pouvoir explicatif des rendements excédentaires et à elle seule cette variable explique 6% de la variation de $r_{m,t+1}^e$ pour cet échantillon.

Dans les deux premiers modèles, on voit que dans chaque cas il y a un facteur qui est statistiquement significatif en présence de la variable cay_t . Cependant, le pouvoir explicatif marginal n'est pas très important pour ces facteurs. Ainsi dans les deux cas l'amélioration du coefficient de détermination est minime et le critère d'information Bayésien préfère un modèle avec seulement la variable cay_t . Ceci est semblable à ce que Ludvigson et Ng (2007) avaient obtenu pour le panel de série macroéconomique standard. Je note par contre que contrairement à leur cas le panel dont je dispose contient en fait certaines séries financières et est donc un panel mixte en regard des leurs.

Pour les facteurs estimés sous le modèle 5, on note qu'un seul facteur contribue suffisamment à expliquer la variation des rendements excédentaires pour être inclus dans la configuration rapportée. Le pouvoir explicatif marginal de ce facteur est cependant important, dans une régression incluant comme variable \hat{F}_{1t} et une constante on obtient un $\bar{R}^2 = 0.067$. Ce qui est plus important que pour les cas avec les facteurs estimés des panels 1 et 2 où $\bar{R}^2 = 0.001$ et $\bar{R}^2 = 0.01$, respectivement. Aussi voit-on qu'une régression avec cay_t et le facteur \hat{F}_{1t} nous donne un coefficient de détermination $\bar{R}^2 = 0.099$. Le critère d'information Bayésien nous indique aussi que ce modèle est préféré au modèle avec la variable cay_t seulement.

Lorsqu'on applique le seuillage dur, on obtient des configurations finales avec des coefficients de détermination nettement supérieurs. Cependant, notons que l'apport marginal de ces facteurs n'est pas aussi grand qu'il y paraît. Lorsque les facteurs sont considérés seuls ils expliquent $\bar{R}^2 = 0.033$

et $\bar{R}^2 = 0.073$ de la variation des rendements excédentaires. Aussi c'est surtout l'apport de cay_t et de DEF_t qui font que le coefficient de détermination soit aussi élevé. Rappelons que lorsque le seuillage dur a été appliqué l'information contenue dans les variables de contrôle était prise en considération puisque celles-ci étaient incluses dans les régressions à la première étape de cette méthode. Par conséquent, les séries choisies par cette méthode ont une information qui est significative au-delà de ces variables. C'est principalement ce fait qui explique que les modèles 3 et 4 aient de tels résultats, surtout que le pouvoir explicatif marginal des facteurs du panel 3 n'est pas si élevé en comparaison du facteur estimé du panel 5.

Le fait d'ajouter le carré des séries aux candidats à être soumis à un seuillage mou a eu un effet considérable. Ainsi le \bar{R}^2 de cette régression pour le cas où seulement les facteurs \hat{F}_{1t} , \hat{F}_{2t} et \hat{F}_{3t} , estimés du panel 6, et une constante sont présents comme variables explicatives dans la régression est de 0.191. En présence de ces facteurs, cay_t et $d_t - p_t$ sont significatifs. Le coefficient de détermination ajusté pour la régression qui inclut ces deux variables et les facteurs comme régresseurs est 0.232. C'est aussi ce modèle qui minimise le critère d'information Bayésien.

On peut conclure que le fait d'avoir des facteurs estimés à partir de panels soumis à un seuillage améliore la performance des facteurs pour les régressions prédictives des rendements excédentaires du marché. Particulièrement, les facteurs estimés de panels issus du seuillage mou ont un pouvoir explicatif marginal de la variation des rendements qui est supérieur à ceux qui sont estimés des panels obtenus par seuillage dur. Même en présence de la variable cay_t , que diverses recherches ont démontré comme ayant un pouvoir explicatif élevé dans des régressions prédictives trimestrielles, l'information contenue dans ces facteurs est importante et contribue grandement à expliquer la variation des rendements. Ainsi il y a un avantage à considérer un modèle qui inclut ces facteurs.

6 Expérience hors-échantillon

L'analyse ci-haut porte sur des régressions prédictives qui utilisent la totalité de l'échantillon. Dans cette section, je rapporte les résultats pour une expérience hors échantillon. Pour chaque cas, je fais des prévisions en utilisant les modèles incluant seulement les facteurs des configurations optimales rapportées dans les tableaux 3 à 6, excluant donc les variables de contrôle. Pour les données mensuelles, les prévisions sont comparées à un modèle qui

inclut seulement une constante. Pour les données trimestrielles, je les compare dans un cas à celles obtenues avec une constante et dans le deuxième cas à celles obtenues avec une constante et cay_t . Les prévisions sont faites sur trois périodes différentes pour les deux cas : 1970 :1-2002 :12, 1980 :1-2002 :12 et 1990 :1-2002 :12 pour les données mensuelles et 1970 :I-2002 :IV, 1980 :I-2002 :IV et 1990 :I-2002 :IV pour les données trimestrielles. L'expérience est faite en calculant récursivement les facteurs et en estimant les coefficients des régressions jusqu'au temps t afin de prévoir les rendements excédentaires pour la période $t + 1$. Pour les cas des facteurs estimés de panels soumis à des seuillages, les séries qui forment les panels restent constantes pour toutes les périodes.

Les tableaux 7 à 9 présentent les résultats de ces expériences. L'erreur quadratique moyenne relative, c'est-à-dire l'erreur quadratique moyenne du modèle avec les facteurs divisé par celle du modèle restreint, est fournie pour chaque cas. La statistique de test de Clark et McCracken (2001), ENC-NEW, pour l'hypothèse nulle que le modèle restreint inclut le modèle non restreint, est calculée ainsi que la valeur critique de cette statistique pour le 95^e percentile de la distribution asymptotique.

Le tableau 7 présente les résultats de l'expérience hors échantillon pour les données mensuelles. Pour la période 1970 :1-2002 :12, on note que sauf pour le cas où les facteurs sont estimés du panel augmenté, les facteurs offrent une erreur quadratique moyenne plus basse que celle du modèle restreint, avec la moyenne historique comme prévision. Le fait que $EQMR > 1$ pour le modèle 2, prouve que d'ajouter des séries à la base de données peut diminuer la qualité des prévisions. D'ailleurs, ce résultat prévaut pour les deux autres périodes considérées. Sur les trois périodes, au moins un des cas où les prévisions sont faites à partir des facteurs estimés des panels 3 à 6 offre de meilleurs résultats que celles des facteurs estimés du panel 1. Les prévisions faites à partir des facteurs estimés des panels augmentés soumis à des seuillages ont une EQMR plus basse que leur équivalent avec les panels standards, sauf pour les prévisions s'échelonnant sur 1980 :1-2002 :12, où les EQMR des modèles 4 et 6 sont supérieures à celles des modèles 3 et 5, respectivement. D'ailleurs sur cette période les prévisions faites par le modèle 6 ont une EQM plus grande que celles du modèle restreint ou du modèle 1. Notons que pour la période 1990 :1 à 2002 :12, les prévisions de tous les modèles sont de moindre qualité que celles du modèle restreint tel que calculé par l'EQM. Cette particularité est bien connue pour les rendements financiers, ceux-ci sont moins prévisibles après les années 1990s.

Le tableau 8 présente les résultats de l'expérience hors échantillon pour

les données trimestrielles avec comme modèle restreint une constante, soit la moyenne historique. Notons encore une fois que le fait d'ajouter des séries au panel standard affecte négativement les prévisions, ainsi l'EQMR du modèle 2 est supérieure à celle du modèle 1 dans tous les cas. À un intervalle trimestriel, les prévisions des modèles avec des facteurs sont en général inférieures. Ainsi, sauf pour la période 1970 :I-2002 :IV, où les modèles 4 et 5 ont des prévisions qui sont statistiquement plus performantes que celles du modèle restreint, les prévisions des modèles 1 à 5 avec les facteurs ne sont pas statistiquement plus performantes que celles de la moyenne historique. Dans plusieurs cas, elles ont en fait une erreur quadratique moyenne plus élevée. Le modèle avec les facteurs estimés du panel soumis à un seuillage mou sur le panel augmenté est le seul pour lequel le résultat soit satisfaisant. En effet, sur les trois périodes considérées ce modèle fournit de meilleures prévisions que celles avec la moyenne historique et ENC-NEW nous indique que ce résultat est statistiquement significatif. Ce modèle domine aussi les autres en ayant des EQMR plus basses sur toutes les périodes.

Pour effectuer le seuillage dur, nous avons pris en considération l'information incluse dans les variables de contrôle. Par conséquent, l'information contenue dans les facteurs estimés à partir de ces panels peut en effet être plus efficace dans le cas où ces variables sont utilisées pour faire des prévisions. Alors, je considère une expérience hors échantillon qui inclut une constante et cay_t comme modèle restreint, et j'ajoute les facteurs dans le modèle non restreint. Dans ce rapport, j'utilise la variable cay_t comme si elle était observée, par conséquent les coefficients de cointégration sont estimés sur toute la période et non récursivement. J'utilise les données qui sont révisées, donc celles pour lesquelles les coefficients de cointégration ont été recalculés depuis 2001. Le tableau 9 présente les résultats de cette expérience. Les prévisions des rendements excédentaires des modèles 1 et 2 sont encore une fois médiocres. Le modèle 5 quant à lui n'est pas nécessairement meilleur que ces derniers. Cependant, on note que les résultats pour les modèles 3 et 4 sont nettement meilleurs sous cette configuration, sur les trois périodes les prévisions présentent des erreurs quadratiques moyennes inférieures au modèle restreint. Indiquant que l'information contenue dans les facteurs est importante au-delà de celle contenue dans cay_t et les prévisions en sont améliorées. De même, le modèle 6 fournit encore des prévisions avec des erreurs quadratiques moyennes moindres que le modèle restreint sur les trois périodes considérées.

7 Coupe Transversale des Rendements

Les sections précédentes montrent que les facteurs estimés de panels ayant subi un seuillage sont plus informatifs sur les rendements excédentaires, en série temporelle, que ceux estimés des panels contenant toute l'information. Qu'en est-il de la coupe transversale des rendements ? La prochaine section répond à cette question en utilisant les facteurs comme instruments dans le cadre d'un MÉDAF conditionnel sur la coupe transversale des 25 portefeuilles de Fama-French triés par taille et valeur. Commençons par le cadre théorique. En absence d'arbitrage on suppose qu'il existe un facteur d'actualisation stochastique M_{t+1} , tel que pour tout actif avec un rendement au temps t , $R_{i,t+1}$, l'équation suivante est vraie :

$$1 = E_t[M_{t+1}(1 + R_{i,t+1})] \quad (21)$$

où E_t dénote l'espérance conditionnelle à l'information disponible au temps t . Puisque je considère un modèle à facteur linéaire conditionnel, le facteur d'actualisation stochastique a la forme suivante $M_{t+1} = a_t + b_t R_{m,t+1}$ où $R_{m,t+1}$ est le rendement du marché, mais ce pourrait être tout rendement sur la frontière efficiente. On modélise la dépendance temporelle avec des variables Z_t qu'on observe au temps t et donc on a :

$$a_t = a_0 + a_1' Z_t; \quad b_t = b_0 + b_1' Z_t \quad (22)$$

Les coefficients a_0 , a_1 , b_0 , et b_1 sont constants à travers le temps, c'est donc à travers les variables de conditionnement que le facteur d'actualisation se trouve à varier dans le temps. Ceci implique donc un modèle inconditionnel avec comme facteurs de risque : les instruments, les facteurs de risque et les produits de ceux-ci. On peut en effet réécrire le facteur d'actualisation stochastique comme suit :

$$M_{t+1} = a_0 + a_1' Z_t + b_0 R_{m,t+1} + b_1' Z_t R_{m,t+1} \quad (23)$$

On peut tester ce modèle à plusieurs facteurs en utilisant les moments inconditionnels, où les coefficients sont constants :

$$1 = E[[a_0 + a_1' Z_t + b_0 R_{e,t+1} + b_1' Z_t R_{e,t+1}](1 + R_{i,t+1})] \quad (24)$$

La méthode que j'utilise est la même que dans Fama et MacBeth (1973). Pour ce faire, j'estime les bêtas en régressant les rendements des portefeuilles sur les facteurs : $R_{it} = \alpha_i + \beta_{1i} R_{mt} + \beta_{2i}' Z_t + \beta_{3i}' R_{mt} Z_t + \varepsilon_{it}$, pour chacun des portefeuilles i individuellement. Par la suite, je fais la régression en coupe

transversale des rendements des 25 portefeuilles sur les bêtas estimés pour obtenir des estimations pour les gammas dans l'équation suivante :

$$E[R_{i,t+1}] = E[R_{0,t}] + \beta' \lambda \quad (25)$$

Étant donné que les valeurs des bêtas sont estimées à la première étape, cette méthode souffre d'un problème d'erreur dans les variables qui entraîne un biais pour les statistiques t . Bien que les statistiques t de Fama-MacBeth soient données au tableau 10, on y retrouve aussi la statistique corrigée de Shanken (1992) pour chaque gamma estimé. Pour fin de comparaison entre les différentes configurations considérées je rapporte aussi le R^2 de la régression en coupe transversale qui montre la partie de la variation sur les rendements moyens que le modèle explique. Celui-ci est calculé comme suit :

$$R^2 = \frac{Var_c(\bar{R}_i) - Var_c(\bar{\epsilon}_i)}{Var(\bar{R}_i)} \quad (26)$$

où \bar{R}_i dénote la moyenne des rendements excédentaires pour chaque portefeuille, et $\bar{\epsilon}_i$ c'est la moyenne des résidus de chaque portefeuille. L'indice c représente le fait que je prends la variance en coupe transversale. Ce R^2 n'est utilisé que pour des fins de comparaisons entre les diverses configurations. La valeur ajustée de cette statistique est aussi fournie. La statistique de test J_α pour l'hypothèse nulle que les moyennes des erreurs d'évaluations des actifs sont nulles est fournie ainsi que sa valeur- p sous celle-ci.

Dans ce rapport, le facteur de risque est le rendement excédentaire du marché. En principe, d'autres facteurs pourraient être utilisés, cependant puisque je considère un modèle conditionnel, le nombre de variables dans l'équation (24) augmenterait assez rapidement. Pour les instruments, j'utilise les deux facteurs estimés à partir de chaque panel contribuant le plus à diminuer les erreurs d'évaluation des actifs. Aux fins de comparaison, j'estime aussi deux modèles inconditionnels, un avec comme facteur de risque le rendement excédentaire du marché, soit celui impliqué par le MÉDAF ainsi que celui avec les facteurs de Fama et French (1993).

Le tableau 10 présente les résultats des régressions en coupe transversale de Fama-MacBeth pour les rendements des 25 portefeuilles de Fama-French. D'abord, je note que comme dans plusieurs autres études sur la coupe transversale des rendements, le rendement excédentaire du marché a un coefficient négatif, et ce à travers tous les modèles considérés. Cependant, la théorie requiert que ce coefficient soit positif. Notons qu'à part pour le modèle à facteurs de Fama-French et pour les modèles avec seuillage dur, ces coefficients

ne sont pas significatifs au seuil de 5%. Le \bar{R}^2 pour le cas du MÉDAF est de 0.20. Déjà avec le modèle inconditionnel de Fama-French, il y a une nette amélioration de l'ajustement, j'obtiens un \bar{R}^2 de 0.74. Pour les deux cas, par contre, on rejette l'hypothèse nulle que les moyennes des erreurs d'évaluation des actifs sont zéro.

Sauf pour le modèle 3, l'ajustement qu'on obtient avec les facteurs dynamiques comme instruments est meilleur que celui obtenu avec les facteurs de Fama-French. Cependant, notons que pour les facteurs estimés à partir des panels soumis à un seuillage mou, on rejette l'hypothèse nulle que les moyennes des erreurs sont zéro. Les deux modèles qui offrent le meilleur ajustement sont le 2 et le 4. Dans les deux cas on ne rejette pas l'hypothèse nulle que les moyennes des erreurs sont nulles et on obtient un $\bar{R}^2 = 0.91$. Cependant, les résultats obtenus pour le modèle 1 sont très bons aussi avec un $\bar{R}^2 = 0.87$ et le fait que ce modèle implique aussi des moyennes des erreurs nulles. Dans le contexte de la coupe transversale des rendements, il n'y a donc pas grand-chose à gagner à utiliser les méthodes de seuillages. Les risques temporels qui ne sont pas capturés par le rendement du marché le sont assez bien avec les facteurs dynamiques du modèle 1.

La figure 1 montre les rendements moyens estimés contre les rendements moyens réalisés. La ligne de 45 degrés montre là où les rendements des portefeuilles devraient se retrouver si l'on réussit à bien estimer leurs moyennes. Pour le cas du MÉDAF je remarque que l'ajustement est médiocre. Les facteurs de Fama-French nous donnent déjà un meilleur ajustement puisque les rendements suivent assez bien la ligne. L'ajustement est cependant remarquablement bon pour les modèles incluant les facteurs dynamiques. Ce résultat est semblable à celui obtenu par Mönch (2006). Selon le coefficient de détermination en coupe transversale l'ajustement est un peu supérieur avec le modèle utilisant les facteurs dynamiques extraits d'un panel augmenté. Cependant, les différences graphiques ne permettent pas de conclure à une amélioration nécessairement.

Conclusion

Les utilisations des modèles à facteur sont nombreuses de nos jours. Ces modèles sont utilisés en économie financière dans la prévision des rendements excédentaires ou de la volatilité. Dans ce travail, j'ai utilisé les modèles à facteurs dynamiques pour expliquer les rendements excédentaires du marché en série temporelle ainsi que la coupe transversale des rendements excédentaires

des 25 portefeuilles de Fama-French. Des méthodes de seuillages ont été utilisées afin de voir l'effet de la composition du panel sur les facteurs qu'on utilise dans nos modèles.

Les méthodes de seuillage permettent en général d'améliorer le pouvoir explicatif des facteurs estimés sur les rendements excédentaires dans le cadre d'une régression prédictive. De plus, augmenter le panel des candidats avec des fonctions non linéaires des séries permet aussi d'obtenir des facteurs qui améliorent l'ajustement du modèle. Même en présence de variables qui ont un fort pouvoir explicatif (cay_t par exemple), les facteurs ont démontré, dans les cas où un seuillage avait été appliqué, que l'apport de ceux-ci au modèle était non négligeable. De plus, les résultats des expériences hors échantillon permettent aussi de conclure qu'il est avantageux de soumettre les séries à un seuillage plutôt que d'utiliser toutes les séries disponibles.

En coupe transversale, l'utilisation des facteurs dynamiques comme variables de conditionnement permet en général un meilleur ajustement que le modèle à trois facteurs de Fama-French. Les facteurs permettent aussi de diminuer les erreurs d'évaluation des actifs tels qu'on peut le voir par les statistiques J_α . Cependant, le fait d'avoir utilisé un seuillage sur les panels avant d'estimer les facteurs diminue en général la qualité de l'ajustement qu'on obtient avec les facteurs du panel standard. Il n'y a donc pas d'avantages à utiliser ces méthodes dans ce cadre.

Références

- [1] Jushan Bai and Serena Ng. Determining the number of factors in approximate factor models. *Econometrica*, 70(1) :191–221, 2002.
- [2] Jushan Bai and Serena Ng. Forecasting economic time series using targeted predictors. *Journal of Econometrics*, 146 :304–317, 2008.
- [3] Jean Boivin and Serena Ng. Are more data always better for factor analysis? *Journal of Econometrics*, 132 :169–194, 2006.
- [4] John Y. Campbell and Motohiro Yogo. Efficient tests of stock return predictability. *Journal of Financial Economics*, 81(1) :27–60, July 2006.
- [5] Todd E. Clark and Michael W. McCracken. Tests of equal forecast accuracy and encompassing for nested models. *Journal of Econometrics*, pages 85–110, 2001.
- [6] John H. Cochrane. *Asset Pricing*. Princeton University Press, revised edition, 2005.
- [7] Bradley Efron, Trevor Hastie, Iain Johnstone, and Robert Tibshirani. Least angle regression. *The Annals of Statistics*, 32(2) :407–499, 2004.
- [8] Eugene F. Fama and Kenneth R. French. The cross-section of expected stock returns. *Journal of Finance*, 47(2) :427–465, 1992.
- [9] Eugene F. Fama and Kenneth R. French. Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, 33 :3–56, 1993.
- [10] Eugene F. Fama and Kenneth R. French. Size and book-to-market factors in earnings and returns. *The Journal of Finance*, 50(1) :131–155, 1995.
- [11] Eugene F. Fama and Kenneth R. French. Multifactor explanations of asset pricing anomalies. *The Journal of Finance*, 51(1) :51–84, 1996.
- [12] Eugene F. Fama and James D. MacBeth. Risk, return, and equilibrium : Empirical tests. *Journal of Political Economy*, 81(3) :607–636, 1973.
- [13] Martin Lettau and Sydney C. Ludvigson. Aggregate Wealth, and Expected Stock Returns. *The Journal of Finance*, 56(3) :815–849, 2001a.

- [14] Martin Lettau and Sydney C. Ludvigson. Resurrecting the (C)CAPM : A cross-sectional test when risk premia are time-varying. *Journal of Political Economy*, 109(6) :1238–1287, 2001b.
- [15] Martin Lettau and Sydney C. Ludvigson. Measuring and modeling variation in the risk-return trade-off. In Yacine Ait-Sahalia and Lars Hansen, editors, *Handbook of Financial Econometrics*, volume 1, chapter 11, pages 619–690. North Holland, 2009.
- [16] Sydney C. Ludvigson and Serena Ng. The empirical risk-return relation : A factor analysis approach. *Journal of Financial Economics*, 83 :171–222, 2007.
- [17] Sydney C. Ludvigson and Serena Ng. Macro factors in bond risk premia. *The Review of Financial Studies*, 22(12) :5027–5067, 2009.
- [18] Sydney C. Ludvigson and Serena Ng. A factor analysis of bond risk premia. In A. Ulah and D. Giles, editors, *Handbook of Empirical Economics and Finance*, pages 313–372. Chapman and Hall, 2010.
- [19] Emanuel Monch. Conditional asset pricing with a large information set. Humboldt University Berlin, 2005.
- [20] Jay Shanken. On the estimation of beta-pricing models. *Review of Financial Studies*, 5(1) :1–33, 1992.
- [21] James H. Stock and Mark W. Watson. Macroeconomic forecasting using diffusion indexes. *Journal of Business & Economic Statistics*, 20(2) :147,162, April 2002a.
- [22] James H. Stock and Mark W. Watson. Forecasting using principal components from a large number of predictors. *Journal of the American Statistical Association*, 97 :1167–1179, 2002b.
- [23] James H. Stock and Mark W. Watson. Implications of dynamic factor models for VAR analysis. *NBER Working Paper No.11467*, 2005.
- [24] James H. Stock and Mark W. Watson. Forecasting in dynamic factor models subject to structural instability. In Jennifer L. Castle and Neil Shephard, editors, *The Methodology and Practice of Econometrics : Festschrift in Honor of David F. Hendry*, chapter 7. Oxford University Press, 2009.

- [25] James H. Stock and Mark W. Watson. Dynamic factor models. In Michael P. Clements and David F. Hendry, editors, *The Oxford Handbook of Economic Forecasting*, chapter 2. Oxford University Press, 2011.
- [26] Robert Tibshirani. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society Series B*, 58(1) :267–288, 1996.
- [27] Yong Wang. A comparison of factor models for explaining the cross section of stock returns. Northwestern University.
- [28] Hui Zou and Trevor Hastie. Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society Series B*, 67(2) :301–322, 2005.

Tableau 1

Panel	Description
1	Panel Standard avec toutes les séries
2	Panel avec toutes les séries et le carré des séries
3	Panel choisi avec le seuillage dur sur les séries
4	Panel choisi avec le seuillage dur sur les séries et leur carré
5	Panel choisi avec le seuillage mou sur les séries
6	Panel choisi avec le seuillage mou sur les séries et leur carré

TABLE 1 – Description des panels utilisés pour extraire les facteurs dynamiques

Notes : Bien qu'il y ait deux panels, un avec des données mensuelles et l'autre avec des données trimestrielles, la même notation est réutilisée. Le panel avec des données mensuelles est celui de Stock et Watson (2005), il contient 132 séries. J'estime les facteurs pour la période s'échelonnant de 1960 :I à 2002 :11. Pour les données trimestrielles, j'utilise le panel de Stock et Watson (2008). Ce dernier contient 144 séries et j'en estime les facteurs pour la période s'échelonnant de 1960 :I à 2002 :III. Les transformations appliquées au préalable aux diverses séries sont les mêmes que celles décrites dans ces deux textes. De plus, les séries sont normalisées avant d'en extraire les facteurs, pour avoir une moyenne de zéro et une variance unitaire.

Tableau 2

Panel	Données mensuelles	Données trimestrielles
	\hat{k}	\hat{k}
1	8	4
2	12	5
3	10	4
4	10	4
5	5	3
6	5	4

TABLE 2 – Nombre de facteurs communs estimés par IC_{p2} pour les divers panels

Notes : Ce tableau rapporte les estimations du nombre de facteurs communs du critère d'information IC_{p2} de Bai et Ng (2002). Pour les données mensuelles, $kmax = 20$ pour les panels 1 et 2 et $kmax = 10$ pour les autres. Pour les données trimestrielles $kmax = 10$ pour tous les panels.

Tableau 3

Variable	Panel					
	1	2	3	4	5	6
const	0.0034 (1.98)	0.0034 (1.92)	0.0034 (1.96)	0.0034 (1.97)	0.0034 (2.02)	0.0052 (2.96)
\hat{F}_{1t}			0.018 (4.77)	0.018 (3.87)	0.018 (4.57)	0.021 (5.92)
\hat{F}_{2t}	-0.031 (-4.09)	-0.0303 (-3.45)	0.012 (2.43)	0.022 (3.95)		
\hat{F}_{3t}					0.027 (4.26)	
\hat{F}_{4t}		-0.025 (-2.74)	0.021 (2.73)			
\hat{F}_{2t}^2						-0.015 (-3.48)
$\hat{F}_{1t} \times \hat{F}_{5t}$					-0.048 (-3.88)	
$\hat{F}_{3t} \times \hat{F}_{4t}$						-0.07 (-6.75)
\bar{R}^2	0.033	0.036	0.066	0.060	0.087	0.10
BIC	-6.29	-6.28	-6.30	-6.31	-6.32	-6.34

TABLE 3 – Régressions des rendements excédentaires mensuels sur les facteurs estimés

Notes : Ce tableau rapporte les estimations des régressions par MCO des rendements excédentaires sur les variables nommées à la première colonne : $r_{m,t+1}^e = \beta_0 + \beta_1' \hat{F}_t + \epsilon_{t+1}$, où \hat{F}_t peut être une fonction des facteurs estimés. La variable dépendante est le log rendement à valeur pondérée du NYSE/AMEX du CRSP moins le rendement des bons du Trésor américain (T-Bill) de 30 jours, pour la période 1960 :2 à 2002 :12. Les variables \hat{F}_{it} sont estimées par la méthode des composantes principales à partir des panels mensuels tels que décrits au tableau 1, pour la période 1960 :1 à 2002 :11. Les statistiques-t corrigées de Newey-West (1987) sont fournies entre parenthèses. Les résultats présentés ici sont ceux pour le sous-ensemble de variables qui contribue le plus à diminuer le BIC pour chaque groupe.

Tableau 4

Variable	Panel					
	1	2	3	4	5	6
const	0.075 (2.80)	0.085 (3.11)	0.08 (3.00)	0.075 (2.91)	0.062 (2.36)	0.063 (2.32)
\hat{F}_{1t}			0.022 (4.61)	0.02 (3.46)	0.018 (4.27)	0.025 (5.28)
\hat{F}_{2t}	-0.052 (-4.52)	-0.058 (-4.29)	0.035 (4.57)	0.044 (5.80)		
\hat{F}_{3t}					0.04 (4.72)	
\hat{F}_{4t}		-0.035 (-3.49)	0.021 (2.93)			
\hat{F}_{2t}^2						-0.017 (-3.52)
$\hat{F}_{1t} \times \hat{F}_{5t}$					-0.048 (-4.00)	
$\hat{F}_{3t} \times \hat{F}_{4t}$						-0.068 (-6.77)
$d_t - pt$	0.019 (2.84)	0.022 (3.21)	0.02 (2.93)	0.019 (2.88)	0.016 (2.51)	0.015 (2.23)
$RREL_t$	1.047 (0.33)	2.150 (0.63)	1.095 (0.35)	4.71 (1.29)	-0.099 (-0.03)	1.306 (0.39)
TRM_t	-0.005 (-2.07)	-0.004 (-2.07)	-0.009 (-3.18)	-0.004 (-2.06)	-0.004 (-1.77)	0.0004 (0.20)
DEF_t	-0.0003 (0.06)	0.001 (0.27)	0.004 (0.69)	-0.0002 (-0.03)	0.003 (0.56)	-0.007 (-1.25)
\bar{R}^2	0.055	0.057	0.089	0.082	0.100	0.107
BIC	-6.26	-6.25	-6.28	-6.28	-6.29	-6.29

TABLE 4 – Régressions des rendements excédentaires mensuels sur les facteurs estimés et Z_t

Notes : Ce tableau rapporte les estimations des régressions par MCO des rendements excédentaires sur les variables nommées à la première colonne : $r_{m,t+1}^e = \beta_0 + \beta_1' \hat{F}_t + \beta_2' Z_t + \epsilon_{t+1}$, où \hat{F}_t peut être une fonction des facteurs estimés. La variable dépendante est le log rendement à valeur pondérée du NYSE/AMEX du CRSP moins le rendement des bons du Trésor américain (T-Bill) de 30 jours, pour la période 1960 :2 à 2002 :12. Les variables \hat{F}_{it} sont estimées par la méthode des composantes principales à partir des panels mensuels décrits au tableau 1, pour la période 1960 :1 à 2002 :11. Les variables de contrôle sont $d_t - pt$ le log du ratio dividende-prix ; $RREL_t$ qui est l'écart entre le rendement du taux sans risque et sa moyenne sur la dernière année ; DEF_t est l'écart entre le rendement des obligations corporatives BAA et AAA ; et TRM_t est l'écart entre le rendement d'un bon du Trésor américain de 10 ans et celui de 3 mois. Les statistiques-t corrigées de Newey-West (1987) sont fournies entre parenthèses.

Tableau 5

Variable	Panel					
	1	2	3	4	5	6
const	0.074 (3.44)	0.089 (3.86)	0.09 (4.01)	0.08 (3.61)	0.058 (2.97)	0.045 (2.36)
\hat{F}_{1t}			0.023 (5.49)	0.016 (3.5)	0.018 (4.78)	0.021 (5.66)
\hat{F}_{2t}	-0.051 (-4.64)	-0.055 (-4.38)	0.033 (4.33)	0.041 (5.38)		
\hat{F}_{3t}					0.031 (4.91)	
\hat{F}_{4t}		-0.034 (-3.49)	0.021 (2.91)			
\hat{F}_{2t}^2						-0.017 (-3.99)
$\hat{F}_{1t} \times \hat{F}_{5t}$					-0.047 (-4.00)	
$\hat{F}_{3t} \times \hat{F}_{4t}$						-0.07 (-6.77)
$d_t - pt$	0.018 (3.14)	0.023 (3.59)	0.022 (3.61)	0.020 (3.26)	0.016 (2.79)	0.011 (2.06)
TRM_t	-0.005 (-2.32)	-0.005 (-2.33)	-0.008 (-3.10)	-0.006 (-2.88)		
\bar{R}^2	0.041	0.060	0.091	0.080	0.098	0.109
BIC	-6.29	-6.28	-6.30	-6.30	-6.32	-6.34

TABLE 5 – Régressions des rendements excédentaires mensuels sur les facteurs estimés et Z_t , configurations optimales

Notes : Ce tableau rapporte les résultats des estimations par MCO de l'ensemble de variables optimales pour chaque panel. Les régressions sont celles des rendements excédentaires sur les variables nommées à la première colonne : $r_{m,t+1}^e = \beta_0 + \beta_1' \hat{F}_t + \beta_2' Z_t + \epsilon_{t+1}$, où \hat{F}_t peut être une fonction des facteurs estimés. La variable dépendante est le log rendement à valeur pondérée du NYSE/AMEX du CRSP moins le rendement des bons du Trésor américain (T-Bill) de 30 jours, pour la période 1960 :2 à 2002 :12. Les variables \hat{F}_{it} sont estimées par la méthode des composantes principales à partir des panels mensuels décrits au tableau 1, pour la période 1960 :1 à 2002 :11. Les variables de contrôle sont $d_t - pt$ le log du ratio dividende-prix et TRM_t est l'écart entre le rendement d'un bon du Trésor américain de 10 ans et celui de 3 mois. Les statistiques-t corrigées de Newey-West (1987) sont fournies entre parenthèses.

Tableau 6

Variable	Panel						<i>cay_t</i>
	1	2	3	4	5	6	
const	0.004 (0.71)	0.004 (0.71)	-0.045 (-3.05)	-0.043 (-3.00)	0.005 (0.90)	0.156 (2.33)	0.004 (0.72)
\hat{F}_{1t}					0.038 (3.23)	0.045 (4.15)	
\hat{F}_{2t}			0.041 (2.64)			0.079 (4.78)	
\hat{F}_{3t}	-0.037 (-2.56)			-0.074 (-4.17)		0.063 (4.58)	
\hat{F}_{4t}			0.11 (4.20)	0.081 (3.56)			
\hat{F}_{5t}		0.053 (1.80)					
<i>cay_t</i>	1.48 (3.55)	1.43 (3.50)	2.16 (5.66)	1.75 (5.13)	1.15 (3.07)	0.87 (2.65)	1.44 (3.44)
$d_t - p_t$						0.044 (2.25)	
<i>DEF_t</i>			0.048 (3.65)	0.047 (3.20)			
\bar{R}^2	0.061	0.066	0.175	0.178	0.099	0.232	0.057
BIC	-4.96	-4.96	-5.03	-5.03	-4.99	-5.07	-4.98

TABLE 6 – Régressions des rendements excédentaires trimestriels sur les facteurs estimés et Z_t , configurations optimales

Notes : Ce tableau rapporte les résultats des estimations par MCO de l'ensemble de variables optimales pour chaque panel. Les régressions sont celles des rendements excédentaires sur les variables nommées à la première colonne : $r_{m,t+1}^e = \beta_0 + \beta_1' \hat{F}_t + \beta_2' Z_t + \epsilon_{t+1}$, où \hat{F}_t peut être une fonction des facteurs estimés. Aucune fonction formée à partir des facteurs estimés n'a été significative pour les données trimestrielles. La variable dépendante est le log rendement à valeur pondérée du NYSE/AMEX du CRSP moins le rendement des bons du Trésor américain (T-Bill) de 3 mois, pour la période 1960 :II à 2002 :IV. Les variables \hat{F}_{it} sont estimées par la méthode des composantes principales à partir des panels mensuels décrits au tableau 1, pour la période 1960 :I à 2002 :III. Les variables de contrôle sont *cay_t* de Lettau et Ludvigson (2001a,b); $d_t - p_t$ le log du ratio dividende-prix et *DEF_t* l'écart entre le rendement des obligations corporatives BAA et AAA. Les statistiques-t corrigées de Newey-West (1987) sont fournies entre parenthèses. Les variables *RREL_t* et *TRM_t* ont aussi été essayées comme faisant partie de Z_t , mais elles n'ont pas exhibé un pouvoir explicatif suffisant pour pouvoir être incluse dans aucune configuration rapportée. Pour servir d'exemple une régression avec une constante et *cay_t* a été estimé, ceci nous donne une idée de l'effet marginal des facteurs pour chaque modèle.

Tableau 7

	Panel					
	1	2	3	4	5	6
1970 :1 - 2002 :12						
EQMR	0.971	1.005	0.959	0.948	0.965	0.95
ENC-NEW	16.29	10.16	35.51	28.53	25.90	31.78
VC 95%	2.42	3.34	4.05	3.34	4.05	4.05
1980 :1 - 2002 :12						
EQMR	0.976	1.01	0.978	0.98	0.962	1.001
ENC-NEW	8.69	6.77	21.42	11.23	19.62	11.15
VC 95%	1.66	2.33	2.84	2.33	2.84	2.84
1990 :1 - 2002 :12						
EQMR	1.033	1.059	1.058	1.032	1.021	1.006
ENC-NEW	-0.58	2.11	6.31	0.43	1.74	3.06
VC 95%	1.11	1.52	1.91	1.52	1.91	1.91

TABLE 7 – Résultat des expériences hors-échantillon Facteurs vs Constante pour les données mensuelles

Notes : Ce tableau rapporte les résultats des expériences hors échantillon pour les données mensuelles. La constitution de chaque panel reste constante dans le temps ainsi que les variables. EQMR réfère à l'erreur quadratique moyenne relative, soit l'erreur quadratique moyenne du modèle non restreint divisée par celle du modèle restreint. La valeur de EQMR la plus petite pour chaque période est présentée en caractères gras. Le modèle non restreint inclut une constante et les facteurs. Le modèle restreint contient seulement une constante. Les prévisions sont faites sur trois périodes : 1970 :1 à 2002 :12, 1980 :1 à 2002 :12 et 1990 :1 à 2002 :12. La statistique de test est celle de ENC-NEW de Clark et McCracken (2001) pour l'hypothèse nulle que le modèle restreint inclus le modèle non restreint. Les valeurs critiques VC 95% sont les valeurs du 95^e percentile de la distribution asymptotique de la statistique de test, que l'on trouve dans Clark et McCracken (2001).

Tableau 8

	Panel					
	1	2	3	4	5	6
	1970 :I - 20062 :IV					
EQMR	1.01	1.013	1.002	0.986	0.975	0.917
ENC-NEW	0.32	1.85	5.54	15.57	13.25	41.59
VC 95%	2.43	2.43	3.36	3.36	2.43	4.05
	1980 :I - 2002 :IV					
EQMR	0.99	1.002	1.056	1.001	1.007	0.938
ENC-NEW	0.82	2.03	0.14	6.1	5.7	13.93
VC 95%	1.67	1.67	2.34	2.34	1.67	2.39
	1990 :I - 2002 :IV					
EQMR	1.00	1.026	1.009	1.02	1.049	0.945
ENC-NEW	0.12	-0.33	0.38	3.27	0.2	5.13
VC 95%	1.11	1.11	1.53	1.53	1.11	1.92

TABLE 8 – Résultat des expériences hors-échantillon Facteurs vs Constante pour les données trimestrielles

Notes : Ce tableau rapporte les résultats des expériences hors échantillon pour les données trimestrielles. La constitution de chaque panel reste constante dans le temps ainsi que les variables. EQMR réfère à l'erreur quadratique moyenne relative, soit l'erreur quadratique moyenne du modèle non restreint divisée par celle du modèle restreint. La valeur de EQMR la plus petite pour chaque période est présentée en caractères gras. Le modèle non restreint inclut une constante et les facteurs. Le modèle restreint contient seulement une constante. Les prévisions sont faites sur trois périodes : 1970 :1 à 2002 :12, 1980 :1 à 2002 :12 et 1990 :1 à 2002 :12. La statistique de test est celle de ENC-NEW de Clark et McCracken (2001) pour l'hypothèse nulle que le modèle restreint inclus le modèle non restreint. Les valeurs critiques VC 95% sont les valeurs du 95^e percentile de la distribution asymptotique de la statistique de test, que l'on trouve dans Clark et McCracken (2001).

Tableau 9

	Panel					
	1	2	3	4	5	6
1970 :I - 2002 :IV						
EQMR	1.006	1.023	0.966	0.95	0.998	0.962
ENC-NEW	0.75	0.89	7.61	18.344	8.69	33.96
VC 95%	2.43	2.43	3.36	3.36	2.43	4.05
1980 :I - 2002 :IV						
EQMR	0.99	1.019	0.968	0.934	0.988	0.925
ENC-NEW	0.95	0.87	3.99	9.26	4.87	13.817
VC 95%	1.67	1.67	2.34	2.34	1.67	2.39
1990 :I - 2002 :IV						
EQMR	0.98	1.004	0.908	0.923	1.043	0.925
ENC-NEW	0.56	0.22	3.35	6.16	-0.06	5.85
VC 95%	1.11	1.11	1.53	1.53	1.11	1.92

TABLE 9 – Résultat des expériences hors-échantillon Facteurs vs Constante et cay_t pour les données trimestrielles

Notes : Ce tableau rapporte les résultats des expériences hors échantillon pour les données trimestrielles avec la variables cay_t . La constitution de chaque panel reste constante dans le temps ainsi que les variables. EQMR réfère à l'erreur quadratique moyenne relative, soit l'erreur quadratique moyenne du modèle non restreint divisée par celle du modèle restreint. La valeur de EQMR la plus petite pour chaque période est présentée en caractères gras. Le modèle non restreint inclut une constante, cay_t et les facteurs qu'on peut voir au tableau 6. Le modèle restreint contient seulement une constante et cay_t . Les prévisions sont faites sur trois périodes : 1970 :1 à 2002 :12, 1980 :1 à 2002 :12 et 1990 :1 à 2002 :12. La statistique de test est celle de ENC-NEW de Clark et McCracken (2001) pour l'hypothèse nulle que le modèle restreint inclus le modèle non restreint. Les valeurs critiques VC 95% sont les valeurs du 95^e percentile de la distribution asymptotique de la statistique de test, que l'on trouve dans Clark et McCracken (2001).

Tableau 10

Modèle		Facteurs				$R^2(\bar{R}^2)$	J_α
Médaf	const	R_m					
$\hat{\lambda}$	1.82	-0.65				0.23 (0.2)	56.44
Stat-t	4.86	-1.54					0.00
Shanken-t	4.77	-1.52					
FF	const	R_m	SMB	HML			
$\hat{\lambda}$	1.77	-0.83	0.15	0.46		0.77 (0.74)	42.18
Stat-t	5.6	-2.24	1.02	3.45			0.00
Shanken-t	5.32	-2.15	1.02	3.44			
Panel 1	const	R_m	\hat{F}_2	\hat{F}_7	$\hat{F}_2 \times R_m$	$\hat{F}_7 \times R_m$	
$\hat{\lambda}$	1.06	-0.12	-0.09	0.03	0.90	-0.71	0.90 (0.87)
Stat-t	2.82	-0.30	-1.16	0.56	2.38	-4.03	0.15
Shanken-t	2.71	-0.29	-1.12	0.54	2.29	-3.88	
Panel 2	const	R_m	\hat{F}_6	\hat{F}_{11}	$\hat{F}_6 \times R_m$	$\hat{F}_{11} \times R_m$	
$\hat{\lambda}$	1.52	-0.57	-0.07	-0.07	0.47	-0.04	0.93 (0.91)
Stat-t	5.27	-1.63	-1.91	-2.22	1.90	-0.30	0.35
Shanken-t	5.16	-1.60	-1.87	-2.18	1.86	-0.29	
Panel 3	const	R_m	\hat{F}_1	\hat{F}_5	$\hat{F}_1 \times R_m$	$\hat{F}_5 \times R_m$	
$\hat{\lambda}$	1.82	-0.83	0.22	0.27	-0.32	0.23	0.74 (0.68)
Stat-t	5.65	-2.22	2.12	3.51	-0.45	0.57	0.00
Shanken-t	5.43	-2.16	2.05	3.38	-0.43	0.55	
Panel 4	const	R_m	\hat{F}_4	\hat{F}_6	$\hat{F}_4 \times R_m$	$\hat{F}_6 \times R_m$	
$\hat{\lambda}$	1.72	-0.77	0.17	0.22	-0.38	-0.01	0.93 (0.91)
Stat-t	6.71	-2.43	1.77	4.25	-0.96	-0.05	0.19
Shanken-t	6.57	-2.39	1.73	4.16	-0.94	-0.05	
Panel 5	const	R_m	\hat{F}_1	\hat{F}_5	$\hat{F}_1 \times R_m$	$\hat{F}_5 \times R_m$	
$\hat{\lambda}$	1.37	-0.39	0.20	-0.33	-0.57	-0.14	0.81 (0.76)
Stat-t	3.73	-0.97	1.94	-4.86	-0.88	-0.34	0.01
Shanken-t	3.60	-0.95	1.87	-4.69	-0.85	-0.33	
Panel 6	const	R_m	\hat{F}_2	\hat{F}_4	$\hat{F}_2 \times R_m$	$\hat{F}_4 \times R_m$	
$\hat{\lambda}$	1.09	-0.22	-0.01	-0.18	0.68	-2.07	0.86 (0.83)
Stat-t	3.01	-0.56	-0.16	-2.23	1.46	-4.72	0.04
Shanken-t	2.90	-0.54	-0.16	-2.15	1.41	-4.56	

TABLE 10 – Tests Fama-MacBeth du MÉDAF conditionnel avec deux Instruments

Ce tableau présente les résultats de régressions Fama-MacBeth pour le MÉDAF inconditionnel, le modèle à 3 facteurs de Fama-French, et pour un MÉDAF conditionnel avec deux facteurs comme instruments pour chacun des panels. Le modèle est :

$$E(R_{i,t+1}) = E(R_{0,t}) + \beta_{i,f_1} \lambda_1 + \dots + \beta_{i,f_k} \lambda_k \quad (27)$$

Pour chaque estimation de λ_i les statistiques-t de Fama-MacBeth sont fournies ainsi que les statistiques-t corrigées de Shanken (1992). Le R^2 est le coefficient de détermination de la régression en coupe transversale et \bar{R}^2 la valeur correspondante ajustée. J_α est le test que la moyenne des erreurs d'évaluation est zéro, la valeur-p de cette statistique est fournie en dessous.

Figure 1

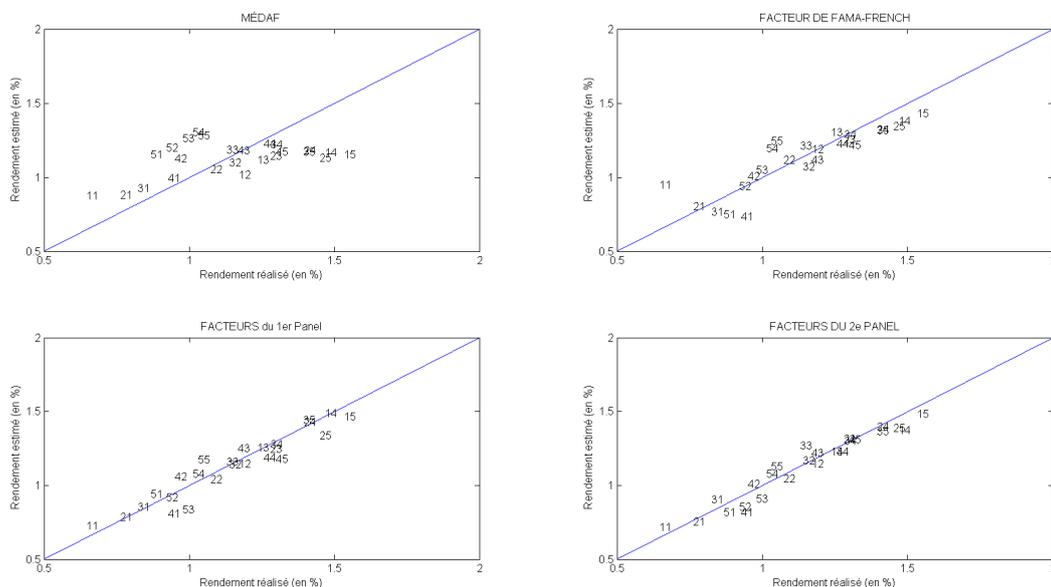


FIGURE 1 – Graphique des rendements estimés sur les rendements réalisés

Ce graphique présente la moyenne des rendements réalisés pour les 25 portefeuilles de Fama-French sur l'axe horizontal contre la moyenne des rendements estimés par quatre modèles sur l'axe vertical : le MÉDAF inconditionnel ; le modèle à trois facteurs de Fama-French ; un MÉDAF conditionnel avec comme instrument les deux facteurs estimés à partir des 132 séries, \hat{F}_2 et \hat{F}_7 ; et un MÉDAF conditionnel avec comme instrument les deux facteurs estimés à partir du panel augmenté du carré des 132 séries, \hat{F}_6 et \hat{F}_{11} . Chaque portefeuille est identifié avec un nombre à deux chiffres, le premier correspond au quintile de la taille, 1 = petit et 5 = grand, et le deuxième correspond au quintile du ratio de la valeur comptable à la valeur marchande, 1 = basse et 5 = élevée.